Miloje M. Rakočević

HARMONY OF GENETIC CODE
Previous works

Volume 1

Belgrade, 2016

A short Foreword

> *"In determination of the genetic code, except two inherent alphabets – twenty amino acids and four amino bases (two pyrimidines & two purines – is involved still one "hidden alphabet", the series of natural numbers, with all its regularities and laws" (Chapter 10 in this book, p. 4).*

This book contains my works published in the period 2005-2013 on my website (also in arXiv). The concept of "harmony" in the title refers to the determination of the genetic code by golden mean, generalized golden mean and harmonic mean. Some parts of the contents, in the meantime are published in some of the official journals, but most are not, and this was the reason for my decision to publish all papers here in their entirety. [The work, which here is given as the third chapter, previously is published in the Proceedings of the Montenegrin Academy of Sciences, together with academician Zvonimir Damjanović] The quote above shows best in what is the difference between my insights (into the essence of the genetic code) in relation to the insights determined by the paradigm in the current science.

Belgrade, 31.12.2016

Miloje M. Rakočević
www.rakocevcode.rs
mirkovmiloje@gmail.com

CONTENTS

GENETIC CODE AS A HARMONIC SYSTEM

Miloje M. Rakočević

Department of Chemistry, Faculty of Science, University of Niš, Ćirila i Metodija 2, Serbia[*]

Abstract.

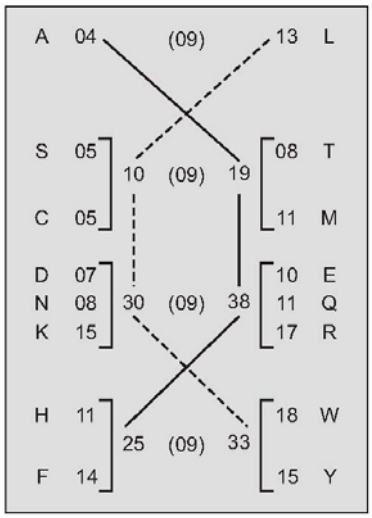
In a certain way, this paper presents the continuation of the previous one which discussed the harmonic structure of the genetic code (Rakočević, 2004). Several new harmonic structures presented in this paper, through specific unity and coherence, together with the previously presented (Rakočević, 2004), show that it makes sense to understand genetic code as a set of several different harmonic structures. Thereby, the harmonicity itself represents a specific unity and coherence of physico-chemical properties of amino acid molecules and the number of atoms and/or nucleons in the molecules themselves (in the form of typical balances). A specific Gauss' arithmetical algorithm has the central position among all these structures and it corresponds to the patterns of the number of atoms within the side chains of amino acid molecules in the following sense: G+V = **11**; P+I = **21**; S+T+L+A+G = **31**; D+E+M+C+P = **41**; K+R+Q+N+V = **61**; F+Y+W+H+I = **71**; (L+M+Q+W) + (A+C+N+H) = **81**; (S+D+K+F) + (T+E+R+Y) = **91**; (F+L+M+S+P) = (T+A+Y+H+I) = (Q+N+K+D+V) = (E+C+W+R+G) = **51**. Bearing in mind all these regularities it make sense to talk about genetic code as a harmonic system. On the other hand, such an order provides new evidence supporting the hypothesis established in the previous paper (Rakočević, 2004) that genetic code has been complete from the very beginning and as such was the condition for the otigin and evolution of life.

1 INTRODUCTION

In the previous paper we have presented a strictly harmonic structure of the genetic code (Table 1 in Rakočević, 2004), consisting of 4 x 5 canonical amino acids (AAs), which follows from two arrangements of amino acid pairs; the first arrangement as presented in Table 1 and the second one as presented in Table 2. In this paper however we will present several other harmonic structures such that they altogether show that the genetic code is a kind of harmonic system. Thereby, the harmonicity itself represents a specific unity and coherence of physico-chemical properties of amino acid (AA) molecules and of the number of atoms and/or nucleons in them (in the form of typical balances).

[*] E-mail: m.m.r@eunet.yu

The starting point in this new analysis of harmonicity of genetic code is the system of 16 AAs of alanine stereochemical type (Survey 1.1 in Rakočević and Jokić, 1996 and Table 1 in this paper). That system presents the relations between amino acid *pairs* through a natural (chemical) distinction into 1-2-3-2 pairs, with the balance of the number of atoms (86:86) along two zigzag lines. [The system of 1-2-3-2 pairs: aliphatic AAs A & L, chalcogene S-T & C-M, basic and acidic plus acid (amide) derivatives K-R, D-E & N-Q, and aromatic AAs: H-W & F-Y; Zigzag line actually represents the first possible periodicity within a „periodic system" consisting of two columns („groups") and several rows („periods").]



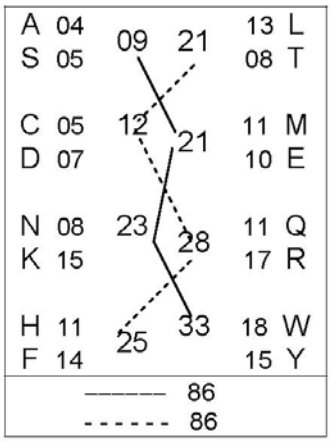
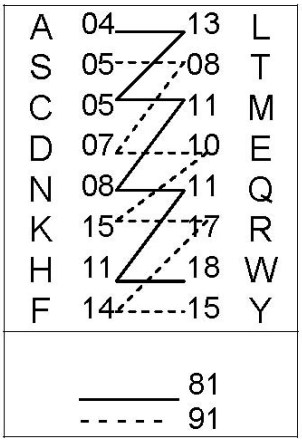


Table 1 (left). The 16 amino acids of alaninic stereochemical type (I). This Table is presented in our previous paper (Survey 1.1 in Rakočević & Jokić, 1996). The 8 amino acid pairs belong to the alaninic stereochemical type (Popov, 1989), that is to the class of „non-contact" amino acids. (About „contact" and „non-contact" amino acids *see* in Remark 1.) Hierarchy observes chemical classification into aliphatic (as the first) and aromatic (as the second) amino acids, whereby the order of pairs has been determined by the mass of the first member in the pair.
Table 1.1 (right). The 16 amino acids of alaninic stereochemical type (II). This Table originates from Table 1, in the way as explained in the text.

The order of the pairs itself has been determined so that the subclass of aliphatic AAs comes first, followed by the subclass of aromatic AAs. Within each subclass the order is dictated by the number of atoms in the side chain of each first member of an amino acid pair.



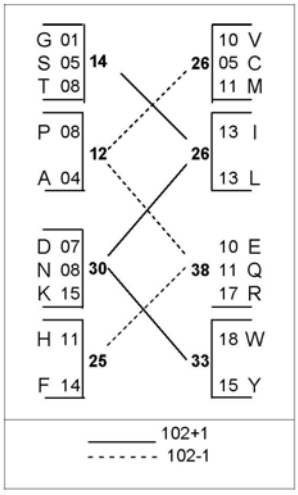


Table 1.2 (left). The 16 amino acids of alaninic stereochemical type (III). This Table originates from Table 1.1. On two zigzag lines here we have the connection between amino acid singlets whereas in Table 1.1 between the doublets.

Table 2 (right). The atom number balance directed by two classes of enzymes aminoacyl-tRNA synthetases (I). This arrangement of amino acid pairs follows from the first one, presented in Table 1 after hard regularities, given in a previous paper (Rakočević, 1998b). On the first (full) zigzag line, there is 102+1 whereas on the second (dotted) line 102-1 atoms. Arithmetic mean for both: 102±1. Class I of enzymes aminoacyl-tRNA synthetases handles the larger amino acid molecules (on the right) whereas the smaller molecules (on the left) are handled by class II (after Rakočević, 1998b).

2 A NEW ARRANGEMENT OF PAIRS

The first question (in this analysis) is the issue of possible rearrangement of the system of eight molecule pairs of AAs of alanine stereochemical type (Table 1), so that the formally unbalanced arrangement of molecule pairs (1-2-3-2), reflecting a possible chemical classification, is transformed into a balanced one (2-2-2-2), maintaining the balance expressed also in the number of atoms along two zigzag lines, regardless of the fact that then the initial classification will be „disrupted“. An affirmative answer to this question (repeatedly obtained ratio 86:86 in the atom number along two zigzag lines) is presented in Table 1.1.

Where then did such result come from? The only acceptable explanation is that the said „disruption“ of classification is still only an illusion. Once established hierarchy of pairs, through strictly determinated neighbourhood, both chemically (physico-chemical charactiristics) and formally (atom number) must

possess at least a unit degree of freedom, so new associations/dissociations of molecules, i.e. molecule pairs, are possible. Thus, from the chemical point of veiw, the said associations/dissociations in a new rearrangement look like this: A-L & S-T, D-E & C-M, K-R & N-Q. By this, in the subclass of aromatic AAs the order is not changed, in the transition from the arrangement given in Table 1 to the one given in Table 1.1.

Indeed, it makes sense, i.e. chemical justification, for *the first* – the oxygen substituted derivatives (S-T) to go together with the original aliphatic molecules (A-L); and also for *the second* – the oxygen substituted compounds to go together with *the first* – the oxygen analogue molecules: D-E with C-M. Finally, it is entirely logical that nitrogen AAs are also found together: K-R with N-Q. [If S-T is the first substituted pair, then D-E is the second; besides, the oxygen pair S-T and sulphuric C-M are analogue in a class of chalcogene AAs.]

3 A GAUSS' ARITHMETICAL ALGORITHM

The presence of the sulfur AAs in a set of 20 canonical AAs represents a "nonserial" substitution in the sense that these are the only two AAs whose functional group owns a chemical element located lower than the second period of Periodic system of chemical elements, in the third period more precisely. The question is also raised here why the sulphur was "selected" and not phosphorus, for example? Why (–SH), and not ($–PH_2$) functional group? The answer is that 2 atoms (–SH) fit together, while 3 atoms ($–PH_2$) do not fit into a very specific *arithmetical algorithm*, which we will designate as "Gauss' algorithm"[1] (for reasons in footnote 1 stated) (Figures 1-2 in relation to Tables 3-4**)**.

[1] There are several anegdotes about Gauss as a boy and his understanding of mathematics at his earliest age. One of them say that the teacher gave an exercize to the class of nine year olds in which Gauss was a pupil, to add all numbers from 1 to 100, thinking he will have enough free time to do his "personal things", but also giving the opportunity to potential geniuses... Indeed, less than three minutes passed when little Gauss came with the result: 5050. Asked how calculated it, he explained that he added the first and the last number (1+100 = 101), then the following and the first preceding one (2+99 = 101) and so on. Since there are 50 such pairs, he multiplied 50 x 101 and obtained the requested result. Now, if the teacher had given the exercize to add all numbers not from 1 to 100, but from 1 to 101, and that the results are distributed in 10 columns, Gauss would get the results as in Tab. 3, whose first order presents, as we can see, the "patterns" of the number of atoms in the rows and columns of the system "5 x 4" of AAs in genetic code (Figure 1).

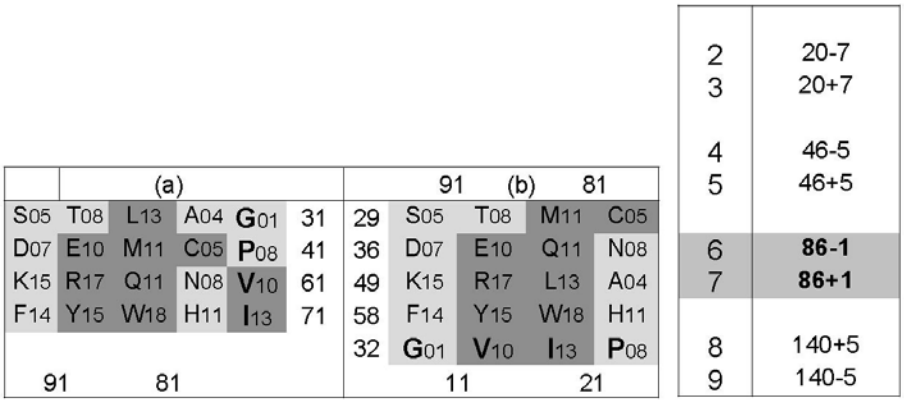


Figure 1 (left). The distribution of amino acids according to Gauss' algorithm; (a) the distribution of amino acids has been derived from Table 1.1, by presenting two pairs here in one row as the "pair of pairs". In the beginning of each row obtained in this way one "contact" amino acid has been associated, by the increasing molecule mass (about "contact" and "non-contact" amino acids see in Remark 1). Atom number (in amino acid side chains) in the rows and columns generated in this way corresponds, one hundred per cent, to the quantums from the first row of Gauss' algorithm of adding numbers from 1 to 101 (Table 3). Dark tones: Class I of amino acids handled by class I of enzymes aminoacyl-tRNA synthetases; light tones: Class II. Obviously, chemically related groups of AAs have been "taken off" by 0, 1 and 2 steps, respectively. By zero "steps" in aromatic; by one step in chalcogenic AAs (M, C in relation to S, T) and carboxylic (carboxylic AAs D & E in relation with their amides N & Q); by two steps in source aliphatic AAs: A, L & K, R. (b) The distribution is the same as in (a) but "leveling" has been conducted here according to the chemical properties of molecules, so there is no more taking off. The class of contact AAs has been added to the beginning of columns, instead of to the beginning of rows. Shading is the same as in (a). Chemical "leveling" conducted in this way is correspondent with the determination of structure with a specific algorithm – algorithm of symmetry through cyclization (Figure 1.1).

Figure 1.1 (right). Arithmetical algorithm of symmetry through cyclization, correspondent to amino acid splitting into four chemical classes, as it is shown in Figure 1, on the right. Within two inner classes (AA side chains) there is 81-1 whereas within two outer classes 81+1 of atoms. The steps of the algorithm: 1. Choose two adjacent numbers from decimal scale; 2. Take the squares of both; 3. Move one modular cycle more (in module 9), for example from 49 to 58; 4. Go back for one half, for example from 58 to 29. From the Figure it is clear that only one solution is possible (balance in accordance to principle of minimum change).

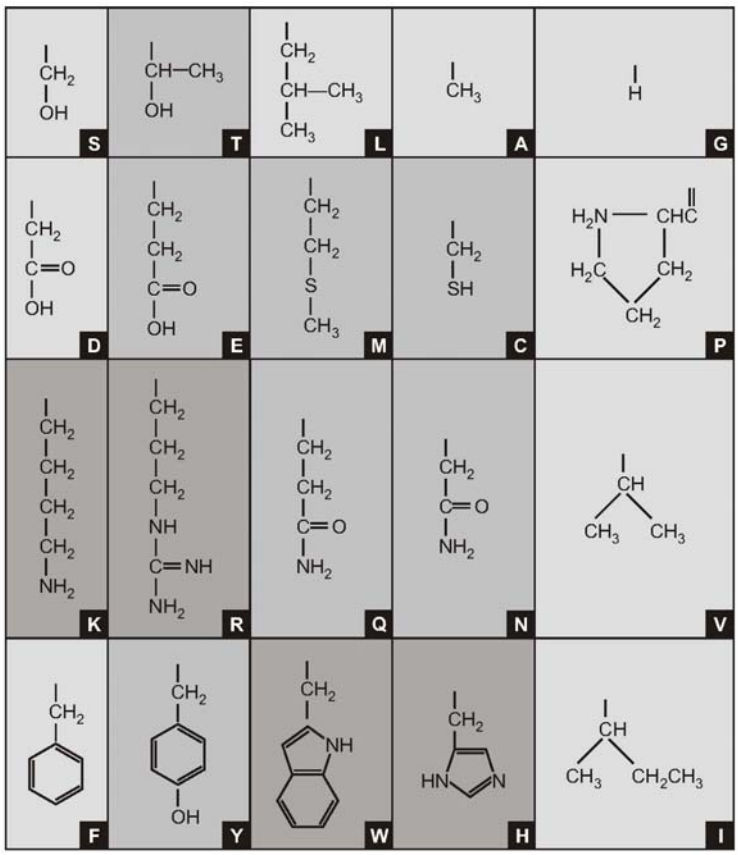


Figure 1.2. The structure of amino acid molecules. The simplest amino acid is glycine (G) whose side chain is only one atom of hydrogen. It is followed by alanine (A) whose side chain is only one CH_3 group, which is the smallest hydrocarbon group. There are total of 16 amino acids of alaninic stereochemical type ("non-contact" amino acids) with one CH_2 group each between the "body" and the "head". The glycinic type contains glycine (G) only; valinic type contains valine and isoleucine (V, I); The last stereochemical type is prolinic with proline (P) which represents the inversion of valine in the sense that the "triangle" of three CH_2 groups for the "head" is not bound by the basis, therefore not only with one but with two CH_2 groups (Popov, 1989; Rakočević & Jokić, 1996). Light tones (G, P, V, I & A, L, S, D, F): invariant AAs; most dark tones (K, R, W, H): most variant AAs; less dark tones (T, E, Q, M, N, C): less variant AAs. (Cf. Section 4.1.)

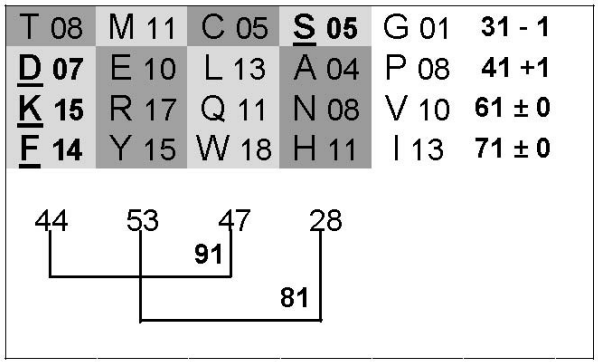


Figure 2. The distribution of amino acids according to Gauss' algorithm, with a minimal change: for ±1 atom in first two rows and for one "take of" in columns. The distribution of amino acids has been derived from Table 2.1 in the same manner as Figure 1 from Table 1.1. Dark and light tones show the changes within the AAs columns going from Figure 1 to Figure 2 (for example: T,E,R,Y in column in Figure 1).

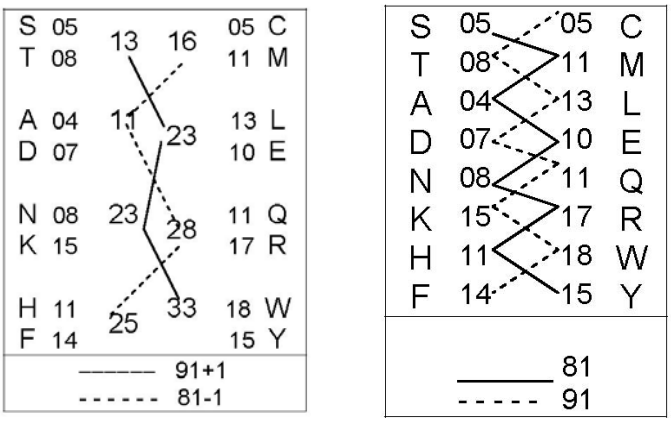


Table 2.1 (left). The atom number balance directed by two classes of enzymes aminoacyl-tRNA synthetases (II). All is the same as in Table 2, except the fact that contact AAs (G-V and P-I) are excluded. The arrangement itself is analog to this one in Table 1.1, the AAs are splitting into 4 x 4 sets.

Table 2.2 (right). The atom number balance directed by two classes of enzymes aminoacyl-tRNA synthetases (III). This Table follows from Table 2.1 in the same manner as Table 1.2 from Table 1.1.

However, apart from the atom number, the type of functional group (actually the unity of these two factors) determines priority in the „selection" of one or the

other amino acid, in the following sense. If we can consider hydrocarbon molecules as main organic compounds, which is a fact actually, then (–OH) derivatives, with 2 atoms in a functional group, are the first possible ones in the act of „copying“ the head of an amino acid to the body (in which copying the principle of *self-similarity* is realized, moving from a part of a molecule to another); only then does the copying of ($–NH_2$) group with 3 atoms follow and finally, the copying of (–COOH) group with 4 atoms, which represents the highest degree of oxydation (and substitution) in the main aliphatic chain. All three copyings are related to functional groups with a single covalent bond, while the copying of (=CO) group which possesses double bond does not occur. However, it is not „neglected“, since this group is the one through which connection with the amino acid precursors is made. [The amino acid precursor pyruvate, a derivative of acetic acid, possessing a (=CO) functional group, is the starting and central precursor of an amino acid (Rakočević, 1998a).]

The finding that the atom number in side chains of 20 canonical AAs, when they are organized in a strict system, the *pairs of pairs* system, is one hundred percent determined by Gauss' algorithm[2], is demonstrated with the main argument which supports not only the *hypothesis* on the complete genetic code, but also the *thesis* that GC is not of a random but strictly determined nature. After such finding it even makes sense to raise a new hypothesis, for further researches, that this and such genetic code represents a general pattern of the *life code*, valid for the entire universe, wherever the existence of water in its liquid aggregate state is possible, with other corresponding conditions.

3.1 Further determinations

The system of eight pairs of AAs of alanine stereochemical type (Table 1) should also be tested from the aspect of possible maintenance of balance of the number of atoms when this system – the „non-contact“ amino acids class – is associated with the remaining three stereochemical types, with the „contact“ AAs class. [*Remark 1*: „Contact“ AAs are those AAs in which there is a direct

[2] The rules of Gauss' algorithm [algorithm of calculating the sum of all numbers in a row in the sequnce of natural numbers, starting from 1 to a number *n* (1 + 2 + 3 + ... + n)] are obvious: I. *Pairs* of numbers (1, n), (2, n-1), (3, n-2), etc. are formed, for which for the even *n* there are n/2, and for the odd *n* there are (n-1)/2 (Tab. 3), each pair with the sum n+1; II. The value of product (n+1)(n/2), i.e. (n+1)(n-1)/2 is calculated, for even, i.e. odd *n*, respectively; III. The central member of the sequence is added to the obtained products, which is 0 for even *n*, and [(n-1)/2] +1 for odd *n*.

contact of the „body“ and the „head“, i.e. the side chain with the amino acid functional group, while „non-contact“ AAs are the ones in which contact is mediated by a CH_2 group (only with threonine that group is metyl-substituted: $H - C - CH_3$). Contact AAs include three stereochemical types: glycinic (with glycine), prolinic (with proline) and valinic (with valine and isoleucine), while the remaining 16 AAs belong to non-contact AAs, all of alaninic sterechemical type. (For details about stereochemical types, *see* Popov, 1989; Rakočević & Jokić, 1996).]

An affirmative answer to the requested testing is presented in Figure 2. How did we come to this result? We have demonstrated that the system of AAs pairs, presented in Table 1, can be „opened“ and with strict rules translated into the system – „the pairs of pairs“ (Figure 1). In one of our earlier works (Rakočević, 1998b) we showed that the same system (Table 1), also according to strict rules, can also be incorporated into the system of two AAs classes, handled by class I and class II of enzymes aminoacyl-tRNA synthetases, respectively (in further text: First enzyme handled system, EHS1, Table 2).

Once we know that, it makes sense to „open“ this system as well (by means of excluding contact AAs, G-V and P-I), and to translate it into „the pairs of pairs“ system (Figure 2). As we can see, only with a small degree of freedom (±1), i.e. with two minimal exceptions from the Gauss' algorithm, a whole new (complete) system is created (Second enzyme handled system, EHS2, as presented in Figure 2).

3.2 Enzyme determination and polarity

The second enzyme handled system, EHS2 (Figure 2), can be expanded by adding contact AAs not only to rows but also to columns as presented in Figure 2.1. Nevertheless, what is of special importance here is a new atom number balance within the columns (102:102 through multiples of number 6) we also have a visible distinction of AAs from the aspect of polarity/nonpolarity. Thereby, it is evident from Figure 2 that all these balances and distinctions are possible through one exception in EHS2 in comparison with EHS1: the original order F-Y-W-H must be inverted into order F-Y-H-W.

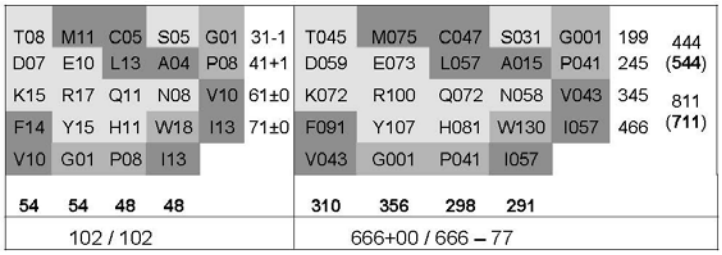


Figure 2.1. This Figure follows from Figure 2 through an adding of contact AAs not only into rows but also into the columns. The details in the text (Section 3.2).

Polar AAs are positioned as a separate entity, in the form of a specific "island" surrounded by non-polar AAs (Figure 2.1). Within the "island" there are all polar AAs, all except serine, which is separate, whereby the existence of a "commodity" with a certain degree in expressing a degree of freedom in polarity as well. Interestingly enough, there are also three ambivalent AAs (glycine, proline and tryptophan)[3] positioned in a snug "string" at the very edge of the system. The distinction is even more complete when semi-ambivalent histidine is added to these three ambivalent AAs. [We can really regard histidine as a „semi-ambivalent" amino acid since it has neither positive nor negative value in cloister energy, but its value is equal to zero (Figure 5 in Swanson, 1984).]

3.3 Atom number multiples and bioprecursors

The presented determination by number 6 (Figure 2.1)[4] can be brought in connection with the determination of canonical AAs (on the binary-code tree) by Golden mean (Survey 2.1 in Rakočević, 1998b)[5]; also with the determination

[3] Glycine: after hydropathy is polar; after cloister energy and polar requirement is non-polar; Proline: after hydropathy and cloister energy is polar; after polar requirement is non-polar; Tryptophan: after hydropathy and polar requirement is polar; after cloister energy is non-polar. Hydropathy (Kyte & Doolittle, 1982; Doolittle, 1985); cloister energy (Swanson, 1984); polar requirement (Woese et al., 1966; Konopel'chenko and Rumer, 1975). About the pairing process of AAs through Hydropathy and Cloister energy *see* Survey 1 in Rakočević & Jokić, 1996, and about ambivalence of glycine and proline see Section 3.3 in Rakočević, 2004.

[4] As a noteworthy is the fact that number 6 is the first perfect number. A hypothesis that perfect numbers can appear as determinant of genetic code we published ten years ago (Rakočević, 1997, pp. 60-64).

through a specific splitting of AAs in relation to their biosynthetic precursors (Figure 3) as well as with the determination of their positions in the Genetic code table (GCT) (Figure 4).

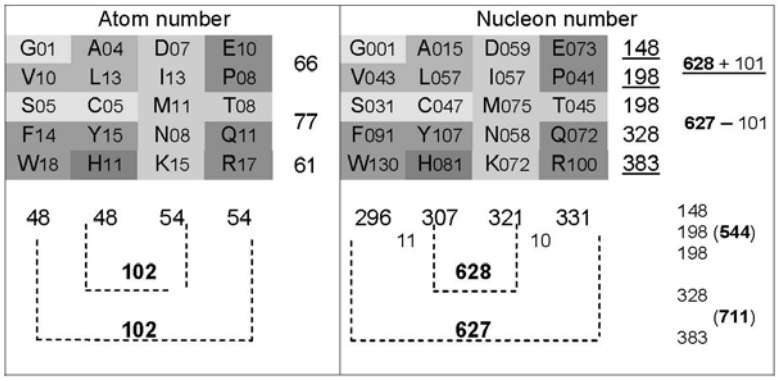


Figure 3. The number of atoms and nucleons within side chains of AA molecules in correspondence with the classification of their bioprecursors (*see* text, Section 3.3). On the left side (in both parts of the Figure): very light tones – AAs synthesized through first bioprecursor (3-Phosphoglycerate: G, S, C); middle light tones – AAs synthesized through second bioprecursor (Pyruvate: A, L, V); very dark tones – amino acid synthesized by fifth bioprecursor (Ribose-5-phosphate: H); dark tones – AAs synthesized by sixth bioprecursor (Phosphoenolpyruvate plus eritrose-4-phosphate: F, Y, W). On the right side (in both parts of the Figure): light tones – AAs synthesized by third bioprecursor (Oxaloacetate: T, M, I, D, N, K); dark tones – AAs synthesized by fourth bioprecursor (2-Oxoglutarate, i.e. α-Ketoglutarate: P, E, Q, R).

Figure 3 shows atom and nucleon number balances in correspondence to classification of amino acid bioprecursors into two classes: the first class with four "outer" precursors and the second class with two "inner" precursors. [*Remark 2*: "Outer" bioprecursors are: 1st, 2nd, 5th and 6th and the "inner" ones 3rd and 4th. The order and hierarchy of AAs biosynthetic precursors are given as in our previous work (Table 1 in Rakočević & Jokić, 1996): 1. 3-Phosphoglycerate (G, S, C); 2. Pyruvate (A, L, V); 3. Oxaloacetate (T, M, I, D, N, K); 4. 2-Oxoglutarate, i.e. α-Ketoglutarate (P, E, Q, R); 5. Ribose-5-phosphate (H); 6. Phosphoenolpyruvate plus eritrose-4-phosphate (F, Y, W). Thereby, the first

[5] Rakočević, 1998b, p. 289: „Within seven 'golden' amino acids there are 60 atoms; within their seven complements there are [60+(1 x 6)] and within six non-complements there are {[60+(1 x 6)] + (2 x 6)} of atoms."

three precursors are *less complex* and last three are *more complex*. At the same time, the first and the two last ones are phospho-precursors, while other three are non-phospho-precursors.]

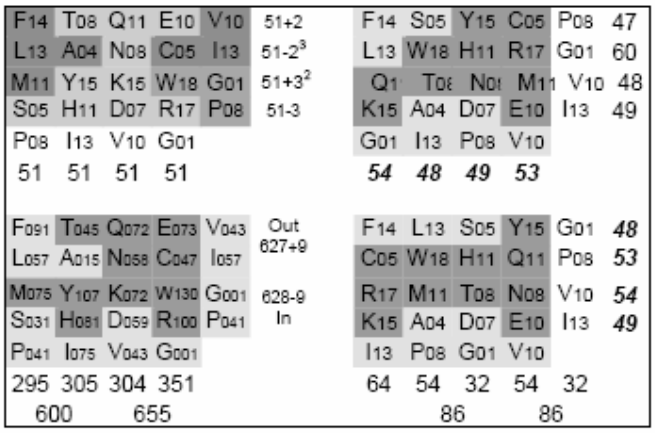


Figure 4. Three "readings" of Genetic Code Table; first on the left side and last two on the right side of Figure. *On the left, up*, the reading through columns: if "contact" AAs (G-V and P-I) are excluded from GCT, the rest of "non-contact" AAs exists in an arrangement of 4 x 4: first quartet (F+L+M+S), second (T+A+Y+H), third (Q+N+K+D+V) and forth (E+C+W+R); if "non-contact" AAs are excluded from GCT, the rest of "contact" AAs exists in an arrangement of 4 x 1: regarding GCT in up-down direction, the first amino acid is P, second I, third V and the fourth G. Altogether: (F+L+M+S+**P**) = (T+A+Y+H+**I**) = (Q+N+K+D+**V**) = (E+C+W+R+**G**) = 51 of atoms each (within side chains); light tones – polar AAs; middle light tones – ambivalent AAs (G, P, W); dark tones – non-polar AAs. *On the left, down*: the order is the same as previous, but with nucleon number balances. *On the righ, up*, the reading through rows – *Py* level. *On the right, down*, the reading through rows – *Py-Pu* level.

3.4 Three "readings" of Genetic Code Table

The order of AAs in the very GCT is also determined by multiples of number 6 (Figure 4). We have come to this result with the finding that the order of non-contact AAs should be read after the exclusion of contact ones and vice-versa.

The upper left corner of Figure 4 shows the order of reading by columns, in the classes of four AAs (with association of contact AAs at the beginning of each column and each row, in the adequate sequence). As we can see, atom number in

all columns (51 each) is exactly as much as the central point[6] of Gauss' algorithm of „adding numbers to 101“ (Table 3). We can say that these are 102 atoms each in the first and the second half of GCT, where number 102 is also multiple of number 6 (102 = 17 x 6).

01+101	**11+91**	**21+81**	**31+71**	**41+61**
02+100	12+90	22+80	32+70	42+60
03+099	13+89	23+79	33+69	43+59
04+098	14+88	24+78	34+68	44+58
05+097	15+87	25+77	35+67	45+57
06+096	16+86	26+76	36+66	46+56
07+095	17+85	27+75	37+65	47+55
08+094	18+84	28+74	38+64	48+54
09+093	19+83	29+73	39+63	49+53
10+092	20+82	30+72	40+62	50+52
01+100	11+90	21+80	31+70	41+60
01+101	**11+91**	**21+81**	**31+71**	**41+61**
01+102	11+92	21+82	31+72	41+62

$0\underline{0}1_1$	01_{10}				101_1	002_{10}			
013_4	07_{10}	06	6	1	101_4	017_{10}	15	18	1
025_7	19_{10}	12	6	1	101_7	050_{10}	33	18	1
037_{10}	37_{10}	18	6	1	101_{10}	101_{10}	51	18	1
049_{13}	61_{10}	24	6		101_{13}	170_{10}	69	18	
$05B_{16}$	91_{10}	30			101_{16}	257_{10}	87		
.....									

Table 3 (left). The Gauss' algorithm for n = 101. The calculating the sum of numbers from 1 to 101 by determining the number of pairs with the same sum. Since here n is an odd number, a middle member appears here, which is 51. In the down part of the Table, the first row has been given for the Table with n = 100, then the first row for the Table with n = 101 and, finally, the first row for the Table with n = 102. (The explanation in the text, especially in footnotes 1 and 2.)
Table 4 (right). The Analogues of number 037 and number 101. Shcherbak (1994; 2003) demonstrated that the number of nucleons in two classes of amino acids (in four-codon and non-four-codon AAs) has been determined with "prime quantum 037", and he also established that out of all double digit numbers written in three positions, number 037 in the decimal numbering system has specific characteristics: all three digits, if they are not reduced to the same digit (as in 3 x 037 = 111), are preserved by multiplying numbers in the limits of module 9 (for example: 1 x 037 = 037; 10 x 037 = 370; 19 x 037 = 703). Shcherbak presented numbering systems with the same characteristics; in other words, he presented analogues of number 037 in these systems, and here they are given on the left side of the Table. With all this, we should note that here determination of continuously generated "steps" is "hidden" here (1 x 1 = 1; 1 x 11 = 11; 3 x **037** = 111; 11 x **101** = 1111). With this insight, it becomes clear why number 101 is also specific and why it has been "selected" to be the key determinant of the genetic code.

Nevertheless, there is a determination with the number 6 also by reading the order of AAs by rows, of which (orders) there must inevitably be two: right up and right down in Figure 4. The *right up* arranging of AAs is realized in sach a way that *Py* is read first, followed by the reading of *Pu* (in third codon positions); the *right down* arrangement of AAs represents a „leap“ in the sense that one must first „read“ all AAs in one four-codon family (*Py-Pu* in the third codon position)

[6] We should take note that this is the only point which has been excluded from the numbering system presented in Table 3.

in one row, then in the next row, etc. Of course, here we also should add contact AAs at the beginning of columns and rows in adequate permutations – adequate from the aspect of achieving balance in the atom number.

Somewhat unexpectedly, we find the same patterns of atom number, up in the columns and down in the rows: 48-49 and 53-54, which are the same multiples of the atom number we find in the system in Figure 3, with differentiations ±0 and ±1. Moreover, nucleon number in three systems (in Figures 2.1, 3 and 4), also in the relation columns-rows, realizes a specific balance (655 – 544 = 111 and 711- 600 = 111)[7].

Take note that in the system presented in Figure 4, determination with number "51" is shown, not only by columns, but by rows as well, with a typical change, correspondent to module 3/2. [We know that the ratio 3:2 is "the limit of the golden numbers" (Moore, 1994) and a key relation in the triadic Cantor set. (Falconer, 1990, p. XIII: "the middle third Cantor set is one of the ... most easily constructed fractals").]

4 LOGIC OF CYCLICITY

The prerequisite and the possbility of unity of physico-chemical characteristics and arithmetical structures, i.e. arithmetical algorithms, are contained in an authentic logic – in *the logic of cyclicity* of the molecules themselves. Namely, by the insight into the validity of Gauss' algorhythm, as a strict determinant of the genetic code, it becomes clear that our earlier presentation (Rakočević & Jokić, 1996) on generating amino acid constituents on the principle of open/closed condition of the side chain now has to be understood as an inevitable *logic of cyclicity*, whereby an open chain presents the zeroth cyclicity.

4.1 Variant and invariant Amino acids

In the analysis of these purely chemical relations, we start from Figure 1.2 which graphically presents amino acid molecules according to the same distribution as in Figure 1, which corresponds to Gauss' algorithm. Out of 20 AAs nine appear as *invariant*, in the sense that these are the first possible amino

[7] Number 111 is the first possible Shcherbak's same-digit pattern, determined by Prime quantum 037 (Shcherbak, 1994) and correspondent to the pattern 1111, determined by Prime quantum 101 (cf. Table 4 where is shown that 3 x 037 = 111 and 11 x 101 = 1111).

acid molecule-structure solutions; the remaining 11 are *variant*, but the variability is limited by the Gauss' algorithm. Namely, the first row is followed by the second one with only slightly more massive molecules, such that they provide the transition from the quantum-level „31" to the quantum-level „41". After the second order significantly more massive molecules must be „selected" so they could achieve the quantum-level „61", at the same time „overlapping" the quantum-level „51". In the next row the „tempo" is slower again – again we have a leap of about ten units (to quantum-level „71"). But to make the whole thing more strict and more specific, the columns must be fitted into the algorithm too, two columns into quantum-level „81" and the remaining two in the quantum-level „91".

The presented double "jump" from quantum-level "41" to quantum-level "61" could explain why, for example, in the third row of the system presented in Figure 1.2, there is a chemically less cognate arginine (with 17 atoms in the side chain) along with lysine, and not the chemically much more cognate ornithine. (Jukes, 1973, p. 24: „I have suggested that arginine displaced ornithine during the evolution of protein synthesis".) With 12 atoms in the side chain ornithine would find itself in the „baned" area between two „allowed" levels. However, bearing in mind the fact that in such a significant biochemical process, such as the biosynthesis of urea, there are interproducts as citrulline, apart from ornithine and arginine, the question is raised why it could not have been the candidate instead of arginine?

In the adequacy of arginine and inadequacy of citrulline lies the entire sense of the logic of choice: apart from formal correspondence (citrulline also has exactly 17 atoms in the side chain, just like arginine) the correspondence in chemical characteristics is also required, as much as possible. So, arginine, with one amino and one *imino* group, is adequate for to make a pair with lysine, while citrulline with one amino and one *carbonyl group* is not. The choice of amino acid molecule with any other number of atoms in the side chain except 17 (such as it is the case with arginine, in correspondence with 15 atoms in the side chain of lysine)[8] is additionally enabled by the fact that fitting into the adequate quantum-level is necessary within both the columns and rows, as has already been said. But what is the most interesting here is the fact that the level has been achieved by „pure" chemistry at the same time, as will be explained hereinafter.

---

[8] From the aspect of validity of the principle of self-similarity lysine which, apart from an NH_2 group, also has four CH_2 groups in the side chain, is more correspondent to the leucine (the first possible case of branched molecule) than are the derivatives with one, two, or three CH_2 groups.

4.2 The generation of the entire amino acid system

In order to give an explanation for just above presented question, let us take another look at Figure 1.2 and try to analyze the generation of the entire system of canonical AAs. The first possible amino acid must be Glycine (glycinic stereochemical type) with side chain of only one and that the smallest possible atom („copied" from the „head" of amino acid). But if that amino acid is an element of a system which is starting to generate itself, the following amino acid must be related to it and must be the first possible case of the zeroth cyclicity – Alanine, with CH_3 group as a side chain; more accurately with CH_2-**H** group, if CH_2 group is a characteristic of the subsystem (subsystem of alaninic stereochemical type) and if H represents the cognate „clasp" both with the Glycine, and with the first possible substituent, with serine (**H** in CH_2-**H** group substituted by **OH** group). (With the generation of *oxygen* substituent – serine, in a „parallel" process its *sulfurous* analogue, cysteine, is generated.)[9] The next two steps are: the first possible semi-cyclicity – valine (valinic stereochemical type: V & I) and the first possible full cyclicity – proline; proline as a prolinic stereochemical type (Rakočević and Jokić, 1996).

By generating the diastereoisomeric isoleucine with the side chain of a butyl group, leucine was inevitably generated with isobutyl group (which is actually, the first possible *branching*), as well as diastereoisomeric treonine ([H-C(OH)-CH3]) instead of the potential serine derivative ([H-C(OH)-H $\rightarrow$ H-C(CH_2-OH)-H]). [Correspondent and parallel with the generation process of S-C, is the generation process of T-M.]

Upon the generation of serine, in the next step (parallel) there is a generation of aspartic acid (by „copying" carboxylic group from the amino acid „head") and its first possible derivative – glutamic acid; followed by the generation of amide derivatives of these two carboxylic amino acids (N & Q).

Along the other line, by generating leucine, phenylalanine is also generated since it possesses the same structural motif; phenylalanine, followed by its hydroxide derivative, tyrosine. [Phenylalanine possesses the structure of an

---

[9] All these chemical connections are naturally realized through bioprecursors, namely in the relations presented in three previous papers (Rakočević and Jokić, 1996; Rakočević, 1998a, 2004).

isobutyl grupe, in other words it is generated from toluen and not from benzene, which is the reason that toluen's ring is the starting molecule for aromatic canonical amino acids.]

Hereby, all *invariant* AAs and their first possible derivatives and/or analogs – the *less variant* AAs – are generated. The remaining ones are only four AAs – two aliphatic and two aromatic ones, all four – the *more variant* AAs (Figure 5). Here we now return to the discussion on fitting arginine and lysine, so only two aromatic AAs remain which must be selected so that they fit into the "narrow passage", created by rows and columns in the quantum-level "71" and "81", respectively. [We should take note that here, when fitting into this „narrow passage", tryptophan follows the logic of generating phenylalanine (though only with one of its two rings), and the last amino acid, histidine, follows the logic of the second tryptophan ring (raising the degree of freedom further, exactly be how much pyrole differs from imidasole).]

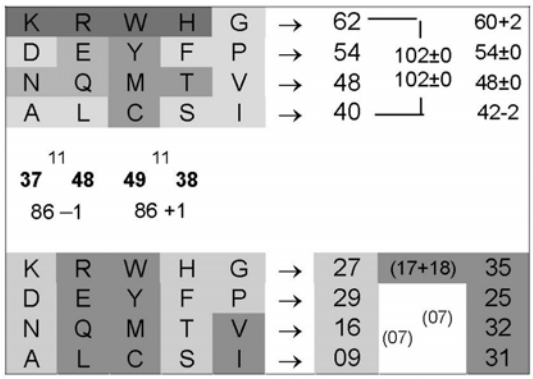


Figure 5. If we take as the "starting point" (above: dark-light as in Fig. 1.2) the "most variant" AAs, in the order which they have in the system of Gauss' algorithm (K,R,W,H) (Figure 1), and then if the remaining AAs are arranged according to chemical characteristics, a new arrangement is created. In this new arrangement the number of atoms in the columns is determined by Shcherbak's Prime quantum 37 (Shcherbak, 1994, 2003), while in the rows it is determined with the multiples of number 6. [Whereas the *number of atoms* is determined there by quantum 037, in the Shcherbak's system the *number of nucleons* is determined by that same quantum (in both amino acid classes, in four-codon and non-four-codon AAs) (Fig. 1 and Tab. 1 in Shcherbak, 1994).] Down: dark-light tones are given as in Figure 1; if so, then the determination through the minimum change principle is obvious (the change for a unit).

5 A NEW CLASSIFICATION OF AMINO ACIDS

Starting from analyzed relations in the amino acid system, presented in Figure 1.2, as well as from the finding that a „half" (10-1) of the set of canonical AAs is made as invariant, and the other „half" (10+1) as variant AAs, we come to a new classification of AAs, through a source splitting into „contact" and „non-contact" AAs (cf. Remark 1).

Class 1: *Invariant AAs* [**subclass 1**: contact AAs in two families; family 1: singlet AAs (G, P) and family 2: non-singlet AAs (V, I)[10]; **subclass 2**: non-contact AAs also in two families; family 1: aliphatic AAs (A, L, S, D) and family 2: aromatic amino acid (F); within family 1 there are two sub-families; sub-family 1: source aliphatic AAs (A, L) and sub-family 2: aliphatic derivatives (S, D)].

Class 2: *Variant AAs* [**subclass 1**: more variant AAs (family 1: aliphatic AAs K & R, and family 2: aromatic AAs W & H); **subclass 2**: less variant AAs (family 1: chalcogene AAs: Y & C and M & T; and family 2: carbonyl AAs: E, Q, N)]. (Cf. Figure 5 and Table 5.)

(13)	**W, H**	13	10	Y, C	(8)
(5)	T, M	12	15	E, Q, N	(8)
(10)	*G, P,V,I*	*22*	*25*	*A, L, S, D , F*	(15)
			20	**K, R**	(8)

Table 5. The determination by the number of carbon and hydrogen atoms. The number of atoms in the given sets of amino acid molecules (correspondent to the classification given in Section 5): carbon (in brackets) and hydrogen - within two columns, in the center of the Table (rows *b* and *c* in Table 7, respectively).

[10] Singlet amino acids – because a single amino acid is contained within one stereochemical type (Glycine in glicinic and Proline in prolinic type).

The Family of chalcogene AAs can be further classified; subfamily 1: OH and SH derivatives: Y & C, respectively; and subfamily 2: C-CH_3 and S-CH_3 derivatives: T & M, respectively.

The Family of carbonyl AAs also in two subfamilies classified: subfamily 1: non-nitrogen, E; and subfamily 2: nitrogen AAs, N & Q.

6 VARIABILITY IN RELATION TO POLARITY

Table 6 shows that above presented variability exists in a strict relation to polarity. In column with the designation "more", both invariant AAs (G & P) are semi-polar whereas four variant AAs appear to be in proportion 2:2; two aromatic AAs (W & H) are also semi-polar and two aliphatic (K & R) are polar. On the other hand, in column with the designation "less", we have a symmetric AAs distribution in proportion **5**:**2** / 2:5. Five invariant AAs as well as two variant (bold) are non-polar; in contrary, two invariant as well as five variant AAs (non-bold) are polar. It is important to say that four AAs (G, P & W, H), on the left side in Table 6, we take here as semi-polar through their ambivalence (*see* Footnote 3).

more	less	
G P	**V I**	invariant
	A L F	
	S D	
W H	T N Y	variant
K R	E Q	
	C M	

Table 6. Variability in relation to the polarity. Explanation in the text (Sections 5 and 6)

7 SPECIFIC ATOM NUMBER DETERMINATIONS

7.1 Determination with prime quantum 037

Once we know that out of 20 canonical AAs there are exactly four AAs with higher degree of freedom (more variant) than the remaining 16, then it is possible

to organize the entire system in view thereof, as was shown in Figure 5. Thereby these four AAs should observe fundamental chemical hierarchy – the first ones are aliphatic (K, R), followed by aromatic (W, H) in the order we have in the system presented in Figure 1. Thereafter, there can be no more ambiguities. Two acidic AAs (D & E) follow two basic ones (K & R), followed by amide derivatives (N & Q) and the remaining two source aliphatic AAs (A & L). Yet, in continuation of two giving aromatic AAs (W, H), there are two remaining aromatic ones, two larger aliphatic chalcogenic ones (M, T) and finally two smaller chalcogenic ones (C, S). The position of contact AAs remains the same as in the system in Figure 1.

Atom number in this new system, as we see, is also[11] determinated by Shcherbak's prime quantum 037 (Shcherbak, 1994; 2003), with adequate other balances.

7.2 Determinations through atom number of carbon and hydrogen

A deeper analysis reveals the sense of the balance of number of carbon atoms in the sets of amino acidic molecules, given in Tab. 5. If we mark the numerical basis of a numbering system with the q, then the proportion $q/2 : q = 1: 2$ represents "the symmetry in the simplest case" (Marcus, 1989) for any q. However, only in the case when $q = 10$, therefore in the decimal numbering system, we also have the correspondence with the Golden mean as the best possible proportion and harmony. Namely, in the decimal numbering system, the $q/2$ is 5, and exactly through the square root of number 5 we come to two solutions of the square equation of the Golden mean. Even more, the „quantum 5" appears also within a „p-adic genetic information space" as an adequate „5-adic model for ... genetic code" (Dragovich and Dragovich, 2006).

Knowing this, it becomes understandable why number 5 is the starting quantum here (in the system presenting in Table 5). Starting from the number 5, there are two possible paths which stick to harmony. The first one is the realization of the following steps of Fibonacci sequence (which in itself is in correspondence with the Golden mean), achieving number 8 and number 13, which is the case here. The other one is the realization of multiples of number 5, for example, achieving number 10 and 15, which is also the case here.

---

[11] We say „also", because so far Shcherbak's "Prime quantum 037" was only known as a determinant of the number of nucleons in four-codon and non-four-codon AAs (Figure 1 in Shcherbak, 1994).

Certainly, the sense of the correspondence of the number of carbon atoms in a set of 20 canonical amino acids, with the numbers of Fibonacci sequence **5**, 8 and 13, becomes clear only when we know that the number of carbon atoms in two classes of that set (class I and class II as in Table 7) corresponds to the numbers of Fibonacci sequence 1, 2, 3 and **5** (cf. Table 2, last row, in Yang 2004, p. 1253).

(i)	117	131	121	147	146	149	131	204	181	174	1501
(h)	43	57	47	73	72	75	57	130	107	100	761
(g)	**451**	**653**	**227**	**244**	**239**	**129**	**341**	**137**	**235**	**687**	**3343**
(f)	**432**	**631**	**213**	**225**	**219**	**109**	**319**	**110**	**211**	**661**	**03130**
(e)	**19**	**22**	**14**	**19**	**20**	**20**	**22**	**27**	**24**	**26**	**213**
(d)	00	00	01	02	02	01	00	01	01	03	(11)
(c)	05	09	03	05	06	07	09	08	07	10	69
(b)	03	04	01	03	03	03	04	09	07	04	41
(a)	**10**	**13**	**05**	**10**	**11**	**11**	**13**	**18**	**15**	**17**	**0123**
I	**V**	**L**	**C**	**E**	**Q**	**M**	**I**	**W**	**Y**	**R**	**I**
II	**G**	**A**	**S**	**D**	**N**	**T**	**P**	**H**	**F**	**K**	**II**
(a)	**01**	**04**	**05**	**07**	**08**	**08**	**08**	**11**	**14**	**15**	**81**
(b)	00	01	01	02	02	02	03	04	07	04	26
(c)	01	03	03	03	04	05	05	05	07	10	46
(d)	00	00	01	02	02	01	00	02	00	01	09
(e)	**10**	**13**	**14**	**16**	**17**	**17**	**17**	**20**	**23**	**24**	**171**
(f)	**448**	**436**	**639**	**219**	**217**	**432**	**424**	**213**	**205**	**223**	**3456**
(g)	**458**	**449**	**653**	**235**	**234**	**449**	**441**	**233**	**228**	**247**	**3627**
(h)	01	15	31	59	58	45	41	81	91	72	494
(i)	75	89	105	133	132	119	115	155	165	146	**01234**

Table 7. The most important parameters for 20 canonical amino acids. I. The first class of amino acids, handled by the first class of enzymes aminoacyl-tRNA synthetases; II. The second class of amino acids, handled by the second class of enzymes aminoacyl-tRNA synthetases. Both classes in the AAs pairs as in Rakočević, 1998b (Survey 4); the order is dictated by the number of atoms in the side chain of each first member of an amino acid pair; (a) number of all atoms in the "bodies" i.e. side chains of amino acid molecules; (b) number of carbon atoms in the "bodies", i.e. side chains of amino acid molecules; (c) number of hydrogen atoms in the "bodies" i.e. side chains of amino acid molecules; (d) number of non-hydrocarbon atoms in the "bodies", i.e. in the side chains of amino acid molecules; (e) number of all atoms in the whole amino acid molecules ("head" plus "body"); (f) number of all atoms in the amino acid belonging codons; g = e + f; (h) number of nucleons in the "bodies", i.e. side chains of amino acid molecules; (i) number of nucleons in the whole amino acid molecules ("head" plus "body").

The number of hydrogen atoms (two columns of numbers in the central part of Table 5) is also determined by the multiples of number 5 and only by the final

Fibonacci's number (final in the Fibonacci sequence 1-13), as well as the numbers which are created with the unit change in the first and/or second position of that number (numbers 12 and 22 as 13 – 01 and 12+10, respectively).

7.3 Determinations through atom number of non-hydrocarbon elements

Apart from the determinations with the atom number of carbon and especially the atom number of hydrogen, the determination with atom number of non-hydrocarbon elements appear to be exact and relevant too. Namely, the system of amino acid doublets such as we find in Table 1.1, organized in the system of quartets (pairs of pairs), by preserving original hierarchy (Figure 1), "hides" in itself not only the Gauss' algorithm, but another specific determination which essentially comes down to determination by the sequence of even numbers 0-2-4-6-8 (Figure 6 in relation to row *d* in Table 7): 2 & 4 in outer and 6 & 8 in inner rows (10 in odd and 10 in even rows). [*Remark 3*: As a noteworthy is the fact that the sum of the squares of the number of atoms within side chains of 20 amino acid molecules is determined also with the sequence 0-2-4-6-8; equals exactly 02468: G(01) + A(16) + S(25) + C(25) + D(49) + N(64) + T(64) + P(64) + V(100) + E(100) + Q(121) + H(121) + M(121) + L(169) + I(169) + F(196) + Y(225) + K(225) + R(289) + W(324) = 02468.]

(a)					10	08	11	(b)	09	
S01	T01	L00	A00	G00	**02** (02) 04		01S	01T	01M	01C
D02	E02	M01	C01	P00	**06** (02) 08		02D	02E	02Q	02N
K01	R03	Q02	N02	V00	**08** (04) 04		01K	03R	00L	00A
F00	Y01	W01	H02	I00	**04** (00) 04		00F	01Y	01W	02H
04	07	04	05	00			00G	00V	00I	00P
10+1		**10-1**			10	12	00 (10+1)		00 (10-1)	
(light tones: **10-1**; dark tones: **10+1**)										

Figure 6. Relations of variability and polarity expressed through the number of non-hydrocarbon atoms. Everything is the same as in Figure 1, the only difference is that in the former case all atoms within amino acid side chains are given, and in the latter case only the atoms which are neither the atoms of carbon nor the ones of hydrogen.

8 DETERMINATION WITH THE SEQUENCE OF NATURAL NUMBERS

We can understand the deeper sense of determination with the sequence 0-2-4-6-8, i.e. sequence of even numbers, only when we notice that other sequences generated from the sequence of natural numbers are also in the "game". [By this one must bear in the mind that in GCT not only codons but also amino acids possess a strict order in correspondence to natural numbers sequence, from 0 to 19 (Damjanović, 1998; Damjanović & Rakočević, 2005).] So, *the atom number* in class I of AAs, in side chains of 10 AAs, is determined by the sequence "0123" (row *a* in Table 7, above). On the other side, *the nucleon number* in class II of AAs in whole molecules of 10 AAs, is determined by the sequence "01234" (row *i* in Table 7, down). (Class I and class II of AAs in correspondence to two classes of enzymes aminoacyl-tRNA synthetases.)

The number of atoms within the 32 codons coding for AAs in class II equals 3456 (row *f* in Table 7, down). [The difference 3456 – 1234 = 2222 shows that the number of *atoms* within the codons is greater for 2222 than the number of *nucleons* within corresponding amino acid molecules.] Except that, there is the same number of atoms (3456) in codon nucleotides, within two and two columns of GCT (Figure 7, down). Further, atom number within side chains of 23 amino acid molecules, in a symmetric arrangement through rows and columns in GCT, is determined by the sequence "456/789" (Figure 7, on the right side; *see* also these relations in Verkhovod, 1994, Figure 2).

All these insights come down to the fact that we must comprehend (and understand) that in the genetic code, apart from two known alphabets (4 amino basic and 20 amino acidic molecules), there is also another one, the original and fundamental (hidden?) alphabet – the sequence of natural numbers. What else could be expected if we know that all molecules are made of atoms of chemical elements (within Periodic system) whose structural hierarchy is harmonized with that very sequence of natural numbers.

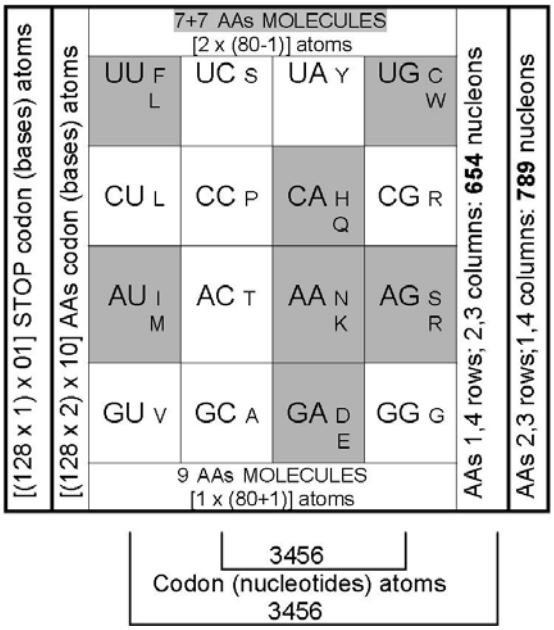


Figure 7. The number of atoms and nucleons. In the GCT the codon doublets code for one or two amino acids. The number of atoms and nucleons in the classes and subclasses of amino acids corresponds to the parts of sequences of natural numbers, with principle of minimum (unit) change. The external vertical row on the left: the number of atoms in three "stop" codons, calculated to bases (U = 12; C = 13; A = 15; G = 16). The internal vertical row on the left: number of atoms in the remaining 61 codons, which have amino acid meaning. Two vertical rows on the right: the number of nucleons in the GCT according to Verkhovod, 1994, Figure 2. Dark tones: double meaning codon doublets, each with two encoded amino acids (total: two times 8-1). Light tones: single meaning codon doublets, each with one encoded amino acid (total: 8+1 molecules, with 81 atoms as in class II, in row *a* of Table 7). The bottom of Figure shows that in two and two columns, i.e. in two and two rows, there are 3456 atoms in the codons, calculated to nucleotides (UMP = 34; CMP = 35; AMP = 37; GMP = 38).

9 CONCLUSION

From the presented facts and given discussion in previous eight sections follows that it make sense to speak about genetic code as a harmonic system. On the other side, presented harmonic structures which appear as the unity and coherence of *form* (atom and nucleon number balances) and *essence*

(physicochemical properties) provide evidence to support the hypothesis, given in a previous paper (Rakočević, 2004), that genetic code was complet**e** from the very beginning as the condition for origin and evolution of the life.

REFERENCES

Damjanović, Z., 1998. Logic core of genetic code. Glasnik Sect. Nat. Sci. Montenegrin Acad. Sci. Art (CANU) 12, 5-8.

Damjanović, M. Z., Rakočević, M. M., 2005. Genetic code: an alternative model of translation. Ann. N. Y. Acad. Sci. 1048, 517-523.

Doolittle, R.F., 1985. Proteins. Scientific American, 253, 74-85.

Dragovich, B., Dragovich, A., 2006. A p-Adic Model of DNA Sequence and Genetic Code. q-bio. GN/0607018.

Falconer, K., 1990. Fractal geometry. John Wiley, New York pp. 34-52.

Jukes, T. H., 1973. Possibilities for the evolution of the genetic code from a preceding Form. Nature 246, 22-27.

Konopel'chenko, B.G., Rumer, Yu. B., 1975. Klassifikaciya kodonov v geneticheskom kode. Dokl. Akad. Nauk. SSSR. 223, 471-474.

Kyte, J., Doollittle, R. F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Marcus, S., 1989. Symmetry in the simplest case: the real line. Comput. Math. Appl. 17, 103-115.

Moore, G.A., 1994. The limit of the golden numbers is 3/2. The Fibonacci Quaterly, June-July, 211-217.

Popov, E. M., 1989. Strukturnaya organizaciya belkov. Nauka, Moscow (in Russian).

Rakočević, M. M., 1997. Genetic code as a unique system. SKC, Niš pp. 60-64.

Rakočević, M. M., 1998a. Whole-number relations between protein amino acids and their biosynthetic precursors. J. Theor. Biol. 191, 463 – 465.

Rakočević, M. M., 1998b. The genetic code as a Golden mean determined system. Biosystems 46, 283-291.

Rakočević, M. M., 2004. A harmonic structure of the genetic code. J. Theor. Biol. 229, 463 – 465.

Rakočević, M. M., Jokić, A., 1996. Four stereochemical types of protein amino acids: synchronic determination with chemical characteristics, atom and nucleon number. J. Theor. Biol. 183, 345 – 349.

Shcherbak, V. I., 1994. Sixty-four triplets and 20 canonical amino acids of the genetic code: the arithmetical regularities. Part II. J Theor. Biol. 166, 475-477.

Shcherbak, V. I., 2003. Arithmetic inside the universal genetic code. Biosystems 70, 187-209.

Swanson, R., 1984. A unifying concept for the amino acid code. Bull. Math. Biol. 46, 187-207.

Verkhovod, A. B., 1994. Alphanumerical divisions of the universal genetic code: new divisions reveal new balances. J. Theor. Biol. 170, 327-330.

Woese, C.R., et al., 1966. On the fundamental nature and evolution of the genetic code. In: Cold Spring Harbor Symp. Quant. Biol., 31, 723-736.

Yang, C.M., 2004. On the 28-gon symmetry inherent in the genetic code intertwined with aminoacyl-tRNA synthetases – The Lucas series. Bull. Math. Biol. 66, 1241-1257.

THE FACTORS OF THE CLASSIFICATION OF PROTEIN AMINO ACIDS

Miloje M. Rakočević

Department of Chemistry, Faculty of Science, University of Niš, Ćirila i Metodija 2, Serbia (E-mail: m.m.r@eunet.yu)

Abstract.

In this work it is shown that three pairs of the factors appear to be the key, i.e. main factors of a natural classification of protein amino acids, i.e. canonical amino acids within the amino acid (genetic) code. First pair: the factors of the **habit** (habitus) of an amino acid molecule (size and polarity). Second pair: the factors of the **association** (type of the amino acid – enzyme reactivity and degree of the hydrophobicity–hydrophilicity of an amino acid molecule). Third pair: the factors of the **dissociation** (degree of the acidity-basicity, over acidic COOH group and degree of the basicity–acidity over the basic NH_2 group). As a result of the influence and interdependence of all six factors (measured through correspondent valid parameters) it appears still one natural classification into polar and non-polar amino acids, where polar amino acids possess negative and non-polar, the positive values of hydropathy index.

Key words: Genetic code, Watson-Crick Table, Canonical amino acids, Polarity, Hydrophobicity-hydrophilicity, Acidity-basicity, The habitus factors, The association factors, The dissociation factors, The First Damjanović's systems, The second Damjanović's system of the amino acids positions.

1. Introduction

Doolittle (1985, p. 76) has shown that a possible classification of protein amino acids (AAs) must be "based on the size of the amino acid's side chain and on the degree to which it is polarized". In this work, however, we start with a *working hypothesis* that three pairs of the factors appear to be the key, i.e. main factors of *a natural classification* of protein amino acids, i.e. canonical amino acids within the amino acid (genetic) code (Tables 1 & 2). First (the Doolittle's) pair: the factors of the **habit** (*habitus*) of an amino acid molecule [1. Size, in relation to the position, i.e. *topos* within genetic code (*t* in Table 2) and 2. Polarity, i.e. hydropathy (*p*)]. Second pair: the factors of the **association** [1. Type of the amino acid – enzyme (aminoacyl-tRNA synthetases) reactivity (*e*) and 2. Degree of the hydrophobicity–hydrophilicity of an amino acid molecule (*h*)]. Finally, as a third pair: the factors of the dissociation [1. Degree of the acidity-basicity, measured over acidic COOH group within the amino acid molecule (*a*) and 2. Degree of the basicity–acidity, measured over the basic NH_2 group (*b* in Table 2)][1].

In a proof proceeding of the working hypothesis, we start from the classification of protein amino acids into ten pairs (Rakočević & Jokić 1996; Rakočević, 1998)[2], as it is presented in Table 1 and Table 2. By this to the class of valine (in six different variants)

[1] The classification into two classes (I and II in Table 1), accordingly with two classes of the enzymes aminoacyl-tRNA synthetases.

[2] Dlyasin (1998) also speaks about ten pairs. Eight pairs are identical by us and by him. The difference in two pairs (By Dlyasin: G-A/V-L, and by us: A-L/G-V) appears because Dlyasin does not include the classification of AAs into four stereochemical types (Popov, 1989; Rakočević & Jokić 1996).

belong AAs designated with "+" whereas to the class of glycine belong AAs designated with "–" (Table 2)[3].

pK_{COOH}	2.3	1.8	2.3	2.4	2.4	2.2	2.2	1.8	2.4	2.2
pK_{NH2}	9.6	10.0	9.2	9.7	9.6	9.7	9.1	9.0	9.4	9.1
N_2	19	14	20	22	22	19	20	26	27	24
N_1	10	05	11	13	13	10	11	17	18	15
Vol.	36	30	52	46	46	41	47	70	83	69
Mass	117.15	121.16	149.21	131.18	131.18	147.13	146.15	174.20	204.23	181.19
H.ph.	0.825	0.680	0.738	0.943	0.943	0.043	0.251	**0.000**	0.878	**0.880**
	+	+	+	+	+	–	–	–	–	–
I	V	C	M	I	L	E	Q	R	W	Y
								|		|
II	G	S	T	P	A	D	N	K	H	F
	–	–	–	–	+	–	–	–	–	+
H.ph.	0.501	0.359	0.450	0.711	0.616	0.028	0.236	**0.283**	0.165	**1.000**
Mass	075.07	105.09	119.12	115.13	089.09	133.10	132.12	146.19	155.16	165.19
Vol.	03	21	32	31	14	30	36	58	50	62
N_1	01	05	08	08	04	07	08	15	11	14
N_2	10	14	17	17	13	16	17	24	20	23
pK_{NH2}	9.8	9.2	10.4	10.6	9.9	10.0	8.8	9.2	9.2	9.2
pK_{COOH}	2.4	2.1	2.6	2.0	2.3	2.0	2.0	2.2	1.8	1.8

Table 1. Two classes of protein amino acids. This Table shows ten pairs of amino acids, classified in accordance to Golden mean and to module 3:2 (Rakočević, 1998). Ten pairs within two classes of protein amino acids, the constituents of the genetic code [the sign "–" for polar and "+" for non-polar amino acids (after: Keyte & Doolittle, 1982 and Doolittle, 1986)]. Class I contains the larger amino acids (larger within the pairs), all handled by class I of enzymes, the aminoacyl-tRNA synthetases. Class II contains the smaller amino acids, all handled by class II of synthetases. Within the rows are given the values for the factors/parameters: hydrophobicity (H.ph.) and molecule mass (Both parameters after: Black & Mould, 1991); then Volume (Swanson, 1984), atom number N_1 as the number of atoms within the amino acid side chain, and N_2 as the number of atoms within the whole amino acid molecule; finely, the values for two constants: pK_{COOH} pK_{NH2} (Cantor & Schimmel, 1980). The correction of the pK_{NH2} value for phenylalanine from 9.1 to 9.2 after: Pine, Hendrickson, Cram & Hammond (1980). Notice an interesting balance in the polarity: in relation to two identical AAs (both non-polar) A-L there are four non-identical pairs on the left. On the other hand, in relation to two non-identical AAs (one polar and other non-polar) F-Y there are four identical pairs (identical from the aspect of polarity) on the left.].

With the working hypothesis it is assumed that from the experimentally determined valid parameters (Table 1)[4], as a measure of the influence of the factors, follow the natural classes of amino acid molecules, such classes as aliphatic and aromatic compounds, chalcogenide derivatives (oxygen and sulphur derivatives), carboxylic amino acids and their amide derivatives etc. On the other hand, it must be that the polar and non-polar amino acid molecules also follow as separate classes.

[3] The valine class is identical with class I, and glycine class with class II in Tables 1 & 2 only in the case of the factor of amino acid - enzyme reactivity (column *e*).

[4] The belonging to two classes of amino acids in all cases is determined by degree – the major and less value of the parameter within the amino acid pair.

I	t	p	h	e	b	a			a	b	e	h	p	t	II
V	+	+	+	+	+	+	6:0	0:6	–	–	–	–	–	–	G
C	+	+	+	+	–	+	5:1	1:5	–	+	–	–	–	–	S
M	+	+	+	+	+	+	6:0	0:6	–	–	–	–	–	–	T
I	+	+	+	+	+	–	5:1	1:5	+	–	–	–	–	–	**P**
L	–	+	+	+	+	–	4:2	3:3	+	–	–	–	+	+	**A**
E	–	–	+	+	+	–	3:3	1:5	+	–	–	–	–	–	D
Q	–	–	+	+	–	–	2:4	2:4	+	+	–	–	–	–	N
R	±	–	–	+	+	+	4:3	1:5	–	–	–	+	–	–	K
W	+	–	+	+	–	–	3:3	2:4	+	+	–	–	–	–	**H**
Y	+	–	–	+	+	–	3:3	3:3	+	–	–	+	+	–	**F**

Table 2. Three pairs of amino acid classification factors. The plusses and minuses designate the degree of the values ("+" more; "–" less) for the parameters: a. Acidity-basicity through pK_{COOH} ("+" more acidic; "–" less acidic; b. Basicity- acidity through pK_{NH2} ("+" more acidic; "–" less acidic); e. Amino acid relation to class I ("+") and class II ("–") of enzymes aminoacyl-tRNA synthetases; h. hydrophobicity (Black & Mould, 1991); p. Polarity for polar ("–") and nonpolar ("+") amino acids; t. "Topos" – the position within Watson-Crick Table: within inner three diagonals ("–"), and within outer four diagonals ("+"). Notice that within glycine class there is 66 molecules (66 minuses) and within valine class 55 molecules (55 plusses). If so, then the sum equals $121 = 11^{2}$, whereas the difference equals $11 = 11^{1}$.

2. TWO CLASSES OF AMINO ACIDS / ENZYMES

2.1. AA - enzyme reactivity and AA molecule size

All AAs within the row I in Table 1, and within column I in Table 2, are handled by class I enzymes aminoacyl-tRNA synthetases **(**plusses in column *e*), whereas all AAs within row/column II are handled by class II synthetases **(**minuses in column *e*). "The class I enzymes chare with dehydrogenases and kinases the classic nucleotide binding fold called the Rossmann fold, whereas the class II enzymes possess a different fold, not found elsewhere, built around a six-stranded antiparallel beta-sheet. The two classes differ … as to where on the terminal adenosine of the tRNA the amino acid is placed: class I enzymes act on the 2' hydroxyl whereas the class II enzymes prefer the 3' hydroxyl group." (Eriani et al., 1995 p. 499). On the other hand, the "class II synthetases are considered to be the more primitive of the synthetases" (Hartman, 1995, p 541).

Except to two classes of enzymes aminoacyl-tRNA synthetases, the plusses/minuses in column *e* are related to the molecule size from three aspects: molecule mass, volume and the number of particles within AA molecule (the number of atoms including the size of each atom, and the number of nucleons[5]; *see* legend to Table 1). In particular, that means that within each pair, the amino acid of class I is a larger, whereas of class II a smaller molecule. By this, bearing in mind that Class II of enyzmes is "more primitive" (Hartman, 1995, p 541), we can conclude that smaller molecules respond to more primitive enzymes, going from one to another AA pair (10 second members within 10 pairs).

[5] Only in the case of the pair C-S it exists, in the first slight, an uncertainty because both side chains possess the same atom number (5 atoms). However, the sulphur atom is larger than the oxygen atom.

2.2. A splitting within two AA classes by the module 3/2

The system of 20 canonical AAs, presented in Table 2 represents a perfect coherence between the physico-chemical properties of AA molecules and the balance of the number of atoms existing within the molecules. The balance appears to be self-evident if one can reveal the existence of a specifying splitting into two classes by module 3/2. [*Note about symmetry and harmony*. In a previous work (Rakočević, 1997), we have shown that the module 3/2 and/or 2/3 is the basic fractal motive within amino acid (genetic) code. The relation 2:3 and 3:2 are unique within the system of natural numbers series, because 2 is harmonic mean of the adherent number 3 and its half. On the other hand the number 3 represents a sum of the precursor 2 and its half. By all these relations one must bear in mind that the existence of a whole and its half is a condition for "the symmetry in the simplest case" (Marcus, 1989)]. The balance realizes itself through *four penthaplets*[6] and over a specific crossing (where the crossing exists also in the module 3/2) as follows. The number of atoms within 3+2 and 3+2 of AA molecules in upper half of the system differs for ±1 of atoms in relation to the arithmetic mean (first penthaplet VCM/PA with 39 -1 and second penthaplet GST/IL with 39+1 atoms). In lower half of the system within 3+2 and 3+2 molecules the difference is ±0 of atoms in relation to arithmetic mean (third penthaplet EQR/HF with 63±0 and fourth penthaplet DNK/WY with 63±0 of atoms also). Altogether, on the first zigzag line (invisible line, which connects AAs designated bold in Table 2) there are 102–1, whereas on the second zigzag line (AAs designated non-bold) there are 102+1 of atoms. [*Note.* The number of nucleons on the first line is 628 + 10 and on the second line: 627 – 10. Cf. with nucleon number in Surveys 1-4. The nucleon number in all cases is as in Shcherbak (1993,1994), that means calculated only for the first nuclide within all five bioelements (H, C, N, O, S)][7]. In a previous paper, however, we have shown (Koruga, Rakočević et al, 1997) a different splitting into 10+10 of AA molecules, also in module 3/2, (and also with a crossing) with two zigzag lines, each line not with 102±1 but with 102±0 of atoms. By that, the vertically arranged amino acid doublets and triplets make not the horizontally arranged pairs.

3. The interdependence among position, size and polarity

The column *t (topos)* in Table 2 is related to position of AAs within Watson-Crick Table (WCT) of the genetic code. By the sign "–" are designated AAs which exist within

[6] Notice that in the starting AA system (Table 2) we have four AA pentaplets, whereas at the ending AA system (Table 4; Section 4.2) – five quadruplets, what means a specific balance also.

[7] This arithmetical balance exists in a correspondence with chemical characteristics of AA molecules within two and two pentaplets. Therefore, within first two pentaplets, the doublets (PA and IL) are aliphatic or alicyclic AAs, whereas within second two pentaplets, the doublets (HF and WY) are aromatic AAs. The triplets existing within first two pentaplets (VCM and GST) contain two (ST) and two (CM) chalcogenide AAs, what means more diversity (OH and SH groups in relation to COOH or NH_2 groups). On the other hand, the triplets existing within second two pentaplets (EQR and DNK) contain two carboxylic AAs and their amide derivatives, plus one (R) and one (K) basic AAs. In first case in the 'game' is COOH group, and in second case NH_2 group, both existing within the AA 'head', what means less diversity. Finally, there are G-V in first two pentaplets without CH_2 group between the 'head' and side chain, and K-R in second two pentaplets with some CH_2 groups between the 'head' and side chain.

inner space of WCT (**i**), that means within three middle diagonals, whereas by the sign "+" are designated AAs which exist within outer space of WCT (**o**), that means within two upper and two lower diagonals. The most AAs within outer space are non-polar (**n**) with the exceptions of Y, W and R, which are polar. On the other hand, the most AAs within inner space are polar (**p**) with the exceptions of F and L, which are non-polar.

These strictly arranged positions of AAs within WCT (column *t* in Table 1: plus for outer and minus for inner AAs) appear to be in a strong correspondence with the polarity of AA molecule and its size (atom number) through the validity of the principle of minimum change as follows:

(n) 4V+1M+3I+4A+**2L**+**4L**+**2F**+2C = 22 molecules

40+ 11+39+16+26 + 52 +28+10 = 2**2**2 atoms (42**0**) (1)

(o) 4V+1M+3I+4A+**2Y**+**4R**+**1W**+2C = 21 molecules

40+ 11+39+16+ 30 + 68 +18+ 10 = 2**3**2 atoms (42**1**) (2)

(p) 4G+2K+2N+4P+**2Y**+**4R**+**1W**+2E+2D+4T+2R+2S+2Q+2H+4S = 39

04+30+16+32+30 + 68+ 18+ 20 +14+ 32+ 34+10+22+22 + 20 = 3**7**2 (72**3**) (3)

(i) 4G+2K+2N+4P+**2L**+**4L**+**2F**+2E+2D+4T+2R+2S+2Q+2H+4S = 40

04+30+16+32+26+52 + 28+20+ 14+32+ 34 +10+22+ 22+ 20 = 3**6**2 (72**2**) (4)

From the above presented equations follows that in 22 molecules of non-polar AAs there are 222 of atoms within their side chains, and 420 of atoms within the whole molecules; in outer space, however: 01 molecule less, 10 of atoms more within side chains and 01 atom more within the whole molecules. On the other hand, in 39 molecules of polar AAs there are 372 of atoms within their side chains and 723 of atoms within the whole molecules[8]; in inner space, however: 01 molecule more, 10 of atoms less within side chains and 01 atom less within the whole molecules. In reality, all these balances are directed by the "hidden" balances of the number of atoms within the AAs-exceptions in this manner: (2Y+1W = 48 atoms within side chains) + (4R = 68) = **7** molecules and 17**9** atoms; (2F = 28) + (6L = 78) = **8** molecules and 17**8** atoms.

4. THE NATURAL CLASSIFICATION OF AAS

4.1. The classification into polar and nonpolar AAs

The belonging of AAs to two classes (plusses for valine class and minuses for glycine class in Table 2) generates a new symmetric system correspondent with the natural numbers series: 3, 4, 5, 6 (Table 3)[9].

[8] Notice the cyclity of the permutations (372 and 723), which cyclity is found also by the balances in the nucleon number (by classification into four-codon and non-four-codon AAs) (Shcherbak, 1993, 1994).

6	(-)	**G**	**T**	—	—
5	(-)	**S**	**D**	**P**	**K**
4	(-)	**Q**	**N**	**H**	—
3	(±)	**E**	**Y**	**W**	**R**
3	(±)	A	F	—	—
4	(+)	L	—	—	—
5	(+)	C	—	I	—
6	(+)	V	M	—	—

IV	III	II	I
G-V	**T-M**	**P-I**	**K-R**
S-C	D-N	H-W	
Q-E	Y-F		
A-L			

Table 3 (left). Amino acid relations. The number of plusses and minuses follows from Table 2. The positions follow from the position in Table 2 as well as from the physico-chemical properties. Bold: polar, and non-bold: nonpolar amino acids.
Table 4 (right). Four Amino acids classes. If amino acids pairs must be as in Rakočević (1998) then four classes follow from the system presented in Table 3. Non-bold: amino acids of the alanine stereochemical type, i.e. the non-etalon amino acids; bold: etalon amino acids of glycine stereochemical type (G), proline stereochemical type (P) and valine stereochemical type (V, I). About etalon and non-etalon amino acids see in: Rakočević & Jokić (1996).

At the same time this generated system of plusses and minuses (as result of the influence of three pairs of the factors, i.e. parameters) exists in a strict interdependence with polar (bold in Table 3) and non-polar AAs. [*Note about plusses and minuses*: To be "positive," that means to be more "out," and to be "negative" that means to be more "in." Therefore, the larger molecules within class I in Table 1 are more "out" while smaller molecules within class II are more "in." The non-polar AAs (designated with plusses in Table 1) are more "out", that means near to the sphere from the aspect of electron density; the polar AAs are more "in", that means further from the said sphere. The more hydrophobic (less hydrophilic) AAs are more "out," that means not in contact with the water; the less hydrophobic (more hydrophilic) AAs are more "in," that means in close contact with water. The more acidic (less basic) AAs, from the aspect either COOH or NH_2 group, are more "out" because the hydrogen ion, i.e. proton is leaving; in contrary, the more basic (less acidic) AAs are more "in", that means designate with the sign minus].

4.2. The classification into five quadruplets and four classes of AAs

Bearing in mind that AA pairs are such pairs as we have been found within Class I and Class II in Tables 1 & 2, and knowing that a strict arrangement-order of AAs which we found in Table 3 is result of their physico-chemical properties, then unambiguously follows the arrangement-order of AAs given in Table 4.

Within first row of Table 3 must be only two AAs (with 6 minuses): G in first column and T in second column. Within second row there must be four amino acids (5 minuses): S in first column as a derivative of G (atom H is substituted with an OH functional group) and D in second column as an analog of T. [Aspartic acid as HOOC-CH_2-CH(NH_2)-COOH and threonine as HO-$CHCH_3$-CH(NH_2)-COOH. Cf. side chains: in first case **HO**-CO and in second case **HO**; then H-C-H in first case and H-C-CH_3 in second case].

[9] One can notice here (for the sequence 3-4-5-6) the validity of the Pythagorean law ($3^2+4^2=5^2$) and Plato's law ($3^3+4^3+5^3 = 6^3$). About the fact that nucleon number within AA molecules (within four-codon AAs) is determined by the Pythagorean law, see in Shcherbak (1994, Fig. 1).

Further, it must follow P with **three** CH_2 groups and, finally, K with **four** CH_2 groups. Within third row there must be three AAs (with 4 minuses): N must go with D as its amide derivative; H with P (heterocyclic AAs) and Q with S. (The NH_2 amide group in glutamine comes as a result of a substitution of an OH group, which group exists in serine too).

If within first three row there are only polar amino acids, then within fourth row there must be also polar AAs – exactly four of them: E with Q (as carboxylic AA and its amide derivative), W with H (heterocyclic aromatic AAs) and R with K (two basic AAs). If so, then Y must go with N, D and T (the connection goes through OH function).

All other AAs are non-polar: A and F with 3 plusses, L with four, C and I with five, V and M with six plusses. If alanine, as first, must be in first column, phenylalanine must be in second column (together with tyrosine, which is its derivative). Leucine as aliphatic must be with alanine. Then comes C–I and V–M. Where? From pairing process, represented in Table 1, we understand that C must go into the column of S; I in column of P; V in column of G and, finally, M in column of T.

From above follows the next arrangement-ordering: one pair in fourth column in Table 3; one outer (P-I) and one inner pair (H-W) in third column; one outer (T-M) and two inner pairs (D-N and F-Y) in second column; two outer (G-V and S-C) and two inner pairs (E-Q and A-L) in first column. Altogether, an unambiguous arrangement-ordering as it is represented in Table 4.

Reading from a diagonal arrangement we find *four quadruplets* (see footnote 6) of 16 AAs of alanine stereochemical type in an original natural classification. Within first diagonal there are two quadruplets: aliphatic AAs (A-L and K-R), as 1^{st} quadruplet and aromatic AAs (F-Y and H-W), as 2^{nd}. Within second diagonal are carboxylic AAs, i.e. carboxylic AAs and their amide derivatives (D-N and E-Q), as 3^{rd} quadruplet. Finally, within third diagonal exist chalcogenide amino acids (S-C and T-M), as 4^{th} quadruplet.

Bearing in mind that side chain represents a 'copy' of the AA 'head' we can speak about AAs which possess less diversity and about AAs which possess more of the diversity. Thus, the four quadruplets appear to be two octets: first octet with less diversity (aliphatic and carboxylic AAs, i. e. 1^{st} and 3^{rd} quadruplets, with 86-1 of atoms within their side chains); and second with more diversity (aromatic and chalcogenide AAs, i.e. 2^{nd} and 4^{th} quadruplets, with 86+1 of atoms within AA side chains)[10]. [*Note about diversity*. The 'copy' of a COOH or NH_2 (or NH) group means less diversity, whereas of an OH group means more diversity, because the OH group generates a very new type of organic compounds].

In a vertical-horizontal reading of the system presented in Table 4 we find four classes of AAs arranged hierarchically from two aspects. In first case, hierarchically from the aspect of the correspondence to the series of natural numbers 1, 2, 3, 4 in the next sense: 1 pair within first column, 2 pairs within second column, 3 pairs within third column and 4 pairs within fourth column. In second case, hierarchy follows from the aspect of the possession of a functional group. So, within 4^{th} column we have less-more-less complex

[10] First octet with less diversity possesses less nucleons within AA side chains (506), whereas second octet with more diversity possesses more of nucleons (607). Notice that digital patterns of the nucleon number notations (506-607) is a noteworthy fact in the respect of discussed particles balances, the principle of minimum change and in an analogy with the quantum physics (607 – 506 = 0101)(Shcherbak, 1994, p. 476: "The laws of additive-position notation of numbers ... have analogies with quantum physics").

molecules. But within 3rd column it exists a vice versa situation: more-less-more complex molecules. On the other words, aliphatic pair A-L is less complex than aromatic F-Y; the carboxylic pair E-Q is more complex than D-N; finally, the chalcogenide S-C pair is less complex than T-M. [*Note*. In relation to this last point (T-M) there are two pairs of etalon amino acids, existing within the remaining three stereochemical types: G-V as less complex on the left, and P-I as more complex on the right].

Within 2nd and 1st column, the functions realize themselves as more-more complex, the more complex aromatic H-W (more complex than aromatic F-Y) and the more complex aliphatic K-R (more complex than aliphatic A-L). With the comparison of 4th and 2nd column, one can understand the position of 5th quadruplet, the etalon quadruplet, with the functions less-less and more-more respectively. Namely, In the beginning of 4th column there is the less complex aliphatic pair (A-L), and at the end the less complex etalon pair (G-V). On the other hand, in the beginning of 2nd column, there is the more complex aromatic pair (H-W) and at the end, the more complex etalon pair (P-I). [About four stereochemical types and etalon AAs, see in Rakočević and Jokić (1996)].

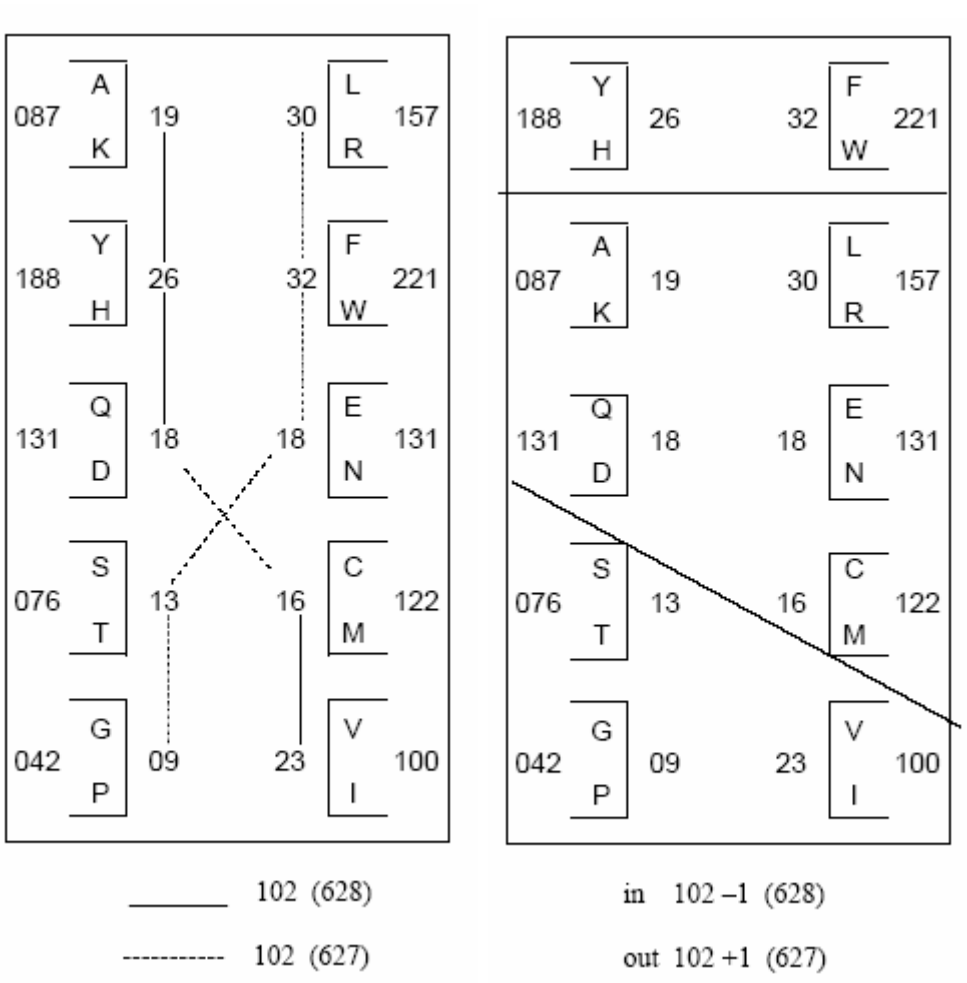
4.3. The splitting within five quadruplets by the module 3/2 (I)

From the system, presented in Table 4, follow two step-by-step AA orders. The first, when the sequencing process begins with outer quadruplet from first diagonal (Model 1 in Survey 1); and the second, when the sequencing process begins with the inner quadruplet from the first diagonal (Model 2 in Survey 2). In Survey 1 it is shown that five quadruplets from Table 4 can be split into two groups through a crossing and by [2 x (3+2)] of pairs in each. In such a manner, all five quadruplets are splitting into two halves with 1:1 pair (reading vertically) and with a strict balance in the particles number: (102 & 102) atoms and (628 & 627) nucleons on each of two zigzag lines. In Survey 2, however, it is shown that five quadruplets can be splitting also into two groups by [2 x (3+2)] of vertical pairs, but into one inner and one outer group with (102±1) of atoms and (628 & 627) of nucleons within the groups.

5. THE INTERDEPENDENCE BETWEEN POSITIONS AND FACTORS

5.1. A symmetric arrangement-ordering of AA quadruplets

All classification and balances presented in four previous Sections represent interdependence between AA positions within WCT and the influence of three pairs of the factors. By this the word is about the positions determined by one and only distinction: to be within inner or within outer space in WCT (cf. the equations 1 – 4). In some way, this is very surprisingly knowing that it is possible to establish the positions for all canonical AAs within WCT, minimum in two manner: as in First Damjanović's system (DS1) – the positions from 0 to 19 (Fig. 2 in Damjanović, 1998) and as in Second Damjanović's system (DS2) (*personal comunication*) – the positions from 1 to 24 (Fig. 1 in this paper).



Survey 1 (left). The splitting of AA quadruplets (Model 1); **Survey 2 (right).** The splitting of AA quadruplets (Model 2)

In such a situation, it makes sense to research a possible interdependence among two Damjanović's systems and our system presented in Table 4. So, Table 5 represents the connection between DS1 and our system. [At Fig. 2/a in Damjanović (1998) the position number for an one-meaning AA, such as alanine (A) equals 12_4 or 6_{10}; for methionine (M) 30_4 or 12_{10} etc. for other one-meaning AAs; the position number for a two-meaning AA, such as arginine (R), represents a middle value for two numbers: 1/2 of (20_4 + 21_4) or 1/2 of (08_{10} + 09_{10}) etc. for other two-meaning AAs]. On the other hand, Table 6 represents the connection between DS2 and our system.

As we can see, the results of two connections are two very symmetrical systems, first in Table 5 and second in Table 6. The middle position in both possesses the quadruplet of aromatic AAs. Out of its there are 4±0 outer pairs within first and 4±1 within second system, all (outer) pairs arranged in four quadruplets. So, at the arrangement-ordering within first system (Table 5) the quadruplet of aliphatic AAs is above and quadruplet of carboxylic AAs (and their amide derivatives) below; then follow two quadruplets in middle positions: etalon AAs exactly very close to aromatic AAs, and one step further – the chalcogenide AAs.

K	Q	E	*	N	H	D	Y
1	2	3	4	17	18	19	20
T	P	A	S				
5	6	7	8				
R	G	L	V				
9	10	11	12				
R	W	M	L	S	C	I	F
	*	I					
13	14	15	16	21	22	23	24

Figure 1. The amino acid order after the second Damjanović's model (personal communication)

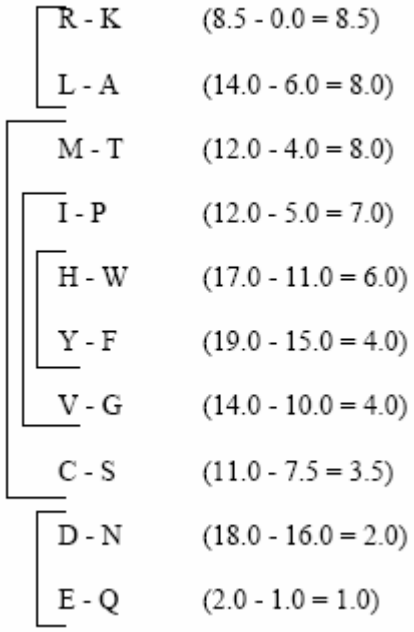
Within second system (Table 6), the arrangement-ordering realizes itself with one crossing step within four outer quadruplets. By this, carboxylic quadruplet is also below, but above exists a vice versa situation in relation to the system in Table 5: I-P, M-T, L-A versus L-A, M-T, I-P etc.

5.2. The splitting within five quadruplets by the module 3/2 (II)

For further comparison of systems DS1 and DS2 with our system, we must make a rearrangement of first system (Table 5) and of second system (Table 6); such a rearrangement in which on the top must be aromatic AAs and other quadruplets must go pair-by-pair and step-by-step as it is shown in Surveys 3 & 4 (Models 1' and 2', respectively). In the comparison of Models 1 & 1' and then of Models 2 & 2' we see that the balance of atom number is changed, accordingly to the principle of the minimum change: from 102 ± 0 to 102+1, whereas the balance of nucleon number is not changed (628 & 627); in second case the balance is the same, except of a change in the direction: the state *in/out* within Model 2 is changed in the state *up/down* within Model 2'.

5.3. The distinction by the module 3/2

If we designate the full zigzag line with roman number I and dotted zigzag line with II, then very interesting relations follow. In the comparison of the subsystem 1(I) in Model 1 with the subsystem 2(in) in Model 2, we see that three pairs are the same and two different. The pairs A-K, Q-D and C-M appear in both subsystems. By this within the subsystem 1(I) exist still Y-H / V-I with 49 atoms and 288 nucleons, whereas within the subsystem 2(in) exist E-N / L-R with 48 atoms and 288 nucleons. On the other hand, in comparison of 1(II) with 2(out) we see that also three pairs are the same and two different. So, the pairs F-W, S-T and G-P appear in both subsystems. By this within subsystem 1(II) exist still L-R / E-N with 48 atoms and 288 nucleons, whereas within the subsystem 2(out) exist V-I / YH with 49 atoms and 288 nucleons.



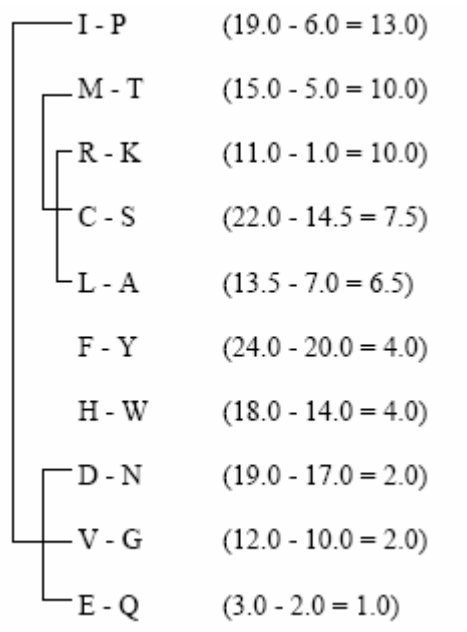
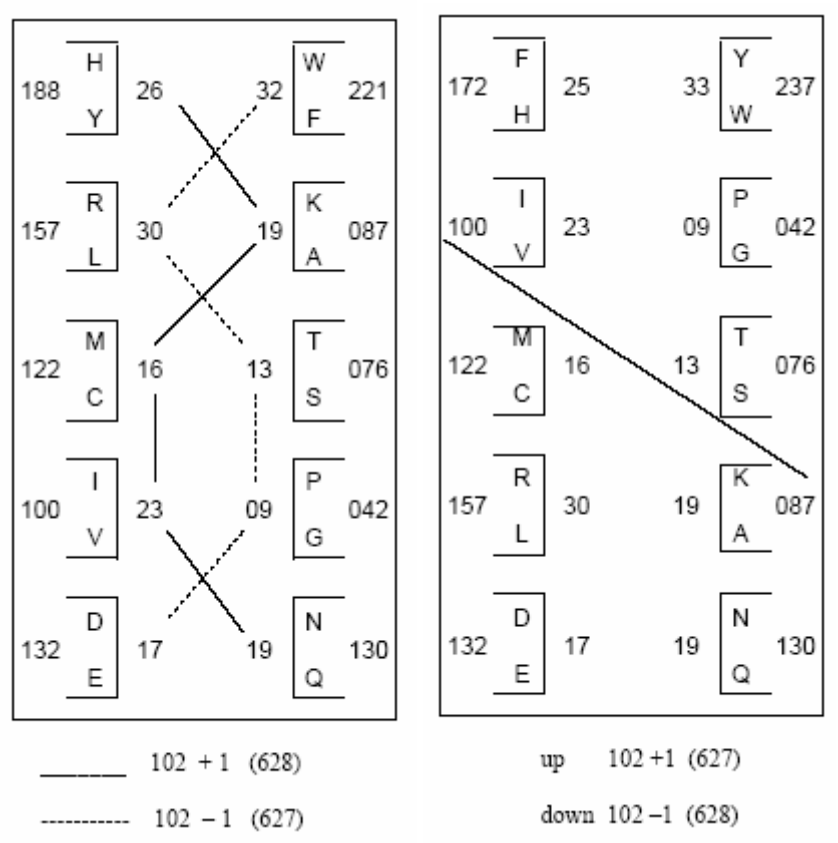


Table 5 (left). Five amino acid quadruplets (System I). In the middle are four aromatic amino acids. Above: four aliphatic, A-L as source aliphatic amino acids and K-R as amine derivatives. Below: carboxylic amino acids with their amide derivatives (altogether: four). Within the first concentric cycle: four etalon amino acids. Within second concentric cycle: four chalcogenide amino acids. The order as well as the numbers follow from the First Damjanović's system (Damjanović, 1998). Notice that in relation to aromatic amino acids, four pairs are above and four below (4±0 in relation to the arithmetic mean).
Table 6 (right). Five amino acid quadruplets (System II). All quadruplets as in Table 5, instead the order and numbers that follow from Second Damjanović's system (Figure 1). Notice that in relation to aromatic amino acids, five pairs are above and three below (4±1 in relation to the arithmetic mean).

However, with a cross in comparison occures a vice versa situation: two pairs of AAs are the same and three different. So, from the comparison of two subsystem 1(I): 2(out) distinctions are as follows: within the subsystem 1(I) A-K / Q-D / C-M with 52+1 atoms and 340 nucleons; within subsystem 2(out) G-P /F-W /S-T with 52-1 atoms and 339 nucleons. From the comparison 1(II): 2(in) the same distinctions follow, but in a vice versa order: within subsystem 1(II) there are G-P / F-W / S-T and within subsystem 2(in) there are A-K/Q-D/C-M.

Identical results can be obtained with a comparison of subsystems contained within Models 1' i 2' with only and one difference: instead the pair E-N the pair E-D appears, and instead the pair Q-D it appears the pair Q-N (the change for 1 atom and 1 nucleon).



Survey 3 (left). The splitting of AA quadruplets (Model 1'); **Survey 4 (right).** The splitting of AA quadruplets (Model 2')

6. CONCLUSION

The argumentation given through five previous Sections of this paper provide evidence to support the working hypothesis, given in the Introduction, that three pairs of the factors, and corresponding parameters appear to be key determinants of a natural classification of protein amino acids. These three pairs unambiguously are: 1. Size and polarity of the AA molecule, 2. Type of the amino acid – enzyme reactivity and degree of the hydrophobicity–hydrophilicity of an amino acid molecule, and 3. Degree of the acidity-basicity, measured over acidic COOH group within the amino acid molecule, and degree of the basicity–acidity, measured over the basic NH_2 group. On the other hand, the natural amino acid classes are: 1. Aliphatic AAs, 2. Aromatic AAs, 3. Chalcogenide AAs and 4. Carboxylic amino acids and their amide derivatives – all four classes within alanine stereochemical type, and 5^{th} class etalon AAs (G-P, V-I)("etalon": after Rakočević and Jokić, 1996) existing within glycine stereochemical type, proline stereochemical type, and valine stereochemical type, respectively. Independent of this classification, from the influence and interdependence of all six factors, follows still one natural classification into two classes: 1. Polar AAs, and 2. Non-polar AAs (after positive and negative values of hydropathy index).

ACKNOWLEDGEMENTS

I am grateful to my professors Bojana Grujić-Injac and Zvonimir Damjanović, and to my colleges Radivoje Papović, Aleksandar Tančić and Djuro Koruga, also to my postgraduate and doctoral studies students Lidija Matija, Nataša Mišić, Tanja Andjelković and Mateja Opačić, for stimulating discussions and benevolent critique.

REFERENCES

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

Cantor, C. R., Schimmel, P. R. (1980) *Biophysical chemistry*, W.H. Freeman and Comp., San Francisco.

Damjanović, Z. (1998). Logic core of genetic code. *Glasnik of the Section of natural sciencies of Montenegrin Academie of Sciences and art (CANU).* **12**, 5-8.

Dlyasin, G.G. (1998) *Azbuka Germesa Trismegista, ili molekulyarnaya tainopis' myshleniya*, Belye Al'vy, Moskva.

Doollittle, R. F. (1985). Proteins. *Scientific American*. **253**, 74-85.

Doollittle, R. F. (1986). *Of URFS and ORFS – a primer on how to analyze derived amino acid sequences* p. 53. Oxford university press, Oxford.

Eriani, G., et al. (1995). The class II aminoacyl-tRNA synthetases and their active-site-evolutionary conservation of an ATP binding-site. *J. Molec. Evol.* **40**, 499-508.

Hartman, H. (1995). Speculations on the origin of the genetic code. *J. Mol. Evol*. **40**, 541-544.

Koruga. Dj., et al. (1997). A new classification of protein amino acids by module 3/2. *Archive of Oncology*. **5**, 137-138.

Kyte, J., Doollittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132.

Marcus, S. (1989). Symmetry in the Simplest Case: the Real Line, *Computers Math. Applic.* **17**, 103-115.

Pine, S. H., Hendrickson, J. B., Cram. D. J., Hammond, G. S. (1980) *Organic chemistry, fourth edition*, McGraw-Hill Kogakusha, Tokyo.

Popov, E.M. (1989) *Strukturnaya organizaciya belkov*, Nauka, Moskva.

Rakočević, M.M. (1997) *The genetic code as a unique system*, SKC, Niš.

Rakočević, M.M. (1998) The genetic code as a Golden mean determined system, *Biosystems*, **46**, 283-291.

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

Shcherbak, V.I. (1993). Twenty canonical amino acids of the genetic code: the arithmetical regularities. Part I. *J Theor. Biol.* **162**, 399-401.

Shcherbak, V.I. (1994). Sixty-four triplets and 20 canonical amino acids of the genetic code: the arithmetical regularities. Part II. *J Theor. Biol.* **166**, 475-477.

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

GENETIC CODE: A NEW UNDERSTANDING OF CODON – AMINO ACID ASSIGNMENT

Zvonimir M. Damjanović* and Miloje M. Rakočević**

**Montenegrin Academy of science and arts (CANU), Podgorica, Montenegro*
*** Department of Chemistry, Faculty of Science, University of Niš, Serbia*
(E-mail: m.m.r@eunet.yu)

Abstract.

In this work it is shown that 20 canonical amino acids (AAs) within genetic code appear to be a whole system with strict AAs positions; more exactly, with AAs ordinal number in three variants; first variant 00-19, second 00-21 and third 00-20. The ordinal number follows from the positions of belonging codons, i.e. their digrams (or "doublets"). The reading itself is a reading in quaternary numbering system if four bases possess the values within a specific logical square: A = 0, C = 1, G = 2, U = 3. By this, all splittings, distinctions and classifications of AAs appear to be in accordance to atom and nucleon number balance as well as to the other physico-chemical properties, such as hydrophobicity and polarity.

K e y w o r d s: Genetic code, Genetic code table, Translation, Numbering system, Spiral model of Genetic code, Canonical amino acids, Logical square, Hydrophobicity, Polarity, Hydropathy, Perfect numbers, Friendly numbers.

1 INTRODUCTION

Gamow was the first who attempted to resolve the problem of codon – amino acid assignment, i.e. to answer the question how 64 codons can have the possible meanings for 20 canonical amino acids (AAs). His solution was the so called "diamond code" (Gamow, 1954) in which 64 codons were classified into 20 classes, corresponding to 20 AAs (Hayes, 1998: "Symmetries of the diamond code sort the 64 codons into 20 classes ... All the codons in each class specified the same amino acid"; cf. legends to Figures 2 and 5 in cited paper). Unfortunately, the real genetic code (Crick, 1966, 1968; Patel, 2005) appears to be less regular than Gamow's. Indeed, the experimentally observed results showed that the codon – amino acid assignment realizes through the relationships very different of those, postulated by the diamond code (Tables 1 and 2 in relation to Table 3).

Table 1
Amino acid properties

Amino acid	R-group property	Mol. weight	Class	Secondary propensity
Gly	Non-polar	75	II	turn
Ala	aliphatic	89	II	α
Pro		115	II	turn
Val		117	I	β
Leu		131	I	α
Ile		131	I	β
Ser	Polar	105	II	turn
Thr	uncharged	119	II	β
Asn		132	II	turn
Cys		121	I	β
Met		149	I	α
Gln		146	I	α
Asp	Negative	133	II	turn
Glu	charge	147	I	α
Lys	Positive	146	II	α
Arg	charge	174	I	α
His	Ring/	155	II	α
Phe	aromatic	165	II	β
Tyr		181	I	β
Trp		204	I	β

Table 1. Amino acid properties. This Table is downloaded from Table 1 in Patel (2005): "Properties of amino acids depend on their side chain R-groups. Larger molecular weights indicate bigger side chains. The 20 naturally occurring amino acids divided into two classes of 10 each, depending on the properties of aminoacyl-tRNA synthetases that bind the amino acids to tRNA. The dominant properties of amino acids for forming secondary protein structures are also listed".

At the same time when Crick brought out two possible hypotheses on the codon – amino acid assignment (genetic code was frozen in an evolution process on the level "64 codons : 20 AAs", or this ratio is result of the stereochemical conditions), a special understanding came from Y. Rumer (Rumer, 1966; Konopel'chenko & Rumer, 1975) [Rumer, 1966, p. 1393: "Considering the group of codons, that relates to one and the same amino acid, shows that within

every codon (z | yx) (it should be read from right to left side) it is expedient to separate two-letter 'root' | yx) of the 'end' (z |. So, every amino acid, in a general case, has a corresponding and specific root, and degeneration of the code appears as consequence of exchanging of the endings."]. Unfortunately, this understanding was forgotten two next decades, when it was restored by Shcherbak still once (Shcherbak, 1989, 1993, 1994).

Table 2
Mitochondrial genetic code

UUU Phe	**UCU Ser**	UAU Tyr	UGU Cys
UUC Phe	**UCC Ser**	UAC Tyr	UGC Cys
UUA Leu	**UCA Ser**	UAA Stop	UGA Trp
UUG Leu	**UCG Ser**	UAG Stop	UGG Trp
CUU Leu	CCU Pro	**CAU His**	CGU Arg
CUC Leu	CCC Pro	**CAC His**	CGC Arg
CUA Leu	CCA Pro	CAA Gln	CGA Arg
CUG Leu	CCG Pro	CAG Gln	CGG Arg
AUU Ile	ACU Thr	**AAU Asn**	AGU Ser
AUC Ile	ACC Thr	**AAC Asn**	AGC Ser
AUA Met	ACA Thr	AAA Lys	AGA Stop
AUG Met	ACG Thr	AAG Lys	AGG Stop
GUU Val	GCU Ala	**GAU Asp**	GGU Gly
GUC Val	GCC Ala	**GAC Asp**	GGC Gly
GUA Val	GCA Ala	GAA Glu	GGA Gly
GUG Val	GCG Ala	GAG Glu	GGG Gly

Table 2. Mitochondrial genetic code. This Table is downloaded from Table 2 in Patel (2005): "The (vertebrate) mitochondrial genetic code differs slightly from the universal genetic code. The wobble rules are exact for the mitochondrial code, so the third codon position has only a binary meaning. Class II amino acids are indicated by boldface letters". (Note: The differences in standard genetic code: AUA – Ile, UGA – Stop, AGA – Arg, AGG – Arg).

2 PRELIMINARIES

Our today's understanding of codon–amino acid assignment relies on the said Rumer's conception; on the other words, our understanding follows, in principle, from a specific manner of reading the codons and their digrams, i.e. „doublets", in which AAs are given in the strict ordinal numbers through three variants (presented in this paper in quaternary and/or in decimal numbering system).

First variant with ordinal number of AAs **00-19** (cf. Damjanović, 1998, and Appendix 1, Surveys 1 & 2 in this paper): 00 K; 01 Q; 02 E; 03 ⊗; 04 T; 05 P; 06 A; 07 S; 08 S, R; 09 R; 10 G; 11 C, W, ⊗; 12 I, I[1], M; 13 L; 14 V; 15 F, L; 16 N; 17 H; 18 D; 19 Y; **Second variant** with ordinal number of AAs **00-21** (cf. Solutions 1- 4 in Section 3, Figure 1 and Tables 4 and 5): 00 K; 01Q; 02 E; 03 ⊗; 04 T; 05 P; 06 A; 07 S; 08 ∅; 09 R; 10 G; 11 C; 12 I; 13 L; 14 V, 15 F; 16 N; 17 H; 18 D; 19 Y; 20 W, 21 M; **Third variant** with ordinal number of AAs **00-20** (cf. Figure 2 and Tables 3, 6, 7 and 8): 00 K, 01Q; 02 E; 03 ⊗; 04 T; 05 P; 06 A; 07 S; 08 ∅ ; 09 R; 10 G; 11 C; 12 I; 13 L; 14 V, 15 F; 16 N; 17 H; 18 D; 19 Y, W, 20 M[2].

This way, it is presented a logic from which a series of AAs 00-19 follows, with the interruption of ordinal number 3 (for all three "stop" codons as a "stop" command within AAs alphabet), and also the logic from which follows a series of AAs 00-21, with interruption of ordinal numbers 3 and 8 (3 as a "stop" command and 8 as a "phantomic" interruption, an empty space) (Damjanović, 1998; Rakočević, 2004). In the other words, in the first case of reading, a "two-meaning" logical pattern is presented while in the second case a "one-meaning" logical pattern. For example, the serine is located on two locations (7 and 8) in the "two-meaning" logical pattern and only on one location (7) in the "one-meaning" logical pattern. Accordingly, for arginine, in the second case, we must assume that it is located only in the position 9, since otherwise it would be "mixed" with serine on the position 8. On the other hand, since position 11 must be occupied with priority by cysteine, then tryptophan must be moved for one cycle, according to the module 9, and it should appear on the position 20 (11+ 9 = 20) as a neighbor of tyrosine on the position 19 (neighbors in GCT also through a "stop command" loop, encoding by three "stop" codons). Similar case happens with methionine which, in relation to isoleucine, moves for one modular cycle further, on the position 21 (12+9 = 21) as a "neighbor" of tryptophan;

[1] In Shcherbak's four-codon/non-four-codon AAs system (Fig. 1 in Shcherbak, 1994) there are only three AAs as "duplicates" (L, S and R), whereas here appears isoleucine as a fourth (through Pu / Py coding codons).

[2] Within decimal numbering system (q = 10), the triplet 19-20-21 appears to be very adequate from still one very specific manner. Namely, if we exclude the zeroth amino acid (K), then the sum of ordinal numbers of other 19 AAs equals 190 – 10 units pro each one of amino acids. On the other hand, the sum from 0 to 20 (20 AAs plus one "stop' command) equals 210 – 10 units pro each one entity, amino acid or "stop' command. [Cf. the determination through "the symmetry in the simplest case" (Marcus, 1989), through the pair (q= 10 and q/2 = 5), presented in legend of Table 8].

"neighbor", from the aspect of the existence of one-meaning, i.e. one-codon amino acids. In the third case of reading (Damjanović and Rakočević, 2005) we have the appearance of a specific "mobile loop". Regarding Figure 2 we see that tryptophan comes one step back "in order" to be together with tyrosine (cf. legend of Table 3) and, at the same time, methionine comes at the former position of tryptophan. This "mobile loop" follows from a "theory of ribosomal code" (Section 4), from amino acid positions in Genetic Code Table (GCT), and from physico-chemical properties of AAs. [W and M as only two one-codon AAs; M and T as horizontal neighbors in GCT and only two AAs (within alanine stereochemical type) with a CH_3 atom group, etc.].

The codons for W and M in Table 3 are given in a vice versa position in order to signalize the mobile loop valid for the codons and not only for AAs. (last column: O.Nr.): The amino acid Y is 103_4 in a normal reading: UAC/CAU $\rightarrow$ 103_4; the M possesses the same ordinal number, reading from anticodon: AUG/GUA $\rightarrow$ CAU $\rightarrow$ 103_4; on the other hand, the M possesses the ordinal number 111_4, reading from the syn-codon: AUG/CCC $\rightarrow$ 111_4. The arithmetical mean for M equals $103_4 + 111_4 = 110_4$. which ordinal number possesses also the W, reading from its anticodon: UGG/GGU $\rightarrow$ CCA $\rightarrow$ 110_4. Thus, methionine – the first amino acid in protein biosynthesis – possess all three meanings (103_4, 110_4. and 111_4), valid for the whole AAs system.

3 A HOLISTIC APPROACH

For a better understanding the process of codon – amino acid assignment, a holistic approach to genetic code is needed (Rakočević, 1998, 2005). In such an approach genetic code appears to be a harmonic structure – a whole (and full) system, determined by Golden mean as well as by Generalized Golden Mean [GGM: three equations ($x^2 \pm 1 - 1 = 0$; $x^n + x - 1 = 0$; $x^2 + x - m/2 = 0$; $n = 1, 2, 3, ...$; $m = 0, 1, 2, 3 ...$)(cf. math.GM/0611095) in correspondence to positions of 20 canonical AAs and 64 codons on the segment-line 0-63, within a binary-code tree]; also as the first and only one possible case from many aspects. So, within a system of "letter-root-word-alphabet", the first possible case is the system "1-2-3-4" (64 *three-letter* words, changeable exactly for *one* letter accordingly to Gray code, 16 *digrams* and *four-letter* alphabet). From the aspect of validity of principle of minimum change and principle of continuity, any other case is not possible.

The same follows from the aspect of information theory, accordingly to principle of symmetry and the principle of self-similarity, in relation to the real

three-dimensionality. Namely, only a 6-bit binary-code tree with 64 words within a B^6 Boolean hyper-cube is possible (Rakočević, 1998).

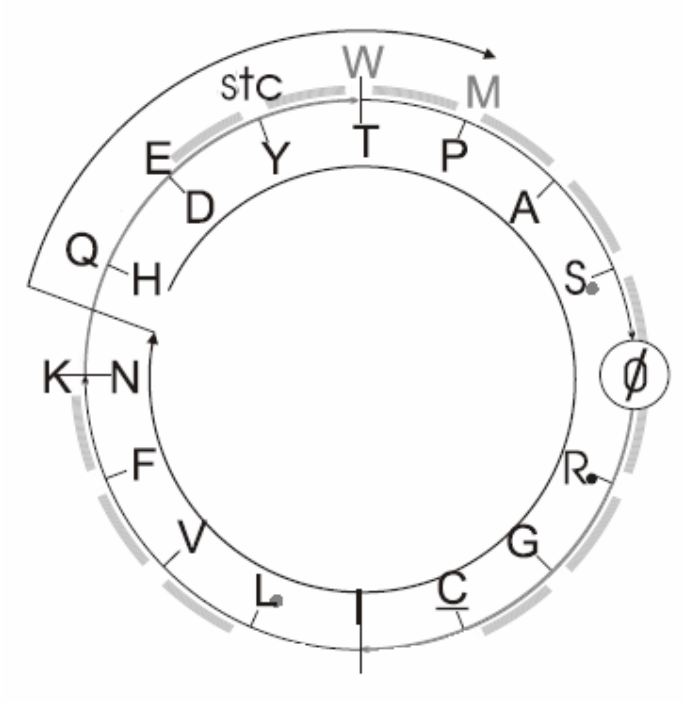


Figure 1. Spiral model of Genetic code (I). The spiral order of amino acids in accordance to their ordinal number, given by first and second variant of reading from codons and/or their digrams (doublets); the reading in quaternary numbering system, as it is presented in Remark 1 on p. 13 (first variant: exactly in relation to App. 1, Survey 2; second variant: exactly in relation to Solutions 1 – 4 and Table 4). An analogous figure for third variant is Figure 2; while a cross arrangement of this spiral is given in Figures 3 and 4.

[Distribution on a 7-bit tree: (B^7- B^6- B^5- B^4- B^3- B^2- B^1- B^0 / 1 x 128, 2 x 64, 4 x 32, **8 x 16**, 16 x 8, 32 x 4, 64 x 2, 128 x 1); on a 6-bit tree: (B^6- B^5- B^4- B^3- B^2- B^1- B^0 / 1 x 64, 2 x 32, 4 x 16, **8 x 8**, 16 x 4, 32 x 2, 64 x 1); on a 5-bit tree: (B^5- B^4- B^3- B^2- B^1- B^0 / 1 x 32, 2 x 16, **4 x 8**, 8 x 4, 16 x 2, 32 x 1)]. By this, the principle of self-similarity (B^3 Boolean real cube, expressed through a 3D model, and/or only through a holographic model, with **8** classes, each class with **8** words i.e. codons) is related to a determination by perfect and friendly numbers as follows (about perfect and friendly numbers as determinants of genetic code, see in Rakočević, 1997 and www.sponce.net).

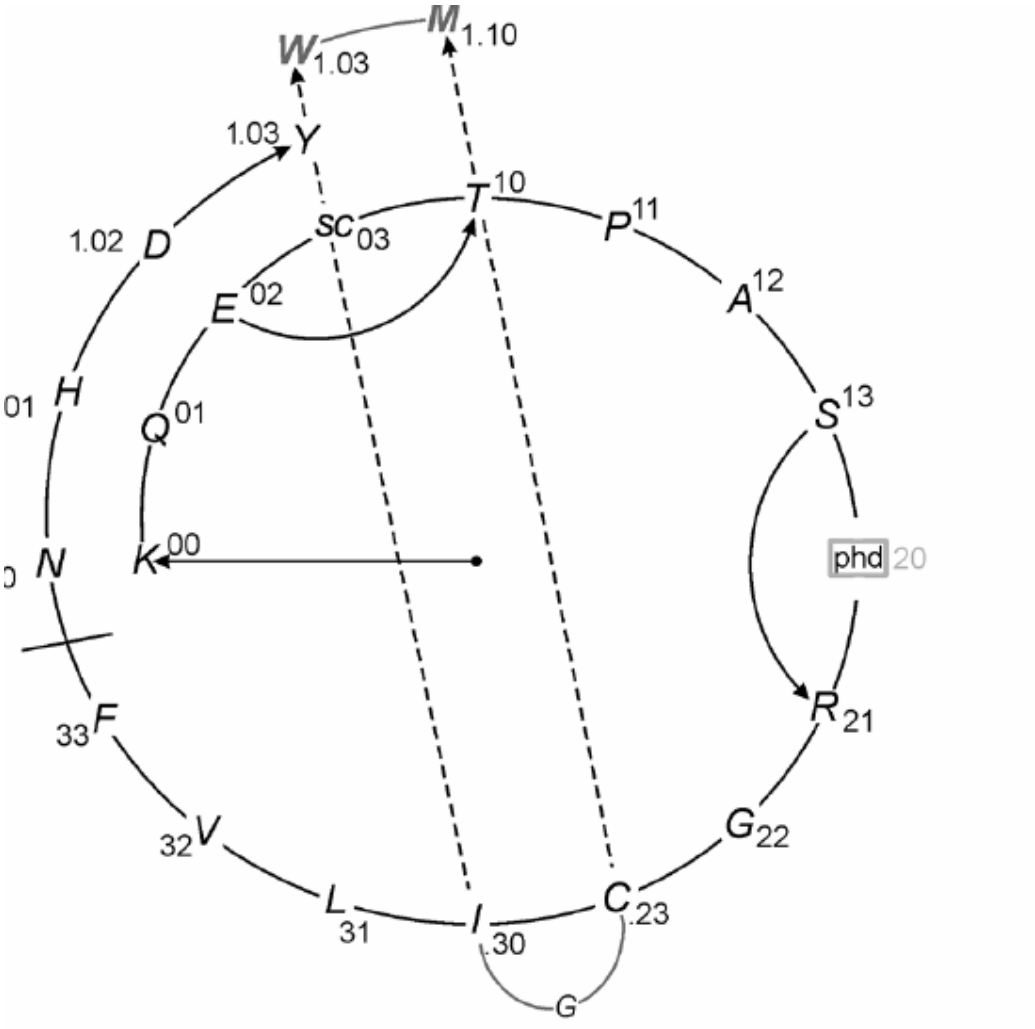


Figure 2. Spiral model of Genetic code (II). The spiral order of amino acids in accordance to their ordinal number, given by third variant of reading from codons and/or their digrams (doublets); the reading in quaternary numbering system, as it is presented in Remark 1 on p. 13; the arrangement itself exactly in relation to Tables 3, 6 and 8. A cross arrangement of this spiral is given in Figures 5.

m/codon	rs/codon	c/ama	O.Nr.
pu /AA	/00	K	0
pu /AC	/01	Q	1
pu /AG	/02	E	2
.yx	/yx	T	4
		P	5
		A	6
		S	7
py /GC	/21	R	9
py /GG	/22	G	10
py /GU	/23	C	11
notG /UA	/30	I	12
.yx	/yx	L	13
		V	14
py/UU	/33	F	15
py/AA	1/00	N	16
py/AC	1/01	H	17
py/AG	1/02	D	18
py/AU	1/03	Y	19
G /UA	1/03	W	19
G /GU	1/10	M	20

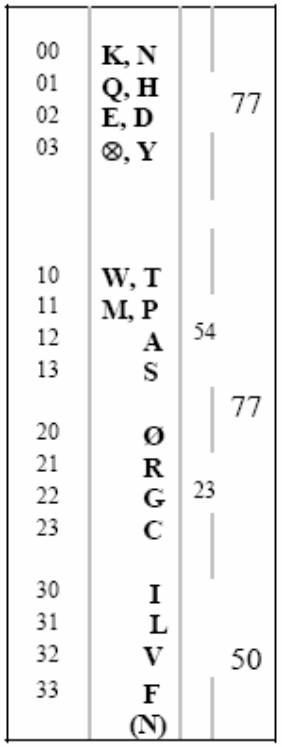


Table 3 (left). The relations among codons and canonical amino acids. The "m/codon" designates messenger RNA codons. In relation to Genetic Code Table (GCT) all codons are read from right to left; "pu"/"py" – purine or pyrimidines in third codon position; the point within codon "**.**yx" designates that both purine or pyrimidines can be in third codon position; "notG" means that only guanine is not possible in third position of codons, coding for amino acid isoleucine; The "rs/codon" – ribosomal codons and/or the conditions within ribosome system, enable to read the ordinal numbers in quaternary numbering system; the digram, i.e. doublet "/yx" is an analog of digram "**.**yx" in column "m/codon" (/yx equals 31_4 for leucine and 32_4 for valine). The "c/ama" – canonical amino acids; at the end of this column there are three multi-meaning AAs: Y, W and M, multi-meaning in terms of ordinal numbers. Last column: O.Nr. – Ordinal number.

Table 4 (right). Ordinal number of amino acids in relations to codon digrams. The splitting (classification and/or grouping) of canonical amino acids in accordance to their ordinal number,

given by Figure 1, i.e. by second variant of spiral model of genetic code, as it is presented in Remark 1 on p. 13, and to logical square (0-1-2-3) at the same time. An analogous table for third variant is Table 6, while (table building) for first variant is leaving to the readers (by this one must consult Surveys 1 and 2 in Appendix 1). A "mirror image" table (analogous to Table 7) also is leaving to the readers. Atom number balance, presented here, and valid for three AAs groups (50-77-77) stay in correspondence with Golden mean balance (60-66-78) as it is shown in Survey 3 in Appendix 2; in correspondence by principle of minimum change as well as of continuity, through an evident symmetry.

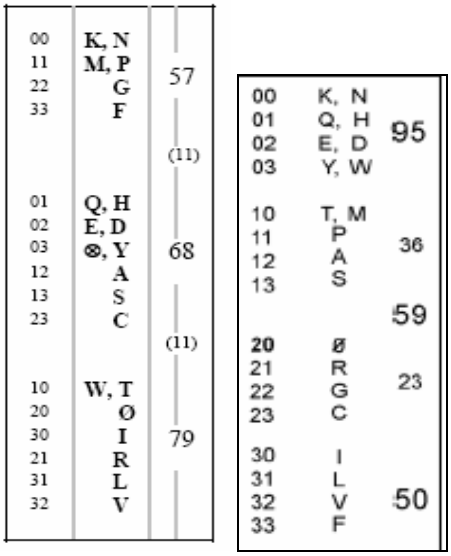


Table 5 (left). Ordinal number of amino acids in relations to codon digrams. The splitting (classification and/or grouping) of canonical amino acids in accordance to their ordinal number, given by Figure 1, i.e. by second variant of spiral model of genetic code, as it is presented in Remark 1 on p. 13, and to logical square (0-1-2-3) at the same time. An analogous table for third variant is Table 6, while (table building) for first variant is leaving to the readers (by this one must consult Surveys 1 and 2 in Appendix 1). A "mirror image" table (analogous to Table 7) also is leaving to the readers. Atom number balance, presented here, and valid for three AAs groups (50-77-77) stay in correspondence with Golden mean balance (60-66-78) as it is shown in Survey 3 in Appendix 2; in correspondence by principle of minimum change as well as of continuity, through an evident symmetry.

Table 6 (right). The relations among canonical amino acids (I).The splitting (classification and/or grouping) of canonical amino acids in accordance to their ordinal numbers, given by Figure 2, i.e. by third variant of spiral model of genetic code, as it is presented in Remark 1 on p.

13, and to logical square (0-1-2-3) at the same time. An analogous table for second variant is Table 4. The atom number balance, presented here, and valid for three AAs groups (95-59-50) is explained in the text (Section 4).

00	K, N	
10	T, M	**56-1**
20	***#***	
30	I	
01	Q, H	
11	**P**	
21	**R**	60
31	L	
02	E, D	**56+36**
12	**A**	
22	**G**	60-28
32	V	
03	Y, W	
13	**S**	
23	**C**	**56+1**
33	F	

000	**K**	
001	**Q**	36(06)
002	**E**	
003	*	
		61(28)
010	*T*	
011	*P*	
012	*A*	25(22)
013	*S*	
020	***#***	
021	*R*	
022	*G*	23(30)
023	**C**	
030	**I**	**73(84)**
031	**L**	
032	**V**	50(54)
033	**F**	
100	*N*	
101	*H*	**70(90)**
102	*D*	
103	*Y*,**W**	
110	**M**	

Table 7 (left). The relations among canonical amino acids (II). This Table is a "mirror image" of Table 6 {[(00-01-02-03) (10-11-12-13) (20-21-22-23) (30-31-32-33)] / [(00-10-20-30) (01-11-21-31) (02-12-22-32) (03-13-23-33)]}. The atom number balance appears to be in correspondence with first and second perfect number (56 = 2 x 28; 36 = 6^2; 60 = 6 x 10 etc.).

Table 8 (right). The splitting of amino acids in correspondence to hydrophobicity. The splitting (classification and/or grouping) of canonical amino acids in accordance to their ordinal number, given by Figure 2, i.e. by third variant of spiral model of genetic code, as it is presented in Remark 1 on p. 13. An analogous table for first and second variant is leaving to the readers. Atom number balance, presented here, and valid for three AAs groups (*70-61-73*) stay in correspondence with Golden mean balance by an evident symmetry ["the symmetry in the simplest case" (Marcus, 1989), through the base of decimal numbering system, q = 10; q/2 = 5 (*70*-60 = **10**; 66-*61* = **5**; 78 – *73* = **5**). The splitting of AAs corresponds also with hydrophobicity and polarity of AAs (see the text: Section 4). Bold amino acids as in Figure 5.

The sum of the ordinal numbers on two middle and neighbor branches on the 6-bit binary-code tree is 220 and 284, respectively, which two numbers represent

first pair of friendly numbers. Their sum equals 504, as in two second, two third and two fourth branches (reading one branch from left and second branch from right side at the same time, in relation to middle point of tree). On the other hand, we have the realization each of four possible letters, minimum once, through the realization of first four words (0. UU**U**, 1. UU**C**, 2. UU**A**, 3. UU**G**) on the binary-code tree and/or in GCT (cf. Rumer's and Shcherbak's idea about the four-codon AAs in next Section). The sum of their ordinal numbers equals 6, which is the first perfect number. After the realization of first eight words (0. UU**U**, 1. UU**C**, 2. UU**A**, 3. UU**G,** 4. **C**U**U**, 5. **C**U**C**, 6. **C**U**A**, 7. **C**U**G**) occurred the determination of upper half of GCT (4 half-columns, each half-column with 8 codons), and the sum of all eight ordinal number equals 28, which is the second perfect number. Stepping to the ordinal (codon) number 31, we have the determination of the left half of GCT (2 columns, each column with 16 codons) and the realization of third perfect number, because the sum of all numbers from 0, i.e. from 1 to 31 equals 496. Finely, with a full cycle (from 0 to 63 and back, from 64 to 127) we have the determination of the full GCT system and the realization of fourth perfect number at the same time (8128), because the sum from 1 to 127 equals 8128.

The holistic approach to genetic code comes also from some specific chemical aspects. First of all, from the aspect of classification of 20 canonical AAs into four stereochemical types (Popov, 1989; Rakočević and Jokić, 1996) – glycine type (with only G amino acid), proline type (only P), valine type (V & I) and alanine type (the rest of 16 AAs, each amino acid with a H-C-H "screen" between the "head" and "body", i.e. side chain; the exception is threonine with an H-C-CH_3 "screen"). The appearance of glycine corresponds to the appearance of the first possible non-hydrocarbonity (H as the side chain); the appearance of alanine corresponds to the appearance of the first possible non-cyclic hydrocarbonity (CH_3 as side chain); the appearance of valine corresponds to the appearance of the first possible hydrocarbon half-cyclity (isopropyl atom group as side chain), and the appearance of proline corresponds to the appearance of the first possible hydrocarbon cyclicity (–CH_2–CH_2–CH_2– group in side chain and in contact with the "head").

As a second holistic chemical aspect is the splitting of 20 canonical AAs into two classes, 10+10, in correspondence to two classes of enzymes aminoacyl-tRNA synthetases. The simpler and/or smaller AAs (within AAs pairs) are handled by less complex enzymes of class II, whereas larger (more complex) AAs molecules are handled by more complex enzymes of class I (bold underlined) through a pairing process: I. aliphatic AAs – Ia. hydrocarbon non-polar AAs (G-**V**, P-**I**, A-**L**); Ib. chalcogene polar AAs (S-**C**, T-**M**, N-**Q**); Ic. polar

charged AAs (D-**E**, K-**R**); II. aromatic AAs (F-**Y**, H-**W**). [Cf. this classification with a similar but more global, presented here in Table 1, in relation to Table 2 (Patel, 2005, p. 529: "A closer inspection of Table 2 shows that all class-II amino acids, except Lys, can be coded by the codons NNY") (Y = U or C)]. A further classification is also a proof for the wholeness within a holistic system of genetic code, the classification in relation to the base type in third position of the belonging codon. Thus, in class II there are AAs whose codons do not possess purine in third position (first subclass): N, D, F, H (neighbors within the system presented in Figures 1 and 2 as F, N, H, D) with 40 atoms within their side chains; then AAs whose codons possess purine in third position (second subclass): **K**, P, A, S, T, G (also as neighbors in Figures 1 and 2, as T, P, A, S, excluding K, which is an exception with the location on the other side in Figures 1 and 2; the exeption still once as in above given Patel's comment; also as a zeroth amino acid in three variants of reading within AAs system, presented in first paragraph of Section 2) with 40+01 atoms. On the other hand in class I there are AAs whose codons possess either pyrimidine or purine in third codon position (first subclass): V, L, R, also with 40 atoms within their side chains; then AAs whose codons possess only purine in third codon position (second subclass): M, Q, E, W with 40+10 atoms; as a third subclass there are AAs whose codons possess only pyrimidine in third codon position: C & Y with 40-20 atoms within their side chains. Out of the classification within class I there is isoleucine which belongs to the first subclass within standard genetic code and to the third subclass within mitochondrial genetic code, presented in Table 2 (cf. isoleucine positions in Tables 3-5 and in App. 1, Surveys 1 and 2).

The presented holistic approach can be also interesting in the study of possible analogy with other natural codes, esspecialy with visual code. In shortest words, the genetic code alphabet UCAG can be analogue with visual alphabet UBGR (U-Union of all rainbow colors, i.e "white" color (light); B – Blue; G – Green, and R – Red). Namely, as it is known, each human cone cell absorbs light in only one of three bands of the spectrum: blue, green and red. This follows from the fact that there exist three types of the genes, coded for the three color receptor proteins. [Note: Cone cell is one of specialized, photosensitive cells in the retina of the eye concerned with the perception of color and with daylight vision]. Within 64 possible "words" from the alphabet UBGR there are exactly 28 (second perfect number!) words which possess U, as white light, and $6^2 = 36$ "color words" without U (number 6 as first perfect number). The hypothesis about a possible analogy of genetic code and visual code is only a part of a larger idea on the analogy of genetic code and neuro code, and separatly – genetic code and a sensory code (Damjanović, 1998).

4 DIVERZITY OF CODON – AMINO ACID ASSIGNMENT

Our earlier studies (Damjanović, 1998; Rakočević, 2002, 2004; Damjanović & Rakočević, 2005) of codon – amino acid assignment have lead us to a specific numbering coding of nucleotides in correspondence with a logical square (A = 0, C =1, G =2, U =3)[3], and to a possible their reading through quaternary numbering system; also to an ordering of digrams (base doublets within the codon) and codons (reading right-to-left: **.**yx & Z**.**yx) as ordinal numbers from 000_4 to 111_4 for belonging canonical AAs. In this way, sixteen digrams and belonging AAs appear as a parallel discrete array (the nucleotide-letters within the codons must be read from right-to-left and the numbers, in quaternary numbering system, from left-to-right): *AA*(00) – K, N; *AC* (01) – Q, H; *AG*(02) – E, D; *AU*(03) – stc,Y; *CA*(10) – T; *CC*(11) – P; *CG* (12) – A; *CU*(13) – S; *GA*(20) – S, R; *GC*(21) – R; *GG*(22) – G; *GU*(23) –C, W, stc; *UA*(30) – I, I, M; *UC*(31) – L; *UG*(32) – V; *UU*(33) – F, L; The ordinal numbers from 100_4 to 103_4 must be read from respective codons as follows (Tables 3, 4 and 5): *C.AA*(1.00) – N; *C.AC*(1.01) – H; *C.AG*(1.02) – D; *C.AU* (1.03) – Y.

(*Remark* 1: We use here the point in designations accordingly to the convention, given obove as „**.**yx" & „Z**.**yx". By this one must notice that at last four AAs the codon letter "C" has a meaning of "Z" in „Z**.**yx").

The digram *GA*, staying for both S & R, appears as an "phantomic" digram, because these two AAs possess their "first" digrams from the codon families: *CU* and *GC*, respectively. Otherwise said, here there are two possibilities for reading. As a first we have a "two-meaning" logical patterns, where both serine and arginine are located on two locations each [*CU*(13) – S; *GA*(20) – S, i.e. 7 and 8 in decimal numbering system; *GA*(20) – R; *GC*(21) – R, i.e. 8 and 9 in decimal numbering system]. In the second case – an "one-meaning" logical patterns, where both serine and arginine are located on one location each, and between them appears a "phantomic" digram (phd) with an empty space for non-existing amino acid [*CU*(13) – S; *GA*(20) ∅, *GC*(21) – R].

[3] As to now we founded only one paper with the same numbering coding for bases, i.e. nucleotides; however, not for RNA but for DNA: A → 0, C → 1, G → 2, T → 3 (Sirakoulis et al, 2004).

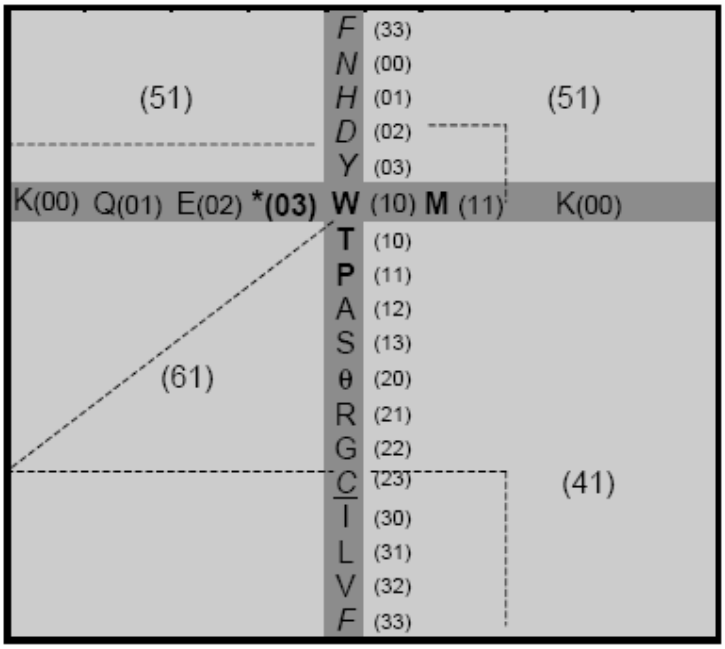


Figure 3. Cross model of Genetic code (I). The cross arrangement of amino acids displayed from Figure 1 in accordance to an atom number balance: above/down 102/102 of atoms within amino acids side chains.

In such an understanding we find that serine must be located only on one location (position 7) and arginine only in one position (the position 9), since otherwise these two AAs would be "mixed" on the position 8. From this reason it is clear why position 8 must be empty (designation ∅ in corresponding Figures and Tables). Ordinal numbers for W and M and their respective analogs deserve a particular consideration: (I**)**. W, T → 10 and/or 1.10 and M, P → 11 and/or 1.11, in Figures 1, 3 and 4 and in Tables 4 and 5; (II). W, Y → 03 and/or 1.03 and M, T → 10 and/or 1.10, plus P → 11 in Figures 2 and 5, and in Tables 3, 6, 7 and 8. In order for a better understanding, the first case is displaying in Solutions 1 – 4 still once:

K(00)-Q(01)-E(02)-*(03)-T(10)-P(11)-A(12)-S(13)- ∅ **(20)**
R(21)-G(22)-*C(23)*-I(30)-L(31)-V(32)-*F(33)* (1)

*F-N(00)-H(01)-D(02)-**Y(03)***-**W(10)-M(11)** (2)
(100)- (101)-(102)-(103)- (110)-(111)

K(00)-Q(01)-E(02)-*(03)- W(10)-M(11)-A(12)-S(13)- ∅ **(20)**
R(21)-G(22)-*C(23)*- I(30)-L(31)-V(32)-*F(33)* (3)

*F-N(00)-H(01)-D(02)-**Y(03)***- **T(10)-P(11)** (4)
(100)- (101)- (102)-(103)- (110)- (111)

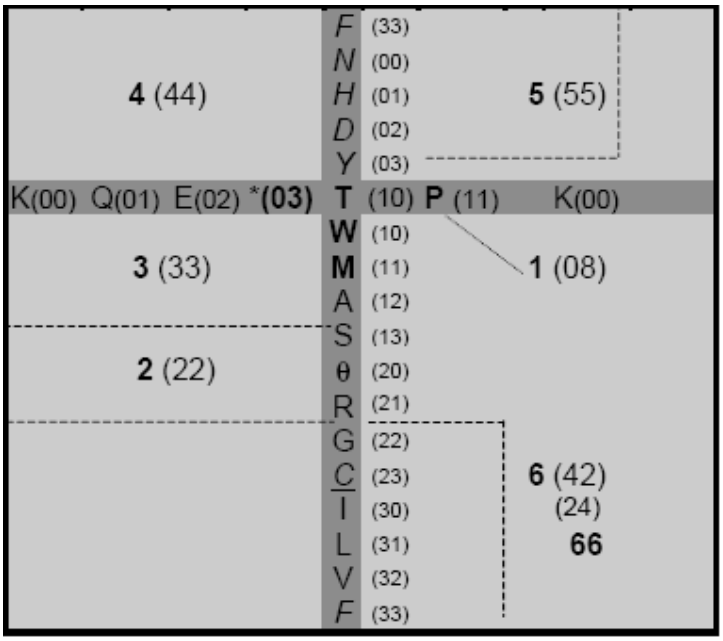


Figure 4. Cross model of Genetic code (II). The cross arrangement of amino acids displayed from Figure 1; the same as in Figure 3 except a vice versa position for the pairs W-M / T-P. The atom number balance follows a molecule number balance. [Notice that in a vice versa position for P/M we have the situation 1(11) – 1 molecule with 11 atoms].

As we can see from Solutions (1) and (2) the spiral model of genetic code, relating to Siemion and Siemion's rule, as well as Davidov's rule (Siemion and Siemion, 1994; Davidov, 1998; cf. Rakočević, 2004)[4] can be given in form of "a

[4] Classification of canonical AAs derived from our dynamic model brings about clarification of physicochemical criteria, such as purinity, pyrimidinity – and, particularly, codon rules. The system implies both rules of Siemion and Siemion and of Davidov, as well as balances of atom and nucleon numbers within groups of AAs. Formalization in this way opens fruitful chances of extrapolating backwards, to initial organization of heredity.

cross" too. [The Solutions (3) and (4) are the same as Solutions (1) and (2), respectively, with a position changing for two AAs pairs: T-P/ W-M]. Bearing in mind that both variants – spiral and cross – seek a connection "head to tile" (F-F), two intersecting lines appear (cf. Figures 1, 3 & 4 with the above given Solutions); the horizontal (shorter) leg of the cross consists of AAs of Pu type (K, Q, E, W, M), while the vertical (longer) leg contains two sub-classes: up there are AAs of Py type (italic: F, N, H, D, Y), and down there are AAs of "Py or Pu" type (from T to V-F), with an exeption of cysteine which is of Py type.

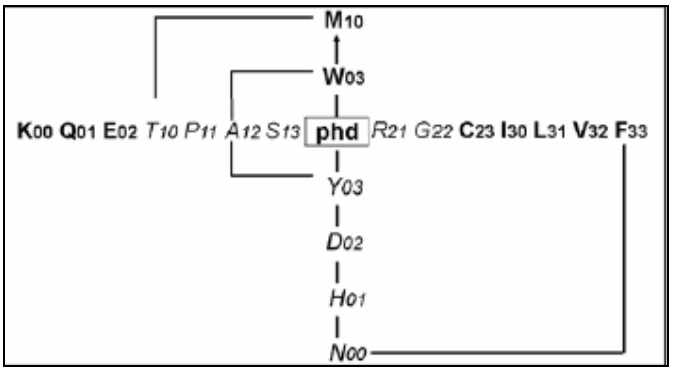


Figure 5. Cross model of Genetic code (III). The cross arrangement of amino acids displayed from Figure 2 in accordance to physico-chemical parameters, hydrophobicity and polarity (explanation in the text: Section 4). Bold amino acids as in Table 8.

4.1. The reading of codons and their digrams

The reading of codons and their digrams in quaternary and/or decimal numbering system (according to Damjanović, 1998) starts with "zero" column in GCT where are the codons with middle base "A". Accordingly, we read zeroth "digram" AA as 00 within the codon AAA that is coding for lysine. Subsequently, the neighboring codon AAC, in reverse case as CAA (that is as 100_4), which is the number 16_{10}, as ordinal number for asparagine, etc. In such a manner the 18 AAs can be read from the codons; the first to last, the tryptophan, is read from the anticodon (anticodon ACC, read from right to left as CCA) with ordinal number 110_4, that is 20_{10} (Damjanović, 1998, p. 6: "the cycle of digrams is presented, and the spiral of codons ... with the 'inverse' appearance of number

20"). Despite the last amino acid, methionine, can be read in a specific way from the "ribosomal code", with ordinal number 111_4, that is 21_{10}, it is interesting that ordinal number of methionine can be read also from the "syn-codon" (Rakočević, 2004, p. 226 down).

[*Remark* 2. The codons with the largest diversity (all three bases are different) together with codons of the smallest diversity (all three bases are the same) contain a balanced number of nucleons: within side chains of AAs assigned to 12 codons with a clockwise direction (U $\rightarrow$ C$\rightarrow$ A) plus codons UUU & AAA there are 703 nucleons; the same result is in side chains of AAs assigned to 12 codons with an anti-clockwise direction plus the codons CCC & GGG (cf. Fig. 4 in Rakočević, 2004, p. 226). From that it is a reason for introduction of the term "syn-codon" on the following way: UUU is a syn-codon for all six permutations of codon CAG, with the designation UUU/CAG. The same goes for the remaining three syn-codons: CCC/UAG (where one of the permutations is methionine-codon AUG), than AAA/UCG and, finally, GGG/UCA].

4.2. Theory of a ribosomal code

As an extract from translation, it is possible a specific mapping of codons onto canonical amino acids (m/codons onto c/ama in Table 3). By this all "capital" and "stop" codons, that impact on ribosome assembling, as well as the duplicates of canonical amino acids (S, R, L and I), are left aside from the model, presented in Fig. 2 & 5). Modeling itself runs, as we said, with help of quaternary numeric transforms of 4 nucleotides: A=0, C=1, G=2, U=3. Codons **xy/Z,** if read from right, expose 16 digrams **•yx** (cf. Remark 1 on p. 13) as a smooth array of ordinal numbers (O.Nr. in Table 3), where 4 groups of digrams, dominated by A, C, G and U in central **y** codon position represent *complemental* pairs **Ax-Ux** and **Cx-Gx** (Table 3) making a quasi cycle, mapped onto canonical amino acids from K to **F** (Fig. 2).

The fact that ribosome is ubiquitous medium of convergence of 61 messenger RNA codons to 20 crucial tRNA canonical AAs blocs is reflected in a *Theory of ribosomal coding* which includes the following principles:

(a) Numeric values of **Z** are restricted to Ø and 1, corresponding to A and C. This, for example, allows interpretation of "ribosomal" codons mapped onto **N, H, D** and **Y** (cf. Remark 1 on p. 13) through possible "para-codons" (De Duve, 1988).

(b) Sliding of ribosome along mRNA is discrete, elementary step covering 3 nucleotides (i.e. codon); the adherence of **y** to its complement (**y**|cy) makes coupling center of codon to tRNA.

(c) The event of translation happens within polar space of ribosome. So the numeric code includes space angles: as mRNA is, by no means, a straight line, a linear arrangement is approximated by a series of **y**|cy, which is the center of polar space.

(d) Translation is made readable with help of an abstract of "ribosomal" codon ("rs" codon) as well as rs**Z** (Ø: 0 or absent; 1: C engaged as **Z** in third codon position; C, writing in "cy" as a small letter, instead a large letter because "C", i.e. "c" is related here to a coordinate – the Z coordinate; cf. Remark 1 on p.13). This applies to 18 canonical AAs in correspondence to 18 hypothetical "ribosomal" codons (cf. Fig. 2, which depicts the basic spiral of canonical AAs).

(e) Purines are seen as more complex than pyrimidines[5]; a G-C base pair is held together by three hydrogen bonds and provides a greater stabilizing influence than A-U pair, which has only two hydrogen bonds. G in position **Z** leads to the digrams - **.**UA and **.**GU – to inversion, i.e. transformation to complemental rs/codon (cf. last two codons in column "m/codon" in Table 3, including the explanation in legend). With this in view, Figure 2 depicts a hyper-spiral K to **M**. On the other hand, Table 3 summarizes the relations of (numeric) m/codon to rs/codon transformation, as well as mRNA - canonical AAs mapping. (Ordinal numbers are given in decimal numbering system.). In addition it is important to say that Figure 5 suggest two ideas: (1) the canonical AAs spiral creeps under "phantomic digram" •GA **(**phd**)**, making evident a spiral in space; (2) it evokes analogy with *sensory code*, which is the matter of further researches (cf. last paragraph in Section 3).

5 PHYSICO-CHEMICAL PARAMETERS

Physico-chemical parameters of canonical amino acids such as hydrophobicity and polarity, are closely related to the parameters of the above mentioned mapping hyper-spiral. So, Tables 4 and 5 are related to Figure 1, 3 and 4. As it is self-evident from the illustrations and their legends, there is a full balance; the balance between ordinal order of AAs, atom number, as well as nucleon number, within side chains of amino acid sub-classes.

[5] Purines are seen as energents too, with a role analogous to that in mitochondrial "electronic respiration"; also their "H-potential" makes Guanine, with 3 H-bridges, in a way dominant.

Between all other evident we reveal some hidden balances. So, the number of atoms within amino acids (their side chains) displaying in two last vertices of logical square in Table 4 ("0" – KN,QH,ED,Y and "3" – I,L,V,F) equals 77+50 = 127, which number is last point in a 7-bit binary-logical tree. On the other hand, within AAs displaying in left-directed complement (**32-31-21-30-20-10**), in Table 5, there are 79 atoms, which number is Golden mean point in a 7-bit binary-logical tree; in right-directed complement (**01-02-03-12-13-23**) there are 11 atoms less, and in non-complement (00-11-22-33)[6] still 11 atoms less, what means a determination by both principles – of minimum change and continuity.

Tables 6 shows the grouping of amino acids accordingly to ordinal number and to two patterns of logical square at the same time. On the other side, in Table 7 it plays only the logical square (in a vice versa position in relation to table 6) but not the ordinal numbers system.

Tables 6 and 7 represent amino acid arrangement based on the order of mRNA *digrams* (Damjanović, 1998); two tables in complemental order, grouped in four logical square patterns (0,1,2,3). The pairs in Table 6, **K**-N, **Q**-H, **E**-D, **Y**-W, **T**-M, **P**-R, **A**-G, **S**-C, **I**-V, **L**-F, with 102 atoms in 10 first pair-members (bold) and 102 atoms in 10 second pair-members, reveal a full symmetrical and proportional balance. There are other balances as well. For example, in three designated amino acid classes there are: 50, 59 and 95 of atoms. The number 59 represents an increase for exactly one modular cycle (in module 9) in relation to number 50; the number 95 is an inversion of 59.

In Table 7 the same pairs appear in a different order: **K**-N, **T**-M, **I**-V, **Q**-H, **P**-R, **L**-F, **E**-D, **A**-G, **Y**-W, **S**-C. Except for the proportional balance 102:102 = 1:1, the system is balanced by number of atoms within amino acid molecules, located on odd and even positions in both lines, the first as well as second pair-members, 61/60 and 41/42, respectively.

These new amino acid pairing agree with physico-chemical properties. Thus, the pairs K-N and Q-H come from a specific crossing: the pair K-H consisting of two basic AAs (the third R is paired with P, both with untypically bonded nitrogen); N-Q as a classic pair – all with nitrogen. Further, from classic H-W, follows Y-W, rather than F-W, because F must go with L, where both molecules possess the same structural motif (isobutane type of branching and H-C-H group between "haed" and "body"). Finelly, both T-M molecules are methyl-

[6] This classification into two complemental and one non-complemental amino acid classes one must cf. with an analog classification determined by Golden mean (Rakočević, 1998, Scheme 2, p. 289); cf. also App. 2, Survey 3 in this paper.

derivatives: threonine possesses H-C-CH_3 group derived from H-C-H; methionine possesses S-CH_3 derived from S-H. The rest of four pairs (E-D, A-G, S-C and I-V) represent the four classical pairs.

Table 8 shows a strict distinction in hydrophobicity among canonical AAs in classic AAs pairs (Black and Mould, 1991; Rakočević, 2000). Namely, more hydrophobic (bold) and less hydrophobic (non-bold) canonical AAs appear alternatively, in separate groups, knowing that in the system of classical amino acid pairs (Dlyasin, 1998; Rakočević, 2004), the order of canonical AAs, in correspondence with the hydrophobicity, is as follows from Fig. 5: (**K**-R, **Q**-N, **E**-D, **C**-S, **I**-P, **L**-A, **V**-G, **F**-Y, **W**-H, **M**-T) / (**K**, **Q**, **E**) (T, P, A, S, R, G) (**C**, **I**, **L**, **V**, **F**) (N, H, D, Y) (**W**, **M**). At the same time, the more hydrophobic canonical AAs (first members) are handled by class I enzymes, aminoacyl-tRNA synthetases (all but lysine and phenylalanine; lysine as a "pure" amino derivative and phenylalanine as a "pure" aromatic amino acid), whereas less hydrophobic (second members) canonical AAs are handled by class II enzymes (all but arginine and tyrosine). It is also self-evident, from Table 8, that these regularities are followed by a balance of the number of atoms within amino acid side chains (numbers for the brackets) and of the sums of ordinal numbers (numbers within the brackets) (cf. legend of Table 8).

The splitting into AAs groups after hydrophobicity corresponds with the splitting after polarity (Figure 5); after polar requirement, cloister energy and hydropathy (Kyte and Doolittle, 1982; Rakočević, 2004). Canonical AAs on the left side of horizontal cross leg, in Fig. 5, are polar (all but alanine), whereas canonical AAs on the right side are nonpolar (all but arginine). On the other hand, at the vertical cross leg only two outer canonical AAs (M and F) are nonpolar, whereas all other – the inner canonical Aas – are polar. (Certainly, one must bear in mind that G and P are ambivalent).

6 CONCLUSION

A new understanding of codon – amino acid assignment, displayed through previous five Sections, appears – through presented regularities – to be very adequate from the aspect of core essence of coding process within genetic code. Between all others aspects, it is showed that positions of 20 canonical amino acids and belonging codons are arranged through a specific – spiral as well as a cross model of Genetic Code, such a model which stay in correspondence with physico-chemical properties of canonical amino acids and with the sequence of natural numbers at the same time.

REFERENCES

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

Crick, C. H. F. (1966) The Genetic Code. *Cold Spring Harbor Symp. Quant. Biol*., 31, published in: The chemical basis of life – an introduction to molecular and cell biology, *Scientific American*, 1973, 192-198.

Crick, C.H. F. (1968) The Origin of the Genetic Code. *J. Mol. Biol*. 38, 367-379.

Damjanović, Z. (1998) Logic core of genetic code. *Glasnik of the Sec. of nat. sci. of Montenegrin Academy of Sciences and art (CANU),* 12, 5-8.

Damjanović, Z. M, Rakočević, M. M. (2005) Genetic code – An alternative model of translation. *Annals of New York Academy of Sciences*, 1048, 1-7.

Davydov, O. V. (1998) Amino acid contribution to the genetic code structure: end-atom chemical rules of doublet composition. *J Theor Biol.* 193, 679-690.

De Duve, C. (1988) The Second Genetic Code. *Nature,* 333,117-118.

Dlyasin, G. G. (1998) *Azbuka Germesa Trismegista, ili molekulyarnaya tainopis' myshleniya*. Belye Al'vy, Moskva.

Gamow, G. (1954) Possible relation between DNA and protein structure. *Nature*, 173, 318.

Gilbert, W. (1986) The RNA World. *Nature* 319, 618.

Hayes, B. (1998) The invention of the genetic code, *Americ. Scient.,* 86, 1-13.

Konopel'chenko, B. G, Rumer, Yu. B. (1975) Klassifikaciya kodonov v geneticheskom kode. *Doklady Akad. Nauk*, SSSR, 223, 471-474.

Kyte, J., Doollittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol*. 157, 105-132.

Marcus, S. (1989) Symmetry in the Simplest Case: the Real Line. *Computers Math. Applic.* 17, 103-115.

Patel, A. (2005) The triplet gen. code had a doublet predecessor. *J. Theor. Biol*. 233, 527-532.

Popov, E. M. (1989) *Strukturnaya organizaciya belkov* (in Russian). Nauka, Moscow.

Rakočević, M. M., 1997. Two classes of the amino acyl-tRNA synthetases in correspondence with the Codon path cube. Bull. Math. Biol. 59, 645-648 (www.sponce.net).

Rakočević, M. M. (1998). The genetic code as a Golden mean determined system, *Biosystems*, 46, 283-291.

Rakočević, M. M. (2000) The factors of the classification of protein amino acids, *Glasnik of the Section of Natural Sciences of Montenegrin Academy of Sciences and art (CANU)*, 13, 273-294

Rakočević, M. M. (2002) Genetic Code: Alphanumerical Regularities Within a Codon Determined Amino Acid Order, Glasnik of the Section of Natural Sciences of Montenegrin Academy of Sciences and art (CANU), 14, 183-197.

Rakočević, M.M (2004) A harmonic structure of the genetic code, *J. Theor. Biol.* 229, 221 – 234.

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

Rakočević, M.M. (2005) A holistic approach in understanding of genetic code (in preparation).

Rumer, Yu, B. (1966) O sistematizacii kodonov v geneticheskom kode, Doklady Akad. Nauk. SSSR, 167, 1393-1394.

Shcherbak, V. I. (1989) Rumer's rule and transformation in context of the co-operative symmetry of the genetic code. *J Theor. Biol.* 139**,** 9271-401.

Shcherbak, V.I. (1993). Twenty canonical amino acids of the genetic code: the arithmetical regularities. Part I. *J Theor. Biol.* 162, 399-401.

Shcherbak, V. I. (1994) Sixty-four triplets and 20 canonical amino acids of the genetic code: the arithmetical regularities. Part II. J Theor. Biol. 166, 475-477.

Siemion, I. Z., Siemion, P. J. (1994) The informational context of the third base in amino acid codons. Biosystems, 33, 139-148.

Sirakoulis, G.Ch. et al. (2004) An algorithm for the study DNA sequence evolution based on the genetic code.*Biosystems*, 77, 11-23.

APPENDICES

	I	II	III	IV	V		
IV	03 ⊗	04 **T**	11 **C**,**W**, ⊗	12 **I**,**M**,**I**	19 **Y**	**83** (518)	**126+10**
III	02 **E**	05 **P**	10 **G**	13 **L**	18 **D**	**39** (231)	(750 – 1)
II	01 **Q**	06 **A**	09 **R**	14 **V**	17 **H**	**53** (311)	**126-10**
I	00 **K**	07 **S**	08 **S**,**R**	15 **F**,**L**	16 **N**	**77** (440)	(750 + 1)
	36	**25**	**63**	**87**	**41**		
			(88)	(87)			
			(77)				

App. 1, Survey 1. Amino acid order read from GCT (I). The splitting (classification and/or grouping) of canonical amino acids in accordance to their ordinal number, given by first variant of spiral model of genetic code, as it is presented in Remark 1 on p. 13. The ordinal numbers are reading from codons and/or their digrams (doublets) in the quaternary numbering system. A strict balance in atom number and nucleon number is self-evident. The three AAs groups (77-78-88), read from columns, stay in correspondence with Golden mean balance (60-66-78) as it is shown in App. 2, Survey 3. The reading from the rows gives an atom number balance: 83+53 = 126+10 and 39+77= 126-10; then, a nucleon number balance: 518+231=750-1 and 311+440 = 750+1. (Note: the number of atoms within amino acid molecules as in Rakočević and Jokić, 1996; and nucleon number as in Shcherbak, 1993, 1994).

	I	II	III	IV	V		
IV	03 ⊗	04 **T**,**W**	11 **C**	12 **I**	19 **Y**	**59** (386)	**102-1**
III	02 **E**	05 **P**,**M**	10 **G**	13 **L**	18 **D**	**50** (306)	(628+10)
II	01 **Q**	06 **A**	09 **R**	14 **V**	17 **H**	**53** (311)	**102+1**
I	00 **K**	07 **S**	08 ∅	15 **F**	16 **N**	**42** (252)	(627-10)
	36	**54**	**23**	**50**	**41**		
		(77)		(50)			
			(77)				

App. 1, Survey 2. Amino acid order read from GCT (II). All as in previous Survey except the amino acids are given as one-meaning. A strict balance in atom number and nucleon number is also self-evident. The three AAs groups (77-77-50), read from columns, stay in correspondence with Golden mean balance (60-66-78) as it is shown in App. 2, Survey 3. The reading from the rows gives an atom number balance: 59+42 = 102-1 and 50+53= 102+1; then, a nucleon number balance: 386+252=628+10 and 306+311 = 627-10.

5	02 13 24 35 46	120 (084)	102
3	**57** **68** **79**	204	
4	90 101 112 123	(222) 426	

6	06 12 24 30 42 48	**162** (42)	
3	**60** **66** **78**	**204**	402
6	84 96 102 114 120 132	(444) **648**	

App. 2, Survey 1 (left). Arithmetical regularities as determinants of Genetic code (I). The arithmetical regularities that determine the splitting of amino acids into classes possesses 57-68-79 of atoms (Table 5); a determination through the connection with the

total number of atoms (204) within 20 AAs molecules, i.e. their side chains. Notice also a parallel determination through Pythagorean pattern: 3-4-5.

App. 2, Survey 2 (right). Arithmetical regularities as determinants of Genetic code (II). The arithmetical regularities that determine the splitting of amino acids into classes possesses 60-66-78 of atoms (cf. App. 2, Survey 3); a determination through the connection with the total number of atoms (204) within 20 AAs molecules (their side chains), and through Golden mean at the same time. Notice also a parallel determination through first perfect number (6) and its half (3).

60	66	78	87	87
–	+	–	–	–
10	11	01	10	01
50	77	77	77	88

App. 2, Survey 3. Arithmetical regularities as determinants of Genetic code (III). The connection of arithmetical regularities presented in two previous Surveys.

FURTHER GENERALIZATION OF GOLDEN MEAN IN RELATION TO EULER'S "DIVINE" EQUATION

Miloje M. Rakočević

Department of Chemistry, Faculty of Science, University of Niš, Ćirila i Metodija 2, Serbia(E-mail: m.m.r@eunet.yu)

Abstract.

In the paper a new generalization of the Golden mean, as a further generalization in relation to Stakhov (1989) and to Spinadel (1999), is presented. Also it is first observed that the Euler's "divine equation" $(a + b^n)/n = x$ represents a possible generalization of Golden mean.

Key words: Golden mean, Generalized Golden mean, Metallic mean, Stakhov's generalization, Spinadel's generalization, Euler's "divine equation".

1. INTRODUCTION

The Golden mean canon (GM) or ratios close to it are found in the linear proportions of masterpieces of architecture, human, animal, and plant bodies. In last decades the canon is extended to the periodic system of chemical elements (Luchinskiy & Trifonov, 1981; Rakočević, 1998a, Djukić & Rakočević, 2002) and to genetic code (Rakočević, 1998b), as well as to the different natural and artificial structures, especially to nanotehnology (Koruga et al., 1993; Matija, 2004). As a noteworty fact, the Golden mean is found in masterpieces of classic literature (Stakhov, 1989; Freitas, 1989; Rakočević, 2000; Rakočević, 2003).

In the present day there are minimum two generalisations of Golden mean. First, a "vertical" generalization with x^n instead x^2 in the equation of Golden mean [Equations (2) and (4) in the next Section] (Stakhov, 1989); second, a "horizontal" generalization with $p > 1$ and/or $q > 1$ instead $p = q = 1$ [Equation (2) in the next Section] within a "family of mettalic means" (Spinadel, 1998, 1999).

2. BASIC CONCEPTS

The GM arose from the division of a unit segment line AB into two parts (Fig. 1b): first x and second $1 - x$, such that

$$\frac{x}{1-x} = \frac{1}{x} \ . \qquad (1)$$

On the other hand, one can say that GM follows from a square equation

$$x^2 \pm px - q = 0 \,, \qquad (2)$$

where $p = 1$, $q = 1$, which solutions are:

$$x_{1,2} = \frac{-1 \pm \sqrt{5}}{2} \,, \quad \text{or} \quad x_{1,2} = \frac{1 \pm \sqrt{5}}{2} \,. \qquad (3)$$

Stakhov (Stakhov, 1989, Figure 7 and Equation 26) revealed a possible generalization of GM, from which it follows:

$$x^n + x = 1 \,, \qquad (4)$$

where $n = 1, 2, 3, \ldots$.

3. A NEW GENERALIZATION

In this paper we reveal, however, a further generalization, such that in equation (2) $p = 1$ and $q = m/2$; thus, we consider the following equation:

$$x^n + x = \frac{m}{2} \ , \qquad (5)$$

where $n = 1, 2, 3, \ldots$; and $m = 0, 1, 2, 3, \ldots$.

For $n = 2$ and $m = 2$, we have the well-known GM (Fig. 1b), and for other values – the generalized GM's (Fig. 1a, c, d and Table 1). By this, the values of m correspond to the square roots of odd positive integers ($r = 1, 3, 5, 7, \ldots$), through the generalized formula (3) as:

$$x_{1,2} = \frac{-1 \pm \sqrt{r}}{2} \,, \quad \text{or} \quad x_{1,2} = \frac{1 \pm \sqrt{r}}{2} \,. \qquad (6)$$

In the following Fig. 1 we shall give the geometric and algebraic interpretations for $m = 1, 2, 3, 4$.

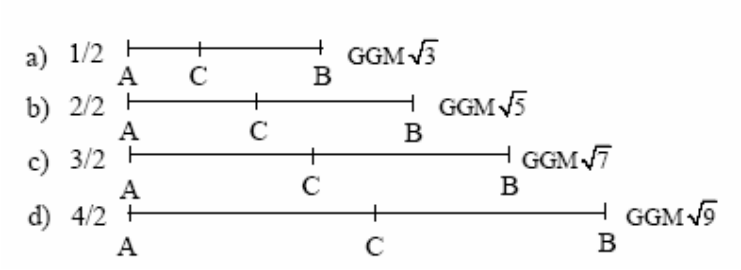


Figure 1. Generalized Golden mean by equation (5)

Consequently, it makes sense to speak about "Golden mean per root" of 1, of 3, of 5, of 7, of 9, and so on, respectively. Also, it makes sense to see the GM as an example of "the symmetry in the simplest case" (Marcus, 1989), just in the case when $n = 1$, $m = 2$. (Notice that this case is equivalent with the case $n = 2$, $m = 4$ as it is self-evident from figure 1d.)

3.1. Integer and non-integer solutions

In the following scheme (Table 1) we shall give the integer and non-integer solutions of Generalized GM.

From Table 1 it is evident that the sum of absolute values of solutions x_1 and x_2, to equation (5) equals, is $\sqrt{r}$, which represents the first cathetus (first leg) of triangle, $\sqrt{r} - m - h$. In such a triangle m is the second cathetus and h the hypotenuse. All such triangles on the left side in Table 1 appear as Diophantus' (Pythagorean) triangles (*see* Box 1), and on the right side their corresponding triangles. According to equations (2) and (6) there are four solutions, two positive and two negative, with two absolute values, as it is given in Table 1. Notice that $r - m - h$ triplets on the right side in Table 1 correspond to the Fibonacci triplets in first three cases (with $\underline{\boldsymbol{h}}$ as an ordinal number)(Mišić, 2004): 0−1−1, 1−2−3, 2−3−5 through a growth for Fibonacci distance triplet 1−1−2. In next (forth) step, with the same distance 1−1−2, the Lucas' triplet 3−4−7 appears, which grows in all further steps just for one Fibonacci distance triplet 1−1−2. Notice also the next relations: on the left side in Table 1 the left-$\underline{\boldsymbol{h}}$, as well as the left-*m*, grows for 4k units (k = 0, 1, 2, 3, ...) whereas on the right side the right-$\underline{\boldsymbol{h}}$ and right-*m* grow just for one unit; the *r* on the left corresponds with r^2 on the right; the left-N^{th} triangle appears in the right sequence through this "4k" regularity. (*Remark* 1: From the "4k" regularity follow triangles $0^{th} - \mathbf{0}^{th}$,

$1^{st} - \mathbf{4}^{th}$ (1^{st} on the left, and $\mathbf{4}^{th}$ on the right in table 1), $2^{nd} - \mathbf{12}^{th}$, $3^{rd} - \mathbf{24}^{th}$, $4^{th} - \mathbf{40}^{th}$ etc., with next solutions: $[0+(4\times 0)=\mathbf{0}]$, $[0+(4\times 1)=\mathbf{4}]$, $[4+(4\times 2)=\mathbf{12}]$, $[12+(4\times 3)=\mathbf{24}]$, $[24+(4\times 4)=\mathbf{40}]$, etc.)

N	x_1	x_2	h	m	$\sqrt{r}$		N	x_1	x_2	h	m	$\sqrt{r}$
0.	$0^2 + 1^2 = \underline{1}$ $(0 + 1)^2 = 1$			0	$\sqrt{1}$		0.	$0^2 + 1^2 = \underline{1}$ $(0 + 1)^2 = 1$			0	$\sqrt{1}$
1.	$1^2 + 2^2 = \underline{5}$ $(1 + 2)^2 = 9$			4	$\sqrt{9}$		1.	$(x_1)^2 + (x_2)^2 = \underline{2}$ $(x_1 + x_2)^2 = 3$			1	$\sqrt{3}$
2.	$2^2 + 3^2 = \underline{13}$ $(2 + 3)^2 = 25$			12	$\sqrt{25}$		2.	$(x_1)^2 + (x_2)^2 = \underline{3}$ $(x_1 + x_2)^2 = 5$			2	$\sqrt{5}$
3.	$3^2 + 4^2 = \underline{25}$ $(3 + 4)^2 = 49$			24	$\sqrt{49}$		3.	$(x_1)^2 + (x_2)^2 = \underline{4}$ $(x_1 + x_2)^2 = 7$			3	$\sqrt{7}$
4.	$4^2 + 5^2 = \underline{41}$ $(4 + 5)^2 = 81$			40	$\sqrt{81}$		4.	$1^2 + 2^2 = \underline{5}$ $(1 + 2)^2 = 9$			4	$\sqrt{9}$
5.	$5^2 + 6^2 = \underline{61}$ $(5 + 6)^2 = 121$			60	$\sqrt{121}$		5.	$(x_1)^2 + (x_2)^2 = \underline{6}$ $(x_1 + x_2)^2 = 11$			5	$\sqrt{11}$
	(...)							(...)				

Table 1. The integer and non-integer solutions of Generalized Golden Mean

In the following example we shall consider the cases $n = 2$ and $m = 1, 2, 3$.

In the first case with $n = 2$ and $m = 1$, we have q = 0.5, the first case in Table 1 on the right (the first, not the zeroth), and in Figure 1a, as the case of "GM" per $\sqrt{3}$ with the two solutions given by equations (2) and (5):

$$x_1 = \left(-1+\sqrt{3}\right)/2 = 0.3660254...$$

and

$$x_2 = \left(-1-\sqrt{3}\right)/2 = -1.3660254... \ .$$

The satisfactory solution is positive solution x_1.

In the second case with $n = 2$, and $m = 2$ we have $q = 1$, the second case in Table 1 on the right, and in Figure 1b, as the case of GM per $\sqrt{5}$ with the two solutions:

$$x_1 = \left(-1+\sqrt{5}\right)/2 = 0.6180339...$$

and

$$x_2 = \left(-1-\sqrt{5}\right)/2 = -1.6180339... \ .$$

In the third case with $n = 2$ and $m = 3$, we have $q = 1.5$, as the case of "GM" per $\sqrt{7}$, with the two solutions:

$$x_1 = \left(-1+\sqrt{7}\right)/2 = 0.8228756...$$

and

$$x_2 = \left(-1-\sqrt{7}\right)/2 = -1.8228756... \text{ , etc.}$$

4. THE METALLIC MEANS FAMILY

As it is known, it is very easy to find the members of "the metallic means family" (MMF) (Spinadel, 1999) as solutions of the equation (2), for various values of the parameters p and q. In fact, if $p = q = 1$, we have the GM. Analogously, for $p = 2$ and $q = 1$ we obtain the Silver mean; for $p = 3$ and $q = 1$, we get the Bronze mean. For $p = 4$; $q = 1$ we have the next metallic mean, etc. On the other hand, if $p = 1$ and $q = 2$, we obtain the Copper mean. If $p = 1$ and $q = 3$, we get the Nickel mean and so on. Thus, we obtain all members of the MMF, which follow from square equation (2).

Box 1. Diophantus' triangles

$1^2 = 0^2 + 1^2$

$5^2 = 4^2 + 3^2$

$13^2 = 12^2 + 5^2$

$25^2 = 24^2 + 7^2$

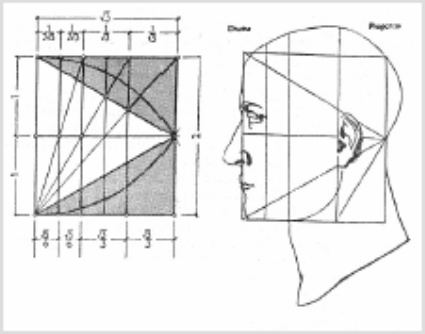
$41^2 = 40^2 + 9^2$

$61^2 = 60^2 + 11^2$

$85^2 = 84^2 + 13^2$

Diophantus' triangles are in fact the Pythagorean triangles with first cathetus "a" taken from the odd natural number series (r = 1, 3, 5, 7, 9, ...), second cathetus "b" calculated through the "4k regularity" presented in Remark 1, within Section 3.1, and hypotenuse "c", where c = b+1.

Box 2. The first Luca Pacioli's triangle



This Luca Pacioli's triangle appears as the first triangle in the right side of Table 1. The right side of Figure follows from original Pacioli's picture, while on the left side there is a Zloković's mathematical analysis (Zloković, 1955).

However, if we by (2) and (5) form the follow equation

$$x^n \pm px = \frac{m}{2}, \qquad (7)$$

where n = 1, 2, 3, ... and p = 1, 2, 3, ... , then we have a generalization of MMF; furthermore, we have a unification of "vertical" and "horizontal" generalization of GM.

Observe that De Spinadel (1999) found "the integer metallic means", for q = 2, 6, 12, 20, 30, ..., which solutions (x_1, x_2), given by equation (2), are positive integers: (1, 2), (2, 3), (3, 4), (4, 5), (5, 6) ... (Spinadel, 1999, Section 3: "Furthermore, it is very easy to verify that ... the integer metallic means, $[2, \overline{0}]$, $[3, \overline{0}]$, $[4, \overline{0}]$, ..., appear in quite a regular way")(cf. Tables 2-3).

0	**0**	0	0	0	0	0	0	0	0
0	1	**2**	3	4	5	6	7	8	9
0	**2**	4	**6**	8	10	12	14	16	18
0	3	**6**	9	**12**	15	18	21	24	27
0	4	8	**12**	16	***20***	24	28	32	36
0	5	10	15	***20***	25	**30**	35	40	45
0	6	12	18	24	**30**	36	**42**	48	54
0	7	14	21	28	35	**42**	49	**56**	63
0	8	16	24	32	40	48	**56**	64	**72**
0	9	18	27	36	45	54	63	**72**	81

(0 x 0) + 0 = 00	0 x 1 = 00
(1 x 1) + 1 = 02	1 x 2 = 02
(2 x 2) + 2 = 06	2 x 3 = 06
(3 x 3) + 3 = 12	3 x 4 = 12
(4 x 4) + 4 = 20	4 x 5 = 20
(5 x 5) + 5 = 30	5 x 6 = 30
(6 x 6) + 6 = 42	6 x 7 = 42
(7 x 7) + 7 = 56	7 x 8 = 56
(8 x 8) + 8 = 72	8 x 9 = 72
(9 x 9) + 9 = 90	

Table 2 (left). The harmonic multiplication Table of decimal numbering system. This Table contains "the integer metallic means", for q = 0, 2, 6, 12, 20, 30, 42, 56 and 72 on the diagonal in form of doublets (pairs): 0-0, 2-2, 6-6, 12-12, 20-20, 30-30, 42-42, 56-56 and 72-72.
Table 3 (right). The key of the harmonic multiplication Table. This key is related to positive integers: (0, 1), (1, 2), (2, 3), (3, 4), (4, 5), (5, 6) which appear as solutions (x_1, x_2), given by equation (2); as solutions for above given q (q = 0, 2, 6, 12, 20, 30, ...).

From Table 1 it is self-evident that Spinadel's "integer metallic means" are related to the Diophantus' triangles too (*see* Box 1), as well as to the square roots of positive integers which are squares of odd integers; thus, r = 1, 9, 25, 49, 81, 121, etc. On the other hand, the generalization, given by equation (5) is related to the square roots of all odd integers; thus, r = 1, 3, 5, 7, 9, 11, 13, etc.

5. THE EULER'S GENERALIZATION

In the history of mathematics it was known a conflict between the famous atheist philosopher Diderot and the famous religious mathematician Euler. ... One day Euler stepped for Diderot and stated: *"Sir,* $(a+b^n)/n = x$*, hence God exists; reply!"* (Eves, 1976). Diderot, as well as any one up to these days had no idea what Euler was talking about. We start here with the hypothesis (for further investigations) that Euler had the idea about "De divina proportione" of Luca Pacioli (1509) (*see* Box 2). However, after presented discussion in previous Sections of this paper we can suppose that this Euler's "divine equation" can be interpreted as a most possible generalization of GM for all cases discussed in this paper. Namely, in the case $a = b$, $n = 2$ and if and only if $x = 1/2$, we have, by equations (2) and (4) just the GM; moreover, GM stand then (accordingly to the principle "if one, then all") the case of one more extended generalization:

$x^n + x^{n-1} = 1$ (Stakhov, 1989, Equation 25), and/or of $x^n + px^{n-1} = m/2$, but that is the subject of a separate work.

6. CONCLUSION

As we can see from the discussion in the previous five Sections, some known generalizations of Golden mean, and this new one, given here, appear to be the cases of one more extended generalization, given first by Luca Pacioly, and then by Leonhard Euler. On the other hand, bearing in mind that genetic code is determined by Golden mean (Rakočević, 1998b) one must takes answer to a question (for further researches), is or is not that determination, valid for Generalized golden mean too. Certainly, the same question it takes set and for other natural and artificial systems.

ACKNOWLEDGEMENTS

I am grateful to professors Ljubomir Ćirić and Djuro Koruga, and to the doctoral dissertation candidate Nataša Mišić for stimulating discussions and benevolent critique.

REFERENCES

1. Djukić, S.T., Rakočević, M. M., The idea of three-dimensionality of periodic system in original Mendeleev's works, Flogiston, 12, pp.109-146, 2002 (in Serbian with an extended English summary).
2. Eves, H., *An Introduction to the History of Mathematics*, 4th ed., Holt, Rinehart and Winston, DTH, 1976.
3. Freitas, L. de., 515 – A symmetric number in Dante, Computers Math. Applic. 17, pp. 887-897, 1989.
4. Koruga, Dj., Hameroff, S., Withers, J., Loutfy, R., Sundershan, M., *Fullerene C_{60}: History, Physics, Nanobiology, Nanotetechnology*, Elsevier, Amsterdam, 1993.
5. Luchinskiy, G.P. & Trifonov, D.N., Some problems of classification of chemical elements and the structure of periodic system of elements, in:

The lerning of periodicity - *the hystory and the contemporaryty,* Edit. D.N. Trifonov, Nauka, Moskva, 1981 (in Russian).

6. Marcus, S., Symmetry in the Simplest Case: the Real Line. Computers Math. Applic. 17, pp. 103-115, 1989.
7. Matija, L., Reviewing paper – Nanotechnology: Artificial Versus Natural Self-Assembly, FME Transactions, Vol. 32, No 1: pp.1-14, 2004.
8. Mišić, N., Personal comunication, March, 2004.
9. Pacioli, L., *De divina proportione*, Venezia, 1509.
10. Petrović, Dj., *The theoretician and of proportions*, Vuk Karadžić, Belgrade, 1967 (in Serbian).
11. Rakočević, M.M., The harmony of periodic system of chemical elements, Flogiston, 7, pp.169-183, 1998a (in Serbian with an extended English summary).
12. Rakočević, M. M., The genetic code as a Golden mean determined system, *Biosystems*, 46, pp. 283-291, 1998b.
13. Rakočević, M.M., *Njegosh's primeval logos*, I, II, Interpress, Belgrade, 2000, 2003 (in Serbian).
14. Spinadel, V. W. de., *From the Golden Mean to Chaos,* Buenos Aires, 1998.
15. Spinadel, V.W. de., The family of metallic means, Visual Mathematics,1, No. 3, 1999. http//members.tripod.com/vismath1/spinadel/.
16. Stakhov, A. P., The Golden section in the measurement theory, *Computers Math. Applic.* 17, pp. 613-638, 1989.
17. Zloković, M., Divina proportione ≠ sectio aurea, *Pregled arhitekture,* Belgrade, 4-5, 1955 (in Serbian).

GENETIC CODE AS A HARMONIC SYSTEM: TWO SUPPLEMENTS

Miloje M. Rakočević

Department of Chemistry, Faculty of Science, University of Niš,
Ćirila i Metodija 2, Serbia (E-mail: m.m.r@eunet.yu)

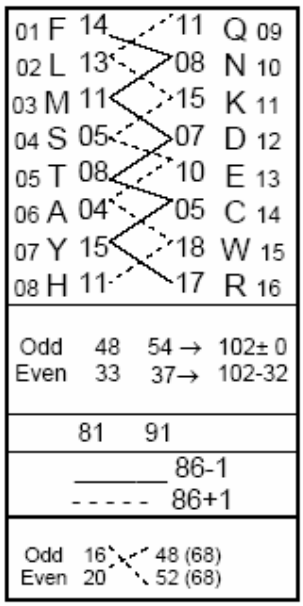
Abstract.

The paper represents two supplements to the source paper, q-bio.OT/0610044, with two new series of harmonic structures of the genetic code, determined by Gauss arithmetical algorithm through atom number balances. By this the determination itself appears as a connection between Gauss algorithm and polarity of molecules (about polarity: in Suppl. 2).

SUPPLEMENT 1

1. Introduction. In this supplement we give a new series of harmonic structures of the genetic code, determined by Gauss arithmetical algorithm as it is shown in the source paper (Rakočević, 2006a), q-bio.OT/0610044. In source paper beside others, we showed that 16 non-contact canonical amino acids (AAs), in an amino acid molecule *size order*, are determined by a Gauss' arithmetical algorithm. In this supplement however we show that the same is valid – for the same 16 non-contact amino acids – for their *coding order* (ordinal number) in Genetic Code Table (GCT). The difference is in the fact that here are valid only two last steps from the said Gauss' algorithm. Namely, from all Gauss' algorithm quantums – 11, 21, 31, 41,( ), 61, 71, 81, 91 – only two last steps are here in the „game"; the quantums „81" and „91" as well as their arithmetical mean 86±0. (Cf. source paper-Figure 1 and see the positions of quantums 86±1 in source paper-Figure 1.1.) As in source paper, all determinations are realized through principle of minimum change, i.e. through the unit atom number balances in first or in the second position of the digit-number-notation, respectively (x±00, x±01; x±10), (y±00, y±01; y±10).

2. Results and Discussion. In following 16 illustrations (Tables) are given the results of calculations of atom number within amino acid side chains; the calculations, related to the amino acid system built either from only 16 non-contact or from all 20 AAs (16 non-contact plus 4 contact AAs). By this the rest of four contact AAs make: Glycine (G), Proline (P), Valine (V) and Isoleucine (I). After our hypothesis (Hypothesis 1) there are some other possibilities of the amino acid splitting into 4 special and 16 other AAs; such a splitting which is related to the atom number balances (x±00, x±01; x±10), (y±00, y±01; y±10). The possible examples are: Serine (S), Threonine (T), Cisteine (C) and Methionine (M) as chalcogene AAs (chalcogene because they possess oxygen or sulfur in molecule side chains); then carboxylic AAs and their amide derivatives: Aspartic acid (D), Glutamic acid (E), Asparagine (N) and Glutamine (Q); four aromatic versus 16 aliphatic etc.



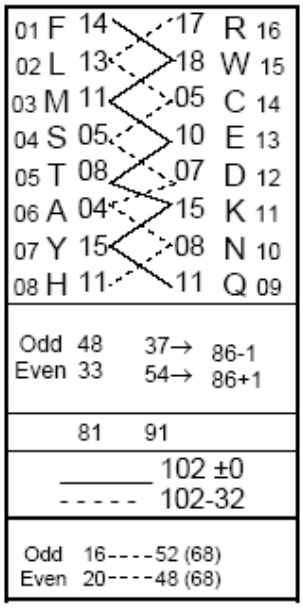
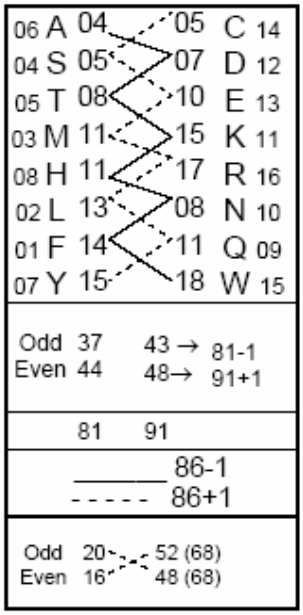


Table 1 [left]. This Table is analog with Table 1.2 in source paper. The 16 non-contact AAs arranged into two octets, correspondingly to their ordinal number in GCT, i.e. correspondingly to ordinal number of belonging codons; the first octet on the left and the second one (in the sequence up/down) on the right. The calculation for the ordinal numbers for all 16 amino acids (their sums) is given just down (16+20 = 36 and 48 + 52 = 100).
Table 2 [right]. All is the same as in previous Table, except the second octet (on the right) is given in a vice-versa sequence (down/up).

Remark 1: For the terms „contact“ and „non-contact“ AAs as well as for all other new terms see source paper q-bio.OT/0610044.

Remark 2: Hypothesis 1 follows from the idea that quantum „4“ represents the first possible case for the existence of *one pair of pairs* (cf. legend given for Figure 1 in source paper). On the other hand the equation 16 + 4 = 20, corresponding to square equation $x^2 + x = 20$, is one and only case from the family of metalic means, in relation to the Golden mean (Rakočević, 1998, 2004a; 2006b – Table 2), which corresponds with the middle point in the Harmonic multiplication table.

The first two Tables (Tables 1 & 2) show a correspondence with the natural numbers series from one side and with the Gauss' arithmetical algorithm from the other side (with quantums "81" and "91" and their arithmetic mean 86±0 and/or 86±1).



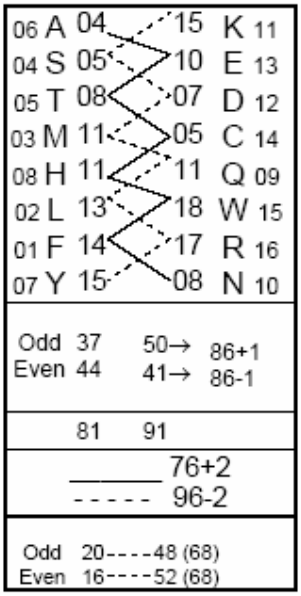
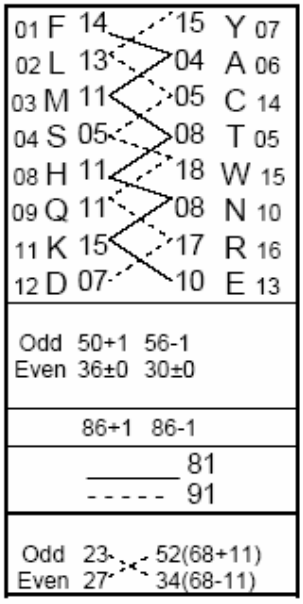


Table 3 [left]. The Table follows from Table 1, so that the left octet is given here in a sequence of molecule sizes. In spite of the fact that the sequence of ordinal numbers is disrupted, the odd/even quantums are the same as in Table 1, that means: 16, 20, 48 and 52, but in a vice versa arrangement.

Table 4 [right]. This Table follows from Table 2 at the same manner as the Table 3 follows from Table 1. Notice that 76 equals 86 – 10 and 96 equals 86 +10.

The amino acid pairs in odd row positions, in Table 1, are larger than the pairs in even row positions. (For example, the amino acid pair F-Q with 25 atoms is larger than L-N with 21 atoms; then the amino acid pair M-K with 26 atoms is larger than S-D with 12 atoms, etc.) Notice also that atom number within side chains of "odd" AAs is determined with multiples of number 6 (8 x 6 = 48 and 9 x 6 = 54). [Cf. these two quantums (48 and 54) with the same quantums in Figure 2.1 in source paper.]



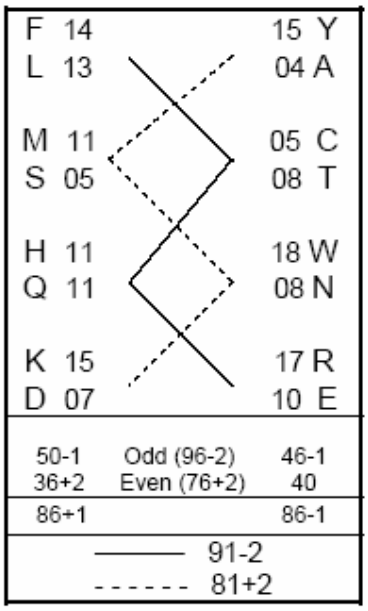
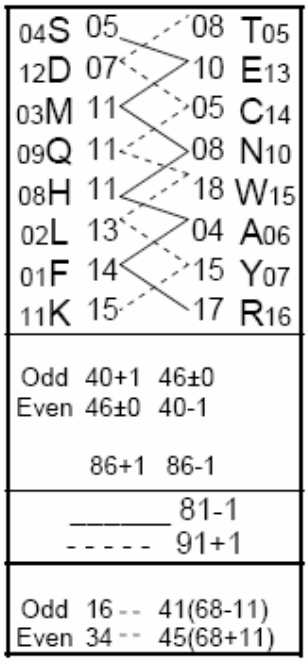


Table 5 [left]. The choice of non-contact AAs from GCT. Each next amino acid chooses its own pair-member from all four columns of GCT; on the other words, each amino acid from the left octet is chosen together with its chemically corresponding pair-member in the right octet. Except a determination with quantums "81" and "91" there is a determination with "half" quantums (81 – 1):2 = **40** and (91 + 1):2 = **46** in forms: 40±10 and 46±10. In spite of the fact that the sequence of ordinal number is disrupted, the odd/even quantums (sums) are only in a balance change (±11)
Table 6 [right]. All is the same (*mutatis mutandis*) as in previous Table, except the pairing process which is arranged horizontally as well vertically. This Table is analog with source paper-Tables 1.1 and 2.1.

L13	A04	Y15	F14		46 (36+10)
S05	T08	C05	M11		29
Q11	N08	W18	H11		48 (58-10)
D14	E15	R17	K18		49
36	30	56-1	50+1		86+1 / 86-1

L13	A04	Y15	F14	G01	47
S05	T08	C05	M11	I13	42
Q11	N08	W18	H11	P08	56
D14	E15	R17	K18	V10	59
V10	I13	G01	P08	102+1 / 102-1	
46	46-3	56	56+3	102 / 102	

Table 7 [left]. This Table follows from Table 6 at the same manner as source paper-Figure 1 follows from source paper-Table 1.1. (Cf. quantums 29, 36, 49 and 58 in source paper-Fig. 1.)
Table 8 [right]. The same is valid as for previous Table; except, the contact AAs are added.



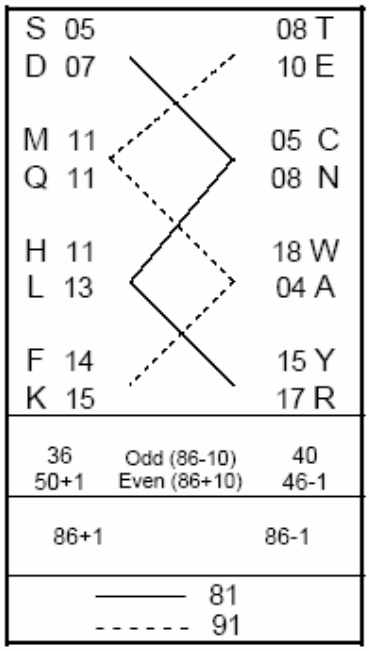


Table 9 [left]. This Table follows from Table 5 (as Table 3 from Table1).
Table 10 [right]. This Table follows from Table 9 at the same manner as Table 6 from Table 5.

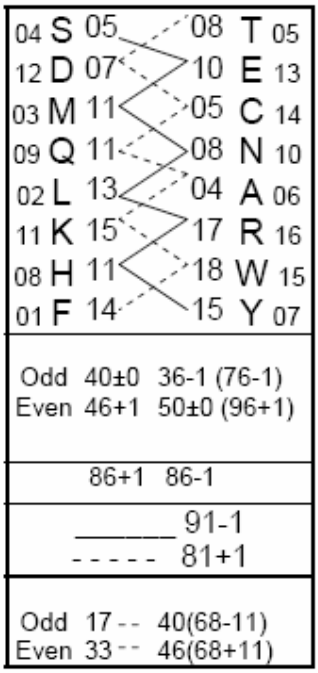
The sum of two (above mentioned) quantums in Table 1 equals: 48+54 = 102 atoms within 8 molecules, where 102 represents a half of total atom number

within 20 canonical AAs, i.e. within their side chains. On the other hand "even" AAs in Table 1 possess 70 atoms which quantum together with the quantum of 32 atoms, existing in four contact AAs (G 01+ P 08 + V 10+ I 13 = 32) equals 102 atoms still once; 102 atoms within 12 molecules. (Proportion 1:1 for atom number and 2:3 for molecule number.) [For proportion 1:1 cf. Marcus (1989) and Stakhov (1989); for proportion 2:3 cf. Moore (1994).] All other illustrations follow analogously and logically – next from previous – as it is shown in their legends.

D07	E10	T08	S05		30
Q11	N08	C05	M11		35 91
L13	A04	W18	H11		46 81
K15	R17	Y15	F14		61
46	40-1	46	40+1		86-1 / 86+1

D07	E10	T08	S05	V10	40
Q11	N08	C05	M11	I13	48
L13	A04	W18	H11	P08	54
K15	R17	Y15	F14	G01	62
V10	I13	G01	P08	102/102	
56	52	47	49	102+1/102-1	

Table 11 [left]. This Table follows from Table 10 at the same manner as source paper-Figure 1 follows from source paper-Table 1.1.
Table 12 [right]. The same is valid as for previous Table; except, the contact AAs are added.



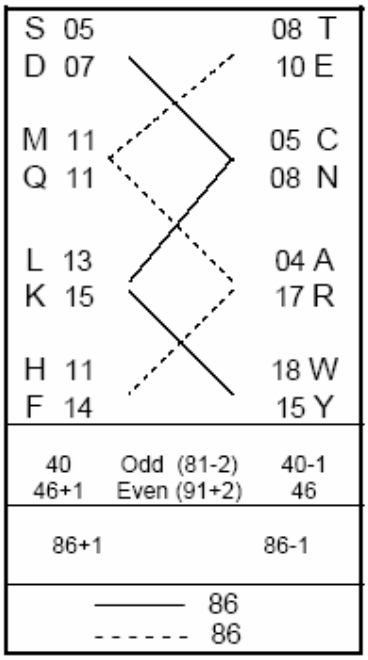
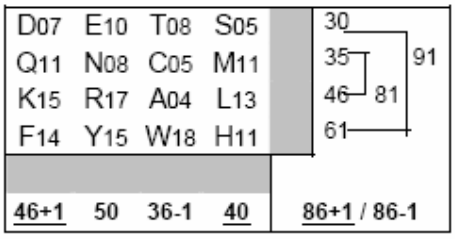


Table 13 [left]. The order as in Table 5 except a distinction into aliphatic versus aromatic AAs (the first must be aliphatic AAs as in source paper-Table 1.1)
Table 14 [right]. This Table follows from Table 13 at the same manner as Table 10 follows from Table 9.



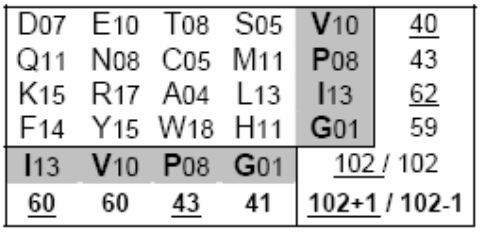


Table 15 [left]. This Table follows from Table 14 at the same manner as source paper-Figure 1 follows from source paper-Table 1.1.
Table 16 [right]. The same is valid as for previous Table; except, the contact AAs are added.

SUPPLEMENT 2

1. Introduction. In this supplement we give still a set of harmonic structures of the genetic code, determined by Gauss arithmetical algorithm as it is shown in the source paper (Rakočević, 2006a), q-bio.OT/0610044. In source paper beside others, we showed that polarity of amino acids (AAs) indirectly is determined by Gauss arithmetical algorithm. Namely, the polar AAs are positioned (within a 4 x 5 amino acid system) as a separate entity, in the form of a specific "island" surrounded by non-polar AAs (source paper – Figure 2.1). By this, the four ambivalent AAs, i.e. polar and nonpolar at the same time (glycine, proline, tryptophan and histidine)[1] are positioned in a snug "string" at the very edge of the system.

In this supplement, however, we show a direct determination, i.e. a direct connection between Gauss arithmetical algorithm and polarity of amino acid molecules, including their positions in standard Genetic Code Table (GCT). In such a case the connection itself appears to be followed by a strict atom number balance through the existence of new four amino acid classes.

2. Results. The starting step is the choice of one and only single amino acid pair, G-P (Table 1); single pair because glycine is one and only amino acid within glycinic stereochemical type and proline also one and only, but within prolinic stereochemical type (about four stereochemical types see in Popov, 1989 and Rakočević & Jokić, 1996). Follow two pairs of aromatic AAs, first H-W and second F-Y. As the first pair is H-W because the quartet G-P/H-W represents ambivalent AAs – polar and nonpolar at the same time (cf. Footnote 1). The next steps are comming in relation to the AAs sequence, such as it occurs in the first column of GCT.

As a noteworthy is the fact that after four ambivalent come four extrem AAs, extrem just from the aspect of polarity: F-Y/L-R. Phenylalanine is the most nonpolar amino acid of all four aromatic AAs and Leucine (together with

[1] Glycine: after hydropathy is polar; after cloister energy and polar requirement is non-polar; Proline: after hydropathy and cloister energy is polar; after polar requirement is non-polar; Tryptophan: after hydropathy and polar requirement is polar; after cloister energy is non-polar. Hydropathy (Kyte & Doolittle, 1982); cloister energy (Swanson, 1984); polar requirement (Woese et al., 1966; Konopel'chenko and Rumer, 1975). Really regarding, histidine as a „semi-ambivalent" amino acid since it has neither positive nor negative value in cloister energy, but its value is equal to zero (Figure 5 in Swanson, 1984). About the pairing process of AAs through Hydropathy and Cloister energy *see* Survey 1 in Rakočević & Jokić, 1996, and about ambivalence of glycine and proline see Section 3.3 in Rakočević, 2004b.

isoleucine) is the most nonpolar of all aliphatic AAs (after hydropathy). On the other hand, Tyrosine is extrem because its polaryty comes not only from the aromatic ring but from a polar functional group (OH); and arginine is extrem through its very massive guanidino group – very massive, and very polar at the same time.

As we can see from equations 1-4 and from Table 2, the four extreme AAs are the same AAs that we know from the balance distinction between **p**olar/**n**onpolar[2] and **i**nner/**o**uter AAs in GCT, with the reading: 22 molecules, 222 atoms within amino acid side chains and 420 atom within whole molecules, etc. (Rakočević, 2000, 2006c):

(n) 4V+1M+3I+4A+**2L**+**4L**+**2F**+2C = 22
40+ 11+39+16+26 + 52 +28+10 = 2**2**2 (42**0**) (1)

(o) 4V+1M+3I+4A+**2Y**+**4R**+**1W**+2C = 21
40+ 11+39+16+ 30 + 68 +18+ 10 = 2**3**2 (42**1**) (2)

(p) 4G+2K+2N+4P+**2Y**+**4R**+**1W**+2E+2D+4T+2R+2S+2Q+2H+4S = 39
04+30+16+32+30 + 68+ 18+20+14+32+34+10+22+22+20 = 3**7**2 (72**3**) (3)

(i) 4G+2K+2N+4P+**2L**+**4L**+**2F**+2E+2D+4T+2R+2S+2Q+2H+4S = 40
04+30+16+32+26+52 + 28+20+ 14+32+ 34 +10+22+ 22+ 20 = 3**6**2 (72**2**) (4)

(p) 6S+4T+2N+2Q+2D+2E+2K+6R+2Y = 28
30+32+16+ 22+ 14+20+ 30+102+30 = 297–1 (549–1) (5)

(n) **4G**+**4P**+**1W**+**2H**+4A+6L+4V+3I+2C+1M+2F = 33
04+32+ 18+ 22+ 16 + 78 +40 +39+10+11+ 28 = 297+1 (594+1) (6)

Equations (5) and (6) show the relations which are more than a balance; a specific relationship of the amino acid "heads" and "bodies" (all "heads" and all "bodies") from one side and the wholeness of molecules from the second side. So, the distribution of atom number is the following. In polar AAs molecules there are exactly as many atoms as in the heads of all molecules, minus one atom (61 x 9 = 549). In the "nonpolar" molecules (nonpolar plus ambivalent) there are as many atoms as in the bodies, i.e. side chains of all 61 molecules, plus one atom [(297+1) + (33 x 9 = 297) = 594 +1]. On the other hand, in the bodies of

[2] Polar/nonpolar AAs after their hydropathy (cf. Footnote 1).

polar as well as of non-polar molecules, there is 297±1 atoms, exactly as in first (297-1) and second half (297+1) of standard GCT[3]. Thirdly, in the heads of non-polar molecules there are exactly as many atoms as it is the half of the atom number in the bodies of all molecules (33 x 9 = ½ 594), while in the heads of polar AAs there are as many as it is the half of 5**0**4, which number is the modular pair (in module 9) of number 5**9**4.

Remark 1. Number 504 represents the sum of each two and two (out of eight in total) branches on the 6-bit binary tree (Figure 1 in Rakočević, 1998), or of two and two octets in GCT as follows: in two central octets there are: 24+25+ ...+31 = **220** and 32+33+ ...+39 = **284** (220+284 = 504); in next two octets there are: 16+17+ ...+23 = 156 and 40+41+ ...+47 = 348 (156+348 = 504); within the first to last pair of octets there are: 8+9+ ...+15 = 92 and 48+49+ ...+55 = 412 (92+412 = 504); Finely, within the last octet pair we have: 0+1+ ...+7 = 28 and 56+57+ ...+63 = 476 (28+476 = 504)(cf. Table 3).

Remark 2. The number 504 represents the sum of the first two friendly numbers (the first pair): 220+284 = 504. On the other hand, Shcherbak has shown (1994) that within the set of 23 AAs, the eight four-codon AAs possess 592+333 = 925 nucleons, where 592 is a half of the third friendly number (1184 = 2 x 592)[4]. But Shcherbak also showed that within side chains of all 23 AAs there are 1443 of nucleons (333 nucleons within four-codon AAs plus 1110 within non-four-codon AAs), where 1443 is a sixth part of the sum of first four perfect numbers [1443 x 6 = 8658 = (7770+0888) = 6+28+496+8128] (About determination of the genetic code with the perfect and friendly numbers see in Rakočević, 1997).

Parallel with the atom number balance there is a molecule number balance as follows: the unit distances between the number of molecules: 22-21 = 1 and 40-39 = 1 in equation (1) in relation to equation (2) and equation (4) in relation to equation (3), respectively; then the double units distances between the number of

[3] Table 2 appears to be the standard GCT if the texture (dark tones) is excluded. In such a manner the left half of GCT make 32 amino acid molecules encoded by 32 NYN codons, whereas the right half make 29 amino acid molecules encoded by 29 NRN codons, plus three stop codons. Within 32 amino acid molecules (side chains) there are 297-1, and within 29 molecules 297+1 of atoms.

[4] The forth friendly number is the number 1210 as a product of 10 x 11^2. The third and forth friendly numbers make the second friendly number pair.

molecules: 33-22 = 11 and 39-28 = 11 in equation (6) in relation to equation (1) and equation (3) in relation to equation (5), respectively.

G 01	H 11	**F** 14	**L** 13	**I** 13	**M** 11	**V** 10
P 08	W 18	Y 15	R 17	K 15	C 05	A 04
			Q 11	N 08	S 05	T 08
			E 10	D 01		

Table 1. First four AAs are ambivalent, next four "extreme" (as it is explained in the text) and other AAs in a chemically determined order and arrangement.

After first two AAs quartets in Table 1 (G-P/H-W and F-Y/L-R) follow other AAs through a chemically relevant relation: two source aliphatic AAs (L-I)[5] and two amino derivatives (R-K); then two sulfur AAs (M-C) and two source aliphatic still once (V-A). The next two are the chalcogene AAs: S-T in a continuation to M-C; amino acid T also in contact with A as two methyl derivatives. At the end come two carboxylic AAs (D-E) whose amide derivatives, as nitrogen compounds, hold a connection with other two nitrogen derivatives (R-K).

[5] Notice that F-L make also a chemical pair through the same structural motive – the first possible branching (iso-butane in relation to toluen structural motive). Here lies the reason why benzene ring is axcluded from the set of aromatic AAs.

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

Table 2. The amino acids within three diagonals are inner, and other – outer AAs. On the other side, within bordered space are polar AAs and other – nonpolar AAs (polar/nonpolar after hydropathy: Kyte & Doolittle, 1982).

/00 - 07/08 - 15/16 - 23/24 - 31//32 - 39/40 - 47/48 - 55/56 - 63/
28 92 156 220 284 348 412 476
64 64 64 64 64 64 64

/00 - 07/00 - 15/00 - 23/00 - 31//00 - 39/00 - 47/00 - 55/00 - 63/
28 120 276 496 780 1128 1540 2016
92 156 220 284 348 412 476

Table 3. The eight octets within 6-bit binary-code tree (Rakočević, 1998) as well as within GCT are determined with the first pair of friendly numbers (220 & 284) and third perfect number (496). For details see the text, especially Remarks 1 & 2.

G	01	08	**P**
H	11	18	**W**
F	14	15	**Y**
L	13	17	**R**
I	13	15	**K**
M	11	05	**C**
V	10	04	**A**
T	08	05	**S**
Q	11	08	**N**
E	10	07	**D**
Odd	**49**	**50**	
Even	~~**53**~~	~~**52**~~	
	102	102	

G	001	041	**P**
H	081	130	**W**
F	091	107	**Y**
L	057	100	**R**
I	057	072	**K**
M	075	047	**C**
V	043	015	**A**
T	045	031	**S**
Q	072	058	**N**
E	073	059	**D**
O	**264**	**293**	
E	~~**331**~~	~~**367**~~	
	595	660	

Table 4 [left]. This Table follows from Table 1. Atom number determination in relation to Gauss' algorithm (explanation in the text)
Table 5 [right]. All is the same as in previous Table, except the determination by nucleon number. Notice a symmetry determination through module 9: 5**9**5 versus 66**0**.

From Table 1 follows Table 4 (in relation to Table 5), first row on the left and second row on the right, plus AAs from third and fourth rows – three outer (T, Q, E) on the left and three inner (S, N, D) on the right. As we see AAs in odd and even positions make four AAs groups with atom number directly determined by Gauss' arithmetical algorithm (Table 6). Namely, in source paper (source paper – Figure 1) we showed that within four amino acid rows there are so many atoms as in 10^{th} and 20^{th} Gauss' pair (in relation to middle point "51": 10^{th} pair as 41-61 and 20^{th} pair as 31-71 atoms). And here we see that within four amino acid rows are so many atoms as in 01^{st} and 02^{nd} Gauss' pair (01^{st} pair as 50-52 and 02^{nd} pair as 49-53 atoms).

G	P	W	**H**
F	Y	R	**L**
I	K	C	**M**
V	A	S	**T**
Q	N	D	**E**
49	50	52	**53**

G	M	P	**C**
H	V	W	**A**
F	T	Y	**S**
L	Q	R	**N**
I	E	K	**D**
52	50	73	**29**

Table 6 [left]. The Table follows from Table 4: two outer columns as odd/even positions in left column of table 4, and two inner columns as odd/even positions in right column of table 4, respectively.

Table 7 [right]. As previous one, this Table also follows from Table 4: first five and last five AAs as two outer columns, and 5 & 5 amino acids as two inner columns. The atom number quantums "50" and "52" are the same as in previous Table (Table 6), whereas two other quantums (73 = 71+ 2 and 29 = 31 – 2) correspond to the Gauss pair 31-71 through a deviation of ±2 (minus first and second step; plus first and second step).

The sum of two quantums in Table 1 equals: 48+54 = 102 atoms within 8 molecules, where 102 represents a half of total atom number within 20 canonical AAs, i.e. within their side chains. On the other hand "even" AAs in Table 1 possess 70 atoms which quantum together with the quantum of 32 atoms, existing in four contact AAs (G 01+ P 08 + V 10+ I 13 = 32) equals 102 atoms still once; 102 atoms within 12 molecules. (Proportion 1:1 for atom number and 2:3 for molecule number.) [For proportion 1:1 cf. Marcus (1989) and Stakhov (1989); for proportion 2:3 cf. Moore (1994).] All other illustrations follow analogously and logically – next from previous – as it is shown in their legends.

3. Conclusion for both Supplements. Bearing in mind that the order of amino acids in presented Tables, in both Supplementsis, is given in correspondence with the order of codons in GCT, it makes sense to speak about genetic code as a harmonic system. On the other side, presented harmonic structures provide evidence to support the hypothesis, given in a previous paper (Rakočević, 2004b), that genetic code was complete from the very beginning as the condition for the origin and evolution of the life.

REFERENCES

Konopel'chenko, B.G., Rumer, Yu. B., 1975. Klassifikaciya kodonov v geneticheskom kode. Dokl. Akad. Nauk. SSSR. 223, 471-474.

Kyte, J., Doollittle, R. F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Marcus, S., 1989. Symmetry in the simplest case: the real line. Comput. Math. Appl. 17, 103-115.

Moore, G.A., 1994. The limit of the golden numbers is 3/2. The Fibonacci Quarterly, June-July, 211-217.

Popov, E. M., 1989. Strukturnaya organizaciya belkov. Nauka, Moscow (in Russian).

Rakočević, M. M., 1997. Genetic code as a unique system. SKC, Niš pp. 60-64.

Rakočević, M. M., 1998. The genetic code as a Golden mean determined system. Biosystems 46, 283-291.

Rakočević, M. M., 2000. The factors of the classification of protein amino acids. GLASNIK of Montenegrin Academy of Science and arts, 13, 173-194.

Rakočević, M. M., 2004a. Further generalization of Golden mean in relation to Euler's „divine" equation. FME Transaction (Faculty of Mechanical Engineering, Belgrade, Serbia), 32, 95-98.

Rakočević, M. M., 2004b. A harmonic structure of the genetic code. J. Theor. Biol. 229, 463 – 465.

Rakočević, M. M., 2006a. Genetic code as a harmonic system. q-bio.OT/0610044.

Rakočević, M. M., 2006b. Further Generalization of Golden Mean in Relation to Euler Divine Equation. math. GM/0611095.

Rakočević, M. M., 2006c. The factors of the classification of protein amino acids. q-bio.BM/0611004.

Rakočević, M. M., Jokić, A., 1996. Four stereochemical types of protein amino acids: synchronic determination with chemical characteristics, atom and nucleon number. J. Theor. Biol. 183, 345 – 349.

Shcherbak, V. I., 1994. Sixty-four triplets and 20 canonical amino acids of the genetic code: the arithmetical regularities. Part II. J Theor. Biol. 166, 475-477.

Stakhov, A. P., The Golden section in the measurement theory, Computers Math. Applic. 17, pp. 613-638, 1989.

Swanson, R., 1984. A unifying concept for the amino acid code. Bull. Math. Biol. 46, 187-207.

Woese, C.R., et al., 1966. On the fundamental nature and evolution of the genetic code. In: Cold Spring Harbor Symp. Quant. Biol., 31, 723-736.

A NEW GENETIC CODE TABLE

Miloje M. Rakočević

Department of Chemistry, Faculty of Science, University of Nis;
Home address: Milutina Milankovica 118, Belgrade: 11070, Serbia
(E-mail: m.m.r@eunet.yu ; www.sponce.net)

Abstract

In this paper it is shown that within a Combined Genetic Code Table, realized through a combination of Watson-Crick Table and Codon Path Cube it exists, without an exception, a strict distinction between two classes of enzymes aminoacyl-tRNA synthetases, corresponding two classes of amino acids and belonging codons. By this, the distinction itself is followed by a strict balance of atom number within two subclasses of class I as well as two subclasses of class II of amino acids.

1 INTRODUCTION

Among the representations of the genetic code, two retain our attention here: the Watson-Crick Table, WCT (the standard genetic code Table) (Watson & Crick., 1953a,b; Crick, 1966, 1968), arranged through an ordering of codon letters that shows the 1st /2nd/ 3rd letters in the order UCAG/ UCAG/ UCAG (first permutation in Table 1), and the so-called Codon Path Cube, CPC (Swanson, 1984) with the order UCGA/UCGA/CUAG (second and third permutations in Table 1, respectively).

On the other hand, Wetzel (1995) has shown that a strict distinction of canonical amino acids (AAs) within the WCT in correspondence with two classes (I & II) of enzymes aminoacyl-tRNA synthetases (aaRS) is valid only for the X**U**X-X**C**X, but not for the X**A**X-X**G**X codon space[1]. Thus, as the AAs, handled by class I of aaRS within the X**U**X-X**C**X space, appear V, M, I, L whereas the AAs, handled by class II of aaRS are F, S, P, T, A. [Eriani et al., 1995, p 499: "The class I enzymes chare with dehydrogenases and kinases *the classic* nucleotide binding *fold* called the Rossmann fold, whereas the class II enzymes possess *a different fold, not found elsewhere*, built around a six-

[1] Wetzel, 1995, p. 545: "all XCX codons code for amino acids handled by class II synthetases, and all but one of the XUX codons code for amino acids handled by class I synthetases"; cf. also Fig. 1 in the cited work on the page 546.

stranded antiparallel beta-sheet. The two classes of synthetases ... differ as to where on the terminal adenosine of the tRNA the amino acid is placed: class I enzymes act on the 2′ hydroxyl whereas the class II enzymes prefer the 3′ hydroxyl group". (My *italics* and my comment: the C atom on the position 2′ is a stronger nucleophile whereas the C atom on the position 3′ is a less strong nucleophile).]

1	**U C A G** / G A C U	**4**
2	U C G A / A G C U	**3**
3	C U A G / G A U C	**2**
4	C U G A / A G U C	**1**
1	**U A C G** / G C A U	**4**
2	U A G C / C G A U	**3**
3	A U C G / G C U A	**2**
4	A U G C / C G U A	**1**
1	U G C A / A C G U	**4**
2	U G A C / C A G U	**3**
3	G U C A / A C U G	**2**
4	G U A C / **C A U G**	**1**

Table 1 The 12+12 permutations of the sequence of four amino bases, two of pyrimidine (Py) and two of purine (Pu) type, arranged through three octets, in accordance to three main characteristics (footnote 4) of amino bases. Only first three permutations are included in CT.

It is also shown (Rakočević, 1997a,b) that within the CPC (Figure 1 in Rakočević, 1997a) the mentioned distinction is valid for the whole X**U**X-X**C**X-X**A**X-X**G**X codon space, *with* only one exception. Namely, within space I there are AAs: M, I, V, L, Y, Q, E, C, W and pyrimidine-coding arginine R, all handled by class I of aaRS, and within space II: F, S, P, A, T, G, K, N, D, H, all handled by class II, and purine-coding arginine R, handled by class I of aaRS.

In this paper, however, we show that within a Combined Genetic Code Table (CT), combined in a specific manner from the WCT and the CPC (Table 2 in

relation to Table 3), the strict distribution of the AAs over the two classes of aaRS occurs *without* an exception.

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

Table 2. A new Table of genetic Code as a Combined Table, realized in a specific combination of Watson-Crick Table and Codon Path Cube (Swanson, 1984), through the first three permutations from the first octet in Table 1. Bold positions (dark tones): the codons coding for AAs handled by class II of aaRS; non-bold positions: the codons coding for AAs handled by class I aaRS plus three "stop" codons, denoted with *.

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

Table 3. A variant of Combined genetic code table, realized through *all four* left permutations from first octet in Table 1. Bold positions (dark tones): the codons coding for AAs handled by class II of aaRS; non-bold positions: the codons coding for AAs handled by class I of aaRS. Notice that fourth permutation (CUGA), in determination of third codon position, can be used only in three rows because the "middle U" is a "gooseneck" – the distinction within three isoleucine codons is not allowed. (About relationships between Table 2 & 3, *see* in Discussion.)

The first subclass of class I (1 in Table 4) is very down in Table 2 and the second one (1' in Table 4) in the top, on the right. On the other hand, the first subclass of class II (2 in Table 4) is in a middle space (below in Table 2) whereas the second one is above on the left (2' in Table 4).

1	11 + 39 + 40 + 78 = 168	
	M + I + V + L = (4)	
2	32 + 16 + 32 + 20 + 28 = 128	
	T + A + P + S + F = (5)	
2'	16 + 14 + 22 + 30 + 10 + 04 = 096	
	N + D + H + K + S + G = (6)	
1'	30 + 22 + 20 + 34 + 68 + 10 + 18 = 202	
	Y + Q + E + R + R + C + W = (7)	

Table 4. The atom and molecule number balance. The two (1 & 1') and two (2 & 2') rows correspond to two and two subspaces (subclasses) within the Combined Table, presented here in Table 2.

Survey 1. *The logic of codon position determination (with I, II and III permutations) in Table 2**

Perm.	**Posit.**	**Rows**
II	2	2 outer
	2	2 inner
	1	2 outer
	3	2 inner
III	3	2 outer
I	1	2 inner

*The order: outer/inner/outer/inner

Survey 2. *Atom number balances*

Nucleotides		Amino acids (Tab. 4)
UMP (34) + **G**MP (38) = 72 ± 0		
CMP (35) + **A**MP (37) = 72 ± 0		
	73 – 1 =	168 – 96 (1 – 2')
UMP (34) + **A**MP (37) = 72 – 1		
CMP (35) + **G**MP (38)= 72 +1		
	73 + 1 =	202 – 128 (1' – 2)

As we can see from Table 4, the calculation of the number of atoms within amino acid molecules (side chains), requires that each molecule be included as many times as there are codons coding for it. By this one must notice the situation R + R (meaning that R occurs two times) in row 1' because arginine possesses two spaces within the WCT; and the situation with only one L in row 1 because Leucine possesses only one space (through six neighbor codons) within the same Table. Certainly, serine also must appear two times (S in space "C" in row 2 and S in space "G" in row 2') in accordance with its two spaces within the WCT[2].

In addition, this CT shows the existence of a strict balance of the number of atoms within the amino acid molecules (their side chains) according to the *principle of minimum change* (Table 4 in relation to Surveys 2 & 3). As it will be shown below, the word is about the changes exactly for ±0 or ±1 of atoms in relation to arithmetic mean for two and two amino acid subclasses.

Together with all these balances, there is a balance of the number of molecules: 1+1' = 2+2', i.e. 4+7 = 5+6 (where the principle of minimum change is also valid through the sequence of the 4-5-6-7).

[2] In contrary, in Shcherbak's nucleon number balance between four-codon and non-four-codon amino acids (Shcherbak, 1994; 2003) Leucine possesses two positions too, because it is distributed into two four-codon families. This fact affirms a conclusion that multi-meaning states are from the inherent characteristics for the genetic code.

Survey 3. *Relations within Table* 4

(1 + 1') 168 + 202 = **297** + 073 = 370

(2 + 2') 128 + 096 = **297** - 073 = 224

(1 + 2') 168 + 096 = 8 x 33

(1' + 2) 202 + 128 = 10 x 33

(1 + 2) 168 + 128 = (9 x 33) - 1

(1' + 2') 202 + 096 = (9 x 33) + 1

9 x 33 = **297**

[(Py-Py)U]: F_2 (28) + L_2 (26) = 54	[(Py-Pu)U]: L_2 (26) + L_2 (26) = 52
[(Py-Py)C]: S_2 (10) + P_2 (16) = 26	[(Py-Pu)C]: S_2 (10) + P_2 (16) = 26
[(Py-Py)A]: Y_2 (30) +H_2 (22) = 52	[(Py-Pu)A]: ** (00) + Q_2 (22) = 22
[(Py-Py)G]: C_2 (10) + R_2 (34) = 44	[(Py-Pu)G]:*W(18) + R_2 (34) = 52
[(Pu-Py)U]: I_2 (26) + V_2 (20) = 46	[(Pu-Pu)U]: IM (24) + V_2 (20) = 44
[(Pu-Py)C]: T_2 (16) + A_2 (08) = 24	[(Pu-Pu)C]: T_2 (16) + A_2 (08) = 24
[(Pu-Py)A]: N_2 (16) +D_2 (14) = 30	[(Pu-Pu)A]: K_2 (30) + E_2 (20) = 50
[(Pu-Py)G]: S_2 (10) + G_2 (02) = 12	[(Pu-Pu)G]: R_2 (34) + G_2 (02) = 36

Table 5. The distribution of the canonical AAs in correspondence to the first-third-letter codon rule (Siemion & Siemion, 1994) in WCT, over the number of atoms within amino acid molecules (side chains). The index designates the number of codons, coding for the given amino acid.

I - III	U	C	A	G	
Py – Py	54	26	52	44	176
Py – Pu	52	26	22	52	152
Pu – Py	46	24	30	12	112
Pu – Pu	44	24	50	36	154
	196	100	154	144	

330-066 **330**

[(330-033) –1] [(330-033) +1]

Table 6. The atom number balances within WCT after first-third-letter codon position rule (Siemion & Siemion, 1994) expressed through the number of atoms within amino acid molecules (side chains). First column (I-III) designates the type of the base in first-third position of corresponding codons. The letters U, C, A, G are related to four columns in WCT. Within two inner and two outer rows as well as within two first and two second columns there are (8 x 33), [(9 x 33)±1], and (10 x 33) of atoms, respectively.

From Survey 2 it follows that *the differences* in the number of the atoms within amino acid subclasses 1' and 2 (202-128) equals 73+1 and within subclasses 1 and 2' (168-96) equals 73 –1 or, in other words, both in relation to arithmetic mean 73±1. On the other hand, *the sums* of the atoms within two and two types of nucleotide molecules equal as follows; in first case: **C**MP + **G**MP = 72+1 and **U**MP + **A**MP = 72 –1, or, both in relation to arithmetic mean 72±1; and in second case: **C**MP + **A**MP = 72±0 and **U**MP + **G**MP = 72±0, or, both in relation to arithmetic mean 72±0.

In Survey 3 are presented some new balances from which it follows that the distribution into 8 x 33; 9 x 33 (±1) and 10 x 33 of atoms appears to be the same as in Py/Pu distinction within WCT from the aspect of "first-third-letter codon position rule" (Siemion & Siemion, 1994)[3] (see Tables 5 and 6 in this paper in relation to Tables 3a and 3b in Rakočević, 2004).

[3] Siemion and Siemion, 1994, p. 139: "It is shown that in the pairs of amino acids coded by the codon possessing identical bases in the first and second positions, the amino acid with R in the third position are of higher structural importance than the amino acids coded with Y" (*Remark* 1:

3 DISCUSSION

The key for understanding the arrangement in this new Genetic Code Table, as it is given in Table 2, is the alphabet-permutation hierarchy, given in Table 1.The first 4+4 permutations are given in relation to the type of the base (to be Py or Pu), next 4+4 in relation to the number of hydrogen bonds (realized between codon – anticodon), and the last 4+4 in relation to the functional group (to be *oxo*, i.e. hydroxyl, or *amino* functional group in terminal molecule position)[4]. By this, one must notice that permutation arrangement is given through the logic of existence *the pairs of the pairs*, and also notice that it makes sense to determine the permutation ordinal number only within each three octets, in two manner: going from up to down, on the left, and in a vice versa direction, on the right. (*Remark* 2: It is necessary to distinguish the principle "*to be the pair of the pairs*" from the principle "*to exchange a pair for a pair*"; cf. Remark 3.)

That follows from the mirror symmetry, which can be realized through two logics: 1. Single letter to single letter, and 2. Letter doublet to letter doublet. Thus, as the first permutation in third octet, we have: in first logic UGCA / ACGU and in second one: UGCA / CAUG. That is the reason why the permutation CAUG is the first, and not – the fourth[5].

In the other words, the key problem in this analysis of the relationships in CT is the question which is related to the status of permutations within Table 1. Bearing in mind this aspect of the said question, and in order for better understanding, we mention here some additional facts. So, the second permutation in first octet (UCGA) is also used as a codon arrangement in WCT, when occurs a strict distinction between four-codon and non-four-codon AAs (Yang, 2003).

Except with Pu, a purine can be denoted with the letter R, and a pyrimidine, Py, with the letter Y as it is here.)

[4] Rakočević, 1988, p. 112: „A system of four „small" molecules is made up of two purine and two pyrimidine bases. ... The four [molecules] are mutually distinguishable by three main characteristics: the type of base (purine, Pu, or Pyrimidine, Py); the type of functional group in the terminal position (position 6 in purine, position 4 in pyrimidine) – either oxo or amino; and the number of hydrogen bonds linking them in the system codon – anticodon."

[5] This cyclization itself is also valid for all "8 x 3" quadruplets in three permutation octets. The quadruplet UUXX is realized three times going in up/down, and three times in a vice versa direction; so on the left, and so on the right, regarding the whole permutation system (two and two outer columns). The same is valid for the quadruplets XYXY.

On the other hand, two permutations from third octet appear also as very characteristic, and in a specific connection through Py-pair / Pu-pair inversion, both on the right and in ordering read from down to up; the first and fourth permutations, CA-UG and AC-GU, respectively. Namely, the CA-UG permutation is used in a p-adic mathematics application to the genetic code structure (Dragovich & Dragovich, 2006), whereas from the AC-GU it follows an ordinal number of AAs within WCT from 0 to 19 (Damjanović, 1998; Damjanović & Rakočević, 2005, 2006).

From this follows that two permutations, CA-UG and AC-GU appear as a new pair (through an inversion), the outer pair, complementary with the inner pair CA-GU and AC-UG. The same is valid for the first quadruplet (and for all other), and that is the reason why the fourth permutation (CUGA) is so power for a new splitting into the CT, such a splitting from that follows Table 3 as "a freedom degree" for CT, presented in Table 2. (*Remark* 3: The existence of one outer and one inner pair within eight quadruplets, in Table 1, one can understand as the principle – *to exchange a pair for a pair*; cf. Remark 2.)

In accordance to this discussion there are some possible chemical reasons, from which we give here only a choice. Namely, we will give only an explanation for the ordinal number 1, going from the simplest to the most complex of amino acid molecules; complex from different aspects (three toned spaces in Table 1).

*****

First octet, permutation UCAG: From the aspect of the type of the base, pyrimidine is simpler than purine (*one* ring versus *two* rings). On the other hand, U and A with *two* hydrogen bonds both, are simpler than C and G, which possess *three* hydrogen bonds each.

Second octet, permutation UACG: From the aspect of the number of hydrogen bonds, U and A are simpler than C and G. On the other hand U is simpler than A as a pyrimidine molecule.

Third octet, permutation CAUG: From two purines A is simpler than G, because it possesses an amino functional group only; in the other words, G is more complex through its two functional groups – amino and oxo[6]. (In addition, one must notice that nitrogen is simpler than oxygen.) In a pairing process, the base U, possessing two more complex functional groups (the oxo group, two times!) must go with G, and simpler C, with only one oxo group, must be with A. Certainly, in both cases of pairing, the pyrimidine must be before purine as the simpler molecule.

[6] The base A, as we know, is simpler than G through the number of hydrogen bonds too.

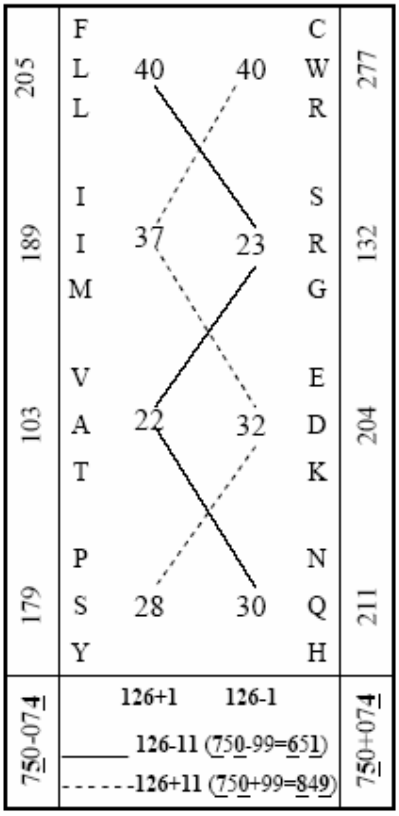


Table 7. The distribution of the canonical AAs in standard Genetic Code Table, in correspondence with the Py-Pu distinction, atom and nucleon number (unit change) balance (including the balance of the isotope number as it is shown in Remark 5)[7]. Notice that AAs are given in ordering as they exist within standard GCT, going from column to column. Notice that amino acid arrangement exists as 12±1 AAs between two "stop" codon positions. First 12 AAs before UAA (ochre) and UAG (amber); and next 12-1 before UGA (opal). Reading from C over F there are 12+1 AAs. Notice that quantum "074" is a head-nucleon-number.

In connection with the first, there is a second question which is related to the hierarchy of the three permutation octets. After our hypothesis, the hierarchy must be – as it is given. That is so, because such a hierarchy is found in Gray

[7] The reading logic, e.g. for Proline: CCY and CCR – one and the same amino acid, P; e.g. for Isoleucine: AUY one amino acid – I; AUR – two different AAs: I and M; in total: I, I, M. (The letters Y & R as said in Remark 1.)

code model of genetic code (Swanson, 1984) as well as on the genetic code binary tree (Rakočević, 1998)[8].

*

Remark 4: Atom number and nucleon number within amino acid side chains: F 14 (91), L 13 (57), I 13 (57), M 11 (75), V 10 (43), A 04 (15), T 08 (45), P 08 (41), S 05 (15), Y 15 (107), H 11 (81), Q 11 (72), N 08 (58), K 15 (72), D 07 (59), E 10 (73), G 01 (01), R 17 (100), W 18 (130), C 05 (47).

Remark 5: The isotope (nuclide) number balance in Table 7: the difference between left and right half equals **±0**, as follows: F28 + L26 + L26 + I26 + I26 + M24 + V20 + A08 + T17 + P16 + S11 + Y31 = 259; 259 = H22 + Q23 + N17 + K30 + D16 + E22 + G02 + R34 + S11 + R34 + W36 + C12.

Remark 6: The isotope (nuclide) number balance in Table 8, in a zigzag (periodic) ordering: the difference between two halves equals **±1**, as follows: (F28 + L26 + S11) + (A08 + V20 + R34) + (L26 + P16 + H22) + (T17 + M24 + I26) = 259–1; 259 + 1 = (G02 + E22 + D16) +(Y31 + C12 + W36) + (S11 + K30 + N17) + (Q23 + R34 + I26).

Remark 7: The number 259 is the first permutation in column 7 of Shcherbak's modular system (in module 9) of multiples of number 037 (Table 1 in Shcherbak, 1994, p. 476); moreover, the nucleon number within four-codon and non-four-codon AAs is determined only with 7th column, as multiples of 259, and with 1st row, as multiples of 111.

Remark 8: An example of calculation of particles number for serine side chain: atom number: CH_2 + OH = 5 atoms; isotope number: C-12 and C-13 equals 2 isotopes; H-1 and H-2 (three times) equals 6 isotope; O-16, O-17 and O-18 equals 3 isotopes; in total: 2 + 6 + 3 = 11 isotopes; nucleon number within first isotope, i.e. nuclide: 12 + (3 x 1) + 16 = 31.

---

[8] On the Gray code model as well as on the genetic code binary tree, the first codon classification is given after type of the base (Py/Pu distinction) and second one after the number of hydrogen bonds. The logic existing within third characteristics (functional group *oxo* or *amino* in terminal Py/Pu base position) parallel with these two classifications is not found. From this it follows: neither to be the first, nor the second one, then must be – the third (the third octet of permutations in Table 1).

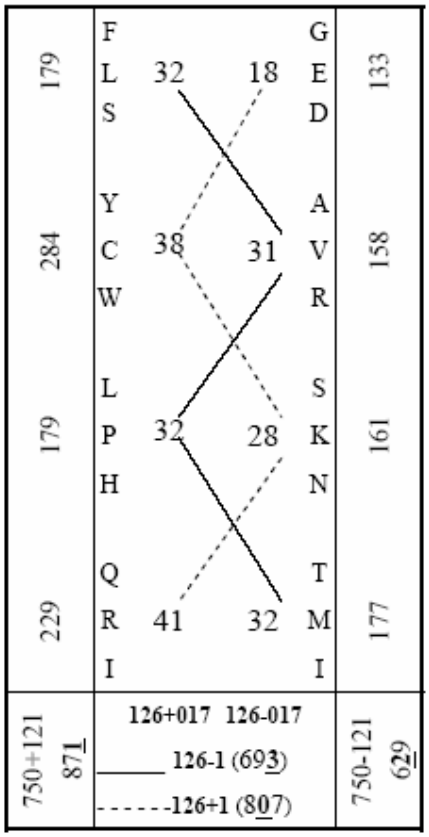


Table 8. The distribution of the canonical AAs in standard Genetic Code Table, in correspondence with the Py-Pu distinction, atom and nucleon number (unit change) balance, including the balance of the isotope number as it is shown in Remark 6. Notice that AAs are given in ordering as they exist within standard GCT, going from row to row. Notice $121=11^2$.

In some manner, independently from discussed questions, there is a fundamental question, which is related to the "choice" of the number of letters within a code-alphabet and, *per se*, the number of permutations. It is self-evident that from the aspect of the principle "to be pair of pair", the 3-letter, 5-letter, 6-letter and 7-letter alphabets are not possible. The next, after the 4-letter alphabet, which seems to be possible, is just the 8-letter alphabet. But that is an illusion. Why? In 4-letter alphabet each permutation really contains a pair of pair of letters: (a-b, c-d) whereas in 8-letter alphabet – more than a pair of pair. From this aspect, and from the aspect of the above discussed two principles "the pair of

the pairs" and "a pair for a pair" (Remarks 2 & 3), it makes sense to conclude that the system of 24 permutations (12 pairs), given in a manner as in Table 1, is one and only possible solution. Bearing this state in mind, it is clear why the amino acid corresponding Py-Pu codon splitting, in standard Genetic Code Table, is also given in 24 amino acid meanings (Tables 7 and 8), i.e. as 12 pairs in both cases[9].

CONCLUSION

In conclusion we express a hope that the revelation of these regularities can help in a better understanding of codon - amino acid assignment over two classes of aaRS (De Duve, 1988; Hipps, 1995; Schimmel, 1995). On the other hand, such strict connections between two classes of enzymes, two classes of AAs and belonging codons, provide evidence to support the hypothesis, given in a previous paper (Rakočević, 2004), that genetic code was complete from the very beginning as the condition for origin and evolution of the life.

REFERENCES

Crick, C. H. F., 1966. The genetic code yesterday, today and tomorrow. Cold Spring Harbor Symposia on Quantitative Biology., 31, 3-9.

Crick, C. H. F., 1968. The Origin of the Genetic Code. J. Mol. Biol. 38, 367-379.

Damjanović, Z., 1998. Logic core of genetic code. Glasnik Sect. Nat. Sci. Montenegrin Acad. Sci. Art (CANU), 12, 5-8.

Damjanović, M. Z., Rakočević, M. M., 2005. Genetic code: an alternative model of translation. Ann. N. Y. Acad. Sci., 1048, 517-523.

Damjanović, M. Z., Rakočević, M. M., 2006. Genetic Code: a new understanding of codon - amino acid assignment, q-bio.OT/0611033.

De Duve, C., 1988. The Second Genetic Code. Nature, 333, 117-118.

[9] Rakočević & Jokić, 1996, p. 346: "Notice that out of all doublet-triplets systems, this is the only and one with two possible distinctions for doublets (to be six and six, and then, to be three and three doublets) and three possible distinctions for triplets (to be four and four, then two and two, and, finally, to be one and one triplet)".

Dragovich, B., Dragovich, A., 2006. A p-Adic model of DNA sequence and genetic code, q-bio.GN/0607018.

Eriani, G. et at., 1995. The Class II Aminoacyl-tRNA Synthetases and Their Active-Site-Evolutionary Conservation of an ATP Binding-Site. J. Molec. Evol., 40, 499-508.

Hipps, D. et al., 1995. Operational RNA Code for Amino-Acids-Species-Specific Aminoacylation of Minihelices Switched by a Single Nucleotide. Proc. Natl. Acad. USA., 92, 5550-5552

Rakočević M.M., 1988. Three-dimensional model of the genetic code. Acta biologiae et madicinae experimentalis, 13, 109-116.

Rakočević M.M., 1997a. Two classes of the aminoacyl-tRNA synthetases in correspondence with the codon path cube. Bull. Math. Biol., 59, 645-648.

Rakočević M.M., 1997b. Genetic code as a unique system. SKC, Niš, p. 118 (www.sponce.net).

Rakočević, M. M., 1998. The genetic code as a Golden mean determined system. Biosystems 46, 283-291.

Rakočević, M. M., 2004. A harmonic structure of the genetic code. J. Theor. Biol., 229, 463 – 465.

Rakočević, M. M., Jokić, A., 1996. Four stereochemical types of protein amino acids: synchronic determination with chemical characteristics, atom and nucleon number. J. Theor. Biol. 183, 345 – 349.

Schimmel, P., 1995. An Operational RNA Code for Amino Acids and Variations in Critical Nucleotide-Sequences in Evolution. J. Mol. Evol., 40, 531-536

Shcherbak, V. I., 1994. Sixty-four triplets and 20 canonical amino acids of the genetic code: the arithmetical regularities. Part II. J Theor. Biol., 166, 475-477.

Shcherbak, V. I., 2003. Arithmetic inside the universal genetic code. Biosystems 70, 187-209.

Siemion, I.Z., Siemion, P.Z., 1994. The informational context of the third base in amino acid codons. Biosystems., 33, 139-148.

Swanson, R., 1984. A unifying concept for the amino acid code. Bull. Math. Biol., 46, 187-207.

Watson, J. D., Crick, F. H. C., 1953a. Molecular structure of nucleic acids. A structure of deoxyribose nucleic acid. Nature, 171, 737-738.

Watson, J. D., Crick, F. H. C., 1953b. Genetical implications of the structure of deoxyribonucleic acid. Nature, 171, 964-967.

Wetzel, R., 1995. Evolution of the Aminoacyl-tRNA Synthetases and the Origin of the Genetic Code. J. Mol. Evol., 40, 545-550.

Yang, C. M., 2003. The bi-pyramidal nature, the Lucas series in the genetic code and their relation to aminoacyl-tRNA synthetases, q-bio.BM/0309008.

GENETIC CODE: FOUR-CODON AND NON-FOUR-CODON DEGENERACY

Miloje M. Rakočević

Department of Chemistry, Faculty of Science, University of Niš, Ćirila i Metodija 2, Serbia

(milemirkov@nadlanu.com; www.sponce.net)

Abstract.

In this work it is shown that 20 canonical amino acids (AAs) within genetic code appear to be a whole system with strict distinction in Genetic Code Table (GCT) into some different quantums: 20, 23, 61 amino acid molecules. These molecules distinction is followed by specific balanced atom number and/or nucleon number distinctions within those molecules.

One of the most important question that has been made since creation of Genetic Code Table (GCT) (Crick, 1966, 1968) is the question of degeneration of the genetic code itself. Even then Rumer has noticed (Rumer, 1966) that the key aspect of degeneration could be found in the codon clasification (and corresponding amino acids, AAs) according to the degeneration of type IV, when all four codons have the same aminoacidity meaning[1], and to the degeneration of type non-IV or type I-II-III. Bearing in mind those facts, Scherbak has shown (Shcherbak, 1994) that classification of AAs into four-codon and non-four-codon AAs is followed by a specific balance of nucleons number, expressed by multiplying the number 037, which is named as Prime quantum (PQ).

It has been showen that this trend exists when calculation is performed with 8 AAs from eight quarter "cells" with codons of the same aminoacidity meaning and 15 AAs from eight also quarter "cells" but with codons of different aminoacidity meaning; all this, in total, is 23 amino acid molecules (Table 1 in relation to Table 2). [**Remark 1**: It is important here to notice the existance of two principes: the principle of continuity and the principle of minimum change,

[1] For 4-letter alphabet the key distinction must be the degeneration of the type IV and non-IV; for the 5-letter alphabet - the type V and non-V, etc.

which follow from the existance of specific proportions (30:30 = 1: 1; 30 : 60 = **1** : **2**; 9:16:25 = **3** : **4** : **5**).]

	G		A		S		P		V	T		L		R		
	01		15		31		41		43	45		57		100		09 x 037
	74		74		74		74		74	74		74		74		**16 x 037**
	75		89		105		115		117	119		131		174		25 x 037
C	I	S	L	N	D	Q	K	E	H	F	R	Y	M	W		
47	57	31	57	58	59	72	72	73	81	91	100	107	75	130		30 x 037
74	74	74	74	74	74	74	74	74	74	74	74	74	74	74		**30 x 037**
121	131	105	131	132	133	146	146	147	155	165	174	181	149	204		60 x 037

Table 1 The four-codon amino acids (above) and non-four-codon amino acids (down) as in Figure 1 in Shcherbak, 1994: nucleon number balances.

However, this paper deals with the fact that the determination law apllying the multiplying number 037 can be used even for the maximum possible aminoacidic distinction, when calculation is performed using all 61 amino acid molecules. Also, we show that Scherbak's system of multiplying number 037 (Table 1 in Shcherbak, 1994) is within one wider system of multiplying numbers 666 and 777 (Table 3), where it could be found that PQ 037 is generated from fraction 111/3. Bearing in mind those facts it is obvious that system with multiplying numbers 66 & 77 (with PQ 11/3) and 6 & 7 (with PQ 1/3) go before the system with multiplying numbers 666 and 777. Thus, those systems show that they are deteminants of classification into four-codon and non-four-codon AAs, if it is taken into account not only the number of nucleons (result 2 x 666 in Table 2, in relation to Table 4)[2] , but also the number of atoms (Tables 5 & 6).

[2] Table 4 (in a connection with Table 3) shows that, if in the coding system there are 8 x 8 codons, then must be 8 + 8 four-codon families with 8+15 = 23 amino acid molecules (neither 24 nor 25) [**8** + (8 – 1) = **15**].

	4G	4A	4S	4P	4V	4T	4L	4R	
	04	60	124	164	172	180	228	400	2 x 666 **36 X 037**
	296	296	296	296	296	296	296	296	64 x 037
	300	356	420	460	468	476	524	696	100 x 037
2C	3I	2S	2L	2N	2D	2Q	2K	2E	2H	2F	2R	2Y	1M	1W	
94	171	62	114	116	118	144	144	146	162	182	200	214	75	130	**56 x 037**
148	222	148	148	148	148	148	148	148	148	148	148	148	74	74	58 x 037
242	393	210	262	264	266	292	292	294	310	330	148	362	149	204	114 x 037

Table 2. The four-codon amino acids (above) and non-four-codon amino acids (down) as in Figure 1 in Shcherbak, 1994, all amino acids multiplied with coding codon number: nucleon number balances.

It can be seen, by comparing the Tables, still other different interesting balances, proportions and symetries. Thus, it can be found in Table 1 that determination using PQ 037 is realized in accordance with „the symmetry in the simplest case“ (Marcus, 1989). The symetry is for amino acid „heads“ deducing on proportion 1:2 (8 AAs molecules : 16 PQ and 15 AAs molecules : 30 PQ)[3], where the total quantum number is 46 (16+30). On the other hand, in Table 2 it can be found that in „bodies“, within AAs side chains there are 46±10 quantums (36 x 037 in singlemeaning and 56 x 037 in multimeaning cells).

The sense of determination by quantums „64“ and „58“, when dealing with nucleon number in the "heads" of AAs, is evident bearing in mind the following equation: **64** + *x* = 2 x **61**, where *x* = 58, and „64“ & “61” stand for known quantums in distinction of codon number, because 64 – 61 = 3, where quantum "3" regards to the number of stop codons. The total number of all quantums in Table 2 (100 + 114 = 214) corresponds on one side with the number of atoms in side chains of 20 AAs (2**1**4:2**0**4), while on the other side with the sum of the first four perfect numbers (2**1**4 : 2**3**4)(234 – 214 = 2**2**4). The change in the second position of number for 0, 1, 2, 3 shows that principes of continuity and minimum

[3] Result 16 PQ, i.e. 16 x 037 = 592 represents a half of the number 1184 which number is the first memeber of the second pair of friendly numbers. [Friendly number pairs: (220, 284), (1184, 1210), (17296, 18416) etc.]

1	**037**	666	777
2	074	**1332**	1554
3	111	1998	**2331**
4	148	2664	3108
. .	. . .	. . .	. . .
13	481	**8658**	**10101**
14	518	9324	10878
. .	. . .	. . .	. . .
16	592	10656	12432
17	629	11322	13209
. .	. . .	. . .	. . .
27	**999**	17982	20979

Table 3. The four-codon amino acids (above) and non-four-codon amino acids (down) as in Figure 1 in Shcherbak, 1994, all amino acids multiplied with coding codon number: nucleon number balances.

change stand here, as well[4]. [**Remark 2.** Notice the correspondence of quantums 214, 224 and 234 with the same quantums of atom number within 4 + 4 + 4 + 4 dinucleotides in Table 1 in Rumer, 1966; within 4 + 4 upper dinucleotides there are 116+108 = 224, and within 4 + 4 lower: 106 + 118 = 224 atoms. On the other hand, we have: 118 + 116 = 234 and 108 + 106 = 214 atoms. (As an interesting result is that, that the balance of atom number within dinucleotides is followed by a balance of atom number within 11 upper amino acid molecules (119) and 12 lower (120), all in Rumer's Table 1).]

[4] The sum of the first four perfect numbers is: 6+28+496+8128 = 8658 = 7770+0888 = 2**3**4 x 037; the 13th case of multiples of 666 in Table 3. (About the determination of genetic code with perfect and friendly numbers see in Rakočević, 1997, p. 60, or in: www.sponce.net).

8 + 7 = 15 8+ 15 = 23	8 + 8 = 16 8+ 16 = 24	8 + 9 = 17 8+ 17 = 25
16 x 37 = 592 30 x 37 = 1110	16 x 37 = 592 32 x 37 = 1184	16 x 37 = 592 34 x 37 = 1258
46 x 37 = 1702 **56** x 37 = 2072 **36** x 37 = **1332**	48 x 37 = 1776 58 x 37 = 2146 38 x 37 = 1406	50 x 37 = 1850 60 x 37 = 2220 40 x 37 = 1480

Table 4. The quantums 8, 15 and 23 (above left) correspond with the patterns of 8 four-codon, 15 non-four-codon (and 23 in total) amino acids. The double values only in first column correspond to quantums in last column of Table 2 (36 as 46-10 and 56 as 46+10).

It is evident through the comparison of tabulated values in Tables 5 and 1, that in Table 5 determination regards the real 20 canonic AAs, while in Table 1, it regards with 23 AAs. To put it in other way, 6-codon AAs (L, S, R) play twice only in the case of nucleon number determination, but do not play in atom number determination. By comparing Table 5 and Table 6 it is shown that in the first case determination regards on the real 20 canonic AAs, while in the second case it regards on 61 AAs; in the first case with the determination through the multiplying number 6, while in the second case through the both multiplying numbers 6 and 66.

G	A	S	P	V	T	L	R	
01	04	05	08	10	08	13	17	11 x 6
09	09	09	09	09	09	09	09	**12 x 6**
10	13	14	17	19	17	22	26	23 x 6

C	I	N	D	Q	K	E	H	F	Y	M	W	
05	13	08	07	11	15	10	11	14	15	11	18	23 X 6
09	09	09	09	09	09	09	09	09	09	09	09	**18 x 6**
14	22	17	16	20	24	19	20	23	24	20	27	41 x 6

Table 5. The four-codon amino acids plus six-codon amino acids (above) and non-four-codon amino acids (down): atom number balances.

4G	4A	4S	4P	4V	4T	4L	4R	
04	16	20	32	40	32	52	68	**4 x 66** 44 x 6
40	52	56	68	76	68	88	104	92 x 6 184 x 3

2C	3I	2S	2L	2N	2D	2Q	2K	2E	2H	2F	2R	2Y	1M	1W	
10	39	10	26	16	14	22	30	20	22	28	34	30	11	18	**5 x 66** 55 x 6
28	66	28	44	34	32	40	48	38	40	46	52	48	20	27	197 x 3

Table 6. The four-codon amino acids plus six-codon amino acids (above) and non-four-codon amino acids (down), all amino acids multiplied with coding codon number: atom number balances.

In Conclusion we can say that all those tabulated results, reveals that all mentioned arithmetic regularities are of systematic character and stands for the whole code, as it is in the form of "standard genetic code". By this, those regularities comprise our hypothesis of complete genetic code (Rakočević,

2004), by which the standard genetic code would be complete from the very beginning, as condition for life existence, while all irregularities from standard code are only the expression of the freedom rate in its minimum. But, there is still enigma, which Shcherbak pointed out in his paper, cited here, that is "the physical nature" of these arithmetical regularities "is so far not clear". (Shcherbak, 1993, p. 401). Even though, this should not be the threat to reveal all these classical as well as new p-adic arithmetical regularities (Dragovich and Dragovich, 2006), because it is the only way to investigate their not only physical but also biological sense.

REFERENCES

Dragovich, B., Dragovich, A., 2006. A p-Adic Model of DNA Sequence and Genetic Code. q-Bio. GN/0607018.

Marcus, S., 1989. Symmetry in the simplest case: the real line. Comput. Math. Appl. 17, 103-115.

Rakočević, M. M., 1997. Genetic code as a unique system. SKC, Niš pp. 60-64.

Rakočević, M. M., 2004. A harmonic structure of the genetic code. J. Theor. Biol. 229, 463 – 465.

Shcherbak, V. I., 1993. Twenty canonical amino acids of the genetic code: the arithmetical regularities. Part I. J Theor. Biol. 162, 399-401.

Shcherbak, V. I., 1994. Sixty-four triplets and 20 canonical amino acids of the genetic code: the arithmetical regularities. Part II. J Theor. Biol. 166, 475-477.

Genetic Code Table: A note on the three splittings into amino acid classes

Miloje M. Rakočević

Faculty of Science, University of Niš (now retired, on the Address: Milutina Milankovica 118, 11070 Belgrade, Serbia (e-mail: m.m.r@eunet.yu; or: milemirkov@nadlanu.com; www.sponce.net or www.rakocevcode.rs)

Abstract

This note represents the further progress in understanding the determination of the genetic code by Golden mean (Rakocevic, 1998). Three classes of amino acids that follow from this determination (the 7 "golden" amino acids, 7 of their complements, and 6 non-complements) are observed now together with two further possible splittings into 4 x 5 and 5 x 4 amino acids.

In one of the previous works (Rakočević, 1998) was shown that the Genetic Code Table (GCT) of the standard genetic code (Table 1) can be developed in a six-bit binary-code tree (Figure 1). Thus, the order of the eight codon octets is: YUN, YCN, RUN, RCN, YAN, YGN, RAN, RGN. [One-letter abbreviations: Y from pYrimidine; R from puRine; and N from aNy (of bases).] (see Comment 1).

With respect to the Golden mean (Table 2), in such a binary tree one can establishe the positions of amino acids (AAs). As result of the Golden mean determination there are three classes of AAs: the 7 "golden" AAs, the 7 their complements (complements to the full pairs) and the 6 non-complements, all as in Table 3. Then the number of atoms (60, 66 and 78) within the three classes differs for 1 x 6; 2 x 6 and 3 x 6 of atoms, respectively (see Comment 2).

In this note, however, we show that the determination with the number 6 (the first perfect number!)[1] has new aspects related to positions (odd/even) of AAs within the AAs classes (Tables 4-9).

[1] The factors of perfect numbers are in correspondence to the binary sequence 2^n (n = 0, 1, 2, 3, ...) as follows. For the first (6): 1, 2, (4 -1 = 3), (8 – 2 = 6); for the second (28): 1, 2, 4, (8 -1 = 7), (16 – 2 = 14), (32 – 4 = 28); for the third (496): 1, 2, 4, 8, 16, (32 -1 = 31), 64 – 2 = 62), (128 – 4 = 124), (256 – 8 = 248) , (512 – 16 = 496); for the forth (8128): 1, 2, 4, 8, 16, 32, 64, (128 -1 = 127), (256 –

If the first two classes of AAs (in the upper part of Table 3) are seen as one class, then it can be reclassified into two new classes, depending on the question – whether AAs are on the even, or on the odd positions; all these in the order dictated by GCT, or by the hierarchy of number of atoms in the amino acid molecules (their side chains), either within the set of "golden" amino acids or their complements.

The separation of the third class of AAs (in the lower part of Table 3) has its own chemical as well as arithmetical justification. Chemically, the four aliphatic AAs (D, E & K, R)[2] are specific in relation to all other aliphatic AAs, because two of them contain the same functional group (carboxylic) twice – in the "head" and in the "body" of an amino acid molecule. The same is valid for K and R, but in relation to the basic amine functional group, NH_2. (Notice that within N and Q there is an amide functional group.)

For two aromatic AAs (H and W) it can not be said that they contain the same functional group just twice, but it can be said that they contain a double function (aromatic and heterocyclic), while the remaining two aromatic AAs (F and Y) contain only a single function (aromatic).

This sample of chemical hierarchy has a match in arithmetical hierarchy. So, while the two first classe determination by number 6 appears only once (10 x 6 and 11 x 6, respectively), so far for the third class the determination is seen by twice: the original determination with 13 x 6 is "split" into two determinations: 5 x 6 and 8 x 6 (see Comment 3). Really, within four aliphatic AAs (D, E & K, R) there are (8 x 6)+1, whereas within two aromatic, H & W, one has (5 x 6) -1 of atoms.

The Table 10 contains a splitting into three amino acid classes (according to Table 3), while Table 11 contains a new splitting into five classes: one class of aromatic and four classes of aliphatic AAs. If, however, in Table 11 we join aromatic AAs to aliphatic ones, one gets four classes: the first (F) to the first class, the second (Y) to the second one (going from top to down); and also the first (H) to the first class, the second (W) to the second one (going from bottom to up). The

---

2 = 254), (512 – 4 = 508), (1024 – 8 = 1016), (2048 -16 = 2032), 4096 – 32 = 4064), (8192 – 64 = 8128): etc. Here one must notice that there is a connection with the friendly numbers through third friendly number, i.e. through the first member of second friendly pair (1184), and a connection with Shcherbak's system of multiples of number 037, valid for GC (Shcherbak, 1994, Figure 1 and Table 1) at the same time: 1, 2, 4, 8, 16, 32 / 37, 74, 148, 296, **592**, 1184 (see position of the number **592**, as the 16th case in Table A.1, in Appendix). The first three friendly pairs are: (220, 284), (1184, 1210), (17296, 18416).

[2] Starting from this chemically strict defined class of AAs it makes sense all the other "quadruplets" also seen as possible chemical classes, what means a splitting into five amino acid classes (5 x 4) as it will be showed in further text.

number of atoms within obtained four classes is also determined by the multiples of number 6, with the deviations for ±0 and/or ±1 (Table 12). Particularly, it is interesting the possibility of a new splitting also into the three classes as it is shown in Table 13. In relation to Table 3 we have a change for ± 1 and/or ± 0 molecules as well as for the ± 1 and/or ± 0 of atoms within the amino acid classes.

From the original "Golden Table" (Table 3), taking the sequence of "golden" AAs in molecule mass ordering, i.e. in hierarchy over atom number (Table 9), follows still one new amino acid splitting – the splitting into 5 x 4 AAs (Table 14).

In the middle position of this Cyclic Invariant Periodic System (CIPS)[3], presented in Table 14, there are chalcogene AAs (S, T & C, M)[4], and then in a cyclic arrangement follow the "contact" AAs (G, P & V, I)[5], two double acidic AAs with two their amide derivatives (D, E & N, Q)[6]; the two original aliphatic AAs with two amine derivatives (A, L & K, R)[7]; and, finally, four aromatic AAs (F,Y & H, W). Notice that within CIPS each amino acid position is strictly determined and none can be changed.

As it is immediately obvious from CIPS, the nature of the genetic code again points out the validity of Aristotle sentence[8], and, on the other hand, necessarily

---

[3] Cyclicity and periodicity through the positions of two and two amino acids – up/down – in relation to middle chalcogene AAs.

[4] The elements from the sixth group of Periodic system are the chalcogenes, and that is valid for oxygen and sulfur. Hence follows that these AAs are just such – the chalcogene AAs! Of course, for still two amino acids (D and E), it can be concluded that from their nature, they are also chalcogene AAs. However, when we see that here is in the question a system *per se*, a system of "5 x 4" elements, then we can say that four amino acids (S, T & C, M) are the source chalcogene AAs, and the other two (D, E) are derived; they must go into a separate class with its derivatives, two amides: N, and Q. [Cf. legend to Table B. 9, where it is shown that (S, T & C, M) and (D, E) are together.]

[5] After Popov (Popov 1989; Rakočević & Jokić, 1996) the four "contact" AAs (G, P & V, I) are of the non-alanine stereochemical types: G from glycine type, P from proline, and V, I from valine stereochemical type. The term "contact" is our , and the explanation is in Comment 4 (Rakočević, 2006).

[6] The amino acids D & E are double acidic because carboxylic gropu exists in both „head" and „body" of molecule at the same time.

[7] The source, i.e. original aliphatic compounds are nonpolar; their amine derivatives are less polar then hydroxy derivatives (NH_2 less polar than OH, and N less than O).

[8] "The existence of such a harmonic structure with unity of a determination with physical–chemical characteristics and atom and nucleon number ... appealed to Aristotle and to his idea of unity of form and essence" (Rakočević, 2004, p. 233). From the aspect of this paper, the form make the

leads to the conclusion that all three main hypotheses about the origin of the genetic code are *mutatis mutandis* valid; two given by Crick (the genetic code as a result of stereochemical interactions[9], or as a result of pure chance[10]) (Crick, 1968) and one by Wong, the theory of co-evolution[11]. (Wong, 1975, p. 1909: "The theory is proposed that the structure of the genetic code was determined by the sequence of evolutionary emergence of new amino acids within the primordial biochemical system".)[12]

As we can see from Tables which follow from CIPS (see Appendix B), the key principles valid for this system are: 1. Principle of minimum change, 2. Principle of simple proportion, 3. Principle of multi meaning diversity, 4. Block-aufbau principle, and 5. Principle of self-similarity. All together these are the principles of building of specific harmonic structures[13] or, shorter, the harmonic principles.

Golden mean determined amino acid positions within CIPS, and the essence is realized through two realities: physically, through molecule mass; and chemically – through the expression of the five revealed chemical classes (see Comment 5).

[9] Crick, 1968, p. 369: "This theory states that the code is universal because it is necessarily the way it is for stereochemical reasons".

[10] Crick, 1968, p. 370: "In its extreme form, the theory implies that the allocation of codons to amino acids at this point was entirely a matter of 'chance' ". For the aspect of CIPS existence that means that distribution of codons to amino acids must be in relation to a pure form, such as it is Golden mean; then, in relation to size of amino acid molecule and in relation to physical–chemical properties at the same time; by this, it is matter of a pure chance to "find" such "golden" conditions in a prebiotic area (see Comment 6).

[11] The act of a "co-evolution" itself is not possible without a "co-influence", i.e. without an interactive influence of amino acid and/or nucleotide components of genetic code (Rakočević, 2004, p. 233: "The word can be only about a 'co-evolution' based on 'co-influence' of more factors").

[12] On a slightly different way, the same idea is expressed by Sukhodolec (Sukhodolec, 1981, p. 499: „The basis of the hypothesis is referred to the idea that during the prebiotic evolution amino acids and nitrogenous bases existed in the form of complex crystal structures and that in the construction of these structures, as components of their elements, the chain units - analogues alternately have been used: the amino acids and duplexes from the first two bases of the codons". Our idea is similar to this idea of Sukhodolec (Rakočević, 2004, p. 232: "At a later stage many nucleotide/amino-acid aggregations, similar to aggregations of Miller's type, or to not much different aggregations of Murchison–meteorite's type, had been realized").

[13] The structures like this one, presented in our previous paper (Rakočević, 2004) (see Table D.3).

1st lett.	2nd letter								3rd lett.
	U		C		A		G		
U	**00**. UUU	**F**	**08**. UCU	**S**	**32**. UAU	**Y**	**40**. UGU	**C**	*U*
	01. UUC		**09**. UCC		**33**. UAC		**41**. UGC		*C*
	02. UUA		**10**. UCA		**34**. UAA		**42**. UGA		*A*
	03. UUG	**L**	**11**. UCG		**35**. UAG	CT	**43**. UGG	CT **W**	*G*
C	**04**. CUU	**L**	**12**. CCU	**P**	**36**. CAU	**H**	**44**. CGU	**R**	*U*
	05. CUC		**13**. CCC		**37**. CAC		**45**. CGC		*C*
	06. CUA		**14**. CCA		**38**. CAA		**46**. CGA		*A*
	07. CUG		**15**. CCG		**39**. CAG	**Q**	**47**. CGG		*G*
A	**16**. AUU	**I**	**24**. ACU	**T**	**48**. AAU	**N**	**56**. AGU	**S**	*U*
	17. AUC		**25**. ACC		**49**. AAC		**57**. AGC		*C*
	18. AUA		**26**. ACA		**50**. AAA		**58**. AGA		*A*
	19. AUG	**M**	**27**. ACG		**51**. AAG	**K**	**59**. AGG	**R**	*G*
G	**20**. GUU	**V**	**28**. GCU	**A**	**52**. GAU	**D**	**60**. GGU	**G**	*U*
	21. GUC		**29**. GCC		**53**. GAC		**61**. GGC		*C*
	22. GUA		**30**. GCA		**54**. GAA	**E**	**62**. GGA		*A*
	23. GUG		**31**. GCG		**55**. GAG		**63**. GGG		*G*

Table 1. The Table of the standard genetic code. Ordinal number of codons after the order-key: YYN, RYN, YRN, RRN, in correspondence with the hierarchy on the binary-code tree in Figure 1.

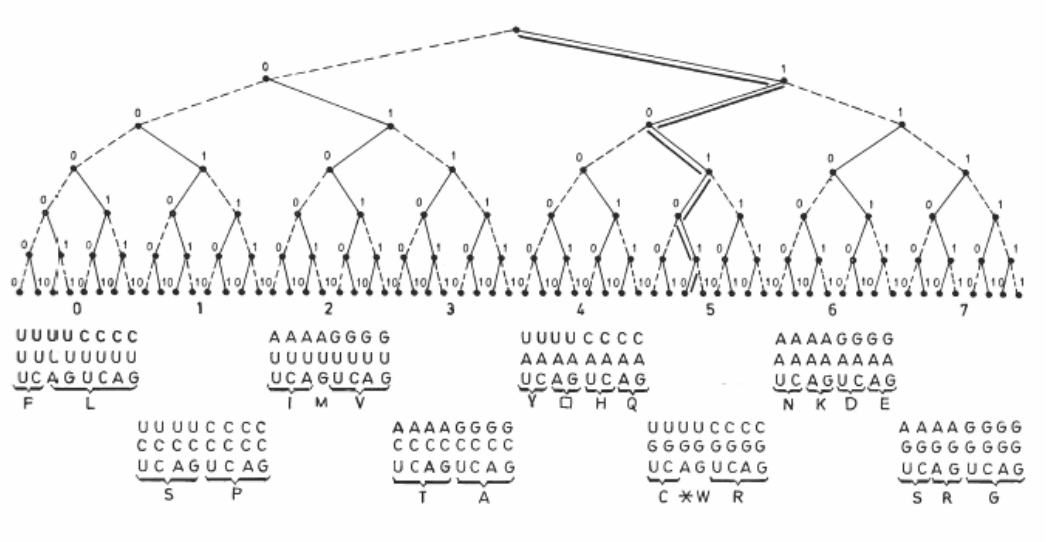


Figure. 1. Genetic code as a binary-code tree. The full lines: the routes of the greater changes, from 0 to 1 and vice versa; the dotted lines: the routes of the less changes, from 0 to 0, as well as from 1 to 1 going from a higher into a lower level. The double full line: the route of the maximum possible changes: from 0 to 1 and vice versa in any step (the route corresponding to the 'Golden mean route' on the Farey tree, Figure 2). Asterisks: 'stop' codon UGA; square: 'stop' codons UAA and UAG. The codon distribution and order after the rules given by R. Swanson (Swanson, 1984, Figure 1): 1 for purine and 0 for pyrimidine; 1 for three and 0 for two hydrogen bonds.

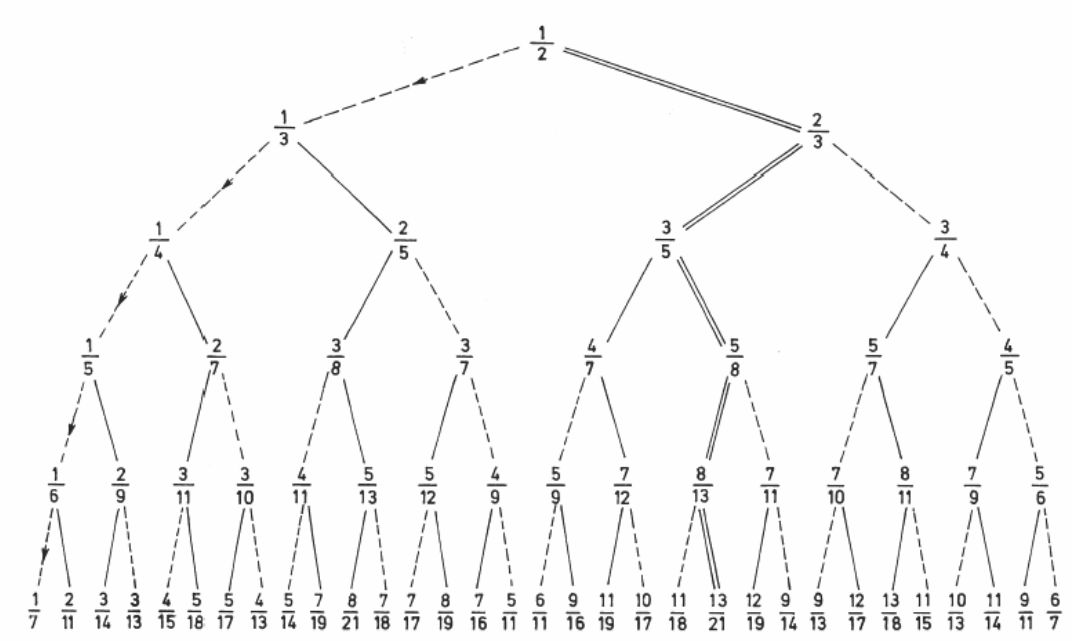


Figure 2. The Farey binary tree. It has a special application in physics of the deterministic chaos, and theory of fractals. At the same time at this tree it is possible to give a proof that the set of rational numbers is countable. The double line represents the "golden route" because it possesses Fibonacci numbers (related to Golden mean) in numerators as well as denominators. [This graph is the same as in original paper (Rakočević, 1998) with a minor error: the line 3/5 – 4/7 must be a dotted one.]

ϕ^0	ϕ^1	ϕ^2	ϕ^3	ϕ^4	$\phi^{5\text{-}7}$	ϕ^8	ϕ^9
G	Q	T	P	S	L	L	F
63	39-38	25-24	15-14	10-09	06-02	02-01	01-00
63	38.94	24.06	14.87	9.19	5.68 – 2.17	1.34	0.83

Table 2. The amino acids in Golden mean power positions within the sequence 0–63 on the binary-code tree in Fig. 1. First row: Golden mean powers within first 'cycle' in

module 9. Second row: amino acids in the positions marked in third row, taken from the binary-code tree in Fig. 1. Fourth row: the values of the Golden mean powers within the interval 0–63. The calculations: 0.618033 x 63 = 38.94; 0.618033 x 0.618033 x 63 = 24.06 etc.

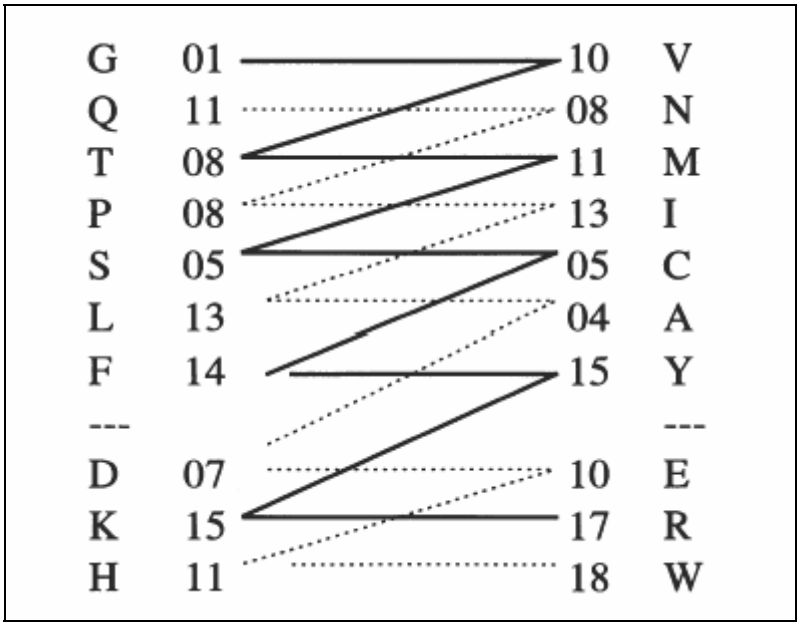


Table 3. Atom number balance directed by Golden mean on the binary-code tree (Scheme 2 in Rakočević, 1998). First seven amino acids on the left are 'golden' amino acids with 60 atoms within side chains; on the right are their (chemically) pairing complements with [60 + (1 x 6)] = 66 atoms; below are three amino acid pairs as non-complements with [66 + (2 x 6)] = 2 x 39 = 78 atoms; Notice that within aliphatic non-complements there are 39+10, whereas within aromatics (H & W) 39-10 of atoms; in the other words: (8 x 6) +1 and (5 x 6) -1, respectively. On the first zigzag (full) line there is 102-1 whereas on the second (dotted) line 102+1 atoms. Arithmetic mean for both: 102±1. Notice also that the arrangement-ordering of "golden" AAs is the same as in Table 2.

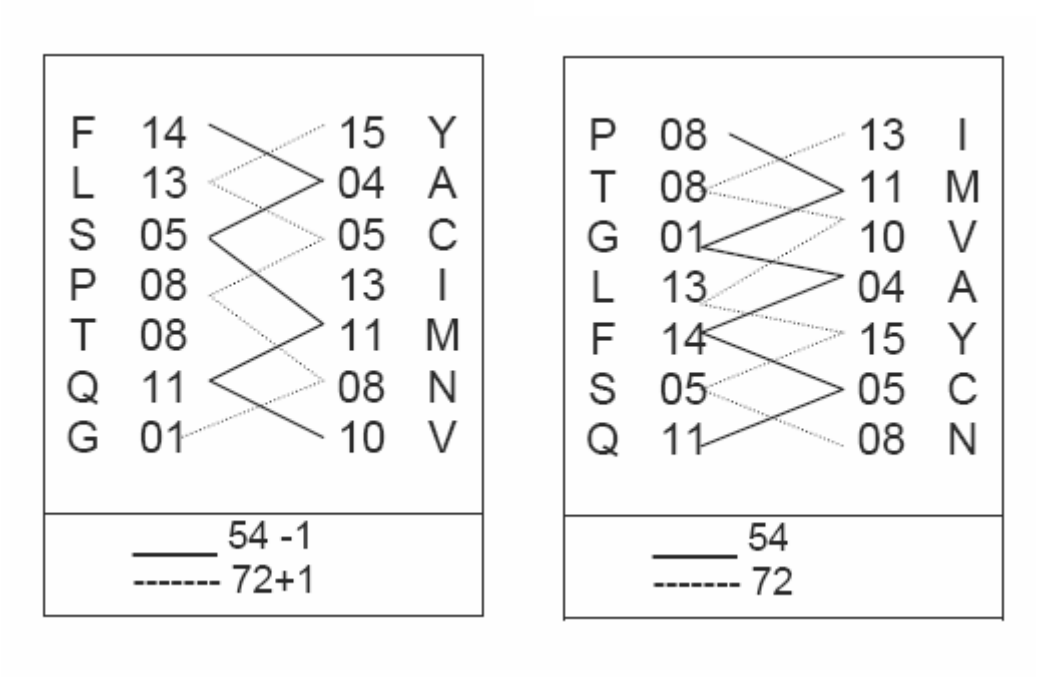


Table 4 (left). The atom number balance within two classes of AAs. Both classes as in Table 3, in a reverse order, from the aspect of the position of "up-down"; there: G, Q, T, P, S, L, F; and here: F, L, S, P, T, Q, G. On the first zigzag (full) line there are [(9 x 6)-1] whereas on the second one (dotted line) [(12 x 6) +1] of atoms.

Table 5 (right). The atom number balance within two classes of AAs: all the same as in Table 4 except the order of AAs is given according to the positions of complements (AAs on the right) on the binary tree, as well as in GCT; on the first zigzag (full) line there are exactly 9 x 6 whereas on the second one (dotted line) 12 x 6 of atoms. (Cf. Table A.5, where aromatic AAs are included with a new type of balance.)

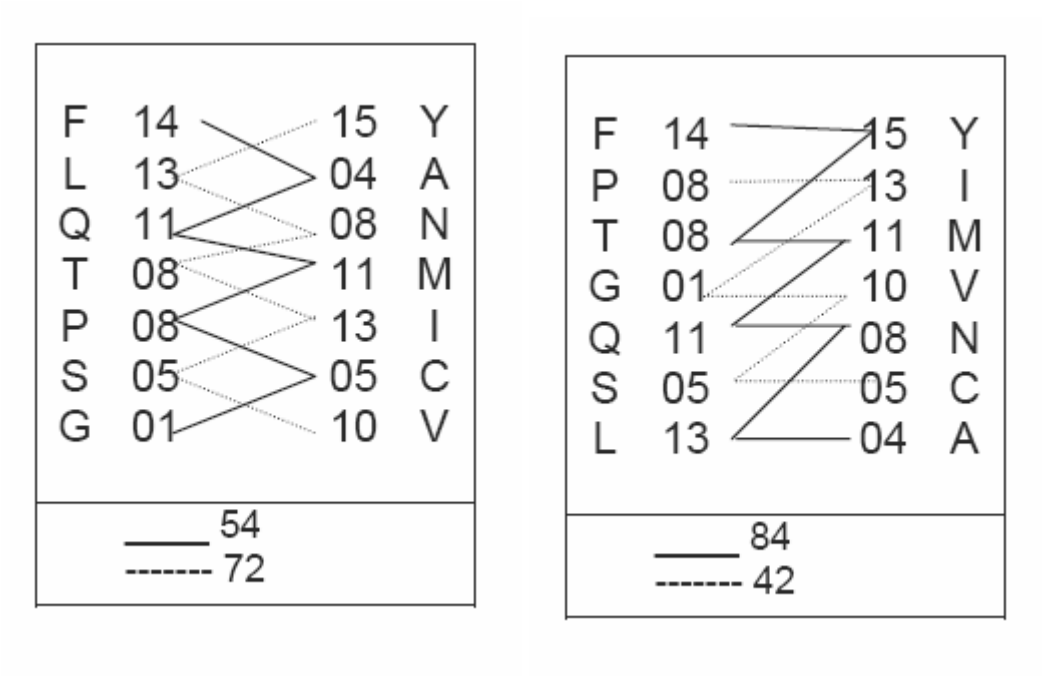
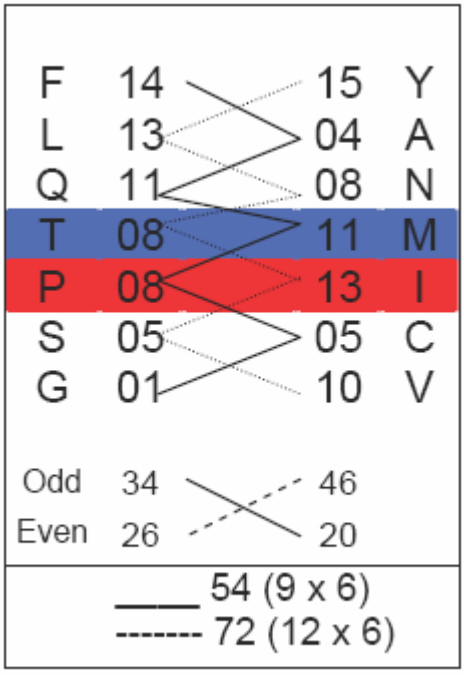


Table 6 (left). The atom number balance within two classes of AAs. Both classes as in Tables 4 and 5, but the order of AAs is given according to the number of atoms in side chains of "golden" AAs (AAs on the left). On the first zigzag (full) line there are exactly 9 x 6 whereas on the second one (dotted) 12 x 6 of atoms.

Table 7 (right). The atom number balance within two classes of AAs. Both classes as in Tables 4 and 5, but the order of AAs is given according to the number of atoms in side chains of "complement" AAs (AAs on the right). On the first zigzag (full) line there are exactly 14 x 6 whereas on the second one (dotted) 7 x 6 of atoms; both lines in proportion 2:1 what is "the symmetry in the simplest case" (Marcus, 1989).



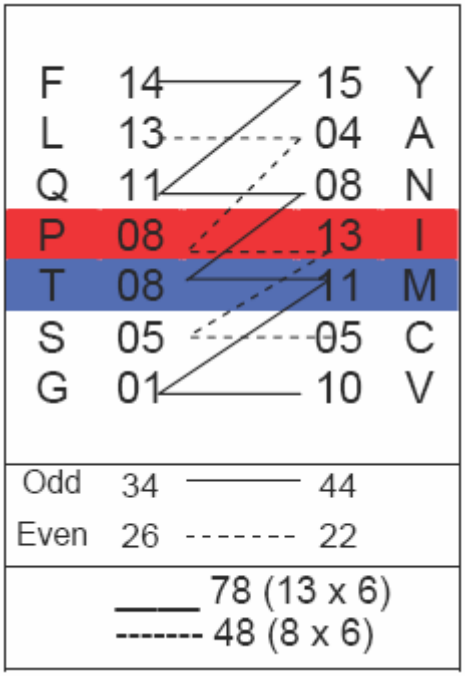


Table 8 (left). The Table as Table 6, but here with the results of calculation the number of atoms at even and odd positions, and with the mark of positions of P and T.

Table 9 (right). The Table as Table 8, but with a change of the position of P in relation to the T; so the S and the T are now in the contact, i.e. in neighborhood. For P is taken that it possesses 8 and not 9 atoms within the side chain because the "head" must be taken with a complete amine (NH_2) group, as by Shcherbak (1994) in nucleon number calculation. Realistically perceived, the order of P-08 here is determined by the order of his mate I-13.

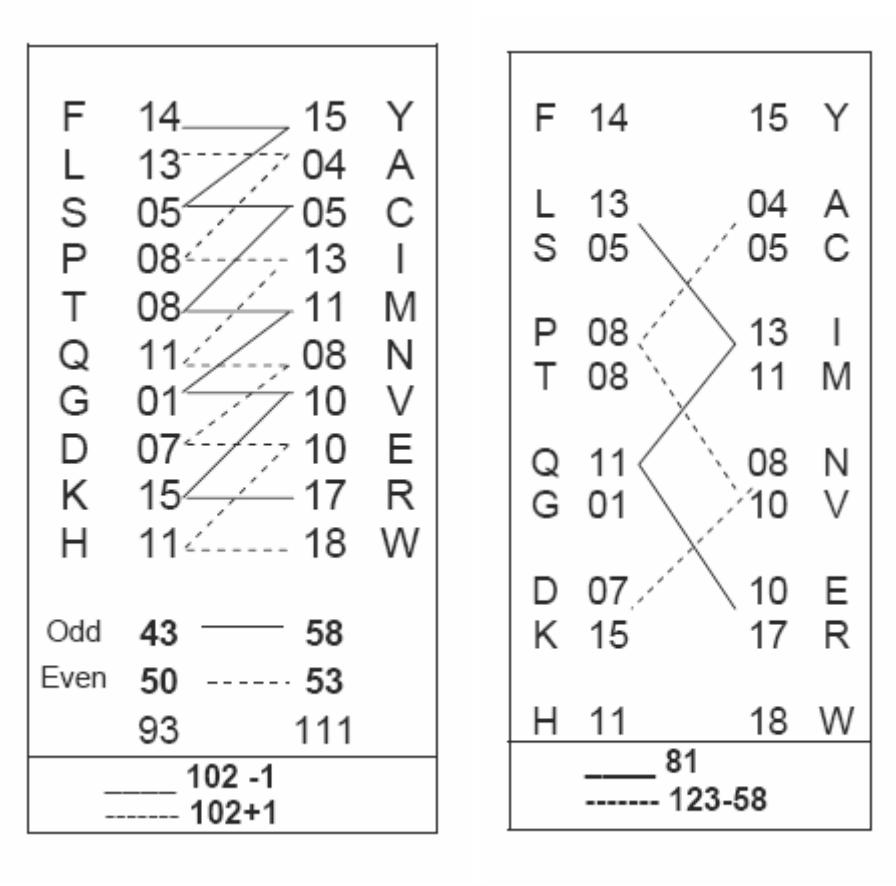


Table 10 (left). This Table is the same as Table 3 in all, except the sequence of "golden" AAs (as well as their complements) is in a vice versa order; still, the results of atom number calculations at even/odd positions. Notice that within aliphatic AAs on the left (column) there is 68 x 1, whereas on the right 78 atoms. Notice also that 78 + 58 (in aromatics) equals 68 x 2 (cf. Table B.2).

Table 11 (right). The same Table as Table 10, but given in amino acid pairs within the "quadruplets". The balance for aliphatic AAs corresponds to the balance valid for two amino acid classes handled by two classes of synthetases (Tables D.1 and D.2) [81 versus 123 atoms, where the aromatic AAs (their side chains) possess 58 of atoms]

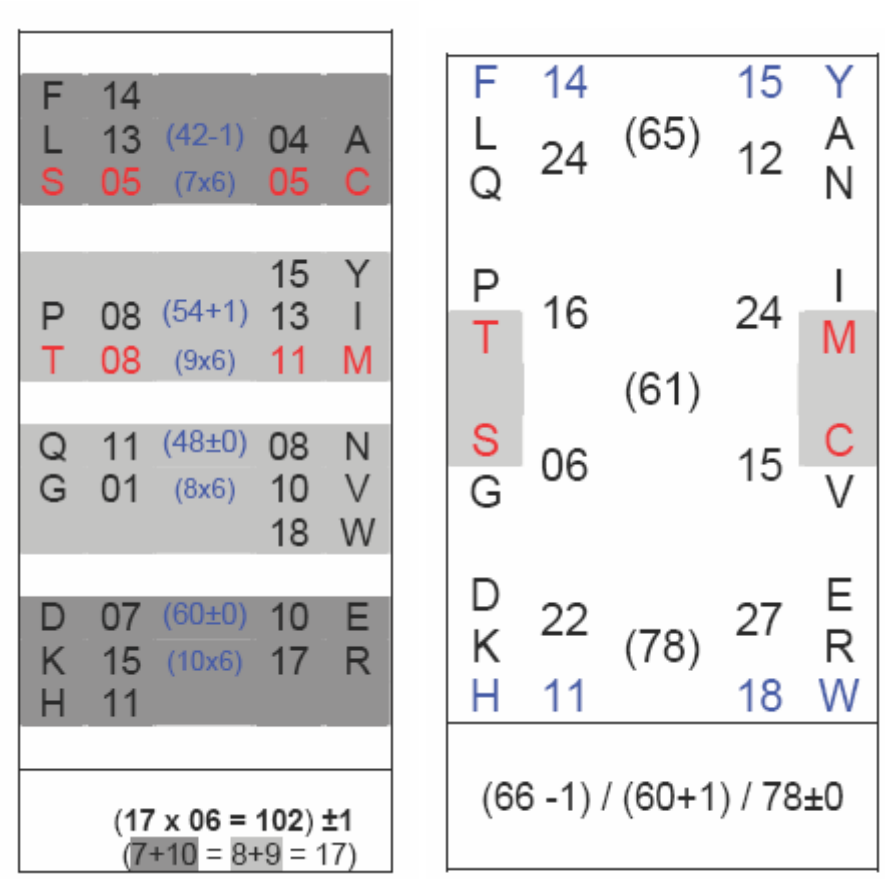


Table 12 (left). Table follows from Table 11 joining aromatic AAs to aliphatic ones: F and Y up, H and W down. Within two outer quintets (of dark tones) there are (17 x 06) - 1, whereas within two inner with light tones (17 x 6) + 1 of atoms.

Table 13 (right). Table is generated from Table 9 with an adding of non-complements and a new amino acid grouping. In relation to Table 3, instead 60, 66, 78 atoms here are (60+1), (66-1) and 78±0 of atoms.

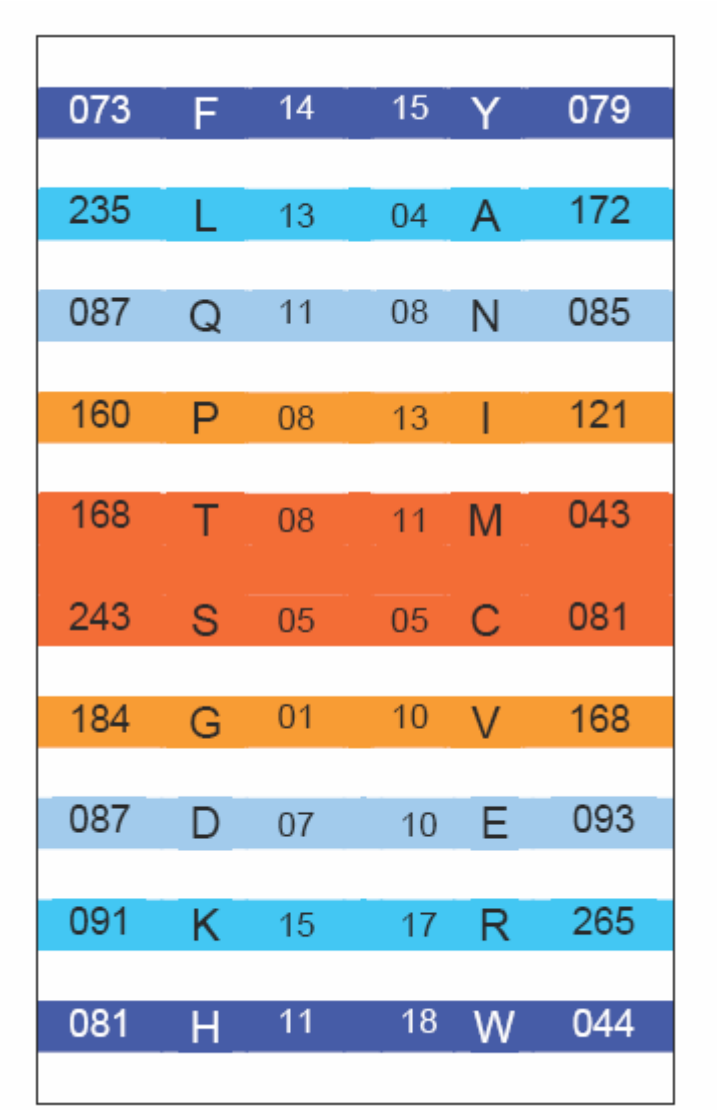


Table 14. The Cyclic Invariant Periodic System (CIPS) of canonical AAs. At the outer side, left and right, it is designated the number of atoms within coding codons; more exactly, in the Py-Pu bases (U = 12, C = 13, A = 15 and G = 16) (cf. Table 3 in Rakočević, 1997b, p. 648); at the inner side – the atom number within amino acid side chains (see Tables B.4 – B.9). In the middle position there are chalcogene AAs (S, T & C, M); follow - in next „cycle“ - the “contact” AAs (G, P & V, I), then two double acidic AAs with two their amide derivatives (D, E & N, Q), the two original aliphatic AAs with two amine derivatives (A, L & K, R); and, finely, four aromatic AAs (F,Y & H, W) – two up and two down.

COMMENTARIES

1. The order of codons in Table 1 corresponds to the chemical complexity of molecules: purine, as imidazole derivative of pyrimidine is more complex then pyrimidine. From two pyrimidines cytosine is more complex because it has two different functional groups (amino and oxo), while uracil two the same (oxo); in the other hand, cytosine participate in the pairing process with three hydrogen bonds, and uracil with two. From two purines guanine is more complex because it contains two different functional groups (amino and oxo), while adenine only one (amino); on the other hand, guanine is involved in pairing process with three hydrogen bonds, and adenine with two. From the said it follows that from the 24 possible permutations of the four amino bases (two of Py and two of Pu type) with chemical complexity corresponds only one, the first: UCAG, the same one that one can find in the original Crick's paper (Crick, 1968, Table 1, p . 368). (In the legend of this Table Crick says that it gives the "best allocations of the 64 codons".)[14] All other permutations are related to some other important properties of GC (Damjanović, 1998; Qiu and Zhu, 2000; Yang, 2004; Damjanović and Rakočević, 2005, 2006; Dragovich & Dragovich, 2006, 2007a, 2007b).

2. For now there is no explanation of what is the meaning of this determination with the number 6. Our opinion is that it may be related to the hypothesis that perfect and friendly numbers are determinants of the genetic code (Rakočević, 1997a, pp. 60-61) (Table A.1). Among other things, the six-bit binary tree is a particular and specific, in addition to everything else, because only at it the sum of numbers in the two internal branches (two octets) corresponds to the sum of the first pair of friendly numbers (220 + 284 = 504); and each two neighboring branches give the same result (504); all together as a realization of a logical square: (0) 220 + 284 = 504; (1) 156 + 348 = 504; (2) 92 + 412 = 504; (3) 28 + 476 = 504 (Table A.2). On the other hand, the sum of the first quartet of numbers (on the binary tree "0-63") is 6, the first octet 28, the first half numbers (from 0 to 31) 496, which is a realization of the first three perfect numbers; finally, if we count all the numbers 0-63, and return back (ciclicity!) (when the number 63 becomes 64, and 0

[14] If we calculate, for example, the positions of the golden mean, as in Table 3, but according to permutation GACU, therefore according to the latest permutation in the set of 24, then as the "golden" AAs appear: F, T, Q, N, D, E, S, R, G. Obviously, adequate chemical pairing with the "complements" is no longer possible. By this, it is important to mention that this AAs sequence, as well as that in Table 3 both follow from the standard (Table 1) as well as mitochondrial genetic code. (About differences between two codes see Comment 8).

becomes 127), we will receive as a result 8128, which is actually the fourth perfect number. Such a realization is possible only on the six-bit binary tree, not on any other. [Hint: The sum of the first four perfect numbers takes 13th place in the system of multiples of **the number 666**, which subsystem consists of multiples of the number 037, valid for the genetic code (Shcherbak, 1993, 1994, 2003, 2008; Rakočević, 2008) (Table A.1)[15].]

3. Having in mind that within the Sukhodolec's system of the distribution of hydrogen atoms to amino acids (Sukhodolec, 1985) the "key" classification is realized through the numbers 5 and 8 (as here), may be presented the hypothesis that – that comes from the specific position of these numbers in the Fibonacci series (5:8 = 0.625 and 8:5 = 1.6 versus for example 8:13 = 0.61538461 ...). The Sukhodolec's system: The **5** hydrogen atoms in G; 7 in A, S, D, C; 9 in P, T, Q, H; 11 in V, F, M, Y and 13 in L, I. The **8** hydrogen atoms in N; 10 in Q, 12 in W and 14 in K,R.

4. In a first attempt (Rakočević & Jokić, 1996) we have named the „contact amino acids" as "etalonic" ("alanine stereochemical type as a measurement subject, and three other types as measurement *etalons* and measurement subject at the same time") and later (Rakočević, 2006) as "contact" AAs; contact, because the amino acid side chain ("body") is directly connected with amino acid functional group ("head"), while by AAs of alanine type this is not the case. At amino acids of this type the connection is mediated by an H-C-H group; in all except in threonine, where the connection is realized through the H-C-CH_3 group[16]. And, just this situation could deny our concept that four stereochemical types, as it was originally indicated by Popov (1989), are at the same time four types not only of configuration but of constitution (in the plane) too. Namely, at the first glance it appears to be the same by isoleucine as by threonine (the same "screen": H-C-CH_3)[17]. Yes, but the binding within threonine, as well as within all the remaining 15 AAs of alanine stereochemical type is realized through a primary carbon, and within isoleucine through the secondary one.

---

[15] „Here is wisdom. Let him that hath understanding count the number of the beast: for it is the number of a man; and his number is Six hundred threescore and six" (The New Testament, Revelation, 13: 18).

[16] Obviously, here is a hydrogen atom from the H-C-H group replaced by a methyl one.

[17] From this objective chemical similarity follows that only these two AAs (T & I) have not only enantiomers but still diastereomers.

5. By itself it is understandable that the existence of a such CIS (Cyclic Invariant System) assumes the existence of a corresponding PIS and NIS (Protein Interactive System, and Nucleotide Interactive System, respectively), as in the act of origin of the genetic code, and in all stages of evolution of the life on Earth.

6. In their certain setting, all three hypotheses on the origin of the genetic code, including the Sukhodolec's one, imply the existence of a prebiotic evolution of the genetic code only, and not the post biotic (see footnote 12); in other words, all these hypotheses imply a complete genetic code from the very beginning[18]. The so-called exceptions to the standard genetic code, represent only the degree of freedom. Unfortunately, in the current science the majority of research people is still of the opinion, that it dates from the early seventies, that the life started with an incomplete genetic code of the seven-eight, then ten, fifteen, eighteen, to be stopped - as "frozen" – to the "today's" twenty amino acids[19].

7. Within decimal numbering system there are the same-two-digit numbers: 11, 22, ..., 55, 66, ... 99; the first 11, the central 55 and the last 99. If one takes a rule: to be a change for 1 in the first, and then in second digit-position, that leads to the system, which middle case is this pattern in the genetic code (Table B.9).

8. The differences in mitochondrial code in relation to standard one are as follows: Isoleucine (I) AUU, AUC; Methionine (M) AUA, AUG; Tryptophan (W) UGA, UGG; „Stop" codons: UAA, UAG (as in standard one) and AGA, AGG, instead these last two codons to be coded for arginine (R) as in standard genetic code.

---

[18] The hypothesis on *a complet genetic code* the reader can find in our previous work (Rakočević, 2004, p. 231: „By this hypothesis, derived from presented facts as we understand them, we support the stand point that genetic code is one and unique, universal, valid for everything living, in fact, it is the condition for origin and evolution of life").

[19] Jukes, 1973, p. 24: „The idea that the number of amino acids in earlier codes was less than twenty is favoured by the fact that some of 'simple' amino acids, such as serine, glycine and alanine, have four of more codons, and that the codons may be arranged in 'quartets' ... in which only the first two bases confer specificity on the amino acid". As we now see from CIPS (Table 14) the key principle, valid for the genetic code is a *Block aufbau principle* which requires the presence of each amino acid, each codon, and each base in it, in an exact and correct position within the system itself.

9. From regularities, harmonic structures and proportions containning within mitochondrial genetic code (see Appendix C) it follows that standard and mitochondrial code posses a paralel and colective prebiotic evolution; colective in the sense of a unity and harmonized dynamic of PIS and NIS (Protein Interactiv System and Nucleotide Interaktive System).

Acknowledgement

I am grateful to Branko Dragovich for the indication of the Sukhodolec's works and helpful, stimulating discussion.

REFERENCES

Crick, F.H.C (1968) The origin of the genetic code, J. Mol. Biol., 38, 367-379.

Damjanović, M. Z. (1998) Logic core of genetic code, Glasnik of Montenegrin Academy of Sciences and Arts (CANU), 12, 5-8.

Damjanović, M. Z., Rakočević, M. M. (2005) Genetic code: an alternative model of translation, Ann. N. Y. Acad. Sci. 1048, 517-523.

Damjanović, M. Z., Rakočević, M. M. (2006) Genetic Code: A new understanding of codon - amino acid assignment, arXiv:q-bio/0611033v1 [q-bio.OT].

Dragovich B., Dragovich, A. (2006) A p-Adic model of DNA sequence and genetic code, p-Adic Numbers, Ultrametric Analysis and Applications 1, 34-41; arXiv:q-bio/0607018v1 [q-bio.GN].

Dragovich B., Dragovich, A. (2007a) p-Adic Degeneracy of the Genetic Code. arXiv:0707.0764v1 [q-bio.GN].

Dragovich, B., Dragovich, A. (2007b) p-Adic modeling of the genome and the genetic code, Computer Journal, doi:10.1093/comjnl/bxm083; arXiv:0707.3043v1 [q-bio.OT].

Jukes, T. H. (1973) Possibilities for the evolution of the genetic code from a preceding Form, Nature, 246, 22-27.

Kyte, J., Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein, J. Mol. Biol., 157, 105-132.

Marcus, S. (1989) Symmetry in the simplest case: the real line, Comput. Math. Appl., 17, 103-115.

Popov, E. M. (1989) Strukturnaya organizaciya belkov, Nauka, Moscow.

Qiu Y, Zhu L. (2000) The rearranged genetic code and its implications in evolution and biochemistry, Biosystems, 56 (2-3), 139-144.

Rakočević, M. M. (1997a) Genetic code as a unique system, SKC Niš, Bina Beograd; www.sponce.net

Rakočević, M. M. (1997b) Two classes of the aminoacyl-tRNA synthetases in correspondence with the codon path cube, Bull. Math. Biol., 59, 645-648.

Rakočević, M. M. (1998) The genetic code as a Golden mean determined system, Biosystems 46, 283-291.

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

Rakočević, M. M. (2006) Genetic code as a harmonic system, arXiv:q-bio/0610044v1 [q-bio.OT]

Rakočević, M. M. (2008) Genetic Code: Four-Codon and Non-Four-Codon Degeneracy, arXiv:0802.1056v1 [q-bio.BM]

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

Shcherbak, V. I.(1993) Twenty canonical amino acids of the genetic code: the arithmetical regularities, part I, J Theor. Biol., 162, 399-401.

Shcherbak, V. I. (1994) Sixty-four triplets and 20 canonical amino acids of the genetic code: the arithmetical regularities. Part II, J Theor. Biol., 166, 475-477.

Shcherbak, V. I. (2003) Arithmetic inside the universal genetic code, Biosystems 70, 187-209.

Shcherbak, V. I. (2008) The arithmetical origin of the genetic code, in: M. Barbieri (ed.), The codes of life: the rules of macroevolution (pp. 153-181), Springer, Berlin.

Sukhodolec, V. V. (1981) Evolucionie preobrazovaniya geneticheskogo koda, predskazivanie iskhodya iz gipotezy o fizicheskom predopredelenii naborov osnovaniy v sostave kodonov, Genetika, XVIII, 3, 499-501.

Sukhodolec, V. V. (1985) Smisl geneticheskogo koda: rekonstrukciya etapa predbiologicheskoy evolucii, Genetika, XXI, 10, 1589-1599.

Swanson, R. (1984) A unifying concept for the amino acid code, Bull. Math. Biol., 46, 187-207.

Wetzel, R. (1995) Evolution of the aminoacyl-tRNA synthetases and the origin of the genetic code, J. Mol. Evol., 40, 545-550.

Wong, J. T-F. (1975) A co-evolution theory of the genetic code, Proc. Nat. Acad. Sci. USA, 72, 1909-1912.

Yang, C.M. (2004) On the 28-gon symmetry inherent in the genetic code interwined with aminoacyl.tRNA synthetases-the Lucas series, Bull Math Biol., 66 (5), 1241-1257.

Appendix A: Some additional relationships within three splittings

x	6x	(1:3)x	66x	(11:3)x	666x	(111:3)x
27	162	9.000	1782	99.000	17982	999
26	156	8.666	1716	95.333	17316	962
25	150	8.333	1650	91.666	16650	**925**
24	144	8.000	1584	88.000	15984	888
23	138	7.666	1518	84.333	15318	851
22	132	7.333	1452	80.666	14652	814
21	126	7.000	1386	77.000	13986	777
20	120	6.666	1320	73.333	13320	740
19	114	6.333	1254	69.666	12654	703
18	108	6.000	1188	66.000	11988	666
17	102	5.666	1122	62.333	11322	629
16	96	5.333	1056	58.666	10656	**592**
15	90	5.000	990	55.000	9990	555
14	84	4.666	924	51.333	9324	518
13	78	4.333	858	47.666	**8658**	481
12	72	4.000	792	44.000	7992	444
11	66	3.666	726	40.333	7326	407
10	60	3.333	660	36.666	6660	370
09	54	3.000	594	33.000	5994	**333**
08	48	2.666	528	29.333	5328	296
07	42	2.333	462	25.666	4662	**259**
06	36	2.000	396	22.000	3996	**222**
05	30	1.666	330	18.333	3330	185
04	24	1.333	264	14.666	2664	148
03	18	1.000	198	11.000	1998	**111**
02	12	0.666	132	7.333	1332	074
01	**06**	0.333	**66**	3.666	**666**	**037**

Table A.1. The integer multiples of 6, 66, 666 and 037. This Table is a variant of Table 3 in Ref. (Rakočević, 2008) as well as of Table 2 in Ref. (Rakočević, 1997a, p. 61). The multiples of 6, 66 and 666 correspond to the multiples of 1/3, 11/3 and 111/3 respectively. As it is directly visible only

in third step all multiples are integer, starting with Shcherbak's "Prime quantum 037". [In 6th step the "Prime quantum" is 37037, in 9th 37037037 an so on.] The red patterns in last column correspond to the patterns of nucleon number in Genetic code as follows: 10 x 111 of nucleons within 15 side chains as well as 15 "heads" of non-four-codon AAs, total 10 x 222. The 333 of nucleons within 8 side chains of four-codon AAs. The 592 (a half of third friendly number) of nucleons within 8 "heads" of four-codon AAs, in total: 333 + 592 = 925; and 592 – 333 = 259. Notice here a determination with Pythagorean law, because 333 = 9 x 37; 592 = 16 x 37 and 925 = 25 x 37. The number 8658 = 7770+088 in 13th row represents the sum of the first four perfect numbers (6+28+496+8128 = 8658). Within the set of 6, 66, 666, 6666, 66666 etc., only the number 666 corresponds to the "Prime quantum" 037, valid for the genetic code.

/00 - 07/08 - 15/16 - 23/24 - 31//32 - 39/40 - 47/48 - 55/56 - 63/
28 92 156 **220** **284** 348 412 476
64 64 64 64 64 64 64
/00 - 07/00 - 15/00 - 23/00 - 31//00 - 39/00 - 47/00 - 55/00 - 63/
28 120 276 **496** 780 1128 1540 2016
92 156 220 284 348 412 476

Table A.2. The determination of the series of the numbers 0-63 (in correspondence with the six-bit binary tree) by perfect numbers (here visible: 28 and 496) and by the sum consisted of the first pair of the friendly numbers: 220+284 = 504 (after: Rakočević, 1997a, Figure 7, p. 60).

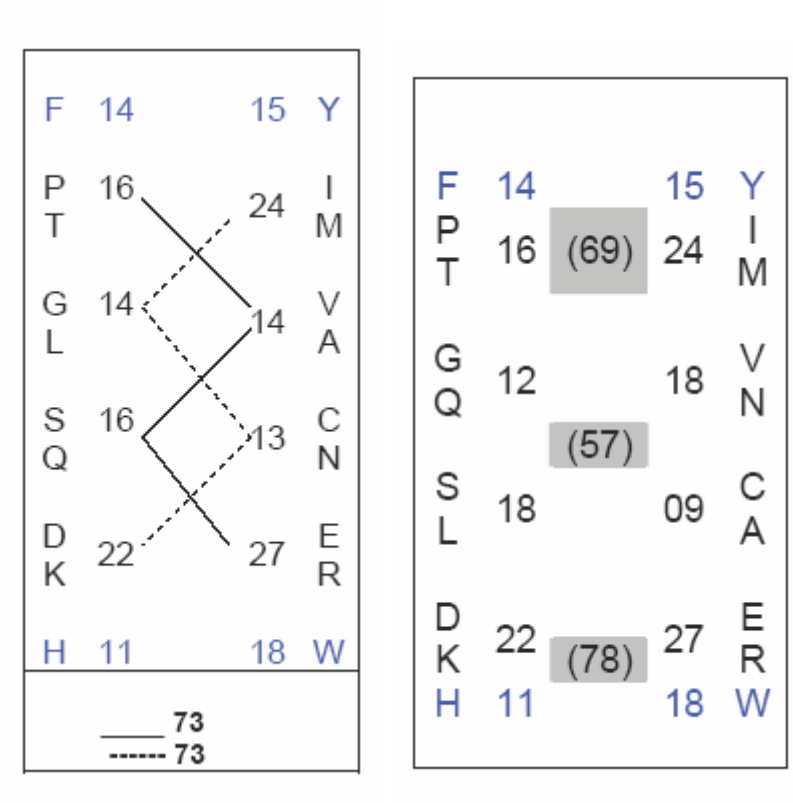


Table A.3 (left). This Table is the same as Table 5, except three pairs of amino acids are adding, and aromatic AAs are up and down. The order follows the positions of complements on the binary-code tree (Fig. 1).

Table A.4 (right). The order follows the size of complement-molecules, with a new amino acid grouping. For the digital patterns of the atom number notations the principle of minimum change is valid in a specific manner: on the second position 5-6-7 and on the first one 7-8-9.

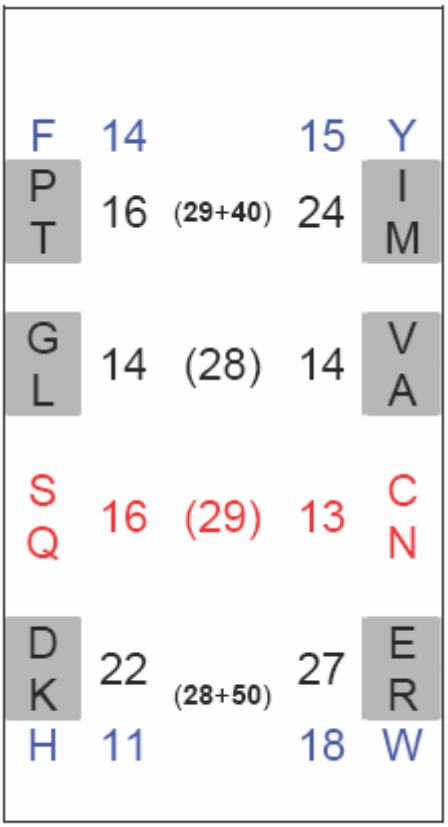


Table A.5. This Table is the same as Table A.3 with a new amino acid grouping. The effect of the principle of minimum change as well as the principle of block-aufbau is obvious: 28-29; 40-50.

Appendix B: Some additional relationships within Cyclic Invariant Periodic System (CIPS) of standard genetic code

54	F 14 L 13 Q 11 P 08 T 08	15 Y 04 A 08 N 13 I 11 M	51
39	S 05 G 01 D 07 K 15 H 11	05 C 10 V 10 E 17 R 18 W	60
Odd Even	49 44	61 50	

Table B.1. Atom number within four quarters and two halves of CIPS. The effects of harmonic principles are self-evident; the multiples of number 6 also: 54 = 9 x 6 (44 = 54 – 10); 60 = 10 x 6 (50 = 60 – 10); 49 = [(8 x 6) + 1] (39 = 49 – 10); 61 = [(10 x 6) + 1] (51 = 61 – 10).

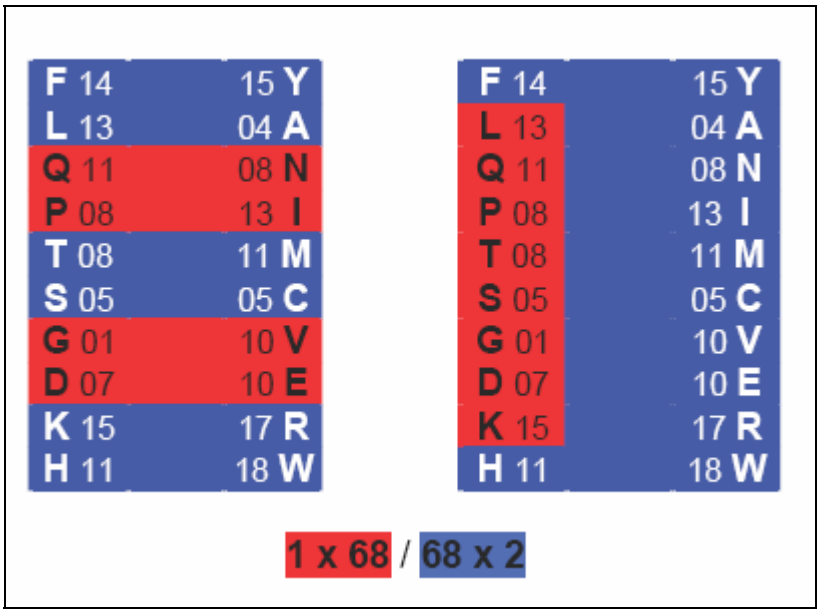


Table B.2. The number of amino acid molecules in proportion 2:3 and the number of atoms within side chains 1:2. On the right: the quantum "1 x 68" make the aliphatic "golden" AAs plus two smaller aliphatic non-complements (D & K), and the quantum "2 x 68" all others. On the left: the quantum "1 x 68" make four AAs (G, P and V, I) of non-alanine stereochemical type (larger diversity!) plus two and two AAs – two double acidic (D, E) and only two amides (N, Q), what means a very specific diversity; and the quantum "2 x 68" all other AAs.

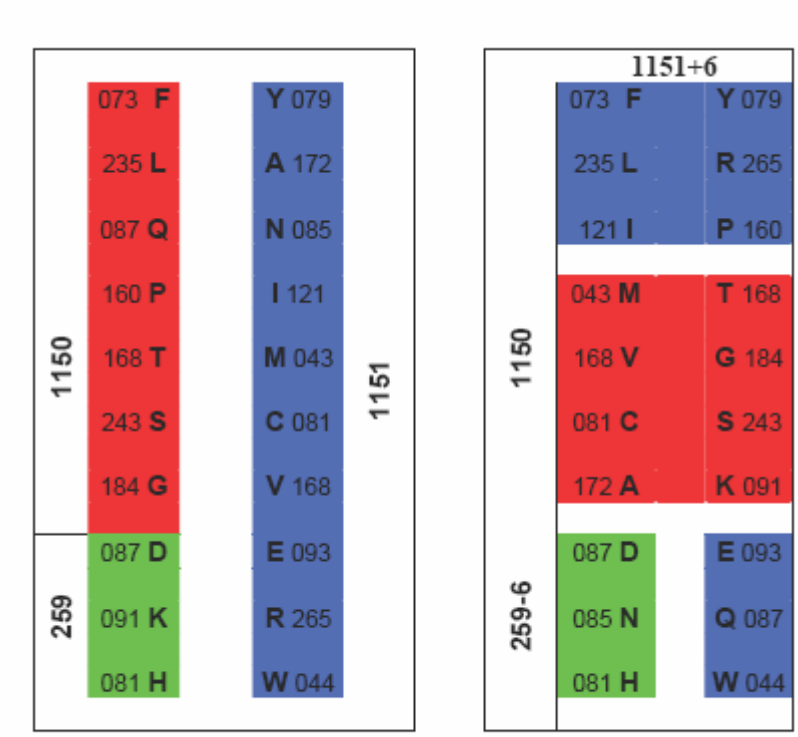


Table B.3 (left). The block-aufbau principle as well as the principle of self-similarity are on the scene: the quantum of 1150 atoms within coding codons (their Py-Pu bases) for seven „golden“ AAs (red block) in relation to the same such quantum valide for eight AAs in the system over there in Table B.4. At the same time the principle of minimum change is valid: 1150 versus 1151 (blue block). Notice also these relations: 1151 + 259 = 01410 („golden“ AAs versus „non-golden“ AAs in relation to 11**5**0 : 1**4**10). Quantum „1150“ make all seven „golden“ AAs.

Table B.4 (right). The self-similarity through the same patterns: the seven nonpolar AAs (F, L, I, M, V, C, A) play the role of seven “golden” AAs (F, L, Q, P, T, S, G) in previous Table (Table B.3). The polar/nonpolar AAs after (Kyte & Doolittle, 1982). The quantum „1150“ make three „golden“ AAs (T, G, S) together with three their complements (M, V, C) plus two other – A & K (Alanine as first possible hydrocarbon amino acid; lysine as a simple amine derivative – the simpler of total two: K and R).

Table B.5. The arrangement (as in Table B.4) into two and two areas; in the first of blue and green tones there are 1443 of atoms within coding codons (Py-Pu bases); and in second one of orange and pink tones, there are 1443 – 326 atoms. (The quantum "326" is the same as the number of atoms within nine nucleotides in three stop codons). At the same time the 23 amino acid molecules (Table 1) possess exactly 1443 of atoms. [The six-codon AAs (L,S,R) are included two times, as by Shcherbak (1994).] Notice that1443 is 1/6 of 8658 where 8658 is the sum of the first four perfect numbers (6+28+496+8128 = 8658 = 7770+0888) (cf. the 13th position within the system, presented in Table A.1.

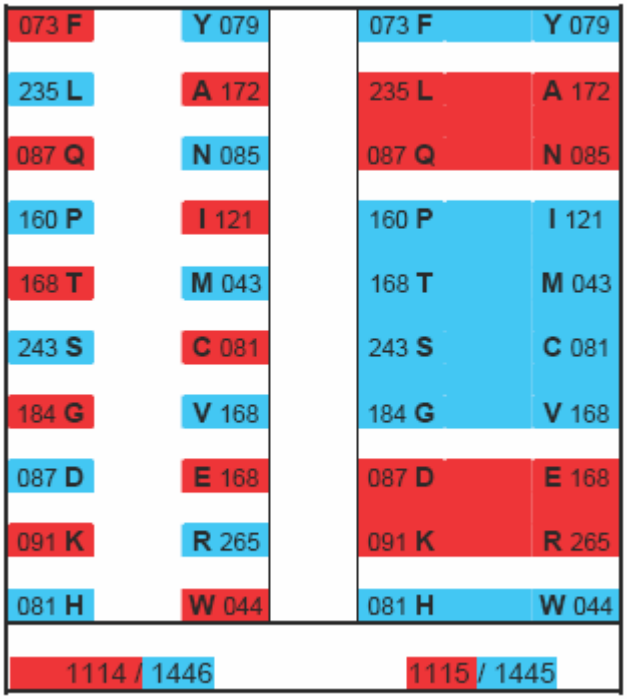


Table B.6 The number of molecules in proportion 1:1 on the left and 2:3 on the right. The atom number balance within coding codons (Py-Pu bases), regarding both arrangements, as 1114:1115 and 1446:1445.

87 - 1	14 **F**	**Y** 15		073 **F**	**Y** 079	1012
	13 **L**	**A** 04		235 **L**	**A** 172	
	11 **Q**	**N** 08		087 **Q**	**N** 085	
	08 **P**	**I** 13		160 **P**	**I** 121	
29	08 **T**	**M** 11		168 **T**	**M** 043	535
	05 **S**	**C** 05		243 **S**	**C** 081	
88+1	01 **G**	**V** 10		184 **G**	**V** 168	1013
	07 **D**	**E** 10		087 **D**	**E** 168	
	15 **K**	**R** 17		091 **K**	**R** 265	
	11 **H**	**W** 18		081 **H**	**W** 044	

Table B.7 The atom number balance within amino acid molecules (side chains) on the left as well as within coding codons (Py-Pu bases) on the right, in relation to middle amino acid class.

n		a	b	AAs
1	Ribose-5-phospate	11	081	H
2	3-Phosphoglycerate	11	508	G S C
3	Pyruvate	27	575	A L V
4	2-Oxoglutarate	46	605	P E Q R
5	Phosphoenolpyruvate	47	196	W F Y
6	Oxaloacetate	62	595	T M I D N K
	Odd	85	[(853-1) x 1]	
	Even	119	[(853+1) x 2]	

Table B.8. The six amino acid biosynthetic precursors. The precursors order (n) after number of atoms (a) within the sets of belonging AAs and the sets of belonging coding codons (b). Notice a balance in column "b" (the proportion 1:2 with a deviation for ±1) and disbalance in column "b" (85:119).

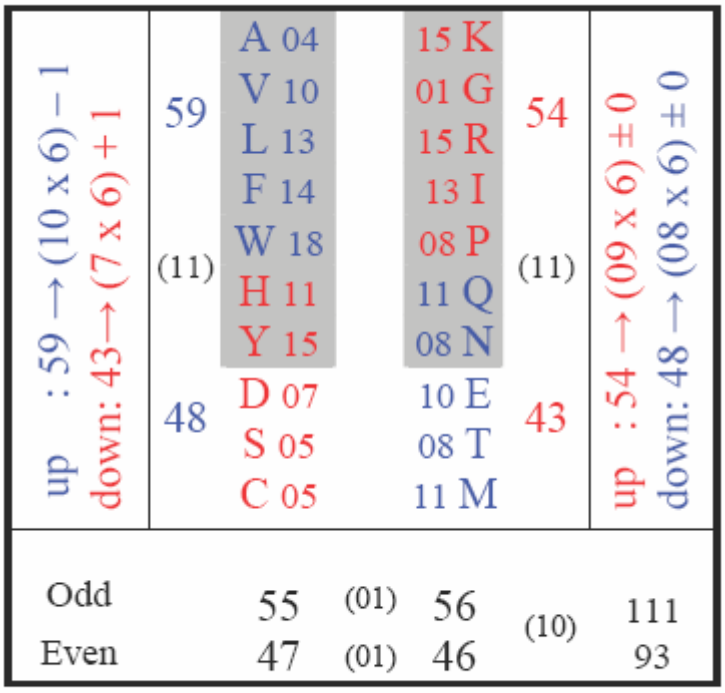


Table B.9. As in Table B.4, with seven nonpolar AAs, the self-similarity through the same patterns is still once on the scene: the seven AAs from the odd positions in precursor system (Table B.8) (H; A, L, V; W,F,Y) play the role of seven "golden" AAs in Table B.3. The system is starting by source aliphatic AAs (A-L and K-R) with an involving of two simpler AAs from class (G-V, P-I); follow four and four AAs with a large scale of diversity; on the left four aromatic AAs, two inner (heterocyclic) and two outer ("pure" aromatic), and on the right: two more complex (N, Q) from class (D-E & N-Q) and two more complex from the class (G-V, P-I). The next, key position, keep two double acidic AAs (D,E), belonging to two classes at the same time: they belong to the class of chalcogene AAs (down) and to class with two amides. The effects of harmonic principles are self-evident, including a special algorithm (see Comment 7). Notice also a full balance for atom number (55 + 47 = 56 + 46 = 102) in contrary to disbalance (85:119) in Table B.8.

Appendix C: Some additional relationships within Cyclic Invariant Periodic System (CIPS) of mitochondrial genetic code

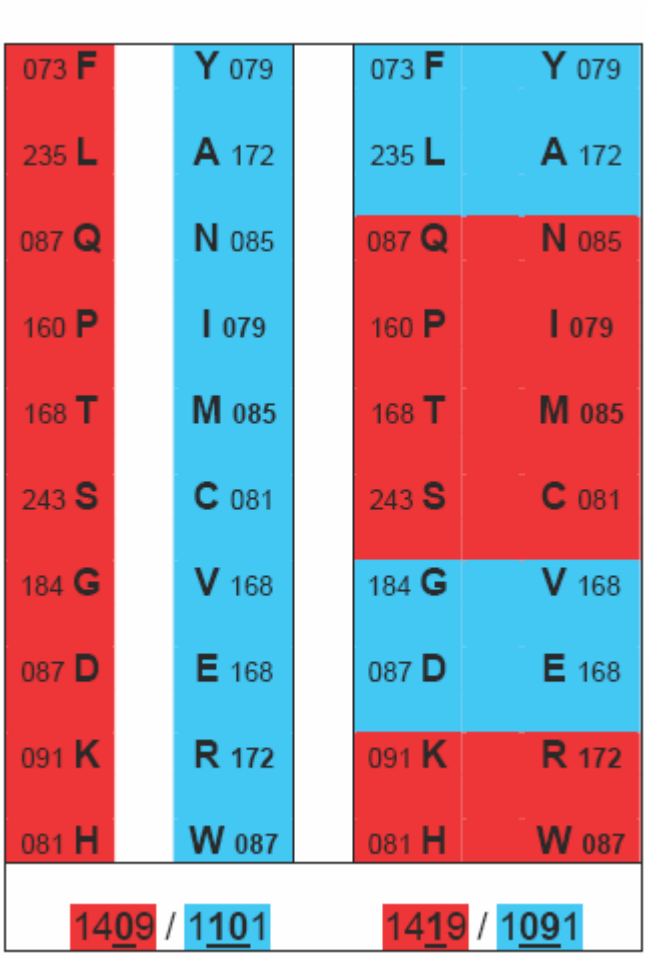


Table C.1. The mitochondrial genetic code in relation to the standard one (cf. Table B.6 and see Comments 8 and 9). As in standard code (Table B.6) here we have the same proportions: the number of molecules in proportion 1:1 on the left and 2:3 on the right. The atom number balance within coding codons (Py-Pu bases), regarding both arrangements, as 14**0**9:14**1**9 and 1101:1**09**1.

073	F	91	107	Y	079
235	L	57	15	A	172
087	Q	72	58	N	085
160	P	41	57	I	079
168	T	45	75	M	085
243	S	31	47	C	081
184	G	01	43	V	168
087	D	59	73	E	093
091	K	72	100	R	172
081	H	81	130	W	087

550 + (550+155) = 1255

1255 1255

Table C.2. Two columns of AAs with two and two columns of numbers; outer: number of atoms within coding codons (Py-Pu bases); inner: number of nucleons within amino acid molecules (side chains). As in Table C.1 here we have the same proportion for the number of molecules, 2:3, but the proportion for the number of atom is 1:1. But relations between atom number within codons and nucleon number within amino acid molecules is noteworthy in a special respect: the number of atom 2 x 1255 and the number of nucleons 1 x 1255 ?! (Within the columns, on the left: 550 and on the right 550+155.) If the Nature is still of a great "N", and then is – too much!

	U	C	A	G	G	A	C	U	
F	5	1	0	0	0	2	1	3	Y
P	1	9	1	1	0	2	1	3	I
T	1	5	5	1	1	3	0	2	M
D	1	1	2	2	3	3	0	0	E
K	0	0	5	1	5	1	5	1	R
H	1	3	2	0	3	1	0	2	W
	09	19	15	05	12	12	07	11	
	(U+A / C+G) (24 / 24) 48 (8 x 6)				(G+C / A+U) (19 / 23) 42 (7 x 6)				

Table C.3. The number of nucleotides within coding codons (Py-Pu bases) in blue areas of previous Table (Table C.2).

	U	C	A	G	G	A	C	U	
L	9	5	2	2	5	1	5	1	A
Q	0	2	3	1	0	4	1	1	N
S	6	6	3	3	2	0	1	3	C
G	1	1	1	9	5	1	1	5	V
	16	14	09	15	12	06	08	10	
	(U+A / C+G) (25 / 29) 54 (9 x 6)				(G+C / A+U) (20 / 16) 36 (6 x 6)				

Table C.4. The number of nucleotides within coding codons (Py-Pu bases) in red areas of Table C.2. Regarding two Tables (C.3 & C.4) together we see a determination with multiples of number 6 in accordance with the continuity principle: (**6** x 6), (**7** x 6), (**8** x 6) and (**9** x 6).

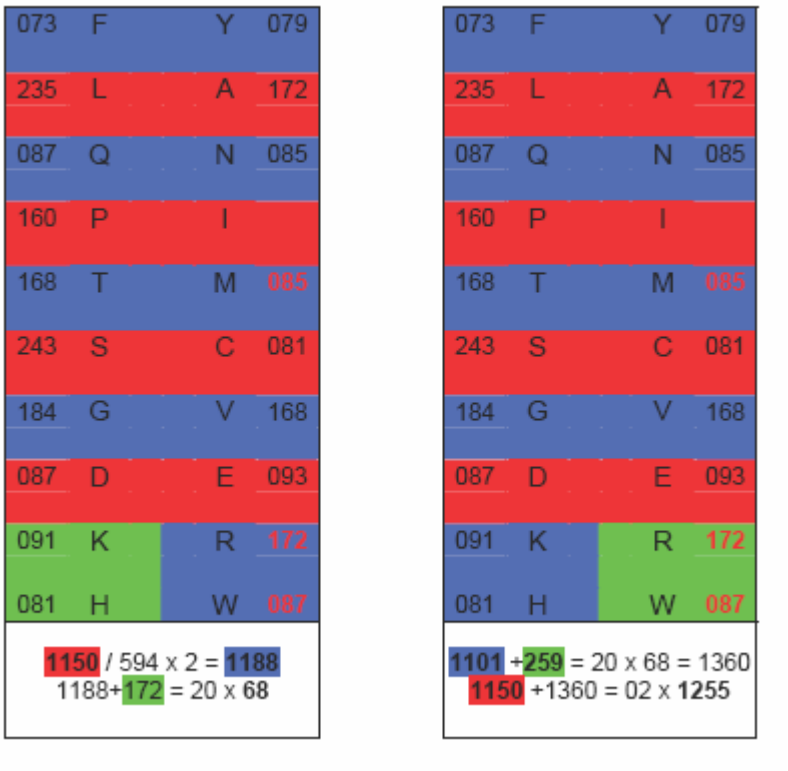


Table C.5 (left). The quantum of "1150" atoms within coding codons (Py-Pu bases) we had two times in standard genetic code, in Tables B.3 and B.4. The quantum "594" corresponds to the number of atoms within 61 amino acid molecules (side chains) in standard genetic code (Table 1).

Table C.6 (right). This Table, in relation to previous one, shows that two basic AAs (K & R) together with two heterocyclic AAs (H & W) make "a tab on the scale" in the balance.

Appendix D. Some harmonic structures from previous works

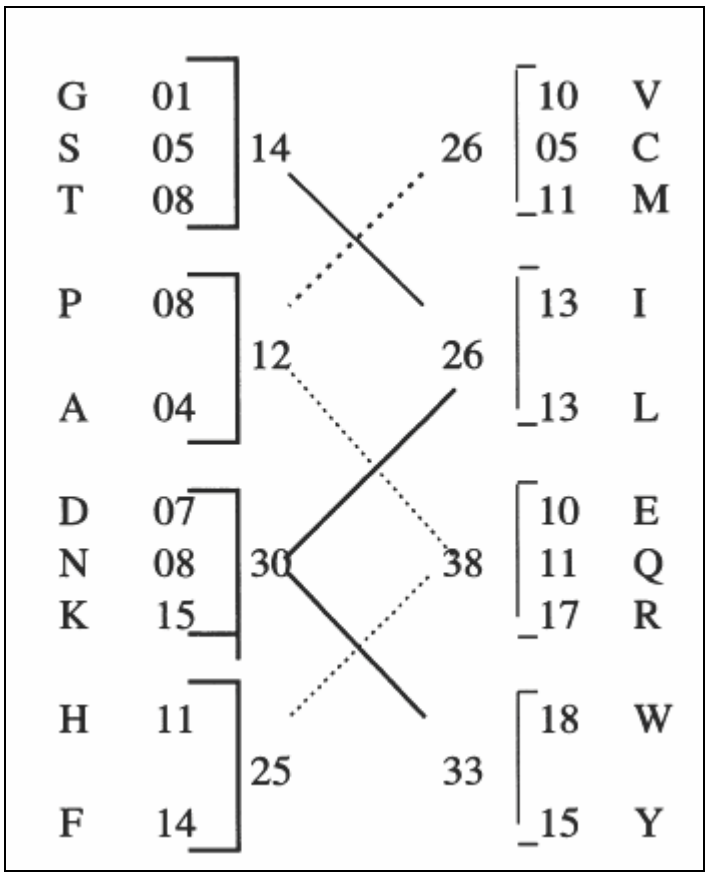


Table D.1. Two amino acid classes generated through the influence of two catalysts (Scheme 5 in Rakočević, 1998). On the left the smaller molecules of AAs, handled by class II of aminoacyl-tRNA synthetases, whereas on the right the larger molecules of AAs, handled by class I of synthetases. On the full line there are 102+1 and on the dotted one 102-1 of atoms. The distinct arrangement of five classes from CIPS (Table 14) is self-evident: the contact AAs (G-V and P-I) surround the chalcogene AAs (S-C and T-M). In the other hand four source aliphatic AAs (A-L and K-R) surround the class of two double acidic AAs (D & E) and their amides (N & Q).

28	09 19	G P A K		23 30	V I L R	53	**81**
53	13 15 25	S T D N F H		16 21 33	C M E Q Y W	70	**123**
81						**123**	204

Table D.2. The arrangement in accordance to the principle: "a little" and "full" in relation to "small" and "large". So, on the left there are AAs (81 atoms) from the left side of previous Table (class II), on the right from right side (class I of AAs with 123 atoms within side chains). At the same time very up there are AAs (81 atoms) just aliphatic and nonpolar (A,V, L, I) and "a little" polar (G, P, K, R) (hydrogen and nitrogen are less polar then oxygen!); in the other hand, except aromatic and sulfur AAs, down are AAs (the row with 123 atoms) also aliphatic, but "full" polar.

					a	b	c	d	M
D	N	A	L	→	**189**	**189**	**221**	**221+3**	485.49 ≈ **485**
R	F	P	I	→	*289*	*289*	*341*	*341+0*	*585.70 ≈ 586*
K	Y	T	M	→	299	299	351	351+2	595.71 ≈ 596
H	W	S	C	→	*289*	*289*	*331*	*331+1*	*585.64 ≈ 586*
E	Q	G	V	→	**189**	**189**	**221**	**221+3**	485.50 ≈ **485**
					1255	1255	1465	1465+9	2738.04

Table D.3. Rakočević, 2004, Table 1, p. 223: "Four choices after four types of isotopes: (a) the number of nucleons within 20 AAs side chains, calculated from the first, the lightest nuclide (H-1, C-12, N-14, O-16, S-32). (b) The number of nucleons within 20 AAs side chains, calculated from the nuclide with the most abundance in the nature [the same patterns as in (a): H-1, C-12, N-14, O-16, S-32; at heavier nuclides of other bioelements the data by (a) and (b) are not the same]. (c) The number of nucleons within 20 AAs side chains, calculated from the nuclide with the less abundance in the nature (H-2, C-13, N-15, O-17, S-36); (d) The number of nucleons within 20 AAs side chains, calculated from the last, the heaviest nuclide (H-2, C-13, N-15, O-18, S-36). (M) The AAs molecule mass. Notice that (d) is greater from (c) for exactly one modular cycle (in module 9) and that total molecules mass is equal to 2×37^2. Notice also that molecule mass within five rows is realized through the same logic-patterns of notations as the first nuclide, i.e. isotope."

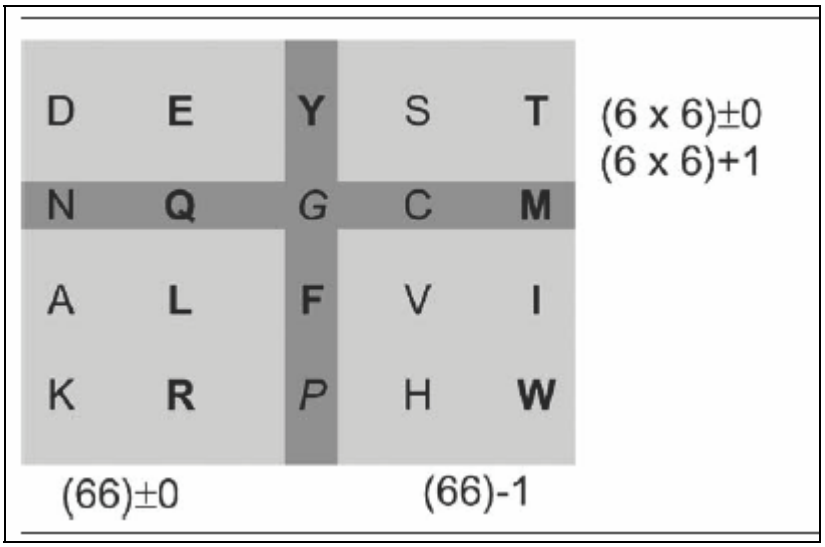


Table D.4. Rakočević, 2004, Table 9, p. 229: "This system follows from the system in Table 4. First row (down): N-ended AAs. Second row: solely C-ended AAs. Last row (up): O-ended AAs. First to last row: remaining five AAs (one solely H-ended, two S-ended and two N-, O-ended, all five as a ''combination''. Within the cross there are only the exceptions: horizontally five the mentioned combining AAs; vertically: Y as aromatic within aliphatic AAs; G without carbon; F as aromatic within aliphatic AAs; and, finally, P as cyclic aliphatic amino acid. In the system there is a balanced proportionality as follows: within horizontal leg of the cross there are (6 x 6) ±0 of atoms, and within vertical leg (without glycine), there are (6 x 6) +1. Without cross: on the left there are (66) ±0 and (66)-1 on the right. [cf. this ''combination-cross'' sub-system with ''four-codon-AAs-cross'' sub-system in Codon path cube (Swanson, 1984; Rakočević, 1997b)]." Now four CIPS amino acid classes are self-evident. Notice that aromatic AAs make a "crossing" with the contact ones, but in Table B.9 only with two contact AAs (P, I) and still two amides (N, Q).

IV	III	II	I
G-V	T-M	P-I	K-R
S-C	D-N	H-W	
Q-E	Y-F		
A-L			

Table D.5. Rakočević, 2000, Table 4, p. 281 and Rakočević, 2006a, Table 4, p. 6: "Four Amino acids classes. If amino acids pairs must be as in (Rakočević, 1998), then four classes follow from the system presented in Table 3. Non-bold: amino acids of the alanine stereochemical type, i.e. the non-etalon amino acids; bold: etalon amino acids of glycine stereochemical type (G), proline stereochemical type (P) and valine stereochemical type (V, I). About etalon and non-etalon amino acids see in: Rakočević & Jokić (1996)." Here we designated four CIPS amino acid classes in four colors.

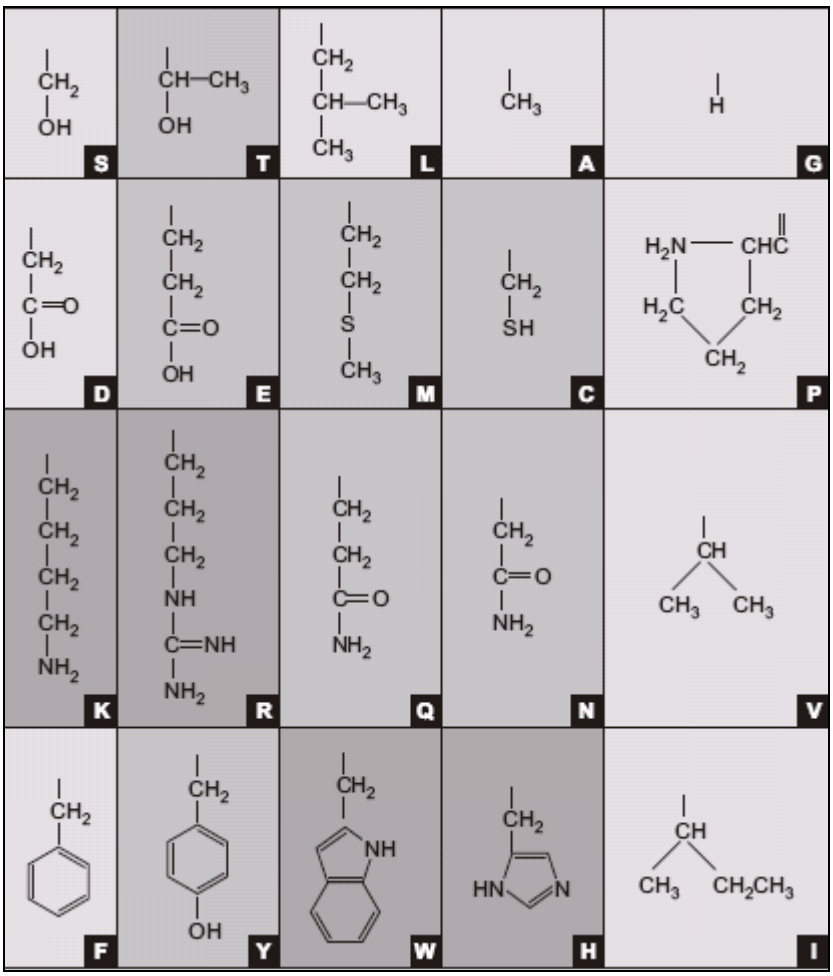


Table D.6.1. The arrangement AAs in accordance to Gauss' algorithm (Rakočević, 2006, Table 1.2, p. 6): "The structure of amino acid molecules. The simplest amino acid (G) … It is followed by alanine (A) whose side chain is only one CH_3 group … There are total of 16 amino acids of alaninic stereochemical type with one CH_2 group each between the "body" and the "head". The glycinic type contains glycine (G) only; valinic type contains valine and isoleucine (V, I); The last stereochemical type is prolinic with proline (P) which represents the inversion of valine in the sense that the "triangle" of three CH2 groups for the "head" is not bound by the basis, therefore not only with one but with two CH2 groups (Popov, 1989; Rakočević & Jokić, 1996). Light tones (G, P, V, I & A, L, S, D, F): invariant AAs; most dark tones (K, R, W, H): most variant AAs; less dark tones (T, E, Q, M, N, C): less variant AAs."

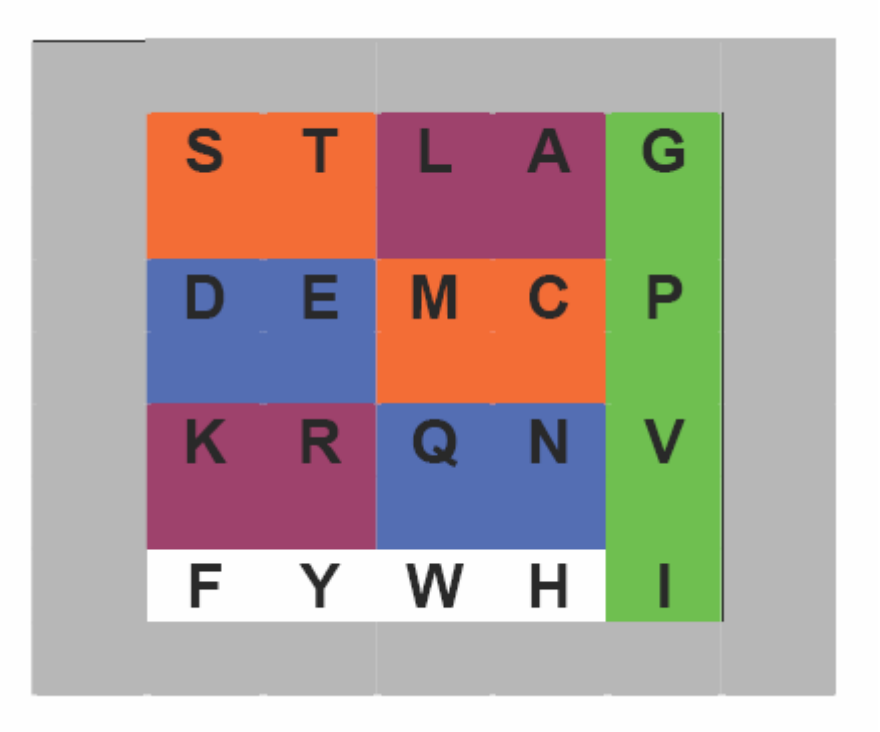


Table D.6.2. This Table is the same one as previous, but without formulas and with a new classification in accordance to the CIPS: in orange color the chalcogene AAs (S, T & C, M), blue – double acidic and their amide derivatives (D, E & N, Q); as pink the two original aliphatic AAs with two amine derivatives (A, L & K, R), green are the "contact" AAs (G, P & V, I); and, finely, four aromatic AAs (F,Y & H, W).

Genetic Code: The unity of the stereochemical determinism and pure chance

Miloje M. Rakočević

Faculty of Science, University of Niš (now retired, on the Address: Milutina Milankovica 118, 11070 Belgrade, Serbia (e-mail: m.m.r@eunet.yu; or: milemirkov@nadlanu.com; www.rakocevcode.rs)

Abstract

It is presented that the positions of amino acids within Genetic Code Table follow from strict their physical and chemical properties as well as from a pure formal determination by the Golden mean.

In this short paper the main result of our previous work (Rakočević, 2009) is presented in a new light; in the light of Crick's two hypotheses, put forward more than forty years ago (Crick, 1968). Crick asked then: Is the genetic code the result of strict stereochemical determinism, or of pure chance? Since those times many different answers have replaced one another. However, we can say that the first hypothesis has prevailed, though occasionally there still appears the second one as a possibility. [Crick himself left this question as an open question for further research.]

Our main result from the previous work shows that there follows from the Genetic Code Table (GCT) (Table 1) and the corresponding binary-code-tree (Figure 1 in relation to Table 2) a strict classification of amino acids (AAs) into five purely chemical classes (Table 3). And that is the said strict chemical determinism; moreover, that determinism is *stereochemical* precisely because the organization of amino acids into pairs must take into account the fact that the nature of the canonical twenty amino acids is such that they necessarily exist in the form of four stereochemical types (Popov, 1989; Rakočević and Jokić, 1996);[1] glycine within the glycinic type, proline within the prolinic, valine and isoleucine within the valinic, and the remaining 16 amino acids within the alaninic stereochemical type – the 12 aliphatic and four aromatic amino acids. [*Remark* 1: "The same four stereochemical types, two and two (G, A and V, P) ... also come

[1] It is simply amazing that, in genetic code research, no one else but us has referred to this work by Popov.

from the amino acid constitution structures, in the following manner. The side chain of glycine (-H) comes from the shortest possible hydrogen chain (H-H) and none of 19 other amino acids has a hydrogen chain of this kind. The side chain of alanine ($-CH_3$), or, in relation to glycine, ($-CH_2-H$), follows from the shortest possible non-cyclic hydrocarbon chain (CH_4), and still 15 amino acids have the alanine-analogue side chain in the form ($-CH_2-R$). The side chain of valine ($H_3C-CH-CH_3$) follows from the shortest possible cyclic hydrocarbon, from cyclopropane, with a permanent openness and with a linkage to the "head" of the amino acid through only one vertex of the cyclopropane "triangle". ... The proline type (only with proline) follows from the same source (cyclopropane), but with a permanent non-openness and with a linkage to the "head" through two vertices of the cyclopropane "triangle"" (Rakočević and Jokić, 1996, p. 345)[2].]

Besides, one step before the system "5 x 4" there is a system ["(2x7)"+"(1x6)"] (Survey 1) with seven "golden" AAs, the order of which follows from the strict golden mean determination; a determination which, in itself, is not physical, or chemical, but of a purely formal nature. It follows from this, as direct evidence, that it is the question of pure chance[3] whether or not (in the prebiotic states of matter) the conditions can be created for the appearance of an adequate molecular aggregation; adequate with respect to golden mean determination[4].

But, let's look once again at what follows from what. First we look at Survey 1 in relation to Table 2. Exactly seven AAs are on the positions of the golden mean, ϕ^n (n = 0, 1, 2, 3, ... , 9). They are G, Q, T, P, S, L, F. Across from them there are seven of their pairing "partners" (within the pairs), as the complements, and below are the six remaining AAs, the three pairs of the noncomplements: D-E, K-R, H-W.

---

[2] Giving to all this the presence of the first possible branching (coming from isobutane) in leucine and isoleucine, one can say that the genetic code is the result of the first possible cases: the first possible openness, ciclicity, half-ciclicity and branching. If so, if these extreme cases represent the condition for the origin of the code, then the possibility of origin of the life is significantly reduced anywhere in the universe.

[3] In this moment I bear in mind the famous book written by Jacques Monod (Chance and necessity, Vintage books, ISBN 0-394-71825-9) in which on the page V is cited the great Democritus of Abdera, living in the 4th century BC: "Everything existing in the Universe is the fruit of chance and of necessity".

[4] This question has also an exobiological aspect. Namely, if the determination by golden mean (the correspondence with the best possible symmetry, proportion and harmony) is the condition for the origin of the code, then it significantly reduces the possibility of origin of the life anywhere in the universe.

[And it makes sense that aliphatic AAs come first (D-E/K-R), followed by aromatic ones (H-W), all of them in the order determined by the size of the molecules, i.e. by the number of atoms within the molecule (side chain).]

In order to see all this, one must previously establish the order of codons in the GCT. Immediately it is obvious that the codon UUU is the zeroth one, and the codon CUG the seventh (8^{th} in the series of codons). After that, the problem is whether AUU or UCU is the eighth codon (9^{th} in the series)? In other words, what type of the base precedes in the chemical hierarchy: purine-pyrimidine-any (Pu-Py-X) or pyrimidine-pyrimidine-any (Py-Py-X)? Over ten years ago, I showed (Rakočević, 1998) that the determination by the golden mean coincides with the chemical order in which the Py-Py-X codons must precede the Pu-Py-X ones (as it is presented in Figure 1 in correspondence to Table 1). [*Remark* 2: The presented order of codons is based on the codon octets, but for a better viewing how the system, presented in Table 3, follows from GCT, it is also necessarily to have an order based on the codon quartets, such as it is presented in Table 4 (Negadi, 2009). Also to have a very symmetric form of GCT, valid for mitochondrial genetic code (Dragovich & Dragovich, 2006, 2007a, 2007b]

So, from the codon order follows the order of AAs, an order where seven of them are found at the positions of the golden mean. But, if the sequence of the seven AAs was turned through 180 degrees, and then the arrangement was made by amino acid molecule size, first glycine with one atom in the side chain, serine with 5, threonine with 8 atoms (the pair T 08 / M 11 must be before the pair of P 08 / I 13) etc., we would get the order as in Table 3: G, S, T, P, Q, L, F. We actually got a strict system that we in the previous paper (Rakočević, 2009) called CIPS (Cyclic Invariant Periodic System)[5].

The following analysis shows that the order of five of the amino acid classes makes sense. On the position "1" came AAs of non-alaninic type, which are aliphatic by origin: G-P and V-I; followed (on the position "2") by aliphatic AAs of the alaninic type (A-L and K-R), two of which are amine derivatives (K-R), with a lower degree of polarity (nitrogen is less polar then oxygen!)

From pure chemical reasons it makes sense to say that these two classes (light tones in Table 3) belong to *a primary superclass*, with original aliphatic AAs (and/or derivatives of lower level), whereas the three remaining classes (dark tones) to *a secondary superclass*, with the derivatives of a higher level.

[5] Cyclicity and periodicity through the positions of two and two amino acids – up/down – in relation to middle chalcogene AAs, on the position "3".

On the position "3" within the system, presented in Table 3, came chalcogene AAs (S, T & C, M); followed (on the position "4") by two double acidic AAs with two their amide derivatives (D, E & N, Q); finally came four aromatic AAs (F,Y & H, W) on the position "5".

The splitting into two superclasses exists in a strict correspondence to the splitting into two cathalitic directed classes (amino acids handled by class I or by class II of aminoacyl-tRNA synthetases), as it is shown in Table 5, in relation to Survey 2. From the said correspondence can follow *a prediction* for further researches: the demonstrated crossing of two classes and two superclasses must be reflected (in some way?) in the protein structures and functions.

As it is immediately obvious from CIPS[6], the nature of the genetic code again points out the validity of Aristotle sentence[7], and, on the other hand, necessarily leads to the conclusion that both Crick's hypotheses about the origin of the genetic code, *mutatis mutandis*, are valid.

[6] Which system is the "fruit" of a unity of strict stereochemical determinism and pure chance, a unity of physical–chemical characteristics and pure formalism (cf. footnote 3).

[7] "The existence of such a harmonic structure with unity of a determination with physical–chemical characteristics and atom and nucleon number ... appealed to Aristotle and to his idea of unity of form and essence" (Rakočević, 2004, p. 233).

1st lett.	2nd letter								3rd lett.
	U		C		A		G		
U	**00.** UUU	**F**	**08.** UCU	**S**	**32.** UAU	**Y**	**40.** UGU	**C**	*U*
	01. UUC		**09.** UCC		**33.** UAC		**41.** UGC		*C*
	02. UUA		**10.** UCA		**34.** UAA		**42.** UGA		*A*
	03. UUG	**L**	**11.** UCG		**35.** UAG	CT	**43.** UGG	CT **W**	*G*
C	**04.** CUU		**12.** CCU		**36.** CAU	**H**	**44.** CGU		*U*
	05. CUC		**13.** CCC		**37.** CAC		**45.** CGC		*C*
	06. CUA	**L**	**14.** CCA	**P**	**38.** CAA		**46.** CGA	**R**	*A*
	07. CUG		**15.** CCG		**39.** CAG	**Q**	**47.** CGG		*G*
A	**16.** AUU		**24.** ACU		**48.** AAU	**N**	**56.** AGU	**S**	*U*
	17. AUC	**I**	**25.** ACC		**49.** AAC		**57.** AGC		*C*
	18. AUA		**26.** ACA	**T**	**50.** AAA		**58.** AGA		*A*
	19. AUG	**M**	**27.** ACG		**51.** AAG	**K**	**59.** AGG	**R**	*G*
G	**20.** GUU		**28.** GCU		**52.** GAU	**D**	**60.** GGU		*U*
	21. GUC	**V**	**29.** GCC	**A**	**53.** GAC		**61.** GGC	**G**	*C*
	22. GUA		**30.** GCA		**54.** GAA	**E**	**62.** GGA		*A*
	23. GUG		**31.** GCG		**55.** GAG		**63.** GGG		*G*

Table 1. The Table of the standard genetic code. Ordinal number of codons after the order-key: YYN, RYN, YRN, RRN, in correspondence with the hierarchy on the binary-code tree in Figure 1. [One-letter abbreviations: Y from pYrimidine; R from puRine; and N from aNy (of bases).]

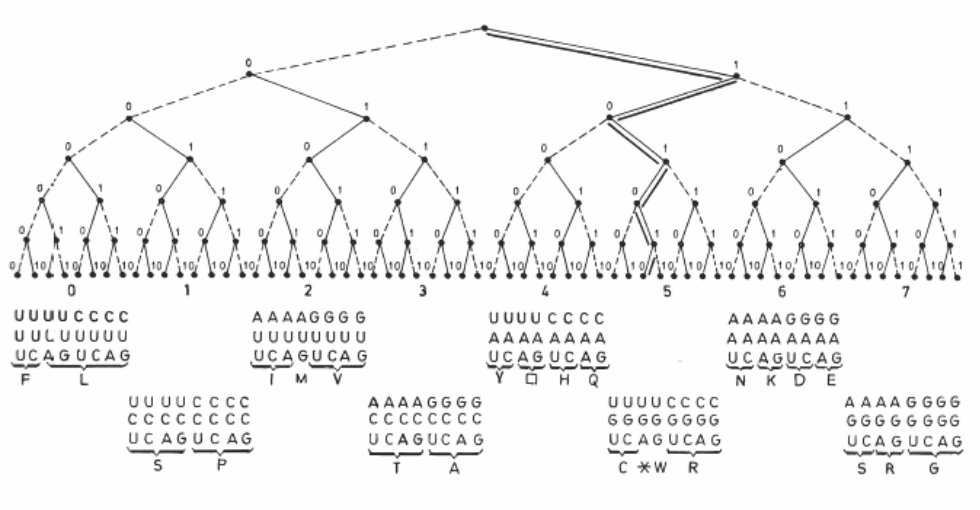


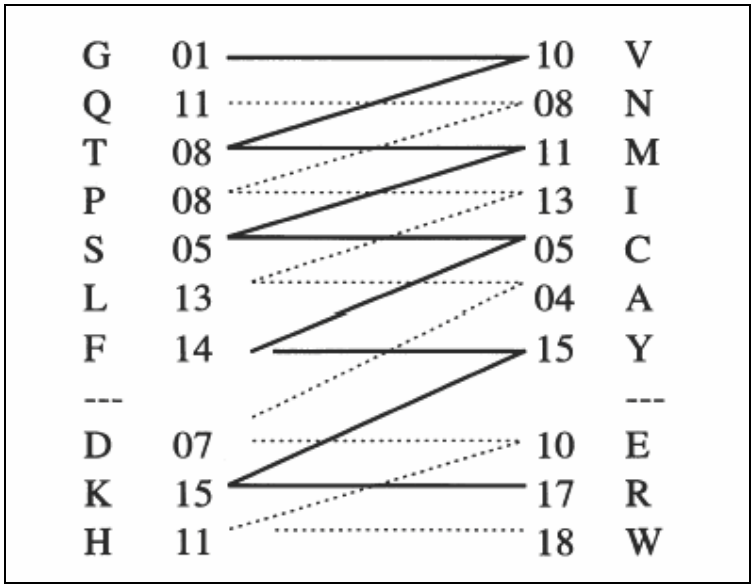
Figure. 1. Genetic code as a binary-code tree. The full lines: the routes of the greater changes, from 0 to 1 and vice versa; the dotted lines: the routes of the less changes, from 0 to 0, as well as from 1 to 1 going from a higher into a lower level. The double full line: the route of the maximum possible changes: from 0 to 1 and vice versa in any step (the route corresponding to the 'Golden mean route' on the Farey tree). The codon order after the rules given by R. Swanson (Swanson, 1984, Figure 1): 1 for purine and 0 for pyrimidine; 1 for three and 0 for two hydrogen bonds.

ϕ^0	ϕ^1	ϕ^2	ϕ^3	ϕ^4	$\phi^{5\text{-}7}$	ϕ^8	ϕ^9
G	Q	T	P	S	L	L	F
63	39-38	25-24	15-14	10-09	06-02	02-01	01-00
63	38.94	24.06	14.87	9.19	5.68 – 2.17	1.34	0.83

Table 2. The amino acids in Golden mean power positions within the sequence 0–63 on the binary-code tree in Fig. 1. First row: Golden mean powers within first 'cycle' in module 9. Second row: amino acids in the positions marked in third row, taken from the binary-code tree in Fig. 1. Fourth row: the values of the Golden mean powers within the interval 0–63. The calculations: 0.618033 x 63 = 38.94; 0.618033 x 0.618033 x 63 = 24.06 etc.

5	073	F	14	15	Y	079
4	235	L	13	04	A	172
3	087	Q	11	08	N	085
2	160	P	08	13	I	121
1	168	T	08	11	M	043
1	243	S	05	05	C	081
2	184	G	01	10	V	168
3	087	D	07	10	E	093
4	091	K	15	17	R	265
5	081	H	11	18	W	044

Table 3. The Cyclic Invariant Periodic System (CIPS) of canonical AAs. At the outer side, left and right, it is designated the number of atoms within coding codons; more exactly, in the Py-Pu bases (U = 12, C = 13, A = 15 and G = 16); at the inner side – the atom number within amino acid side chains. In the middle position there are chalcogene AAs (S, T & C, M); follow - in next „cycle“ - the AAs of non-alaninic stereochemical types (G, P & V, I), then two double acidic AAs with two their amide derivatives (D, E & N, Q), the two original aliphatic AAs with two amine derivatives (A, L & K, R); and, finely, four aromatic AAs (F,Y & H, W) – two up and two down. The said five classes belong to two superclasses: primary superclass in light areas and secondary superclass in dark areas. Notice that each amino acid position in this CIPS is strictly determined and none can be changed. [For details see the text; for arithmetical regularities, atom and nucleon number balances within this CIP system see in our previous paper (Rakočević, 2009).]



Survey 1. Atom number balance directed by Golden mean on the binary-code tree (Scheme 2 in Rakočević, 1998; Table 3 in Rakočević, 2009). First seven amino acids on the left are 'golden' amino acids with 60 atoms within side chains; on the right are their complements with [60 + (1 x 6)] = 66 atoms; below are three amino acid pairs as non-complements with [66 + (2 x 6)] = 2 x 39 = 78 atoms; Notice that within aliphatic non-complements there are 39+10, whereas within aromatics (H & W) 39-10 of atoms; in the other words: (8 x 6) +1 and (5 x 6) -1, respectively. On the first zigzag (full) line there is 102-1 whereas on the second (dotted) line 102+1 atoms. Arithmetic mean for both: 102±1. Notice also that the arrangement-ordering of "golden" AAs is the same as in Table 2.

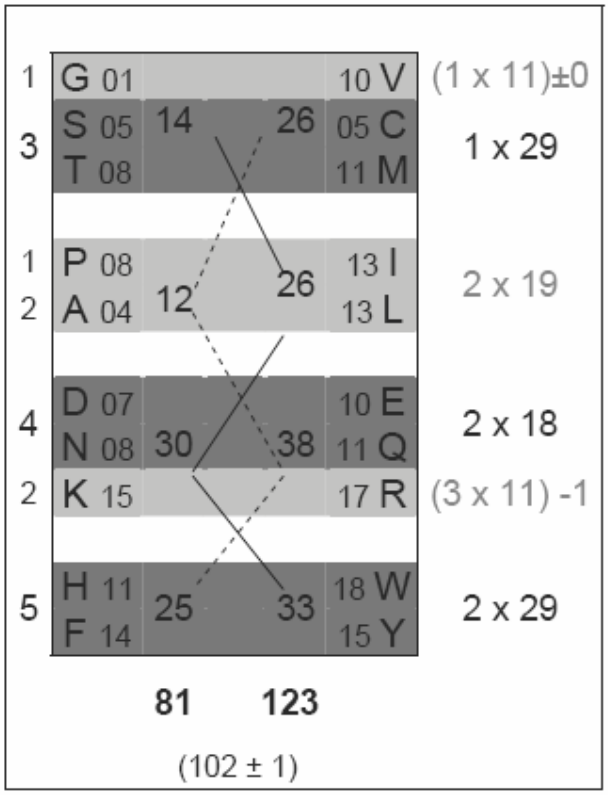
U-U-X	U-C-X	C-U-X	C-C-X
U-A-X	U-G-X	C-A-X	C-G-X
A-U-X	A-C-X	G-U-X	G-C-X
A-A-X	A-G-X	G-A-X	G-G-X

Table 4 . The codon order in Genetic Code Table, based on the codon quartets in first and second half of the Table. [After Table 1 in (Negadi, 2009), here simplified and generalized.]

28	09 19	G **P** A **K**	(1) (2)	23 30	V **I** L **R**	53	**81**
53	13 15 25	S **T** D **N** **F** H	(3) (4) (5)	16 21 33	C **M** E **Q** Y **W**	70	**123**
81						**123**	204

Table 5. The arrangement in accordance to the principle: "a little" and "full" in relation to "small" and "large". So, on the left there are AAs (81 atoms) from the left side of Survey 2 (class II, with smaller molecules within the pairs); and on the right there are AAs (123 atoms) from the right side of Survey 2 (class I, with larger molecules within the pairs). At

the same time very up there are AAs from primary superclass (81 atoms), just aliphatic and nonpolar (A,V, L, I) and "a little" polar (G, P, K, R) (hydrogen and nitrogen are less polar then oxygen!); in the other hand, except aromatic and sulfur AAs, down are AAs from secondary superclass (the row with 123 atoms), also aliphatic, but "full" polar. (This Table is the same as Table D.2 in Rakočević, 2009.)



Survey 2. Two amino acid classes generated through the influence of two catalysts (Scheme 5 in Rakočević, 1998 and Table D.1 in Rakočević, 2009). On the left the smaller molecules of AAs, handled by class II of aminoacyl-tRNA synthetases, whereas on the right the larger molecules of AAs, handled by class I of synthetases. On the full line there are 102+1 and on the dotted one 102-1 of atoms. The distinct arrangement of five classes (and two superclasses) from CIPS (Table 3) is self-evident. By this one must notice the existence of a crossing valid only for the primary superclass; the crossing in relation to original system in Table 3. (Why P together with A, L, I one can read in Remark 1.)

Acknowledgement

I am grateful to Branko Dragovich and Vladimir Ajdačić for helpful, stimulating discussion.

REFERENCES

Crick, F.H.C (1968) The origin of the genetic code, J. Mol. Biol., 38, 367-379.

Dragovich B., Dragovich, A. (2006) A p-Adic model of DNA sequence and genetic code, p-Adic Numbers, Ultrametric Analysis and Applications 1, 34-41; arXiv:q-bio/0607018v1 [q-bio.GN].

Dragovich B., Dragovich, A. (2007a) p-Adic Degeneracy of the Genetic Code. arXiv:0707.0764v1 [q-bio.GN].

Dragovich, B., Dragovich, A. (2007b) p-Adic modeling of the genome and the genetic code, Computer Journal, doi:10.1093/comjnl/bxm083; arXiv:0707.3043v1 [q-bio.OT].

Negadi, T. (2009) The genetic code degeneracy and the amino acids chemical composition are connected, Neuroquantology, Vol. 7, 1, 181-187; arXiv:0903.4131v1 [q-bio.OT]

Popov, E. M. (1989) Strukturnaya organizaciya belkov, Nauka, Moscow.

Rakočević, M. M. (1998) The genetic code as a Golden mean determined system, Biosystems 46, 283-291.

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

Rakočević, M. M. (2009) Genetic Code Table: A note on the three splittings into amino acid classes, arXiv:0903.4110v1 [q-bio.BM].

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

Swanson, R. (1984) A unifying concept for the amino acid code, Bull. Math. Biol., 46, 187-207.

GENETIC CODE: FOUR DIVERSITY TYPES OF PROTEIN AMINO ACIDS

Miloje M. Rakočević*

Department of Chemistry, Faculty of Science, University of Niš, Serbia

Abstract

This paper presents, for the first time, four diversity types of protein amino acids. The first type includes two amino acids (G, P), both without standard hydrocarbon side chains; the second one four amino acids, as two pairs [(A, L), (V, I)], all with standard hydrocarbon side chains; the third type comprises the six amino acids, as three pairs [(F, Y), (H, W), (C, M)], two aromatic, two hetero aromatic and two "hetero" non-aromatic); finally, the fourth type consists of eight amino acids, as four pairs [(S, T), (D, E), (N, Q), (K, R)], all with a functional group which also exists in amino acid functional group (wholly presented: H_2N-ĊH-COOH; separately: OH, COOH, $CONH_2$, NH_2). The insight into existence of four types of diversity was possible only after an insight into the existence of some very new arithmetical regularities, which were so far unknown. Also, as for showing these four types was necessary to reveal the relationships between several key harmonic structures of the genetic code (which we presented in our previous works), this paper is also a review article of the author's researches of the genetic code. By this, the review itself shows that the said harmonic structures are connected through the same (or near the same) chemically determined amino acid pairs, 10 pairs out of the 190 possible.

---

* Now retired, on the address: Milutina Milankovica 118/ 25, 11070 Belgrade, Serbia; e-mail: milemirkov@open.telekom.rs; website: (http://www.rakocevcode.rs).

1. INTRODUCTION

Eighteen years ago V. Shcherbak pointed out that the classifications and systematizations of the genetic code constituents (protein, i.e. canonical amino acids, AAs) "appears to be loaded with elegant arithmetical regularities, which are so far unknown", and also that "the physical nature of such a phenomenon is so far not clear" (Shcherbak, 1993).

By this, the mentioned classifications-systematizations and the arithmetical regularities in which "the amino acids appear to be divided into two parts" are realized through the number of nucleons just in these two parts (Shcherbak, 1993, 1994). Surprisingly, since then no attempt has been made to answer two questions which follow from Shcherbak's finding: 1. how to explain such regularities, which had never before been known to mathematics; and in particular, how to explain their presence in the genetic code; and 2. what is the physical (and chemical) nature of the presence of these regularities in the genetic code?

Hence, this paper is an attempt to make the first step toward the search of correct answers to the above questions, if not completely, then to present other possible arithmetical regularities, as well as other possible classifications-systematizations of AAs, which would take us, *per se*, at least closer to an adequate response.

2. PRELIMINARIES

We start from Shcherbak's basic Table (here: Table 1)[1], given here with our shading, which indicates that the determination of the number of nucleons involves only one column (the seventh) and only one row [a part of the first row: the numbers with the same digits (111, 222 and 333)]; in other words, the determination involves only the digital patterns of the

[1] Cf. Table 1 in this paper with Shcherbak's Tables 1 and 2, in relation to Figure 1 in his source work (Shcherbak, 1994).

nucleon number notations that are mediated by the Pythagorean triplet, one and only existed within natural numbers series: 3-4-5.

In fact, only the number of nucleons in four-codon AAs ("Group IV" in Shcherbak's labelling) is directly determined by the Pythagorean triplet: 3^2 x 037 = 333, 4^2 x 037 = 592, 5^2 x 037 = 925. However, the fact that the numerical notation "333" participates in this Pythagorean determination is sufficient reason to conclude that the determination of the number of nucleons in the non-four-codon AAs ("quasi group III-II-I" in Shcherbak's labelling) is also a Pythagorean determination; albeit indirect, through a connection between notations "111", "222" and "333" (cf. Figure 1 in Shcherbak, 1994).

1	2	3	4	5	6	7	8	9
037	074	111	148	185	222	259	296	333
370	407	444	481	518	555	592	629	666
703	740	777	814	851	888	925	962	999

111	259
222	592
333	925

Table 1. Shcherbak's Table of multiples of number 037, a little simplified and with our appendix on the right. (Table 1 in Shcherbak, 1994.) For details see the text.

When it comes to the Pythagorean triplet, in the said context, it is important to know that the Pythagorean triangle 5-4-3 is actually the first in the series of the so-called Diophantus' triangles (not counting the zeroth triangle, which is not really a triangle because it lacks one straight line). It should also be noted that one cathetus of any Diophantus' triangles belongs to the series of odd integers, and the other is generated from the series whose first term is zero, and the following members are respectively increased by 4n (n = 1, 2, 3, 4, ...); the same applies to the hypotenuse, but

it generates from the series whose initial member is number 1 (Figure A1 on the left, in Appendix A).

Finally, the Pythagorean 5-4-3 triangle belongs to the series of triangles that begins with Luca Pacioli's triangle. The series itself results from a specific generalization of the golden mean (Figure A1, on the right, in relation to Table A1, also on the right).

Having in mind all these regularities, valid for the Pythagorean triplet 3-4-5 it is difficult to accept that the determination of the number of nucleons in the genetic code, just through the Pythagorean triplet, is only a mere coincidence; or, moreover, that the referring to the Pythagorean triplet is a scholastic approach, a numerology and nothing more. On the contrary, these regularities may indicate that it makes sense *a working hypothesis (I)*, according to which in determination of the genetic code, except two inherent alphabets – 20 amino acids and four amino bases (two pyrimidines & two purines) – is involved still one "hidden alphabet", a series of natural numbers, with all its regularities and laws.

Besides starting from Shcherbak's work mentioned above, in this paper we also start from Sukhodolets' basic Table, presented 26 years ago (Sukhodolets, 1985), which demonstrated that the number of hydrogen atoms ("hydrogen" protons) in protein AAs is also accompanied by specific arithmetical regularities [Table 2 in relation to Table A2 in Appendix A; then Table 3 in relation to Tables 3.1 and 3.2, including Solutions (1) and (2) which explained in legend to Table 3.1].

$$11\underline{7} + (20 \times 4) = 19\underline{7}$$
$$08\underline{7} + (20 \times 5) = 18\underline{7} \qquad (1)$$

$1 \rightarrow 38 = 741$ (741 : 2 = 370.5)
$1 \rightarrow \mathbf{39} = 780$ (780 : 2 = **39**0) (2)
$1 \rightarrow 40 = 820$ (820 : 2 = 410)

n	Amino acids	Codon root	
		First letter	Second letter
5	Gly	G	G
7	Ala Ser Asp cys	G U | G U	C C | A G
8	Asn	A	A
9	Pro Thr Glu His	C A | G C	C C | A A
10	Gln	C	A
11	Val Phe Met Tyr	G U | A U	U U | U A
12	Trp	U	G
13-14	Leu Ile Arg Lys	C A | C A	U U | G A

Table 2. The Sukhodolets' Table: the number of hydrogen atoms (n) within amino acid molecules, in relation to natural numbers series: 5, (6), 7, 8, 9, 10, 11, 12, 13, 14 (Sukhodolets, 1985) (Cf. Table A2 in App. A). First letter plus second letter equals a codon root. The codon root plus third letter equals a complet codon. Within 64 codons (192 nucleotides) there are 2 x 3456 atoms; 3456 in two inner and 3456 in two outer columns of standard Genetic Code Table (GCT), all calculated as in Figures A2 and A3.

The number of H atoms (in brackets) and conformations						
G (**05**) (04)	A (**07**) (03)	S (**07**) (09)	D (**07**) (10)	C (**07**)(21)	153	
N (**08**) (16)	P (**09**)(02)	T (**09**) (08)	E (**09**) (20)	H (**09**) (24)	298	
Q (**10**) (38)	V (**11**) (08)	F (**11**) (12)	M (**11**) (20)	Y (**11**) (12)	388	569/686
W (**12**) (24)	R (**14**) (66)	K (**14**) (66)	I (**13**) (20)	L (**13**) (22)	416	Nucleon number
GW + AC + PH + VY + RL = 58H+ 44 nonH = 102 (622 nucl.) Out						
NQ + SD + TE + FM + KI = 59H+ 43 nonH = 102 (633 nucl.) (In)						
The number of conformations:						
GW28 + **AC24** + PH26 + **VY20** + RL88 = 142 / **44** (Outer)						203 –10
NQ54 + SD19 + **TE28** + FM32 + **KI86** = **168** / 51 (Inner)						**202 +10**
Dark / Light: (210 / 195) // (210 + 100) / (195 - 100) Odd / Even						
Dark: (41+1) x 5 = 210; Light: (40-1) x 5 = 195						

Table 3. The Sukhodolets' Table, with a minimal modification. Above: the system of 4 x 5 AAs. The shadow space: AAs with even number of hydrogen atoms (8, 10, 12, 14); the non-shadow space: AAs with odd number of hydrogen atoms (5, 7, 9, 11, 13). In brackets: number of hydrogen atoms – bold (after Sukhodolets, 1985) and then the number of conformations – no bold. [Number of conformations after (Popov, 1989); the calculations for the number of conformations are given in three last rows, and also in fourth, the additional row.] Nucleon number through a specific "simulation": 569 within two outer rows (as the number of neutrons, 569, in all 20 AAs: within their side chains); and 686 nucleons within two inner rows (as the number of protons, 686, in all 20 AAs: within their side chains). Nucleon number through a specific balance: 622/633 within outer/inner amino acid pairs, respectively (cf. with result 622/633 in Table 6.1; then with result 633+11 / 622 – 11 in Tables 6.1 & 7; and with result 633+10 / 622 – 10 in Fig. 1). All others as it is written.

(a)	G (05)	A (07)	S (07)	D (07)	C (07)	33	13
	N (08)	P (09)	T (09)	E (09)	H (09)	44	24
	Q (10)	V (11)	F (11)	M (11)	Y (11)	54	34
	W (12)	R (14)	K (14)	I (13)	L (13)	66	46
	17/18	21/20	21/20	20/20	20/20	**98**	**58**
	35	41	41	40	40	**99**	**59**
			81/81				
(b)	G (05)	A (06)	S (07)	D (09)	C (07)	34	09
	N (09)	P (08)	T (08)	E (10)	H (11)	46	21
	Q (10)	V (08)	F (12)	M (09)	Y (13)	52	27
	W (15)	R (12)	K (10)	I (09)	L (09)	55	30
	20/19	18/16	17/20	18/19	16/24	**98**	**48**
	39	34	37	37	40	**89**	**39**
			74/74				
(c)	G (05)	A (06)	S (07)	D (09)	C (07)		
	N (09)	P (08)	T (08)	E (10)	H (11)	**88**	**38**
	Q (10)	V (08)	F (12)	M (09)	Y (13)	99	49
	W (15)	R (12)	K (10)	I (09)	L (09)		

Table 3.1. Amino acid arrangement as in modified Sukhodolets' Table (as in upper part of Table 3) with hydrogen atom number in (a) and non-hydrogen atom number in (b & c). The last column is valid for the amino acid side chains, and first to last for whole molecules. Total number of hydrogen as well as non-hydrogen atoms: 58 + 59 = 117 / 48 + 39 = 087 in side chains and 98 + 99 = 197 / 98 + 89 = 187 in whole molecules, respectively [cf. Solution (1)]. The sums of non-hydrogen atoms in two and two rows in last column of (b) are 39 and 39+9, respectively (the change for the number "9" means to make a cycle more, through module q-1 in decimal numbering system). [The uniqueness of the number 39 (through the sum from 1 to 39) is presented in Solution (2).] All other balances are self-evident.

(a)	G (02)	A (03)	S (03)	D (04)	C (03)	**15**	**05**
	N (04)	P (05)	T (04)	E (05)	H (06)	24	14
	Q (05)	V (05)	F (09)	M (05)	Y (09)	**33**	**23**
	W (11)	R (06)	K (06)	I (06)	L (06)	35	25
	13/09	09/10	09/13	10/10	09/15	**48**	**28**
	22	19	22	20	24	59	39
			42/43				
(b)	G (**02**)	A (**03**)	S (**03**)	D (**04**)	C (**03**)		
	N (**04**)	P (**05**)	T (**04**)	E (**05**)	H (**06**)	**54** / 53	
	Q (**05**)	V (**05**)	F (**09**)	M (**05**)	Y (**09**)	**34** / 33	
	W (**11**)	R (**06**)	K (**06**)	I (**06**)	L (**06**)		

Table 3.2. The amino acid arrangement as in the modified Sukhodolets' Table (as in upper part of Table 3) with the number of carbon atoms; (a) The balances in two odd and two even rows; within whole molecules: 59 – 48 = 11 and within their side chains: 39 – 28 = 11; (b) The balances between two halves of the system; within whole molecules: 54 – 53 = 1 and within their side chains: 34 – 33 = 1.

3. NEW ARITHMETICAL REGULARITIES

Among the arithmetical regularities valid for number 037 in the decimal numbering system, Shcherbak has shown that this number is also present in the system of numbers which are analogs of 037 in some other numbering systems; their values for q (numbering system basis), going from one to another, differ by three units. Written together without specifying the numbering basis, these numbers-analogs (Shcherbaks'

"Prime Quantums", PQ) are as follows: 13, 25, 37, 49, ... As we can see, the first digit belongs to the series of natural numbers, and the second to the series of odd integers.[2]

Knowing this, it is easy to see the possibility of a specific presentation of a series of natural numbers, in the form of a matrix scheme (n x 11_q) (Scheme 1 in Table 4 and Scheme 2 in Table 5); such a scheme along whose first diagonal appear numbers in the form of Shcherbak's numbers, analogs of number 37 (N = 13, 25, 37, 49, 5B, ...).[3]

The system of numbers in the two presented schemes (Tables 4 and 5) is a totally regulated system on several grounds; primarily by the logic of close proximity, and in accordance with two important principles: the principle of minimum change (unit change in one or both digit positions)[4] and the principle of continuity. Going from left to right, and vice versa, the change is by 01; along the second diagonal[5] by 10 (after a start with number q-1); moving vertically: by 11; and along the first diagonal (containing Shcherbaks' numbers and starting with number 01), the changes are by 12.

When it comes to changes to the vertical, then, besides being different by 11, the numbers form ordered tuples of odd and even natural numbers

---

[2] "In the close vicinity of the decimal system, for example, some number systems with the bases 4 (Quantum $13_4 = 7$), 7 ($25_7 = 19$), 10 (37), 13 ($49_{13} = 61$), etc, have the same periodic features for three-digit numbers" (Shcherbak, 1994).

[3] For N = 13, q = 1+3 = 4; for N = 25, q = 2+5 = 7; for N = 37, q = 3+7 = 10 etc. These are the real analogues, in contrast to the "simulation" analogs in the first diagonal of the arrangement of natural numbers in Tables 4 & 5. There are, in fact, the numbers that possess the form of Shcherbak's numbers, not for various, but for the one and the same numerical basis, *q*. (For example in Table 4, q = 10 and in Table 5, q = 16.) Such a simulation means a specific self-similarity at the same time.

[4] We say "in both positions" because we are here dealing with two-digit numbers, which (some of them) turn out to be the key determinants of the genetic code. This means that the "diagonal law" mentioned above is also valid for two-digit numbers, while future research is to show if there is such a law for n-digit numbers (n > 2) as well.

[5] The second diagonal contains numbers the first digits of which are generated – respectively, out of the sequence of natural numbers (0, 1, 2, 3, ...), while the second digit is number q-1.

series, alternately and periodically, after modulo q-1, as it is shown in Solution (3):

01	02	03	04	…
03	04	05	06	
05	06	07	08	
07	08	09	10	
09	10	11	12	
11	12	13	14	
13	14	15	16	
15	16	17	18	
17	18	19	20	
...				

(3)

However, since the changes in the vertical are only by a unit in both positions (11_q), they are also related to the last column of the Scheme, consisting of notations 11_q, 22_q, 33_q, ... and so on. In the case of a decimal numbering system, the number 11_{10}, as such maintains the best connection to the Golden mean (Table B1 in relation to Table B2 in Appendix B).

Within decimal numbering system, as we know from Shcherbak's works (Shcherbak, 1993, 1994), the number 37 is characterized by specific and unique arithmetical regularities (Table 1). Here, however, we have shown that in Scheme 1 (Table 4) specific and unique arithmetical regularities characterize not only number 37 but also its neighbours, numbers 26 and 48, each in its own way.

3.1. The uniqueness of the number 26

Again we consider Scheme 1 (Table 4). If the first diagonal neighbour of the number 26, the number 16 (26 – 16 = 10) is added to number 26 and its two followers (17 and 18) are successively added to the obtained result, we get the results as in Solution (4.1):

26 = 26	26 + 42 + 59 + 77 = Y	**16** +**17** + **18** = Z	
26 + **16** = 42	Y = 204	Z = 51	(4.1)
42 + **17** = 59		Z = Y/4	
59 + **18** = 77	Y/4 = 51		

With three addings (16 +17 +18 = 51 = Z) we obtained three new results, and with the inclusion of the initial number 26 – four results. Their sum is 204 (26 + 42 + 59 + 77 = 204 = Y = 4Z), exactly four times greater than the sum of the three addings (16+17+18 = 51 = Z). But this connection of two equalities is a single and unique case in the entire system of numbers within Scheme 1 (Table 4), in other words within the set of natural numbers. Moreover, it appears that this is the zeroth case in the 4th column of Scheme 1 in the decimal numbering system (Table C1 in Appendix C), and also the zeroth case within all Shcherbak's numbering systems (q = 4, 7, **10**, 13, 16, …) (cf. Table C2).

3.1.1. The uniqueness of number 26 through unique pair 25-36

The uniqueness of the number 26 is expressed not only through the difference 26 - 10 = 16, but also through the sum 26 + 10 = 36, where number 36, as the second diagonal neighbour of the number 26, appears to be the member of a unique pair 25-36; unique case in the entire system of numbers within Scheme 1 (Table 4), and that means within the set of natural numbers. Namely, the numbers 25 and 36 are neighbours in third column of Table 4 with a difference of 11 as in all other cases, in all columns. But their square roots, as integers, appear to be also neighbours, and that is the said uniqueness [Solution (4.2) in relation to Table C3]:

$$\begin{aligned} x_1 + y_1 &= 36 = 6^2 \quad (x_1 = 26;\ y_1 = 10) \\ x_2 + y_2 &= 25 = 5^2 \quad (x_2 = 17;\ y_2 = 08) \\ x_1 - y_1 &= 16 = 4^2 \\ x_2 - y_2 &= 09 = 3^2 \end{aligned} \qquad (4.2)$$

3.2. The uniqueness of the number 48

Now we will demonstrate the uniqueness of number 48 within the system of numbers in Scheme 1 (Table 4) with one exercise: Find a number in the system of numbers of Scheme 1, in series of odd or even numbers, which with its three followers gives the same result (204) which we obtained with number 26, as it is shown above. It is immediately obvious that this is number 48, as it is shown in Solution (5):

48+50+52+54 = 204 (5)

	...										
(-2)	...									...	-22
(-1)	-21	-20	-19	-18	-17	-16	-15	-14	-13	-12	-11
(0)	-10	-09	-08	-07	-06	-05	-04	-03	-02	-01	00
(1)	01	02	03	**04**	05	06	07	08	09	10	11
(2)	12	13	14	15	**16**	**17**	**18**	19	20	21	22
(3)	23	24	25	**26**	27	28	29	30	31	32	33
(4)	34	35	36	37	38	39	40	41	42	43	44
(5)	45	46	47	48	49	50	51	52	53	54	55
(6)	56	57	58	59	60	61	62	63	64	65	66
(7)	67	68	69	70	71	72	73	74	75	76	77
(8)	78	79	80	81	82	83	84	85	86	87	88
(9)	89	90	91	92	93	94	95	96	97	98	99
(A)	A0	A1	A2	A3	A4	A5	A6	A7	A8	A9	AA
(B)	B1	B2	B3	B4	B5	B6	B7	B8	B9	BA	BB

Table 4 (Scheme 1). A specific arrangement of natural numbers in decimal numbering system with Shcherbak's "simulation" analogs (13, 25, 37, 49 ...) within the first diagonal (cf. footnote 3). [The numbers 3, 4 and 5 appear to be unique through Pythagorean law; the uniqueness of the number 78 one can see in Solution (2) and legend to Table 3.1]. For all other details see the text.

It is also immediately obvious that this is the middle row in the system of numbers (in decimal number system) of Scheme 1, in Table 4. Thus, it is inevitable that every two rows at the same distance from the middle row (in analogous positions) yield twice the value of number 204 [Example: (37 + 39 + 41 + 43) + (59 + 61 + 63 + 65) = 2 x 204).] In addition, in relation to the two end, and the two central numbers of the sequence 48-50-52-54, there is a symmetrical and cross-regularity (Table C4).

(1)	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	10	11
(2)	12	13	14	15	16	17	18	19	1A	1B	1C	1D	1E	1F	20	21	22
(3)	23	24	25	26	27	28	29	2A	2B	2C	2D	2E	2F	30	31	32	33
(4)	34	35	36	37	38	39	3A	3B	3C	3D	3E	3F	40	41	42	43	44
(5)	45	46	47	48	49	4A	4B	4C	4D	4E	4F	50	51	52	53	54	55
(6)	56	57	58	59	5A	5B	5C	5D	5E	5F	60	61	62	63	64	65	66
(7)	67	68	69	6A	6B	6C	6D	6E	6F	70	71	72	73	74	75	76	77
(8)	78	79	7A	7B	7C	7D	7E	7F	80	81	82	83	84	85	86	87	88
(9)	89	8A	8B	8C	8D	8E	8F	90	91	92	93	94	95	96	97	98	99
(A)	9A	9B	9C	9D	9E	9F	A0	A1	A2	A3	A4	A5	A6	A7	A8	A9	AA
(B)	AB	AC	AD	AE	AF	B0	B1	B2	B3	B4	B5	B6	B7	B8	B9	BA	BB

Table 5. (Scheme 2). A specific arrangement of natural numbers in hexadecimal system with Shcherbak's "simulation" analogs (13, 25, 37, 49 ...) within first diagonal (cf. footnote 3). For details see the text.

4. NEW CLASSIFICATIONS AND SYSTEMATIZATIONS

4.1. Genetic code as an "imitation" of arithmetical regularities

Through a purely chemical analysis it is possible to find such arrangements of protein amino acids that fully correspond to the observed arithmetical regularities, related to the uniqueness of numbers 26 and 48; as if we are dealing with an "intelligent" imitation, in one possible classification and systematization.

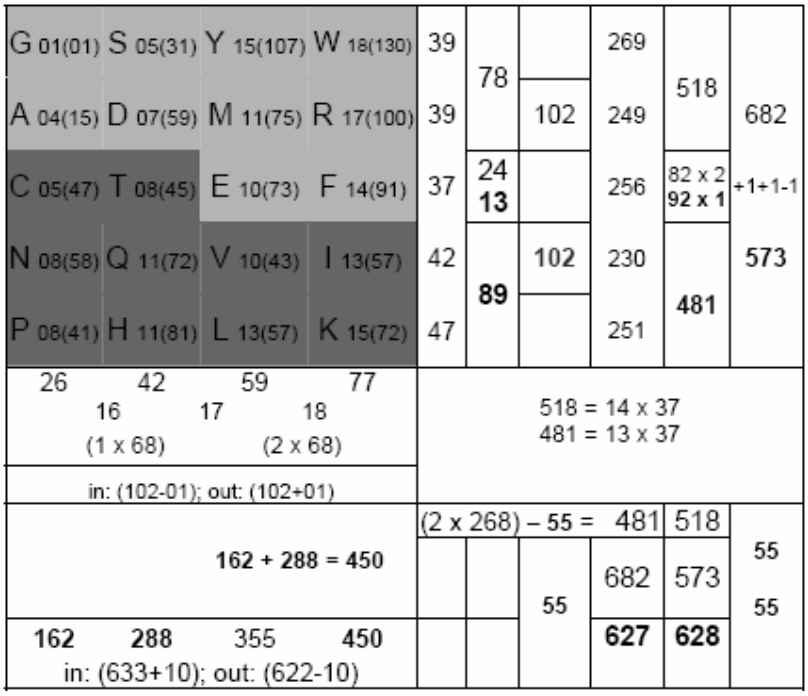


Figure 1. A specific classification and sistematizacion of amino acids, which follow from four diversity types (Figure 2). In the shadow space there are 20 AAs with nucleon number in the brackets and atom number without the brackets, both times in molecules side chains. Regarding at the columns: 26, 42, 59 and 77 atoms, the quantums appear to be as in Solution (4.1) and Table C1. Within first two and last two columns: 1 x 68 and 2 x 68 atoms, respectively. Within two inner and two outer columns: 102 ± 1 atoms and 633 +10 & 622-10 nucleons, respectively. Regarding at the rows: there are 78 atoms within first two and 78 +11 = 89 within last two rows (about the uniqueness of the number 78 see Solution (2) and legend to Table 3.1]; within first half of the middle row 13, and within the second one 13 +11 atoms. Within two halves of shadow space (light and dark) there are also specific balances: 102 ± 00 atoms and 627/628 ± 55 nucleons (682+573 = 1255). From possible six permutations: **<u>268</u>**, 286, **<u>628</u>**, **<u>682</u>**, 826, 862, the three are in the "game" (bold and underlined) and three are not; the sums of the "chosen" three [(268+628+682 = 2 x 789); (286+826+862 = 2 x 987)] correspond to nucleon number in an indirect way, as it is presented in Tables D3 and D4, in relation to Tables D1 & D2 in Appendix D. [Note: All the amino acid sequences are of the growing series from the aspect of number of atoms; all but one, in which Q-11 precedes V-10, because different stereochemical types have be distinguished: N-Q belong to alanine but V-I to valine type. (Hint: obviously, a species of principle of "uncertainty" is valid here: the order by types excludes the order by the number of atoms and vice versa.).]

Figure 1 shows an arrangement of AAs (in the 4 x 5 system), with the number of atoms as in Solution (4.1) and in Table C1 (in the row starting with 26). On the other hand, in Table 6 we can see the arrangement of AAs (in the 2 x 10 system), with the number of atoms at odd and even positions as in the unique sequence, discussed above (48-50-52-54) [cf. Tables 6.1 & 6.2, and Solution (5)].

	out	in		out	in
	G (01)	N (08)		G (01)	S (05)
	W (18)	Q (11)		A (04)	T (08)
	A (04)	S (05)		L (13)	I (13)
	C (05)	D (07)		V (10)	D (07)
	P (08)	T (08)		P (08)	E (10)
	H (11)	E (10)		R (17)	K (15)
	V (10)	F (14)		Y (15)	F (14)
	Y (15)	M (11)		W (18)	Q (11)
	R (17)	K (15)		H (11)	N (08)
	L (13)	I (13)		C (05)	M (11)
Odd	40	50		48	50
Even	62	52		54	52
	102	102		102	102

Table 6. The outer/inner amino acid pairs which follow from Sukhodolets' Table (Table 3). On the left: the original order as in Table 3; on the right: the chemical order of singlet AAs as it is explained in Section 4.2.1. The sense and generating logic for the sequence 48-50-52-54 see in Table 6.1, and for the sequence 40-50-52-62 in Table 6.2.

G	S	T	A	G	S	M	C	22	154	102 / 102	633+11 / 622-11
L	I	D	V	L	I	Q	W	55	316		
P	E	K	R	P	E	K	R	50	286		
Y	F	Q	W	Y	F	D	V	46	300		
H	N	M	C	H	N	T	A	31	199		
48	50	52	54	48	50	52	54				
				287	310	323	335				
				102 / 102 622/ 633							

Table 6.1. The Table follows from the right subsystem in Table 6; the four non-shadow and four shadow columns of AAs follow from odd-even positions in that subsystem. In shadow space two right columns are in a vice-versa order in relation to non-shadow space. The sequence 48-50-52-54 (here as the number of atoms) is the 24th consecutive quad in a series of even natural numbers quads, as it is explained in the first paragraph of chapter 4.3. The balances of atom number (102:102) and nucleon number [622:633 and (622-11):633+11) are self-evident. (Note: the number of atoms and nucleons within amino acid side chains as in Figure 1.)

(1)	10	20	22	32
(2)	20	30	32	42
(3)	30	40	42	52
(4)	**40**	**50**	**52**	**62**
(5)	50	60	62	72
(6)	60	70	72	82
(7)	70	80	82	92
(8)	80	90	92	102
...				

Table 6.2. The generating logic of the sequence 40-50-52-62. Knowing that this sequence must be in relation to sequence 48-50-52-54 in Table 6, it is clear that after the "start" with first possible two-digit number (the digits "0" and "1" in 10), it follows that just the sequence 40-50-52-62 must be in the middle position of a symmetrical two-digit numbers system.

4.2. "Outer" and "inner" amino acids in Sukhodolets' system

Let us now raise the question how – through an exact chemical analysis – we arrive at the two said arrangements of AAs, the arrangement in Figure 1, and the arrangement in Table 6? To answer this question let us go in reverse order; first – how do we get to Table 6?

In analyzing the relationship between AAs and codons, Sukhodolets started from the premise that the (standard) genetic code had to be fully completed even at the prebiotic stage (Sukhodolets, 1985). The fact that this idea is in full accordance with our hypothesis on the (prebiotic) complete genetic code (Rakočević, 2004a) was a sufficient reason for

giving full faith to his arrangement (Table 2)[6], with a minimum necessary modification (Table 3)[7]; and it further meant that the analysis of Sukhodolets' system should start from several major obvious facts:

1. The positions of AAs within the system in Table 3 are strictly determined;

2. The system in Table 3 consists of two subsystems; the first subsystem: the left (small) subsystem consisting of $4^1 = 4$ AAs (G, N, Q, W)[8] in the form of four singlet sequences [(G), (N), (Q), (W)]; the second subsystem: the right (large) subsystem of $4^2 = 16$ AAs, in the form of four quaternary sequences [(ASDC), (PTEH), (VFMY), (RKIL)];

2.1. In the subsystem with 4 AAs, the outer are G and W; the inner ones N and Q;

2.2. In the subsystem with 16 AAs, alanine (A) and cysteine (C) are the outer; serine (S) and aspartic acid (D) the inner ones, and so on;

3. Besides the two subsystems referred to in the preceding paragraph, there are two sub-systems in terms of the existence of sets of even and odd numbers of hydrogen atoms in the amino acid molecules (the "even" subsystem with shading, and the "odd" subsystem, without shading).

After this analysis we can generate the left subsystem of Table 6: within the first column ("out") are the outer, and within the second one ("in") the inner AAs, given in the same order as in Sukhodolets' system in

---

[6] "В настоящей работе определенный порядок в дублетах оснований в кодонов объясняется на основании гипотезы о предсуществовании кристаллических ассоциатов из свободных молекул оснований и аминокислот" (Суходолец, 1985, с. 1589) {"In this paper, a certain order in the doublets of bases in the codons [and correspondent amino acids] is explained on the basis of the hypothesis on the pre-existence of crystalline associates of the free molecule bases and amino acids" (Sukhodolets, 1985, p. 1589)}.

[7] In the first step the exchange L, I / R, K → R, K / L, I was completed, for distinguishing AAs with even and odd numbers of hydrogen atoms; in the second step: L, I / I, L, for harmonizining the number of conformations in the column with the initial aspartic amino acid (D).

[8] It should be noted that these four AAs represent characteristic extremes: glycine (G) is the only amino acid that does not contain carbon in its side chain; tryptophan (W) is the only one out of 20 AAs with two rings; asparagine (N) and glutamine (Q) are the only two AAs which contain amide groups in their side chains.

Table 3. Now the next question makes sense: is it possible to sort AAs in both columns in accordance with their fundamental chemical nature? The answer is affirmative, as it is demonstrated in the right subsystem of Table 6 and in the next Section (in Section 4.2.1).

4.2.1. The chemical hierarchy of amino acids

Regarding at the Table 6 (on the right), it is self-evident that the simplest glycine (G) must be followed by a little more complex alanine (A). After alanine leucine (L) follows as its counterpart within the alanine stereochemical type, rather than valine (V), which belongs to the valine stereochemical type. Valine comes in the next step, as the first possible semicyclic amino acid. Then comes the first possible cyclic amino acid, the proline (P), as a counterpart to valine[9].

In this way we determined the order for five AAs (G, A, L, V, P), while the remaining five were the following: R, Y, W, H, C. When it is known that 18 AAs are non-sulfur ones, while only two AAs belong to the sulfur type, it is to be expected that sulfur AAs must be at the end.[10] If so, then the nitrogen amino acid proline (P) is followed by another nitrogen amino acid, arginine (R). Of the remaining three AAs (the forth, sulfur C, is at the end, as we concluded), it makes sense for an "ordinary" aromatic one (Y) to come first, and to be followed by two aromatic hetero-cyclic AAs; W should be the first one because it has a "normal" aromatic ring (one of two rings) and thus agrees with Y. Finally, there is a "pure" hetero-cyclic amino acid, H, in contact with sulfur amino acid C, which is "hetero"-non-cyclic.

---

[9] "The side chain of valine … follows from the shortest possible cyclic hydrocarbon, from cyclopropane, with a permanent openness and with a linkage to the 'head' of the amino acid through only one vertex of the cyclopropane 'triangle' … The proline … follows from the same source (cyclopropane), but with a permanent non-openness and with a linkage to the "head" through two vertices of the cyclopropane 'triangle'" (Rakočević & Jokić, 1996).

[10] In relation to the set made of four kinds of atoms (C, N, O, H), the S atom features as a "hetero" atom.

Having determined the order of the outer AAs, we must also determine the sequence of the inner AAs (the rightmost column in Table 6). What is immediately obvious is the (chemically) possible pairing of singlets into doublets: L-I, R-K, Y-F and C-M. It is also obvious that the remaining pairings must be double pairings, exactly as they are presented: a simple pair of S-T with a simple pair of G-A; the non-nitrogen D-E with the non-nitrogen V- P[11]; finally, the nitrogen Q-N with the nitrogen W-H.

With this we have completely generated the system in Table 6, with strictly defined positions of AAs, such that the even and odd amino acid positions (through atom numbers) correspond to the required sequence of 48-50-52-54 atoms.

4.3. Toward four diversity types

Regardless of the existence of the sequence 48-50-52-54 in the system of numbers within Scheme 1 (Table 4), as well as in Sukhodolets' "In-Out" system (Table 6), it makes sense to examine this sequence in a series of natural numbers, such that it traces the "journey" of quads. It turns out to be the 24th consecutive case [**(1) 2-4-6-8**; (2) 4-6-8-10; (3) 6-8-10-12; ...; (23) 46-48-50-52; **(24) 48-50-52-54**].[12]. In this state of affairs it makes sense to set up a *working hypothesis (II)* with the following premises:

1. If the 24th case refers to the number of atoms in 20 AAs (in their side chains), then perhaps the first case relates precisely to the 20 amino acid

---

[11] The nitrogen in the proline derives from the "head" of the amino acid, and only partially belongs to the side chain (cf. footnote 9).

[12] In one of the previous works (Rakočević & Jokić, 1996) we have shown that 24 amino acid singlets within Genetic Code Table appear in the form of 12 doublets (pairs) and 8 triplets, and at the same time we pointed out that the uniqueness of number 24 is manifested precisely through that distinction: "Notice that out of all doublet-triplet systems, this is the only one with two possible distinctions for doublets (i.e. six and six, and then, three and three doublets) and three possible distinctions for triplets (i.e. four and four, then two and two, and, finally, one and one triplet)" (Rakočević & Jokić, 1996).

molecules. That confirms that the 20 amino acid molecules can be classified into four diversity types, in terms of their chemical differences: 2, 4, 6 and 8 AAs.

2. The order of AAs in the 2-4-6-8 sequence must be such that the simpler of the two ones is at the beginning, while the simplest of eight AAs comes at the end.

3. It follows from the previous premise that the arrangement of 20 AAs in the four diversity types inevitably takes a linear as well as circular form, i.e. system, with strictly determined positions of AAs in it (Fig. 2).[13]

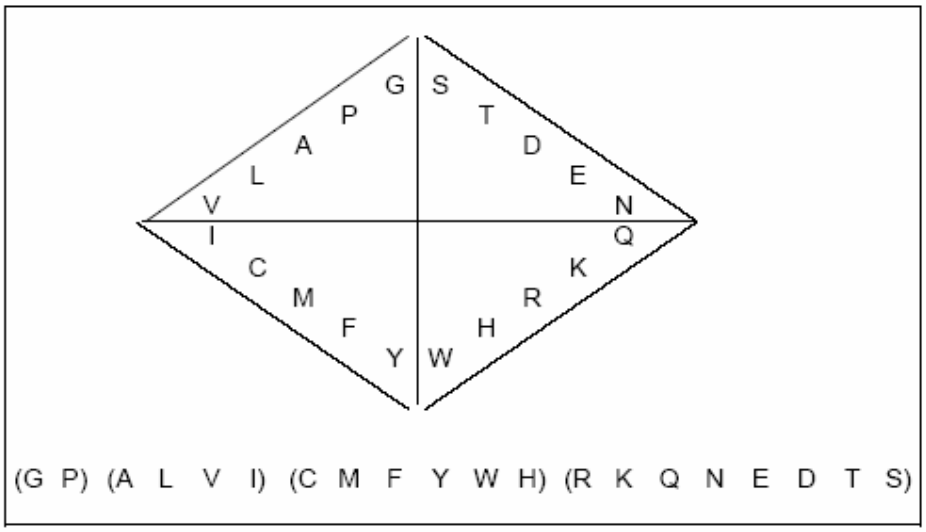


Figure 2. Four diversity types of protein amino acids in a linear arrangement in form of the sequence 2-4-6-8; then in form of a circular arrangement.

[13] Future research should show weather this circular arrangement is in any way related to the circular arrangements of amino acids, mediated by codons (Swanson, 1984; Castro-Chavez, 2010).

4.3.1. Chemical justification for the four diversity types

In classifying of 20 protein AAs into four diversity types, as shown in the linear model in Figure 2, what is immediately apparent is the splitting into 2, 4 and 8 of AA. The first type consists of two AAs (G; P), both without standard hydrocarbon side chains; the second one includes four AAs, as two pairs [(A, L), (V, I)], all with standard hydrocarbon side chains; the fourth type comprises eight AAs, as four pairs [(S, T), (D, E), (N, Q), (K, R)], all with a functional group which also exists in the amino acid functional group (wholly: H_2N-CH_2-COOH; separately: OH, COOH, $CONH_2$, NH_2). In addition, two AAs with the hydroxyl group in the side chain (S, T) come first, followed by two AAs with the carboxyl group in the side chain (D, E) and two of their amide derivatives (N, Q); and at the end of the set are two AAs with the amino group in the side chain (K, R).

It is a little harder to account for a set of 6 AAs, i.e. for the third diversity type, in the sense that it can represent a separate type. Following extensive analysis, the logic (and chemistry) must be as follows: the third type consists of six AAs, as three pairs [(F, Y), (H, W), (C, M)], two aromatic, two hetero aromatic and two “hetero” non-aromatic.

4.3.2. Four diversity versus four stereochemical types

Regarding at the Figure 2, the first diversity type is a doublet (GP) consisting from two stereochemical types [G of glycine stereochemical type, and P of proline stereochemical type]; the second one is a quadruplet (A, L, V, I); (A, L) of alanine stereochemical type and (V, I) of valine stereochemical type); then follow third diversity type (with six AAs) and fourth diversity type (with eight AAs), both of alanine stereochemical type. Precisely, because of these relationships between four diversity types and four stereochemical types, it makes sense an additional pairing between the first and second as well as third and fourth type. By this, the additional pairing between AAs of the first and second type can only participate in (G, P) and (V, I) because none of these four AAs does not belong to the alanine stereochemical type. As two new pairs appear (G,V) and (P,I), which pairs we also find in determination by two classes of

aminoacyl-tRNA synthetases (Table E3 in relation to Tables E1 and E2, in Appendix E). On the other hand, in the case of an additional pairing between the third and fourth diversity types, the possibilities are even less: instead of (S, T) and (C, M) the obtained pairs are (S, C) and (T, M); such the pairs as in the determination by two classes of enzymes aminoacyl-tRNA synthetases, as it is presented in Table E3 and Solution (6). [Note: Within all pairs in two rows of Solution (6) the first member is smaller and the second one (underlined) the larger molecule. In the second (lower) row all first members are handled by the class II of enzymes aminoacyl-tRNA synthetases, while the second members – by class I.]

1 2 3 4 5 V IV III II I
G<u>P</u>, [(A<u>L</u>), (V<u>I</u>)], C<u>M</u>, F<u>Y</u> / <u>W</u>H, <u>R</u>K, <u>Q</u>N, <u>E</u>D, <u>T</u>S

(6)

1,3 I,4 I,4 1,3 2 II III IV V 5
G<u>V</u>, S<u>C</u>, T<u>M</u>, P<u>I</u>, A<u>L</u>, D<u>E</u>, N<u>Q</u>, K<u>R</u>, H<u>W</u>, F<u>Y</u>

4.3.3. Toward an adequate amino acids arrangement

If the positions of AAs in the linear as well as the circular system, in Figure 2, are very strictly established, then it makes sense to test several possible new orders and/or arrangements of AAs (4 x 5 or 5 x 4 AAs in all of the connected sub-sets). For example, starting from the beginning we can take first four, second four, third four AAs, and so on, as it is presented in Solution (7) and, *per se*, related to our CIPS[14], presented in Solutions (8) and (9).[15]

[14] **<u>C</u>**yclic **<u>I</u>**nvariant **<u>P</u>**eriodic **<u>S</u>**ystem, arranged in cycles, starting with first cycle in middle position, as follows: 1st ST/CM, 2nd GP/VI, 3rd DE/NQ, 4th AL/KR and 5th FY/HW. [Note: The number-order in Solution (7) make the numbers taken from Solution (6), but the number-order in Solution (8) make the numbers taken directly from CIPS. But anyway, whether it comes from one or the other sequence, the same pairs of AAs appear in all three systems.] (Rakočević, 2009).

[15] As we can see, in all three systems, (6), (7) and (8), AAs appear to be in the pairs that are chemically reasonable, as we pointed out in our previous papers (Surveys 1.1 & 1.2 in Rakočević

1 2 3 4 5 V IV III II I
GP/AL, VI/CM, **FY/WH**, RK/QN, ED/TS (7)

1 3 2 IV 4 I III II 5 V
GP/VI, AL/RK, CM/TS, QN/ED, **FY/WH** (8)

2 II 4 IV 1 I III 3 5 V
GP/VI, AL/RK, CM/TS, QN/ED, **FY/WH** (9)

From the circular system in Figure 2 it is also possible to take four AAs from the central vertical line (GSYW); then, four AAs from the central horizontal line (NQVI); four from the middle point of each of the four sequences (ADRM); then, four and four from the left/right side, respectively, in relation to the middle point (PKHL/CTEF). As we can see, the five selected sequences correspond to the five sequences in Figure 1.

Each sequence in Figure 1 is arranged by the size of molecules, i.e. by the number of atoms in the side chain,[16] going from left to right; and the order of sequences is determined by the size of the first amino acid molecule in the sequence.[17] Only with such a precise and strictly regulated system can we get the desired result, the sequence 26-42-59-77, signifying the number of atoms in the four columns of AAs (in their side chains) [cf. Table 4, Solution (4.1) and Figure 1]. The obtained result also proves that our *working hypothesis (II)* is fully confirmed.

---

and Jokić, 1996; Survey 4 in Rakočević, 1998), all as the pairs in harmonic structures, presented in Appendix E.

[16] The only exception is valine, which is understandable enough when we know that valine and isoleucine belong to the same stereochemical type, the valine type. That sequence should, therefore, be understood as follows: two AAs of the alanine type (N, Q) are followed by two AAs of the valine type (V, I).

[17] The dilemma whether before N or P is resolved by the following pairs: N is followed by a smaller pair V(10) – I(13), while P by a larger one L(13) – K(15).

4.3.4. Toward a corresponding codons arrangement

If the *working hypothesis (II)*, related to Solution (4.1), is fully confirmed, then it makes sense to set up a *working hypothesis (III)*, related to Solution (4.2): it must be that the quantities, given in Solution (4.2), $x_1 = 26$, $y_1 = 10$ as well as $x_2 = 17$, $y_2 = 08$ in a way also contained in the genetic code. Figure 3 is an obvious and direct evidence for this. The first diversity type of AAs and corresponding **08** codons appears to be diagonally on the right within Genetic Code Table (GCT), designated here in light tones; the second one with **17** codons on the left (in dark tones); altogether in two of low-level-function diversity types there are 25 codons.

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

Figure 3. The standard Genetic Code Table with designation of four diversity types of protein amino acids and corresponding codons: first and second type without color (in light and dark tones, respectively), but third and forth in color. The codon number: first **08**, second **17**, third **10** and fourth **26** [just as in algebraic system in Solution (4.2)]. The roman numbers designate class I and class II of aminoacyl-tRNA synthetases as in Table E3. The details see in the text.

The third diversity type of AAs and corresponding **10** codons follows, within GCT, in the next order (light blue): in column "U" *up* and *down*, and in columns "A" and "G" only *up*. The fourth type, with **26** codons (dark blue), in column "C" *up* and *down*, and in columns "A" and "G" only *down*; altogether in two of high-level-function diversity types there are 36 codons.

G P	A L V I	C M F Y W H	R K Q N E D T S
9+18=27	40+36 =76	74 + 54 = 128	81+ 72 = 153
Atom number: (27 + 153 = **180**); (76 + 128 = **204**)			
G P **A L** V I **C M** F Y **W H** R K **Q N** E D **T S**			
Pairs on the linear arrangement in Fig. 2: Odd (633+11) / (**622 - 11**) Even			
V I **L C** A M **P F** G Y **S W** T H **D R** E K **N Q**			
Pairs on the circular arrangement in Fig. 2: Odd (569 / **686**) Even			
G **P**	A **L** V **I**	C **M** F **Y** W H	R **K** Q **N** E **D** T **S**
Singlets on the linear arrangement in Fig. 2: Odd (627-10) / (**628 + 10**) Even			

Table 7. The relationships between four diversity types of protein amino acids, in the linear as well as circular arrangement. The calculations above: within 10 AAs of two inner types there are 180 atoms, just as within 20 amino acid "heads", i.e. 20 amino acid functional groups. On the other hand, within 10 AAs of two outer types there are 204 atoms, just as within 20 amino acid side chains. This specific "simulation" is analogue to the "simulation", valid for the number of protons and neutrons in Table 3 as well as in two last rows here. The balances within all other calculations are self-evident.

By all these results one must notice a very specific self-similarity, realized through arithmetical and/or algebraic regularities: within the system in Figure 1, related to the arithmetical system in Solution (4.1), the number 26 appears to be a minimal atom number whereas within the

system in Figure 3, related to the algebraic system in Solution (4.2), the number 26 as a maximal codon number. On the other hand, since the number of atoms within codons nucleotides in two inner as well as two outer columns in GCT is 3456 both times, we can say that there is a correspondence with four exponents (3, 4, 5, 6)[18] which appear in algebraic system, given in Solution (4.2). All together, we can speak about a specific self-similarity triad (SST). [*Note*: The calculation of the number of atoms within 64 x 3 = 192 nucleotides in GCT after data: UMP 34, CMP 35, AMP 37, GMP 38; as it is presented in Figures A2 and A3.]

With these observations and, in addition, the observations marked into relationships, presented in Table 7, we conclude the analysis of relations within the four diversity types of protein amino acids and their corresponding codons.

5. FINAL COMMENT

As we have said in the Introduction, the intention of this paper was an attempt to answer the question why chemical classifications-systematizations of protein amino acids are followed by specific arithmetical regularities; or on the other words – what is the "physical nature" of such a phenomenon. Our answer is given in the *working hypothesis (I)*, in Preliminaries, according to which such determination follows from the fact that the genetic code is (just prebiotic) complete system, whose (harmonic) structures appear to be in a strict correspondence to the series of natural numbers, with all its regularities and laws. Evidences which are supporting *the working hypothesis (I)* we gave in the third and fourth chapter. With the knowledge that this is just

[18] In the set of all possible exponent quads [(1,2,3,4), (2,3,4,5), **(3,4,5,6)**, (4,5,6,7), ...] only two, the second and the third, contain the Pythagorean triplet. However, in the second case the number of possible "codons", 41, does not follow toward a logical Boolean cube (41"codons" + 8 stop "codons" = 49 = 7^2). In the third case the number of possible codons, 61, follows toward a logical Boolean cube (61 codons + 3 stop codons = 64 = 8^2). (About genetic code as a Boolean space see in Rakočević, 1997.)

so, it makes sense set out two predictions that could be useful for further researches.

Prediction I. The correspondence of possible classification structures of amino acids (and corresponding codons), especially the classification into four diversity types, with the series of natural numbers, should be also reflected on the entire proteome (and thus on the entire genome), measured on a representative sample of organisms.

Prediction II. If the first prediction be proved true, it would then mean that the statement of prediction I is also valid, *mutatis mutandis*, on the proteome and genome of individual organisms.

However, with the first prediction we actually knocking on the open door. Namely, the existing samples of proteins in which were analyzed amino acids relationships (Swanson, 1984, Taylor 1986, 1997, Kosiol et al, 2004) show that within so called Mutation ring the relationships are just such that „amino acids that are close together exchange frequently“ in the evolution process (Kosiol, 2004, Figure 4 on p. 104). But now we see that interchangeable amino acids in the Mutational ring are actually these amino acid pairs that follow from the classification into four diversity type; in other words, follow from the correspondence between amino acid classification structures and the series of natural numbers [Appendix F, especially Figure F7 and Solution (F2)]. Hence, the first prediction should be understood so that it will be applied to all future representative samples of organisms.

Of course, in the above answer (to two Shcherbak's questions)[19] is hidden a new question - why is it so that the genetic code is arranged just as it is. Our answer to this "hidden" question is a *hypothesis* (for further researches) after which the solution should be related to the Mendeleyev periodic system of chemical elements (PSE). For a complete answer to this new question it is necessary, however, to resolve previously all doubts

[19] That in connection of amino acid classification structures and arithmetical regularities appear to be the series of natural numbers.

concerning to the PSE. For example, is the hydrogen really in the first and the seventh group at the same time, or only in the seventh, as we understand it (cf. Table G2 and G3, in relation to Table G1).[20]; that the PSE, in the form of short periods, appear to be a logical cube; or, in the form of long periods, a logical hypercube, as it is indicated in the original Mendeleyev (manuscript) works[21], and so on.

Revealing that in determination of the genetic code, except two inherent alphabets, is involved still one "hidden alphabet" (a series of natural numbers) our intention was not only to shed light on some new classifications and new arithmetic regularities in the number of atoms and/or nucleons (which could be of use in future research in finding answers to the questions Shcherbak raised (Shcherbak, 1993, 1994)), but to provide broader answers regarding the role of numbers in the genetic code (Maddox, 1994; Dragovich & Dragovich, 2006, 2007; Dragovich, 2009; Negadi, 2009, 2011).

Apart from that, even regardless of answering these questions, it is our belief that the presented classification of protein AAs into the four diversity types will be the *key to understanding* the genetic code on a general plane as well.[22]

Acknowledgement

I am grateful to Branko Dragovich for pointing out to me Sukhodolets' works and for a helpful, stimulating discussion about the genetic code.

[20] Similarly, the question of whether the third period is double or triple (the latter is also our idea; cf. Table G2 and G3).

[21] Cf. Photocopies X and XI, as well as Tables 13 and 16 in (Kedrov, 1977), that show the three-dimensionality and indicate the four-dimensionality. Cf. also photocopies III, VIII and IX in which Mendeleev demonstrated that the periodicity of elements properties are associated with the corresponding properties of their compounds (Kedrov, 1977; Rakočević, 2011: Mendeleev's Arhive) (http://www.rakocevcode.rs).

[22] Besides this, in my book (GENETIC CODE – Keys to Understanding), soon to be installed on my website (http://www.rakocevcode.rs), I will point out eight more keys to understanding the genetic code.

I would also like to thank Dragiša Janković for a wholehearted support – through many years – in a sophisticated analysis of the key elements of mathematics, especially arithmetic and number theory.

REFERENCES

Castro-Chavez, F. (2010) The rules of variation: amino acid exchange according to the rotating circular genetic code, J. Theor. Biol., 264, 711-721.

Dragovich B., Dragovich, A. (2006) A p-Adic model of DNA sequence and genetic code, p-Adic Numbers, Ultrametric Analysis and Applications 1, 34-41; arXiv:q-bio/0607018v1 [q-bio.GN].

Dragovich, B., Dragovich, A. (2007) p-Adic Modelling of the Genome and the Genetic Code, The Computer Journal 53 (2010) 432-442; arXiv:0707.3043v1 [q-bio.OT].

Dragovich, B. (2009) Genetic Code and Number Theory, in "Modern Topics in Science", a book of invited papers (Eds. R. Constantinescu, G. Djordjevic, Lj. Nesic), Niš, Serbia; arXiv:0911.4014v1 [q-bio.OT].

Kedrov, B.M. (1977) Prognozy D.I. Mendeleyeva v atomistike, Atomizdat, Moskva.

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

Kosiol, C. et al. (2004) A new criterion and method for amino acid classification, J. Theor. Biol., 228, 97–106.

Maddox, J. (1994) The genetic code by numbers, Nature, 367, p. 111.

Negadi, T. (2009) The genetic code degeneracy and the amino acids chemical composition are connected, Neuroquantology, Vol. 7, 1, 181-187; arXiv:0903.4131v1 [q-bio.OT].

Negadi, T. (2011) The multiplet structure of the genetic code, in one and small number, arXiv:1101.2983v2 [q-bio.OT] (Submited to Neuroquantology Journal).

Popov, E. M. (1989) Strukturnaya organizaciya belkov, Nauka, Moskva.

Rakočević, M. M. (1988) Three-dimensional model of the genetic code, Acta Biologiae et Medicinae Experimentalis, 13, 109-116.

Rakočević, M. M. (1991) The coherence of the chemical and genetic code, Proceedings of Faculty of science, Chemistry Section, 2, 1-29, Niš.

Rakočević, M. M. (1997) Genetic code as a unique system, SKC, Niš. (http://www.rakocevcode.rs)

Rakočević, M. M. (1998) The genetic code as a Golden mean determined system, Biosystems, 46, 283-291.

Rakočević, M. M. (2004a) A harmonic structure of the genetic code, J. Theor. Biol. 229, 463 – 465.

Rakočević, M. M. (2004b) Further Generalization of Golden Mean in Relation to Euler Divine Equation, FME Transactions, 32, 95-98; arXiv:math/0611095v1 [math.GM].

Rakočević, M. M. (2006) Genetic Code as a harmonic system, arXiv: q-bio/0610044v1 [q-bio.OT]

Rakočević, M. M. (2009) Genetic Code: The unity of the stereochemical determinism and pure chance, arXiv:0904.1161v1 [q-bio.BM].

Rakočević, M.M. (2011) Mendeleyev's Archive (http://www.rakocevcode.rs).

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

Shcherbak, V. I.(1993) Twenty canonical amino acids of the genetic code: the arithmetical regularities, part I, J Theor. Biol., 162, 399-401.

Shcherbak, V. I. (1994) Sixty-four triplets and 20 canonical amino acids of the genetic code: the arithmetical regularities. Part II, J Theor. Biol., 166, 475-477.

Суходолец, В.В. (1985) Смысл генетическово кода: реконструкция этапа предбиологической эволюции, *Генетика*, XXI, 10, 1589 – 1599.

Sukhodolets, V. V. (1985) A sense of the genetic code: reconstruction of the prebiological evolution stage, *Генетика*, XXI, 10, 1589 – 1599.

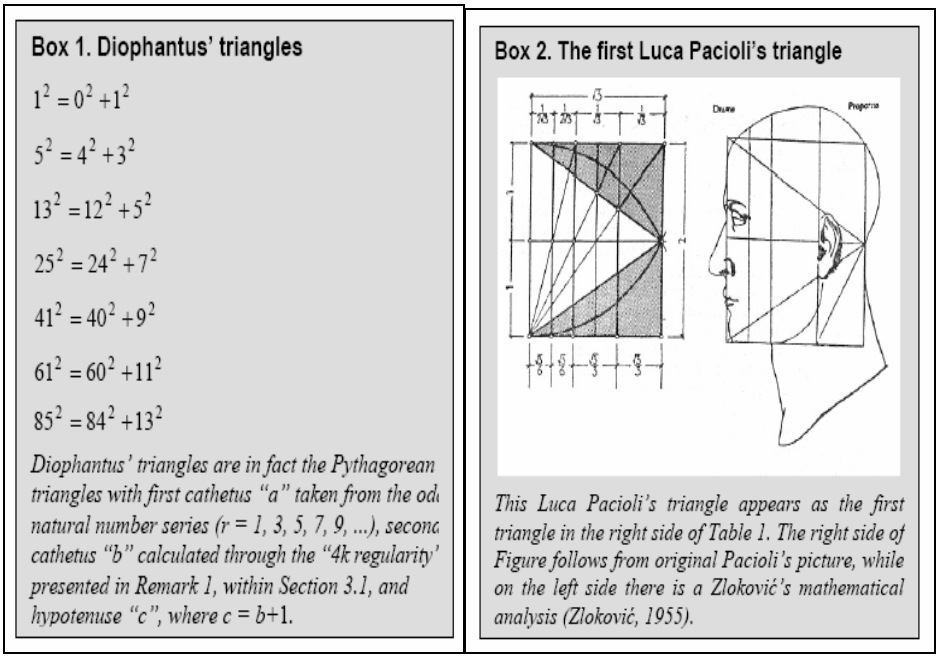
Swanson, R. (1984) A unifying concept for the amino acid code, Bull. Math. Biol., 46, 187-207.

Taylor, W.R. (1986) The classification of amino acid conservation. J. Theor. Biol. 119, 205–218.

Taylor, W.R. (1997) Residual colours: a proposal for aminochromography. Prot. Eng. 10, 743–746.

Verkhovod, A. B. (1994) Alphanumerical divisions of the universal genetic code: new divisions reveal new balances, J. Theor. Biol., 170, 327–330.

Appendix A: Harmonic structures from original works (I)



Box 1. Diophantus' triangles

$1^2 = 0^2 + 1^2$

$5^2 = 4^2 + 3^2$

$13^2 = 12^2 + 5^2$

$25^2 = 24^2 + 7^2$

$41^2 = 40^2 + 9^2$

$61^2 = 60^2 + 11^2$

$85^2 = 84^2 + 13^2$

Diophantus' triangles are in fact the Pythagorean triangles with first cathetus "a" taken from the odd natural number series (r = 1, 3, 5, 7, 9, ...), second cathetus "b" calculated through the "4k regularity" presented in Remark 1, within Section 3.1, and hypotenuse "c", where c = b+1.

Box 2. The first Luca Pacioli's triangle

This Luca Pacioli's triangle appears as the first triangle in the right side of Table 1. The right side of Figure follows from original Pacioli's picture, while on the left side there is a Zloković's mathematical analysis (Zloković, 1955).

Figure A1. The Diophantus' triangles (on the left) and Luca Paciolis' triangle (on the right), the first triangle in the system presented in Table A1, on the right (Rakočević, 2004b).

N	x_1	x_2	$\underline{h}$	m	$\sqrt{r}$		N	X_1	x_2	$\underline{h}$	M	$\sqrt{r}$
0.	$0^2 + 1^2 = \underline{\mathbf{1}}$ $(0 + 1)^2 = 1$			0	$\sqrt{1}$		0.	$0^2 + 1^2 = \underline{\mathbf{1}}$ $(0 + 1)^2 = 1$			0	$\sqrt{1}$
1.	$1^2 + 2^2 = \underline{\mathbf{5}}$ $(1 + 2)^2 = 9$			4	$\sqrt{9}$		1.	$(x_1)^2 + (x_2)^2 = \underline{\mathbf{2}}$ $(x_1 + x_2)^2 = 3$			1	$\sqrt{3}$
2.	$2^2 + 3^2 = \underline{\mathbf{13}}$ $(2 + 3)^2 = 25$			12	$\sqrt{25}$		2.	$(x_1)^2 + (x_2)^2 = \underline{\mathbf{3}}$ $(x_1 + x_2)^2 = 5$			2	$\sqrt{5}$
3.	$3^2 + 4^2 = \underline{\mathbf{25}}$ $(3 + 4)^2 = 49$			24	$\sqrt{49}$		3.	$(x_1)^2 + (x_2)^2 = \underline{\mathbf{4}}$ $(x_1 + x_2)^2 = 7$			3	$\sqrt{7}$
4.	$4^2 + 5^2 = \underline{\mathbf{41}}$ $(4 + 5)^2 = 81$			40	$\sqrt{81}$		4.	$1^2 + 2^2 = \underline{\mathbf{5}}$ $(1 + 2)^2 = 9$			4	$\sqrt{9}$
5.	$5^2 + 6^2 = \underline{\mathbf{61}}$ $(5 + 6)^2 = 121$			60	$\sqrt{121}$		5.	$(x_1)^2 + (x_2)^2 = \underline{\mathbf{6}}$ $(x_1 + x_2)^2 = 11$			5	$\sqrt{11}$
	(...)							(...)				

Table A1. The relationships within Generalized Golden Mean, in relation to the natural numbers series (Rakočević, 2004b).

Порядок в генетическом коде, выявляемый при разделении аминокислот на группы по числу протонов в молекуле

Число протонов — номер группы	Аминокислоты	Основание кодона *	
		первое	второе
5	глн	Г	Г
7	ала сер асп цис	Г У | Г У	Ц Ц | А Г
8	асп	А	А
9	про тре глу гис	Ц А | Г Ц	Ц Ц | А А
10	глн	Ц	А
11	вал фен мет тир	Г У | А У	У У | У А
12	трп	У	Г
13-14	лей иле арг лиз	Ц А | Ц А	У У | Г А

Table A2. The original Sukhodolets' Table (Sukhodolets, 1985): the numbr of hydrogen atoms within amino acid molecules, in relation to natural numbers series: 5, (6), 7, 8, 9, 10, 11, 12, 13, 14.

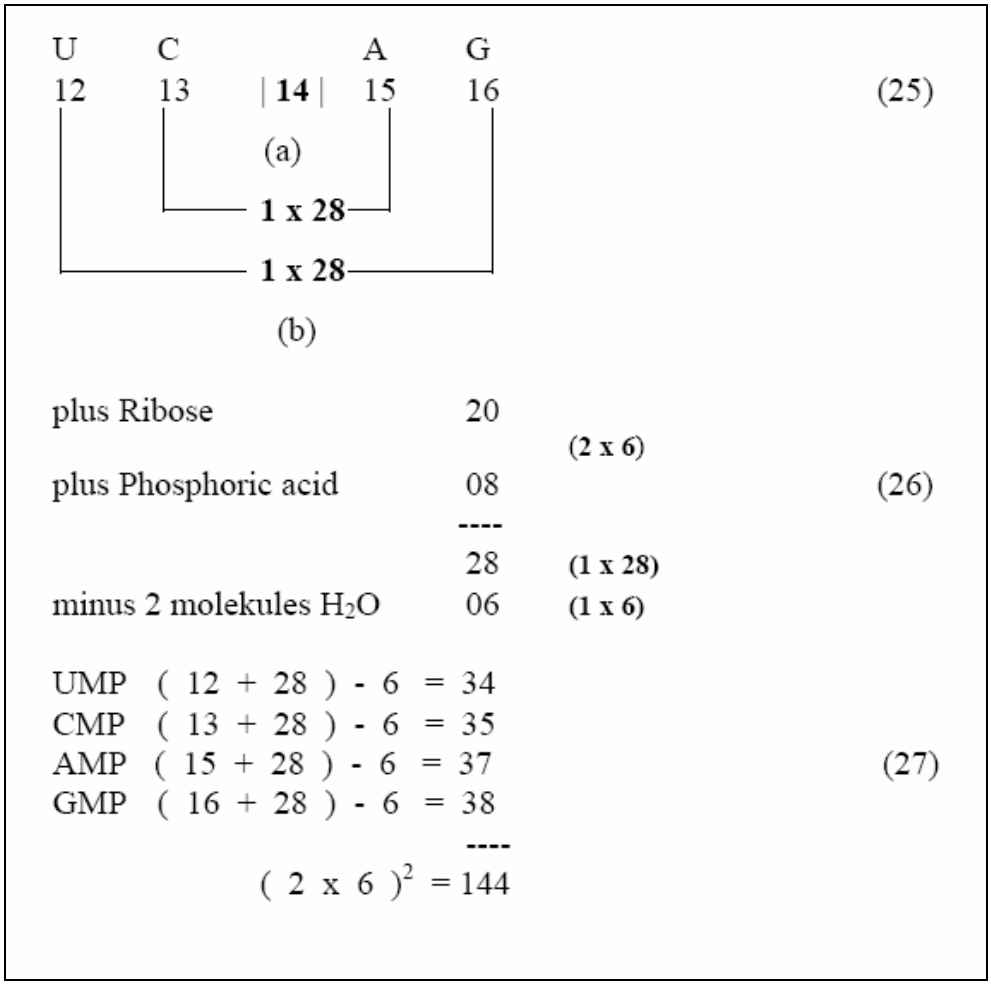


Figure A2. Atom number within four bases as well as four nucleotides (Rakočević, 1997) [Cf. legend to Table 2.]

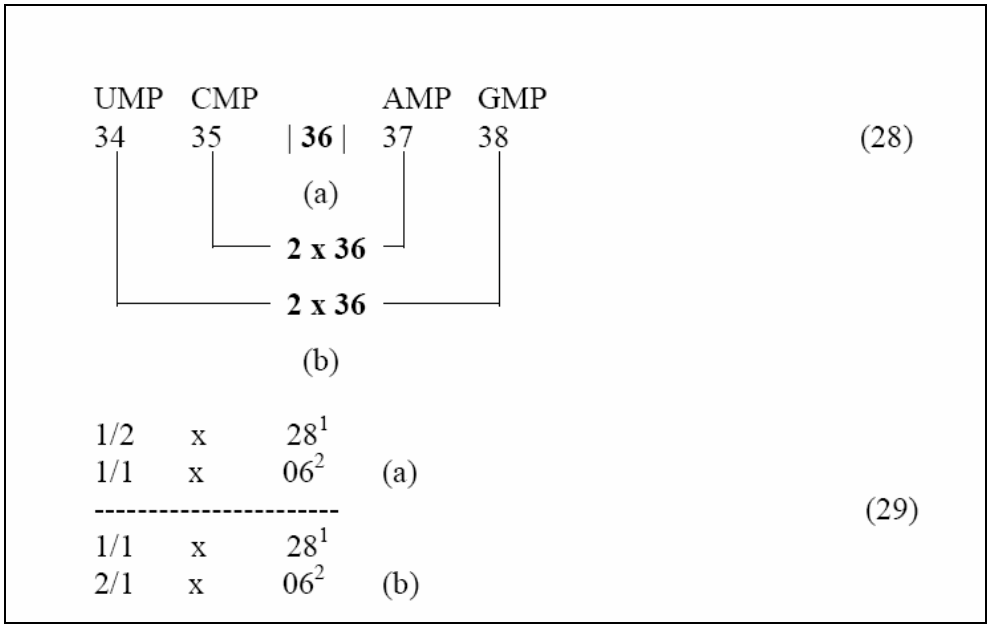


Figure A3. Atom number within four nucleotides (Rakočević, 1997). [As a curiosity: the numbers 6 and 28 are the first and second perfect number, respectively.]

Appendix B: Golden mean of natural numbers

	a	b	a – b
1	0.61803 ...(ϕ^1)	0.38196 ...(ϕ^2)	0.23606... (ϕ^3)
2	1.24	0.76	0.48
3	1.85	1.15	0.70
...	...	...	...
9	5.56	3.44	2.12
10	6.18	3.82	2.36
11	**6.80**	**4.20**	**2.60**
12	7.42	4.58	2.84
...			

a:b = 1.618033 ...

$(1.618033 ...)^2 = 2.618033 ...$

Table B1. The number 11 is only and one number within the set of two-digit numbers, whose golden mean is the most "golden": through the shortest distance (0.02) between (a – b) and $(a:b)^2$.

	a	b	a – b
1	0.61803 ...(ϕ^1)	0.38196 ...(ϕ^2)	0.23606...(ϕ^3)
11	**6.8** (6.79837 ...)	**4.2** (4.20162 ...)	**2.6** (2.59674 ...)
111	**68.6** (68.60177 ...)	**42.4** (42.39822 ...)	**26.2** (26.20354 ...)
1111	**686.6** (686.63577 ...)	**424.4** (424.36422 ...)	**262.2** (262.27154...)

a:b = 1.618033 ...

$(1.618033 ...)^2 = 2.618033 ...$

Table B2. The number 111 is more "golden" than its neighbors 11 and 1111, measured through the distance between (a – b) and [n (a : b)^2] (n = 1, 10, 100, respectively).

Appendix C: The specific relations within the system of numbers in Shemes 1 & 2 (Tables 4 & 5)

X_{10}		I	II	III	IV	Y	Y/4	Z	Z - Y/4
04	→	04	-02	-07	-11	-16	-04	-15	-(5.5 x 2)
15	→	15	20	26	33	094	23.5	18	-(5.5 x 1)
26	→	**26**	**42**	**59**	**77**	**204**	**51**	**51**	**±(0.0)**
37	→	37	64	92	121	314	78.5	84	+(5.5 x 1)
48	→	48	86	125	165	424	106	117	+(5.5 x 2)
59	→	59	108	158	209	534	133.5	150	+(5.5 x 3)
70	→	70	130	191	253	644	161	183	+(5.5 x 4)
81	→	81	152	224	297	754	188.5	216	+(5.5 x 5)
. . .									

Table C1. This Table follows from Table 4. The number 26, from 4th column of Table 4, appears here to be in a zeroth position; (Y = I + II + III + IV; Z = 16 + 17 + 18); (the sequence 16, 17, 18 appears to be in a diagonal neighborhood to the number 26 within Scheme 1 in Table 4). Other explanations in the text.

q	a	X_q	X_{10}	Y_{10}	Y/4	Z_{10}	Z - Y/4	
04	13_4	02	02	00	00	- 003	- (3 x 1)	
07	25_7	14	11	72	18	015	- (3 x 1)	0
10	**37_{10}**	**26**	**26**	**204**	**51**	**051**	**± 0.0**	1
13	49_{13}	38	47	396	99	105	+ (3 x 2)	2
16	$5B_{16}$	4A	74	648	162	177	+ (3 x 5)	3
...								

1 1 0 2 5 9 14 20 ...
0 1 2 3 4 5 6 ...

Table C2. This Table shows the relationships in Table 4 for decimal numbering system and analog Tables (not given here) for Shcherbak's numbering systems (q) and Shcherbak's numbers, analogs of number 37_{10} (a). As it is clear, decimal numbering system is a zeroth system within this arrangement, and other numbering systems (within column "q"), through their distances in relation to zeroth point, make a strict natural numbers series.

$2^2 - 1^2$	3
$3^2 - 2^2$	5
$4^2 - 3^2$	7
$5^2 - 4^2$	9
$6^2 - 5^2$	11
$7^2 - 6^2$	13
$8^2 - 7^2$	15
$9^2 - 8^2$	17

Table C3. The uniqueness of the pair 25-36 in Scheme 1 (Table 4). Details in the text.

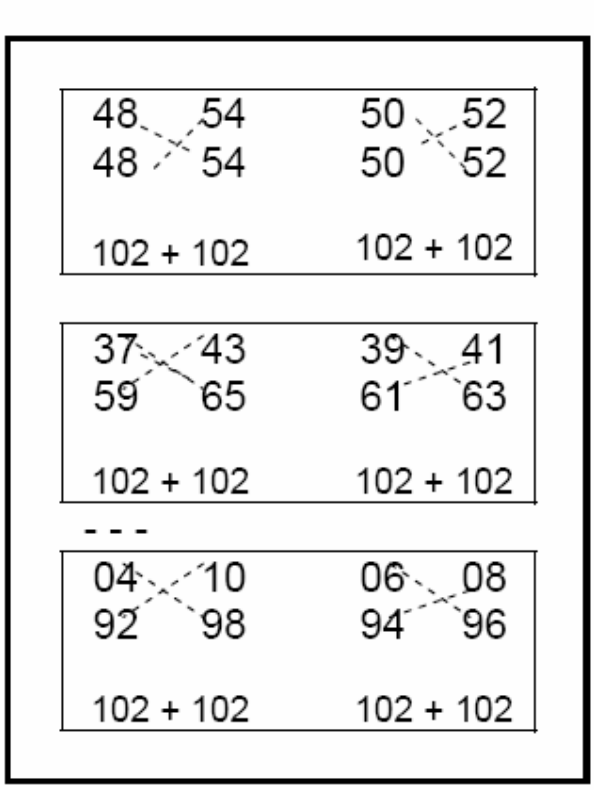


Table C4. From all crossing connections within Scheme 1 (Table 4) follows result 102 that we have in genetic code, as a half of all atoms in 20 AAs (their side chains).

Appendix D: *Relationships within standard genetic code Table*

1. F 91 2. L 57	1. S 31	6. Y 107 stop	5. C 047 stop 6.W 130
2. L 57	2. P 41	7. H 81 8. Q 72	7. R 100
3. I 57 4.M 75	3. T 45	9. N 58 10. K 72	1. S 031 7. R 100
5. V 43	4. A 15	10. D 59 9. E 73	8. G 01
Odd 627-1/ Even 628+1			
1.F 2.L 3.I 4.M 5.V 6.Y 7.H 8.Q 9.N 10.K			
1.S 2.P 3.T 4.A 5.C 6.W 7.R 8.G 9.E 10.D			

Table D1. In the standard Genetic Code Table (GCT) inevitably there are two amino acid sequences; the sequence of those AAs encoded by weaker codons (the middle base U or A) and of those AAs that are encoded by stronger codons (middle base C or G). One obtains the two "necklaces", with ten amino acids each, connected at the loop K-D (very bottom within two weak columns, U & A, there are strong codons because they possess strong bases in the first codon position). The obtained two "necklaces" are in correspondence with two strings in Table D2. Color marked amino acid pairs in correspondence with two classes of amino acids, handled by two classes of enzymes aminoacyl-tRNA synthetases (as in Table E3). It is obvious that the members within any pair are in neighborhood or near to the neighborhood.

Conf. N	12	22	20	20	08	12	24	38	16	66
Isot. N	28	26	26	24	20	31	22	23	17	30
PN	49	33	33	41	25	57	43	39	31	41
NN-1	91	57	57	75	43	107	81	72	58	72
NN-T	196	127	127	231	96	247	173	173	142	159
M. Mass	165.19	131.18	131.18	149.21	117.15	181.19	155.16	146.15	132.12	146.19
AN	14	13	13	11	10	15	11	11	08	15
	+	+	+	+	+	−	−	−	−	−
	F	L	I	M	V	Y	H	Q	N	K
	S	P	T	A	C	W	R	G	E	D
	−	−	−	+	+	−	−	−	−	−
AN	05	08	08	04	05	18	17	01	10	07
M. Mass	105.09	115.13	119.12	089.09	121.16	204.23	174.20	075.07	147.13	133.10
NN-T	85	90	116	34	169	278	217	03	192	161
NN-1	31	41	45	15	47	130	100	01	73	59
PN	17	23	25	09	25	69	55	01	39	31
Isot. N	11	16	17	08	12	36	34	02	22	16
Conf. N	09	02	08	03	21	24	66	04	20	10

	AN	M. Mass	NN-T	NN-1	PN	Isot. N	Conf. N
Odd	102-1	1369-1	15*13*	627-1	343-1	210-1	203+1
Even	102+1	1369+1	15*03*	628+1	343+1	211+1	202-1

Table D2. This table corresponds to Table 7 in Rakočević, 2004a. Plus and minus marks for non-polar and polar amino acids, respectively, according to the hydropathy index (Kyte and Doolittle, 1982). Designations are as follows; AN - number of atoms, NN1 - the number of nucleons in the first nuclide; NNT - total number of nucleons; PN - the number of protons; Isot. N - number of isotopes; Conf. N - number of conformations (Popov, 1989); M. Mass - molecular mass. The balances are selfevident.

F 91 L 57	S 31	Y 107 stop	C 47 stop W 130	
L 57	P 41	H 81 Q 72	R 100	789
I 57 M 75	T 45	N 58 K 72	S 31 R 100	
V 43	A 15	D 59 E 73	G 01	
789				

Table D3. The number of nucleons within two outer columns and two inner rows in GCT (Verkhovod, 1994). Bearing in mind nucleon number in contra-spaces (Table D4), the correspondence with natural number sequence is self-evident (**789**/987 and 456/**654**).

F 91 L 57	S 31	Y 107 stop	C 47 stop W 130	
L 57	P 41	H 81 Q 72	R 100	654
I 57 M 75	T 45	N 58 K 72	S 31 R 100	
V 43	A 15	D 59 E 73	G 01	
654				

Table D4. The number of nucleons within two outer rows and two inner columns in GCT (Verkhovod, 1994). Bearing in mind nucleon number in contra-spaces (Table D3), the correspondence with natural number sequence is self-evident (456/**654** and **789**/987).

Appendix E: *Harmonic structures from original works (II)*

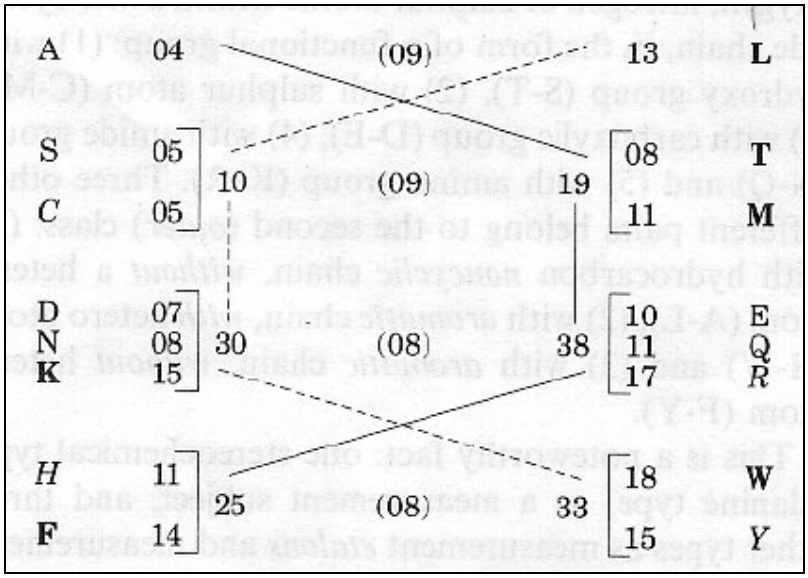


Table E1. This Table corresponds to Survey 1.1 in: (Rakočević & Jokić, 1996). The 16 AAs (8 pairs) of alanine stereochemical type (in further communication as the first subsystem, SubS1) are arranged in accordance to natural numbers series, as it is presented in Box 1. When this subsystem is connected with the subsystem in Table E2, then we obtain a complete overview of ten amino acid pairs; the same pairs as in upper row of Solution (e1). In the bottom row of Solution (e1) there are the pairs as in the linear arrangement of the system consisting of four diversity types, shown in Figure 2.

(GV, PI); **(**AL, ST, CM, DE, NQ, KR, HW, FY)

(e1)

(GV, PI); (AL, CM, FY / WH, RK, QN, ED, TS)

Box E1. *Molecule pairs hierarchy in relation to natural numbers series (I)*

1. Aliphatic AAs (simpler than aromatic)

1.1. Hydrocarbon AAs (start: 04 atoms);
1.2. First possible OH derivatives (start: 05 atoms);
1.3. Sulfur OH analog, i.e. SH derivatives (start: 05 atoms; S > O);
1.4. Carboxyl group derivatives (start: 07 atoms);
1.5. Amide derivatives (start: 08 atoms);
1.6. Amino derivatives (start: 15 atoms);

2.2. Heteroatom derivatives (start: 11 atoms);
2.1. Aromatic hydrocarbon and its OH derivative (start: 15 atoms);

2. Aromatic AAs (more complex than aliphatic)

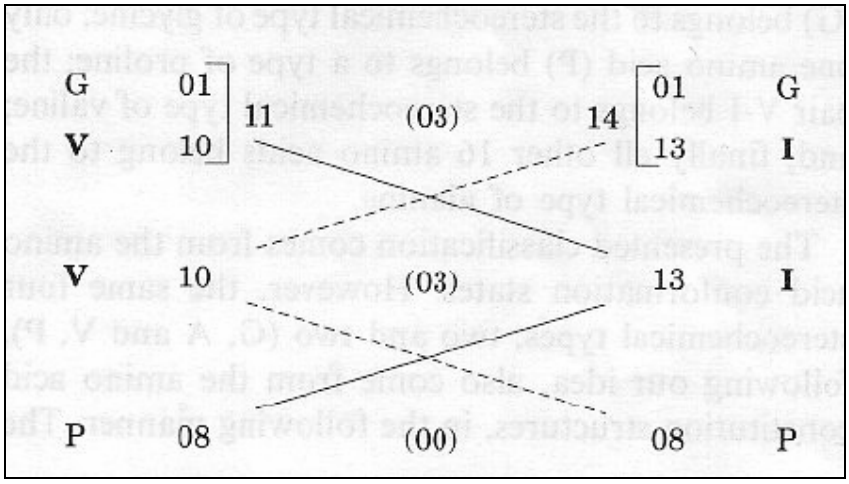


Table E2. This Table corresponds to Survey 1.2 in: (Rakočević & Jokić, 1996) (the second subsystem, SubS2). Here there are four AAs of non-alaninic (non-alanine) stereochemical types (G in glycinic, P in prolinic, and V-I in valinic stereochemical types) [The four stereochemical types after: (Popov, 1989; Rakočević & Jokić, 1996).] Except G-G, P-P and V-I, the possible pairs are G-V and P-I, as it is presented in the system in Table D3, in accordance to two classes of aminoacyl-tRNA synthetases. (Cf. G-P as a pair, in Fig. 2.)

Box E2. *Molecule pairs hierarchy in relation to natural numbers series (II)*

1. Vertical pairs in SubS1 (Fig. D1) and SubS2 (Fig. D2)

1.1. Glycine as the simplest (G-V) (start: 01 atom);
1.2. Serine as next (S-C) (start: 05 atoms);
1.3. Threonine as next (T-M) (start: 08 atoms);
1.4. Proline as next (P-I) (start: 08 atoms; M < I);

2. Horizontal pairs in SubS1 (Fig. D1) and SubS2 (Fig. D2)

2.1. Hydrocarbons (A-L) (start: 04 atoms);
2.2. Carboxyl group derivatives (D-E) (start: 07 atoms);
2.3. Amide derivatives (N-Q) (start: 08 atoms);
2.4. Amino derivatives (K-R) (start: 15 atoms);

3.2. Heteroatom derivatives (H-W) (start: 11 atoms);
3.1. Aromatic hydrocarbon and its OH derivative (F-Y) (start: 15 atoms);

3. Aromatic AAs (more complex than aliphatic)

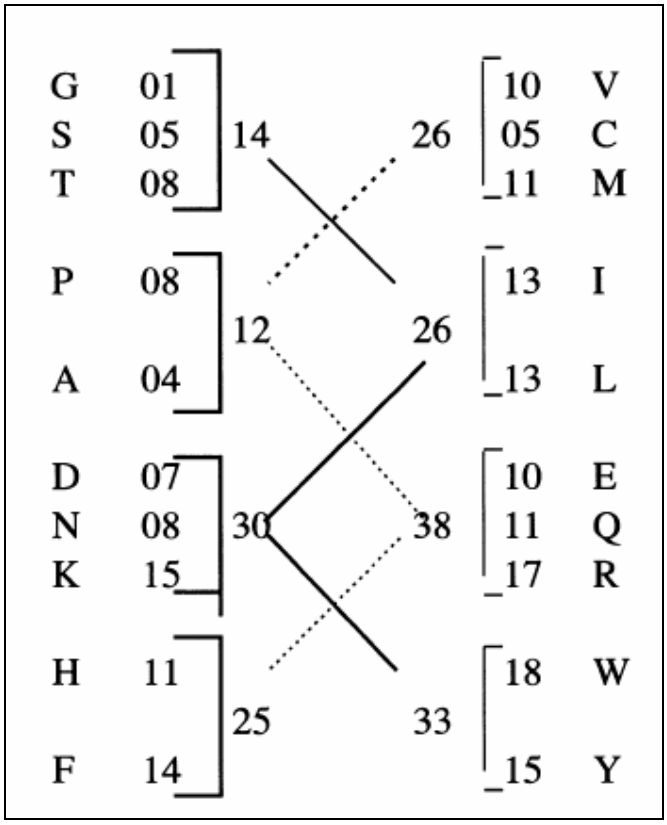


Table E3. This Table follows the previous two tables, the Table E1 and E2 from. When one keeps in mind the determination by two classes of aminoacyl-tRNA synthetases, it is seen that some pairs in Table E1 and Table E2 must be read as horizontal or as vertical pairs. In this Table come first vertical, then horizontal pairs, so as stated in box 2; the pairs as in upper row of Solution (e2). In the bottom row there are the pairs as in the linear arrangement of the system consisting of four diversity types, shown in Figure 2. The pairs are consistent in all but apart within the analogues ST and CM exchange occurred: the small molecule with the small, the large with the large.

(GV, PI); (SC, TM, AL, DE, NQ, KR, HW, FY)

(e2)

(GV, PI), (AL, CM, FY / WH, RK, QN, ED, TS)

0.0. **D**	0.1. **N**		5. A	5. L
1. R	1. F		4. P	4. I
2. K	2. Y		3. T	3. M
3. H	3. W		2. S	2. C
0.2. **E**	0.3. **Q**		1. G	1. V

Table E4. This Table represents the connection between the Table E3 and Table E5. In Table E3 the molecule pair D-E make the only two charged acidic molecules. In their neighborhood are two their amide derivatives. The following six molecules, in a strict order, appear to be two triplets; the first (three basic molecules), the triplet RKH and the second one (three neither acidic nor basic molecules), the triplet FYW. Actually, the first triplet ends with the sole aromatic amino acid, which is basic charged (H), and continue with three aromatic non-electrified; and then: between the first two amino acids (D-E) comes the first triplet, and between the second two (N-Q) comes the second triplet. Thus, in such a manner is generated the left half of the system in Table E5; then comes the right half from the upper part of the system in Table E3. [Note: The correspondence with the natural numbers series is visible, on the one hand, through the sequences of two three-member chains (RKH, FYW) and two five-member chains, and on the other hand, through a specific logical (or "logical"?) square: 00 D, 01 N, 10 E, 11 Q. The chemical sense is this: the source molecule D with three derivatives; in terms of logical dimensions of the square, the first derivative is the E; then both acids generate their own amide derivatives (on the second logical dimension of square), N and Q.]

					a	b	c	d	M
D	N	A	L	→	**189**	**189**	**221**	**221 + 3**	485.49 = ***485***
R	F	P	I	→	*289*	*289*	*341*	*341+0*	*585.70=586*
K	Y	T	M	→	299	299	351	351 + 2	*595.71=596*
H	W	S	C	→	*289*	*289*	*331*	*331+1*	*585.64=586*
E	Q	G	V	→	***189***	***189***	***221***	***221 + 3***	485.50 = ***485***
					1255	1255	1465	1465 + 9	2738.04

Table E5: This table corresponds to Table 1 in our previous paper (Rakočević, 2004a); a. number of nucleons within twenty amino acid molecules (side chains), calculated after the first, i.e. the lightest nuclides (H = 1, C = 12, N = 14, O = 16, S = 32); b. the same as "a"; c. number of nucleons, calculated according to the nuclides with lowest abundance in nature (H = 2, C = 13, N = 15, O = 17, S = 36); d. number of nucleons, calculated according to the latest, i.e. heaviest nuclides (H = 2, C = 13, N = 15, O = 18, S = 36); M. molecular mass. It should be noted that for nucleon number (within all nuclides) as well as for molecule mass, two principles are valid: the principle of continuity and unit change principle (e4 - e6). That means that other possible balances of nucleon number as well as molecule mass, in accordance with a-b-c-b-a model, can not save neither the neighborhood of pair-members nor the validity of these two principles. By this, the pairs are as in upper row of Solution (e3). In the bottom row there are the pairs as in the linear arrangement of the system consisting of four diversity types, shown in Figure 2.

(GV, PI); (SC, TM, AL, DE [DN, EQ], NQ, KR, HW, FY)

(GV, PI), (AL, CM, FY / WH, RK, QN, ED, TS) (e3)

$189 - {}_{100} - 289 - {}_{10} - 299 - {}_{10} - 289 - {}_{100} - 189$ (e4)

$221 - {}_{110} - 331 - {}_{10} - 341 - {}_{10} - 351$ (e5)

$485 - {}_{101} - 586 - {}_{10} - 596 - {}_{101} - 586 - {}_{101} - 485$ (e6)

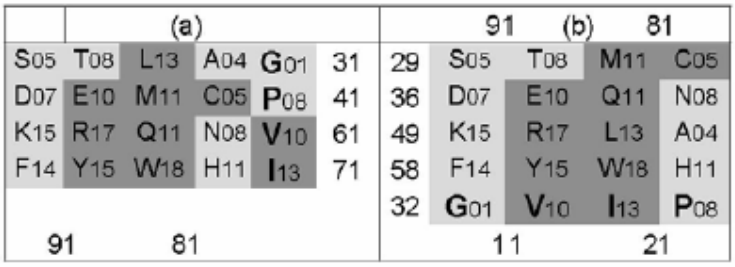


Table E6: Distribution of amino acids according to the Gaussian algorithm. According to an anecdote the small Gauss (only 9 years of age) summed up all the numbers from 1 to 100. If, however, he summed from 1 to 101, and then selected all the tens at the top to get the numbers 11, 21, 31, 41 – 61, 71, 81 and 91, then he got the patterns that are now found in the genetic code (Rakočević, 2009). The pairs as in upper row of Solution (e7). In the bottom row of Solution (e7) there are the pairs as in the linear arrangement of the system consisting of four diversity types, shown in Figure 2.

(GV, PI); (AL, CM, NQ, HW, FY, KR, DE, ST)

(e7)

(GV, PI); (AL, CM, FY / WH, RK, QN, ED, TS)

Appendix F: *Harmonic structures from original works (III)*

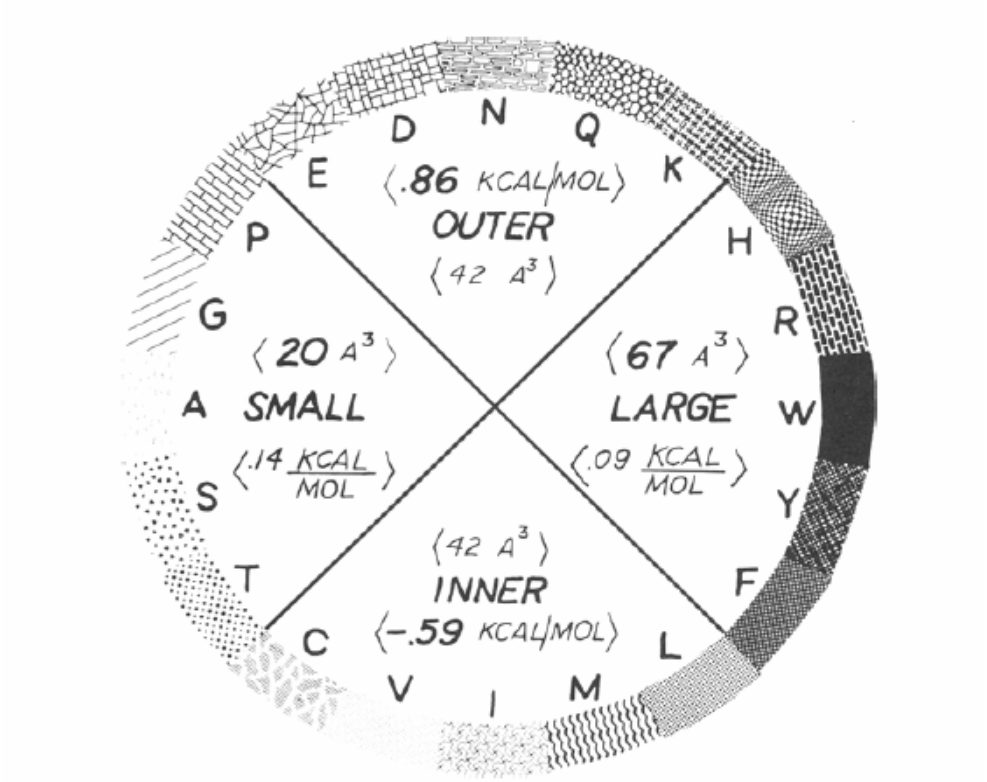


Figure 3. Mutation ring. As the codon ring expresses the minimum change relation among codons, so the mutation ring expresses the minimum change relation among the amino acids. The mutation ring shows the broader relationships among the amino acids, as well as the detailed ordering. For this the ring is quartered into the four groups, and for each group the average values of energy and volume are entered. The amino acids are marked by their one-letter codes and textured patterns. The patterns represent progressions in the physical properties of the amino acids. Dark tones are for large residues, light for small. Coarse, checkered or blotchy textures signify external residues, and smooth, delicate or even-textured patterns designate internal.

Figure F1. This is Figure 3 in (Swanson, 1984, p. 191)

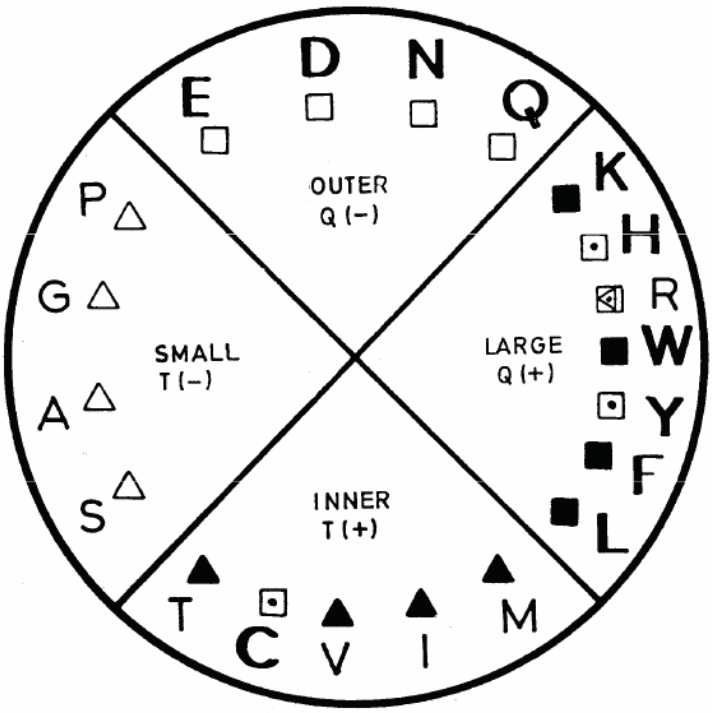


Figure 6 *Mutation ring II. This Ring could be regarded the Mutation ring II provided that R. Swanson's Mutation Ring (Swanson, 1984, Fig. 2) is regarded the Mutation Ring I; Everything is the same as on Mutation Ring I, only the S.T.-Q.K. line is shifted by one step on both ends in relation to Mutation Ring I; and P.E.-M.L. line is shifted only on one (the other) end. The squares designate the amino acids from Space-4 and triangles designate the amino acids from Space-3. The empty squares and empty triangles designate the nonessential amino acids, otherwise they designate the essential amino acids; the dots designate the semi-essential amino acids. The lines strictly separate non-essential from yes-essential amino acids; then the lines strictly separate the Space-3 amino acids from Space-4 amino acids. There are the two exceptions: C is full-strayed; R is semi-strayed. One should note that the complementarity principle is applied as follows: outer-inner: non-essential amino acids from Space-4 are complementary with the essential amino acids from Space-3, etc.*

Figure F2. This is Figure 6 in (Rakočević, 1997, p. 28). Space-3 and Space-4 are from the Boolean cube; Space-3: (7) SRG, **(3) TA**, (2) IMV, (1) SP; Space-4: (6) NKDE, (5) CWR, **(4) YHQ**, (0) FL (Rakočević, 1998, Fig. 1, p. 284).

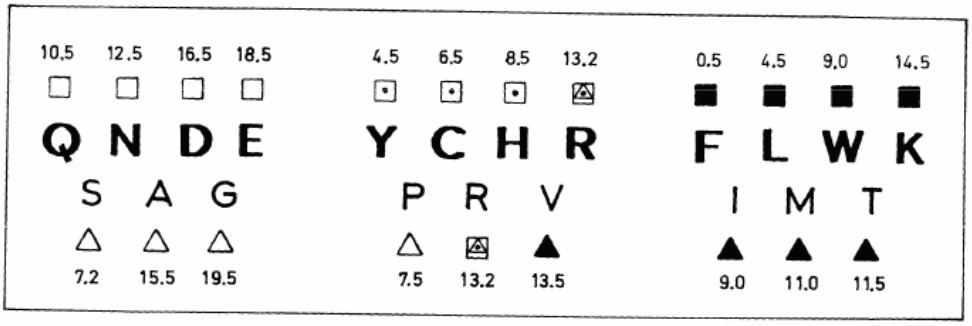


Figure F3. This is Figure 5 in (Rakočević, 1997, p. 27).Non-essential, semi-essential and essential protein amino acids, respectively. The order through the binary values calculated by models presented in Figures F5 and F6. The "reading" of the pairs, in relation to Mutation ring in Figure F2, is presented in Figure F4. Notice, for example the sequence QNDEPGAS**T**/VIM**T** in both Figures (F3 and F4).

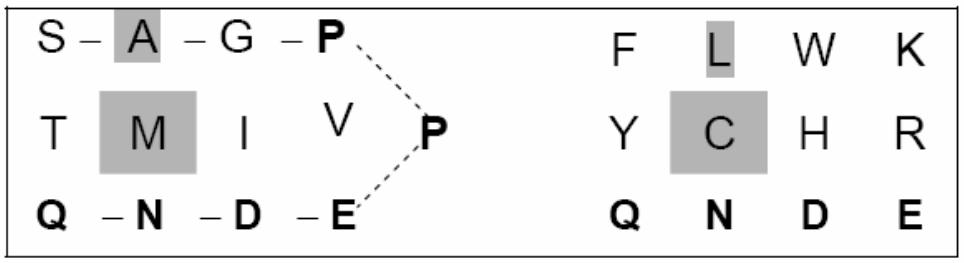


Figure F4. The "reading" of the pairs positions in the system presented in Figure F3. The pairs as in second (bottom) row of Solution (f1). In the first (upper) row of Solution (f1) there are the pairs as in the linear arrangement of the system consisting of four diversity types, shown in Figure 2.

1 2 3 4 5 V IV III II I
GP/AL, VI/CM, **FY/WH**, RK/QN, ED/TS

(f1)

G-P/A-L, VI/C-M, **FY/WH**, RK/QN, ED/TS

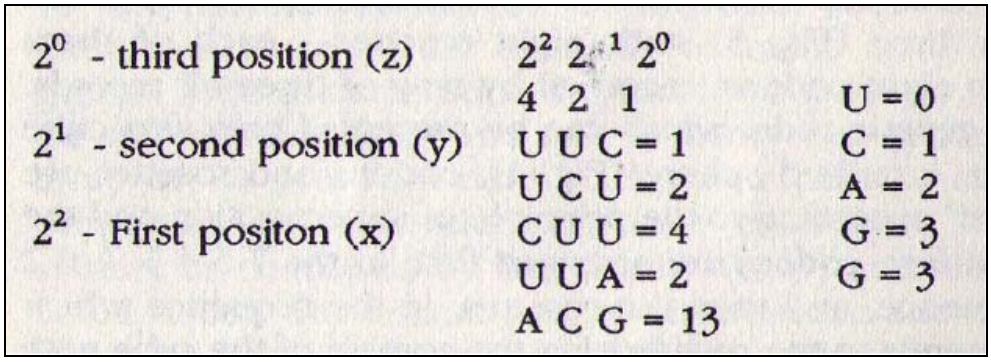


Figure F5. The method of calculating the binary values of the codons (Rakočević, 1988)

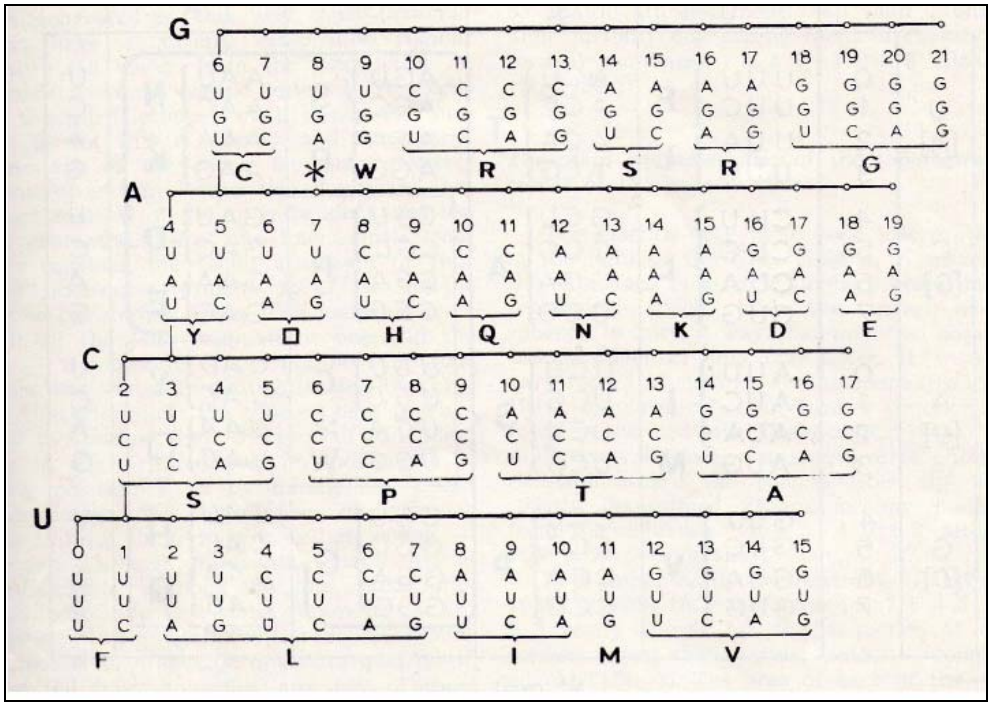


Figure F6. The calculated binary values of the codons (Rakočević, 1988).

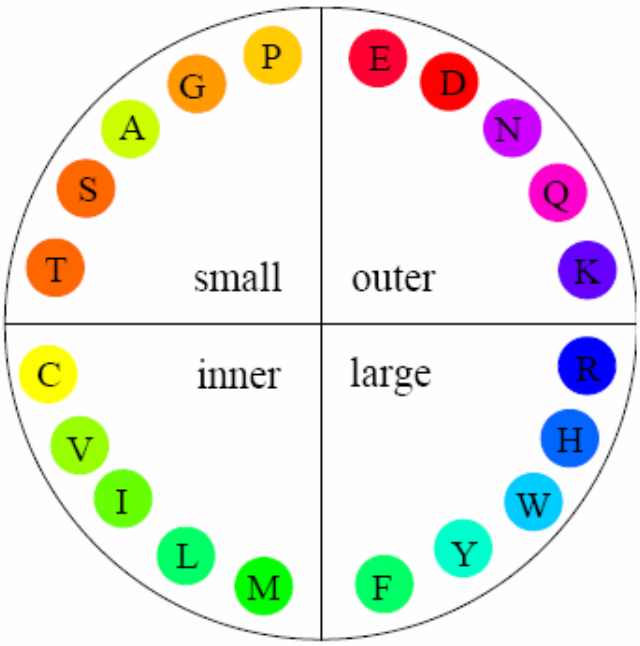


Fig. 4. Representation of the PAM matrix. This projection of the matrix by multidimensional scaling is an idealization adapted from French and Robson (1983) by Taylor (1986). The vertical axis of the circle corresponds to hydrophobicity, and consequently to whether the amino acid is mostly found in the inner or outer parts of proteins, and the horizontal axis corresponds to the molecular volume (small or large) of the amino acid. Amino acids that are close together exchange frequently. Colours used are those proposed by Taylor (1997).

Figure F7. This is Figure 4 in (Kosiol, 2004, p. 28). The pairs as in second (bottom) row of Solution (f2). In the first (upper) row of Solution (f2) there are the pairs as in the linear arrangement of the system consisting of four diversity types, shown in Figure 2.

1 2 3 4 5 V IV III II I
GP/AL, VI/CM, **FY/WH**, RK/QN, ED/TS

(f2)

GP/A-L, VI/C-M, **FY/WH**, RK/QN, ED/TS

Appendix G: *Periodic system of chemical elements*

	III	IV	V	VI	VII	0
					$_{1}$**H**	$_{2}$He
…	$_{5}$B	$_{6}$**C**	$_{7}$**N**	$_{8}$**O**	$_{9}$F	$_{10}$Ne
	$_{13}$Al	$_{14}$Si	$_{15}$**P**	$_{16}$**S**	$_{17}$Cl	$_{18}$Ar
	$_{31}$Ga	$_{32}$Ge	$_{33}$**As**	$_{34}$**Se**	$_{35}$Br	$_{36}$Kr
	$_{49}$In	$_{50}$Sn	$_{51}$**Sb**	$_{52}$**Te**	$_{53}$I	$_{54}$Xe
	…					

Table G1. A part of the periodic system of chemical elements. Color indicates the possible vertically directed neighborhood from the aspects of participation in the construction of the constituents of the genetic code. [Sulfur is involved in the construction of two sulfur amino acids; under nitrogen phosphorus as builder of DNA and RNA. Blue marks a next step: selenocysteine in the set of protein amino acids; arsenic instead phosphorus in DNA and RNA of some few very specific organisms that live in extreme conditions. ["Dec. 2, 2010: NASA-supported researchers have discovered the first known microorganism on Earth able to thrive and reproduce using the toxic chemical arsenic. The microorganism, which lives in California's Mono Lake, substitutes arsenic for phosphorus in the backbone of its DNA and other cellular components." (www.science.nasa.gov)] Finally, the color green – a hypothesis for organisms in the maximum possible extreme environments. (Note: hydrogen can not be in the first group but only in seventh, as a neighbor of helium; *see* next two tables, G2 and G3.)

rows	periods	groups	I			II			III			IV			V			VI			VII			0	VIII		IX		X		XI	XII	XIII	XIV
			a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	b	c	b	c	c	c	c	c
1	1	a																			H 1			He 2										
2	2	a	Li 3			Be 4			B 5			C 6			N 7			O 8			F 9			Ne 10										
3	3	a	Na 11			Mg 12			Al 13			Si 14			P 15			S 16			Cl 17			Ar 18										
4	4	a	K 19			Ca 20			b	Sc 21			Ti 22			V 23			Cr 24			Mn 25				Fe 26		Co 27		Ni 28				
5		b		Cu 29			Zn 30		a	Ga 31		Ge 32			As 33			Se 34			Br 35			Kr 36										
6	5	a	Rb 37			Sr 38			b	Y 39			Zr 40			Nb 41			Mo 42			Tc 43				Ru 44		Rh 45		Pd 46				
7		b		Ag 47			Cd 48		a	In 49		Sn 50			Sb 51			Te 52			I 53			Xe 54										
8	6	a	Cs 55			Ba 56			b	La 57		c	Ce 58			Pr 59			Nd 60			Pm 61				Sm 62		Eu 63		Gd 64		Tb 65	Dy 66	Ho 67	Er 68
9		c			Tm 69			Yb 70	c		Lu 71	b	Hf 72			Ta 73			W 74			Re 75				Os 76		Ir 77		Pt 78				
10		b		Au 79			Hg 80		a	Tl 81		a	Pb 82			Bi 83			Po 84			At 85			Rn 86										
	7	a	Fr 87			Ra 88			b	Ac 89		c	Th 90			Pa 91			U 92			Np 93				Pu 94		Am 95		Cm 96		Bk 97	Cf 98	Es 99	Fm 100
		c			Md 101			No 102	a		Lr 103	a	Ku 104			Ns 105			106			107				108		109		110					
		b																																	

Table G2. For [(s & p), d, f] elements, to the border of stability ($_{84}$Po), we have **8** times the pattern 5-3-1; then **2** times 0-3-1 and **4** times 0-0-1. All together: 9-4-1 elements. [8-4-2 and 9-4-1 as $2^{\mathbf{3}}$-$2^{\mathbf{2}}$-$2^{\mathbf{1}}$ and $\mathbf{3}^2$-$\mathbf{2}^2$-$\mathbf{1}^2$, respectively] The first three periods are single, each with a single row; fourth and fifth are doubles, each with two rows; sixth period is threefold, has three rows [Table 4.2 in (Rakočević, 1991), or Table 18 in (Rakočević, 1997)].

rows	periods	groups	I a	II a	III b	IV c	V c	VI c	VII c	VIII c	IX c	X c	XI c	XII c	XIII c	XIV c	I c	II c	III c	IV b	V b	VI b	VII b	VIII b	IX b	X b	I b	II b	III a	IV a	V a	VI a	VII a	0 a
1	1	a																														H 1	He 2
2	2	a	Li 3	Be 4																								B 5	C 6	N 7	O 8	F 9	Ne 10
3	3	a	Na 11	Mg 12																								Al 13	Si 14	P 15	S 16	Cl 17	Ar 18
4	4	a	K 19	Ca 20	Sc 21															Ti 22	V 23	Cr 24	Mn 25	Fe 26	Co 27	Ni 28	Cu 29	Zn 30	Ga 31	Ge 32	As 33	Se 34	Br 35	Kr 36
5		b																															
6	5	a	Rb 37	Sr 38	Y 39															Zr 40	Nb 41	Mo 42	Tc 43	Ru 44	Rh 45	Pd 46	Ag 47	Cd 48	In 49	Sn 50	Sb 51	Te 52	I 53	Xe 54
7		b																															
8	6	a	Cs 55	Ba 56	La 57															Hf 72	Ta 73	W 74	Re 75	Os 76	Ir 77	Pt 78	Au 79	Hg 80	Tl 81	Pb 82	Bi 83	Po 84	At 85	Rn 86
9		b				Ce 58	Pr 59	Nd 60	Pm 61	Sm 62	Eu 63	Gd 64	Tb 65	Dy 66	Ho 67	Er 68	Tm 69	Yb 70	Lu 71															
10		c																															
	7	a	Fr 87	Ra 88	Ac 89															Ku 104	Ns 105	106	107	108	109	110	111	112	113	114	115	116	117	118
		b				Th 90	Pa 91	U 92	Np 93	Pu 94	Am 95	Cm 96	Bk 97	Cf 98	Es 99	Fm 100	Md 101	No 102	Lr 103															
		c																															

Tabela G3: Table of long periods which follows from Table G2 [Table 4.3 in (Rakočević, 1991), or Table 19 in (Rakočević, 1997)].

	DIADS																							
					TRIADS																MONAD			
	I				I				II				III				II							
1	1	H	(2)	VII	2	He	(2)	VIII	3	Li	(2)	I	4	Be	(1)	II	5	B	(2)	III	6	C	(2)	IV
2	7	N	(2)	V	8	O	(3)	VI	9	F	(1)	VII	10	Ne	(3)	VIII	11	Na	(1)	I	12	Mg	(3)	II
3	13	Al	(1)	III	14	Si	(3)	IV	15	P	(1)	V	16	S	(4)	VI	17	Cl	(2)	VII	18	Ar	(3)	VIII
4	19	K	(3)	I	20	Ca	(6)	II	21	Sc	(1)	III	22	Ti	(5)	IV	23	V	(2)	V	24	Cr	(4)	VI
5	25	Mn	(1)	VII	26	Fe	(4)	VIII	27	Co	(1)	IX	28	Ni	(5)	X	29	Cu	(2)	I	30	Zn	(5)	II
6	31	Ga	(2)	III	32	Ge	(5)	IV	33	As	(1)	V	34	Se	(6)	VI	35	Br	(2)	VII	36	Kr	(6)	VIII
7	37	Rb	(2)	I	38	Sr	(4)	II	39	Y	(1)	III	40	Zr	(5)	IV	41	Nb	(1)	V	42	Mo	(7)	VI
8	43	Te	(0)	VII	44	Ru	(7)	VIII	45	Rh	(1)	IX	46	Pd	(6)	X	47	Ag	(2)	I	48	Cd	(8)	II
9	49	In	(2)	III	50	Sn	(10)	IV	51	Sb	(2)	V	52	Te	(8)	VI	53	I	(1)	VII	54	Xe	(9)	VIII
10	55	Cs	(1)	I	56	Ba	(7)	II	57	La	(2)	III	58	Ce	(4)	IV	59	Pr	(1)	V	60	Nd	(7)	VI
11	61	Pm	(0)	VII	62	Sm	(7)	VIII	63	Eu	(2)	IX	64	Gd	(7)	X	65	Tb	(1)	XI	66	Dy	(7)	XII
12	67	Ho	(1)	XIII	68	Er	(6)	XIV	69	Tm	(1)	I	70	Yb	(7)	II	71	Lu	(2)	III	72	Hf	(6)	IV
13	73	Ta	(2)	V	74	W	(5)	VI	75	Re	(2)	VII	76	Os	(7)	VIII	77	Ir	(2)	IX	78	Pt	(6)	X
14	79	Au	(1)	I	80	Hg	(7)	II	81	Tl	(2)	III	82	Pb	(4)	IV	83	Bi	(1)	V	84	Po	(5)	VI
15	85	At		VII	86	Rn		VIII	87	Fr		I	88	Ra		II	89	Ac		III	90	Th		IV
16	91	Pa		V	92	U		VI	93	Np		VII	94	Pu		VIII	95	Am		IX	96	Cm		X
17	97	Bk		XI	98	Cf		XII	99	Es		XIII	100	Fm		XIV	101	Md		I	102	No		II
18	103	Lr		III	104	Ku		IV	105	Ns		V	106			VI	107			VII	108			VIII
19	109			IX	110			X	111			I	112			II	113			III	114			IV

Table G4. If the elements are arranged „six pro period“, then it is reached the boundary of stability/instability, the 84-element, Polonium. This table is very like to the first Mendeleyev's Table which he sent to the most famous chemists of the world [the first and second photocopies in Mendeleyev's Archive, presented in Kedrov, 1977 (16 photocopies between p. 128 and p. 129)]. The classification into monads, the two groups of dyads and three of triads is determined by the number of isotopes, through the validity of two principles: the principle of continuity and minimum change principle.

HARMONIC MEAN AS A DETERMINANT OF THE GENETIC CODE

Miloje M. Rakočević

Department of Chemistry, Faculty of Science, University of Niš, Serbia
(E-mail: milemirkov@open.telekom.rs; www.rakocevcode.rs)

Abstract. It is shown that there is a sense in splitting Genetic Code Table (GCT) into three parts using the harmonic mean, calculated by the formula H (a, b) = 2ab / (a + b), where a = 63 and b = 31.5. Within these three parts, the amino acids (AAs) are positioned on the basis of the validity of the evident regularities of key parameters, such as polarity, hydrophobicity and enzyme-mediated amino acid classification. In addition, there are obvious balances of the number of atoms in the nucleotide triplets and corresponding amino acid groups and/or classes.

1. Introduction

In a previous work we showed (on the model of a binary tree, 000000-111111, that is 0 - 63) that the *golden mean* is a characteristic determinant of the genetic code (GC) (Rakočević, 1998), and in this paper we will show that the same is valid for *the harmonic mean*, if Genetic Code Table (GCT) is divided into three equal[1] parts and if the ordinal codon number is the same as in R. Swanson's concept (Swanson, 1984; Rakočević, 1998, Fig. 1, p. 284).

According to the concept of R. Swanson, the order follows this sequence through the ordinal numbers: YUN, YCN, RUN, RCN / YAN, YGN, RAN, RGN. [Perhaps Negadi's order (UUN, UCN, …, GAN, GGN) (Negadi, 2009), as a sub-variant, corresponds with this order, but that is a question for further researches.] However, as we can see from GCT, the other (chemical) possibility is: NUN, NCN / NAN, NGN in accordance with the ordinal numbers 0-15, 16-31, 32-47 and 48-63 within four columns. In this paper it will become evident that this second order does not allow a chemically adequate division into three harmonic parts. (Note: Y for pYrimidine, R for puRine and N for aNy of four bases.)

In fact, two concepts of R. Swanson are taken into account - Gray Code Model of the genetic code as the first, and the Codon Path Cube as the second one. Both models start with binary values of nucleotide bases: U = 00, C = 01, A = 10 and G = 11. In the first concept, R. Swanson showed that codon binary values (and the codon ordinal number) are derived from the binary values of bases, from zeroth codon UUU = 000000 to the last one GGG = 111111; and we have shown that the same records can be applied to the binary-code tree (Rakočević, 1998), as well as in the GCT as the ordinal number of the codons (Table 1 and Appendix A).

On the other hand, we have also shown (Rakočević, 1988) that on the Codon Path Cube a special kind of quantitative "weight" of codons can be calculated, following R. Swanson's idea that "the three edges of the cubes represent the three positions in a codon" (Swanson, 1984, p. 189). According to these "weights", calculated on the basis of the binary numeral system (q = 2), codons are arranged in four "floors" of GCT with the following ordinal numbers: 0-15 in the first column, 2-17 in the second, 4-19 in the third and 6-21 in the fourth

[1] The codon space is made of 64 codons with 63 intervals between them (Table 1). When we say "three equal parts" that means 63 : 3 = 21. In the first part there are 22 codons, from zeroth UUU to 21st, GUC. In the second part there are 21 codons, from 22nd, GUA to the 42nd, UGA. Finally, within the third part there are also 21 codons, from 43rd, UGG to the last GGG.

column. For better understanding, we provide an example of calculation for the last codon, the codon GGG: $(3 \times 2^2) + (3 \times 2^1) + (3 \times 2^0) = 21$ (Rakočević, 1988). In this way, the number 21 appears to be a specific and important determinant of the genetic code one more time.

1st lett.	2nd letter								3rd lett.
	U		C		A		G		
U	00. UUU 01. UUC 02. UUA 03. UUG	**F**-II L-I	08. UCU 09. UCC 10. UCA 11. UCG	**S**-II	32. UAU 33. UAC ~~34. UAA~~ ~~35.UAG~~	**Y**-I CT	40. UGU 41. UGC ~~42. UGA~~ 43. UGG	**C**-I CT **W**-I	*U* *C* *A* *G*
C	04. CUU 05. CUC 06. CUA 07. CUG	 L-I	12. CCU 13. CCC 14. CCA 15. CCG	**P**-II	36. CAU 37. CAC 38. CAA 39. CAG	H-II Q-I	44. CGU 45. CGC 46. CGA 47. CGG	 R-I	*U* *C* *A* *G*
A	16. AUU 17. AUC 18. AUA 19. AUG	 I-I M-I	24. ACU 25. ACC 26. ACA 27. ACG	T-II	~~48. AAU~~ ~~49. AAC~~ ~~50. AAA~~ 51. AAG	N-II K-II	56. AGU 57. AGC 58. AGA 59. AGG	S-II R-I	U C A G
G	20. GUU 21. GUC 22. GUA 23. GUG	V-I **V**-I	28. GCU 29. GCC 30. GCA 31. GCG	**A**-II	52. GAU 53. GAC 54. GAA 55. GAG	D-II **E**-I	60. GGU 61. GGC 62. GGA 63. GGG	 **G**-II	*U* *C* *A* *G*

Table 1. The Table of the standard genetic code (GCT). Total codon space is divided into three parts in correspondence with the harmonic mean (H) of the whole codon space sequence (a) and its half (b), where a = 63, b = 31.5 and H = 42. Bold and underlined amino acids are those that are at the end of a sequence (outer AAs), and they are different from most AAs in the sequence. In the central area, the three stop codons (CT, codon terminations) are crossed out.

2. Polarity of amino acids

As there are 63 intervals between 64 points (64 codons), the harmonic mean, calculated by the formula H (a, b) = 2ab / (a + b), (a = 63, b = 31.5), is positioned on the number 42, i.e. on the position of the UGA stop codon. This enables us to formulate *a working hypothesis*, according to which, for better understanding of the relations within the GCT, it is necessary to divide the total codon space into three parts: 0-21, 22-42 and 43-63. This proves that the relation between the number 21 and number 42 (the 42 as harmonic mean, H = 42 in the sequence 0-63) appears to be a realization of "the symmetry in the simplest case" as the ratio 1:2 (Marcus, 1989, p. 103). In Table 1, the first part and the third part of the space are presented in dark tones and the second one in lighter tones.

Indeed, only through such a demarcation of the codon space, can we perceive a characteristic inter-relation between the arrangements of polar and non-polar amino acids (AAs). [Polarity is here taken as hydropathy (Kyte & Doolittle, 1982); more precisely, as the hydropathy index: non-polar amino acids are marked by positive and polar by negative index values (Remark 1).] So, we first perceive that on the right side (in the third part of codon space), all AAs are polar. However, on the left side (the first shaded area) both are present – nonpolar and polar AAs, but in an order imposed by a specific logic. There is, in fact, a spatial sequence (independent of the ordinal numbers of the codons) of non-polar AAs, and only at the end of the sequence (at only one end!), two polar AAs (S and P) have been "added":

[(V–M–I–L–F) – (**S**–**P**)]. (In a linear reading within a linear sequence: all AAs in the first column of Table 1 are non-polar, and first two AAs in second column are polar.) In addition, from a chemical point of view, the added AAs are "the first possible cases": serine as the first possible oxygen derivative, and proline as the first possible cyclic molecule (iso-propyl group bound, by two ending carbon atoms, to the amino acid functional group).

Remark 1. The hydropathy index: R = –4.5, K = –3.9, D = –3.5, E = –3.5, N = –3.5, Q = –3.5, H = –3.2, P = –1.6, Y = –1.3, W = –0.9, S = –0.8, T = –0.7, G = –0.4; A = + 1.8, M = + 1.9, C = + 2.5, F = + 2.8, L = + 3.8, V = + 4.2, I = + 4.5 (Kyte & Doolittle, 1982).

The same logic, but with the opposite chemical meanings, applies to the central area (marked by lighter tones). Here, there appears a spatial sequence of polar AAs and in this sequence the three non-polar AAs are also added at the ends (at both ends!); V and A at the beginning, and C at the end of the spatial sequence: [(**V–A**) – (T–Q–H–Y) – (**C**)]. In addition, from a chemical point of view, the added AAs here are also "the first possible cases": valine as the first possible semi-cyclic molecule (cyclo-propyl group bound, by the central carbon atom, to the amino acid functional group), alanine as the first possible carbon derivative, and cysteine as the first possible sulfur derivative. [Carbon, oxygen and sulfur, as derivatives in relation to glycine as the simplest possible amino acid: the substitution of H atom with corresponding functional groups (CH3, OH, SH), respectively. Only nitrogen has remained, as a possible derivative. However, the first possible nitrogen derivative (– CH2–NH2) is not a constituent of the genetic code, but the fourth, in the form of lysine (K) molecule. (For more details of the relationship between valine and proline, in terms of the binding of the isopropyl group, see: Rakočević & Jokić, 1996).]

In connection with polarity of the AAs it is necessary to see their positions in the GCT also from the aspect of "cloister energy" (Swanson, 1984), as well as from the aspect of "polar requirement" (Woese, 1966; Konopel'chenko & Rumer, 1975). Thus, from the aspect of cloister energy, glycine (G) and tryptophan (W) are not polar, but nonpolar AAs. [Glycine (G) is nonpolar also from the aspect of polar requirement.] According to one kind of spatial reading of amino acid sequence (a "diagonal" reading with the last amino acid G on the right in Table 1), in the second shaded area, such a status is expected; expected from the aspect of "sequence logic", presented above. Namely, a possible reading through a spatial sequence is as follows: **W**-(R-N)-(S-K)-R(D-E)-**G**. In such a case, nonpolar AAs are at the ends – at both ends; more exactly, at the beginning and at the end. [At the 43rd codon position there is a nonpolar AA (W); at the 44-59th positions there are polar AAs and at the 60-63rd positions there is a nonpolar AA (G).] In addition, from a chemical point of view the AAs added at the ends are two extreme cases: glycine, as the one and only amino acid without carbon atom in the side chain; tryptophan (W), as the one and only amino acid with two aromatic rings in the side chain.

As for "polar requirement", not only glycine but proline as well is nonpolar and at the same time an extreme amino acid: one and only alicyclic molecule within the set of 20 AAs; as such it is added at the end of the sequence [(V–M–I–L–F) – (**S**–**P**)] within the first shaded area.

3. Two classes of AAs handled by two classes of enzymes

The same logic (or similar logic) of "adding at the end" is also valid for the arrangement of AAs, classified into two classes, corresponding to two classes of enzymes aminoacyl-tRNA synthetases (Tables 2.1 & 2.2 in relation to Tables 3.1 & 3.2). It can be said that the

corresponding analysis, which was once presented by R. Wetzel (Wetzel, 1989; Rakočević, 1997a), is now shown in a new, clearer light. Bearing this in mind, one must notice that the distribution of AAs into two classes, consistent with enzyme-synthetases classes, corresponds to the hierarchy of their molecule sizes, without any exception (*m* and *v* in Table 2.1): within 10 amino acid pairs, larger molecules (+) belong to class I and the smaller (–) to the class II. On the other hand, two systems – "20" (Table 2.2) and "61" (Table 3.2) amino acid molecules – appear to be in very symmetrical and balanced relationships (Box 1 & 2).

h	+ 0.825	+ 0.943	+ 0.680	+ 0.043	+ 0.251	+ 0.738	+ 0.943	– **0.000**	+ 0.878	– 0.880
v	+ 36	+ 46	+ 30	+ 41	+ 47	+ 52	+ 46	+ 70	+ 83	+ 69
m	+ 117.15	+ 131.18	+ 121.16	+ 147.13	+ 146.15	+ 149.21	+ 131.18	+ 174.20	+ 204.23	+ 181.19
I	V_{10}	L_{13}	C_{05}	E_{10}	Q_{11}	M_{11}	I_{13}	R_{17}	W_{18}	Y_{15}
II	G_{01}	A_{04}	S_{05}	D_{07}	N_{08}	T_{08}	P_{08}	K_{15}	H_{11}	F_{14}
m	75.07 –	89.09 –	105.09 –	133.10 –	132.12 –	119.12 –	115.13 –	146.19 –	155.16 –	165.19 –
v	03 –	14 –	21 –	30 –	36 –	32 –	31 –	58 –	50 –	62 –
h	0.501 –	0.616 –	0.359 –	0.028 –	0.236 –	0.450 –	0.711 –	0.283 +	0.165 –	**1.000** +

Table 2.1. Two classes of amino acids handled by two classes of enzymes. (Class II with 81 and Class I with 123 atoms.) The ten amino acid pairs, *natural* pairs from the chemical aspect, are classified into two classes. Class I contains larger amino acids (larger within the pairs), all handled by class I of enzymes aminoacyl-tRNA synthetases. Class II contains smaller amino acids, all handled by class II of synthetases. Within the rows the given values for the parameters are the following: molecule mass (m), volume of amino acid molecules in angstroms (v) and hydrophobicity (h) on a natural scale (0-1). [Molecule mass and hydrophobicity after: Black & Mould, 1991; volume after: Swanson, 1984.] The order follows the number of atoms within side chains of class II AAs (given here as index); from left to right: first there are aliphatic, and then aromatic AAs. The largest aliphatic molecule of class I (R) is of a zeroth hydrophobicity (h = 0), while the smaller aromatic molecule, F (smaller within the pair F-Y) is of a maximal value (h = 1). [Notice that the pair F-Y is simpler as only aromatic and H-W is more complex as aromatic heterocyclic.]

As we can see, in the first area of the codon space there are AAs from the first class, with three AAs from the second class, added at the end (at one end): [(V–M–I–L) – (**F–S–P**)]; all three as the first possible cases: phenylalanine as the first possible aromatic derivative[2], serine and proline in the manner as we explained above. The opposite situation applies to the right side (second shaded area). As added (at both ends) there are three AAs of the first class, and AAs of the second class are between them; all together: {(**W**, **R**) – [(S, N), (K, R), (G, D)] – **E**)}. (There is also a spatial "diagonal" reading here, but with the last amino acid **E** on the left in Table 1, instead of G as was the case in polarity, in the sequence, shown above.) Thus, all three added AAs are – "the last cases", i.e. the extreme cases: tryptophan as the only amino acid with two rings (within the set of four aromatic AAs), arginine as more complex in the set of two amino derivatives (K, R), and glutamic acid as more complex in the set of two carboxylic AAs (D, E). [This logic is valid for all ten amino acid pairs: more complex (larger) molecule is handled by the first, while the simpler one - by the second class of aminoacyl-tRNA synthetases.]

[2] If we know that phenylalanine belongs to a set of 16 amino acids of alanine stereochemical type (with a CH_2 group between amino acid "head" and "body"), it becomes clear why a toluene and not benzene derivative must be 'the first possible case'. [About four stereochemical types see: (Popov, 1989; Rakočević and Jokić, 1996).]

Atom number of 20 AAs (AAs positions in Table 2.1)			
Class	Odd	Even	Sum
I	57	66	→ 123
II	33	48	→ 81 (9 × 9)
	90	114	→ (10 × 9), (123 – 9)
	Odd	Even	Sum
I	349	412	→ 761 (661+100)
II	212	282	→ 494 (594-100)
	561	694	→ (661-100), (594+100)
(661 – 628 = 33), (627 – 594 = 33)			
628+627 = 1255 (total nucl. numb. in side chains)			
Nucleon number of 20 AAs			

Table 2.2. Balances of the number of particles within 20 AAs. The balances are given here concerning the relations between parameters in Table 2.1. For atom number: being 9 × 9 in 10 molecules of class II and 10 × 9 at odd positions for both classes; there are differences of 9 atoms etc. For nucleon number: through a specific balance at odd/even positions the difference of 33 atoms appears in relation to middle value (627-628); one balance more: the relations to the number 594 are given and this number also represents the number of atoms within 61 amino acid molecules (cf. Table 3.2).

a	+	+	–	+	+	–	+	+	–	+
	40	78	10	20	22	11	39	102	18	30
I	$4V_{10}$	$6L_{13}$	$2C_{05}$	$2E_{10}$	$2Q_{11}$	$1M_{11}$	$3I_{13}$	$6R_{17}$	$1W_{18}$	$2Y_{15}$
II	$4G_{01}$	$4A_{04}$	$6S_{05}$	$2D_{07}$	$2N_{08}$	$4T_{08}$	$4P_{08}$	$2K_{15}$	$2H_{11}$	$2F_{14}$
a	4	16	30	14	16	32	32	30	22	28
	–	–	+	–	–	+	–	–	+	–

Table 3. 1. Atom number within 61 amino acid molecules. Two classes of AAs as in Table 2.1, except that they are relating only to the number of atoms; (**a**) Number of atoms within codon corresponding 61 amino acid molecules; the number of atoms within amino acid side chains, per "sets" (V = 4, L = 6 etc.), is such that (+/–) designations, i.e. quantities, follow a logic of the strict symmetry: one minus follows two pluses, then two pluses follow one minus etc., for aliphatic AAs; and one plus follows one minus for aromatic AAs.

Through two positions in third part of Table 1, the arginine (R) appears to be the only exception, which "spoils" the presented logic. Arginine also appears as such an exception in the same way as in the Codon Path Cube, as we have previously shown. Namely, in Codon Path Cube two classes of AAs, handled by two classes of enzymes aminoacyl-tRNA synthetases, are splitting into two separated areas, with only one exception, the exception of arginine (Rakočević, 1997a).

In the second part of codon space (in the central codon-space area) the logic is as follows: in the lower part of the space, in the sequence: T, A, **V**, the "different" molecule is the last one (**V**), whereas in the upper part [C–Y–(**H**)–Q] it is – the first to the last (**H**). As we know, this amino acid is also a special case with the value H = 0 in cloister energy. From both these specificities follows that Y – C are at the end of whole amino acid sequence.

Molecule number				
Class	Odd	Even		
I	12	17	→	**29**
II	18	14	→	**32**
	30	**31**	(29, 30, 31, 32)	
Class	Odd	Even		
I	129	241	→	**370**
II	104	120	→	224
	233	361	→	**594**
Atom number				

Table 3. 2. Balances of the number of particles within 61 amino acid molecules, given here concerning the relations between parameters in Table 3.1. The total atom number within 61 amino acid molecule (594) appears to be in relation to the Shcherbak's "Prime Quantum 37". On the other hand, the number of molecules (through odd/even positions) makes a segment of natural numbers series: 29, 30, 31, 32.

Box 1

20 AAs	→	90 (Odd)	–	81 (Class II)	=	9
61 AAs	→	233 (Odd)	–	224 (Class II)	=	9
		110		110		
20 AAs	→	123 (Class I)	–	114 (Even)	=	9
61 AAs	→	370 (Class I)	–	361 (Even)	=	9

Box 2

20 AAs		61 AAs	
90 (Odd)	+	361 (Even)	= 451
81 (Class II)	+	370 (Class I)	= 451
114 (Even)	+	233 (Odd)	= 347
123 (Class I)	+	224 (Class II)	= 347

3A + 3T + D + K + F = 72
3V + 2I + C + Q = 72

L + M + E + R + Y = 66
A + T + D + K + F = 48 (18)
V + I + C + Q = 39
G + P + S + N = 22 (17)

48 + 39 = 87
66 + 22 = 88

H + W = 29
29 -18 = 11
18 – 17 = 01

Box 1. This arrangement presents the relationships between 20 and 61 amino acid molecules in the form of atom number difference in the sets of AAs at odd / even positions and / or within the class I / II, of the systems presented in Tables 2.1 and 3.1.

Box 2. The same (*mutatis mutandis*) arrangement as in Box 1, but through "plus" instead of "minus" operations. Most AAs on the odd/even positions and/or within the class I/II are the same, but some are diferent (thus, the balance and symmetry are realized) as it is presented on the right side. [The different AAs are: A, T, D, K, F from the first class, and V, I, C, Q from the second one.]

4. Arithmetical balances

All presented distinctions, at the same time logical and chemical, are accompanied by the appropriate arithmetical balances (Table 4). Thus, in the first part of the system (the shaded left side) there are 87 atoms in the 8 amino acid molecules, while on the right side there are 98 atoms in 9 amino acid molecules; the difference: for molecules 01 and for atoms 11. In the second (the lighter shaded) part there are 64 atoms: in the middle area (column "A", Table 1) there are 37, while within two ending areas (columns "U-C" and "G" in Table 1) there are 27 atoms; the difference 37 – 27 is 10 (11- 01). [The atom number in column "A": Y15 + H11 + Q11 = 37; in columns "U-C": V10 + A04 + T08 and in column "G": C05; all together equals 27 atoms.]

These results follow if they are calculated only for the amino acid side chains. If, however, the amino acid functional groups (the "heads" of molecules) are taken into account (9 atoms per AA), then instead of 87 and 98 there are 159 and 179 atoms, respectively, which means ± 10 in relation to the arithmetic mean, which is 169. On the other hand, based on the relations to the number of atoms in the second part (127 atoms) and on the differences 159-127 = 2 × 16 and 179-127 = 2 × 26, numbers 16 and 26 appear as the key determinants of the codon distribution (per amino acid) into four classes, based on the four types of amino acid diversity (Rakočević, 2011: Table 4, Eq.4 and Fig. 3 on pp. 826-828; arXiv:1107.1998 [q-bio.OT]).

Besides the balances shown above, there is a balance of the number of atoms in the corresponding codons, and in their nucleotides, but only when the stop codons are excluded (cf. Remark 2). Thus, on the left side there are 2321, whereas on the right side there are 2322 atoms (row "H", Table 4). The difference is exactly one atom (2322 – 2321= 1). In the central area, within the pyrimidine nucleotides (UMP and CMP) there is 971-2, and in the purine ones (AMP and GMP) there is 972 +2 of atoms. A similar balance appears to the number of hydrogen bonds, in the nucleotide bases: there are 161 on the left, and 162 on the right side. In the central area, in the pyrimidine nucleotides, (within their U and C bases) there are 68+5, and 69-5 in the purine type of nucleotides (within their A and G bases).

With the breakdown of the whole codon space into three parts, there inevitably appears one additional amino acid molecule, valine, resulting in the total of 24 amino acids within 16 quadruplets. (Notice that the side chain of valine represents a crossroads between openness and cyclicality.) Only this insight enables us to see new balances, for example, the balance of the number of atoms as well as of isotopes in just 24 molecules, placed into odd-even positions, counted – one at a time or two at a time (Tables 5 & 6).

Remark 2. Within U, C, A, G bases there are 12, 13, 15, 16 atoms, respectively. Within nucleotides: UMP, CMP, AMP, GMP → (12, 13, 15, 16), plus 20 atoms in ribose molecule and 8 atoms in phosphoric acid, minus 6 atoms in 2 water molecules; all together equals: UMP-34, CMP-35, AMP-37, GMP-38. This implies that the two inner as well as the two outer columns of GCT (Table 1) contain 3456 of atoms each (including three stop codons) (cf. Remark 3). Thus, within whole GCT there are 3456 + 3456 = 6912. In order to interpret the three parts, determined by the harmonic mean (Table 1), this sum must be reduced by as much as 326 atoms in the three stop codons: 6912-326 = 6586, and this number is actually the product of the Shcherbak's "Prime Quantum 37" (6586 = 37 × 178). [About Shcherbak's "Prime Quantum 37", see in (Shcherbak, 1994), and its geometrical interpretation, see in (Mišić, 2011).] In relation to the row H, in Table 4, we have: 2321+1943+2322 = 6586 (1943 = 969 + 974).

Remark 3. The "stop" codons are not the same in standard and mitochondrial Genetic codes. On the other hand, the mitochondrial code, in a way, is more symmetrical than the standard one; and thus more adequate, for example, for p-adic mathematics determination (Dragovich & Dragovich, 2010). Hence, in further research it is necessary to check whether it is possible that the mitochondrial genetic code Table – via the harmonic mean – can also be divided into three equal parts; i.e. are we any closer to answer the question which code is more original, and which occurred in the course of evolution.

	I	II	III
A	8	7	9
B	87	64	98
C	159	127	179
D	22	18	21
E	66	54	63
F	50/16	28/26	15/48
G	u 28/9 a c 22/7 g	u 11/14 a c 17/12 g	u 6/21 a c 9/27 g
H	2321	u,c 971-2 972+2 a,g	2322
K	161	u,c 68+5 69-5 a,g	162

Table 4. Distributions and distinctions within three parts of GCT. Roman numerals I, II and III correspond to the three shaded areas in Table 1. A. Number of molecules; B. Number of atoms in the side chains of amino acids; C. Number of atoms in the entire amino acid molecules: amino acid functional groups plus side chains ("head" + "body"); D. Number of codons in the three shaded areas; E. Number of nucleotides; F. Number of pyrimidine/purine nucleotides; G. Number of UMP, CMP, AMP and GMP nucleotides, respectively; H. Number of atoms in nucleotides; K. Number of hydrogen bonds.

F	14	11	H	F	27	22	H
L	13	11	Q	L			Q
L	13	08	N	L	26	23	N
I	13	15	K	I			K
M	11	07	D	M	21	17	D
V	10	10	E	V			E
V	10	05	C	V	15	23	C
S	05	18	W	S			W
P	08	17	R	P	16	22	R
T	08	05	S	T			S
A	04	17	R	A	19	18	R
Y	15	01	G	Y			G
	60	65	**125**		64	61	**125**
	64	60	**124**		60	64	**124**
	124	**125**			**124**	**125**	

Table 5. Number of atoms in the side chains of "24" amino acid molecules. The sequence of amino acid molecules is given in Table 1: on the left side one by one at odd/ even positions; on the right side two by two. The result is the same: 124 and 125 atoms in either horizontal or vertical reading. That means that the number of atoms in the first twelve molecules equals the number of atoms in 12 molecules at even positions (124); and, the number of atoms in the last twelve molecules equals the number of atoms in 12 molecules at odd positions (125).

F	28	22	H	F	54	45	H
L	26	23	Q	L			Q
L	26	17	N	L	52	47	N
I	26	30	K	I			K
M	24	16	D	M	44	38	D
V	20	22	E	V			E
V	20	12	C	V	31	48	C
S	11	36	W	S			W
P	16	34	R	P	33	45	R
T	17	11	S	T			S
A	08	34	R	A	39	36	R
Y	31	02	G	Y			G
	122	135	**257**		131	128	**259**
	131	124	**255**		122	131	**253**
	253	**259**			**253**	**259**	

Table 6. Number of isotopes in side chains of "24" amino acid molecules. Amino acids are counted in the same way as in the previous Table (Table 5). The result shows the following number sequences 253-255-257-259 and 253-259/253-259 of the number of isotopes (in correspondence with continuity and minimum change principle). The number of isotopes is calculated as follows (in the example for serine): [(3 × 2) H] + [(1 × 2) C] + [(1 × 3) O] = 11. (These are the only stable isotopes, i.e. stable nuclides: 2 for hydrogen, 2 for carbon and 3 for oxygen.)

5. Hydrophobicity of amino acids

The next parameter which corresponds to the splitting of the GCT (Table 1) into three parts is hydrophobicity (*h* in Table 2.1). For eight out of ten amino acid pairs everything is the same as in the distribution of AAs in relation to two classes of synthetases. The remaining two pairs, F-Y and K-R, are in a vice versa position (through +/– designations in Table 2.1). The epilogue is that in the shaded area on the left side, there is, again, the same situation as with polarity: [(V–M–I–L–F) – (**S**–**P**)]. Each of the first five amino acids appears to have a higher value for the hydrophobicity within their own pairs. The other two, which are added at the end, have lower values. [In connection with the hydrophobicity, it is important to recall that just this variant of parameter, as a "natural scale" (Black and Mould, 1991), is the best one among a host of other parameters which also measure the hydrophobicity of AAs (Chechetkin and Lobzin, 2011).]

In the shaded area on the right side of Table 1, in this parameter, arginine (R) is not an exception any more. All AAs [R, (S, N), (**K**, R), (G, D)], which are located between the two ends (**W**-**E**), have lower values within their own pairs; all but one: lysine (K) appears now to be an exception, instead of arginine (R). The rest of two AAs (at two ends) have higher values (W-E).

The same logic of "adding at the end" is valid also for the arrangement of AAs in the middle area if we understand that this area consists of two parts: more exactly two sub-parts: the sub-part up and sub-part down. In the first case we have: (**Q**–H–Y–**C**), where two central AAs have lower, whereas the two AAs at two ends have higher values in their own pairs; and, in the second case (**V**–A–T), the T and A have lower values but V, at the end, possesses a higher one.

6. Cube equation as a model for GC

The splitting of the total codon space into three parts: 0-21, 22-42 and 43-63, i.e. the triple multiplication of the number 21, inevitably gives us the idea (the hypothesis for further researches) that, perhaps, it makes sense to do the fourth multiplication, i.e. to add the amino acid space – the 20 amino acid molecules plus one stop signal (stop signal, encoded by three stop codons) to the codon space. In support of this idea, there is the fact that there is such a

justification for splitting the set of the constituents of genetic code into three parts that corresponds to general form (model) of a cubic equation ($21 \times 4 = 84$) (Table 7).

Table 7 gives the most general form of a cubic equation. If one of the three possible real solutions for x (in the case when a = b = c = 1) takes values from the series of natural numbers (x = 1, 2, 3 ...), it is immediately obvious that the fourth case (x = 4) is the model (corresponding to the structure) of the genetic code; there are 64 codons, 16 aliphatic and 4 aromatic amino acids, with a total of 84 quantities.

In addition, there are some other regularities. Thus, through validity of the similarity principle and self-similarity principle there is a chemical correspondence between four aromatic amino acids and four nucleotide bases, also of aromatic nature. Moreover, there is a balance in the number of atoms: U12+G16 = C13+A15 = 28; and in the amino acid side chains: F14+Y15 = H(14-3) + W(15+3) = 29. On the other hand, the total of 84 quantities is a "missing link", because, in previous papers, we showed that the GC is determined by a quadratic equation in the form of golden mean (Rakočević, 1998, Table 2, p. 288), as well as by two systems of linear algebraic equations with solutions x_1, y_1 and x_2, y_2 (Rakočević, 2011, Eq.4, p. 827). However, this is also the "missing link" from a different point of view. Namely, in a previous work, we showed that the GC is determined by geometric progressions with quotient 2 and 3, respectively (Rakočević, 2011, Tables A.1 & A.2, p. 839), and now it appears that the determinant of GC is still the geometric progression with the quotient of number 4. Finally, the 64 codons exist as eight 8-membered octets (the principle of self-similarity!): YUN-YCN-RUN-RCN / YAN-YGN-RAN-RGN.

Table 7. The cube equation as a model for the structure of genetic code. Second case corresponds with eight codon octets (YUN, RUN, YCN, RCN, YAN, RAN, YGN, RGN), four quartets of aliphatic AAs (G-P,V-I), (A-L, K-R), (S-T, C-M), (D-N, E-Q) and two doublets of aromatic AAs (F-Y, H-R). Fourth case corresponds with 64 codons, 16 aliphatic AAs and 4 aromatic AAs. Two last columns: the correspondence with the octal and binary numeral systems, respectively (cf. Tab. B.1).

$ax^3 + bx^2 + cx + d = 0$ If a = b = c = 1, then: x^3 + x^2 + x = d							Octal	Binary
If (x = 1, 2, 3, 4, 5, ...), then:								
1	+	1	+	1	=	03		
8	+	4	+	2	=	14	8^1	2^3
27	+	9	+	3	=	39		
64	**+**	**16**	**+**	**4**	**=**	**84**	$\mathbf{8^2}$	$\mathbf{2^6}$
125	+	25	+	5	=	155		
...								

7. Outer and inner AAs in GCT

Amino acids that have the status "added at the end" can be designated as *outer*[3] and all other as *inner* (Table 8.1, in relation to Box 3). But also without a division of codon space into three parts, it is self-evident from GCT that eight AAs (F-S-Y-C and V-A-E-G) have the status of "outer AAs". And, as we said in chapter 2, all these AAs are „extreme" from a chemical point of view: glycine (G) as the simplest possible amino acid, alanine (A) as the first possible carbon derivative, serine (S) as the first possible oxygen derivative, cysteine (C) as the first possible sulfur derivative, valine (V) as the first possible semi-cyclic molecule, phenylalanine (F) as the first possible aromatic amino acid, and tyrosine (Y) as its first possible oxygen derivative. The only exception is glutamic amino acid (E), which is not "the first", but "the last" case in the set of only two AAs, which possess carboxylic functional group in the side chain.

[3] This logic is valid only for the "ends" in the three shaded parts of Table 1, and not for T and Q which have the status "be at the end" within two sub-parts of the middle lower shaded part of Table 1.

However, only with the division of GCT into three parts, we can see that two AAs more also have the same status ("outer AAs"): triptophan (W) and proline (P); triptophan as the only amino acid with two aromatic rings and proline as the only amino acid, which possesses one non-aromatic ring (Tab. 8.1).

If we compare two sets of AAs in Table 8.1 with two sets in Box 4 and with two sets, i.e. two classes in Table 2.1, it becomes evident that both sets in Table 8.1 contain two or three amino acid doublets and six or four singlets each. Three doublets with **6** AAs and **4** singlet AAs (the third case in Box 3) if AAs are chosen from the system in Table 2.1; two doublets with **4** AAs and **6** singlet AAs (the fourth case in Box 3)[4] if AAs are chosen from the system in Box 4. In such an arrangement the balance of the number of particles is self-evident: either for the number of atoms (Table 8.1)[5] or for the number of protons and nucleons (Box 5).

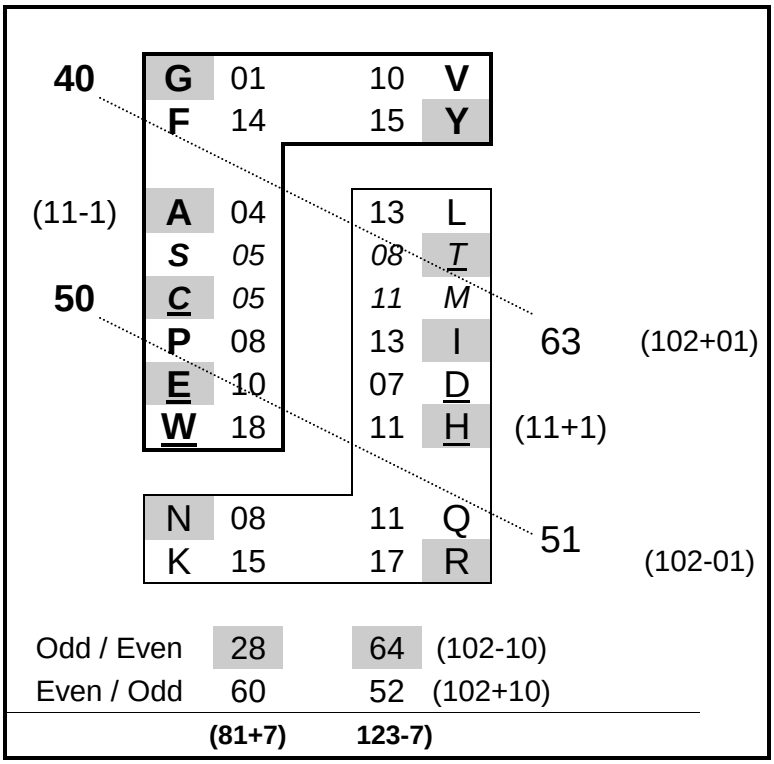


Table 8.1. The atom number balance between outer and inner AAs. Up and on the left there are outer AAs. Down and on the right there are inner AAs. The choice from the system in Table 2.1, valid for outer AAs: (G-V, F-Y, S-C) plus (A, P, E, W); and for inner AAs: (K-R, N-Q, T-M) plus (L, I, D, H). The choice from the system in Box 4, valid for outer AAs: (G-V, F-Y) plus (S, C, A, P, E, W); and for inner AAs: (K-R, N-Q) plus (T, M, L, I, D, H).

With this knowledge it makes sense to ask the following question: which amino acid pairs are present in only one of the three areas and which are present in two areas (first pair-member in one, and second pair-member in another area). With the answer to this question we have a new insight in a perfect balance of the number of atoms within the set of 16 aliphatic AAs as well as within the set of all 20 AAs (Table 8.2). Therefore, three pairs are present in only one area (D-E, P-I, K-R), and five pairs are present in two different areas (G-V, A-L, S-C, T-M, N-Q). As for two aromatic pairs (such pairs as in Table 2.1: F-Y and H-W), they are present in two different areas each. However, as we can

[4] The splitting of 10 amino acid pairs into 5 ± 1 doublets and singlets, in this way, appears to be the best solution and it is in accordance with "the symmetry in the simplest case" (Marcus, 1989).

[5] With insight into the results, shown in Tables 8.1 and 8.2, one is forced to propose a hypothesis (for further researches) that here, there really is a kind of *intelligent design*; not the original intelligent design, dealing with the question - intelligent design or evolution (Pullen, 2005), which is rightly criticized by F.S. Collins (2006). Here, there could be such an intelligent design, which we could call "Spontaneous Intelligent Design" (SPID) that is consistent with that design which was presented by F. Castro-Chavez (2010), and is also in accordance with the Darwinism. [F. Castro-Chavez (2010, p. 718): "We can conclude that the genetic code is an intelligent design that maximizes variation while minimizing harmful mutations."] Actually, it can be expected that the hypothetical SPID, contained in the results shown in Table 8.1, is in accordance with an identical (or similar?) SPID, presented in the only diagram, in Darwin's book "Origin of Species" (Darwin, 1996), as we have shown through an analysis of that diagram in one of our books (Rakočević, 1994; www.rakocevcode.rs). [In the case of the statement that *spontaneity* and *intelligent design* are mutually opposite, one must ask the question: isn't it true that human intelligence is the result of a spontaneous evolutionary process?]

see from Table 8.2, for a full balance, the amino acid pairing must be different, but also with a chemical justification: F-W and H-Y (F-Y because Y is an oxygen derivative of F; F-W because both contain the original benzene ring; H-W because both AAs are heterocyclic and H-Y because both AAs are more polar than F-W). With this change there is also a balance in the number of molecules: 5-1 molecule pairs are present in GCT within only one area (of three possible areas) (D-E, P-I, K-R, **H-Y**), and 5+1 molecule pairs are present in two different areas (G-V, A-L, S-C, T-M, N-Q, **F-W**).

G	01	10	V	G	01	10	V
A	04	13	L	A	04	13	L
S	05	05	C	S	05	05	C
T	08	11	M	T	08	11	M
N	08	11	Q	N	08	11	Q
				D	07	10	E
D	07	10	E	P	08	13	I
P	08	13	I	K	15	17	R
K	15	17	R	H	11	15	Y
				F	14	18	W
	22	51	**73**		33	69	**102**
	34	39	**73**		48	54	**102**
					81	123	**204**

Table 8.2. The distribution of amino acid pairs into three parts of GCT. On the left: first five aliphatic amino acid pairs appear to be together each, within one of the three areas in Table 1. The rest of three pairs are impaired: first member in one area and the second one in another area. On the right: the 8 aliphatic plus two aromatic pairs of AAs. The atom number balance is evident in both halves of the system.

In Table 8.1, up and left, there are the AAs with the status "added at the end" (outer) within the three parts of GCT (Table 1), while all the others are down and on the right (inner). Two columns correspond to two rows in Table 2.1. In the first column there are AAs handled by class II of aminoacyl-tRNA synthetases, all but three – C, EW – from the first class; in the second column there are those AAs which are handled by class I aminoacyl-tRNA synthetases, except three – T, DH – from class II.

The balances of the number of atoms in Table 8.1 are directly indicated. However, there is a hidden balance between the number of atoms within two columns. Originally, first column (first row – Class II – in Table 2.1) contains 81 atoms, which means 102 – 21, where 102 is the arithmetic mean, while the second column contains 123, i.e. 102 + 21 atoms. Furthermore, the difference 123 – 81 = 42, as a quantity "42", corresponds to quantity "42" which also appears to be harmonic mean of the codon space 0-63 (see Chapter 2).[6]

Remark 4. The proton number within amino acid side chains: (**G = 01, V = 25; F =49, Y = 57**), (**A = 09, S = 17, C = 25, P = 23, E = 39, W = 69**) / (L = 33, T = 25, M = 41; I = 33, D = 31, H = 43), (N = 31, Q = 39; K = 41, R = 55).

Remark 5. The nucleon number within amino acid side chains: (**G = 01, V = 43; F =91, Y = 107**), (**A = 15, S = 31, C = 47, P = 41, E = 73, W = 130**) / (L = 57, T = 45, M = 75; I = 57, D = 59, H = 81), (N = 58, Q = 72; K = 72, R = 100).

Box 3		Box 4	
Singlets	Doublets	a	b
inner + outer	inner + outer		
10 + 10 (5 + 5)	0 + 0 (5 - 5)	A – L	G – V
		S – T	P – I
8 + 8 (5 + 3)	1 + 1 (5 - 3)	C – M	
		D – E	
6 + 6 (5 + 1)	**2 + 2 (5 - 1)**	N – Q	
		K – R	
4 + 4 (5 - 1)	**3 + 3 (5 + 1)**	H – W	
		F – Y	
2 + 2 (5 - 3)	4 + 4 (5 + 3)		
0 + 0 (5 - 5)	5 + 5 (5 + 5)		

Box 3. Possible "choices" of amino acid singlets or doublets from the pairs given in Table 2.1 are presented here. In Table 1 they take outer or inner positions; that means that in some cases AAs appear separately, in other cases they appear in pairs, as shown in Table 8.1. If we compare Box 3 with Table 8.1, it is clear that the two central cases are"chosen".

Box 4. Going from the system presented either in Table 2.1 or in Box 4, to the system in Table 8.1, the amino acid pairs can be split

[6] The splitting of atom number space into three parts ($3 \times 14 = 42$) appears to be a simulation of such a splitting of codon space also into three parts ($3 \times 21 = 63$).

into AAs as singlets or doublets, in the framework of these possibilities. Column *a* as in Survey 1.1 and column *b* as in Survey 1.2, both in: Rakočević and Jokić, 1996. In column *a* there are AAs from the alanine stereochemical type, while in column *b* there are AAs from three non-alanine stereochemical types: glycine type (G), proline type (P) and valine type (V-I).

Box 5. Number of protons and nucleons within outer and inner AAs within the system in Tab. 8.1

Proton number

Number of protons in outer AAs: Singlets (18**2**) – Doublets (13**2**) = 50
Number of protons in inner AAs: Singlets (20**6**) – Doublets (16**6**) = 40
Difference of differences: 50 – 40 = 10

Positions in both columns: 36**8** – 31**8** = 50
Even + Odd = 368
Odd + Even = 318

Nucleon number (Differences: ± 20)
Outer / Inner AAs: 57**9** / 67**6**
Left / Right column: 55**9**/ 69**6**

Box 5. All data are related to the system in Table 8.1

Table. 9. The harmonic mean of harmonic mean. Harmonic mean of 63 and its half is H = 42. The half of 42 is H/2 = 21. Finely, harmonic mean of 21 and its half is h = 14 (cf. Table B.1).

n	2^n	Intervals	H/2	h/2
1	2	0 – 1		
2	4	0 – 3	1	
3	8	0 – 7		
4	16	0 – 15	5	
5	32	0 – 31		
6	**64**	**0 – 63**	**21**	**(7)**
7	128	0 – 127		
8	256	0 – 255	85	
9	512	0 – 511		
10	1024	0 – 1023	341	
11	2048	0 – 2047		
12	4096	0 – 4095	1365	(455)
...				

The changes caused by a direct EW/DH "crossing-over", and an indirect (diagonal) C/T "crossing-over", the atom number within columns becomes 102 ± 14, respectively, where 14 is the harmonic mean of the number 21 and its half (h = 14, in relation to Table 9). Thus, bearing in mind that (102 +14) - (102 -14) = 28, where 28 is the harmonic mean of 42, it is self-evident that the relationships in Table 8.1 (results 81 + 7 and 123 – 7) correspond with the 6th case in the number arrangement in Table 9. In other words, the EW/DH and C/T "crossing-overs" occur correspondingly with the harmonic mean of the harmonic mean (H = 42 and 2h = 28; "H" and "h" in relation to the number arrangement in Table 9). [Cf. shaded rows with columns in Table B.1.]

8. Three parts of GCT by the codons

In addition to previously presented large amount of regularities, which justify the codon space division into three equal parts, it makes sense to examine the validity of this division from the aspect of relations between the codons, as well. Beside other possible relations, here we present only those relations that are in connection with the nucleotide number interdependence – relations of lower and higher molecule complexity – in the three specified parts ("0" for lower and "1" for higher molecule complexity in binary records, given inTables A.1, A.2 and A.3 in Appendix A). In addition, two pyrimidines in relation to two purines are of lower complexity, as well as two nucleotides with two hydrogen bonds in relation to two nucleotides with three hydrogen bonds: (U, C → A, G), (U, A → C, G) .

The hierarchy which we have just presented follows the original concept of R. Swanson (1984), which means that the hierarchy is derived from the number of units in binary codon record, in the Gray code model, on the binary tree as well as in the GCT (Table 1).

As we can see from Tables 10 & 11 (in relation to the Appendix A), on the 6-bit binary tree, i.e. within 64 words in GCT there are exactly 6 +1 patterns that are related to the occurrence of a unit (a nucleotide of higher complexity) in binary records (column *a* in Table 10 in relation to Table 11). In the six patterns it is inevitable for such a change to occur i.e. it is inevitable that at least one nucleotide of higher complexity appears.[7] However, there is exactly one pattern (the seventh) in which there is no hierarchy change, a nucleotide of higher complexity does not appear, and that is zeroth codon UUU (000000) (cf. "0" in column "a_1" of Table 10 and "F" in column "1" of Table 11).

I		II		III	
a_1	b_1	a_2	b_2	a_3	b_3
0	**(1)**	1	**(1)**	2	**(1)**
1	**(5)**	2	**(5)**	3	**(5)**
2	**(9)**	3	**(9)**	4	**(9)**
3	**(6)**	4	**(5)**	5	**(5)**
4	**(1)**	5	**(1)**	6	**(1)**
	22		**21**		**21**

Table 10. Codon hierarchy within three parts of GCT. The significance of columns is as follows: a. Number of occurrences of nucleotides of higher complexity (purinees and/or nucleotides with three hydrogen bonds); b. Number of codons in the patterns provided in "a".

b	1	5					9									6 / 5						1
b_1	F	F	L	L	S	I	L	L	L	S	S	P	I	I	V	L	S	P	P	M	V	P
b_2	Y	T	Y	*	H	C	V	T	T	A	*	H	Q	C	*	V	T	A	A	Q	-	A
b_3	N	R	N	K	D	S	W	R	R	K	D	E	S	R	G	R	E	R	G	G	-	G

Table 11. AAs within three parts of GCT after Table 10. Explanation in the text.

9. Concluding remarks

The presented results confirm *the working hypothesis* according to which it is necessary to divide the total codon space within GCT into three parts: 0-21, 22-42 and 43-63, and also the results go in favor of our earlier hypothesis that the genetic code was already complete in prebiotic conditions (Rakočević, 2004). However, to what we have said in previous work about the prebiotic completeness of GC[8], we now add one idea more (hypothesis for further research): only such an aggregation of AAs that can generate all regularities and interrelationships that are presented here can also generate life as such. But, bearing in mind the relationships, presented in Tables B.1, C.1 and C.2, it makes sense to formulate a hypothesis about possible extraterrestrial life. By doing so, we start from the following facts: four nucleotide bases (U, C, A, G) and 18 non-sulfur amino acids are made out of the first

[7] The sequence of digits in the record excludes the dilemma when, for example "C", should be read as a pyrimidine nucleotide (0), and when, as a nucleotide of higher complexity with three hydrogen bonds (1), which was otherwise specified in the regularities presented by R. Swanson (1984).

[8] "At a later stage many nucleotide/amino-acid aggregations … had been realized … Each of those aggregations could (and must) have its own ''evolution'', but only one could have been selected — the one that gained the characteristic of self-reproduction …" (Rakočević, 2004, p. 232).

possible, i.e. of the simplest non-metals (H, C, N, O).[9] Functional groups in their simplest compounds are also found in the composition of amino acid functional group: [(H–CH_3), (H–OH), (H–NH_2), (C=O)] → [(H–C(H)–(C=O)–(OH)–(NH_2)]. On the other hand, in Tables B.1, C.1 and C.2 we see that there is a strict relation between the number of atoms in 20 amino acids (384) and in 64 codons, i.e. in 192 nucleotides, through an integer multiplication (384 x 018 = 3456 + 3456) (cf. Table C.1). Thus, in my opinion, the next hypothesis can be formulated based on all these facts: all the planets in the universe, which possess water (and other conditions needed for life) must necessarily have just the same – the simplest possible – terrestrial genetic code.[10]

Without further justification for making the hypothesis, according to which intelligent beings from outer space have been incorporated "intelligent signals" in the terrestrial genetic code (*sh*Cherbak and Makukov, 2013)[11], we believe that the idea about an artificial genetic code can be useful when it comes to stocking the digital information. In such a case, we believe that the presented regularities, contained in the standard genetic code, would benefit, particularly the building of genetic-code- algorithmic structures.

Besides everything stated above, on the basis of these results, it makes sense to give two predictions:

1. The presented relations between amino acids must be, *mutatis mutandis*, expressed in the structure of proteins, especially in terms of determination of invariant, conservative and radical amino acid positions, with distinct biological effects;

2. The presented unity of form and essence (mathematical regularities versus physicochemical properties), together with such a unity presented in our previous paper (Rakočević, 2011, Eq. 3 on p. 826 in relation to Fig. 3 on p. 828), must be reflected – through genes expression – on the relationship between genotype and phenotype in the development and evolution of organisms; everything in accordance with Futuyma's idea that the lack of knowledge of how genotypes generate phenotypes is the greatest gap in our understanding of evolution processes (Futuyma, 1979).

[9] The inclusion of two sulfur AAs occurs through the inclusion of sulfur, the oxygen's first neighbour in the sixth group; a completion of nucleotide molecule through the inclusion of phosphorus, nitrogen's first neighbour in the fifth group of the Mendeleev's Periodic System.

[10] "The simplest possible code" in relation to the following facts: glycine as the first possible amino acid (AA) without carbon in side chain; alanine as the first possible carbon AA with the open side chain; valine as the first possible half-cyclic AA; proline as the first possible cyclic AA; leucine and isoleucine as the first possible branched AAs (first possible isomers); phenylalanine as the first possible aromatic AA; all other AAs as derivatives, realized through the validity of the minimum change principle, continuality principle and self-similarity principle. For example: serine and cysteine in relation to alanine; aspartic AA in relation to serine and glutamic AA in relation to aspartic one (from both these dicarboxilic AAs follow their two amides); lysine with four carbon atoms in side chain in relation to isoleucine, also with four carbon atoms in side chain; tyrosine and tryptophan in relation to phenylalanine etc. From this logic minimally deviate methionine, arginine and histidine, but they, just as such ("clumsy") are necessary for completion the Gaussian algorithm (Rakočević, 2011, Fig. 9).

[11] "As the actual scenario for the origin of terrestrial life is far from being settled, the proposal that it might have been seeded intentionally cannot be ruled out." (*sh*Cherbak and Makukov, 2013).

Appendix A

Tables A.1, A.2 and A.3, provide a detailed disposition (arrangement) of amino acids and codons according to Tables 10 & 11. Codons are arranged in the same way (by the same ordinal numbers) as in Table 1, corresponding to the number of units in the binary record of that arrangement within quantum 1-5-9-6-1 in Table A.1 and within quantum 1-5-9-5-1 in Tables A.2 & A.3 (cf. Survey 1, below). In all three cases that and such disposition is accompanied by a strict atom number balance within the amino acid side chains. Partial balance, through quantum combination in sequence 1-5-9-5-1 is self-evident. However, if we observe the three sums within the three parts (Survey 2 in relation to Survey 3), we realize that the possible meaning of the relation is not visible at first sight. The fact that differences "14" and "62" correspond to the second perfect number (28) and to the third perfect number (496), could be, in further researches, understood only as a curiosity, or as an additional proof that perfect and friendly numbers are indeed the determinants of the genetic code (Rakočević, 2007b).

1 5 9 5 1 **4** **4** **4** **4**	Survey (1)

II III I 152 214 228 62 14	$62 \times \mathbf{8} = 496$ $14 \times \mathbf{2} = 28$	Survey (2)

4^1	$2^{\underline{2}}$		8^1	$2^{\underline{3}}$		16^1	$2^{\underline{4}}$		32^1	$2^{\underline{5}}$		64^1	$2^{\underline{6}}$	Survey (3)
4^2	$2^{\underline{4}}$		$\mathbf{8}^2$	$\mathbf{2}^{\underline{\mathbf{6}}}$		16^2	$2^{\underline{8}}$		32^2	$2^{\underline{10}}$		64^2	$2^{\underline{12}}$	

1		5						9								6						1
F		F	L	L	S	I		L	L	L	S	S	P	I	I	V		L	S	P	P	M	V	P
0		1	2	4	8	16		3	5	6	9	10	12	17	18	20		7	11	13	14	19	21	15
U		U	U	C	U	A		U	C	C	U	U	C	A	A	G		C	U	C	C	A	G	C
U		U	U	U	C	U		U	U	U	C	C	C	U	U	U		U	C	C	C	U	U	C
U		C	A	U	U	U		G	C	A	C	A	U	C	A	U		G	G	C	A	G	C	G
0		0	0	0	0	0		0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0
0		0	0	0	0	**1**		0	0	0	0	0	0	**1**	**1**	**1**		0	0	0	0	**1**	**1**	0
0		0	0	0	**1**	0		0	0	0	**1**	**1**	**1**	0	0	0		0	**1**	**1**	**1**	0	0	**1**
0		0	0	**1**	0	0		0	**1**	**1**	0	0	**1**	0	0	**1**		**1**	0	**1**	**1**	0	**1**	**1**
0		0	**1**	0	0	0		**1**	0	**1**	0	**1**	0	0	**1**	0		**1**	**1**	0	**1**	**1**	0	**1**
0		**1**	0	0	0	0		**1**	**1**	0	**1**	0	0	**1**	0	0		**1**	**1**	**1**	0	**1**	**1**	**1**
14								93 **(114+1)**																8
		58						**(114-1)**										55						

Tab. A.1. Relationships of AAs and codons within first part of GCT after Table 11. Amino acids and codons; the codons with their ordinal numbers and binary records. From left to the right there is a hierarchy of the units (number "1") within binary records of codons. Below, the number of atoms within amino acid side chains and their sums appear to be as balance solutions; within three columns there are 14 + 93 + 8 = 114+1; and within two columns there are 58 + 55 = 114 -1.

1		5						9										5						1
Y		T	Y	*	H	C		V	T	T	A	*	H	Q	C	*		V	T	A	A	Q		A
32		24	33	34	36	40		22	25	26	28	35	37	38	41	42		23	27	29	30	39		31
U		A	U	U	C	U		G	A	A	G	U	C	C	U	U		G	A	G	G	C		G
A		C	A	A	A	G		U	C	C	C	A	A	A	G	G		U	C	C	C	A		C
U		U	C	A	U	U		A	C	A	U	G	C	A	C	A		G	G	C	A	G		G
1		0	**1**	**1**	**1**	**1**		0	0	0	0	**1**	**1**	**1**	**1**	**1**		0	0	0	0	**1**		0
0		**1**	0	0	0	0		**1**	**1**	**1**	**1**	0	0	0	0	0		**1**	**1**	**1**	**1**	0		**1**
0		**1**	0	0	0	**1**		0	**1**	**1**	**1**	0	0	0	**1**	**1**		0	**1**	**1**	**1**	0		**1**
0		0	0	0	**1**	0		**1**	0	0	**1**	0	**1**	**1**	0	0		**1**	0	**1**	**1**	**1**		**1**
0		0	0	**1**	0	0		**1**	0	**1**	0	**1**	0	**1**	0	**1**		**1**	**1**	0	**1**	**1**		**1**
0		0	**1**	0	0	0		0	**1**	0	0	**1**	**1**	0	**1**	0		**1**	**1**	**1**	0	**1**		**1**
15												57 **(76)**												4
				39								**(76)**								**37**				

Tab. A.2. Relationships of AAs and codons within second part of GCT after Table 11. The number of atoms within amino acid side chains and their sums appear to be balance solutions; within three columns there are 15 + 57 + 4 = 76; and within two columns there are 39 + 37 = 76. Everything else as in Table A.1.

1		5						9										5						1
N		R	N	K	D	S		W	R	R	K	D	E	S	R	G		R	E	R	G	G		G
48		44	49	50	52	56		43	45	46	51	53	54	57	58	60		47	55	59	61	62		63
A		C	A	A	G	A		U	C	C	A	G	G	A	A	G		C	G	A	G	G		G
A		G	A	A	A	G		G	G	G	A	A	A	G	G	G		G	A	G	G	G		G
U		U	C	A	U	U		G	C	A	G	C	A	C	A	U		G	G	G	C	A		G
1		**1**	**1**	**1**	**1**	**1**		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**		**1**	**1**	**1**	**1**	**1**		**1**
1		0	**1**	**1**	**1**	**1**		0	0	0	**1**	**1**	**1**	**1**	**1**	**1**		0	**1**	**1**	**1**	**1**		**1**
0		**1**	0	0	0	**1**		**1**	**1**	**1**	0	0	0	**1**	**1**	**1**		**1**	0	**1**	**1**	**1**		**1**
0		**1**	0	0	**1**	0		0	**1**	**1**	0	**1**	**1**	0	0	**1**		**1**	**1**	0	**1**	**1**		**1**
0		0	0	**1**	0	0		**1**	0	**1**	**1**	0	**1**	0	**1**	0		**1**	**1**	**1**	0	**1**		**1**
0		0	**1**	0	0	0		**1**	**1**	0	**1**	**1**	0	**1**	0	0		**1**	**1**	**1**	**1**	0		**1**
												(107)												
8				52								(107)								46				1

Tab. A.3. Relationships of AAs and codons within third part of GCT after Table 11. The number of atoms within amino acid side chains and their sums appear to be as balance solutions; within central column there are 107 atoms as well as within other four columns. All other as in Tab. A.1 and A.2.

Appendix B

Table B.1 coresponds with the 4th case in Table 7, the 6th case in Table 9 and with the second case in Survey 3 (App. A). The multiplies of binary quantums (2 exp *n*, where *n* takes the values from series of even natural numbers) show a specific connection between numbers 6, 64 and 384, and these numbers are significant for the genetic code. [Cf. quantum 384 in Table B.1 with the Plato's harmonic number 384 (in Timaeus), as it is presented in our previous paper (Rakočević, 2011, Appendix, Table A.2) (Cf. also Table B.1 with Table 9.][12]

2^2	x	2	=	4	x	2	=	8
2^4	x	4	=	16	x	4	=	**64**
2^6	x	6	=	**64**	**x**	**6**	**=**	**384**
2^8	x	8	=	256	x	8	=	2048
2^{10}	x	10	=	1024	x	10	=	10240
2^{12}	x	12	=	4096	x	12	=	49152
…								

Table B.1. The three significant numbers for the genetic code: number 6 as the six-bit digital record of codons on the binary genetic code three (Rakočević, 1998); number 64 as the number of codons, and number 384 as the total number of bits on binary genetic code tree, and the total number of atoms within 20 amino acid molecules. (Note 1: First two columns – the realization of the principles of similarity and self-similarity.)

Note 2: (8–16 = - 8); (**64**–**64** = ±0); (384–256 = 2 x 64); (2048–1024 = 16 x 64); (10240–4096 = 96 x 64) …

Appendix C

In this appendix we add Table C.1 (in relation to Table C.2), which is in a way an extension of Shcherbak's Table of multiples of "Prime quantum 037" (which refers to the number of nucleons in the two classes of amino acids; four-codon and non-four-codon AAs) (Shcherbak, 1994, Table 1). We assume that the key features of Shcherbak's Table are the division by integer and the validity of the self-similarity principle, when it comes to the digits in number records ("037 versus 703" as a model). If we look at the first column in Shcherbak's Table (037, 370, 703) it is clear that the first two steps can be realized by all two-digit numbers while the third step (through module 9) is possible only for number 037; for example (037, 370, **703**) versus (038, 380, **722**).

As we can see, only the third step in Shcherbak's Table represents a unique opportunity, and here we show that this number (703) is at the same time the element of another arithmetic system (Table C.1, in relation to Table C.2). In this new system, in terms of division by integer, there appear only three numbers more: 105, 108 and 405; and when it comes to correspondence with the genetic code, the only significant number is 108 (see the section below the diagonal line in Table C.1).

[12] The fact that Nature "designed" everything in such a way that it equalized the number of atoms and bits (384), irresistibly leads us to Nicholas Negroponte (a Greek American architect, best known as the founder and Chairman Emeritus of Massachusetts Institute of Technology's Media Lab, and also known as the founder of the One Laptop Per Child Association, OLPC), who, 18 years ago came up with the idea of an inevitable relation between atoms and bits in the world (Negroponte, 1995). [Negroponte discusses the differences between bits and atoms. Atoms make up tangible physical objects and digital information, on the other hand, is made up of bits. He believes that all forms of information that are now made of atoms will eventually be made into bits.]

102|201	103|301	104|401	**105**|501	106|601	107|701	**108**|801	109|901
	203|302	204|402	205|502	206|602	207|702	208|802	209|902
204-024 =180 204+**180** = 384		304|403	305|503	306|603	307|**703**	308|803	309|903
			405|504	406|604	407|704	408|804	409|904
105 = 015 x 07				506|605	507|705	508|805	509|905
108 = 018 x 06	037 x 01 = 037				607|706	608|806	609|906
405 = 045 x 09	037 x 10 = 370					708|807	709|907
703 = 037 x 19 407 = 037 x 11	037 x 19 = **703**			**108** x 64 = 3456+3456			809|908
[(64 x 6 = 384) (384 x 018 = 3456+3456)][$3^3+4^3+5^3=6^3$ (Plato's law)]							

Table C.1. A possible natural numbers arrangement in relation to the "critical" Shcherbak's number 703 and its inversion 307. From only four division-integer cases, two (108 and 703) correspond to the number of atoms within constituents of the genetic code (384 in 20 amino acid molecules and 3456+3456 in 192+192 nucleotides). [Note 1: The significance of numbers 703 and 407 as it follows from Shcherbak's Table 1 in (Shcherbak, 1994), presented here as Table D.1. Note 2: Underlined patterns, first (102) in first row and second (204) in second row as half and whole number of atoms within 20 amino acid side chains, respectively.]

012|210	013|310	014|410	**015**|**510**	016|610	017|710	**018**|**810**	019|910
	023|320	024|420	025|520	026|620	027|720	028|820	029|920
		034|430	035|530	036|630	037|730	038|830	039|930
			045|**540**	046|640	047|740	048|840	049|940
510 = 015 x 34				056|650	057|750	058|850	059|950
810 = 018 x 45	**810** = 081 x 10				067|760	068|860	069|960
540 = 045 x 12 740 = 037 x 20	**108** x 64 = 3456+3456					078|870	079|970
Class II with 81 and Class I with 123 atoms (cf. Tab. 2.1)							089|980
(A half set of AAs) 12**3** / **3**456 (A half set of nucleotides)							

Table C.2. A possible natural numbers arrangement in relation to the Shcherbak's "Prime quantum 037" and its inversion 730. From only three division-integer cases, two (810 and 037) correspond to the number of atoms within constituents of the genetic code (81/123 and 3456/3456 in 20 amino acid molecules and in 192+192 nucleotides, respectively). [Note 1: The significance of number 037 and 740 as it follows from Shcherbak's Table 1 in (Shcherbak, 1994), presented here as Table D.1.]

Appendix D

In this appendix we add Table D.1 (Table 1 in: Shcherbak, 1994) and Table D.2 as the third possible pattern; all three patterns, corresponding with first three Shcherbak's numbers (first column in Table D.1: 037, 370, 703). The pattern ″703″ as in Table C.1; pattern ″037″ as in Table C.2, and pattern ″370″as in Table D.2.

1	2	3	4	5	6	7	8	9
037	074	111	148	185	222	259	296	333
10	11	12	13	14	15	16	17	18
370	407	444	481	518	555	592	629	666
19	20	21	22	23	24	25	26	27
703	740	777	814	851	888	925	962	999

Table D.1. The Shcherbak's Table of multiples of "Prime quantum 037" (Table 1 in: Shcherbak, 1994).

120|021	130|031	140|041	150|051	160|061	170|071	180|081	190|091
	230|032	240|042	250|052	260|062	270|072	280|082	290|092
		340|043	350|053	360|063	**370**|073	380|083	390|093
			450|054	460|064	470|**074**	480|084	490|094
370 = 037 x 10				560|065	570|075	580|085	590|095
074 = 037 x 02					670|076	680|086	690|096
						780|087	790|097
The numbers – multiples of 10 (and their inversions)							890|098

Table D.2. The third possible Table, corresponding with the Shcherbak's pattern ″370″.

References

Black, S.D., Mould, D.R., 1991. Development of hydrophobicity parameters to analyze proteins, which bear post- or cotranslational modifications. Anal. Biochemistry 193, 72-82.

Castro-Chavez, F., 2010. The rules of variation: Amino acid exchange according to the rotating circular genetic code. J. Theor. Biol. 264, 711-721.

Chechetkin, V.R., Lobzin, V.V., 2011. Stability of the genetic code and optimal parameters of amino acids. J. Theor. Biol. 269, 57-63.

Collins, F.S., 2006. The Language of God, Free Press, Bethesda, Maryland, USA.

Darwin, Ch., 1996. On the Origin of Species, Oxford UP.

Dragovich B, Dragovich A., 2010. p-Adic modeling of the genome and the genetic code. Comput. J. 53(4), 432-442; arXiv:0707.3043 [q-bio.OT].

Futuyma, D.J., 1979. Evolutionary Biology. 1st Ed. Sinauer Associates, Sunderland, Massachusetts.

Konopel'chenko, B.G., Rumer, Yu.B., 1975. Klassifikaciya kodonov v geneticheskom kode. Dokl. Akad. Nauk. SSSR 223, 471-474.

Kyte, J., Doollittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Marcus, S., 1989. Symmetry in the simplest case: the real line. Comp. Math. Applic. 17, 103-115.

Mišić, N. Ž., 2011. Nested numeric/geometric/arithmetic properties of shCherbak's prime quantum 037 as a base of (biological) coding/computing. Neuroquantology 9, 702-715.

Negadi, T., 2009. The genetic code degeneracy and the amino acids chemical composition are connected. Neuroquantology 7, 181-187; arXiv:0903.4131v1 [q-bio.OT].

Negroponte, N., 1995. Being Digital. Alfred A. Knopf, Inc., USA.

Popov, E.M., 1989. Strukturnaya organizaciya belkov. Nauka, Moscow (in Russian).

Pullen, S., 2005. Intelligent design or evolution? Free Press, Raleigh, N. Carolina, USA.

Rakočević, M.M., 1988. Three-dimensional model of the genetic code. Acta Biologiae et Medicine Experimentalis 13, 109-116.

Rakočević, M.M., 1997a. Two classes of the aminoacyl-tRNA synthetases in correspondence with the codon path cube. Bull. Math. Biol. 59, 645-648.

Rakočević, M. M., 1997b. Genetic code as a unique system. SKC, Niš (www.rakocevcode.rs)

Rakočević, M.M., 1998. The genetic code as a golden mean determined system. Biosystems 46, 283-291.

Rakočević, M.M., 2004. A harmonic structure of the genetic code. J. Theor. Biol. 229, 463-465.

Rakočević, M.M., 2011. Genetic code as a coherent system. Neuroquantology 9, 821-841. (www.rakocevcode.rs)

Rakočević, M.M., Jokić, A., 1996. Four stereochemical types of protein amino acids: synchronic determination with chemical characteristics, atom and nucleon number. J. Theor. Biol. 183, 345-349.

Shcherbak, V.I., 1994. Sixty-four triplets and 20 canonical amino acids of the genetic code: the arithmetical regularities. Part II. J. Theor. Biol. 166, 475-477.

*sh*Cherbak, V.I., Makukov, M A., 2013. The "Wow! signal" of the terrestrial genetic code. Icarus, 224, 228-242.

Swanson, R., 1984. A unifying concept for the amino acid code. Bull. Math. Biol. 46, 187-207.

Wetzel, R., 1995. Evolution of the aminoacyl-tRNA synthetases and the origin of the genetic code. J. Mol. Evol. 40, 545-550.

Woese, C.R. et al., 1966. On the Fundamental Nature and Evolution of the Genetic Code. Cold Spring Harb. Symp. Quant. Biol. 31, 723-736.

AUTHOR'S REFERENCES FOR THIS VOLUME

1. Rakočević, M. M. (2006) Genetic code as a harmonic system, arXiv:qbio/0610044 [q-bio.OT].

2. Rakočević, M. M. (2000) The factors of the classification of protein amino acids, *Proceedings, (Glasnik) of the Section of Natural Sciences of Montenegrin Academy of Sciences and art (CANU)*, 13, 273-294. (arXiv:q-bio/0611004 [q-bio.BM])arXiv:q-bio/0703011v2 [q-bio.OT].

3. Damjanović, M. Z., Rakočević, M. M. (2006) Genetic Code: A new understanding of codon - amino acid assignment, arXiv:qbio/0611033v1 [q-bio.OT] (an extended version of paper published in Ann. N.Y. Acad. Sci. 1048: 517-523 (2005)); published also in: Proceedings (*Glasnik*) of the Section of Natural Sciences of Montenegrin Academy of Sciences and art (CANU).

4. Rakočević, M. M (2004) Further generalization of Golden mean in relation to Euler's „divine" equation, *FME Transaction* (FME = Faculty of Mechanical Engineering, Belgrade, Serbia), **32**, 95-98. (arXiv:math/0611095 [math.GM]).

5. Rakočević, M. M. (2007) Genetic code as a harmonic system: two supplements, arXiv:q-bio/0703011v2 [q-bio.OT].

6. Rakočević, M. M. (2007a) A new genetic code table, arXiv:qbio/0703012v1 [q-bio.GN].

7. Rakočević, M. M. (2008) Genetic Code: Four-Codon and Non-Four-Codon Degeneracy, arXiv:0802.1056v1 [q-bio.BM].

8. Rakočević, M. M. (2009) Genetic Code Table: A note on the three splittings into amino acid classes, arXiv:0903.4110v1 [q-bio.BM].

9. Rakočević, M. M. (2009) Genetic Code: The unity of the stereochemical determinism and pure chance, arXiv:0904.1161v1 [q-bio.BM].

10. Rakočević, M. M. (2011) Genetic Code: Four Diversity Types of Protein Amino Acids, arXiv:1107.1998v2 [q-bio.OT].

11. Rakočević, M. M. (2013) Harmonic mean as a determinant of the genetic code, arXiv:1305.5103v4 [q-bio.OT].

AUTHOR'S REFERENCES IN OTHER SOURCES

Rakočević, M. M. (1988) Three-dimensional model of the genetic code, *Acta Biologiae et Medicinae Experimentalis* (Priština), 13, 109-116 [An excerpt in: (http://www.rakocevcode.rs).]

Rakočević, M. M. (1988) *Genes, Molecules, Language* (in Serbian with an English Language Supplement), Naučna knjiga, Belgrade. (http://www.rakocevcode.rs)

Rakočević, M. M. (1990) Information-Topological Concept of the Amino Acid Code, Compendium (*Zbornik radova, sekcija Hemja*) of the Faculty of Science – The Faculty of mathematics and natural sciences (former Faculty of philosophy), Niš, Serbia1, 3-23.

Rakočević, M. M. (1991). The Coherence of the Chemical and Genetic Code, Compendium (*Zbornik radova, sekcija Hemja*) of the Faculty of Science – The Faculty of mathematics and natural sciences (former Faculty of philosophy), Niš, Serbia, 2, 1-29.

Rakočević (Mirkov) Miloje (1994) *Logic of the Genetic Code*, Publisher: Naučna knjiga, Belgrade (Beograd). (http://www.rakocevcode.rs)

Rakočević, M.M, Jokić, A (1996) Four stereochemical types of protein amino acids: synchronic determination with chemical characteristics, atom and nucleon number, *Journal of Theoretical Biology*, 183, 345 – 349.

Rakočević, M. M. (1997) Two classes of the amino acyl-tRNA synthetases in correspondence with the Codon path cube, *Bull. Math. Biol*. 59, 645-648.

Rakočević, M.M. (1997) *The genetic code as a unique system*, SKC, Niš & Bina, Beograd. (http://www.rakocevcode.rs)

Rakočević, M.M. (1998) The genetic code as a Golden mean determined system, *Biosystems*, 46, 283-291.

Rakočević, M. M. (1998) Whole-number relations between protein amino acids and their biosynthetic precursors, *J. Theor. Biol*. 191, 463 – 465.

Rakočević, M.M. (1998) The harmony of periodic system of chemical elements, *Flogiston* (Belgrade), **7**, 169-183 (in Serbian with an extended English summary).

Rakočević, M. M. (2000) The factors of the classification of protein amino acids, *Proceedings (Glasnik) of the Section of Natural Sciences of Montenegrin Academy of Sciences and art (CANU)*, 13, 273-294. (arXiv:q-bio/0611004 [q-bio.BM])

Rakočević, M. M. (2002) Genetic Code: Alphanumerical Regularities Within a Codon Determined Amino Acid Order, *Proceedings (Glasnik) of the Section of Natural Sciences of Montenegrin Academy of Sciences and art (CANU)*, 14, 183 – 197.

Rakočević, M.M (2004) A harmonic structure of the genetic code, *J. Theor. Biol.* 229, 221 – 234.

Rakočević, M. M (2004) Further generalization of Golden mean in relation to Euler's „divine" equation, *FME Transaction* (FME = Faculty of Mechanical Engineering, Belgrade, Serbia), **32**, 95-98. (arXiv:math/0611095 [math.GM]).

Rakočević, M.M. (2005) A holistic approach in understanding of genetic code, Quantum-Informational Medicine, QIM, 2011, *Proceedings of Round Table Knowledge Federation Dialog Belgrade*, 2011: Partial Versus Holistic Oriented Approaches, 25 September 2011, Belgrade, Serbia.
Rakočević, M.M. (2011) Genetic code as a coherent system, *Neuroquantology*, 9 (4), 821-841. (www.rakocevcode.rs)

Rakočević, M.M. (2013) Golden and Harmonic Mean in the Genetic Code, *Proceedings* of the 2nd International Conference "Theoretical Approaches to BioInformation Systems" (TABIS.2013), September 17 – 22, 2013, Belgrade, Serbia.