# System-directed pairing of protein amino acids. Part II

**Miloje M. Rakočević**

University of Niš, Faculty of Sciences and Mathematics, Department of Chemistry, Višegradska 33, 18000 Niš, Serbia

**Abstract.** In this second part of the paper, the pairs of protein amino acids (AAs), canonical in the genetic code (GC) in the three system-arrangements (OS, DS and AS), presented in the first part, are compared with the Table of chemical similarity of AAs (Table 1 in Part I). The obtained results are such that it makes sense to speak about an ORiginal CHEmical STandard (Orchest), valid not only for amino acids but for life complete.

**Key words.** Natural code, genetic code, chemical code, protein amino acids, nucleotides, genetic code Table, amino acids pairing.

1. *Introduction*. In the first part of this paper, three system-arrangements of protein amino acids (AAs) are crossed: orthogonal (OS), diagonal (DS) and angular system-arrangement (AS). The results of this crossing are given in Surveys 1, 2 and 3 of the first part of the paper. The question arises as to how many of these new pairs of AAs, formed by this crossing, correspond, i.e. how much they agree with the amino acid pairs in the Table of their chemical similarity (Table 1 in the first part of the paper and Table G1 in this second part). Here we give the answer to that question.

*2. Elaboration of comparison*. In Table G1, the pairs of amino acids are marked, in the sense that the pairs are vertical neighbors. This marking is given by rounding and the color. Four pairs are marked in yellow, which are found in all three system-arrangements obtained by crossing three systems (OS, DS, AS), which means in Surveys 1, 2 and 3 of the first part of the paper. Blue (S-T) indicates the pair that is in two systems: in the system-arrangements of Survey 1 and Survey 3. Green indicates the pairs of AAs that are in only one of the three system-arrangements.

Tables G2 and G3 indicate the deviations from the similarity Table. In the question are amino acid pairs that were formed by crossing three system-arrangements (OS, DS and AS). Each pair is indicated by the same color. These are actually facts about the relationships of the three system-arrangements on the one hand; and, on the other hand these are also facts that show the degree of freedom in the pairing of protein amino acids. However, whether it is a matter of agreeing with the Table of chemical similarity, or disagreeing, nothing here is random, but everything is a case of the organizing of a possible harmonious structure (Ref. 4). We see this from Survey G1 and G2. Survey G1 refers to the regularities given in the Table of chemical similarity (Table G1) while Survey G2 refers to the exceptions given in Tables G2 and G3.

3. *Results, in relation to one standpoint and one prediction.* The results expressed through the above elaboration are such that it makes sense to speak about an ORiginal CHEmical STandard (Orchest), valid for AAs. In this sense, our *standpoint* is that in future research and practice this Orchest could be useful: in bioinformatics, as in searching of protein and nucleotide databases, as well as in the analysis of proteomes and, indirectly, the genomes (Ref. 1 and 2). This is, therefore, an affirmative attitude. However, here we also present *a prediction*, but in the form of negation. Given that there are more than 900 nonprotein but natural amino acids, the prediction is that it is considered impossible, with the involvement of all computers in the world, to create a new Orchest (to get one new such "orchestration"), from amino acids outside these 20, canonical.

It is not possible to create such a system, in the way presented here; and that means by crossing three potential system-arrangements, "taken" from the Periodic System of Numbers (PSN), and at the same time from the previously "constructed" Table of chemical similarity, of new selected AAs; and, altogether, to be in relation to the Gaussian and Dürer's numbers.

4. *Discussion* (II).[1]

Our standpoint above does not mean that anyone will ever apply it anywhere. However, in case someone wants to, I state here a few more facts for two different system-arrangements; for this one presented in Table G1 and the other one presented in Tables G2 and G3. For amino acids that are in accordance with the Table of chemical similarity (Table G1) we can say that they have the status of something that is the first; while for the amino acids found in Tables G2 and G3, which are therefore not in accordance with the Table of chemical similarity, we can say that they have the status of something that is derived.

In support of this, we look at the amino acids in Table G1. We find the first possible most adequate chemical pairings that follow the course of chemical change, from the aspect of similarity / dissimilarity.[2] After this statement, it is immediately clear that the pairings in Tables G2 and G3 are chemically less adequate.[3] It is up to the experimenters to show how much is this "less", and how it relates to that what is "the most", and is given in the Table of chemical similarity, presented in Table G1.

[1] The first part of the Discussion is given in Part I of this paper.

[2] Here again we see the reason why in understanding the genetic code must differ: what is "the code", what is "the cipher" of the code, and what is "the key" of the cipher (Ref. 12).

[3] It is important to note here that Tables G2 and G3 are two tables for technical reasons only; otherwise it is one Table in the question. Namely, Table G1 is the Table of accordance with the Table of chemical similarity (Table G1), while the remaining two (Table G2 and Table G3) represent the Table of disagreement (in the act of pairing of AAs) with the Table of chemical similarity.

But the pairing of amino acids by chemical similarity has an even deeper meaning than the above said. To make sure of that, we look at Table G1 again. We notice that only four pairs of AAs (marked with yellow rings) appear in all three system-arrangements, derived from the crossing of OS, DS and AS, respectively (Surveys 1, 2 and 3 in Part I of this paper). None of the eight amino acids contained in these four pairs participate in the pairing shown in Tables G2 and G3. The question is what their specificity is, and they are more powerful in pairing than other amino acids. We do not have a complete answer to this question, but we can point out the specifics.

So, glycine is the simplest, and thus the first possible amino acid, and alanine is its first possible derivative. On the other hand, leucine and isoleucine as a side chain have the first possible branched hydrocarbon, iso-butane.[4] On the third hand, phenylalanine is the first possible aromatic amino acid, with a side chain from the set of hydrocarbons, the arenes, and tyrosine is its hydroxide derivative. It is important to note that the first possible aromatic amino acid does not have a benzene ring in the side chain, but a toluene one.[5] Only in this way do we have a branching pattern that corresponds to the pattern found in the two amino acids – leucine and isoleucine, which possess "the structures of the methyl and iso-butyl groups", in the same way as glycine and phenylalanine (Ref. 5, p. 38).

It remains to point out the specificity of amino acids in the K-R pair. The amino acid lysine is not the first possible case analogous to serine and/or cysteine, than is even the fourth case with four instead of one methylene group. However, it is the first possible case from the aspect of similarity with leucine and isoleucine: in all three molecules there are four carbon atoms each. The same goes for arginine, a lysine pairing member.

*

These four pairs of AAs are distinct also in quantitative sense, measured by the number of atoms in the molecules. Thus, we see a difference from the original Chemical Similarity Table (Table 1 in Part I, i.e. Table G1 here). There were 10 vs 10 molecules with 102 versus 102 atoms. And here is the change for ± 2 molecules and ±10 atoms. Again we have balancing and nuancing within the genetic code in a new way (Ref. 5, 6, 7 and 8). [Cf. Table G1 with Survey G1.]

[4] Leucine and isoleucine are the only isomeric pair of molecules in the set of 20 canonical amino acids, in the terrestrial genetic code.

[5] Ref. 15: "Phenylalanine, as its name suggests, is an alanine derivative by replacing a hydrogen atom in the side chain of alanine, in the CH3 group, with a benzene phenyl group. By this act appears the situation which is also readable as so that phenylalanine is formed as a derivative of derivative: in the benzene derivative toluene … one hydrogen atom is replaced by an amino acid functional group."

These four pairs remain the same during the crossing of the systems OS, DS and AS. The amino acids in the remaining six pairs, however, also enter into new pairings as shown in Tables G2 and G3, corresponding to Survey G2. But, despite these new pairings, the *self-similarity* was maintained. The changes that occurred after the crossing (Surveys 1, 2 and 3 in Part I), the changes in relation to the Dürer's number, remained even after comparison with the Table of chemical similarity (Tabs G2 and G3 in relation to Table G1 and Survey G2). [Cf. Table C2 in Part I.]

*

Our standpoint, said above (Section 3) that the "orchestration" given in Surveys 1, 2 and 3 of Part I of this paper, as well as this "orchestration" given in the G-Atlas, applies not only to terrestrial conditions, but also to conditions throughout universe. And the same goes for our prediction. It follows from both that except 20 protein (canonical) AAs, which in terrestrial conditions are constituents of GC, no other amino acid set can build any life, anywhere in the universe. Moreover, canonical amino acids also can be included – all, except the four starting ones,[6] even then life would not be possible. Also, Life would not be possible with the addition of any other molecules. All this presents the idea on the existence of a hypothetical – the Biological Perpetuum Mobile.

**References**


1. Ashburner, Michael et al. (2000) Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium, Nature Genetics, 25(1): 25-29. [PMC free article] [PubMed] [Google Scholar]

2. Brown, Stuart M. (2003) Essential of Medical Genomics, John Wiley & Sons, New Jersey.

3. Popov, E.M. (1989) Strukturnaya organizaciya belkov, Nauka, Moskva. [Евгений Митрофанович Попов: "Структурная организациия белков", Москва: Наука, 1989.]

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

5. Rakočević, M.M. (2018a) The Cipher of the Genetic Code, BioSystems 171 (2018) 31–47.

6. Rakočević, M. M. (2018b) Analogies of Genetic and Chemical Code, Polyhedron, 153, 292–298.

7. Rakočević, M. M. (2019a) Protein amino acids as a complete (periodic) system, Chemia Naissensis, Vol 2, Issue 1, pp. 1-43. https://www.pmf.ni.ac.rs/chemianaissensis/

8. Rakočević, M. M., (2019b) Genetic code as an image of the mirror image. Part I, OSF Preprint DOI 10.31219/osf.io/xht4c

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


[6] The first four amino acids found in the Chemical Similarity Table: glycine, alanine, valine and proline (G, A, V, P in Table G1), are not only starting here but also starting, each in its own stereochemical type (Ref. 3 and 9).

# G – Atlas

**Table G1.** Full accordance of the pairing of AAs in the Chemical Similarity Table with the pairing at the crossing of their three system-arrangements

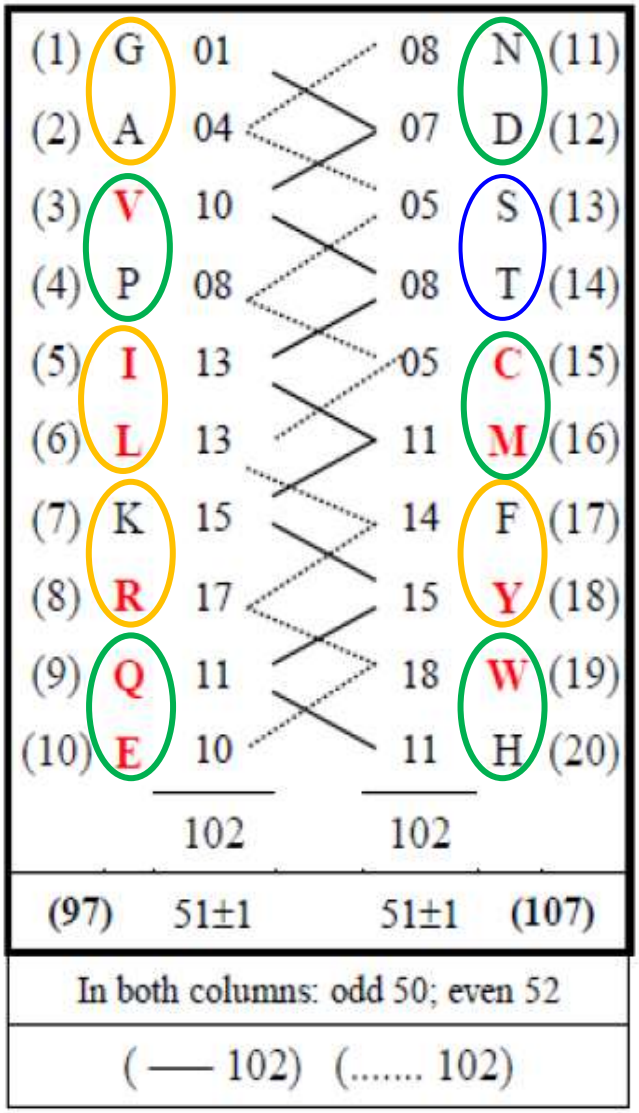

**Table G2.** Exceptions from pairing in Chemical similarity Table (I)

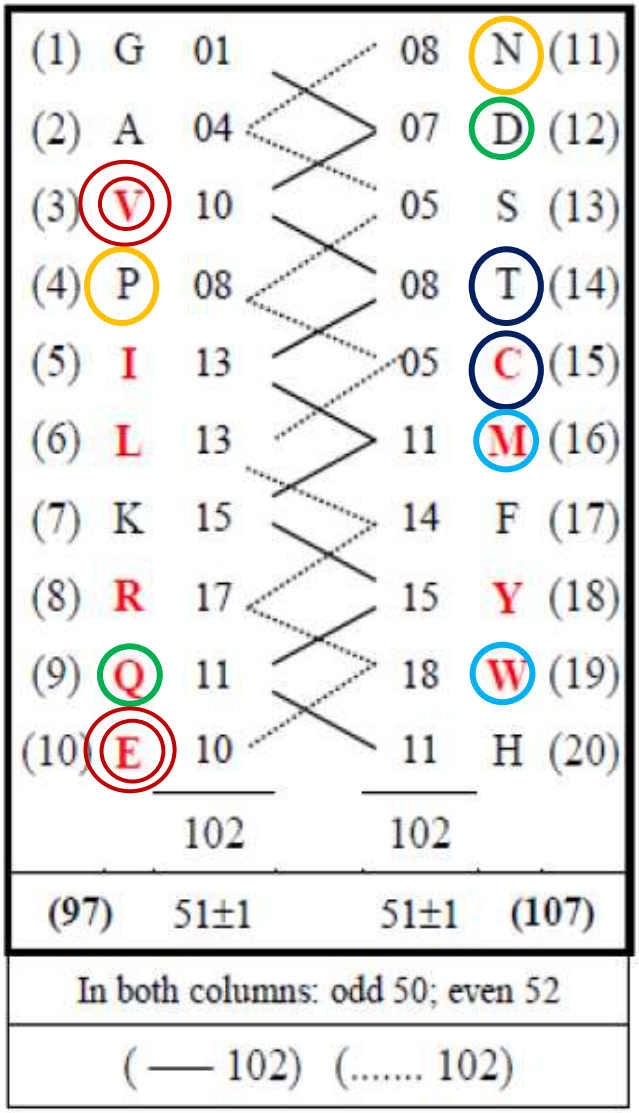

**Table G3.** Exceptions from pairing in Chemical similarity Table (II)

| | | | | | |
|---|---|---|---|---|---|
| (1) | G | 01 | 08 | N | (11) |
| (2) | A | 04 | 07 | D | (12) |
| (3) | V | 10 | 05 | S | (13) |
| (4) | P | 08 | 08 | T | (14) |
| (5) | I | 13 | 05 | C | (15) |
| (6) | L | 13 | 11 | M | (16) |
| (7) | K | 15 | 14 | F | (17) |
| (8) | R | 17 | 15 | Y | (18) |
| (9) | Q | 11 | 18 | W | (19) |
| (10) | E | 10 | 11 | H | (20) |
| | | 102 | 102 | | |
| **(97)** | | 51±1 | 51±1 | | **(107)** |

In both columns: odd 50; even 52

( —— 102) (..... 102)

**Survey G1.** Relationships in Table G1

| 3 times | | Less than 3 times | |
|---|---|---|---|
| G 01 | | | N 08 |
| A 04 | | | D 07 |
| I 13 | | V 10 | S 05 |
| L 13 | | P 08 | T 08 |
| K 15 | | | C 05 |
| R 17 | | | M 11 |
| F 14 | | Q 11 | W 18 |
| Y 15 | | E 10 | H 11 |
| 91 + 01 | | 101 + 11 | |
| 102 - 10 | | 102 + 10 | |

**Survey G2.** Relationships in Tables G2 and G3

| Left decade | | Right decade | |
|---|---|---|---|
| T G2 | T G3 | T G2 | T G3 |
| V 10 | | N 08 | N 08 |
| V 10 | | D 07 | D 07 |
| P 08 | P 08 | T 08 | |
| Q 11 | Q 11 | C 05 | S 05 |
| E 10 | | M 11 | |
| E 10 | | W 18 | H 11 |
| 34 + 05 | | 68 + 05 | |
| 34 + 05 | | 15 | |
| 68 + 10 | | 68 + 20 | |