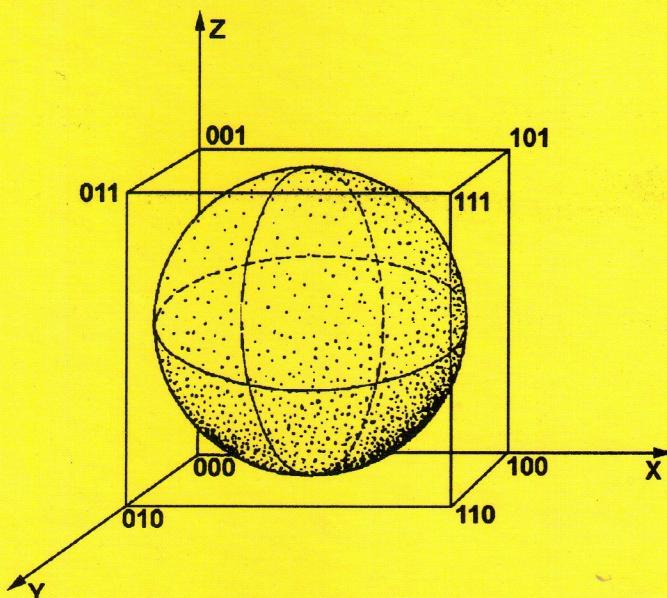


**Miloje M. Rakočević**



**GENETIC CODE  
AS A  
UNIQUE SYSTEM**

**SKC NIŠ, 1997.**

Prof. dr Miloje M. Rakočević

# GENETIC CODE AS A UNIQUE SYSTEM

(The Ray of the Microcosm, II 48-50)

Genetic Code as a Unique System  
Miloje M. Rakočević

Design & Layout  
Miloje M. Rakočević

SKC Niš, 1997.

ISBN 86-7227-082-7

Prof. dr Miloje M. Rakočević  
GENETIC CODE  
AS A UNIQUE SYSTEM

Reviewers

Prof. dr Djuro Koruga  
Prof. dr Radivoje Papović

Publisher

Studentski kulturni centar Niš,  
St. Šumatovačka bb, 18000 Niš

On behalf of the publisher  
Nebojša Randjelović, director

Editor in chief  
Aleksandar Blagojević

Executive publisher  
Goran Ilić, for BINA Beograd

Design & Layout  
Mateja Opačić

Printed in Serbia by BINA-PRESS (Niš-Belgrade)  
500 copies

ISBN 86-7757-037-3

## ACKNOWLEDGEMENTS

I would like to thank to Prof. Dr. Bojana Grupčić Ilijac, Academician Prof. Dr. Zvonimir Damjanović, Prof. Dr. Radivoje Papović and Prof. Dr. Djuro Koruga for their stimulating discussions, strong and strict criticism, merciful at the same time, during many years, from the very beginnings of my researches related to this study.

Many colleagues and friends have helped in the preparation of this study and I express my thanks to them. I have contributed in any way, but especially, to the preparation of this study Dr. Milica Jaređić, Mrs. Dr. Ratko Tadić, Mrs. Dr. Bojana Grupčić Ilijac and Mirela Oparić student Ana Joković, Dr. Bojana Djordjević and Mirela Oparić  
**Вељи јеси, Творче и Господи,  
И чудна су творења Твоја,  
Величеству Твоме краја нејма!**  
(Луч макрокозма, II 48-50)

This book represents a continuation of my book "Genes, Molecules, Languages", in Serbian (Belgrade: Naučna knjiga, 3rd ed. 1997) and "Logic of the Genetic Code", in English (Belgrade: Naučna knjiga, 1997). Benevolent comments of Prof. Dr. Russell P. Doolittle (University of California, San Diego, La Jolla, USA), Prof. Dr. Wilfried W. De Jong (University of Nijmegen, The Netherlands), Prof. Dr. Vladimir I. Shevelev (Russia), Prof. Dr. Mihajlo Raspotnicić (Serbia), Prof. Dr. Jelena Ursić Janović, Prof. Dr. Bojana Grupčić Ilijac (Serbia), Prof. Dr. Radivoje Papović and Prof. Dr. Djuro Koruga (Croatia) were of great importance in the preparation of this book.

**You are great, my Creator and Lord,  
And wondrous are all your works,  
There is no end to your Majesty!**  
(*The Ray of the Microcosm*, II 48-50)

P.P. Njegoš

I would also like to thank to the publishers of this book, to NKG (Studentiški kulturni centar) and BINA, especially to the gentlemen Nebojša Radošević, Aleksandar Blagojević, Predrag Štefanović, Goran Ilić and Tomislav Kostović.

For their technical assistance in preparing this book I am thankful to Mirela Oparić and Tatjana Kovo.

## ACKNOWLEDGEMENTS

I would like to thank to Prof. Dr. Bojana Grujić Injac, Academician Prof. Dr. Zvonimir Damjanović, Prof. Dr. Radivoje Papović and Prof. Dr. Djuro Koruga for their stimulating discussions, strong and strict criticism, merciful at the same time, during many years, from the very beginnings of my researches related to this study.

Many colleagues and friends have helped in the preparation of this study and I express deepest appreciation to all who have contributed in any way, but especially to: Prof. Dr. Rangel Igov, Prof. Dr. Mileta Jaredić, Prof. Dr. Ratko Tošić, Prof. Dr. Aleksandar Tančić and to doctoral studies students: Anja Jokić, Lidija Matija, Nataša mišić and Mateja Opačić.

This book represents a continuation of my previous two books Genes, Molecules, Language, in Serbian (Belgrade: Naučna knjiga, 3rd ed. 1990) and Logic of the Genetic Code, in English (Belgrade: Naučna knjiga, 1994). Benevolent comments of Prof. Dr. Russel F. Doolittle (University of California, San Diego, La Jolla, USA), Prof. Dr. Wilfried W. De Jong (University of Nijmegen, The Netherlands), Prof. Dr. Vladimir I. Shcherbak (Kazakh State University, Alma-Ata), Prof. Dr. Bogdan Šešić, Prof. Dr. Milan Raspopović, Prof. Dr. Randjel Mihajlović, Prof. Dr. Jasna Ursić Janković, Prof. Dr. Radosav Palić, Prof. Dr. Milovan Purenović, Prof. Dr. Blaga Radovanović and Prof. Dr. Velimir Abramović, where an important stimulus and encouragement for what I feel obliged.

I would also like to thank to the publishers of this book, to SKC (Studentski kulturni centar) and BINA, especially to the gentlemen: Nebojša Randjelović, Aleksandar Blagojević, Predrag Jelenković, Goran Ilić and Tomislav Kresović.

For their technical assistance in preparing this book I am thankful to Mateja Opačić and Tatjana Rovo.

Last, but not the least, I would like to thank to my friends without whose helps this book also would not have been written: Ognjen Lakićević, Jovo Kunac, Svetlana Mitrović, Zorica and Pera Rakin, Blažo Perović, Dragiša Janković and Djuro Koruga.

At the end, what means at the beginning, I would like to thank to my Žizela for her patience and understanding, without whose help my whole work would not be possible.

## AUTHOR

## PREFACE

This book is an attempt to focus attention on new understanding of the genetic code. The central theme is the Boolean spaces and genetic coding based on them. The mysteries of the genetic code are ultimately explained in terms of LIGHT (Logical, Information, Geometric, Homeomorphic, Topological) model and System.

After reading this book we can see that, there have been several reasons for its writing. First, until this book there is not good enough mathematical model to make link with physical reality of genetic code. Second, we can see that there is strong determination between atom and nucleon number of amino acids and their physical and chemical parameters. A third motive for this book is to make available a unified resource for teaching Ph.D. students. The book can also serve both researchers and students in the field of biochemistry, molecular biology and interdisciplinary studies.

Dr. Rakočević does a particularly excellent job of creating a working model of nucleon number (mass) and shell properties (electrons). The book thoroughly, clearly, and gently opens the reader's mind to the conclusion that we, as biological begins, are more than classical chemistry and physics entities.

In the past seven years, I have had the pleasure to be associated closely with Dr. Rakočević. I liked the book, have enjoyed reading it and think that is a timely contribution. As usually in science, we may do not agree with all of it but, in the main, Dr. Rakočević model of genetic code is more consistent with physical reality than any other model which I know.

There is no question that this book will have great influence on our thought about genetic code.

Prof. Dr. Djuro Koruga

## 1. INTRODUCTION

In spite of the fact that the genetic code had been practically decoded as early as 1966 (Crick, 1966a), there are still many unanswered questions and controversies even today in relation to it. Among the unanswered questions, that is, the unsolved problems we list the following:

1. Is the genetic code really universal or not (Crick, 1968; Porschke, 1985; Alvager et al., 1989)?
2. If it makes sense, instead of a universal, to talk about the standard code, does an established connection, one defined by law or principle, with nonstandard codes (cf. Attardi, 1985; Alvager et al., 1989) then exist?
3. Being redundant, did the genetic code (in the process of evolution) really degenerate, or was it generated as such? (Caspari, 1968, p 327: "This code was completely degenerate and... each code letter was used in more than one codon.")?
4. As to the interpretations of the origin of the genetic code, to which of the two theories should preference be given: "The Stereochemical Theory" or "The Frozen Accident Theory" (Crick, 1968; Porschke, 1985); moreover, to what degree is the first theory supported by the Watson-Crick rules of base pairing (Watson & Crick, 1953a, 1953b), and to what degree is it thwarted by mispairing in the process of complementary base pairing as a condition for the origin of substitution mutations (Topal & Fresco, 1976a), and in the codon-anticodon interaction (Topal & Fresco, 1976b); and, how much do the facts concerning pairing favor (if at all) the second theory?
5. Does the fact that only the L-amino acids participate in the genetic code favor "The Stereochemical Theory" or not, especially if it can be shown that every codon fulfills the stereochemical conditions for coding the appropriate *L,S-dimer* (Grafstein, 1983, p 157: "An intricately coupled stereochemistry is formulated which displays a binary logic for amino acid-codon recognition")?
6. With the formulation of "the general base-pairing hypothesis" (Topal & Fresco, 1976a) and "the two out of three" hypothesis (Lagerkvist, 1978 and Lagerkvist et al., 1981) has the "Wobble Hypothesis" (Crick, 1966b) been refuted?
7. Has the genetic code been "from the beginning" as it is today: a four-letter alphabet (four amino-imino bases: two purines and two pyrimidines, with at least one base- uracil- which is only an imino base)

from which three-letter words are generated; plus the twenty-letter alphabet (20 amino-imino acids, 19 amino acids and 1 imino acid) from which one-letter words are generated. Or, was the genetic code originally (in the beginning) a four-letter alphabet with two-letter words; or a two-letter alphabet (2 amino-imino bases, 1 pyrimidine and 1 purine) with doublets or with triplets; of course, in all cases, with the suitable (which?) number of amino acids (cf. Eck, 1963; Jukes, 1963, 1966, 1973, 1983; Yockey, 1977; Eigen & Schuster, 1979; Rowe & Trainor, 1983b)?  
(Note: By "the beginning" of the genetic code we mean the origin of life anywhere in the universe; in the sense that, if there's a genetic code, there is life, and if no code exists, then no life exists.)

8. Is the fact that the present day code completely represents the realization of a mathematical model - the third class variations with the repetition from the set of four elements (Gamow, 1954) of great importance (Konopka & Brendel, 1983, p 472: "The theoretical possibility... is a result of the mathematical structure of the genetic code"), or it is not of great importance (Osawa et al., 1992, p 230: "The general pattern of the genetic code results from biochemical properties of nucleotides rather than from any mathematical formula")?
9. Which factors have determined the replacements of amino acids in proteins during the evolutionary process (Dayhoff, 1969, 1972-1978; King & Jukes, 1969; Doolittle, 1981, 1985; Doolittle & Kyte, 1982; Swanson, 1984; Frömmel and Holzhütter, 1985; Taylor, 1986; Prat et al., 1986)?
10. Does Darwin's theory of selection, as a nonrandom process, still hold for the macromolecular level, and for the level of genomes, or are we talking about a "non-Darwinian Evolution" (King & Jukes, 1969) as a random and drift process, and as an indirect result of the existence of neutral mutations (Kimura, 1968)?

With argumentation for one general and several separate hypotheses (bearing the status of working hypotheses), we will show in this study that the answers to the previously asked questions have to be affirmative (the declaration of the position taken in the first part of the question section), except for the third and sixth question; the ninth question will be discussed separately.

## 2. THE HYPOTHETICAL FRAMEWORKS

2.1. *The general hypothesis*: Boolean (logical) spaces are the main determinants and the invariants of the genetic code.

2.2. *The separate hypotheses*:

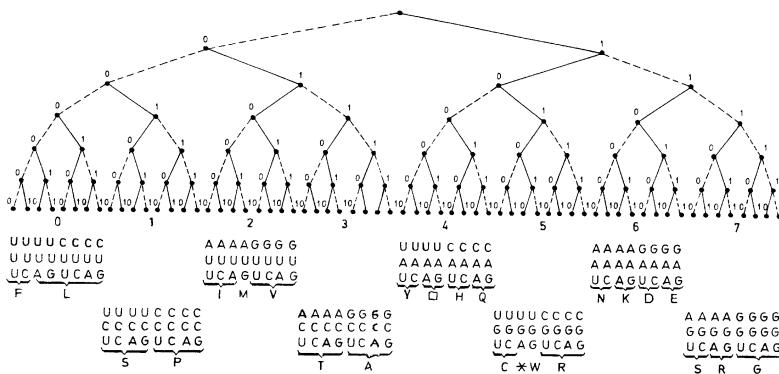
- 1) In answer to the question of whether there is any sense in talking about the evolution of the genetic code, a reliable answer can be found on condition that the following three *input-output* relations are correctly analyzed: I. *Input*: Codon-Anticodon interaction - *Output*: Codon - Amino acid relation; II. *Input*: Codon ring - *Output*: Mutation ring (both rings as in Swanson, 1984, p 188 and p 191; cf Appendix 1) and III. *Input*: Essential amino acids - *Output*: essential, semi-essential and non-essential amino acids;
- 2) *The wobble principle* is a universally - held principle for the genetic code and does not amount to only codon-anticodon interaction (Crick, 1966b);
- 3) The relation of *Strong - Middle (mixed) - Weak (SMW)* (Lagerkvist et al., 1981), that is, *Full-Semi - Empty (FSE)* (Rakoëeviæ, 1994) is a universally - held relation for the genetic code;
- 4) “*The Crossing - over*” principle is a universally - held principle for the genetic code not only for its physical but also for its logical systems (structures).

The presented argumentation for the stated hypotheses proves that the genetic code represents a whole, unique, and unified system with strict relations of binary symmetricality, proportionality and harmoniousness of all its parts (constituents) within the whole; and that not only from the formal aspect (the number of molecules, atoms and nucleons) but also from the essential aspect (the structure and the physical and chemical properties of the constituents). The genetic code must have been in “the beginning” in the same state as it is at present (today) because the generation of such a genetic code is a prerequisite for the origin of life anywhere in the universe. (The chemical evolution of macromolecules, which occurred prior to the genesis of the genetic code, will be considered as being prebiotic in this study; cf Dickerson, 1978, pp 70-86: “One of the fascinating side issues of origin-of-life biochemistry is why the present set of 20 amino acids was chosen”; cf also Pflug, 1984, p 67: “A prebiotic evolution took place on the early earth. The origin of life is open to alternative explanations, including extraterrestrial phenomena”).

**Remark 2.1.** All the three elementary types of symmetries (1. in relation to the point, i.e. center, 2. in relation to the line or axis and 3. in relation to the plane) can be represented by the binary symmetry of the segment line, i.e. by the symmetry of the entities  $A \leftrightarrow A'$  (or  $A \leftrightarrow B$ ) in relation to the central entity  $C$  - the center of symmetry (whether or not it exists, represented only by a point), where  $AC = CA' = n$  and  $AA' = 2n$ . If so, then:

The entity  $C$  contains (in itself) the quantity  $c$  ( $c = n$ ), which represents the arithmetical mean for  $AC$  and  $CA'$ . This is the essence of binary symmetry as the "symmetry in the simplest case" (Marcus, 1989, p 103: "In its simplest form, in a one-dimensional Euclidean space, symmetry is defined with respect to one point. Given two points  $A$  and  $B$  in this space, the symmetric point  $C$  of  $A$  with respect to  $B$  is one such that the distance from  $C$  to  $B$  is equal to the distance from  $A$  to  $B$ , whereas the distance from  $C$  to  $A$  is the sum of these two distances"; Notice that our  $B$  or  $A'$  is Marcus'  $C$  and vice versa). Its simplicity is the reason why we can consider binary symmetry as the best possible symmetry.

1. There is also the entity  $M$  which contains (in itself) the quantity  $m$ , where  $m = 4/3 n$ , which represents the harmonic mean for  $AC$  and  $AA'$  (the essence of binary harmony!);  
 $AA'$  can be, in infinitely numerous ways, divided into two unequal parts;  
 $AA'$  can, in a finite number of ways, be divided into two parts which are proportional to each other (the essence of binary proportionality!);
5. There is exactly one pair of points  $G, G'$  which represent the *golden section* along the segment line  $AA'$  (the essence of infinite division and the best possible proportion!).



**Figure 1** The binary tree of the genetic code. It generates from the Gray code model of the genetic code (Swanson, 1984) by codons arranging according to the natural numbers series “for the numbers 0-63”. The 8 rosettes on the binary tree correspond to 8 codon classes. The broken line is the primary (source) line and the full line is the secondary one; pyrimidine type molecule is the parent molecule while the purine type is its derivative. However, when the purine type is to be selected, it is the primary line for the purine type base but only in one step. The Arabic numerals, as the vector numbers designate the rosettes (classes) of codons that correspond to the vertices of the unit Boolean 3-cube.

**Remark 2.2.** The binary symmetry  $A \leftrightarrow A'$  (or  $A \leftrightarrow B$ ) in this paper is still understood, except by S. Marcus, in the sense used by P. Hilton and J. Pederson (1989, pp 73-74) so that the given entity from the pair  $A, A'$  (or  $A, B$ ) "may admit several different combinatorial structures and each structure will be regarded as combinatorially distinct. You should thus be warned that [entities] which we regard as *the same* (i.e. combinatorially equivalent), you may hitherto have regarded as *different* (see Fig.1(a)); and [entities] which we regard as different (i.e. combinatorially distinct), you may hitherto have regarded as the same (see Fig.1(b))." For a better understanding of this we present a possible correspondence between their Fig.1 and our Fig.1. The entity on the left side of their Fig.1(a) - "the cube" - corresponds to the pyrimidine entity (in 32 combinations, or to be more correct, variations) on the left branch of the binary tree in our Fig.1; their right side entity ("the prism") corresponds to our purine entity on the right side (also in 32 variations). The entity on the left side of any one of the 4+4 "rossete" (or classes) codons on the binary tree in our Fig.1 corresponds to their left entity in their Figure 1(b); a simpler cube model corresponds to our simpler pyrimidine (U) and/or purine (A) entity; their right entity - a more complex cube model - corresponds to our more complex pyrimidine (C) and/or purine (G) entity on the right side (cf analogous "cubes" and "prisms" in Fig. 3.5 in Dubinin, 1985, p 81).

**Remark 2.3.** Since the genetic code can be reduced to the Gray code model (Fig.1 in Swanson, 1984, p 188) and to a binary tree (Fig.1 in this paper), with a starting codon UUU 000000 and a final codon GGG 111111, it follows that as to questions of symmetry in relation to the genetic code *the mathematical group theory* holds only partially (our hypothesis and a prediction of this - *Prediction 1* - remains for further, that is, future research). This results from the fact that set Q of rational numbers, including zero, does not form *a group* with respect to a multiplication operation. The above mentioned is the reason why we won't use the mathematical group theory to research the symmetries of the genetic code in this paper.

### 3. GENETIC CODE AS A BOOLEAN SPACE

A more detailed analysis of (experimental) facts shows that the nature of the genetic code is such that the two contradictory views stated in the 8-th question in the *Introduction* hold true simultaneously: for the characteristics of the genetic code, which, being as they are, are "the result of the mathematical structure", in other words, they are not that, they didn't originate "from any mathematical formula" but are the result of the "biochemical properties of nucleotides". The genetic code, in fact, represents a unity of both one and the other: the relations of the characteristics of the genetic code are such that they correspond to an ideal (one or more) mathematical model; "correspond" in the sense that they are correspondent of and in accordance with the model. (With this a hypothetical answer has been given to question 8 in the *Introduction*, with proof being that of the arguments which, given through all the seven

chapters of this paper, show that Boolean spaces are the main determinants and invariants of the genetic code).

### 3.1. Mathematical and Physical Models

What is an *ideal* mathematical model? To answer this question, we will first show that it makes sense to talk about a mathematical *reality-model*, on the one hand (Douglas et al., 1994, p 3: "Studying a valid model enables us to understand and rationalize the behavior of physical 'reality' and, ideally, to predict its behavior"), and a mathematical *non-reality* model, that is, a mathematical convention model, on the other hand. So, the generation of a binary pair (the Boolean segment line) in a computer will be: (0) no electricity - (1) electricity, or in the Morse code: (0) dot - (1) dash, are examples of conventions. However, the generation of a binary pair (Boolean segment line) in the genetic code: (0) a pyrimidine base type - (1) a purine base type, or: (0) a system with two hydrogen bonds, (1) a system with three hydrogen bonds, represents a reality well. Therefore, it is understood that the status of an "*ideal*" mathematical model can have only a "reality-model". Examples of a binary arranged "quartet" could be the following: From the Morse binary pair, observing as many binary relations among the constituents as possible and the principles of binary symmetry, proportions and harmonies, we can generate a "quartet" (the Boolean square): (00) small dot - (01) large dot - (10) short dash - (11) long dash. Despite the observation of the aforesaid mathematical principles this model remains only a non-reality-model, that is, a model-convention. Within the genetic code we can generate the "quartet" (the Boolean square as a reality-model) by representing the pyrimidine base type with 2 bases, and the purine base type with 2 bases as well: (00) U - (01) C - (10) A - (11) G. At this stage a very important question arises: are we talking about a non-reality-model, that is, a convention, or about a reality-model, in the sense that the relation between important physical and chemical properties of these 4 bases are such that they correspond to the relations between the 4 vertices of the Boolean (logical) square? The second question is: has, in the process of generation, care been given to the strict observation of binary relations between the bases, or in other words, to the observations of the principle of binary symmetry, proportion and harmony?

The answers to both questions are affirmative. If by "physical and chemical properties" we mean the base type and the number of hydrogen bonds (they can, and must!), then an argument is in itself a fact which shows

that the genetic code can be reduced to the Gray code, but only and exclusively by way of the mentioned "quartet" (Swanson, 1984; notice that she does not use the term "Boolean", nevertheless she represents the zeros by dots and the ones by dashes). However, the Boolean logical square of the genetic code follows from other physical and chemical properties, like energy for example (Orgel, 1986, p32:"G.C base pair is held together by three hydrogen bonds and provides a greater stabilizing influence than an A.U pair which has only two. The stacking interaction between contiguous bases is also important; pairs of purines provide the greatest stacking energy and pair of pyrimidine the least"), or like the pKa values: they are almost the same for the pair in the end positions of the Boolean square, for U(00) it is 9.5 and for G(11) it is 9.2; and the values are similar for the pair in the middle-positions, for C(01) it is 4.45 and for A(10) it is 4.15 (cf Fig. 1.1 in Rakoèeviæ, 1994, and Table 3.1 in Dugas & Penney, 1981).

The Boolean logical square, with its binary values - the Boolean vector numbers - 0, 1, 2, 3 for U(T), C, A, G, respectively, follows, in fact, from the positions of these bases in the Watson-Crick Table, as shown in Table 3 in Shcherbak, 1994, p 476 ("The four numbers 0, 1, 2, 3 are used instead of the base symbols T, C, A, G"). Keeping in mind the facts we have given here for energy and the pKa values, and keeping in mind the positions of the four bases in the binary tree (Figure 1), or in the Gray Code model (which have been determined by base type, *pyr* or *pu*, and by the number of hydrogen bonds), it follows that in spite of the fact that "there may be 24 notation versions depending on the assigning of the symbols 0, 1, 2, 3 to the base" (Shcherbak, 1994, p 476), only one permutation, or assigning, makes physical sense and that is: 0 U (T), 1C, 2A, and 3G.

Once the binary values for the four bases are determined, it is then possible to determine even the binary values for the codons, and finally for the amino acids, as we have done in our previous works (Rakoèeviæ, 1990, p 11; 1994, p 72). With the calculated binary values for the amino acids we then find that the codon assignement with amino acids is such that the separation of the codons in the codon "space" strictly corresponds to the separation of the amino acids in the amino acid "space", and do so analogous to the key quantum model, (the magnetic quantum number): -3, -2, -1, 0, +1, +2, +3 (cf. Table 4.1 in Rakoèeviæ, 1994, p 56 and see this model in chapter 6.2 in this study).

However, a more important result which follows from the discovery of the binary values for the amino acids is the fact that amino acids in that way "divide" their space so that they exist in the relations of the members of the

Fibonacci series, from sequency 1 - 21, just in the way that has been determined for the deterministic chaos (see Rakoëeviæ, 1990, pp 11 & 19; 1994, pp 64 & 182).

### 3.2. Boolean Square Relations

Because the set of all Boolean vectors of length  $n$  represents a unit  $n$ -dimensional cube ( $B^n$ ), it follows that within the Boolean square ( $B^2$ ) we have a set of vectors: (00), (01), (10), (11) with their weights: 0, 1, 1, 2 and the vector numbers 0, 1, 2, 3 respectively (about the Boolean vectors see Gavrilov & Sapozhenko, 1989, p10). It is clearly obvious that there are, from the aspect of weight, 3 entities: the entity of being "heavy" (*strong* or *full*) 11, "light" (*weak* or *empty*) 00, being both "light and heavy" (*middle* or *mixed* or *semi*) 01 or 10; from the aspect of the number of the vectors, there are 4 entities: the zeroth, the first, second and third, that is, the 1st, 2nd, 3nd, 4th entity.

Everything that has been discussed so far is a mathematical formalism, but at the same time it also portrays basic relations in the Boolean (logical) square. However, the main question here is whether these relations can be found in the genetic code. The existence of a strong correspondence among the four entities (the four vertices) of the logical square with the 4 amino-imino bases - two purines and two pyrimidines - and that by way of the reality-model has already been shown. However, in order for the SMW-FSE (*Strong - Middle - Weak or Full - Semi - Empty*) relations existing within the Boolean square to be noticed and understood as also being the reality-model of the genetic code, it is necessary to first understand that the Watson-Crick Table of the genetic code represents also the Boolean logical square. Therefore, we will quote the findings of competent researches: Lagerkvist et. al., 1981, pp 2640-2641: "This... reading must be a function of the strength of the interaction between the anticodon and the first two codon nucleotides. It is furthermore assumed that a G.C interaction with three hydrogen bonds is stronger than an A.U interaction with only two. Consequently, the probability of *two out of three* readings would be greatest for codons making only G.C interactions in these positions ('strong codons'), intermediate for codons which make one A.U and one G.C interaction ('mixed' codons), and minimal for codons making only A.U interactions ('weak' codons)"; and further on: "In the left half of the *codon square* all mixed codons appear in families" (*italics*: M.R.)

From the aspect of *strong-mixed-weak* relations, the Watson-Crick Table is characterized by a exact binary symmetry (Fig.4 in Lagerkvist et.al., 1981, p2641), as well as by a exact proportionality. The ratio of codon "quartets" which form families and those which do not is 3:1 in the left half of the codon squares, in the first and third quadrant; in the right half it is the opposite - 1:3.

In order for the Watson-Crick Table to really be a "square" which would correspond (as a reality-model) to the Boolean logical square, it is necessary to restructure it in such a way that instead of a system of  $4 \times 16$  codons it would have a system of  $8 \times 8$  codons (cf Helene, 1987, p 140). In such a table the codons UUU and UUC would be written side by side, instead of one on top of the other; below this pair would be the UUA-UUG pair, and that is how a logical (Boolean) square would be realized by way of a third base. Following this principle the codons would be arranged in both directions - horizontally and vertically. The four quadrants obtained in this manner by way of the Watson-Crick Table correspond to the four vertices of the Boolean logical square, in the following way: I (UUU-CCG) 00, II (UAU-CGG) 01, III (AUU-GCG) 10, IV (AAU-GGG) 11; where the correspondence (as the reality-model) is strongly determined by the *strong-mixed-weak* relations.

The process of the mutual replacement of amino acids in proteins in the process of evolution is also determined by the *strong-weak-middle (mixed)* relation but in such a way that this relation determines even the structure of every tRNA on one hand, and every (any one) protein molecule, on the other. In the molecules of tRNA there are invariant positions (*strong*) semiinvariant (*middle*) and variable positions (*weak*). [Voet & Voet, 1990, p 903: "There are 13 invariant positions (always have the same base) and 8 semiinvariant positions (only a purine or only a pyrimidine) that occur mostly in the loop regions... .The site of greatest variability among the known tRNAs occurs in the so-called variable arm. It has from 3 to 21 nucleotides and may have a stem consisting of up to 7 bp. The D loop also varies in length from 5 to 7 nucleotides"]. In protein macromolecules there are also invariant positions (*strong*), conservative (*middle*) and radical (*weak*) positions (cf. Dickerson, 1972).

Where the determination of the tRNA *strong - weak - middle* (or *mixed*) relation is concerned, a hypothesis and a prediction (*Prediction 2*) has a sense which hypothesis has to do with the hierarchically ordered system of an entire set of relations of this kind. Therefore, if we take "standard" tRNA to have 76 nucleotides, then we can say of it that it is *strong*; having the

most deviations, tRNA-95, it is *weak* (see Appendix 2); the task of future researchers being to find the hierarchical ordering system of the remaining tRNAs which are *middle*, or *mixed* (Voet & Voet, 1990, p 902: "Presently, the base sequence of ~30 tRNAs from a great variety of organisms are known.... They vary in length from 60 to 95 nucleotides... although most have 76 nucleotides"; Lewin, 1987, p 123: "The numbering system for tRNA illustrates the concrancy of the structure. Positions are numbered from 5' to 3' according to the most common tRNA structure, which has 76 residues. The overall range of tRNA insights is from 74 to 95 bases. The variation in length is caused by differences in the structure of only two of the arms"; Champe & Harvey, 1994, p 379: "Transfer RNAs... have between 74 and 95 nucleotide residues"). On the other hand, standard tRNA can behave at least in three ways: its *basic* behaviour is such as it is: with exactly 76 nucleotides (*strong*); then there is its changed , behaviour in the first degree, which is the result of the least possible change in structure (*middle*), having the replacement of only one nucleotide pair in the characteristic position 3 - 70 (cf. Schimmel & Hou, 1988, p 143: "This data suggests virtually complete conversion of the cysteine-specific suppressor to an alanine-specific suppressor through a simple C3:G70 G3:U70 replacement"; cf. another important commentary as to this unexpected experimental result is: De Duve, 1988, p 118: "In any case, the development of the classical code as a "frozen accident"... becomes more easily understandable if it was superimposed on a more fundamental, deterministic code"); and finally, there is its changed behaviour in the second degree, which must be the result of at least two and/or more changes in the nucleotide pairs (*weak*).

The hypothesis-prediction (*Prediction 3*) as to the hierarchically ordered system of the *strong - middle - weak* relation has to *ipso facto* extend onto the structure and organisation of the different types of DNA's (cf 3rd separate hypothesis in chapter 2). Thus, future research should show if, in the basic level of the hierarchy, determination is as follows: B-DNA, as a right-handed form, with 10 base pairs per turn - *strong*; A-DNA, also right-handed, with 11 base pairs - *middle*; and Z-DNA, as a left-handed form, with 12 base pairs per turn - *weak* (cf. Lewin, 1987, p 66: "The Z-form provides the most striking contrast with the classic structure families. It is the only *left-handed* helix... Z-DNA has the most base pairs per turn of any duplex form, and so has the least twisted structure"; and further on, on the same page: "The c form probably does not occur *in vivo*. Two other forms are the D-form and E-form, which may be extreme variants of the same

form; they... are taken up *in vitro* only by certain DNA molecules that lack guanine"). Thus, future research should show if the following levels in the hierarchy, with whom all the other forms of the molecule DNA, whether they are found *in vivo* and/or *in vitro* (cf. Eichhorn, 1993, p 169: "The binding of increments of metal ions to B-DNA can lead to the conversion of B- to Z-DNA, of Z- to "U"-DNA... and from "U"- to  $\psi$  DNA"). Future research should show whether or not within each of the three DNA families (A, B, Z) determination by way of the *strong - middle - weak* relation occurs (cf. Lewin, 1987, p 65: "Within each family, the parameters can vary slightly; for example, for B-DNA  $n$  could be 10.0 - 10.6").

### 3.3. The Alphabet and the Words

Since the possibility of comparison between the logical (Boolean) square and the genetic code exists, it is important to notice another possible relation of correspondence. From the Boolean square with 4 entities (00, 01, 10, 11) (from the aspect of vector number), three entities (0,1,2) from the aspect of vector weights can be obtained as we have seen. On the other hand, the system of the genetic code with 4 entities (U, C, A, G) can function only by way of three-entity aggregations NNN (N = U, C, A or G), of which there are, both theoretically and experimentally, 64. The following question arises: is the presented (possible) correspondence: the Boolean logical square - the genetic code a nonreality-model (a convention) or is it a reality-model? This question is important if we bear in mind the fact that number 4 is the exact harmonic mean of the number 3 and its double value, the number 6. In more general terms, number 4 is the first possible whole-numbered harmonic mean ( $m$ ) in all binary systems ( $n$ ,  $2n$ ) in which  $m = 4/3 n$ . It then follows that for  $m = 4$  we have the simplest, that is, the best possible harmony which would have to be in union with "symmetry in the simplest case", that is, with binary symmetry, or the best possible symmetry (cf Remark 2.1); also bearing in mind that numbers 3 and 4 "maintain" the middle position in the Boolean cube ( $B^3$ ), dividing the space of the Boolean cube into two equal and symmetrical parts (the middle position in  $B^2$  is "maintained" by numbers 1-2, and in  $B^4$  by 7-8, etc.). If we talk about a reality-model, then we could say that the quantity of the genetic code alphabet the number 4 is the harmonic mean of the word length (the number 3) and its double value. But for now we have no answer to this question; we offer only the prediction (*Prediction 4*) that future research will show that a reality-model is really what is being discussed here.

For future researches it is important that we mention the problem of disagreement between the possible potential harmonious systems and the number of amino-acids. The harmony systems mentioned here are those which appear as a result of mapping the original binary system ( $n$ ,  $2n$ ),  $m=4/3$  in the realized binary system ( $n, m$ ):

### Survey 1.

$(1 \times 3) + (1 \times 4)$	I
$(2 \times 3) + (2 \times 4)$	II
$(3 \times 3) + (3 \times 4)$	III
.....	
$(n \times 3) + (n \times 4)$	n-th

It should be noted that case III in Survey 1 is the case of the best possible harmony and from the aspect of a unique situation: only numerical entities 3 and 4 exist here (cf. Figure 5).

According to our prediction (*Prediction 4*), only by understanding the harmony relations (in the sense of the existence of the codon number relations by means of the harmonic mean) will it be possible to solve the problem of "the degeneracy of the codons as described by the genetic code" (Alvager et al., 1989, p 189). Indications for this follow. Degeneracy I is not a degeneracy. Degeneracy II is the first case of degeneracy, that means that it is the first possible case, or the lowest degree of, synonymy. The question as to what should the highest degree of degeneracy, or synonymy be, arises. Many solutions are possible. But there is only one solution, and a unique case at that, if we want all the degeneracies between the two extreme cases (the case of the minimal and maximal degeneracy) to be cases of harmonic means of those two extreme points. That is the case with degeneracy VI. If this is so, then degeneracy III makes sense because the number 3 is exactly the harmonic mean of the numbers 2 and 6 ; also degeneracy IV makes sense because the number 4 is the harmonic mean of the numbers 3 and 6. By the way, the numbers 2 and 3 give the harmonic mean which is a whole number only with the number 6. Degeneracy V, as we can see, makes no sense. Here it is very important to notice an important principle, which holds for the *arithmetic of the genetic code*: when one thing happens, then everything happens! The number 6, namely, is the first perfect number, and by means of this number all relations between it and all the other perfect and

friendly numbers are established (see Appendix 3), as well as all cases of symmetry, harmony and proportionality of the genetic code which we are presenting in this study.

The physical sense of harmony degeneracy, presented in above, is impossible to understand without referring to the *information theory*; of course, on condition that it has been previously proven that the relation between the *information function* and the characteristics of the genetic code makes physical sense (Alvager et al., 1989, p 192: "We have defined the information function for any distribution of degeneracies in the genetic code... .Obviously, the function will have physical meaning... "). The first condition for beginning to study the degeneracies of the genetic code by means of the *information theory* is accurately defined by the following view: "To describe the distribution of degeneracies of the genetic code it is convenient to group all codons together in a single block that code for the same amino acid" (Alvager et al., 1989, p 190). If we linger on this problem a while longer, we will see that this view, as far as it relates to the physical sense of codon grouping, must take for granted "that each of the three codon bases has a general correlation with a different, predictable amino acid property, depending on position within the codon" (Taylor and Coates, 1989, p 177). In other words, if triplet codons are those same codons which code for amino acids, then that function cannot be shared by doublet, quadruplet, nor any other type of codons (Taylor and Coates, 1989, p 185: "Our observations, regardless of interpretation, are highly incompatible with the view that the codon assignments were largely the product of chance"), which is argumented by proof, given in the following chapters. Regarding all that has been said, it makes sense to imagine ideal situations whereby amino acids build themselves into proteins, doing so with equal probability and equal speed, on condition that the codons (for each individual amino acid) fall in or become connected in the same way: with equal probability and equal speed. Such a case of occurrences would resemble the state of greatest *entropy*. All realistic cases of amino acid build-in and the use of codons ("codon bias") represent, then, suitable measureable divergencies (Gatlin, 1972, p 68: "... it is also possible to describe all divergence from the maximum entropy state purely in terms of divergence from equiprobability of the *n*-tuples"). In this manner we come to the physical sense of harmony, which can be most easily understood by means of a concrete example. If, under the above-mentioned conditions of equal probability, we have only two amino acids, one with degeneracy III and the other with degeneracy VI, then for one time interval, for one amino acid, on average "activates" or

includes, not the number of codons as given by the arithmetic mean (4.5), but as many codons as given by the *harmonic mean* (4.0). This is what follows from basic statistics.

## 4. GENETIC CODE NUCLEON NUMBER

Besides the strict, above-shown regularities of the genetic code, regularities characterize the genetic code in other ways as well. If codon systematization is observed not only in quartets but also in octets (Rumer, 1966) we get exactly two classes of separate binary symmetrical codon doublets (the first and second base of the codon), the first class being within the first octet, and the second class within the second octet (Table 1 in Shcherbak, 1989, p 272). The ratio of the number of doublets is 1:1 (or 8:8); which also corresponds to the number of codons in the two classes: 1:1 (32:32). As to the codon-coded entities (for amino acids and/or for termination entity), the ratio of four-codon and non-four-codon entities is 1:2 (that is, 8:16). Finally, the relation of the number of "strong" (C,G) to the number of "weak" bases (U,A) in codon doublets of the first octet is that of 3:1, whereas that relation in the second octet is 1:3.

### 4.1. Union of Chemistry, Physics and Boolean Arithmetic

However, what is in a way unexpected and most surprising is the fact that (binary) symmetricality and proportionality is achieved through the number of nucleons (Fig. 1 in Shcherbak, 1994, p 475). Namely, from the aspect of nucleon number, 16 of the non-four-codon entities are symmetrically separated into the "head" and the "body" (the side chain) in one way, and the 8 four-codon entities in another. The first way "uses the same symbols", and the second way has "the numbers arranged by cyclic permutation", but in both cases the numbers in question are those taken from the table of the multiples of the number 037, which form a system arranged in accordance with module 9 (Table 1 in Shcherbak, 1994, p 476). The relation of the number of nucleons in the "heads" and "bodies" of non-four-codon entities is that of 1:1 (that is, 1110:1110), whereas the relation of the whole (molecule) to that of one of its individual parts "heads" and "bodies" is 2:1. On the other hand, the relation of the number of nucleons in the "heads" to that of the "bodies" of the four-codon entities (amino acids only!) is 16:9; that is, when the wholeness of the molecule is taken into account, proportionality is then reduced to very small numbers, not to any number,

but to those numbers which demonstrate the squares of the first three Pythagorean numbers  $3^2:4^2:5^2$ . (*Hint.* Not only the total amino acid nucleon number, but also the total *pu-pyr* nucleon number is related to the multiples of 037; cf footnote in Shcherbak, 1994, p 476).

Shcherbak's assumption that "the laws of additive-position notation of numbers... have analogies with quantum physics" (Shcherbak, 1994, p 476) make sense and is scientifically based only on condition that "the numbers" generate in the Boolean space: "When the highest state  $a_j = q$  is attained, there is a quantum transition to a further level  $j + 1$ , its value being increased by unity and so forth" ("where  $a$  are natural series symbols...  $q$  is the base of the system"). If this is so, and if for the number  $037_{10}$ , from all the two-digit numbers, the law of modulo ordering by way of "the same symbols" and/or "cyclic permutation" (Shcherbak, 1994, p 475) holds, then we can give an important hypothesis-prediction (*Prediction 5*) which follows. Position 37 and its inverse form 73 must have a specific role in the make-up and organization of all linear-generated bio-macromolecules, as nucleic acids are, for example, and proteins and their aggregations. This must also be the case for positions 03/30 & 07/70, 33 & 77, etc. On the other hand, by means of these positions, relations among all other positions in the bio-macromolecule are established (*see Appendix 2*). (cf Osawa et al., 1992, p 234: "Nucleotide 37, adjoining the third anticodon position in tRNA, is often extensively modified... .In contrast, the tRNA nucleotide at position 33 is an invariant unmodified U"; Saenger, 1984, p 347: "The hypermodified, semi-invariant base at position 37... serves to locate the codon in proper register and to prevent misreading due to frame shifting").

What is especially important is the fact that determination by way of nucleon number holds not only for the categorization of amino acids into four-codon and non-four-codon amino acids, but also for categorization within the four stereochemical types (*see Surveys 2-3 and Solutions 1-7; see in Popov, 1989, p 79, about the categorization of amino acids into four types; see also Discussion*).

### Survey 2

<b>S</b>	31	121
<b>V</b>	43	
<b>C</b>	47	
<b>T</b>	45	177
<b>I</b>	57	
<b>M</b>	75	

<b>N</b>	58	189
<b>D</b>	59	
<b>K</b>	72	
<b>Q</b>	72	245
<b>E</b>	73	
<b>R</b>	100	

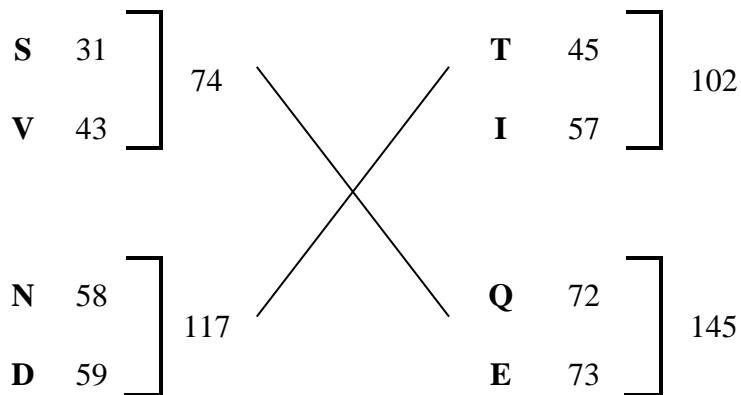
### Survey 2 (I)

<b>S</b>	31	132
<b>V</b>	43	
<b>N</b>	58	
<b>T</b>	45	174
<b>I</b>	57	
<b>Q</b>	72	

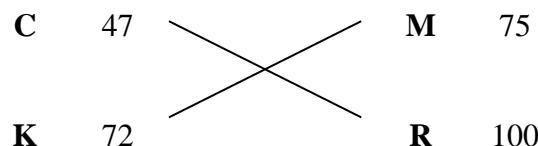
  

<b>D</b>	59	131
<b>K</b>	72	
<b>E</b>	73	
<b>R</b>	100	173

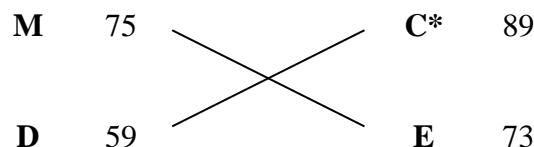
### Survey 2 (II)



### Survey 2 (III)A



### Survey 2 (III)B



### Survey 3

A	15	L	57
H	81	W	130
F	91	Y	107

$$A(15) + Y(107) = 122 \quad (1)$$

$$F(91) + W(130) = 221 \quad (2)$$

$$F(91) + L(57) = 148 \quad (3)$$

$$[2 \times (N, D, K)] + [1 \times (T, I, M ; H)] = 1 \times (S, V, C ; F, Y ; A, L ; Q, E, R) \quad (4)$$

$$(2 \times 189) + [1 \times (258)] = 1 \times (636)$$

$$[2 \times (N, D, K)] = (S, V, C ; F, Y ; A, L ; Q, E, R) - (T, I, M ; H) \quad (5)$$

$$(2 \times 189) = (1 \times 636) - (1 \times 258)$$

$$(1 \times 378) = (1 \times 378) \quad (6)$$

$$(1 \times 378) = (1 + 2 + \dots + 26 + 27) \quad (7)$$

In Survey 2, we can see that the five pairs of the alanine type (S-T, C-M, N-Q, D-E, and K-R) plus one pair of the valine type (V-I) form the specific Boolean logical square: (0) N,D,K - (1) S,V,C - (2) Q,E,R - (3) T,I,M, for which "the law of diagonals" holds: the number of binary values along its diagonals is equal. In the original Boolean square:  $0 + 3 = 1 + 2$ ; here, in this model (reality-model!):  $189 + 177 = 121 + 245 = 366$ , or read according to module nine:  $0 + 6 = 4 + 2 = 6$  (zeroth position for D,N,K is expected from the reasons of chemistry: D is the simplest *acid* amino acid; N is its the simplest derivative; K is the simplest *base* amino acid. This position is expected also from the position D-N-K in Table 1, in contrast to position for E-Q-R). Within this "large" square there are the "small" squares (Surveys 2 - I, II, III) which show a unified categorization from the chemical aspect (the chemical nature of the molecules) and the physical aspect (the number of nucleons in the atom's nucleus which form its molecular composition). That union is achieved, as we can see, by way of strict Boolean arithmetic (including Boolean logic and algebra). What is especially important here is the fact that as the final result we have the pairing off of two *base* amino acids with the pair C-M and of the two *acid* amino acids with the pair M-C\* (*Crossing over! See* the next chapter-section 4.2.), where, in the first case, the "player" is sulpho-cysteine, and in the second case, seleno-cysteine (in the first case with the first i.e. lightest isotope of sulphur S-32, and in the second case with the first isotope of selenium Se-74).

As far as the three pairs of the stereochemical alanine type (A-L, H-W, and F-Y in Survey 3) are concerned, the Boolean square does not exist, or if it does exist, then it is not so easily noticeable; this has to be established by

future research. However, the connectedness of these three pairs with the six pairs in Survey 2 is clearly obvious. Thus in Solution (1) we can see that "the diagonal" A-Y corresponds to the both diagonals of the "large" square in Survey 2, and that in the following manner:  $122 \times 3 = 366$ ; the diagonal F-W in Solution (2) is an inversion of its previous one (122|221); finally, the diagonal F-L in Solution (3) is reduced to the same numerical value which we have found for the diagonal M-E, or D-C\* in Survey 2(III)B. In Solutions (4-7) further proportional relations between the system with *six* and the system with *three* pairs of amino acids are presented (see Survey 1, chapter 3.3, for the way in which quantities 3 and 6 are conditions for the best possible harmony; in chapter 6.3. see why W must be excluded). ... In the final result we can see the connection between the two systems determined by the sequence of numbers 1-27 (Solution 7), by whose sequence the categorization of amino acids into four-codon and non-four-codon amino acids has been determined (cf. Table 1 in Shcherbak, 1994, p 476; as to why the amino acids G, P, and W have been excluded, see Chapter 6.3; the connection between the 18 amino acids from the two yes-pairing types and the two amino acids from the 2 non-pairing types - G & P - remains to be discussed in another article; see Appendix 7).

## 4.2. Amino Acid Systematization

The important fact is that the four-codon entities (through the nucleon number) are separated into a specific system. But the number of nucleons (their order according to quantity) has also been used for the further classification (and systematization) of codon entities into: one-codon, two-codon, and three-codon entities. In the drawing (Fig. 1 in Shcherbak, 1994, p 475) it can be seen that in the non-four-codon entity system in the end positions are located, on one side C-I, and on the other side M-W. Using these same amino acids, the symmetrical division of the Gray genetic code model has been determined (Fig. 1 in Swanson, 1984, p 188). The amino acids from the pair M-W divide in such a way that one occupies the "north" position and the other the "south" position in the codon ring. There are exactly 32 codons to the left of the north-south line, and 32 to the right of this line as well (including M-W). Behind the M-W pair and moving to the right the C-I pair follows, but in such a way that in relation to the nucleon system they realize an interesting "crossing-over". This happens in the following way: In the *nucleon system*, observing the order, we have the C-I and the M-W pair; in other words, the C-M and the I-W pair (the first amino

acid pairs off with the first, the second amino acid with the second); in the codon ring then a "crossing-over" occurs, thus we have the pairs W-C and M-I.

**Remark 4.1.** The classification of *Cys* into that of three-codon amino acids, as has been done in Figure 1 in Shcherbak, 1994, p 475, can be understood only when it is known that there is a non-standard genetic code in which, besides UGU and UGC which code for *Cys*, still UGA codes for the same amino acid (cf. Osawa et al., 1992, p 250: "In *Euplotes octocarinatus*, UGA is a codon for *Cys*, in addition to the universal *Cys* codons UGU and UGC").

In order for us to follow all these relations (but not only that) it is necessary to "transform" the Gray code model into a binary tree (Fig. 1 in this Supplement). This transformation will also entail another "crossing-over", so that the W-C pair becomes the C-W pair, and the M-I pair becomes the I-M pair. The binary tree itself reflects the natural binary division of the Codon ring into 32 codons of the *pyr* type (the left side of the tree) and 32 codons of the *pu* type (the right side of the tree).

**Remark 4.2.** In civilizations in which the writing process moves from left to right the zeroth codon UUU has to be on the far left side; in civilizations in which writing is directed from right to left, it is the opposite; in civilizations where writing moves from the bottom up, the zeroth codon UUU has to be on the bottom; in civilizations where it is the other way around (from top to bottom), the zeroth codon has to be at the top.

Now we can see how the symmetry of the codon ring has been determined by the M-W and I-C pairs (which correspond to the determination of the *nucleon system* by the pairs M-W and C-I). The division of the codon ring into 32 *pyr* and 32 *pu* codons, in comparison to the division of the ring by the M-W line, has been achieved through another binary and symmetrical *crossing-over*, which, with the use of the binary tree of the genetic code, is easily noticeable: at position M the *pyr* half of the codons has been "reduced" by exactly one codon class, or by "rosette" 3 (T,A), whereas at position W the *pyr* half of the codons has been "increased" again by exactly one codon class, that is, by "rosette" 5 (C,W,R).

## 5. HIERARCHY OF BOOLEAN SPACES

When the Watson-Crick table was first presented in the form of a codon cube (Fig. 64 in Eigen & Schuster, 1979), it was not possible then to expect anything in the way of a reality-model, much less the Boolean cube  $B^3$ . However, with the presentation of the "codon path cube" (Fig. 2 in Swanson, 1984, p 189) there was no doubt about it, all the more so since out of six possible choices in the Gray code model ( $B^6$ ), the situation is exactly

balanced with the generation of  $B^3$ : the second base was chosen with both questions being taken into account (base type and number of hydrogen bonds) and the first base was chosen with the first question being considered (base type). Choosing, and doing so on the basis of only one question, means to choose! Choosing the first and second base and not the third, is a strict rule which can be otherwise expressed as choice according to the model of "*two out of three*" (cf. with the reading "*two out of three*" in the codon-anticodon system in Lagerkvist, 1978 and Lagerkvist et al., 1981). Knowing that, in the coding process according to the Watson-Crick Table, *mutatis mutandis*, only the first two bases are coding, and the third is noncoding (Lewin, 1987, p 129: "The pattern of third base degeneracy... shows that in almost all cases either the third base is irrelevant or a distinction is made only between purines and pyrimidines."), we can say that the choice according to the "*two out of three*" model is such that we are talking about a reality-model; therefore, the generated Boolean cube  $B^3$  (generated after the third choice) is also a reality-model. The relations of codon entities and amino acid entities in such a model are in fact shown in the genetic code binary tree (Fig. 1).

The achieved balance after the third choice is one in the sense that besides the number of chosen bases there were exactly the same number of those which were not chosen. The second base based on both questions and the first base based on one question were chosen; but the first base with one question and the third base with two questions were not. It is important to notice that at the realization of the six choices the first and second base are chosen by the essential presence of "*crossing-over*", whereas the choice of the third base takes place without it (Swanson, 1984, p 188: "Note the interleaving of the Gray code bits representing the first and second bases of codon.").

However, the "*two out of three*" and the "*crossing-over*" principles are fully observed with the fourth choice: the first two bases are then fully chosen, while the third base in the codon is not. This is the reason why the genetic code binary tree (Fig. 1), that is, the Gray code model (Fig.1 in Swanson, 1984, p 188) represents the unity of the Boolean cube and the hypercube  $B^3 - B^4$ . In accordance with this, it is understood that besides the eight large rosettes in the genetic code binary tree (0-7), there simultaneously exists 16 small rosettes (0-15).

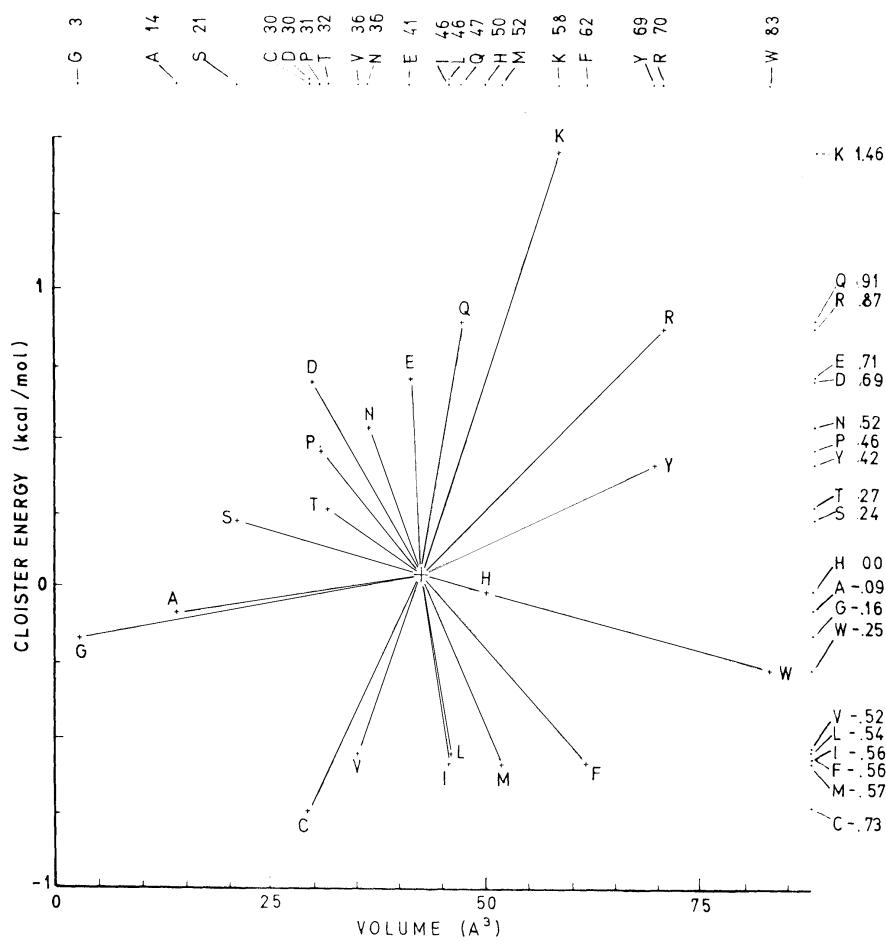
## 6. PARAMETRIC RELATIONS

The informed reader will find it easier to see a hypercube in the binary tree (Fig. 1) than a cube, but the physical and chemical parameters, nevertheless, give priority to the cube. If the whole Boolean space of the cube is divided into two equal (and symmetrical) parts, into space-3 and space-4 (a harmonious division in the sense of the discussion given in chapter 3.3.), we will get two classes of amino acids (*Note*: Vertex 3 and adjoining vertices 1,2 and 7 form space-3; vertex 4 and adjoining vertices 0,5 and 6 form space-4.). Space-3 contains 9 amino acids: T,A; S,P; I,M,V; **R**,G. Space-4 contains 12 amino acids: Y,H,Q; F,L; C,W,**R**; N,K,D,E. Because the amino acid **R** appears in both spaces, the number of amino acids "increases" by 1 so that now there are "21 amino acids".

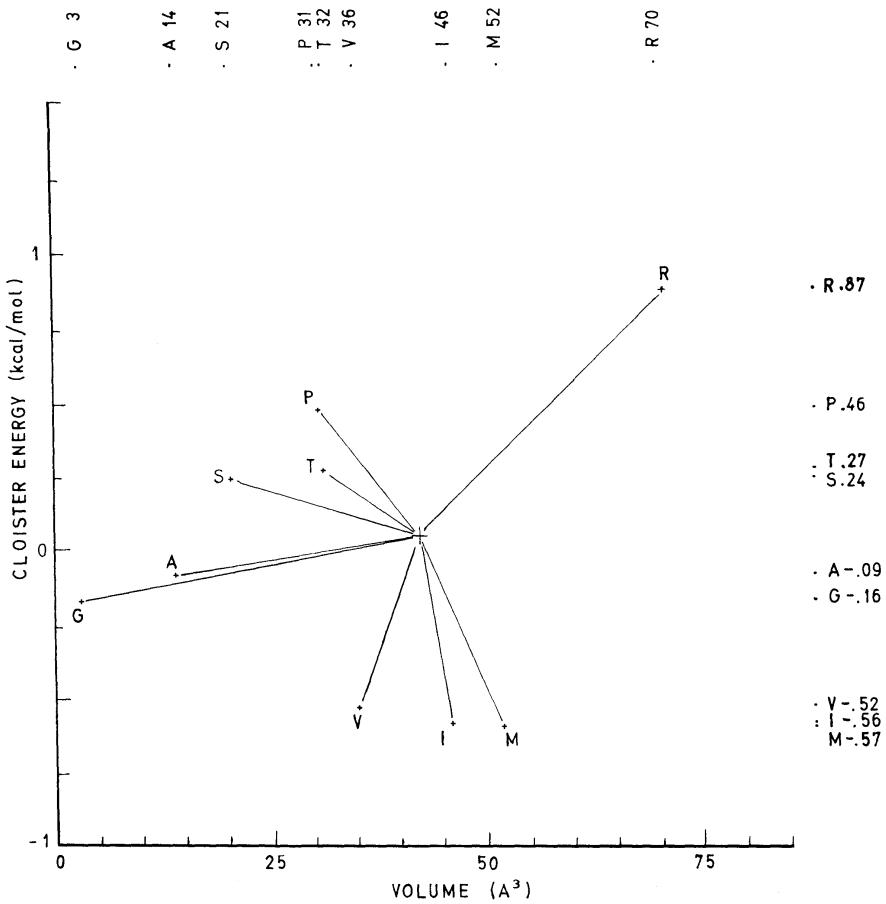
### 6.1. The Three Rings

With division of the amino acids into two classes within Space-3 and Space-4 the existence of the physical properties ring, along with those of the Mutation ring and the Codon ring (Figs. 4,3,1 respectively, in Swanson, 1984) becomes evidently clear.

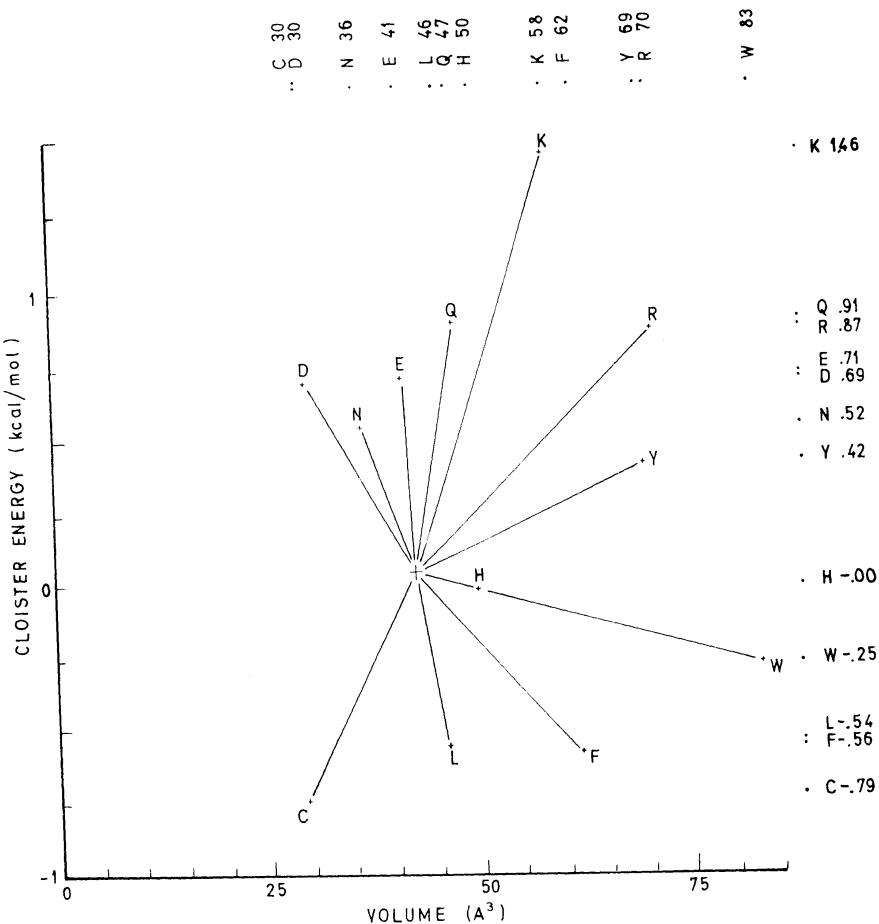
If, in Figure 2 (Fig. 4 taken from Swanson, 1984, p 192) through the arc which passes through the points P-M and the arc which passes through the points D-F we divide the space into two parts - the right and left - then in the upper left part or in its adjoining area can be found amino acids from Space-3 (G,A,S,T,P), whereas amino acids from Space-4 (D,N,E,Q,K,R,H,Y,W) can be found in the upper right part or its adjoining area. After this division, in the lower part of the ring (far from the top part and its adjoining area) remain amino acids which are located exactly on the two arcs: on the arc V-I-M are the amino acids from Space-3 and on the arc C-L-F are the amino acids from Space-4. (Notice that the two arcs intersect! *Crossing-over?*). The total division of the amino acids is shown for Space-3 on



**Figure 2** Physical properties ring. According to R. Swanson, 1984 (Fig. 4, p. 192). Cloister energy vs. volume on exact scales.



**Figure 3** The amino acids from the Boolean Space-3. Those are the amino acids from rosette 3 on the binary tree as well as the amino acids that represent the “adjacent” neighbours. Such amino acids represent one of the two symmetric “halves” originated by “cutting” of the physical properties ring given in previous Figure (Fig. 2).

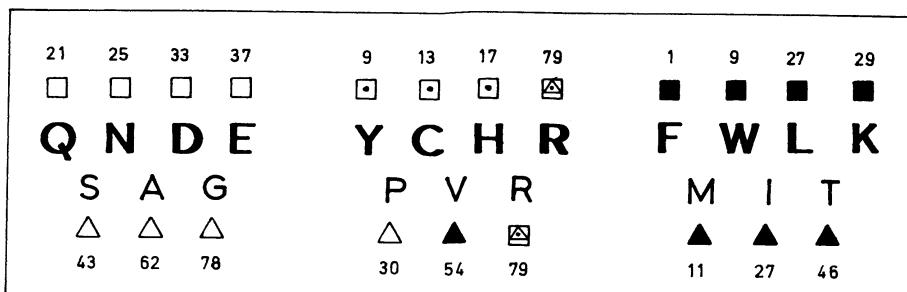


**Figure 4** The amino acids from the Boolean Space-4; they generate analogously to Space-3 amino acids such as represented in the previous Figure (Fig. 3)

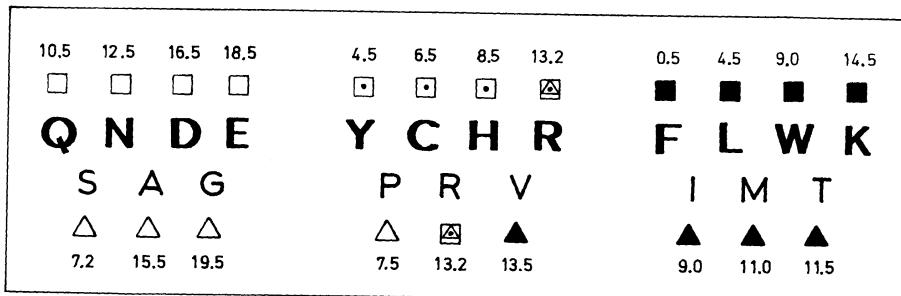
Figure 3, and for Space-4 on Fig. 4. If we bear in mind that the physical properties ring, i.e., plot, represents the distribution of amino acids, determined by "cloister energy vs. volume on exact scales," and that "cloister energy" is essentially "a formal free energy... of transfer of the amino acid from the outside of a protein to the inside" (Swanson, 1984, p 196), then the importance of this result (presented on Figs. 2,3 and 4) becomes clear. Thus we get the answer to the 9th question in the *Introduction* in the sense that we can now more clearly see which *factors* determine the replacement of amino acids in proteins. It has been shown that

they are the size of the amino acid molecule (volume!) and polarity (cf. Doolittle, 1985, p 76), but another factor has been revealed - the position of the amino acid within the Boolean space.

With the discovery that an important factor is the position of the amino acid within the Boolean space, it is simultaneously revealed that an important factor is also the chemical *structure* (correspondent to *essentiality*) of every individual amino acid (Figs. 5-6). It is shown in Fig. 5 that with the discovery of the classification (and systematization) of amino acids into the Space-3 and Space-4 class, it becomes clear that the *essentiality* of amino acids (chemical structure complexity) is determined by the relation of *strong-weak-middle(mixed)*, in the sense of strong (essential) - weak (nonessential) - middle (semi-essential) for amino acids from Space-3 and Space-4 and still "mixed" as well for amino acids from Space-3. Of course, the *full-empty-semi* relation can also be adequately used. But the surprising thing in this is the fact that the classification of the "21 amino acids" in Fig. 5 is in total agreement with the best possible harmony, that is, with system III given in Survey 1 in chapter 3.3. (The criteria for assigning the numerical values that are listed in Fig. 5 can be found in Rakoëviæ, 1990, p 11 and 1994, p 72).



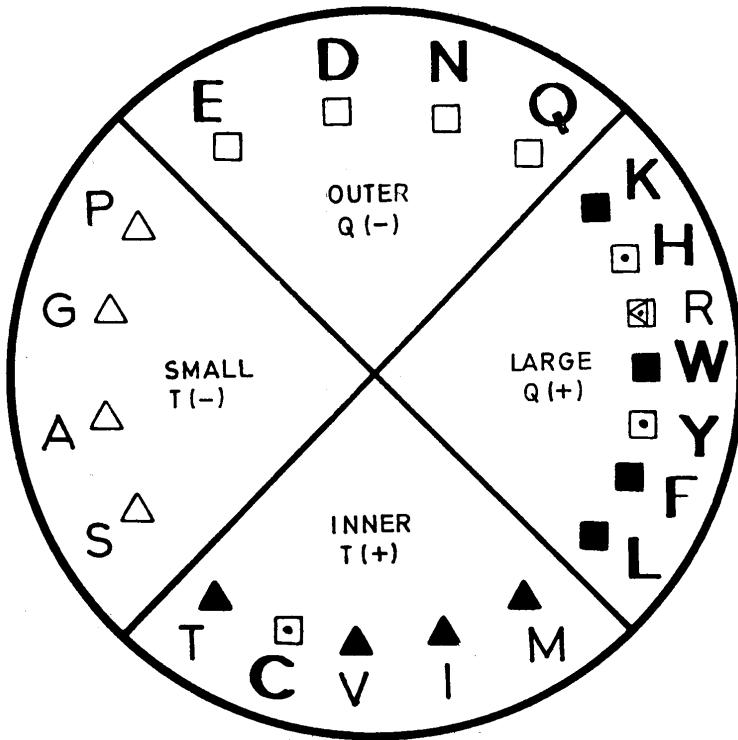
**Figure 5(I)** Here are given the amino acids from Space-4 (squares) and from Space-3 (triangles) depending on the binary value (strict order per rising values in each group) and depending on the best possible harmony. It becomes evident that there exists the categorization of the amino acids under request for the best possible harmony:  $(3 \times 3) + (3 \times 4)$ . The mean binary values are given in Figure according to Rakoëviæ, 1990, p 10-11; 1994, p. 72. Note that the sequences (Q, N, D, E) and (S, A, G, P) are the same as in Mutation ring.



**Figure 5(II)** Here are given amino acids from Space-3 and Space-4 as in previous Figure but here are taken the collective binary values (cf Rakočeviæ, 1980, p 10). In case of nonexistence of such values, the categorization (3 x 3) would not make any sense for Space-3, but only the categorization (2 x 4). Note that the sequences (Q, N, D, E) and (S, A, G, P) are the same as in Mutation ring.

With this we have total and definitive proof for the existence of the *Input* (Essential amino acids) - *Output* (Essential, semiessential and non-essential amino acids) relation presented in the second working hypothesis (Chapter 2). Of course, here it is understood that, for organisms which first came into being, all the amino acids had to be essential (in other words, non-essential, depending on the view; they were non-essential in the sense that the organisms themselves were able to synthesize all of them).

It is important to notice that with the systemization of amino acids, as given in Figure 5, the problem of amino acid classification is solved, and from the aspect of essentiality, that has been achieved according to the model 10 : 10 or the model 8 : 4 : 8 (4 semiessential amino acids). The surprisingly large number of different views about this problem, which we have mentioned in a previous study (Rakoèeviæ, 1994, pp 84-85) now acquires a simple solution: amino acids must first of all separate into those of Space-3 and Space-4; only then can their essentiality be analyzed.



**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.

Surprises, however, do not stop here. We can see in Figure 6 how the arranged system of essential amino acids, determined by Space-3 and Space-4, brings order to the relations among the amino acids within the mutation ring. The essentiality of amino acids and the relation between Space-3 and Space-4, in fact, reveals that the Mutation ring (Fig. 3 in Swanson, 1984, p 191 - Mutation ring I) must exist in yet another form, as

shown in Fig. 6 in this paper (Mutation ring II). It should be noted that in the half of the ring with the non-essential amino acids there is no "Crossing-over", whereas in the half with the essential amino acids the "Crossing-over" exists: *full* (strong) "Crossing-over" for C, *semi* (middle) "Crossing-over" for R (and *empty* "Crossing-over" for other amino acids.). Binary symmetry is evident and so is proportionality: yes-essential: non-essential = 3:2 (or 12:8); non-essential from Space-3: non-essential from Space-4 = 1:1 (or 4:4); yes-essential from Space-4 (including R): yes-essential from Space-3 (excluding R) = 2:1 (or 8:4). Including and excluding the amino acid R is also another specific way of "Crossing-over".

With such a view regarding the structure of the mutation ring, we can be certain that it is the result of *a representative sample* not only from the aspect of the number of analyzed proteins, but also from the aspect of an a long enough passage of time in the process of evolution (Dayhoff, 1969; 1972-1978; Swanson, 1984). Bearing this knowledge in mind, it follows that in the *input* (codon ring) - *output* (mutation ring) relation, *feedback* had to exist, and had to be negative. But what does that practically mean? It means that with a sufficiently large number of "dice throws" (the replacement of amino acids in proteins as a result of mutations), the relations among the amino acids in the proteins have come to be the same as those originally found in the genetic code. There is, therefore, no discontinuation, which means that the genetic code was originally the same as it is today. Mutations, even when they are "obviously" neutral in fact are not neutral. All of them are an indispensable part of the whole, representing at least the smallest pebble which are, one by one, continually and gradually built into the mosaic, which after a long enough time forms in such a way as to be the exact copy of the mosaic originally contained in the genetic code itself. If all this is so (and this follows from the analyzed results), then we have a full and definitive proof for the existence of a Codon ring (*Input*) - Mutation ring (*Output*) relation (First working hypothesis in chapter 2); then not even the Non-Darwinian evolution existed, being that it was based on wrong suppositions about the possibilities of neutrality for mutations. With this the answer to question 10 from the *Introduction* has been given.

With accurately argumented proof that Boolean spaces are actually the main determinants and invariants of the genetic code, then non-Darwinian evolution, *per se*, is not possible. However, independently of this, in the very act of founding the theory of non-Darwinian evolution many methodological mistakes have been made. The main experimental result on which the findings of this theory has been based (Figure 1 and Table 6 in

King & Jukes, 1969, p 796) originated from a *selective*, instead of a *representative* sample ("53 completely sequenced *mammalian* proteins") (*italics* M.R.). On the other hand, basing their theory on the genetic essence of being, the authors of the mentioned theory, have again made a mistake. They have started from both the genotype and the phenotype model, that is, from two entities of the genetic essence of being, as defined by Johannsen (1909, 1913), and which are in their sense non-reality models or conventions; instead of starting from reality-models, as perceived and defined by Mendel (1866) and who have reduced the entity number to two instead of four, with a strict mathematical interdependence: Stammarten - Konstante Formen - Glieder - Individuen,  $1^n - 2^n - 3^n - 4^n$ , respectively. In other words, using modern terminology, we can say the following: Parent type - Phenotype - Genotype - Individual type (see Rako~evi}, 1994, pp175-177 for details).

## 6.2. Codon-Anticodon and Codon-Amino Acid Relations

The results, which we are presenting above in so straightforward a manner, are also strengthened by the results given in Table 1. The relations among the amino acids given in this Table are, in fact, "copied" relations of the amino acids united in the Codon ring - Mutation ring system (see Appendix 1). As we can see, the amino acids are strictly divided on the basis of the "key" of positive and negative values of a very important parameter, that of hydropathy (Doolittle & Kyte, 1982).

The presented conclusion may be surprising for the reader because it has been drawn only from the relation (and interaction) of codons - amino acids (cf Reuben & Plok, 1980, p 111: "The genetic code appears be the 'fossil record' of nucleotide - amino acid interactions in the prebiotic milieu"). At first sight it seems as if the codons - anticodons interactions have not been taken into consideration at all. The contradiction, however, disappears when the following two things are understood:

**Table 1.** Cyclic and Periodic Amino Acid System: Zeroth Step.

	6	7	0	1	2	3	4	5	border
one-meaning	0	D -3.5	N -3.5	G -0.4	A +1.8	W -0.9	Y -1.3	V +4.2	I +4.5
	1	E -3.5	Q -3.5	—	F +2.8	H -3.2	—	M +1.9	—
two- meaning	2	K -3.9	—	—	—	P -1.6	T -0.7	C +2.5	—
	3	R -4.5	—	—	—	—	S -0.8	—	L +3.8

*The System is generated as a result of the mutual relations of Codon ring, Mutation ring and Physical properties ring. The values are given for hydropathy. (See text for details).*

1. The codon - anticodon interactions are also determined by the *strong - weak - middle*, i.e. mixed (*full - empty - semi*) relation. Bearing in mind the idea of "pairing at the third position of the codon" (Table 1 in Crick, 1966b, p 552), as well as the "present status of wobble usage" (Table 2 in Osawa et al., 1992, p 233), we can unambiguously conclude the following: as far as the possibility of one-meaning for the codon - anticodon interaction is concerned, base *A* possesses maximal one-meaning ("Strong" or "Full"), and a minimal multiple-meaning ( $A \rightarrow U$ ), whereas base *U* possesses minimal one-meaning ("Weak" or "Empty"), and a maximal multiple-meaning ( $U \rightarrow U, C, A, G$ ). From the point of view of chemistry, we can see that one-meaning is connected to the presence of *amino* group in the basic state of the base molecule, whereas multiple-meanings are connected to the presence of the *hydroxyl* group. It therefore follows that we have maximal one-meaning when only one amino group is present and a minimal number of groups at that, only one group; and we have minimal one-meaning when only the *hydroxyl* group is present and the maximal number of groups at that, a total of two. Between these two extreme states, we have the "*middle*" ("semi") or "mixed" states. Thus, hypoxanthine, containing only the *hydroxyl* group (and not an amino), has the tendency of being one-meaning, but a little less than *U* (hypoxanthine  $\rightarrow U, C, A$ ) because it has fewer

*hydroxyl* groups. For *C* and *G*, then, everything is clear. These are bases which have been originally "*mixed*"; they contain both an *amino* and a *hydroxyl* group, thus, they are in their codon - anticodon interactions, from the aspect of one-meaning - multiple-meanings, also "*mixed*" states. It is also clear why all the modified forms, for all five bases, have to be "*mixed*" states. It is our prediction (*Prediction 6*) that, for future research, all the "*mixed*" states are in one hierarchical system of meanings, which are further determined by the *strong - middle - weak* (i.e., *full - semi - empty*) relation. When pursuing this line of thought, "the order of things" is confirmed, many controversies will be solved, one of them being, for example, the "Four-way wobbling... versus the two-out-of-three mechanism." (Osawa et al., 1992, p 234).

2. All codon - anticodon interactions (*input*) are strictly (deterministically) "timed" into *output*: codon - amino acid relation (i.e., interaction) in the following sense (see the first working hypothesis in Chapter 2). The codon - amino acid relations have to be such as to be strictly determined by the sequency of a series of numbers within the borders of module 9 (0, 1, 2, 3, 4, 5, 6, 7, 8, 9) with a total of 9 possible congruent classes and being such that they correspond to the key quantum (physical) analogue, with values for the "quantum magnetic number"; and they must be such that they strictly correspond to the *Golden Section*, i.e.. to the Fibonacci series. We shall directly prove that this is, in fact, so.

Within the Standard genetic code there is exactly *one* situation when three codons give the meaning (I), *two* situations when one codon gives the meaning (M,W), *three* situations when one codon gives the meaning (M,W, the *opal* "stop" codon); *four* situations do not exist; there are *five* situations when two codons give the following meaning (Osawa et al., 1992, p 233: "The five sets that end with purines are UUR, CAR, AAR, GAR, and AGR... "); *six* situations when to the previous five we add another two-codon situation (*ochre* and *amber* "stop" codons); there are *seven* situations which form "the seven pyrimidine-terminated sets" (Osawa et al., 1992, p 233); and, finally, we have *eight* situations which form "the eight cases in which a single amino acid has four codons" (Osawa et al., 1992, p 233). As we can see, the sequence of the number series within the borders of module 9 has been realized, but in such a way as not to realize the first and last "point" (0 & 9) and not realize the lower border of the middle point pair of module (the number 4 in the pair 4-5).

On the other hand, we have a realization within the number which on the model of the Boolean hypercube represents the inverse form of the module

9, and that is number 6 (9 1001 / 0110 6). The realization of the inverse form of the module coincides, in fact, with "the number of degeneracies in the SGC and some mitochondrial genetic codes" (Alvager et al., 1989, p 190), those numbers being 1, 2, 3, 4 and 6; thus the number 5 is the only number that is not to be realized as the upper border of the middle point (pair 4-5) of module 9. When everything within module 9 and its inverse form (number 6) is considered, the realization of situations with numbers 1, 2, 3 and 6 (the factors of the first perfect number - the number 6) is "*full*"; the realization of situations with numbers 7 and 8 is "*semi*" or "*middle*" due to the fact that they are realized only within the sequences of the module (with 9 situations), and not within the sequences which form the inverse form of the module (with 6 situations), and, finally, the realization of situations with numbers 0,9 (and 4,5?) is "*empty*", due to the fact that these situations are not realized. The first and last number, i.e. the outer pair (0,9) [and the middle pair (4,5)?] within the module 0-9 are not realized. There is complete realization of the pair (3,6) which "touches" the middle pair (4-5), while the state of the remaining pairs is as follows: in the pairs (1,8) and (2,7) the first member is realized both times, while the second member is realized only once. (*Hint.* The numbers 7-8 represent the middle pair of numbers in the Boolean hypercube). As we can see, the physical sense of the inversion is more than obvious: the number of codons having the same meaning (with how many codons in every individual combination) versus the number of sets of codons having the same type of meaning. (The meaning of codons in one set is the same; but it is different as we move from set to set).

The key quantum model (reality-model) - the quantum magnetic number, holds for stable chemical elements having the form of: -3, -2, -1, 0, +1, +2, +3. In strict analogy to this model, the codons, with their amino acids and/or their "stop" situations within the "Floor-Table" of the genetic code, start separating (Table 4.1. in Rakoèeviæ, 1994, p 56). If we take the "Floor-Table" to have four floors (with the middle base U,C,A and G, respectively), then the simultaneous separation of the amino acids and their codons occurs as follows: Analogue "0" in the quantum model corresponds to the division along "line 9" in the "Floor-Table", where the "line" passes through all the four floors without any interruption. Then, on the left side, we have C, opal, W, Y, ochre and amber, H, S, P, F, L and I (AUU & AUC), and on the right side we have M (AUA & AUG), V, A, T, E, D, K, N, Q, G, R, S and R. In regard to the Watson-Crick Table, "line 9" represents the "ladder" by which codons of the *pyr* type separate from codons of the *pu* type, and they do so

according to a strict rule which can be designated as "the unit change law" (Rakoèeviæ, 1994, p 36). The left side of the "Floor-Table" corresponds to the upper part of the Watson-Crick Table, whereas the right side corresponds to the lower part. The ratio of *pyr* and *pu* two-codon situations from the upper and lower part of the table is then as follows: (5:3), (4:4), (3:5), (2:6).

In area "-1" of the quantum analogue, we have: C, opal, W, ochre & amber, H and P, whereas in area "+1" of the analogue we have: T, N, Q and R (with four codons). The ratio of two-codon situations here is: (2:2), (2:2), (2:2).

In area "-2" we have : Y and S (with four codons), whereas in area "+2" we have: A, M, D, K, R (with two codons) and S (with two codons). The ratio of two-codon situations here is that of: (2:2), (1:2), (0:2). As we can see, the "unit change law" holds here as well.

Finally, in area "-3" we have F, whereas in area "+3" we have E and G, with the following ratios of two-codon situations: (1:0), (0:1), (0:1).

Let us conclude as to the analogy with the quantum model: along the dividing line *zero* amino acids were chosen ("empty"); the amino acids L, I, M, V were chosen only *once* ("middle" or "semi"), when "line 9" divided Floor-Table in left and right side, in other words, when it divided the Watson-Crick Table into the upper and lower part; all the remaining amino acids were chosen *twice* ("strong" or "full"), the first time along the zeroth dividing through, "line 9", and the second time in the appropriate "quantum area".

(*Note*. As to the analogy with quantum physics regarding categorization of amino acids by means of nucleon number, see Shcherbak, 1994, p 476. As to the calculation of number 9 on "line 9", as well as all the other "Binary values", resulting from the right operations, see Rako~evi}, 1990, p 11 and 1994, p 72).

If, besides the previously mentioned results, we bear in mind the result from our previous research works (Figure 5 in Rakoèeviæ, 1990, p 11; then 1990, p 19, and 1994, p 64) which proves that the relations between the amino acids, measured with "binary values", are strictly determined by the Fibonacci series, i.e. the Golden Section, then our conclusion is clear: Codon - anticodon interactions (*Input*) have been "calculated" in such a way that the *Output*, i.e. the relations between codons and amino acids, is such as follows from strict determination by way of binary symmetry, harmony and proportion, including the best possible proportion - the Golden Section. With this, proof of the existence of a relation between *Input* (Codon -

anticodon) - *Output* (Codon - amino acid) from the first working hypothesis (Chapter 2) is total and definitive.

### 6.3. Boolean Spaces Through Hydropathy and Cloisteredness

If the conclusion in chapter 6.1. is true that the key factors which determine the replacement of amino acids during the evolution of protein macromolecules are: molecule size, its structure, its position within the Boolean space and polarity (hydrophilicity - hydrophobicity), then stereochemical determination cannot be independent of determination by way of polarity (hydrophilicity - hydrophobicity). Evidence for the existence of connectedness would be possible only if we possessed fixed parameters for measuring polarity, that is, hydrophilicity - hydrophobicity. Fortunately, such parameters exist: the hydropathy index (Doolittle & Kyte, 1982) and cloister energy - "a formal free energy" (Swanson, 1984, p 196; Rackovsky & Scheraga, 1977); and fortunately they are strictly (precisely and accurately) determined. (The situation is such that the discovered law confirms parameter accuracy, parameter accuracy being the condition for a law to be discovered).

The Boolean square - Boolean cube system ( $B^2$ -  $B^3$ ) in Table 1 can be reduced to the Boolean square - Boolean square system ( $B^2$ -  $B^2$ ), as shown in Table 1(I). Then the hydrophobic amino acids are on the even-numbered vertices of vertical  $B^2$ , excluding G, whereas the hydrophilic are on the odd-numbered vertices. Thus odd vertices 1 and 3 are "strong" hydrophilic, the even vertex 2 is "weak" and the even vertex 0 is "mixed".

Knowing this, we have uncovered one more *crossing-over* in Figure 6, starting from Y, going in a counter-clockwise direction, and ending with T, all amino acids are hydrophilic, excluding two amino acids which have "strayed": G & A, because they are not hydrophilic (A), i.e., are not pure hydrophilic (G); on the other hand, starting from F and going in a clockwise direction, and ending with A, all amino acids are hydrophobic, excluding two hydrophilic amino acids which have "strayed": S & T.

The values of the hydropathy index as we see follow the law of hydrophilic - hydrophobic segregation, with the exception of one case, the case of *Gly*. Thus we can say that, as far as *Gly* is concerned, a deviation occurs, or in other words, that amino acid is a *wobble*. This characteristic becomes obvious once again if, instead of the hydropathy index, we take cloister energy as a parameter. According to the logical relatedness of the

two parameters, it is expected that all the amino acids, which were "positive" by the hydropathy index, acquire "negative" cloister energy values and vice versa (the values for cloister energy *see* in Fig. 2). However, this is not the case for *Gly*. Besides *Gly*, this is also true only for *Trp*. In this way we were able to discover another amino acid with a deviation, which is, in fact, a wobble within the unified Hydropathy - Cloisteredness system (HYCLO system).

Note the logic of the amino acid categorization. From the aspect of stereochemistry, deviant, i.e. *wobble* are *Pro* & *Gly*. But *Gly* is wobble also within the hydropathy system (in the relation to the position in Boolean space); within HYCLO system, however, except *Gly*, still the *Trp* is deviant, i.e. *wobble*.

From this, the following conclusion can be drawn: parametrically (HYCLO), all amino acids, excluding these three deviant ("wobble"), are strictly determined as being one-meaning (*strong*); *Gly* is "weak more", *Trp* is "weak less", and *Pro* is *mixed* - in the categorization from the aspect of stereochemistry, it is *weak*, while, in the HYCLO system, it is *strong* one-meaning. All this strictly follows from the analysis of the parameter relations and the analysis of the relations of the *Input*: Codon ring - *Output*: Mutation ring (the second case in the first separate hypothesis given in chapter 2), but that directly and obviously follows from the basic chemical composition of the molecule. Thus, *Gly* is the only amino acid which, instead of a hydrocarbon has a hydrogen side chain. *Pro* is the only *imino* acid, and *Trp* is the only one with two *fused* rings. But what is in a way surprising is the fact that, by the removal of the deviant ("wobble") amino acids, the system of the 17 remaining amino acids is strictly balanced by means of nucleon number, and that is achieved through the realization of the law of binarity and proportionality (*see* Surveys 2-3 and Solutions 1-7 in Chapter 4).

The reduction of the Boolean system from  $B^2$ - $B^3$  in Table 1 to the Boolean system  $B^2$ - $B^2$  in Table 1 (I), is realized by the boolean logical "mappings", that is, by the function's function, as is shown in Solution (8):

$$\begin{aligned}
 (000, 001) &\leftrightarrow 00 \\
 (010, 011) &\leftrightarrow 01 \\
 (100, 101) &\leftrightarrow 10
 \end{aligned} \tag{8}$$

$$(110, \ 111) \leftrightarrow 11$$

**Table 1(I).** Cyclic and Periodic Amino Acid System: First Step.

	0	1	2	3
0	G -0.4	A +1.8	W -0.9	Y -1.3
1	F +2.8		H -3.2	M +1.9
2	—	P -1.6	T -0.7	C +2.5
3	—		S -0.8	L +3.8
				R -4.5

The Table generates from Table 1 by overlapping of each two adequate vertices of the Boolean cube in accordance with the solution (8).

The next possible step is the mapping of the system  $B^2-B^2$  into  $B^2-B^1$ , as shown in Table 1(II). The amino acids, then, remain on their vertices in space  $B^2$  in a fixed order determined by means of the hydropathy index (by cloister energy as well) in space  $B^1$ , and that separately for vertex “0” and vertex “1”. For these two vertices one order is within three non-zero vertices, but there is a separate order for the zero vertex within space  $B^2$ . All these rules are possible only on condition that mapping  $B^2-B^2$  into  $B^2-B^1$  is realized by the Boolean operation given in solution (9):

$$00 * 11 \leftrightarrow 0 \quad (9)$$

$$01 * 10 \leftrightarrow 1$$

**Table 1 (II).** Cyclic and Periodic Amino Acid System: Second Step.

	0		2			1			3		
0	A +1.8	G -0.4	I +4.5	V +4.2	L +3.8	S -0.8	W -0.9	Y -1.3	D -3.5	N -3.5	R -4.5
1	F +2.8		C +2.8	M +1.9	*	T -0.7	P -1.6	H -3.2	E -3.5	Q -3.5	K -3.9

The Table generates from Table 1 (I) by overlapping of the outer (0-3) and inner (1-2) vertices of the Boolean square. The vertex 0 generates from the first overlapping and the vertex 1 (of 1-dimensional Boolean space) generates from the second overlapping.

The amino acids, then, on the second odd vertex (vertex “3”) from  $B^2$ - $B^2$  pass on to the same vertex in  $B^2$ - $B^1$ , when rotating in a clockwise direction, from top to bottom: (D-N,R), (E-Q,K); on the first odd vertex (vertex “1”) the rotation is in the same direction, but it now moves from the bottom to the top: (S,W-Y), (T-P,H). On the second even vertex the rotation is counterclockwise in one case and that from top to bottom: (I-V,L); in the other case, the order is determined by going from the bottom up - from one molecule to another, either rotating in a clockwise or counterclockwise direction: (C-M); on the first even vertex (vertex “0”), the rotation is also in a counterclockwise direction, going from one molecule to another from top to bottom or the other way around - either way: (A-G); in the second case, it makes no sense to talk about either the direction of movement (clockwise or counterclockwise), or about the direction of approach (from top to bottom or vice-versa): (F). In accordance with this, mapping for (F), as far as determination by means of direction is concerned, is *empty*; mapping for G-A and M-C is *semi* because they have been determined by means of only one direction (only horizontal or vertical, respectively); mapping for the remaining discussed cases is *full* because it is realized by determination from both directions (horizontal & vertical).

With the overlapping of the zero-vertex with the non-zero vertices in the system  $B^2$ - $B^1$  we get a new system,  $B^2$ - $B^1$  in Table 1(III), in which there are no amino acids on the zero vertex, all of them being on the non-zero vertices in a sequence that would be expected by a chemist. Thus we have: three sets of amino acids in Table 1(III) which are the result of strict relations between the Codon-ring, Mutation ring I, and Mutation ring II. On the other hand, those same three sets have been perceived thirteen years earlier by two chemists (Doolittle & Kyte, 1982, p 113) purely from the aspect of chemical and physical properties of the amino acids, first of all from the aspect of the

hydropathy index values. They remarked the following: "As an initial test, the 20 side-chains were assigned to three groups according to their rank on the hydropathy scale. Thus, arginine, lysine, asparagine, aspartic acid, glutamine, glutamic acid and *histidine* were assigned to cluster I; proline, tyrosine, serine, tryptophan, threonine and glycine to cluster II; and alanine, methionine, cysteine/cystine, phenylalanine, leucine, valine and isoleucine to cluster III." (italics M.R.) As far as we can see there is a slight disagreement where *histidine* is concerned.

**Table 1 (III).** Cyclic and Periodic Amino Acid System: Third Step.

		0				2				1				3		
		I	V	L	A	G	S	W	Y	D	N	R				
0		+4.5	+4.2	+3.8	+1.8	-0.4	-0.8	-0.9	-1.3	-3.5	-3.5	-4.5				
1		F	C	M	*	*	T	P	H	E	Q	K				
		+2.8	+2.8	+1.9			-0.7	-1.6	-3.2	-3.5	-3.5	-3.9				

*The table generates from Table 1 (II). For details see the text.*

Concluding our parametric argumentation, it is important to emphasize the following. The system of amino acids in Table 1, with amino acid D in starting, i.e., zeroth position, represents the reality of the Codon-ring and Mutation ring relations. The system in Table 1(I), with G in the zeroth position, represents the ordered system from the aspect of the hydropathy index scale. If this system is designated as sistem 1(I)G, then there must exist the system 1(I)H, with the amino acid H in the zeroth position, i.e., with ordering according to the cloister energy value scale. It is obvious that in the last case zero is really zero (*strong* or "*full*"); in the second case (with G), the value of the hydropathy index is near zero (*middle* or *semi*); and finally, in the first case (with D), zero is not zero at all (*weak*, or more precisely *empty*).

## 7. FINAL COMMENTS

How has the genetic code become "from the beginning" that what it was: why with those bases and that exact number of bases; why with those amino acids and why with that exact number of amino acids; with exactly 3 "stop" codons in the alphabet which functions on the level of words, with exactly one termination situation in the alphabet which functions on the level of letters (as showed graphically in Figure 1 in Shcherbak, 1994, p 475)? All this follows from a strict determination by nucleon number presented in

chapter 4 and a strict determination based on physical and chemical parameters presented in chapter 6. The genetic code, therefore, must be universal for life which could exist anywhere in the universe (with this the answer to question 1 from the *Introduction* has been given). But it is not universal in the sense that there is one-meaning correspondence between the words of one alphabet and the letters of the other alphabet in all cases. On the contrary, for the largest number of cases the correspondence is really one of one-meaning (*strong*), but in a number of cases there must be a deviation from one-meaning, and that by two possible levels: *middle* and *weak*. In a previous work (Rakoèeviæ, 1988, pp 182-183), we have given the following prediction: "the optimal path in the process of coding (insofar as there aren't any anomalies) is realized with at least one binary step and at the most with two! Therefore, all exceptions from the universal code can appear only within these limits... . The phenomenon of multiple-meanings in cases of suppression does not overstep the limits of the two binary steps." And now we can more precisely say: the deviation ("*wobble*" or "*wobbling*") from one-meaning in the coding process is a law, a universal principle, and in the case of the genetic code this principle manifests itself in such a way that there are "one-meaning" limits which is *strong* (in most cases), then a "one-meaning" which is slightly weaker (within the limits of one bit in the Gray code model, or in the genetic code binary tree), *middle*, and an even weaker "one-meaning" (within the limits of two bits), "*weak*" (with this the answer to question 2 from the *Introduction* has been given).

All examples of deviations from the standard genetic code, presented prior to or following 1988, confirm our prediction: they are deviations only within the limits of two bits. We should, however, list some concrete examples: Kuchino et al. (1985) and Horowitz & Gorovsky (1985) report that in the *Tetrahymena thermophily* codon UAA there is no "stop" meaning, but it codes for glutamine. The UAA position in the standard code binary tree is determined by means of the Boolean vector (100010). The position of the first codon which codes for glutamine of codon CAA is (100110). The difference is 1 bit. The position of the second codon, CAG, is (100111). As we can see, the difference is two bits (the difference in the number of ones). And now the conclusion: from the aspect of the first codon, the situation in the change of one-meaning can be described as *middle*, and from the aspect of the second codon, as *weak*; in relation to both codons, the situation is, however, *mixed*. Osawa et al. (1992, p 230) report, however, that "in certain ciliated protozoans, UAR codes for Gln." This means that, besides UAA, UAG codes for Gln. In either case, the limit of

two bits is not violated (the reader can easily convince himself of that by "reading" the six-bit-records of appropriate codons in the binary tree in Figure 1). The second example, as reported by Yamao et al. (1985, p 2306) and Osawa et al. (1992, p 230), refers to the organism *Mycoplasma capricolum* in which "UGA codes for Trp." But, instead of citing examples of particular cases, we can generally conclude the following: in all cases of deviation from the standard code, which have been discussed by a great number of researchers (Sanger et al., 1981; Jukes, 1983; Attardi, 1985; Alvager et al., 1989; Osawa et al., 1992), are such that they do not violate the limit of two bits. With this, our prediction from 1988 forward still holds (now as *Prediction 7*): and in the future there will not appear cases of deviation from the standard code by more than two bits.

The discussed cases of deviation from one-meaning given by the standard code concern homonymy (one and the same codon has different meanings in different systems). However, deviations from one-meaning are determined by the *strong-middle-weak* relation even when the chemical composition of the genetic code constituents (amino-imino acids and amino-imino bases) are considered. The 18 amino acids are strictly one-meaning, in the sense that all are made up from the same 4 kinds of atoms - H, C, N, O. They, therefore, have a *strong* one-meaning. For the remaining 2 amino acids (M & C) a deviation ("wobble" or "wobbling") already appears, and so does the fifth kind of atom (S); this is how "weakening" of one-meaning occurs. In regard to that, methionine (in both forms: sulpho-methionine and seleno-methionine) stays *middle*, while cysteine "weakens" even further to become *weak*. It becomes so in two ways. First, it "multiplies" itself for a whole "step", that is, for one whole "neighborhood" - in proteins it appears in the form of cystine. Secondly, it becomes "weak" by "multiplying" its standard nucleus (atom S) by one whole neighborhood (Se) (by one electron level!), so that it gives rise to a "nonstandard nucleus" (atom Se) and with it, to amino acid selenocysteine (cf. Voet & Voet, 1990, p 912; Osawa et al., 1992, p 254; cf. "The anomalous" behavior of cysteine in Mutation ring II in Figure 6.).

(Osawa et al., 1992, p 254: "One of the most remarkable properties of coding is the occasional incorporation of selenocysteine in polypeptide synthesis in both prokaryotes and vertebrates. Secys has been sometimes termed the '21st amino acid'. It occurs as the active center of a few enzymes... Enzymes containing Secys have not been detected in green plants"; and further at the same page: "Notably, Secys cannot replace cysteine in cysteine tRNA. In this respect, Secys is unlike selenomethionine,

which can become aminoacylated to methionine tRNA and is then incorporated into thiolase of *Clostridium kluyveri*”);

Our prediction for future research (*Prediction 8*) is that an analogous strict determination of one-meaning - multiple-meanings has to exist for amino-imino bases as well, whose determination we can now only hint at. Namely, in the following sense: from the aspect of "standardization", C, A, G have a "strong" one-meaning, while T and U already show a "weakening" - the next step in the "weakening" is represented by different modifications of nonstandard pyrimidine and purine bases (cf. Voet & Voet, pp 902-903).

Strict determination of one-meaning - multiple-meanings of amino acids by way of the *strong-weak-middle (mixed)* relation is important for the pairing and non-pairing of amino acids from the aspect of their stereochemical categorization (cf chapter 4.1.). If it is noticeable that according to the character of the influence of the side chain R on the conformational freedom of the basic monopeptide segment (-CONH - C<sup>α</sup>HR - CONH -), the 20 canonical amino acids can be categorized into 4 stereochemical types: *Gly*, *Ala*, *Val*, and *Pro* (Popov, 1989, p 79), determination then takes place in the following manner. According to E.M.Popov, glycine belongs to type *Gly*, proline belongs to only type *Pro*, Isoleucine, together with valine, belongs to type *Val*, while the remaining 15 (of the total 16) amino acids belong to type *Ala*. Bearing this in mind, we are of the opinion that every chemist can easily see that the 16 amino acids of the *Ala* type are strictly divided into 8 pairs: A-L, S-T, C-M, N-Q, D-E, K-R, H-W, and F-Y. The following conclusion can be drawn from this: from the aspect of strict stereochemical one-meaning pairing (*strong*), there is only one pair of amino acids, and that is V-I; one pair is *weak (empty)*, but it is not really a pair: G-P; finally, then all the remaining 8 pairs within the stereochemical type *Ala*, are *mixed*, in the sense that it has 8 different variations of one and the same stereochemical type. If we add to this the fact that stereochemical pairing - non-pairing is determined, also very strictly, by the number of nucleons, and even by perfect numbers (cf. Appendix 3), then no special discussion is necessary to additionally prove why "The Stereochemical Theory", and not "The Frozen Accident Theory" holds true for the genetic code (with this the answers to question 4-5 from the *Introduction* have been given). Moreover, it becomes obviously clear why the genetic code had to be "from the beginning" the same as it is today (with this the answers to question 7 from the *Introduction* have been given).

When it is once perceived and understood that the Boolean spaces are the main determinants and invariants of the genetic code, as we have shown in

the previous six chapters, then all the other experimental results as to the genetic code have to be perceived in a different light and differently interpreted. We will show this with several examples. It follows from the accurately given “Mutation Data Matrix”, MDM, (Dayhoff et al., 1979; Dayhoff & Orcutt, 1985) that the evolution of proteins was “a random” process (no ordering of amino acid groups in the matrix is preceivable). But that is, in fact, due to the fact that the order of amino acids is not the one that would unavoidably follow from the positions of amino acids in the Boolean space. With such an order, the situation is the opposite (as expected!): a strict ordering by amino acid group exists; in other words, the evolution of proteins must be “a non-random” process (cf. the original order of amino acids in the MDM with our order in Rako~evi}, 1988, p 196 and 197; the table on p 196 is the same one from Figure 4 on p 7 in Dayhoff & Orcutt, 1985). The same holds for “The genetic code matrix” (Dayhoff et al., 1979; Dayhoff & Orcutt, 1985): in the original order of the amino acids, there are as many as 12 mismatchings (the mismatching of number 3 with number 2 along the diagonal, whereas in our order there are only 4 mismatchings (cf. Table on p 193 and Table on p 1995 in Rakoèeviæ, 1988). Of course, in our original matrix, which strictly follows the position of amino acids in the Boolean space , there is not even one mismatching (Table 48 in Rakoèeviæ, 1988, p 192). (Note. Table 49 in our study on p 193 is the same one from Figure 3 in Dayhoff & Orcutt, 1985, p 6. The necessity of matching of numbers 3 and 2 is clear from the explanation given by Dayhoff & Orcutt, 1985, p 6: “Identical amino acids obtain a score of 3; those for which two nucleotides could be identical, 2; one nucleotide, 1; and 0 if no nucleotides are ever shared in the codons for the amino acids”).

The presented “genetic code matrix” can also exist in its inverse form which was used by Fitch & Margoliash (1967, p 280) and Leunissen & De Jong (1986, p 192). In such a case, “the table is symmetrical about the diagonal of zeros.” In any case, there still remains 12 mismatchings in it. However, altered by our (Boolean) order of amino acids (Rakoèeviæ, 1988, p 180) there are no more than 5 mismatchings. Undoubtedly, our original (Boolean) matrix even in this form shows not even one mismatching (Rakoèeviæ, 1988, p 188). There is no need to specifically emphasize the fact that the results of the previously mentioned authors, which follow from the comparisons with “the genetic code matrix”, would have been different had that matrix had the Boolean order of amino acids.

Instead of every researcher having to give his order of amino acids in the mutation matrix, or in the matrix of the genetic code, it is essential that the

order be standardized, and that, only that order which follows from the positions of the amino acids within the Boolean spaces, with respect to “the unit change law”, that is, the allowed change should vary only by one bit going from one amino acid to the next in the genetic code binary tree (Figure 1), perceived as being three-dimensional (three-four-dimensional to be more exact).

Schulz and Schirmer (1979, p 172) changed the order of amino acids in “the Mutation probability matrix for the evolutionary distance of 2 PAM’s” (Dayhoff, 1972, p 92), with the aim of explaining the main result on which the “Non-Darwinian Evolution” theory was based (King & Jukes, 1969). Had they brought the change to its end (reduced it to the Boolean order), their observations would have been more complete, but as it is, because of the good correspondence of their order to the Boolean one, their observations are exceptional. Contrary to the conclusion of King & Jukes, they hold that the result as to “correlation between observed and expected amino acid frequency” (p 173) favors Darwin’s Theory of Selection, and not the other way around (p 174: “Therefore it cannot be deduced from the correlation between such summary values as amino acid frequencies that the evolution is neutral, i.e., non-Darwinian”) (cf. Rakoëviæ, 1988, p 72: “From the experimental results we will here cite those of King and Jukes... In spite of the fact that these authors are using this result to refute Darwinism, facts are facts, and the question of scientific conclusion depends at times on the subject himself - the scientist”).

The complete analysis we have given in this paper confirms that the frequencies expected on the basis of the genetic code cannot at all be random, but are (with the representative sample, not only from the aspect of a sufficiently long evolution period, but also from the aspect of a sufficient number of different kinds of organisms taken for analysis) evidently non-random. With the correct conclusion, therefore, the result of King and Jukes is excellent because it shows that, in spite of the small selective sample (p 796: “Graph showing the similarity between the observed frequencies of amino acids in 53 completely sequenced mammalian proteins”), the *output* is such as expected on the basis of the *input* - the physical and chemical properties of the genetic code constituents and their positions within Boolean space. That this is indeed so is also proven by “the Growth factor for 2 PAM” which was presented by Schulz and Schirmer (Figure 9-1b, p 173), and which corresponds to the graph of King and Jukes. Schulz and Schirmer perceive the agreement but cannot make sense of it (p 174: “Note that no attempt was made to explain the observed correlation of Figure 9-

1b”). And the sense is more than evident. The graph on (their) Figure 9-1b represents, in fact, the symmetrical order of the amino acids from Space-3 and Space-4 and that in the following way: below the line of the graph are the amino acids from Space-3: M, I, P T, S and R; above the line of the graph are the amino acids from Space-4: W, C, H, F, Y, D, K and L; with this another full Crossing over is realized: two amino acids from Space-3 have strayed into Space-4 (A,V), and two amino acids from Space-4 have strayed into Space-3 (N,Q); but a semi Crossing over is also realized: exactly on the line of the graph is one amino acid from Space-3 (G) and one from Space-4 (E). Thus, to conclude: from the aspect of Crossing over, (A,V) and (N,Q) are “full” (complete Crossing over); (G) and (E) are “semi” (semi Crossing over), whereas all the remaining amino acids are “empty” (there is no crossing over).

As to the strict agreement of experimental results with theory, for future research, the following important things must be kept in mind. Selective samples are permissible only in cases when the power and range of the *input - output* relation is examined (in the sense designated in the first working hypothesis in chapter 2), otherwise they are not permissible; they are especially not permissible regarding things which pertain to the question of the existence of a Darwinian or non-Darwinian evolution. In that sense, all criticism directed at King and Jukes by L. Gatlin in the all-embracing polemic is justified:

*King & Jukes (1969), p 789:* “As far as is known, synonymous mutations are truly neutral with respect to natural selection.”

*Gatlin (1972), p 198:* “This is not the case with respect to... selection”; p 180: “King and Jukes (1969) have selected an amino acid composition from a sample of vertebrate proteins which they believe is representative.” As to further debateable aspects of this polemic, see appropriate numbers in *J. Mol. Evol.* (7, 185-195, 1976; 8, 295-297, 1976 and 8, 299-300, 1976).

One of the questions which was a rather polemical subject is “the conspicuous disparity of the observed and expected frequencies of occurrence for arginine” (King & Jukes, 1969, p 797). Not intending to spark off any discussions in regard to this, we will remind ourselves of the fact that arginine is the only amino acid which is simultaneously located in both spaces, Space-3 and Space-4, of the Boolean cube, bringing the number of amino acids to a total of “21”. There is disparity there, and there is disparity here! And to top all surprises: this amino acid deviates from even this deviation - within Mutation ring II in Figure 6 this amino acid is located in only Space-4, and not in Space-3 as would be expected. All in all,

we can see that the behaviour of this amino acid is characterized by a specific "wobbling" (existing to a significant degree). Therefore, in the system of 20 amino acids, it can certainly carry the epithet -"the wobbling of wobbling's wobbling" (I Wobbling: the genetic code, due to the fact that there exist deviations from the standard code within the limits of one and/or two bits; II wobbling: the 20 canonical amino acids, bearing in mind the fact that they can be "forced" to become "21" amino acids; III wobbling: arginine, by means of which this "forcing" is realized.). The reader here probably recalls that the next amino acid which can also carry this epithet is cysteine (see previous discussion and compare with position C and R in Mutation ring II in Figure 6; also notice that C is the only amino acid in the right half of the Watson-Crick Table which has a positive value for the hydropathy index.)

What is in a way paradoxical, however, is the fact that if any of the 20 amino acids can carry the epithet - "the invariant of the invariant's invariant"- then that amino acid is arginine again. That follows from its position in the system in Figure 5. Without arginine that system would be neither symmetrical nor harmonious; and no other amino acid could replace arginine in that role, not even ornithine, despite Jukes' findings (Jukes, 1973, p 24: "I have suggested that arginine displaced ornithine during the evolution of protein synthesis"). Notice here that arginine has a very complex structure and that it is a semi-essential amino acid for most organisms; on the contrary, ornithine has a very simple structure, thus making it a non-essential amino acid (cf. Van Nostrand's Scient. Enc., 1983, p 119) (*Hint.* I invariant: the genetic code, the fact that it is universal, with the permissible 2 steps of freedom; II invariant: the 20 canonical amino acids, the fact being that from the genesis of the genetic code until the present day, there have been 20 amino acids, as there will be in the future, despite the "wobbling" behaviour of arginine and cysteine; III invariant - arginine and/or cysteine, the fact being that without arginine the system in Figure 5 could not exist, and/or the fact that the role played by sulphocysteine and seleno-cysteine cannot be played by any other amino acid.).

It should also be noticed that the three "wobblings", i.e., the three invariants, can be "read" in the opposite direction where I becomes III and vice-versa. In that case, in the role of entity I can be found any one of the 20 amino acids with a precisely defined degree of "wobbling", that is, invariantness.

Finally, it should also be noticed that everything that holds for the system of the 20 canonical amino acids analogously holds for the system of the four

canonical bases (U, C, A, G) as well. This system can also be “forced”, in other words, increased by exactly one base and which can be done in two ways. Accordingly, cysteine’s analogue is uracil, whereas arginine’s analogues are A and G simultaneously. Analogous to the “widening” of sulpho-cysteine into seleno-cysteine, uracil “widens” in the interaction of DNA-RNA (in the transcription process) in such a way that it becomes even thymine. On the other hand, the fact that what is happening to arginine is unreal (“mapping” two unreal entities from Space-3 and Space-4 in a real molecule of arginine), what happens to adenine and guanine is real: these two real entities are “mapping” themselves into a new real entity: hypoxanthine (primarily in the codon-anticodon interaction, in processes of translation. Besides all this, the system of “20 + 1” amino acids is “clean” (less “wobbling”), whereas the system of “4 + 1” bases in one way, and “4 + 1” bases in another way, in other words, the system of “4 + 1 + 1” bases, is “dirty” (more “wobbling”) due to the existence of a great number of modifications. [*Hint.* A maximally widened system of “21” amino acids and a maximally widened system of 6 (4 + 2) bases, exist in a strictly harmonious relationship of the first (6) and the second (28) perfect number; in the sense that 21 is 3/4 of 28, and 6 is 4/4 of 6. The quantities 3 and 4 exist in the relation of the best possible harmony, as we have shown in many instances. Notice, in regard to this, that the quantities 3 and 4 are here connected by the mathematical operation of division, whereas in the system in Figure 5 they are connected by the operation of multiplication, which also represents a special kind of inversion. With this the sense of classification 8 : 4 : 8 in the system in Figure 5 becomes even more clear].

The strict agreement of theory and experimental research, as we have shown in the six chapters and the Discussion of this research paper, demands other requirements. The Codon ring, Mutation ring I and Mutation ring II (as we have presented them in this paper), must be in the future used as standard and referential systems, in the sense that they are reality-models, and changes regarding them are not permissible. Not even minimal changes can be tolerated, like those carried out by Taylor (1986, p 208), who has changed the positions for H and R in the mutation ring; much less greater changes which, (for 8 amino acids) also in the mutation ring, observing it as “the rosette”, were carried out by A. Prat and her associates (1986, p 56, Figure 5) (the very idea of a “rosette” is otherwise an excellent one and it agrees with our own view of the eight rosettes in the binary tree, Figure 1).

On the basis of what we have presented, on the basis of the discussion given in every one of the six chapters, as well as all integral discussions,

the inevitable conclusion is that all the working hypotheses given in chapter 2 have been proved. The general hypothesis, according to which the Boolean spaces are actually the main determinants and invariants of the genetic code, has been therefore proven. The Boolean spaces have been shown to be reality-models! From this it further follows that it makes no sense to talk about the neutrality of mutations, or about a non-Darwinian evolution.

In specific places in our paper answers were directly or indirectly given to all the questions mentioned in the *Introduction*, except for question number 3 and number 6, which were directly answered, through the evidence given for the four separate hypotheses. Thus, “the present status of Wobble usage” or “the general base-pairing hypothesis”, or “the two out of three” hypothesis, do not refute “the wobble hypothesis” but do, in fact, promote it to a generally-held principle for the genetic code. On the other hand, the genetic code, being redundant, did not become degenerate in the process of evolution, but was generated in origin as such. It is today as it was in the beginning and it will remain so in the future anywhere in the universe, because that follows from the positions of the bioelements in the periodic system of elements; bioelements - being the constituents of the genetic code. Accordingly, it makes no sense to talk about the evolution of the genetic code, but it does make sense to talk about the evolution of the macromolecules, that is, the evolution of life which came into being on the basis of just such a code - a universal genetic code.

## Appendix 1

### Generation of Cyclic Amino Acid System

The cyclic and periodic system of amino acids shown in Table 1 is the result of the strict relations of the amino acids in the Codon ring (Figure 1 in Swanson, 1984, p 188), Mutation ring I (Figure 3 in Swanson, 1984, p 191) and Mutation ring II (Figure 6 in this study). How is this system generated? In other words, how is (in sequence) the choice of one amino acid, then another, made? In order to answer this question, it should be at first noticed that there exists at least only one ring (the codon ring) in which amino acids are categorized into two groups: two-positional amino acids (L, S, R) and one-positional amino acids (all the remaining amino acids). It is logical, therefore, to expect that the choice of two-positional amino acids will be allowed only after all the one-positional amino acids have been chosen.

Moreover, we have noticed that the one-positional amino acids, of which there are 17, can themselves be categorized into two groups: multiple-meaning (K, P, T, C) and one-meaning (the remaining amino acids). From the set of four possible meanings: *small*, *large*, *inner* and *outer*, the one-meaning amino acids in all three rings have only one meaning; the multiple-meaning amino acids in one or two rings have one meaning, and in the remaining ring another meaning and vice versa. The next step in our observation shows that the one-meaning amino acids can themselves be further categorized into two groups: into amino acids which are located right next to the borderline, that is, the separation of all the four possible meanings (E, Q, F, H, M), and into amino acids which are not located right next to the borderline (D, N, G, A, W, Y, V, I). For the first type of amino acids we say that they are, from the aspect of possessing a "border", in themselves "*strong*" (or "*full*"), while the second type, which does not possess a "border", is "*weak*" (or "*empty*"). The multiple-meaning amino acids cannot, however, be further categorized because all four amino acids (K, P, T, C) are right next to the borderline in at least one of the three rings; they are, therefore, from the aspect of possessing "borders" in themselves, also "*strong*" (or "*full*").

We can see that within the whole "space" which is occupied by the 20 amino acids, for the one-positional amino acids (17 in all) three sub-spaces are needed, one next to the other along the horizontal line, or the vertical line (which amounts to the same thing), and at least one sub-space is needed

for the two-positional amino acids, a total of four sub-spaces. The largest of the four sub-spaces is the sub-space of the one-positional - one-meaning amino acids (D, N, G, A, W, Y, V, I) because it requires the largest number of "columns".

We can thus see that the meaning of the order of the amino acids in Table 1 is already becoming clear. There is sense, therefore, in starting the choice-making process with the simplest situations, with one of the eight one-positional - one-meaning amino acids, but with which one? To arrive at an answer, it is essential to carry out all possible tests. Should one start from the "*outer*" area or the "*inner*" one, that is, from the "*large*" or "*small*" area? After these tests it is shown that it makes more sense to start from the "*Outer*" area (which is in a certain way logical because access to any object or system is possible, in fact, only from the outside). If this is so, then the choice boils down to only two alternatives: either first choose D, or N. Reasons having to do with chemistry tell us that N as a derivative of D cannot be chosen first; rather the opposite is true. With this we have chosen the first two amino acids: the amino acid D which is really an acid (*strong*) and its derivative N which is *middle* (in relation to the other possible extreme case which could be represented by one of the bases in the set of amino acids). After this, the easiest and most certain choice of the last, the eighth consecutive amino acid, can be made. This amino acid, according to logical assumptions, must be from the "*Inner*" area. And, again, we have only two alternatives: either that of amino acid V, or I; the value of the hydropathy index determines the following: the last amino acid in the first of the four sub-spaces has to be I, and V must be right next to it.

Knowing that, on the basis of V. Shcherbak (quoted in Chapter 4), the number of nucleons is also important for the categorization of amino acids, it makes sense to suppose that in the "*column*" with I there can be only L, because they are the only two amino acids which are isomers, then we have defined the position of L in the cyclic periodic system. It must be at the opposite end of the column in relation to I, because, as a two-positional amino acid, it must be in the fourth sub-space. With this the three "*corner-stones*" have been determined. But, with them being determined, the fourth "*corner-stone*" has also been determined: if L is at the end of the fourth sub-space, then R must be at the beginning, its initial position being required from the aspect of the hydropathy index values.

In the first sub-space between the left "*corner*" (D, N) and the right "*corner*" (V, I), we have the remaining amino acids G, A, W, Y for whose possible order 24 permutations are possible. For now (up to this point in the

analysis), we cannot, namely, know which amino acids follow N from the left side and V from the right side. However, we know (and see) that in the same column after D must come E, and after N, in the neighboring column must come Q. After this, there remains in the "Outer" area only K (from the aspect of position in the Codon ring and in Mutation ring I). It cannot come right under D (where E is) because it is not a one-meaning amino acid; it must, therefore, come after E, but in front of R. It is obviously clear what follows next: if the second sub-space begins with E, the last amino acid in that sub-space must be M, while the last amino acid in the third sub-space (beginning with K) must be C; however, they cannot be in the last column (with I and L) because they would by their presence destroy the unity of the referential isomer system.

If we suppose that the system  $4 \times 8$  (four sub-spaces and eight columns) represents the unity of the Boolean square and the Boolean cube ( $B^2 - B^3$ ), as given in Table 1, then we can by the "function of function" relation reduce this system, at first to  $B^2 - B^2$ , as shown in Table 1(I), and then to  $B^1 - B^2$ , as shown in Table 1(II) in Chapter 6.3. With these reductions we have the possibility of testing the 24 possible permutations for (G, A, W, Y), with the aim of finding out which amino acid makes the most sense from the aspect of chemistry. According to our analysis the most resonable permutation is the one contained in Table 1. If this is so, then it becomes clear that in the column with G there is no place for any other amino acid; F must come under A as its derivative; then H under W as its stereochemical pair (see Survey 3 in Chapter 4.1); T and S must come in the column with the hydroxil function (the place where Y is). The only amino acid remaining to be ordered, in fact, the only imino acid is - prolin. Is proline in the column with A or with W? Its negative value for the hydropathy index gives an unambiguous answer - proline is in the column with W!

With this not only has the choice of one-positional amino acids been made, but also of the two-positional ones as well. In the Codon ring in which the distinction into one-positional and two-positional amino acids has been made, we read the following: S and R as four-codon and L as two-codon amino acids are not on the border; while, on the other hand, we have the opposite: L as a four-codon amino acid and S and R as two-codon amino acids are on the border, namely, on the borderline of the *large, inner, small* and *outer* meanings. It follows from this that these three amino acids, from the aspect of "border" possession, are in themselves simultaneously *strong* and *weak*, which means that they are *mixed* (or *semi*).

The cyclic system in Table 1 is, as we can see, strictly determined by the relations of the amino acids in the three rings, as well as by the parameters of hydrophobicity and cloister energy. But, in a certain way, that system is determined by the two spaces in  $B^3$ , by space-3 and space-4 in Figure 5(I) and Figure 5(II), and by the binary values of the amino acids. The sign for "border" for the second, third and fourth sub-space in Table 1 remains the same, in the sense that all the amino acids in these three sub-spaces are in the far end "border" positions from the aspect of their order and position in their systems in Figure 5(I) and Figure 5(II). However, in the first sub-space (one-positional - one-meaning amino acids) a change occurs: four amino acids remain without "border" in themselves (D, N, A, W), while the remaining four (G, Y, V, I) obtain the position of "border" in the systems in Figure 5(I) and Figure 5(II), if the criterium "to be on the borderline at least in one of the two systems" is used. The final result is, in accordance with this, both important and indicative: in the first sub-space we have 4 + 4, in the second 5, in the third 4 and in the fourth 3 amino acids. The ratio is, therefore, 4 : 5 : 4 : 3, or 5 : 4 : 3 if the reading starts with the second and ends with the fourth sub-space; that is, it is 3 : 4 : 5 if the reading begins with the fourth, then goes to the first (in one cyclical closing), and ends with the second sub-space. This is, as we can see, the relation of the first three Pythagorean numbers. In Chapter 4.1 we have shown, referring to the works of V. Shcherbak, that the relation of the number of nucleons in the four-codon amino acids has been determined by these numbers, or more precisely, by their squares.

All four semi-essential amino acids in space-4 in Figure 5(I) and Figure 5(II) are on the borderline, that is, they occupy the end positions. That is obvious for Y and R. The amino acid C is on the borderline when we take into account the classification of amino acids as 10 : 10 - ten non-essential and ten essential; that, therefore, also holds for H. As to the classification of amino acids into that expressed by the ratio 10 : 10 and 8 : 4 : 8 from the aspect of essentiality. It is important to notice the identical manner of the ordering of the 4 + 4 amino acids in Mutation ring II and in the systems in Figure 5(I) and in Figure 5(II): Q, N, D, E and S, A, G, P. As we see it is the same situation with one "Crossing-over".

## Appendix 2

### Arithmetical regularities within the tRNA molecules

In this Appendix we will show that the hypothesis-prediction , given as Prediction 5, makes sense on the example of "standard" tRNA with 76 nucleotides (cf. e.g. Figure 1 in McClain & Seidman, 1987, p 605, or Figure 6.1 in Lewin, 1987, p 124). After terminus ACC/CCA, and after the "pair" 0-73, another 07 pairs follow in the acceptor arm, with additions up to 73; within the 8 pairs (including the loop) of the D arm addition occurs up to 35, which is half of 70; within the 8 pairs of the anticodon arm addition is up to 70 (with the remaining position 35, which is again half of 70 and which is the central position in the anticodon arm); finally, within the 8 pairs of the TΨC arm addition is up to  $114 = 70 + 37 + 07$ , with the remaining position 57 (notice the logical explanation for this: if the single base position 35 in the Anticodon loop is increased by 011 the single base position 47 in the Extra arm loop is obtained; a further increase by 011 will obtain the single base position in the T  $\psi$  C arm). (*Hint.* The number 08 in its inverse counting, starting from vertex 15 as zeroth, within the Boolean hypercube, is number 07).

It is obvious, from the aspect of the given analysis, that in a molecule of tRNA-76 it is logical to start "counting" from the pair 03-70, with a minus sign for the pairs which are moving towards terminus, and with a plus sign for the pairs which are moving towards the anticodon. In this way, a specific (Boolean) logical square for the pairs from 00-73 to 03-70 is realized. After this observation, it becomes clear why the pair 03-70 has to be unique, about which we have talked in chapter 3.2, referring to the paper of Schimmel & Hou, 1988, p 143. Notice that the number  $03_{10}, 11_2$ , represents the end point in the Boolean space  $B^2$ , while the number  $07_{10}, 111_2$ , represents the end point in  $B^3$  (as to the unity of  $B^2$ - $B^3$  and its importance for the genetic code, see chapter 6.3). (For futher details about numbers 3, then 7 and 37 see in Appendix 4).

Additional facts about tRNAs are as follows. When tRNA-76 (*strong*) is taken into account with tRNA-95 (*weak*), then it is easily evident that the whole tRNA system is really determined only and exclusively by the number 037; not directly by this number as "the Prime quantum (PQ) 037" (Shcherbak, 1994, p 475), but indirectly by the number which represents the number of the pair within the number 037, that is, by the number 19 which

we will label as the second prime quantum (PQ) 19. The starting pair in the number 037 is 0-37, then 1-36, 2-35, etc. follow up to the last (or middle, depending of the reading), the nineteenth pair: 18-19, with members which are (the only) direct neighbours, that is, they are distanced from each other by 1. Knowing this, let us now observe the relations of tRNA-76 and tRNA-95 (in Figure 6.1 in Lewin, 1987, p 124). Between the positions 17-18 in the tRNA-76 a potential position can be occupied by the position 17:1, while in tRNA-95 the position 17:1 - 18 will become positions 18-19. Thus all this is exactly as would be expected from the logical ordering of pairs in the number 037, in other words, as would be expected by the neighborhood criterion, and as follows from "the unit change law" (Rako~evi}, 1994, p 36). With this change the following has also occurred: In the D arm in tRNA-76 there was no single base position, while one now has appeared in tRNA-95. That position is the one which precisely agrees with the result of the second PQ 19 ( $1 \times 19$ ). Everything that occurs from that point on strictly follows from "the unit change law" where one-number steps actually represent module 9, the same one by which the system of the multiples of the number 037 has been ordered (in Table 1 in Shcherbak, 1994, p 476) (notice here the "play" of the binary pair 09-19): in the Anticodon arm the single base position is 38 ( $2 \times 19$ ), in the Extra arm it is 57 ( $3 \times 19$ ), in the TΨC arm it is 76 ( $4 \times 19$ ), and, finally, in the Acceptor arm (last or first) the single base position in terminus is 95 ( $5 \times 19$ ). All together, the "played-out" situations are the following situations for the single base positions within the five arms given in solution (10); within the borders of module 9 the following situations have not been "played out" given in solution (11).

$$(1 \times 19) + (2 \times 19) + (3 \times 19) + (4 \times 19) + (5 \times 19) = (570:2) \quad (10)$$

$$(6 \times 19) + (7 \times 19) + (8 \times 19) + (9 \times 19) = (570:1) \quad (11)$$

As we can see, the ratio of the situations which have been "played out" and those which have not is that of the most simple proportion 1:2.

As far as the double base positions are concerned, the state of things is as follows. The numbers of the positions of all doublets in the D arm fills up to number 38 ( $2 \times 19$ ), in the Anticodon arm it fills up to 76 ( $4 \times 19$ ), in the Extra arm it fills up to 114 ( $6 \times 19$ ), in the TΨC arm it fills up to 152 ( $8 \times 19$ ), and, finally, in the Acceptor arm it fills up to  $73+19=92$ ; as it was

for tRNA-76, 19 is added, bringing the number to 92 (notice the specific "Crossing-over": for the tRNA-76 the result of 92 is found within the Extra arm, while for tRNA-95 this result is found within the Acceptor arm; on the other hand, for the tRNA-76 the result of 114 is found within the TΨC arm, while for tRNA-95 that same result is found within the Extra arm, which all together means that one double "Crossing-over" has occurred here.) Therefore, for the doublets the even situations have been "played out", with an increase of 92 for the Acceptor arm (solution 12); but the following have not been "played out" (solution 13).

$$(0 \times 19) + (2 \times 19) + (4 \times 19) + (6 \times 19) + (8 \times 19) + 92 = (567 - 95):1 \quad (12)$$

$$(1 \times 19) + (3 \times 19) + (5 \times 19) + (7 \times 19) + (9 \times 19) = (567 - 92):1 \quad (13)$$

As we can see, the ratio of those situations which have been "played out" and those which have not is that of the most simple (identically-numbered) proportion 1:1, on condition that the sum of the "played-out" situations is decreased by the value of one singlet of the Acceptor arm loop, and the sum of the situations which have not been "played out" by the value of one doublet of that same Acceptor arm loop.

The result of 570 in the solutions (10-11) and the result 567 in the solutions (12-13) corresponds to each other at the same time: through the number 037, through the sum of the first three perfect numbers ( $6 + 28 + 496 = 530$ ), through the first two friendly numbers, i.e., through the first friendly pair (220 & 284), through the binary pair 09-19, through the end (or middle, depending of the reading) pair 4-5 within module 9 and through all the four basic arithmetic operations (solutions 14-17).

$$570 - (37 \times 2) = 496 \quad (14)$$

$$567 - (37 \times 1) = 530 \quad (15)$$

$$570 + 530 = (220 \times 5) \quad (16)$$

$$567 - 496 = (284 : 4) \quad (17)$$

So far we have considered the nucleotide positions in the region, loops and stems within the three, i.e. four arms on the cloverleaf form of tRNA-76. But, what about the regions which are located between two arms? To answer this question, it is essential to present some important facts about the cyclicity and periodicity.

$$\begin{aligned} 0 + 1 + 2 + \dots + 08 &= 9 \times 4 \\ 0 + 1 + 2 + \dots + 09 &= 9 \times 5 \end{aligned} \tag{18}$$

$$\begin{aligned} 0 + 1 + 2 + \dots + 17 &= 9 \times 17 \\ 0 + 1 + 2 + \dots + 18 &= 9 \times 19 \end{aligned} \tag{19}$$

$$\begin{aligned} 0 + 1 + 2 + \dots + 26 &= 9 \times 39 \\ 0 + 1 + 2 + \dots + 27 &= 9 \times 42 \end{aligned} \tag{20}$$

$$\begin{aligned} 0 + 1 + 2 + \dots + 35 &= 9 \times 70 \\ 0 + 1 + 2 + \dots + 36 &= 9 \times 74 \end{aligned} \tag{21}$$

$$\begin{aligned} 0 + 1 + 2 + \dots + 44 &= 9 \times 110 \\ 0 + 1 + 2 + \dots + 45 &= 9 \times 115 \end{aligned} \tag{22}$$

In the solutions (18-22) addition according to module 9, has been shown for cases when the last added number belongs to the congruent class 8 or 9, or for cases when the result is obtained as integer or whole numbers, which are perceived in the form of the sum of the module and as a whole (integer) number. In all the other cases (for the remaining congruent classes) we do not get whole (or integer) numbers as the result.

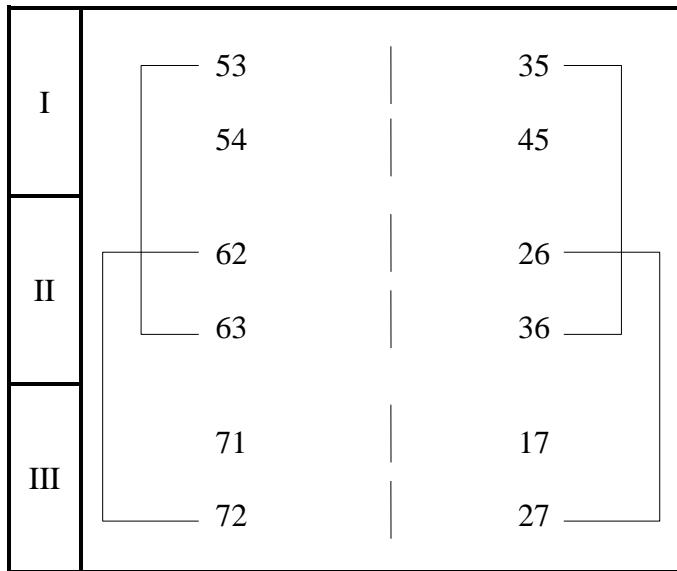
With this knowledge in mind, it becomes evidently clear that the last added numbers are really the determinants of the positions which are located at the top of the loops or between the arms. Thus, the first (8 & 9), the third (26 & 27) and the fifth (44 & 45) pair are the determinants of the positions which are located *between* the arms, while the second (17 & 18) and the fourth (35 & 36) pair are the determinants of the positions which are at the *top* of the loop. The logical explanation for the three odd situations is as follows: the first pair (8 & 9) is really *strong* or *full* located between the arms; the third pair (26 & 27) is partially, *middle* or *semi*, located between the arms, there is only one member of the pair (26) while the second member (27) enters (very slightly) the stem of the arm; the third pair (44 & 45) moves further exactly by one step, thus not one member of the pair has the status of "being between" (and is, therefore, *empty*), both members enter the stem (enter it deep: they occupy the whole stem, because it has no more than two pairs in tRNA-76).

The logical explanation for the three successive even situation is as follows: from the aspect of "being a pair exatly at the top", the second pair (17 & 18) fulfills this condition completely and is *strong* or *full*; the fourth pair (35 & 36) is only partially at the top and is *middle* or *semi*: it occupies a single base (35) and only one member of the doublet (36), while the remaining member of the doublet (34) is not included. There is , in a way, a *mixed* state here. Finally, the next pair in the succession, the sixth pair (53 & 54) is not located at the top any more: the state is *empty*.

The sixth pair, therefore, offers the logical explanation for the remaining three pairs (from the total of 8 which are in tRNA-76): for the sixth pair 53 & 54, the seventh 62 & 63. and the eighth 71 & 72 (in reality, if 08 & 09 is a zeroth pair, here are 5<sup>th</sup>, 6<sup>th</sup>, 7<sup>th</sup> pairs; cf with results 567 in solutions 12-13). The central pair of the above-mentioned three (62 & 63) is really located in the middle position. From the aspect of "being on the border", it is in the *empty* state. The first pair (53 & 54) is located on the border in such a way that it separates the two beginnings (or the two endings?): it separates the loop from the stem; it is, therefore of *semi* or *middle (or mixed)* status. Finally, the last (eighth) pair (71 & 72) is located on the border in such a way that it is only at the beginning: *strong* or *full* - at the beginning of the Acceptor arm.

It is important to notice that (53&54) pair makes a specific *Crossing-over* between two logics: 1. to be on the top, and 2. to be on the border. On the other hand, in the first pair, 53 equals 1/10 from the sum of first three perfect numbers (  $6 + 28 + 496 = 530$  ); in the second pair 62 equals 1/8 of the third perfect number (  $62 \times 8 = 496$  ); and, finally, in the third pair 71 equals 1/4 of the second friendly number (  $284 : 4 = 71$  ). Notice also a system of the further "crossing-overs" in Survays 4-5.

## Survey 4



In addition to what has been said so far, it is important to notice two more characteristic relations:

$$[1 \times (8+9)] + [1 \times (17+18)] + [1 \times (26+27)] + [1 \times (35+36)] + [1 \times (44+45)] = (530:2) \quad (23)$$

$$[1 \times (8+9)] + [2 \times (17+18)] + [3 \times (26+27)] + [4 \times (35+36)] = (530:1) \quad (24)$$

Solution (24) is analogous to solution (10). Together with solution (23) it again leads to the result (530) which represents the sum of the first three perfect numbers, that is, of the numbers which remain inside the symmetrical binary system, as shown in the solutions (16-17). With this the cycle is complete, and it can be considered proven that the system tRNA-76 and tRNA-95, as well as the total nucleon number within the genetic code, is strictly determined with NSM III in Table 2, in Appendix 3. The manner in which other relations in connection to the genetic code as well as to other natural codes are established remains for future researchers to show.

## Survey 5

	meaning (of every two)	meaning (of every one)	Accordance
I	empty (“Top”)	semi (“Top”)	→ full
	semi (“Border”)	empty (“Between”)	
II	empty (“Border”)	semi (“Between”)	→ empty
		semi (“Top”)	
III	full (“Border”)	full (“Top”)	→ semi
		semi (“Between”)	

## Appendix 3

### Perfect and friendly numbers

As to the manner in which perfect numbers are the determinants of Boolean spaces, or, in the other words, as to how perfect numbers are the determinants of the genetic code, we have shown in our previous research works - in Rakoëeviae, 1990, 1991, 1994. In this Appendix some new perfect and friendly number relations within the genetic code is shown (Fig. 7).

After the sum of the first three perfect numbers, in the logic of succession (based on the principle of continuity) the next thing that follows is the sum of the first four perfect numbers (8658); after the realization of the first two friendly numbers comes the realization of the third (1184). In Table 2 we can see that both results correspond to the multiples of the number 037: the first result (8658:1 = 8658) is completely in the position 13d, and the second result (1184:2 = 592) with one of its halves in position 16e. As 8658 is equal to  $7770 + 0888$ , or to  $78 \times 111$ , and, as  $8658 + 592 = 925 \times 10$ , we can see that all the nucleon number patterns for four-codon-amino acids and for non-four-codon-amino acids have been realized (see Figure 1 in Shcherbak, 1994, p 475).

/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      64							
<hr/>							
/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							

**Figure 7** The determination of the series of the numbers 0-63. When we look closely into the structure of the sequence 0-63 of the series of the natural numbers we come to the obvious and self-evident explanation of the reason why the genetic code must be six-bit code, no matter if it is the manifestation in the form of the Gray Code model (Swanson, 1984, p 188), or it is in the form of the Binary tree (Rakoëeviae, 1994, p 38). There must be 8 codon, i.e. amino acid classes. The structure of the sequence 0-63 is strictly determined by third perfect number (496) and the sum consisted of the first pair of the friendly numbers (220+284). Along with this, the specific Boolean square is being made and it is the restrictive factor, in a sense that it is not possible to "go on" any further - not ahead, not back: (0)  $220+284=504$ ; (1)  $156+348=504$ ; (2)  $92+412=504$ ; (3)  $28+476=504$ . The key distinctions within the genetic code are obviously self-evident: entity 64 as a series of continuance (correspondent with 64 codons); entity 20 from 496(III PN)-476=20 (correspondent with 20 amino acids) etc.

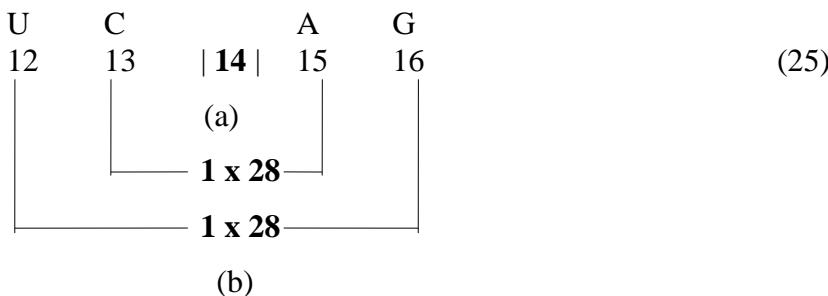
From Table 2 it is obvious that the Number System of Multiples of 037 (NSM 037) is only a sub-system of one extensive system of multiples: the Number System of Multiples of 666 & 777 (NSM III). According to our hypothesis-prediction (*Prediction 9*) all natural codes must be determined

**Table 2 The Number System of Multiples NSM III**

<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>
14	27	20979	17982	999
13	26	20202	17316	962
12	25	19425	16650	925
11	24	18648	15984	888
10	23	17871	15318	851
09	22	17094	14652	814
08	21	16317	13986	777
07	20	15540	13320	740
06	19	14763	12654	703
05	18	13986	11988	666
04	17	13209	11322	629
03	16	12432	10656	592
02	15	11655	09990	555
01	14	10878	<b>09324</b>	518
00	13	10101	08658	481
01	12	<b>09324</b>	07992	444
02	11	08547	07326	407
03	10	07770	06660	370
04	09	06993	05994	333
05	08	06216	05328	296
06	07	05439	04662	259
07	06	04662	03996	222
08	05	03885	03330	185
09	04	03108	02664	148
10	03	02331	01998	111
11	02	01554	01332	074
12	01	00777	00666	037
13	00	00000	00000	000

- The original number, countdown starting from the middle row;
- The original number, countdown starting from starting (zero) point;
- The multiples of the number 777;  $c = 21 \times e$ ;
- The multiples of the number 666;  $d = 18 \times e$ ;
- The multiples of the number 037; they are existing only in NSM III  
(For the details see the text).

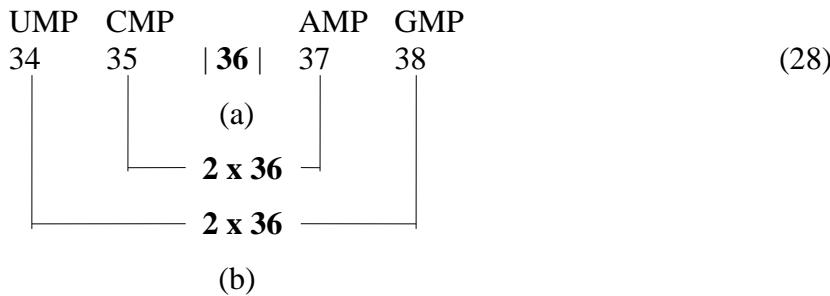
with this system, including its predecessors with the multiples 6 & 7 (NSM I) and 66 & 77 (NSM II), as well as its followers with the multiples 6666 & 7777 (NSM IV), etc. It should be noticed that the whole system is in a certain way determined by the first and second perfect number ( $6 = 4/4$  of 6;  $7 = 1/4$  of 28). With this observation, however, it is easy to perceive that the total atom number within the four pu-pyr bases and their nucleotides is also determined by the relation of the first two perfect numbers (solutions 25-29):



plus Ribose	20	
		$(2 \times 6)$
plus Phosphoric acid	08	
	---	
	28	$(1 \times 28)$
minus 2 molecules $\text{H}_2\text{O}$	06	$(1 \times 6)$

(26)

$$\begin{aligned}
 \text{UMP} \quad (12 + 28) - 6 &= 34 \\
 \text{CMP} \quad (13 + 28) - 6 &= 35 \\
 \text{AMP} \quad (15 + 28) - 6 &= 37 \\
 \text{GMP} \quad (16 + 28) - 6 &= 38 \\
 \hline
 (2 \times 6)^2 &= 144
 \end{aligned} \tag{27}$$



$$\begin{array}{ccc}
 1/2 & x & 28^1 \\
 1/1 & x & 06^2 \quad (a) \\
 \hline
 1/1 & x & 28^1 \\
 2/1 & x & 06^2 \quad (b)
 \end{array} \quad (29)$$

In connection with this, non-existing entities (a) and yes-existing entities (b) exist in strict binary symmetry interrelations (cf. Solutions 25 & 28 with 29).

After our hypothesis and prediction (*Prediction 10*) not only the total *pu-pyr* atom number, but also the total amino acid atom number must be related to the first two perfect numbers in the next sense:

$$\begin{array}{r}
 x + y = 2/1 \quad 28 \\
 x - y = 1/2 \quad 28 \\
 \hline
 x = 35; \quad y = 21
 \end{array} \tag{29-1}$$

<i>T</i>	<i>I</i>	<i>M</i>	<i>K</i>	<i>L</i>	<i>F</i>	<i>V</i>	<i>W</i>	
08 <sub>2</sub>	13	11 <sub>2</sub>	15	13 <sub>2</sub>	14	10 <sub>2</sub>	18	(29-2)
<b>21</b>		26		27		<b>28</b>		

6	<i>Apples</i>	8	6	4	4	<i>Apples</i>	4	4	8	6	<i>Apples</i>	8	<i>out</i>
C	H	R	A	S	D	P	E	G	N	Q	Y		
05	4	2	1	4	3	7	04	4	05	4	07	4	08
	33			34			3	4	3	9	01	4	08

(29-3)

The amino acids in Solution (29-2) are essential: first two (T, I) can make the diastereoisomers, the other can not. The inner amino acids in Solution

(29-3) are nonessential, the outer semi-essential. Notice that  $34 = 06 + 28$  and that  $27 + 33 = 26 + 34 = 10 \times 06$ . Notice also that atom number of 08 amino acids in Solution (29-2) is equal to the atom number of 12 amino acids in Solution (29-3):  $21 + 26 + 27 + 28 = 33 + 34 + 35 = 102$ .

## Appendix 4

### About the number 037

There are 3 and 7 non-zero vertices within the square and cube, respectively ( $3 = 1/2 \cdot 6$  and  $7 = 1/4 \cdot 28$ ; 6 is the first, 28 the second perfect number). On the other hand, if we have a cube corresponding to the binary tree through the four letters as in Fig 1, then there must be exactly 37 three-digit words which contain a determined letter (37 codons with one or more U, C, A, G, respectively). From this it follows:  $37 \times 4 = (2 \times 64) + (1 \times 20)$ . As we see, there exists 64 real and  $(2 \times 64) + (1 \times 20)$  unreal entities. If, from 64 words, 37 contain a determined letter (e.x. U), then the 27 sufficient words contain only the three other letters (e.x. C, A, G). That means: within the Watson-Crick Table there are  $37 + 27 = 64$  codons. On the other hand, in Shcherbak's Table (Shcherbak, 1994, p 476), the end multiple of 037 is the 27th:  $37 \times 27 = 999$ . From this it follows:  $999 - 64 = 057 + 878$  (78 or 87 is the middle pair in the hypercube);  $878 \times 2 = 1756$ ;  $999 + 64 = 0567 + 496$  (496 is the third perfect number; the results 057 and 567 correspond to the numbers 570 and 567 in Solution 16-17 in Appendix 2; on the other hand, number 567 is the (last) case in Survey 6).

### Survey 6

I	II	III	IV	V	VI
(1, 10, 26)	(2, 15, 20)	(3, 4, 30)	(5, 13, 19)	(6, 8, 23)	(7, 33, 34)
027	054	081	135	162	189
270	405	108	351	216	891
702	540	810	513	621	918
VII	VIII	IX	X	XI	XII
(9, 12, 16)	(11, 27, 36)	(14, 29, 31)	(17, 22, 35)	(18, 24, 32)	(21, 25, 28)
243	297	378	459	486	567
324	729	783	594	648	675
432	972	837	945	864	756

But, besides the relations  $27 + 37$  and  $27 \times 37$  there exists more complex relations  $27 \times 37$  as we see in Survey 6. Except for the cyclic permutation system (with the  $6 \times 3 = 18$  permutation triplets) in Table 1 in Shcherbak, 1994, p 476, there is a parallel permutation system in Survey 6 (with  $12 \times 3 = 36$  permutation triplets). Notice that the sum in all triplets in the first row

is  $1 \times 37$  except in the last (6th) where it is  $2 \times 37$ . In the second row there is an inverse case of a Crossing-over: in all triplets the sum is  $2 \times 37$  except in the first, where it is  $1 \times 37$ . The same is valid for the sums of any three cyclic permutations  $1 \times 999 \times 6 = 5994$  ( $9 \times 666$ );  $2 \times 999 \times 6 = 11988$  ( $18 \times 666$ ). The further relations between these two numbers are as follows. **Firstly:**  $3 + 7 = 10$ , which corresponds to the 10 pairs of amino acids: 8 pairs of the alanine stereochemical type, plus 1 pair of the valine type (V-I), and plus 1 non-pairing pair (G within the glycine stereochemical type, and P within the proline stereochemical type). **Secondly:**  $3 \times 7 = 21$ , which corresponds to the three possibilities within the genetic code: 20 amino acids plus selenocysteine; 19 amino acids plus two times R (in Fig. 5 see explanation); 20 amino acids plus 1 "stop" situation (as in Fig. 1 in Shcherbak, 1994, p 475). **Thirdly:**  $3 \times 37 = 111$  and  $7 \times 37 = 259$ , which corresponds to the patterns for the total nucleon number within the 8 four-codon and/or 15 non-four-codon amino acids (Shcherbak, 1994, p 475). Besides this, it is also important to notice the following. If the numbers  $11_2$  and  $111_2$  are read in the decimal numbering system as  $011_{10}$  and  $111_{10}$ , then the first number, itself excluded, has no other factors; the number  $111_{10}$ , however, itself excluded, has the factors 03 and 37. And, finally, where an analogy with quantum physics is concerned, the state of  $111_2$  is an analogue of the Hund semi-full state, while  $011_q$  is the previous state; for the case  $q=10$ , this previous state (as we have seen) is the quantum through which the *strong-middle-weak* relation for the three cases of "single base position" is realized; in the case of the Anticodon arm the state is strong and determined by the numbers 70 ( $35 \times 2 = 70$ ), in the case of the Extra arm we have middle agreement (a deviation of  $011 \times 1$ ), and in the case of the TψC arm we have weak agreement (a deviation of  $011 \times 2$ ). [Hint. The patterns of the "total nucleon number" for four-codon and non-four-codon amino acids in Figure 1 in Shcherbak, 1994, p 475, are exclusively multiples of the number with "the same symbols" 111 ( $03 \times 37 = 111$ ), or they are "cyclic permutations" of the number 259 ( $07 \times 037 = 259$ )].

In position 11 on the Boolean square is the most complex base, guanine, as in Fig. 1 in Rakočeviæ, 1994, p 8; in position 111 on the Boolean cube are the most complex 8 guanine type codons, as in Fig. 1 in this study, coding for aginine - the most complex ("strong") amino acid, then for serine - the less complex ("middle"), and, finally, for glycine - the least complex ("weak") amino acid.

The science-conscious reader, educated in the science of the twentieth century, cannot but conclude at this point that all that is being hopelessly

discussed here is only numerology and nothing else. The reader is separated by twenty-five centuries from Pythagora and his axiom according to which “the harmony of the Uuniverse...depends on the number”, thus it follows from this that the basis of every genuine science, which tends to discover universal laws, has to be “the study of even and odd numbers, simple and complex, figurative and perfect numbers, of arithmetic, geometric and harmonic proportions and means.” (Mathematical encyclopedic dictionary, 1988, p 737). The misunderstanding with Pythagora during the whole twenty-five centuries is first of all in the fact that it was considered that Pythagora took numbers in their “usual” sense of the word, as intuitively “seen” and perceived. Not much attention has been paid to the “figurativeness” of the spaciousness of numbers, from which Pythagora sets out. In this way, the perceived numbers are not only “numbers” but are also relations in space, and represent the relation of the parts within a whole. Understood in this manner, besides being quantities “by meaning of which counting is separated by ones” (as follows from the fifth Peano axiom), the numbeers become quantities by which interconnected ones are counted (such is the case with the numbers generated in the Boolean spaces).

As to the previously cited *Hint* regarding the relations of the universal genetic code, its constituents are such that they are exclusively determined by the relation 03 x 037; 037 x 07, it is not adequate proof for the Pythagorean axiom, or for our main hypothesis according to which Boolean spaces are the main determinants and invariants of the genetic code, then we believe that facts concerning tRNAs we have additionally given here adequate proof for even the most sceptical of science-conscious readers.

## Appendix 5

### Gold and Golden Section as Determinants of Bioelements Choice

After our hypothesis and prediction (*Prediction 11*) the further research will show that the choice of the bioelements is also determined by the *strong-middle-weak* relation. If so, then the first level ("strong") makes a system of 18 bioelements (S-18), that is 10 metals and 8 nonmetals, with a strong distinction between metals and non metals. There are 10 metals with strong order and a strong proportion: (Na, K, Mg, Ca), (Mn, Fe, Co), (Cu, Zn), (Mo) 4:3:2:1. There are 8 nonmetals also with the strong order and proportion: (H), (C, N, O), (P, S, Cl), (I) 1:3:3:1 (cf. Yatsimirskii, 1976). The 10 metals exist in the nature in the form of 33 nuclides and the 8 nonmetals exist in the form of 17 nuclides, so altogether  $(33 + 17 = 50)$  gives 50 nuclides (cf. Van Nostrand's Scient. Enc., 1983, p 601). The chemistry and biology arguments and grounds for a reasonable biosystem S-18 are already established (Yatsimirskii, 1976). We are, on the other hand, introducing the physics and Boolean arguments and reasons in the Table 3 and solution (30) ("physics": nucleon number; "Boolean": all presented numbers must be from the Boolean spaces).

**Table 3 Nucleon number within 18 bioelements**

	a	b	c
(n)	218	255	473
(m)	882	1073	1955
(n+m)	1100	1328	2428

**a.** The number of the protons within the 18 bioelements in all atoms of all nuclides (17 for nonmetals and 33 for metals); **b.** The number of the neutrons; **c.** The number of the nucleons; **(n).** Normal state (nonmetals); **(m).** Meta-state (metals). Characteristic points and situations:  $(2 \times 473) = (1 \times 0946)$ ,  $V$  permutation of the III PN which in the relation with first to last permutation gives the sum of the first four perfect numbers  $(9604 - 0946 = 8658)$  (cf position 13d in the Table 2). The total number of the protons is  $1100 = 5 \times 220$  (I FN).

$$2x \odot a (m) - (n) \otimes = 1x \odot b (m) + b (n) \otimes \quad (30)$$

$$2x (882 - 218) = 1x (1073 + 255)$$

Proton and neutron number, within two families, is in the proportional relation with the small integer numbers, in the simplest different digit form of 2:1, through the two simplest inverse arithmetical operations, and that is self-evidently from the result (30).

After our hypothesis these 18 bioelements make the first system of essentiality (in reality the zeroth system) - "strong" system. After that, there are two systems with the "middle" status, middle 1 and middle 2 (two "mixed" systems), S-24 and S-30. Within the system S-24 there are the said 18 bioelements - Table 3 and Solution 30, plus 6 bioelements - Table 4; within the system S-30 we have the elements presented in Tables 3, 4 and 5. The fourth system with the "week" status is made from the remaining elements from first 5 periods and 10 groups of the periodic table (why there are more than 8 groups see in Martynenko and Spitsyn, 1986, p 26: "New versions of the periodic system are often proposed, in which the lanthanides are not arranged in a separate row but are positioned in several rows under other groups"; cf. Rakoëeviæ, 1991, pp 20-21): Li, Be, Sc, Ti, Ga, Rb, Y, Zr, Nb, Ru, Rh, Pd, Ag, Cd, In, Sb, plus Te and Br, with an exclusion of the elements He, Ne, Ar, Kr, Xe and Te.

**Table 4 Nucleon number within 6 bioelements**

	a	b	c	Element
(n)	57	64	<b>121</b>	F, Si, Se
(m)	97	115	<b>212</b>	V, Cr, Sn
(n+m)	<b>154</b>	<b>179</b>	333	

*The designations as in Table 3; the first (lightest) nuclides.*

**Table 5 Nucleon number within the next 6 bioelements**

	a	b	c	Element
(n)	70	90	<b>160</b>	B, Ge, As
(m)	79	94	<b>173</b>	Al, Ni, Sr
(n+m)	<b>149</b>	<b>184</b>	333	

*The designations as in Table 3; the nuclides with the most abundance.*

As we can see, four systems are included: (3x6), (4x6), (5x6), (3x6). The first system is "strong", the last is "weak"; between them are two "middle" ("mixed") systems. Only one system is excluded: (1x6). Altogether, there are  $9 \times 6 = 54$  elements within 5 periods and 10 groups ( $10:5=2:1$ ).

**Table 6 Proton number within four bioelements systems**

	a(n)	a(m)	a(n+m)	
1	123	241	364	S - 18
2	180	338	518	S - 24
3	250	417	667	S - 30
4	87	568	655	S - 18
	640	1564	2204	
	14 x 66			
	12 x 77			

*The designations as in Table 3.*

The 18 bioelements from the first system are essential for the most of organisms (plants and/or animals). Many fewer organisms make a choice of essential elements from first three systems, and the fewest organisms make a choice from all four systems.

After our hypothesis the four systems make a Boolean logical square through first three Pythagorean numbers and first perfect number: 0. (3x6), 1. (4x6), 2. (5x6), 3. (3x6). The proportion in the solution (30) is valid for the first (zeroth) system. The following two systems, precisely their subsystems (subsystem 1 - Table 4 and subsystem 2 - Table 5), are determined with the Boolean square placed in the "Region of maximum possible inversions within the frame of the decimal number system" (Rakoèeviæ, 1994, p 235 and Survey 14 in this study), that means with the Boolean square: (0) 11, (1) 12, (2) 13, (3) 14. The zeroth vertex (11) corresponds with the fourth friendly number 1210 using the square  $11^2 = 121$  and through this correspondence are the main relations, as follows from the tables 7-10. The number of the proton and neutron pair 154-179 of the first subsystem in the Table 4 is taken from the Table 7 (third position); the number of the nucleon pairs 149-184 of the second subsystem - Table 5 is taken from the Table 10 (second position); the number of the nucleon pair 160 - 173 is taken from the Table 9 (third and fourth position at the same time).

**Table 7 The relations of the number 11**

1	$121 + 11 = 132$ $212 - 11 = 201$	→	$(11 \times 6) + 3$
2	$121 + 22 = 143$ $212 - 22 = 190$	→	$(11 \times 4) + 3$
3	$121 + 33 = 154$ $212 - 33 = 179$	→	$(11 \times 2) + 3$
4	$121 + 44 = 165$ $212 - 44 = 168$	→	$(11 \times 0) + 3$
5	$121 + 55 = 176$ $212 - 55 = 157$	→	$(11 \times 2) - 3$
6	$121 + 66 = 187$ $212 - 66 = 146$	→	$(11 \times 4) - 3$
7	$121 + 77 = 198$ $212 - 77 = 135$	→	$(11 \times 6) - 3$

In the "region of maximum possible inversions within the frame of the decimal number system" (Rakoëeviæ, 1994, p 235) are the numbers 11, 12, 13 and 14 (cf. Survey 14 in this Study). This Table shows the relation with the two complements (121 and 212) of the number 333 ( $121 + 212 = 333$ ) (Notice that the number 121 is  $11^2$  and 1210 is the 4<sup>th</sup> friendly number).

**Table 8 The relations of the number 12**

1	$121 + 12 = 133$ $212 - 12 = 200$	→	$(12 \times 6) - 5$
2	$121 + 24 = 145$ $212 - 24 = 188$	→	$(12 \times 4) - 5$
3	$121 + 36 = 157$ $212 - 36 = 176$	→	$(12 \times 2) - 5$
4	$121 + 48 = 169$ $212 - 48 = 164$	→	$(12 \times 0) + 5$
5	$121 + 60 = 181$ $212 - 60 = 152$	→	$(12 \times 2) + 5$
6	$121 + 72 = 193$ $212 - 72 = 140$	→	$(12 \times 4) + 5$
7	$121 + 84 = 205$ $212 - 84 = 128$	→	$(12 \times 6) + 5$

The number 12 is the second number within the "region" of 11-12-13-14 (cf. the legend to Table 7).

**Table 9 The relations of the number 13**

1	$121 + 13 = 134$ $212 - 13 = 199$	→	$(13 \times 5)$
2	$121 + 26 = 147$ $212 - 26 = 186$	→	$(13 \times 3)$
3	$121 + 39 = \mathbf{160}$ $212 - 39 = \mathbf{173}$	→	$(13 \times 1)$
4	$121 + 52 = \mathbf{173}$ $212 - 52 = \mathbf{160}$	→	$(13 \times 1)$
5	$121 + 65 = 186$ $212 - 65 = 147$	→	$(13 \times 3)$
6	$121 + 78 = 199$ $212 - 78 = 134$	→	$(13 \times 5)$
7	$121 + 91 = \mathbf{212}$ $212 - 91 = \mathbf{121}$	→	$(13 \times 7)$

The number 13 is the third number within the "region" of 11-12-13-14 (cf. the legend to Table 7).

**Table 10 The relations of the number 14**

1	$121 + 14 = 135$ $212 - 14 = 198$	→	$(7 \times 9)$
2	$121 + 28 = 149$ $212 - 28 = 184$	→	$(7 \times 5)$
3	$121 + 42 = 163$ $212 - 42 = 170$	→	$(7 \times 1)$
4	$121 + 56 = 177$ $212 - 56 = 156$	→	$(7 \times 3)$
5	$121 + 70 = 191$ $212 - 70 = 142$	→	$(7 \times 7)$
6	$121 + 84 = 205$ $212 - 84 = 128$	→	$(7 \times 11)$
7	$121 + 98 = 219$ $212 - 98 = 114$	→	$(7 \times 15)$

The number 14 is the fourth number within the "region" of 11-12-13-14 (cf. the legend to Table 7).

One can show that these two situations within two subsystems: the situation of the lightest nuclide and the most abundance nuclide, are in an exact correspondence with other two situations (of the heaviest nuclide and of the least abundance nuclide) just and only through the module 9. However, all four (large) systems ( $6 \times 3$ ;  $6 \times 4$ ;  $6 \times 5$ ;  $6 \times 3$ ) are also in relation that are the exact representations of the binary symmetry, harmony and proportion. For example, the proton-neutron (number) pairs for the first system (18 the lightest nuclides) is 364-417 and for the fourth system 655-824; the difference is 222 ( $222 = 6 \times 037$ ), that means that within 36 ( $36 = 6^2$ ) nuclides there are 222 neutrons more than protons. On the other hand, within 18 heaviest nuclides of the first system S-18 there are 824 nucleons (364 protons plus 460 neutrons) what is equal to the 824 neutrons within 18 the lightest nuclides of the fourth system (ratio 1:1). The further examples: there are 417 neutrons within the 13 most abundance metal nuclides in S-24, then 417 protons within 16 metal nuclides in S-30 and, finally, 417 neutrons within 18 the lightest nuclides in the first system S-18 (ratio 1:1:1).

There are two questions: 1. Why is there the number 417? 2. What is the physical nature of these numbers? The answers are in the Table 6, Tables 11-12 and solutions 31-33.

$$\begin{aligned}
 (1 \times 22) + \textcircled{2}(2 \times 79) - 0\textcircled{2} &= 180 \\
 (1 \times 22) + \textcircled{2}(4 \times 79) - 0\textcircled{2} &= 338 \\
 (2 \times 22) + \textcircled{2}(6 \times 79) - 0\textcircled{2} \\
 (1 \times 22) + \textcircled{2}(5 \times 79) - 0\textcircled{2} &= 417 \\
 (1 \times 22) + \textcircled{2}(6 \times 79) - 0\textcircled{2} &= 496 \\
 (1 \times 22) + \textcircled{2}(3 \times 79) - 9\textcircled{2} \\
 (1 \times 44) + \textcircled{2}(1 \times 79) - 0\textcircled{2} &= 123 \\
 (1 \times 127) + \textcircled{2}(3 \times 79) - 0\textcircled{2} = 364 & \quad (33)
 \end{aligned}$$

In the Table 7 we have number 22 in the second position and 44 in the fourth position (the numbers of the proton and neutron pair 154-179 are in the third position - Table 4). In the Table 6 we have the data about proton number within the four systems and we can see that S-24 is "strong" determined by the pairs 22-44 and with the number 79 like in the Solution (31). The first system (S-18) and third system (S-30) have a "middle" determination (exactly "mixed"). The fourth system (S-18) has a "weak" determination, as we can see in the Solutions (32-33). In the Solution (32) we can see that the number 417 is in a "strong" relation with the number 22 (second position in the Table 7), with the number 496 (third perfect number) and also with the number 79, that means with the 79 protons, i.e. the Gold position in the periodic system of elements.

If we imagine the periodic system as three-dimensional, with the surface, the Gold would be at the top and the element Lithium would be at the bottom between the metals. That would be the case from the aspect of electrode potential within the water milieu (standard potential at 25° C for  $\text{Au}/\text{Au}^+ \cong +1.7$  V and for  $\text{Li}/\text{Li}^+ \cong -3.2$  V). We can see in the Tables 11-12 that Gold position (79) through Golden section (Golden mean, i.e., Golden number is a pair 0.618033 .../1.618033 ...), and its portions, determines the key chemically valid distinctions within periodic system of elements. This determination goes, as we can see, through *strong-middle-weak* relation, in the sense 0,1 and 2 deviation steps.

**Table 11 Golden mean portions in the relation of the Gold position (I)**

(9)	79.000	x	1.618...	=	127.824...	(127 - 128)
(8)	79.000	x	0.618...	=	48.824...	(48 - 49)
(7)	48.824...	x	0.618...	=	30.175...	(30 - 31)
(6)	30.175...	x	0.618...	=	18.649...	(18 - 19)
(5)	18.649...	x	0.618...	=	11.525...	(11 - 12)
(4)	11.525...	x	0.618...	=	7.123...	(7 - 8)
(3)	7.123...	x	0.618...	=	4.402...	(4 - 5)
(2)	4.402...	x	0.618...	=	2.720...	(2 - 3)
(1)	2.720...	x	0.618...	=	1.681...	(1 - 2)

**Table 12 Golden mean portions in the relation of the Gold position (II)**

(9)	d - p	78Pt - 79 Au	80 Hg - 81 Tl	(2)
(8)	d - p	48Cd - 49 In		(0)
(7)	d - p	30Zn - 31 Ga		(0)
(6)	p - s	18Ar - 19 K		(0)
(5)	s - p	11Na - 12 Mg	12 Mg - 13 Al	(1)
(4)	p - p	7N - 8 O	9 F - 10 Ne	(2)
(3)	s - p	4B - 5 B		(0)
(2)	s - s	2He - 3 Li		(0)
(1)	s - s	1H - 2 He		(0)

For a full understanding we should notice that a (n) 640 - Table 6 (proton number of all 16 elements which are nonmetals, i.e. are not pure metals) corresponds with a border, marked with number 64, within the system, in the Figure 7; the system in the Table 11 is determined by the fourth perfect number through the inversion ( $127_{10} = 000000111111_2 / 111111000000_2 = 8128_{10}$ ); the difference 14 x 66, i.e. 12 x 77 corresponds with the positions 14 and 12 in NSM II - Table 2; finally, within all four systems of potentially possible bioelements exists a strong simple proportion 1:2:3, i.e., 16:32:48 (48 elements, 32 pure metals and 16 elements that are not purely metal are, in fact, non metals).

For better understanding cf. the text of Stakhov, 1989, p 614: "According to Johannes Kepler, Geometry has two treasures - the Pythagorean theorem and the golden section. The former can be compared with pure gold, the latter with precious stone. The golden section was known to the ancient Babylonians and Egyptians. It forms the basis for Pythagoras' teaching of the number harmony of the world. The term 'golden section' was introduced by Leonardo da Vinci".

The chemistry and biology argumentation and grounds for the system S-24 can be looked up in the Frieden, 1972 (pp 52-60) and in Dickerson,

1978 (pp 70-86: "Twenty-four elements are now known to be essential for the processes of life"). The chemistry and biology argumentation for the system S-30 does not exist, but there are the facts about the essentiality of the next six elements. For example in Frieden, 1972, p 55: "The most likely candidates are aluminum, nickel and germanium. The element boron already appears to be essential for some plants". The chemistry argumentation and reasons for boron and arsenic can be looked up in Lehninger, 1982, and for strontium in Gilyarov, 1989 (p 361).

**Note 1.** The Golden mean portions in Table 11 one can read as:

$1/\phi^1$  (for 48.824 ... in relation to 79.000 ...),  $1/\phi^2$  (for 30.175 ...) etc. to the  $1/\phi^7$  (for 2.720 ...) and  $1/\phi^8$  (for 1.681 ...).

**Note 2.** The idea that the Gold is in the Golden mean position within Mendeleev's periodic system was given to me by Mr Aleksandar Petroviæ, the responsible editor of PHLOGISTON (Journal for the history of science), and I acknowledge him with gratitude for this.

**Note 3.** This Appendix, as well as the Appendices 6-8 are in a link to the doctoral dissertation of Anja Jokiæ, done in Department of Chemistry, Faculty of Science, University of Niš, in which case I was a mentor.

## Appendix 6

### **Four Stereochemical Types of Protein Amino Acids: Synchronous Determination with Chemical Characteristics, Atom and Nucleon Number (I)**

Shcherbak (1993; 1994) and Verkhovod (1994) have shown that the structural and functional distinction of protein amino acid molecules is followed by a strict balanced proportionality of nucleon number. Therefore, it follows, for example, that the number of nucleons in the "heads" ( $\text{H}_2\text{N}-\text{CH}-\text{COOH}$ ) of all non-four-codon amino acids is equal to the number of nucleons in their "bodies", i.e. in the side chains (1110:1110) (Shcherbak, 1994, p. 475). or, the nucleon number in the "heads" of all amino acids which are coded with codons of the pyrimidine type (with a pyrimidine base in the first codon position) is equal to the nucleon number in their "bodies" (814:814) (Verkhovod, 1994, p. 328). The nucleon number in both cited works applies only to the first, the lightest, nuclide for all bioelements - amino acid constituents.

We show, in this Appendix, that the presented law of balanced proportionality of nucleon number is also valid for the structural distinction into the four stereochemical types, as well as for the functional distinction within the individual stereochemical types of protein amino acids (the nucleon number in presented surveys is given only for the side chains of amino acids).

According to E.M. Popov (1989) only one amino acid (G) belongs to the stereochemical type glycine; only one amino acid (P) belongs to type proline; the pair V-I belongs to the stereochemical type valine; and, finally, to stereochemical type alanine belong the following amino acid pairs: I. S-T, C-M, N-Q, D-E, K-R and II. A-L, H-W, F-Y. (The idea about the pairs within two types and about two groups within alanine type is ours.)

<b>Survey 2'</b>						<b>Survey 2''</b>					
S	31		T	45		-S	31	+T			
V	43	121	I	57		+V	90	+I			177
C	47		M	75		+C			+M		
N	58		Q	72		-N	117	-Q			145
D	59	189	E	73		-D		-E			
K	72		R	100		+K	72	+R			100

<b>Survey 2'.1.</b>						<b>Survey 2'.2.</b>					
S	31		T	45		S	31	T	45		102
V	43	132	I	57		V	43	74	I	57	
N	58		Q	72		N	58		Q	72	
D	59		E	73		D	59	117	E	73	145
K	72	131		R	100						

<b>Survey 2'.3.</b>						<b>Survey 2'.4.</b>					
C	47		M	75		M	75		C*	89	
K	72		R	100		D	59		E	73	

In Survey 2' we show that the presented law of balanced proportionality of nucleon number is valid for the system consisting of six pairs: one pair of amino acid type valine and five pairs of the alanine type  $[(S+V+C) +$

$(Q+E+R)] : [(N+D+K) + (T+I+M)] = 366:366$ ; however, the law is also valid for the five pairs (Survey 2'.1.), when the C-M pair is excluded from the system and is replaced by the N-Q pair -  $[(S+V+N) + (E+R)] : [(D+K) + (T, I, Q)] = 305:305$ . With the exclusion of the C-M and K-R pairs from the system only four other pairs remain (Survey 2'.2.), for which the law of balanced proportionality is again valid -  $[(S+V) + (Q+E)] : [N+D] + (T+I)] = 219:219$ . Moreover, this law is also valid for the newly formed system (Survey 2'.3.), consisting of a pair of sulfuric amino acids, C-M, and a pair of base amino acids, K-R, -  $(C+R) : (K+M) = 147:147$ . Finally, the law is also valid when the newly formed system (Survey 2'.4.) is modified in such a way that in the C-M pair seleno-cysteine ( $C^*$ ) replaces cysteine (C), and the pair of base amino acids, K-R, is replaced by a pair of acid amino acids, D-E, -  $(D+C^*) : (M+E) = 148:148$ . The expected correspondence (*Prediction 12*) with the nucleon systems of the remaining amino acids remains the subject of future research. In favour to this prediction we have an atom number balance valid for all eight pairs of alanine type in next sense: first members of first group plus second members of second group equals to second members of first group plus first members of second group  $[(S\ 05 + C\ 05 + N\ 08 + D\ 07 + K\ 15) + (L\ 13 + W\ 18 + Y\ 15) = (T\ 08 + M\ 11 + Q\ 11 + E\ 10 + R\ 17) + (A\ 04 + H\ 11 + F\ 14)] = [(40) + (46) = (57) + (29)]$ .

But that what is surprising is the fact that nucleon number balance within the system in Survey 1 is followed by a new balance (Survey 2'': bold means essential; + essential; + semi-essential; - nonessential) - the amino acid essentiality balance  $[(T+I+M) + (K)] - [(V+C) + (R)] = [(S) + (Q+E)] - [(N+D)]$ . The nucleon number ratio in this case is 59:59. The amino acids on the left from the equal sign are essential or semi-essential (C & R); on the right are nonessential amino acids (Van Nostrand's Scientific Encyclopedia, 1983, p. 117: "Generally, those amino acids which the human body cannot synthesize... are called essential amino acids... the term nonessential is taken to mean those amino acids that are really synthesized in the body"; and still for a semi-essential amino acid which "is essential for the normal growth of the human infant, but to date it is not regarded as essential for adults").

The expected correspondence (*Prediction 13*) with the essentiality - nonessentiality systems of the remaining amino acids must be the subject of future researches. In favor to this prediction we have a strict essentiality - nonessentiality distinction for all 20 protein amino acids within mutation ring of amino acid (genetic) code (Swanson, 1984, p. 191; Rakoëeviæ,

1994, p. 85). Namely, within the mutation ring, 8 amino acids left from the line S-Q, including these two on the line, are nonessential (S, A, G, P, E, D, N, Q); on the right are 8 essential (K, W, F, L, M, I, V, T) and 4 semi-essential (H, R, Y, C) amino acids. The molecule number ratio is ESS : SESS : NESS = 8 : 4 : 8, i.e. 2 : 1 : 2. The nucleon number difference  $[(8 \text{ ESS} + 4 \text{ SESS}) - (8 \text{ NESS})] = [(570 + 335) - (350)] = 555$  is followed, as we see, by a specific balance expressed through "the digital pattern" in which "the numbers are written using the same symbols" (Shcherbak, 1994, p. 475; cf. this  $1 \times 555 = 15 \times 037$  result with the  $2 \times 555 = 30 \times 037$  result for nucleon number within 15 side chains of 15 non-four-codon amino acids). Also, there is an atom number balance:  $[(8 \text{ NESS} + 4 \text{ SESS}) = 8 \text{ ESS}] = [(54 + 48) = 102]$ .

## Appendix 7

### **Fractal Structure of Amino Acid (Genetic) Code**

#### I

In this Appendix it is shown a fractal organization of amino acid code in which the ratio 3:2 appears to be a basic motive. In other words, twenty canonical amino acids of the genetic code appear within the groups of *two* and *three* at the same time. From a such fractal structure it follows that four stereochemical types of protein amino acids are determined with a synchronical balance of shemical characteristics and of atom and nucleon number within the *singlets*, *doublets* and *triplets* of amino acids. These strict regularities provide a new standpoint for addressing questions of evolution of the amino acid code. The presented facts show namely that it is no any sense to speak about evolution of the code, but only about evolution of macromolecules and organisms.

#### II

Shcherbak (1993, 1994) and Verkhovod (1994) have shown that the structural and functional distinction of canonical amino acids of the genetic code is followed by a strict balanced proportionality of nucleon number for the first (lightest) nuclide.

In this study we show that the presented law of balanced proportionality is also valid for the structural and functional distinction into the four stereochemical types of twenty canonical amino acids of the genetic code, synchronically through chemical characteristics and still through atom and nucleon number balance within a fractal structure which basic motive is the ratio 3:2 (Surveys 7 and 8). (The atom number in presented Survey 7 and nucleon number in Survey 8 are given only for the side chains of amino acids).

According to E.M. Popov (1989) only one amino acid (G) belongs to the stereochemical type of glycine, making a doublet, i.e. a pair (G-G) in itself; only one amino acid (P) belongs to the type proline, making a pair (P-P) in itself; the pair V-I belongs to the stereochemical type of valine; and, finally, to the stereochemical type of alanine belong the following amino acid pairs: I. S-T, C-M, D-E, N-Q, K-R and II. A-L, F-Y, H-W. (The idea about the doublets, i.e. pairs and about two classes within alanine type is ours).

From Survey 7 we see that atom balance law is valid for tow classes within alanine type in next manner: AN of the first members of the first class plus AN of the second members of the second class equals to AN of the second members of the first class plus AN of the first members of the second class (Solution 34):

$$\begin{aligned} (\text{S 05} + \text{C 05} + \text{D 07} + \text{N 08} + \text{K 15}) + (\text{L 13} + \text{Y 15} + \text{W 18}) &= 86 \\ (\text{T 08} + \text{M 11} + \text{E 10} + \text{Q 11} + \text{R 17}) + (\text{A 04} + \text{F 14} + \text{H 11}) &= 86 \end{aligned} \quad (34)$$

From Survey 7 we see also a strict accordance and correspondence between atom number balance and chemical characteristics balance. From the aspect of chemical characteristics (the inductive effect, IE, of atom groups within side chain and electron density, ED, in itself etc.), first class of alanine type with the pair V-I of valine type makes a subsystem; the second class of alanine type makes a second subsystem with the pairs of other three types: V-I, G-G and P-P. These two subsystems make a whole fractal system with 12 doublets (pairs) and 8 triplets presented in Survey 7 (doublets : triplets = 3 : 2). Notice that first subsystem has an inner, but the second subsystem has an outer position within the system (cf with inner and outer amino acids in mutation ring in Swanson, 1984, p 191). [Hint. The 12 doublets corespond to the 12 edges and 8 triplets to the 8 vertices on the LIGHT (Logical - Information - Geometric - Homeomorphic-Topological) model of  $B^3$  unit Boolean cube; cf "LIGHT Model and System" in: Rakoèeviæ, 1994, p 53].

As an etalon of IE-ED, the whole system must use G-G pair for non-cyclic and P-P for cyclic side chains; also V-I pair as an etalon for comparation the two subsystems (inner and outer) within one integral whole system. This is really a noteworthy fact: one stereochemical type (alanine type) as a measurement subject, and three other types as measurement etalnons and measurement subjects at the same time!

Bearing all this in mind, we can see, except a self-evident IE-ED balance, still a strict atom number balance between two subsystems (Solution 35):

$$\begin{aligned} (\text{II 27} - \text{I 15}) + (\text{VIII 41} - \text{VII 33}) &= 20 \\ (\text{IV 32} - \text{III 20}) + (\text{VI 38} - \text{V 30}) &= 20 \end{aligned} \quad (35)$$

But atom number balance is also valid for the whole fractal system (Solution 36) and for its first and last triplet-square (Solution 37):

$$\begin{aligned}
 15 + 32 + 30 + 41 &= 118 \\
 27 + 20 + 38 + 33 &= 118
 \end{aligned} \tag{36}$$

$$\begin{aligned}
 15 + 32 = 20 + 27 &= 047 \\
 30 + 41 = 33 + 38 &= 071
 \end{aligned} \tag{37}$$

As we see atom number balance is not valid for the middle triplet-square; from that follows the fractal motive 3:2 again. However, for this middle square there is a balance through nucleon number (Survey 8): the balance for two squares with realization of the fractal motive 3:2 still once.

But not only this. In the Survey 7 we see that from three triplet - squares, two are with the balance (first and last; the middle square is not with balance). This “two from three” situation we see still once again in Survey 8 (cf. Lagerkvist’s rule “two out of three” in Lagerkvist, 1978 and Lagerkvist et al., 1981). Bearing in mind that ratio 3:2 is the basic fractal motive in the middle third (Fig. 8) and middle ninth Cantor set (Fig. 9); that, on the other hand, “THE LIMIT OF THE GOLDEN NUMBERS IS 3/2” (Moore, 1994; see Addenda), and, on the third side that “dimensionality of [dimension]  $N = 0$  is  $n = 3/2$ ” (Koruga, 1995, p 245), all these regularities are clear and expected (Notice the validity of diagonal balance for all triplet squares within the nucleon system in Survey 8 through modulo 9).

In the connection with the said regularities one must notice the ratio between the atom number and nucleon number as the strict proportionalities. Namely, within  $2 \times 12 = 24$  amino acids in the fractal system of amino acids in Survey 7 (about fractal system see in further text) there are exactly  $1 \times 236$  atoms (cf Solution 36 where  $118 + 118 = 236$ ). On the other hand, within  $1 \times 12 = 12$  amino acid side chains in the fractal system of amino acids in Survey 8 (first 12 amino acids without 12 last) there are exactly  $2 \times 236$  nucleons ( $059 + 177 = 236$ ;  $115 + 121 = 236$ ). Within the last  $1 \times 12 = 12$  amino acid side chains in the Survey 8 there are exactly  $1 \times 925$  nucleons ( $189 + 278 = 467$ ;  $245 + 213 = 458$ ;  $458 + 467 = 925$ ), strictly as within  $1 \times 8$  side chains plus  $1 \times 8$  “heads” of  $1 \times 8$  four-codon-amino acids in the codon amino acid system ( $333 + 592 = 925$ ) (cf. Fig. 1 in Shcherbak, 1994, p 475). The nucleon number ratio in the last case is 2:3 and 3:4 because there are 8:12 and 16:12 amino acid entities at the same time (8 side chains plus 8 “heads” equals 16 entities).

The difference of two systems (fractal amino acid systems minus codon amino acid system) corresponds to “the unit change law” (Rakoèeviæ, 1994,

p 36) ( $24 - 23 = 1$ ), and the sum corresponds to the Golden mean through eight root of 47 ( $23 + 24 = 47$ ); eight root of 47 equals  $1.6181 \dots$ , and Golden mean is  $1.6180 \dots$

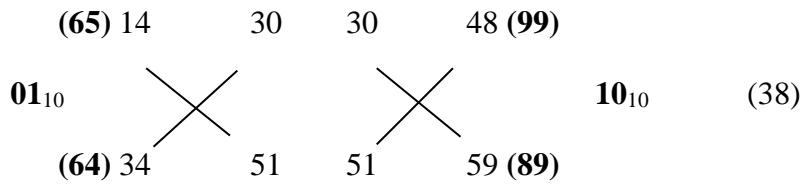
(*Prediction* 14. The total number of conformations for 20 protein amino acids, 405, established by E. M. Popov in Ref. 1989, p 88, must be in a strict relation with atom and nucleon number, in some way).

Knowing all this, now it is self-evidently that amino acid component of the genetic code, like *pu-pyr* base codon component, is arranged as a doublet-triplet system with validity of "strong - weak", i.e. "strong-middle-weak" or "strong-mixed-weak" principle (Lagerkvist et al., 1981); strong-weak effects for doublets, and strong-middle-weak effects for triplets. [Lagerkvist et al., 1981, pp 2640-2641: "reading must be a function of the strength of the interaction between the anticodon and the first two codon nucleotides" (italics M.R); and further on p 2641 see about "strong" codons, "mixed" codons and "weak" codons, that means about such triplets]. For example, in first amino acid triplet in Survey 7 the positive IE follows the next logic: "weak" (side chain H -); "middle" (side chain -  $\text{CH}_3$ ); "strong"(side chain  $\text{CH}_3\text{CHCH}_3$ ); in doublet A - L: "weak" - "strong" etc.

### | III

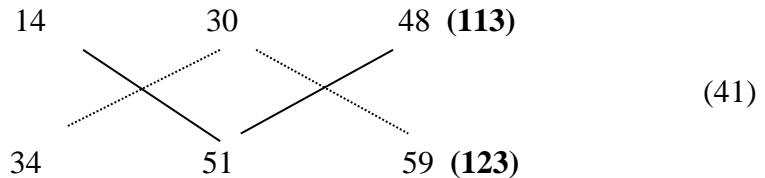
In order to compare two systems - amino acid doublet/triplet system and *pu-pyr* nucleotide doublet/triplet system - one must rearrange the system from Survey 7, as we made in Survey 7.1. and 7.2. The basic fractal motive, the ratio 3:2, for nucleotide system now is clear and self-evident: for doublets two distinctions, for triplets three distinctions (the fractal structure of amino acid system see in the further text). For doublets: from 4 the same nucleotide pairs (U-U, U-U, U-U, U-U) to be 2 and 2, or to be 2 and 2 the same pairs with crossing-over (U-C, U-C and/or A-G, A-G) and then, to be 1 and 1 pair. For triplets: to be 4 and 4 triplets, then 2 and 2, and finally, to be 1 and 1 triplet (as here with middle base U, analogous situations we have with middle base C, A, G respectively).

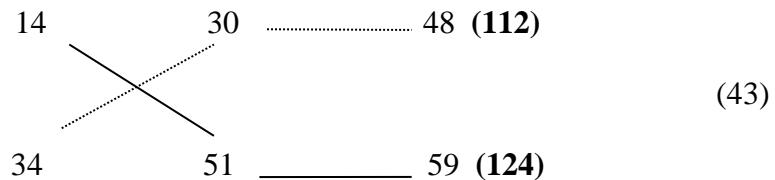
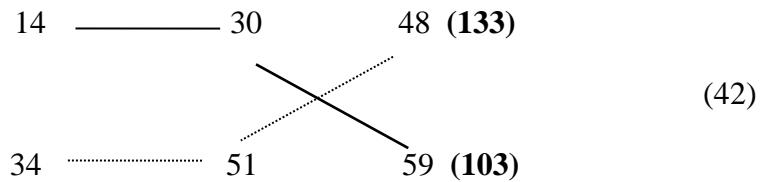
Within the system in Survey 7 doublets are horizontally, but the triplets vertically arranged; within the system in Survey 7.1. and 7.2. in vice versa arrangement. In this vice versa arrangement two systems of amino acid triplets (the first system in Survey 7.1. and the second system in Survey 7.2.) show new arithmetical regularities - new proportionalities and new balances, presented in Solutions 38-45, then in 38'-45', and finally in Solutions 46-50.



$$\begin{array}{ccccc}
 14 - & 16 - & 30 - & 18 - & 48 \\
 | & | & | & | & | \\
 \mathbf{20} & \mathbf{21} & \mathbf{11} & & \\
 | & | & | & | & | \\
 34 - & 17 - & 51 - & 08 - & 59
 \end{array}
 \quad (39)$$

$$\begin{array}{l}
 20 + 17 = (37 + 0) \\
 21 + 16 = (37 + 0) \\
 21 + 08 = (28 + 1) \\
 11 + 18 = (28 + 1)
 \end{array}
 \quad \left. \begin{array}{l} \boxed{\phantom{000}} \\ \boxed{\phantom{000}} \\ \boxed{\phantom{000}} \\ \boxed{\phantom{000}} \end{array} \right\} (67 - 1) \quad (40)$$





$$(4) \ 103 - 9 - 112 (4)$$

$$(5) \ 113 - 0 - 113 (5)$$

(44)

$$(6) \ 123 - 0 - 113 (6)$$

$$(7) \ \underline{133 - 9 - 124} (7)$$

$$2 \times \underline{236} \quad 2 \times \underline{236}$$

$$2 \times \underline{632} \quad 2 \times \underline{632}$$

$$\underline{\quad} \quad \underline{\quad}$$

$$16 \times \underline{79} \quad 16 \times \underline{79}$$

$$4567 \mid 7654 \longrightarrow 12221$$

(45)

$$4567 \mid 7654 \longrightarrow 12221$$

$$\underline{24442}$$

$$\begin{array}{ccccccc}
 & (63) & 15 & & 33 & & 46 (94) \\
 11_2 & & & \times & & \times & 00_2 \\
 & (66) & 33 & & 48 & & 61 (94) \\
 & & & & & & (38')
 \end{array}$$

$$\begin{array}{ccccccc}
 15 & - & \mathbf{18} & - & 33 & - & \mathbf{13} & - & 46 \\
 | & & | & & | & & | & & | \\
 \mathbf{18} & & \mathbf{15} & & \mathbf{15} & & \mathbf{15} & & \\
 | & & | & & | & & | & & | \\
 33 & - & \mathbf{15} & - & 48 & - & \mathbf{13} & - & 61 \\
 & & & & & & & & (39')
 \end{array}$$

$$\begin{array}{c}
 18 + 15 = (37 - 4) \\
 18 + 15 = (37 - 4) \\
 13 + 15 = (28 - 0) \\
 13 + 15 = (28 - 0)
 \end{array}
 \left. \begin{array}{c} \boxed{ } \\ \boxed{ } \\ \boxed{ } \\ \boxed{ } \end{array} \right] \quad (67 - 6) \quad (40')$$

$$\begin{array}{ccc}
 33 & & 46 (113-4) \\
 & & 15
 \end{array}$$

(41')

$$33 \quad 48 \quad 61 (123+4)$$

$$15 \quad \text{---} \quad 33 \quad 46 (133-6)$$

(42')

$$33 \quad \text{.....} \quad 48 \quad 61 (103+6)$$

$$\begin{array}{ccccc}
 15 & & 33 & \dots & 46 \ (112-0) \\
 \diagup \diagdown & & & & \\
 33 & & 48 & \longrightarrow & 61 \ (124+0)
 \end{array}
 \quad (43')$$

$$(1) \ (103+6) \quad — \quad 3 \ — \quad (112-0) \ (4)$$

$$(1) \ (113-4) \quad — \quad 0 \ — \quad (113-4) \ (1)$$

(44')

$$(1) \ (123+4) \quad — \quad 0 \ — \quad (123+4) \ (1)$$

$$(1) \ (133-6) \quad — \quad 3 \ — \quad \underline{(124+0)} \ (7)$$

$$2 \times \underline{236} \quad \quad \quad 2 \times \underline{236}$$

$$\underline{2 \times 632} \quad \quad \quad \underline{2 \times 632}$$

$$16 \times \underline{79} \quad \quad \quad 16 \times \underline{79}$$

$$1111 \mid 1111 \longrightarrow 02222$$

(45')

$$4117 \mid 7114 \longrightarrow 11231$$

$$\overline{13453}$$

$$24442 - 13453 = (7992 + 2997) \ (\text{cf sol. 56}) \quad (46)$$

$$\begin{array}{ccc}
 \mathbf{15} & \mathbf{33} & \mathbf{46} \\
 -1 & -3 & +2 \\
 14 & 30 & 48 \\
 & & \\
 34 & 51 & 59 \\
 +1 & +3 & -2 \\
 \mathbf{33} & \mathbf{48} & \mathbf{61}
 \end{array} \tag{47}$$

$$\begin{array}{ccc}
 33 - 15 & | & 48 - 33 & | & 61 - 46 \\
 & & & & \\
 18 & & 15 & \dots & 15 \ (\mathbf{50}) \\
 \diagup \diagdown & & & & \\
 20 & & 21 & \dots & 11 \ (\mathbf{50}) \\
 & & & & \\
 34 - 14 & | & 51 - 30 & | & 59 - 48
 \end{array} \tag{48}$$

$$\begin{array}{ccc}
 18 & & 15 & & 15 \ (\mathbf{50+4}) \\
 \diagup \diagdown & & \diagup \diagdown & & \\
 20 & & 21 & & 11 \ (\mathbf{50-4}) \\
 & & & & \\
 18 & \dots & 15 & & 15 \ (\mathbf{56\pm 00})
 \end{array} \tag{49}$$

$$\begin{array}{ccc}
 18 & \dots & 15 & & 15 \ (\mathbf{56\pm 00}) \\
 & & \diagup \diagdown & & \\
 20 & \dots & 21 & & 11 \ (\mathbf{56-12})
 \end{array} \tag{50}$$

In Solutions 38 and 38' the diagonal balance law is valid through Boolean square (the differences 00, 01, 10, 11). In Solutions 44 and 44' still once Boolean square determination (0, 1, 2, 3) and the determination through first three perfect numbers at the same time. Namely,

$112 = 4/1 \ 28$ ;  $124 = 1/4 \ 496$ ;  $124 - 112 = 2 \times 6$ . The determination through first and second perfect numbers once more in Surveys 40 and 40'. The number 28 is second perfect number; 37 is the next "28" in modulo 9; within the result 67 the first perfect number exists as  $4/4$  from 6 and second perfect number as  $1/4$  from 28; through number 037 the system in Solution 40 is in relation with Shcherbak's system of multiples of 037 (cf Shcherbak, 1994, p 476, Table 1) and through result  $67 - 1 = 66$  in relation with NMS II, position 1d (NMS III in Table 2 in this book).

From result 41 - 43 follows the result in Solution 44; from 41' - 43' follows the result in Solution 44'. The congruent classes (4, 5, 6, 7 and 1, 1, 1, 1, 1 and 4, 1, 1, 7) taken as four-digit numbers as in Solutions 45 and 45' give a final result in Solution 46 ( $7992 + 2997$ ) which result is the number from position 12d in NMS III in Table 2 (7992) plus its inversion (2997).

The two systems (in Survey 7.1. and 7.2.) are also determined with last three factors of the first perfect number, with 236 in Solutions 44 and 44'; and with first three factors, with 1, 2, 3 in Solution 47 (cf the said about  $1 \times 236$  of atoms and  $2 \times 236$  of nucleons).

The relation between two systems is a balanced proportionality 1:1 in one manner (Solution 48) and in another manner (Solution 49); in the third case (Solution 50) the balance is not 1:1 but the determination is realized through the first  $2 \times 6 = 12$  and the second perfect number  $2 \times 28 = 58$ .

#### IV

The reader must notice still one "hidden" doublet-triplet system within the system in Survey 7. More exactly, that new system is a doublet-triplet and doublet-doublet system at the same time. The doublet-doublet system contains two pairs of alanine-type with cyclic side chains (F-Y and H-W). The rest of six pairs ( $2 \times 3$ ) of alanine-type with non-cyclic side chains (A-L, S-T, C-M, D-E, N-Q and K-R) makes the first subsystem within the new doublet-triplet system; the second subsystem is mixed (cyclic: P-P and non-cyclic: G-G and V-I) and it contains three etalon-pairs, each pair from one of remaining three stereochemical types of canonical amino acids. These "new divisions reveal new balances" (Verkhovod, 1994) also through fractal motiv 3:2 - three balances within atom number system and two within nucleon number system (Surveys 9 and 10).

#### V

But that what is surprising is the fact that atom number balance follows also the essentiality - nonessentiality distinction within the system of twenty

cannonical amino acids of the genetic code, also through the same fractal motive 3:2. In mutation ring of amino acid (genetic) code (Swanson, 1984, p 191; Rakoèeviæ 1994, p 85), 8 amino acids left from the line S-Q, including these two on the line, are nonessential, NESS (S, A, G, P, E, D, N, Q); on the right are 8 essential, ESS (K, W, F, L, M, I, V, T) and 4 semi-essential, SESS (H, R, Y, C) amino acids (Van Nostrand's Scientific Encycl., p 117: "Generally, those amino asids which the human body cannot synthesize ... are called essential amino asids ... the term nonessential is taken to mean those amino acids that are really synthesized in the body"; and still for a semi-essential amino acid which "is essential for the normal growth of the human infant, but to date it is not regarded as essential for adults"). Thus, a strict atom number balance through the fractal motive 3:2 (12:8 amino acids) is presented in Solution 51:

$$[(8 \text{ NESS} + 4 \text{ SESS}) = 8 \text{ ESS}] = [(54 + 48) = 102] \quad (51)$$

## VI

From the Surveys 7-10 and then 11-14 it is self-evidently that the doublet-triplet system of 20 canonical amino acids of the genetic code is arranged as a strict fractal structure (Falconer, 1990); the structure with the form of self-similarity, expressed through always the same and equal ratio 3:2 in the sense to be *three* and *two* at the same time (Falconer, 1990, p XVIII: "The word 'fractal' was coined by Mandelbrot in his fundamental essay from the Latin *fractus*, meaning broken, to describe objects that were too irregular to fit into a traditional geometrical setting...Fractals have some degree of self-similarity - they are made up of parts that resemble the whole in some way"); the fractal structure, expressed through three manner within two realities - phisical and logical reality. First manner, first reality: if in the beginning there are three doublets and two triplets, then: within any of two subsystems (inner and/or outer) there are two times more of doublets and triplets; within whole system there are four times more of doublets and triplets. Second manner, second reality: for the doublets, two binary distinctions are possible: to be six doublets within inner subsystem, and more six doublets within outer subsystem; then, to be three and three doublets within both subsystems; for the triplets, three binary ditinctions are possible: to be four triplets within inner subsystem and four triplets within outer subsystem; then, to be two and two; and, finally, to be one and one

triplet as a whole. Third manner, both realities: the atom number differences ratio for two and two triplets within both subsystems is the same, 12:8, that means 3:2. ( *Prediction 15.* The conformation number differences ratio for triplets must be also 3:2 in some way).

One must notice that the doublet-triplet system 12:8 is a unique and only one system from all doublet-triplet systems; only this system can have a full “3:2 ratio” fractal structure (cf. the preceding system 9:6 and following system 15:10; also double system 24:16 etc.). [Hint. The results 16x79 in Solution 44 cf with the results in Solutions 31-33 and Table 11. For the results in Surveys 11-13 notice that there are 15 non-cyclic and 05 yes-cyclic amino acids and at the same time 18 and 06 after the Survey 7. If so, then 15 entities plus 05 entities equals 20 entities; these 20 entities plus 10 distances ( $15 - 05 = 10$ ) equals 30 ( $30 : 20 = 3 : 2$ , what is the basic fractal motive). Notice also that only two systems are possible, 5th and 6th in Survey 11, because the next connections are valid: 66 is position 1d in NSM II (cf Table 2); 110 is a half of 220 (first friend number); and 11, 12, 13 are from the “region of maximum possible inversions within the frame of the decimal number system” (cf Survey 14). The final result in Survey 13 (38610) is the sum of first six friendly numbers (or: first three pairs):  $(220 + 284) + (1184 + 1210) + + (17296 + 18416) = 38610$ ].

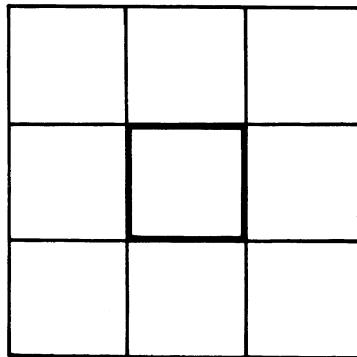
## VII

In comparison Fig. 1 and Fig. 8 we see that the binary tree of the genetic code represents a realization of the middle third Cantor set (the triadic Cantor set) in specific manner: all open middle parts are deleted; that means that middle part is empty. But at the same time genetic code corresponds to the middle ninth Cantor set (the nonadic Cantor set) in Fig. 9.



**Fig. 8.:** "Construction of the triadic Cantor set. The initiator is the unit interval  $[0, 1]$ . The generator removes the open middle third. The figure shows the construction of the five first generations.  $D = \ln 2 / \ln 3 = 0.6309$ " (After: Falconer, 1990).

One must notice that the middle third Cantor set corresponds to the initial Boolean vector (as a middle part) and its maximum possible two neighbors (left part and right part). The initial vector itself is empty, the two parts are full. The 6th generation of this system represents a realisation of 64 codons.



**Fig. 9.** Construction of the middle ninth Cantor set.

The next step within the Boolean spaces,  $B^n$ , after initial vector,  $B^1$ , is a Boolean square,  $B^2$ , with maximum possible still 8 squares as neighbors (Fig. 9). The model in Fig. 9 is also an adequate model for the genetic code: initial square (middle part) is full with 24 amino acids from Survey 7; the neighbor 8 squares are full with 8 classes of codons  $8 \times 3 = 24$  nucleotides from binary tree of genetic code (Fig. 1) (to be "full" it means to be in full correspondence).

As we also see the genetic code is a middle third Cantor set with 6 generations and a middle ninth Cantor set with only 1 generation at the same time (cf the relation between 1 and 6 in Solution 40 and 40'). The both sets are fractal sets *per se* with ratio 3:2 as their basic motive, and with a fractal dimension in the range  $0 < D < 1$  (Falconer, 1990, p XIII: "The middle third Cantor set is one of the best known and most easily constructed fractals") (about correspondence between the ratio 3:2 and the Golden mean see in *Addenda*).

## VIII

Knowing that the amino acid (genetic) code is a fractal structure with basic motive 3:2, and, on the other hand, bearing in mind that "the middle

third Cantor set is one of the... most easily constructed fractals" (Falconer, 1990, p XIII) with logic: to be three and two at the same time; that "the limit of the golden numbers is  $3/2$ " (Moore, 1993, p 211); that "dimensionality of [dimension]  $N = 0$  is  $n = 3/2$ " (Koruga, 1995, p 245), and that "the Hausdorff dimension  $D_H^{C(0)}$  of a randomly Cantor middle third set for  $N = 0$  is  $D_H^{C(0)} = GM$ - or  $\phi$ , where  $GM^- = (\sqrt{5} - 1)/2$  is the Golden Mean" (Koruga, 1995, p 249), all these strict regularities presented in this work provide "a new standpoint for addressing questions of selection vs random drift in the evolution of the code" (Swanson, 1984, p 201). [Hint. To understand why "a randomly Cantor . . . set, cf. the Reference: Mauldin et al., 1986, p 325: "Of course, by a Cantor set we mean a compact, perfect, 0-dimensional metric Space"; and p 342: "with probability one, we obtain a

Cantor set with Hausdorff dimension  $\alpha$ , where . . .  
 $\alpha = (\sqrt{5} - 1)/2$  ].

**Survey 7 (atom number)**

<b>I</b>	<b>II</b>
$G\ 01$	$01\ G$
$A\ 04$	$13\ L$
$V\ 10$	$13\ I$
<b>III</b>	<b>IV</b>
$S\ 05$	$08\ T$
$C\ 05$	$11\ M$
$V\ 10$	$13\ I$
<b>V</b>	<b>VI</b>
$D\ 07$	$10\ E$
$N\ 08$	$11\ Q$
$K\ 15$	$17\ R$
<b>VII</b>	<b>VIII</b>
$P\ 08$	$08\ P$
$F\ 14$	$15\ Y$
$H\ 11$	$18\ W$

(**Note:** One must notice that two last triplets must exist:  
PFH-PYW as here in Survey 7  
and PHF-PWY as there in  
Survey 7').

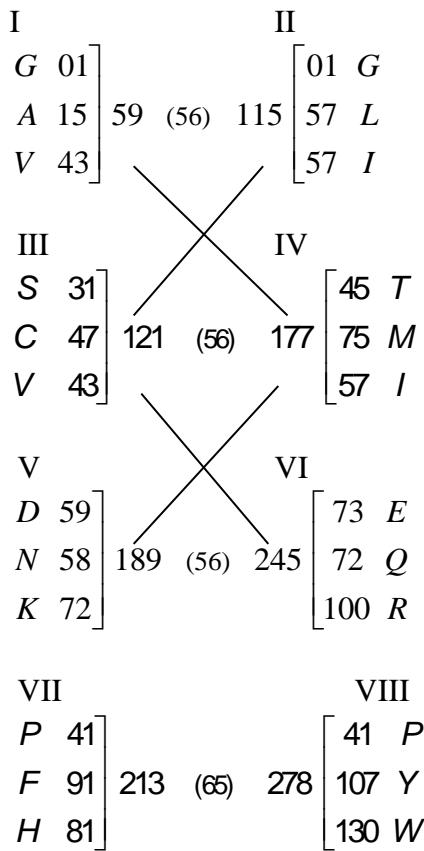
**Survey 7.1.**

15	33	46
G	A	V
G	L	I
S	C	V
T	M	I
<hr/>		
D	N	K
E	Q	R
P	F	H
P	Y	W
33	48	61

**Survey 7.2.**

UUU	14	30	48
UUC	G	A	V
UUA	S	C	V
UUG	G	L	I
<hr/>			
CUU	D	N	K
CUC	T	M	I
CUA	P	F	H
CUG	E	Q	R
<hr/>			
	34	51	59

**Survey 8 (nucleon number)**



**Survey 9 (atom number)**

$G$	01		01	$G$
$P$	08	19	08	$P$
$V$	10		13	$I$
$A$	04		13	$L$
$S$	05	14	08	$T$
$C$	05		11	$M$
$D$	07		10	$E$
$N$	08	30	11	$Q$
$K$	15		17	$R$
$F$	14	25	15	$y$
$H$	11		18	$W$

**Survey 10 (nucleon number)**

$G$	01		01	$G$
$P$	41	85	41	$P$
$V$	43		17	$I$
$A$	15		57	$L$
$S$	31	93	45	$T$
$C$	47		75	$M$
$D$	59		73	$E$
$N$	58	189	72	$Q$
$K$	72		100	$R$
$F$	91	172	107	$y$
$H$	81		130	$W$

VI	V	IV	III	II	I	1	2	3	4	5
0	1	2	3	4	5	V	IV	III	II	I
0	1	2	3	4	5	6	7	8	9	1
00 (00)	03 (02)	06 (04)	09 (06)	12 (08)	15 (10)	18 (12)	21 (14)	24 (16)	27 (18)	30 (20)
00 00	01 02	02 03	03 04	04 05	05 06	06 07	07 08	08 09	09 10	10 10
00 (00)	04 (06)	08 (06)	12 (12)	16 (18)	20 (24)	24 (30)	28 (36)	32 (42)	36 (48)	40 (54)
12 18	20 30	28 42	36 54	44 66	52 78	60 90	68 130	102 150	114 170	190 190
30 + 50 + 70 + 90 = (120 x 2)										
510 - 110 = 400										
510 - 120 = 390										
510 - 130 = 380										
1170										
1170 x 3 = 3510										
3510 + 0351 = 3861										
3861 x 10 = 38610										

## Survey 14

<b>0</b>	$11 \times 1 = 11$ $11 \times 2 = 22$ $11 \times 3 = 33$	$11 \times 1 = 11$ $11 \times 2 = 22$ $11 \times 3 = 33$	$11^2 = 121$
<b>1</b>	$12 \times 1 = 12$ $12 \times 2 = 24$ $12 \times 3 = 36$	$21 \times 1 = 21$ $21 \times 2 = 42$ $21 \times 3 = 63$	$12^2 = 144$ $21^2 = 441$
<b>2</b>	$13 \times 1 = 13$ $13 \times 2 = 26$ $13 \times 3 = 39$	$31 \times 1 = 31$ $31 \times 2 = 62$ $31 \times 3 = 93$	$13^2 = 169$ $31^2 = \overline{961}$
<b>3</b>	$14 \times 1 = 14$ $14 \times 2 = 28$ $14 \times 3 = ?$	$41 \times 1 = 41$ $41 \times 2 = 82$ $41 \times 3 = ?$	$14^2 = 196$

*Region of maximum possible inversions within the frame of the decimal number system (after: Rakoëeviæ, 1994, p 235)*

## Appendix 8

### **Four Stereochemical Types Of Protein Amino Acids: Synchronical Determination With Chemical Characteristics, Atom and Nucleon Number (II)**

Shcherbak (1993; 1994) and Verkhovod (1994) have shown that the structural and functional distinction of protein amino acid molecules is followed by a strict balanced proportionality of nucleon number.

We show, in this Appendix, that the presented law of balanced proportionality of nucleon number (NN), as well as of atom number (AN) corresponds with the structural and functional distinctions, present in the four stereochemical types of protein amino acids, as well as in the six types of their precursors within the biosynthesis paths. The atom number in presented Survey 7' and nucleon number in Survey 8' are given only for the side chains of amino acids: bold - essential (ESS); italic - semiessential (SESS); the other - nonessential (NESS).

According to E.M. Popov (1989) only one amino acid (G) belongs to the stereochemical type of glycine; only one amino acid (P) belongs to type of proline; the pair V-I belongs to the stereochemical type of valine; and, finally, to the stereochemical type of alanine belong all other - 16 amino acids.

The presented classification comes from the amino acid conformation states. However, the same four stereochemical types, two and two (G, A and V, P), following our idea, come also 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 the other 19 amino acids has a hydrogen chain of this kind. The side chain of alanine ( - CH<sub>3</sub>), or, in relation to glycine, ( - CH<sub>2</sub> - H) follows from the shortest possible noncyclic hydrocarbon chain (CH<sub>4</sub>), and still 15 amino acids have the alanine - analogue side chain in the form ( - CH<sub>2</sub> - R). The side chain of valine (H<sub>3</sub>C-CH-CH<sub>3</sub>) follows from the shortest possible cyclic hydrocarbon, from cyclopropane, with a permanent openness and with a linkage to the "head" of amino acid through only one vertex of cyclopropane "triangle"; still only one amino acid, isoleucine, belongs to this type with the side chain H<sub>3</sub>C-CH-CH<sub>2</sub>-CH<sub>3</sub>. The proline type (only with proline) follows from the same source (cyclopropane), but with a permanent non-openness

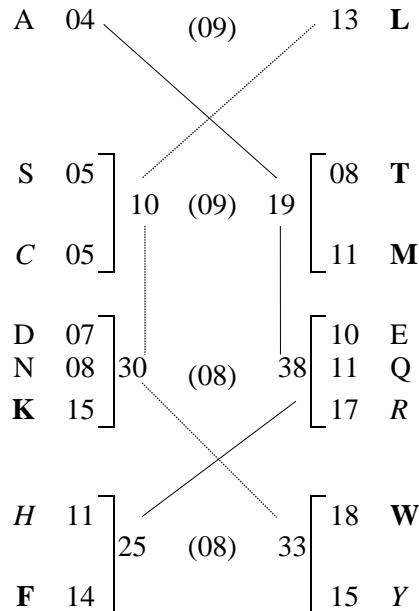
and with a linkage to the "head" through two vertices of cyclopropane "triangle".

### Survey 7' Atom Number within 12 Doublets and 8 Triplets of "24" Amino Acids

CE	HP	I	II	HP	CE
-0.16	-0.4	G 01	01	G	-0.4
-0.09	+1.8	A 04	15	<b>L</b>	+3.8
-0.52	+4.2	<b>V</b> 10	(12)	13	-0.54
				13	-0.56
				<b>I</b>	
III			IV		
+0.24	-0.8	S 05	08	<b>T</b>	-0.7
-0.73	+2.5	C 05	20	<b>M</b>	+1.9
-0.52	+4.2	<b>V</b> 10	(12)	11	-0.57
				13	+4.5
				<b>I</b>	-0.56
V			VI		
+0.69	-3.5	D 07	10	<b>E</b>	-3.5
+0.52	-3.5	N 08	30	<b>Q</b>	-3.5
+1.46	-3.9	<b>K</b> 15	(08)	11	+0.91
				17	+0.87
				<b>R</b>	-4.5
VII			VIII		
+0.46	-1.6	P 08	08	<b>P</b>	-1.6
±0.00	-3.2	H 11	33	<b>W</b>	-0.9
-0.56	+2.8	<b>F</b> 14	(08)	18	-0.25
				15	+0.42
				<b>Y</b>	-1.3

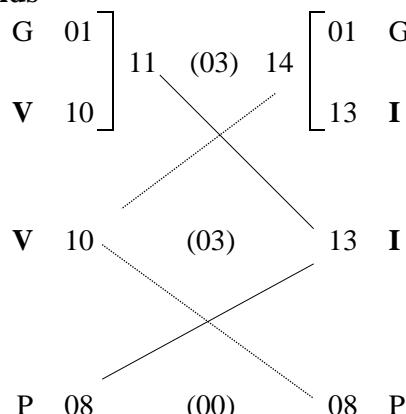
On the full line, as well as on the dotted line, there are 118 atoms; HP: hydropathy index on a number unnamed scale; CE: cloister energy in kcal/mol.

### Survey 7'.1. Atom Number within 8 Pairs of Alanine type amino acids



On the full line, as well as on the dotted line, there are 86 atoms; the differences 8 and 9 ( $9 - 8 = 1$ ) express "the minimum change relation among the amino acids" (Swanson, 1984, p 191). The order follows from the system in Survey 7'.

### Survey 7'.2. Atom Number within 4 Pairs of non-Alanine Type Amino Acids



The two lines have 32 atoms each. The number of amino acids and the order comes from the system in Survey 7'.

## Survey 8' Nucleon Number within 12 Doublets and 8 Triplets of "24" Amino Acids

I		II	
G 01		01 G	
A 15	59	57 L	
V 43	(56)	57 I	
III		IV	
S 31		45 T	
V 43	121	57 I	
C 47	(56)	75 M	
V		VI	
N 58		72 Q	
D 59	189	73 E	
K 72	(56)	100 R	
VII		VIII	
P 41		41 P	
H 81	213	130 W	
F 91	(65)	107 Y	

Within the first triplet square both lines have 236 nucleons each; within the second triplet square 366 each; within the third triplet square, first line: 467 - 0 and second line: 467 - 9 (one cycle less in module 9; cf. nucleon cycles in module 9 on Table 1 in: Shcherbak, 1994, p 476); the two permutations of differences (56 three times and 65 once) express "the minimum change relation" (56 - 56 = 0; 65 is a permutation of 56).

Knowing that valine type is arranged into the pairs (with only one pair) we researched the alanine type following the same concept. The result was as it is shown in the Survey 7'.1. There are eight pairs of alanine type, classified into the two classes. The five amino acid pairs belong to the first (inner) class, with the presence of oxygen, nitrogen or sulphur atoms within a noncyclic side chain, in the form of a functional group: 1. with hydroxy group (S-T), 2. with sulphur atom (C-M), 3. with carboxylic group (D-E), 4. with amide group (N-Q) and 5. with amino group (K-R). Three other and different pairs belong to the second (outer) class: 1. with hydrocarbon noncyclic chain, without a hetero atom (A-L), 2. with aromatic chain, with hetero atom (H-W) and 3. with aromatic chain, without hetero atom (F-Y).

This is really a noteworthy fact: one stereochemical type (alanine type) as a measurement subject, and three other types as measurement etalons and measurement subjects at the same time (Survey 7'.2.).

Having this result and bearing in mind that the interactions between the side chain and "head" of an amino acid come from the possible inductive effect (IE) and/or electron density (ED), the following was clear to us: glycine must be a pair within itself (G-G) as an etalon of IE-ED for all other pairs of noncyclic amino acids; proline must be also a pair within itself (P-P) as an etalon for cyclic amino acids; the pair V-I must be an etalon for the linkage of two subsystems: cyclic and noncyclic.

The Surveys 7'- 8' and Tables (13-15) are providing proofs for this discussion. From the Survey 1 one can see that the whole system of 20 canonical amino acids of the genetic code is arranged as a doublet-triplet system with 12 doublets (pairs) and 8 triplets. [Hint. 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).]. An IE-ED balance within this system is self-evident. For example, through the parameters of hydropathy, HP (Kyte and Doolittle, 1982), and "cloister energy", CE (Rackovsky and Scheraga, 1977; Swanson, 1984), as indirect measurements of IE-ED, the first 10 doublets are "strong" pairs, the last (F-Y) is "weak" ("weak", because both amino acids are with the benzene ring, but only Y is with the hydroxy-group), and the first to last (H-W) is "mixed" pair (because both amino acids are aromatic and heterocyclic, but only W is with the benzene ring) (cf. with "strong", "weak" and "mixed" codons in: Lagerkvist et al., 1981, pp 2640-2641). From the aspect of the triplets, there are 3, 2, 0, 1 of the negative values in CE in both columns; in HP, on the right: 1, 1, 3, 3 and on the left: 1 (why not 0?), 1, 3, 2.

All these IE-ED balances (with only one exception, or more exactly, semi-exception) are followed with a strict atom number balance (Survey 7') and nucleon number balance (Survey 8'). By this, the order of doublets and triplets is established with a strict atom and/or nucleon number increasing from one to another next amino acid in correspondence with the involvement of the biosynthetic precursors one after another (cf the order in Survey 7' with the order in Table 13). This involvement is followed by a balanced proportionality: the less complex precursors (one with phospho

function and two without) appear twice and the more complex precursors (two with phospho function and one without) appear once.

**Table 13 The Order of Involvement of Amino Acid Precursors in Biosynthesis Paths**

1. 3-phosphoglycerate	G			•	•	•
		•⊕	•⊕			
2. pyruvate	A	L	V	•		•
					•⊕	
3. oxaloacetate	•	•⊕	I	•	•⊕	•
1. 3-phosphoglycerate	S	C		•	•⊕	•
			•⊕			
2. pyruvate	•	•⊕	V	•	•⊕	•⊕
3. oxaloacetate	T	M	I	D	N	K
4. 2-oxoglutarate	•⊕	•	P	E	Q	R
5. ribose-5-phosphate	•	•	H		•⊕	•
					•⊕	
6. phosphoenolpyruvate plus eritrose-4-phosphate	•⊕	•	W	F	Y	•⊕

*This Table corresponds with the system in Survey 7'. From the first three (less complex) precursors follow  $9 + 3 = 12$  amino acids. From the last three (more complex) precursors follow  $5 + 3 = 8$  amino acids (5 cyclic, 3 non-cyclic). The ratio is the same as the doublets - triplets ratio within the system in Survey 7' (12:8 = 3:2).*

The presented balances which follow the presented involvement of the biosynthetic precursors, follow also the essentiality (ESS) - semiessentiality (SESS) - nonessentiality (NESS) distinctions within the system of 20 protein amino acids. For example, within the inner subsystem in Survey 8' (triplets III, IV, V, VI) the balance exists not only through the sums, but also through the differences (Solution 52) :

$$[(T+I+M) + (K)] - [(V + C) + (R)] = [(S) + (Q + E)] - [(N + D)] \quad (52)$$

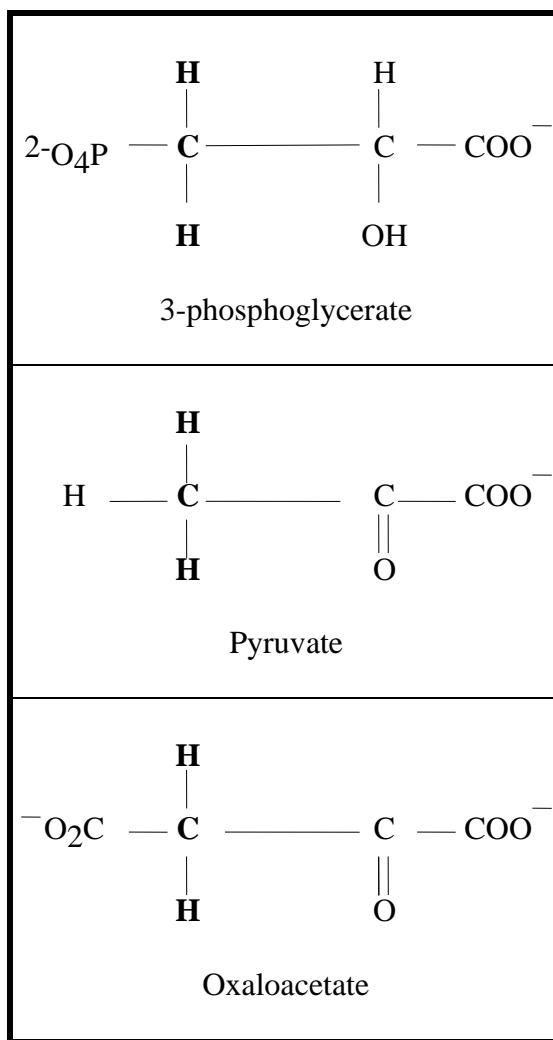
$$[(45 + 57 + 75) + (72)] - [(43 + 47) + (100)] = [(31) + (72 + 73)] - [(58 + 59)]$$

Table 14. First Nine Amino Acids (Side Chains) with the Less Complex Precursors

1.	$  \begin{array}{c}  \text{H} \\    \\  \text{H}_2\text{N} - \text{C} - \text{COOH} \\    \\  \text{H}  \end{array}  $ Glycine (G)		
2.	$  \begin{array}{c}    \\  \text{H} - \text{C} - \text{H} \\    \\  \text{H}  \end{array}  $ Alanine (A)	$  \begin{array}{c}    \\  \text{H} - \text{C} - \text{H} \\    \\  \text{CH} \\  \text{H}_3\text{C} - \text{CH}_3  \end{array}  $ Leucine (L)	$  \begin{array}{c}    \\  \text{CH} \\  \text{H}_3\text{C} - \text{CH}_3  \end{array}  $ Valine (V)
3.			$  \begin{array}{c}    \\  \text{CH} \\  \text{H}_3\text{C} - \text{CH}_2 - \text{CH}_3  \end{array}  $ Isoleucine (I)
1.	$  \begin{array}{c}    \\  \text{H} - \text{C} - \text{H} \\    \\  \text{OH}  \end{array}  $ Serine (S)	$  \begin{array}{c}    \\  \text{H} - \text{C} - \text{H} \\    \\  \text{SH}  \end{array}  $ Cysteine (C)	
2.			$  \begin{array}{c}    \\  \text{CH} \\  \text{H}_3\text{C} - \text{CH}_3  \end{array}  $ Valine (V)
3.	$  \begin{array}{c}    \\  \text{H} - \text{C} - \text{CH}_3 \\    \\  \text{OH}  \end{array}  $ Threonine (T)	$  \begin{array}{c}    \\  \text{H} - \text{C} - \text{H} \\    \\  \text{CH}_2 - \text{S} - \text{CH}_3  \end{array}  $ Methionine (M)	$  \begin{array}{c}    \\  \text{CH} \\  \text{H}_3\text{C} - \text{CH}_2 - \text{CH}_3  \end{array}  $ Isoleucine (I)

Here are first nine amino acids from Table 13 with the less complex precursors. The rest of the other three amino acids (D, N, K) makes a connection with the last  $3 + 5 = 8$  amino acids with more complex precursors (cf. Table 13.).

**Table 15. First Three (Less Complex) Amino Acid Precursors**



*This Table corresponds with Tables 13 and 14. The connection of amino acids D-E, N-Q, K-R (Table 13) is followed by a connection between their precursors: the last of the less complex precursors, oxaloacetate (with one - CH<sub>2</sub> group), is the precursor for D, N, K. On the other hand, the first of the more complex precursors, 2-oxoglutarate (with two - CH<sub>2</sub> groups), is the precursor for E, Q, R.*

As we can see, the SESS amino acids (C and R) are included in this nucleon "balance game", as a whole quantum, together with the ESS amino acids. (Note: According to Van Nostrand's Scientific Encyclopedia, 1983,

p 119, the semiessential, or "quasi-essential", amino acids are R, H, Y, C; furthermore, on the 117 page: "Generally, those amino acids which the human body cannot synthesize... are called essential amino acids... the term nonessential is taken to mean those amino acids that are really synthesized in the body"; and it was also written about a semi-essential amino acid which "is essential for the normal growth of the human infant, but to date it is not regarded as essential for adults").

In one other and different nucleon number balance, all four SESS amino acids, also as a whole quantum, and also together with the ESS amino acids, appear within the set of all 20 protein amino acids (Solution 53) :

$$(8 \text{ ESS} + 4 \text{ SESS}) - 8 \text{ NESS} = 1 \times 555 \quad (53)$$

$$(570 + 335) - 350 = 1 \times 555$$

This specific balance, expressed through "the digital pattern" in which "the numbers are written using the same symbols" (Shcherbak, 1994, p 475) we recognize from the classification into 05 four-codon and 15 non-four-codon amino acids. Within these 15 (within their side chains, or within their "heads") there are exactly  $2 \times 555$  of nucleons (cf. Fig. 1 in: Shcherbak, 1994, p 475). On the other hand, this specific distinction: existence of  $(8 \text{ ESS} + 4 \text{ SESS})$  opposite to  $(8 \text{ NESS})$  amino acids we have in a real protein macromolecules evolution process. Namely, within the mutation ring of the amino acid code (Swanson, 1984, p 191; Rakoëeviæ, 1994, p 85), 8 amino acids left from the line S-Q, including those two on the line, are nonessential (S, A, G, P, E, D, N, Q); on the right of the line there are 8 essential (K, W, F, L, M, I, V, T) and 4 semiessential (H, R, Y, C) amino acids.

Contrary to these essentiality - nonessentiality nucleon number balances, there are atom number balances where all four SESS amino acids appear, also as a whole quantum, not together with the ESS, but with NESS amino acids (Solution 54) :

$$8 \text{ NESS} + 4 \text{ SESS} = 4 \text{ ESS (LTFW)} + 4 \text{ ESS (VMIK)} \quad (54)$$

$$54 + 48 = 53 + 49$$

Notice that first four ESS amino acids (LT-FW) one can read from the triplets which are distant (II, IV - VII, VIII) within the system in Survey 7'; on the other hand, second four ESS amino acids (VM-IK) one can read from the close triplets (III, IV - IV, V). Notice also "the minimum change relation" (Swanson, 1984, p 191) between two and two groups of amino acids:  $54 - 53 = 1$ ;  $49 - 48 = 1$ .

The pairs of amino acids within the Surveys 7', then 7'.1. and 7'.2, and finally within the Survey 8' are also the pairs from the aspect of existence the two classes of the aminoacyl-tRNA synthetases (20 of them for 20 amino acids). [Eriani et al., 1995: "Previous sequence analyses have suggested the existence of two distinct classes of aminoacyl-tRNA synthetase. The partition was established on the basis of exclusive sets of sequence motifs

(Eriani et al., 1990, *Nature* 347: 203-306). X-ray studies have now well defined the structural basis of the two classes: the class I enzymes share 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 of synthetases catalyze the same global reaction that is the attachment of an amino acid to the tRNA, but 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." ].

To understand the rules of the enzymes coding for amino acids within the genetic code, one must previously note the following. Within the inner subsystem in Survey 7'.1. must exist a further distinction into the inner subsystem I, ISS I ( S-T and C-M ), and into the inner subsystem II, ISS II (D-E, N-Q and K-R). For the ISS I, the hetero atoms ( O and S ) come from the same group in periodic system, but for the ISS II, the hetero atoms ( O and N ) come from the different groups in periodic system.

And now, the rules. Within the outer subsystem of alanine type in Survey 7'.1., the first class of enzymes is coding for amino acids that are - the second members in the pairs or doublets ( L, W, Y ); the second class of enzymes - for the first members of amino acids ( A, H, F ). Within ISS II the same rule is valid: first class of enzymes is coding for the second members ( E, Q, R ) and second class of enzymes is coding for the first members ( D, N, K ). Within ISS I, the first class of enzymes is coding for the second pair ( C-M ) and the second class of enzymes is coding for the

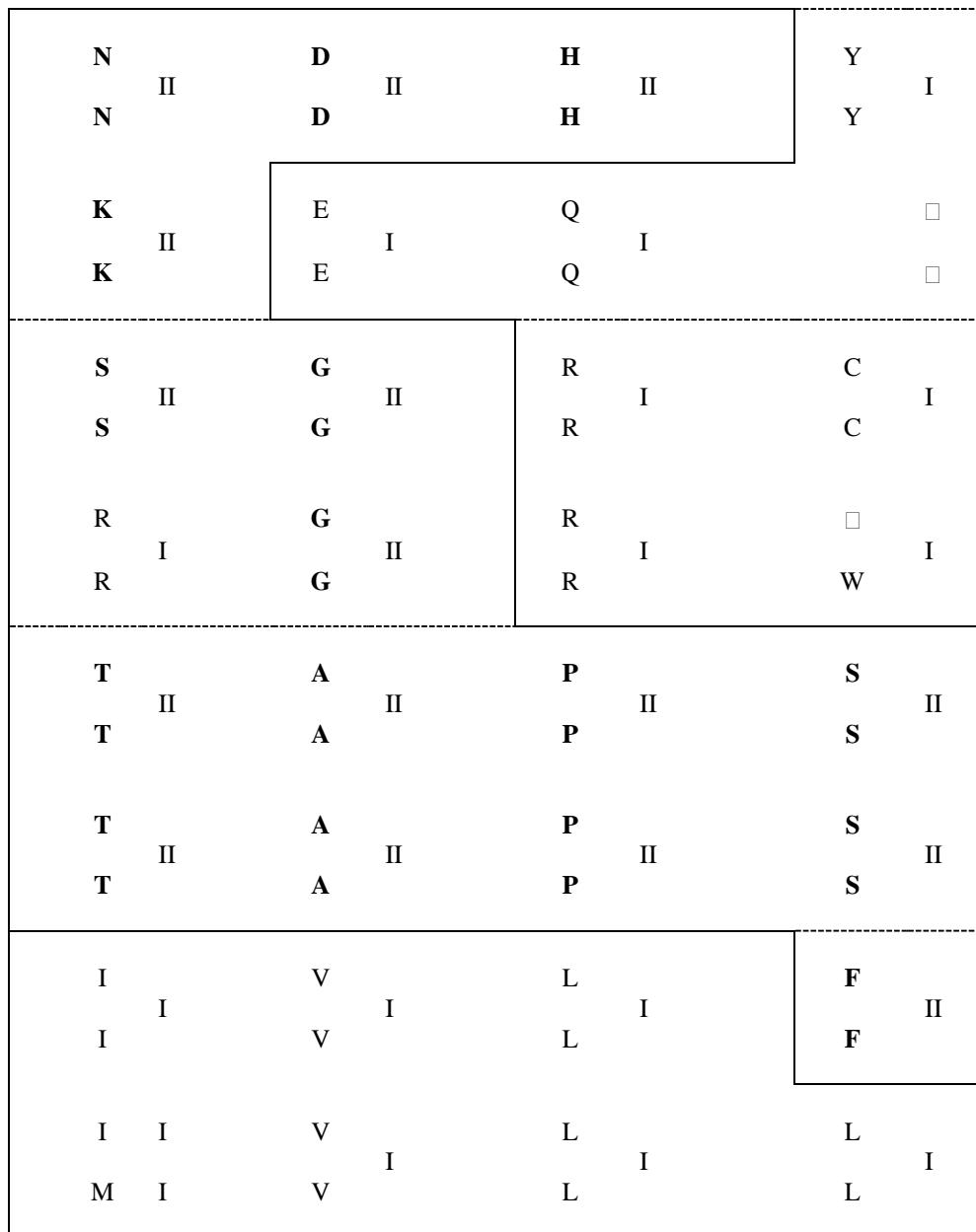
first pair ( S-T ) [Hint. From the three bioelements - the constituents of amino acids - in the second period ( C, N, O ) the most complex is the oxygen; its analogue, sulphur, in the third period, must be still more complex, and that is a reason why S-T as "amino acids handled by class II synthetases" ( Wetzel, 1995, p 545 ) must be less complex than C-M amino acids]. As a conclusion for the alanine type: the less complex amino acids (within the pairs, or from one to another pair) are handled by class II synthetases, and, on the other hand, the more complex amino acids are handled by class I synthetases.

From the aspect of IE-ED, the stereochemical type of alanine must be more complex than that of glycine (hydrocarbon side chain vs hydrogen side chain). Thus, if only one amino acid of the alanine type is handled by class II synthetases, then glycine must be also handled by class II synthetases. From the same aspect, valine type must be more complex than alanine stereochemical type ( a direct IE influence to  $C^\alpha$  atom vs indirect IE influence through a  $-CH_2$  atom group ). Thus, if only one amino acid of the alanine type is handled by class I synthetases, then both amino acids of the valine type ( V and I ) must be handled by class I synthetases. Finally, we have the relation between two stereochemical types which are generated from the simplest possible cyclic side chain ( cyclopropane ) - between the valine and proline type. The IE influence to the  $C^\alpha$  atom must be stronger with an open side chain ( V ) than with a non-open side chain ( P ). Thus, if all amino acids within the valine type are handled by class I synthetases, then all amino acids within proline type ( only one: proline ) must be handled by class II synthetases.

The understanding of all these relations between four stereochemical types of amino acids and two classes of enzymes which are coding for them, is a condition for a better understanding "the existence of a second genetic code, imprinted into the structure of aminoacyl-tRNA synthetases and matching the amino acids with the structural features of tRNAs that are recognized by these enzymes" ( De Duve, 1988, p 117 ), i.e. a condition for better understanding what is "an operational RNA code for amino acids" ( Hipps et al., 1995; Schimmel, 1995 ).

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

**Fig. 10.** The distinctions of the two classes of enzymes - aminoacyl - tRNA synthetases within the classical Watson-Crick's Table of the genetic code ( after : Wetzel, 1995).



**Fig. 11.** The distribution of amino acids within the "Codon path cube" (after: Swanson, 1984). This distribution is followed with a strict distinction of two classes (I and II) of the enzymes - aminoacyl - tRNA synthetases (for details see the text).

Knowing all these presented relations and, on the other hand, bearing in mind that "the genetic code is ... an example of a Gray code" ( Swanson, 1984, p 187 ), Gray code as "the Boolean space genetic code concept as an inevitable reality" ( Rakoèeviæ, 1994), we can have, finally, a key for all atom and nucleon balances within the genetic code. This key is the "Codon path cube" ( Figure 2 in: Swanson, 1984, p 189 ) in a strict correspondence with the LIGHT - Logical, Information, Geometric, Homeomorphic, Topological - model of the genetic code ( Rakoèeviæ, 1994, p 53 ). From the classical Watson-Crick Table of the genetic code it does not follow a full distinction of the aminoacyl-tRNA synthetases into two classes ( Fig.10 in this book, or Fig. 1 in Wetzel, 1995, p 546), but from the "Codon path cube" it follows ( *see* Fig. 11 in this book; cf. the three codon cubes: first in Fig.64 in Eigen and Schuster, 1979, or in Fig. 1.3. in Rakoèeviæ, 1994, p 12; second in Fig. 2 in Swanson, 1984, p 189, or in Fig. 1.4. in Rakoèeviæ, 1994, p 13; and third in Fig. 4.1. in Rakoèeviæ, 1994, p 54); a strict distinction, only with one crossing-over: glycine ( the simplest amino acid ) separates the codons coding for arginine ( the most complex amino acid ) into two classes. ( **Hint.** From the aspect of hydropathy, glycine is the simplest amino acid because it has a minimal negative value, HP = -0.4; arginine is the most complex, with the maximal negative value, HP = - 4.5 ).

Tables 16-17 and Solutions 55-59 show the atom and nucleon number balances within the "Codon path cube" ( Fig. 11) as follows .

The Table 16 shows atom number balances within Fig. 11. From the begining:  $M = (1 \times 11)$ ,  $I = (3 \times 13)$  etc. That are maximal possible values for the amino acids. On the other hand Table 17 shows the nucleon number balances through minimal possible values :  $M = (1 \times 75)$ ,  $I = (1 \times 57)$ ,  $V = (1 \times 43)$  etc. As we see in four cycles ( rows ) there are 4, 5, 6, 7 molecules, respectively ("minimum change relation").The total atom number (  $2 \times 297$  ), Solution 55, corresponds with the position 12d in NSM II through an inversion ( 297/792).The total nucleon number ( Solution 56 ) also, but through an another inversion (594/495 ; notice that 495 is the third perfect number 496 minus 1). At the same time the total nucleon number corresponds to the sum of first six frendly numbers ( exactly to 1/10 of it : the sum 1386 in relation to 3861 ). The Solution 57 shows how the sum of all atoms and all nucleons (  $594 + 1386 = 1980$  ) corresponds to the position 12d in NSM III ( Table 2 ).

$$\begin{array}{rcccl}
 (168 - 0) & + & (129 - 1) & = & 296 \\
 (168 + 0) & + & (129 + 1) & = & 298 \\
 \hline
 & & (1 \times \mathbf{594}) & \longleftrightarrow & (2 \times 297)
 \end{array} \tag{55}$$

$$\begin{array}{rcccl}
 (1 \times 232) & + & (1 \times 223) & = & 455 \\
 & & & & \mathbf{476} \\
 (1 \times 232) & + & (3 \times 233) & = & 931 \\
 \hline
 3861 & - & (5 \times \mathbf{495}) & \longleftrightarrow & 1386
 \end{array} \tag{56}$$

$$\begin{array}{rcccl}
 1386 & + & 594 & = & 1980 \\
 & & & & \mathbf{0891} \\
 495 & + & 594 & = & 1089 \\
 \hline
 \frac{1}{9} [38610 & - & (7992+2997)] & \longleftrightarrow & 3069
 \end{array} \tag{57}$$

$$\begin{array}{rcccl}
 (168 - 0) & - & (129 - 1) & = & 40 \\
 (168 + 0) & - & (129 + 1) & = & \underline{38} \\
 & & & & 78
 \end{array} \tag{55^{\sim}}$$

$$\begin{array}{rcccl}
 (1 \times 232) & - & (1 \times 223) & = & 009 \\
 (3 \times 233) & - & (1 \times 232) & = & \underline{467} \\
 & & & & 476
 \end{array} \tag{56^{\sim}}$$

$$78 + 476 = 322 + 232 \tag{58}$$

$$\begin{array}{rcccl}
 233 & & 233 & & 777 \\
 232 & & 323 & & 888 \\
 \hline
 322 & & 332 & & \underline{\frac{1}{4} 6660} \\
 \hline
 777 & & 888 & & 4
 \end{array} \tag{59}$$

**Table 16. Atom number balance within "Codon Path Cube".**

11	+	39	+	40	+	78	=	168	-	0
M	+	I	+	V	+	L	=	(4)		
32	+	16	+	32	+	20	+	28	=	129 - 1
T	+	A	+	P	+	S	+	F	=	(5)
30	+	22	+	20	+	68	+	10	+	18 = 168 + 0
Y	+	Q	+	E	+	R	+	C	+	W = (6)
16	+	14	+	22	+	30	+	10	+	04 + 34 = 129 + 1
N	+	D	+	H	+	K	+	S	+	G + R = (7)

Four rows in correspondence with the four groups of amino acids within the system in Fig. 11.

**Table 17. Nucleon Number Balance within "Codon Path Cube".**

75	+	57	+	43	+	57	=	(1 x 232)	+	000
M	+	I	+	V	+	L	=	( 4 )		
45	+	15	+	41	+	31	+	91	=	(1 x 223) - 000
T	+	A	+	P	+	S	+	F	=	( 5 )
107	+	72	+	73	+	100	+	47	+	130 = (1 x 232) + 297
Y	+	Q	+	E	+	R	+	C	+	W = ( 6 )
58	+	59	+	81	+	72	+	31	+	01 + 100 = (3 x 233) - 297
N	+	D	+	H	+	K	+	S	+	G + R = ( 7 )

Four rows in correspondence with the four groups of amino acids within the system in Fig. 11.

As we can see the sums of atoms and nucleons are determined with the permutations of numbers 223 and 233 ( Solution 56 ); once again through the fractal motive "to be 3 and 2 at the same time" ! And, not only through the sums, but also through the differences ( Solutions 52' and 53' and 55 ). In the final step there is also a correspondence to the positions 1c and 1d in NSM III ( Table 2; cf. Solution 56 ).

And now, after all this, one question and one answer. Why the "Codon path cube" is a key for all atom and nucleon number balances within genetic code?

Because, if "Codon path cube" exists, then the "numbers" within itself are not the numbers in classical sense, but they are Boolean, i.e LIGHT relations between the entities of the code.

## Appendix 9

### Four Stereochemical Types of Protein Amino Acids in Correspondence with Two Classes of Enzymes

Wetzel (1995) has shown that two classes of enzymes aminoacyl-tRNA synthetases are distributed within the codon - amino acid matrix of the genetic code with a strict distinction only for the XCX and XUX columns, but not for the XAX and XGX columns. From this follows that a strict regularity exists not.

Survey 15 (Atom number)			Survey 16 (Nucleon number)		
I	II	I	II		II
G 01	01 G	G 01	01 G		
S 05	08 T	S 31	45 T		
P 08	08 P	P 41	41 P		
III	IV	III	IV		
C 05	11 M	V 43	57 I		
V 10	13 I	V 43	57 I		
V 10	13 I	C 47	75 M		
V	VI	V	VI		
A 04	13 L	A 15	57 L		
H 11	18 W	H 81	130 W		
F 14	15 Y	F 91	107 Y		
VII	VII	VII	VII		
D 07	10 E	N 58	72 Q		
N 08	11 Q	D 59	73 E		
K 15	17 R	K 72	100 R		

Atom number (on the left) and nucleon number (on the right) within 12 doublets and 8 triplets of "24" amino acids. **Bold** positions: amino acids handled by class I and other positions by class II synthetases.

(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). In this study, however, I show that a such (strict) regularity exists, but expressed through division-distinction of protein amino acids into four stereochemical types. Namely, if we know that protein amino acids are arranged hierarchically as a specific stereochemical doublet-triplet amino acid system, DTAAS (Surveys 7' and 8'), then a general rule is very simple and self-evident: *the first* amino acids are handled by class II (more specific enzymes), whereas *the second* amino acids are handled by class I (less specific enzymes) synthetases (Surveys 15 & 16) (Eriani et al., 1995, p 499: "the class I enzymes share 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 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 more strong nucleophile whereas the C atom on the position 3' is a less strong nucleophile.).

For better understanding the explanations, which follow, I give previously some additional notes. Within two classes of alanine stereochemical type there are two and two subclasses (cf. Survey 7'.1, p 102): I<sub>1</sub>. S-T and C-M with hetero atom O or S within side chain, seeing together, both from different period and the same group in Mendeleev's periodic system; I<sub>2</sub>. D-E, N-Q and K-R with hetero atom O and/or N, seeing together, both from different groups and the same period; II<sub>1</sub>. A-L with hydrocarbon non-cyclic chain, without hetero atom; II<sub>2</sub>. H-W and F-Y with hydrocarbon (cyclic) aromatic side chain. Notice "the minimum change relation among the amino acids" (Swanson, 1984, p 191): in second class 1-2 and in first 2-3 amino acid pairs.

And now the explanations. The correspondence amino acid - enzyme appears in two logics: for ones six amino acid pairs, the first three pairs are "the first" amino acids, and the second three pairs are "the second" amino acids; for other six amino acid pairs, six first members within the pairs are "the first", and six second members are "the second" amino acids (see Surveys 15 & 16).

So, within the system of three non-alanine stereochemical types (cf. Survey 7'.2, p 103), the first pair from above G-G and from below P-P, both are handled by class II synthetases; on the other hand, the second pair from above V-I and from below V-I, "both" are handled by class I synthetases. Within the system of alanine stereochemical type, the first pair of I<sub>1</sub> subclass (S-T) is handled by class II synthetases whereas the second pair (C-M) by class I synthetases.

For the pairs within I<sub>2</sub>, II<sub>1</sub> and II<sub>2</sub> subclasses is valid the said different other logic: all first members within all six pairs are handled by class II, whereas all second members by class I synthetases.

Surveys 15 & 16 are providing proofs for given discussion. They contain a DTAAS corresponding to the classification of enzymes aminoacyl-tRNA synthetases into two classes: (less specific and more specific enzymes), ECO-DTAAS. If so, then the DTAAS within Survey 1 in our previous work (Rakoeviæ & Jokiæ, 1996, or Survey 7', p 102 in this study) is a PECO-DTAAS, i.e. precursor corresponding and enzym corresponding DTAAS at the same time.

These two systems, PECO-DTAAS and ECO-DTAAS, taken together, are providing proofs for the hypothesis of the existence of an optimal unique system: precursors - amino acids - enzymes system (PRACENZ system), arranged through a maximum possible balance, and symmetry expressed through chemical characteristics, atom and nucleon number.

By this one must notice that atom number within 12 doublets and 8 triplets of "24" amino acids, in both systems (ECOS and PECOS) is the same on both lines: on the full line, as well as on the dotted line, there are 118 atoms (cf. Survey 7', p 102 and Survey 15, p 118). The nucleon number on the full line in Survey 16 is 694 + 0 and on the dotted line 694 + 9 (one cycle more in module 9), whereas in Survey 8' we have a vice versa situation: 694 + 9 on the full and 694 + 0 on the dotted line, what means the realization of a maximum possible symmetry.

In conclusion, we can say that all these regularities, presented not only in these Appendix, but also in the whole book, actualize again an old question: to which one of the theories should be given the advantage during the interpretations of the origin of the amino acid (genetic) code, "the stereochemical theory", or the frozen accident theory" (Crick, 1968). In other words, these regularities provide "a new standpoint for addressing questions of selection vs. random drift in the evolution of the code".

## **ADDENDA**

## 1. Excerpt from Reference (Moore, 1993)

# THE LIMIT OF THE GOLDEN NUMBERS $1/2$

**Gregory A. Moore**

West Shore Community College, Scottville, MI 49454-0277

gamoore@aol.com

(Submitted September 1992)

## 1. INTRODUCTION

The “Golden polynomials”  $\{G_n(x)\}$  (defined in [2]) are Fibonacci polynomials satisfying

$$G_{n+2}(x) = x G_{n+1}(x) + G_n(x)$$

for  $n \geq 0$ , where  $G_0(x) = -1$  and  $G_1(x) = x - 1$ . The maximal real root,  $g_n$ , of the function  $G_n(x)$ , can be considered to an  $n^{\text{th}}$  - dimensional golden ratio.

Our concern here is the study of the sequence  $\{g_n\}$  of “golden numbers”. A computer analysis of this sequence of roots indicated that the odd-indexed subsequence of  $\{g_n\}$  was monotonically increasing and convergent to  $3/2$  from below, while the even-indexed subsequence was monotonically decreasing and convergent to  $3/2$  from above.

In this paper, the implications of the computer analysis are proven correct. In the process, a number of lesser computational results are also developed. For example, the derivative of  $G'_n(x)$  is bounded below by the Fibonacci number  $F_{n+1}$  on the interval  $[3/2, \infty)$ ...

## 2. Excerpt from Reference (Shcherbak, 1994)

## Sixty-four Triplets and 20 Canonical Amino Acids of the Genetic Code: the Arithmetical Regularities. Part II

Rumer naturally divided the genetic code triplets into two equal groups: the first one included the complete synonymous series of the same 5' - doublets (i.e. degeneracy IV), the other (quasi-group) a split series (i.e. degeneracy III, II, I). This approach revealed rigorous mathematical rules which govern the degeneracy and composition of triplets in the series (Rumer, 1966; Konopelchenko & Rumer, 1975; Shcherbak, 1989, 1993a)...

The idea of denoting numbers by symbols, assigning to them meanings as to both the form and position, appealed to Laplace as being elegant in its simplicity. However, Nature seems to have left Man behind in using similar formalism. How and why this happened is still to be discovered.

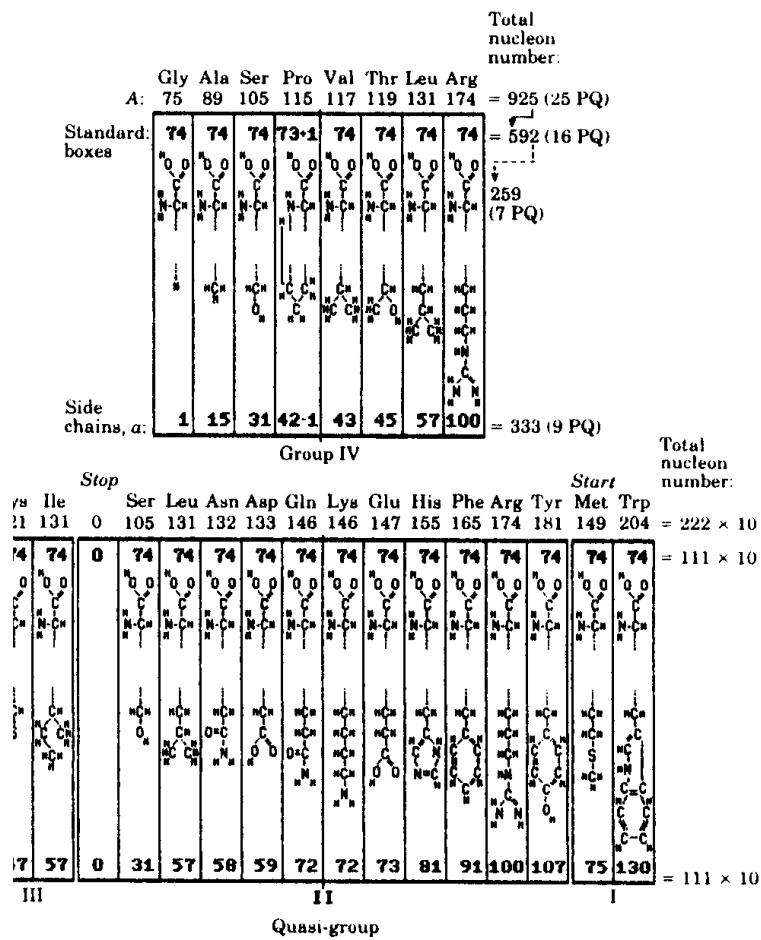


FIG. 1. Arithmetical regularities for the barionic numbers of the free amino acid molecules divided into group IV and quasi-group III-II-I (Shcherbak, 1993b). The sums of nucleons in amino acid side chains are multiples of the Prime Quantum (PQ) 037. Group IV. In the case of Pro, the formal borrowing of one nucleon ensures the sums being multiples of the PQ both for the standard boxes and the side chains. In the PQ dimensions, the sums demonstrate the squares of the first three Pythagorean numbers. Quasi-group III-II-I. The sums of nucleons in the standard boxes and the side chains demonstrate a precision balance. The symbol  $A$  denotes the total nucleon number in the atomic nuclei of neutral molecules of amino acids, the symbol  $a$  denotes the total nucleon number in their side chains. The punctuation triplets *Stop* are assigned a zero nucleon parameter as a formal symbol of the absence (normally) of the corresponding aminoacyl-tRNA.

### 3. The whole text from Reference (Koruga, 1995)

## INFORMATION PHYSICS: IN SEARCH OF A SCIENTIFIC BASIS OF CONSCIOUSNESS

Djuro Koruga

Molecular Machines Research Center, University of Belgrade  
Belgrade, Serbia, Yugoslavia

&

Advanced Biotechnology Lab, Department of Anesthesiology  
University of Arizona, Tucson, USA

**Abstract.** In this paper we consider a scientific interdisciplinary approach to consciousness. Molecular computing, both quantum and classical, has been used as a link between physics and biology. However, connections between biology and consciousness, and consciousness and physics can be explained through a new physical theory - *information physics*. The information physics we propose encounters a synergetic theory of classical mechanics, quantum mechanics and theory of information. Using information physics, as a new scientific paradigm, we have found that microtubules, clathrins and water clusters in living cells are major bimolecular devices which satisfy synergy principles of classical mechanics, quantum mechanics and information theory. Based on information physics we propose that human sub-consciousness has its beginnings in microtubule-water clusters interaction inside of the tube. Through a hierarchy of cytoskeleton network of cell and neural networks activities on synapses becomes consciousness, while in the brain synergetic activities result in self-consciousness. Our model of human consciousness can be tested on both levels molecular, as a general anesthesia experiment, and brain, through mapping conscious contents of the brain into artificial devices. This device has to be self-assembled on the same principles of information physics, which we have used to explain human consciousness.

**Key words:** quantum mechanics, chaos, fractals, Cantor random set, golden mean, information physics, microtubules, clathrin, water, sub-consciousness, consciousness, self-consciousness

## 1. INTRODUCTION

Consciousness is a common thing but is too mysterious to understand indeed. This is a longlasting paradox and calls for scientific investigation. As Einstein said, "The most beautiful thing we can experience is the mysterious. It is the source of all true art and science". It can be said, according to Einstein, that scientific research without mysterious note may result in only local scientific truth. However, results based on local truths are useful for everyday life. For example, we use statement like "... the sun arises in the east and sets in the west", because this "fact" determines our everyday behavior and experience. One of the main reason for this kind of vision is, that we stay on the Earth without sensing that it rotates. Similarly, most of our scientific truths are local ones because we do not include in our approaches consideration of consciousness. Quantum mechanics is one of the first mysterious scientific discipline which covers this problem as a human being-experiment interaction. But in quantum theory, which has more faces, there are (Dirac approach) at least four troubles ("ghosts"): gauge invariance, the fine structure constant, the singularities and the negative energies. For this reason, the relation between local scientific truths and quantum mechanics is similar as the muses position in mysterious ancient mythical world: "We know how to speak of many a lie so akin to truth, but when we so wish, we do know how to speak up the very truth". Local scientific truths look like "many a lie so akin to truth", while quantum mechanics looks like "when we so wish, we do know how to speak up the very truth". In goal to search quantum mechanics' " so wish to speak up the very truth", we propose to overcome some confusion about quantum theory and use its ontological rather than epistemological interpretation. This allows us to propose information physics as a new physical theory.

## 2. FUNDAMENTALS OF INFORMATION PHYSICS

### 2.1 Information theories and physics

In information theories coding is one of the crucial points to define information. Shannon's famous theorem gives us one possible solution which has been exploited in human-made information systems with great success. The basic idea is simple: given a data source which emits letters  $L_1, L_2, L_3, \dots, L_n$  with probabilities  $p_1, p_2, p_3, \dots, p_n$  respectively, each letter emitted being chosen independently of all other symbols, there exists a binary code which gives the best data compression. Bearing in mind that data compression is one of the fundamental questions in communication theory, it is understandable why Shannon's noiseless coding theorem has been so popular in classical communications engineering approach. However, success in the application of this technical approach to the theory of information on biological systems was partial. One of the main reasons for this is quite different biological solutions: self-organization and self-control of living matter versus bulk matter and system controlling from outside.

According to Shannon's coding theorem we can always reliably distinguish between different letters (or symbols). Different symbols mean different states of a physical carrier of information. If the physical carrier of information is macroscopic, Shannons theorem works well, while for quantum mechanical systems it needs to be reformulated. Why do quantum systems make a difference? In quantum systems the physical bit is any quantum object for which the state is described by a vector in two-dimensional Hilbert space. For these systems quantum mechanics states that only states represented by mutually perpendicular vectors can be perfectly distinguished from each other. If symbols-states are non-perpendicular, which is usually the case, any deduction procedure is imperfect. For the link between classical mechanics and quantum mechanics the important question is: which  $N$ -dimensional Euclidian space of coding will give  $90^\circ$  angle between the vector pointed from the origin  $(0, 0, \dots, 0)$  to the point  $(1, 1, \dots, 1)$  - and any coordinate axis? If that one Euclidian space (macroscopic) of coding exist, then correct mapping of information contents between it and quantum systems can exist. In the search for this kind of

Euclidian space, we will consider the problem of unit spheres as the most perfect symmetrical object.

## 2.2 Unit spheres

To define N-dimensional Euclidean space it is necessary to use Pythagorean distance as

$$X_1^2 + X_2^2 + X_3^2 + \dots + X_n^2 = r^2, \quad (1)$$

which is equal to the definition a sphere of radius "r" by this expression. It is well known that the formula for the volume of a sphere is

$$V_n(r) = C_n r^n, \quad (2)$$

where  $C_n$  is a constant (unit sphere):

$$C_n = (2\pi/n)C_{n-2}. \quad (3)$$

To find the values of  $C_n$  it is necessary to use a gamma function and its integral in polar coordinates [1]. Calculations based on Eq. (3) and symmetry gives results [2,3] which are summarized in Table 1.

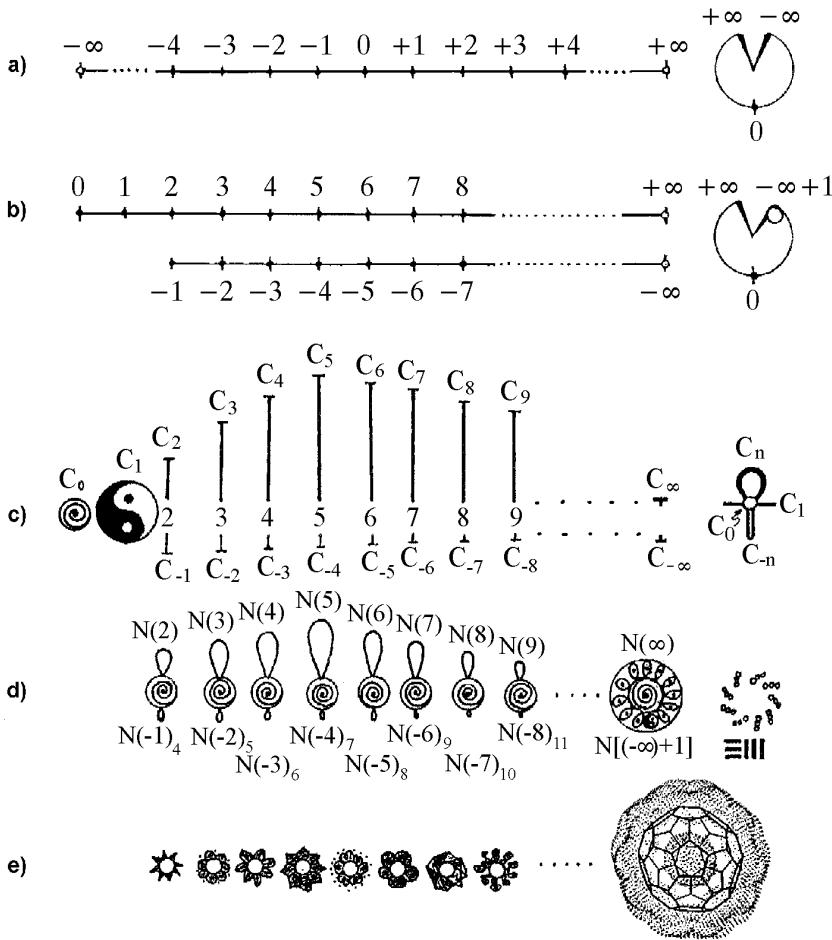
From Table 1 we can see: (1) the volume of the unit sphere comes to a maximum at  $N = n = 5$  and falls off to zero for  $N = n = \infty$ , and for  $N = -\infty$ , (2) there is one-to-one correspondence between dimension (N) and dimensionality (n) for positive "N", while for negative one it is not the case, (3) in spite that negative dimensions (-N) exist there is only positive dimensionality (n), (4) negative dimensions (-N) are inversion of positive dimensions (+N) through dimension  $N = 0$ , except for  $N = 1$ , (5) dimensionality of  $N = 0$  is  $n = 3/2$  (is not integer), and (6) there is a one-to-one correspondence between  $N(2)$ ,  $N(3)$ , ...  $N(m)$  and  $N(-1)$ ,  $N(-2)$ , ...  $N(-\{m-1\})$  in  $N = 0$ , respectively. This gives result (Fig.1) that each pair, including two infinities  $+\infty$  and  $-\infty$ , has their unification in  $N = 0$  (note that for  $N = 0$ ,  $V_0 = C_0 = 1 = 0!$  zero factorial).

**Table 1** Information complementary for systems of all dimensions, based on symmetry and unit spheres.

$N = 6$	$C_6 = \frac{2\pi}{n} C_4 = \frac{2\pi}{6} \cdot \frac{\pi^2}{2} = \frac{\pi^3}{6} = 5.1677$
$N = 5$	$C_5 = \frac{2\pi}{n} C_3 = \frac{2\pi}{5} \cdot \frac{4\pi^2}{3} = \frac{8\pi^3}{15} = 5.2637$
$N = 4$	$C_4 = \frac{2\pi}{n} C_2 = \frac{2\pi}{4} \cdot \pi = \frac{\pi^2}{2} = 4.9348$
$N = 3$	$C_3 = \frac{2\pi}{n} C_1 = \frac{2\pi}{3} \cdot 2 = \frac{4\pi}{3} = 4.1887$
$N = 2$	$C_2 = \frac{2\pi}{n} C_0 = \frac{2\pi}{2} \cdot 1 = \pi = 3.1415$
$N = 1$	$C_1 = \frac{2\pi}{n} C_{-1} = \frac{2\pi}{1} \cdot \frac{1}{\pi} = 2 = 2$
$N = 0$	$C_0 = \frac{2\pi}{n} C_{-2} = \frac{2\pi}{\frac{3}{2}} \cdot \frac{1}{\frac{4\pi}{3}} = 1 = 1$
$N = \bar{1}$	$C_{-1} = \frac{2\pi}{n} C_{-3} = \frac{2\pi}{4} \cdot \frac{1}{\frac{\pi^2}{2}} = \frac{1}{\pi} = 0.3184$
$N = \bar{2}$	$C_{-2} = \frac{2\pi}{n} C_{-4} = \frac{2\pi}{5} \cdot \frac{1}{\frac{8\pi^2}{15}} = \frac{1}{\frac{4\pi}{3}} = 0.2387$
$N = \bar{3}$	$C_{-3} = \frac{2\pi}{n} C_{-5} = \frac{2\pi}{6} \cdot \frac{1}{\frac{\pi^3}{6}} = \frac{1}{\frac{\pi^2}{2}} = 0.203$
$N = \bar{4}$	$C_{-4} = \frac{2\pi}{n} C_{-6} = \frac{2\pi}{7} \cdot \frac{1}{\frac{16\pi^3}{105}} = \frac{1}{\frac{8\pi^2}{15}} = 0.1899$

If we consider the shell fraction of the volume of an  $N$ -dimensional sphere that is within a distance  $[d^*]$  from the surface we can write:

$$\frac{\text{Shell}}{\text{Volume}} = \frac{C_n r^n - C_n (r - d^*)^n}{C_n r^n}. \quad (4)$$



**Figure 1** Cage networks as knot shells: a) Pure mathematical representation of positive and negative numbers and symbolic representation of infinities ( $+\infty$  and  $-\infty$ ); b) Positive and negative dimensions representation (Table 1). One member associates more with negative dimensions than positive; c) Unit sphere at zero ( $C_0$ ) is present as Golden Mean spiral, while unit sphere of a one-dimension is present as two complementary units themselves. Other unit spheres of dimension  $N > 1$  and  $N < 0$  are complements through dimension  $N = 0$ . Maximal value of unit sphere is for  $N = 5$ . Symbolic representation of unit spheres of all dimensions is present too (right); d) Because dimensions  $N > 1$  and  $N < 0$  are complements in  $N = 0$ , their both point of departure and the end is in  $N = 0$  (knots 0!); e) Possible representation of information objects as cage networks.

This is so for no matter how thin a shell we wish to use, and how close we wish to get to \*value one. Now, if we examine the angle between the vector pointed from the origin  $(0, 0, 0, \dots, 0)$  to the point  $(1, 1, 1, \dots, 1)$  - and any coordinate axis, we will find that the projection on each axis is exactly equal to one. The length of the vector is  $\sqrt{N}$ , so the angle between diagonal line and each coordinate axis

$$\cos\alpha = \frac{1}{\sqrt{N}}. \quad (5)$$

As we can see only for sufficiently large  $N$  ( $N \rightarrow \infty$ ) angle  $\alpha$  is  $\pi/2$ , or the diagonal line is perpendicular to each coordinate axis.

According to results from Table 1. systems with  $N > 1$  can not exist independently, each of them has correspondent pair in  $N(-1)$ ,  $N(-2)$ , ..., so that

$$R_w^N S_w^{(1-N)} = N(0), \quad (6)$$

where:  $R_w$  is the real world (macro space-time:  $N(2)$ ,  $N(3)$ , ...  $N(m)$ ), and  $S_w$  we named the world of shadow (micro space-time:  $N(-1)$ ,  $N(-2)$ ,  $N(-3)$ , ...  $N(-[m-1])$ ). So,  $N = \infty$  is orthogonal in  $N = 0$ , and may provide link between classical mechanics and quantum mechanics through coding in  $N = 0$  as knot shall (cage network with value and properties of 0!).

## 2.3 Unit spheres packing and coding

Unit sphere packing in  $N$ -dimensional space is equal to coding and to digital transmission of information [4-6], where face-centered-cubic symmetry, as the symmetry group of the unit spheres packing, was used.

Results based on this research have shown that when packing is constructed from codes for digital transmission of information, dimensions  $N = 11$ ,  $N = 12$  and  $N = 13$  are optimal for lattice packing, while dimensions  $N = 10$  and  $N = 13$  are optimal for non-lattice packing.

These results indicate that one possible information theory may exist based on the unit spheres packing. How unit spheres of positive and negative dimensions correlate to quantum mechanics is an open question, which we will consider using fractal space-time and renormalization quantum field theory approaches.

## 2.4 Information Physics

It was indicated that trajectory behavior of quantum mechanics objects is characterized by a fractal [7]. A "thought experiment" was done, in which objects are confined to move on fractal space-time trajectories, treating the case of a Peano-Moore trajectory in detail [8]. Also, it is known that chaos may generate stochastic and fractal behavior [9]. A connection between chaos and quantum, as "quantum chaos", was noted by research groups independently [10-13]. A disadvantage to these attempts was that quantum stochasticity had not been included in the quantum chaos theory. One approach to include the quantum stochasticity is called "chaos quantum" [9]. A major problem with chaos quantum has been its place in the theory of stochastic mechanisms of an "objective" background radiation ("noise") which have pervaded the Universe, giving every object of mass  $m$  a diffusion-type perturbation of intensity  $\hbar/m$ , where  $\hbar$  is Plank's constant divided by  $2\pi$ . This means that if one can smoothly transit fractal space-time (with Hausdorff dimension) to Minkowskian space-time (with Euclidean dimension) with random perturbation of source of mapping, it will be possible to unify classical mechanics and quantum mechanics. Our approach to solve this problem is through unification dimensions  $N = 0$  and  $N = 1$ , what we named information physics.

To unify  $N = 0$  and  $N = 1$ , an invariant set with the following properties has to exist: (1) the measure is zero; (2) the value is one; (3) the value of the transformation  $T(x)$  map of  $R_w$  (macro world) into  $S_w$  (micro world), and vice versa, has to be  $3/2$  (this is because for dimension  $N = 0$  dimensionality is  $n = 3/2$ , Table 1.). If we use Cantor's random middle third set with  $T(x) = (3/2)(1 - |2x - 1|)$  all our conditions are satisfied [14,15].

One of the most beautiful and sublime results in the last decade in mathematics is the Mauldin-Williams theorem, which shows that the Hausdorff dimension  $D_H^{C(0)}$  of a randomly Cantor middle third set for  $N = 0$  is  $D_H^{C(0)} = GM^-$  or  $\phi$ , where  $GM^- = (\sqrt{5} - 1)/2$  is the Golden Mean. If we extrapolate the random construction of the Cantor set to  $N$  dimensions with  $GM$  properties than the Hausdorff dimensions [16]:

$$D_H^{C(N)} = \left( 1/D_H^{C(0)} \right)^{N-1}. \quad (7)$$

Based on Eq. (7), solution for  $D_H^{C(4)} = 4 + D_H^{C(-2)}$  is  $(1/GM)^3$  or  $(GM)^{-3}$ . In other words we can write Eq. (7) in form

$$D_H^{C(N)} = \left( D_H^{C(0)} \right)^{1-N}, \quad (8)$$

which is same as Eq. (6). If we summarize our calculation based on Eq. (8), Table 2, we see that the results are same as in Table 1.

We found that one of the main properties of dimension  $N = 0$  is the Golden Mean (GM) based on random Cantor set. One more interesting question arises: what does  $n = 3/2$  of  $N = 0$  represent? To answer this question we need to consider our system, in Table 1, from quantum mechanical point of view. It is well known from a quantum field theory [17] that the dimension of mass ( $D_m$ ) is calculated from the expression:

$$D_m = (D/2) - 1, \quad (9)$$

where  $D$  is a space-time dimension value. Only for  $D = 5$  (whose unit sphere  $C_5$  is maximal in system) value of  $D_m$  is  $3/2$ , what indicates: (1) mass is manifestation of  $N = 0$  in 5-D space structure; (2) we see our world as 3-D (in spite of it being 5-D), because we are space-time entities of  $N(-2)$  properties ( $N(3) \cdot N(-2) = N(0)$ ); (3) solutions of fundamental questions, including the question about consciousness, may exist ( $N(5)$  and  $N(-2)$  have similar 5-D space structure, and consciousness may explain the Universe and Itself), and (4)  $N = 0$  is the Nothing, which, through 5-D space and  $[-4]_7$ , as monad of mass, becomes Everything.

**Table 2** Values of Hansdorff dimensions of a randomly Cantor middle third set for different dimension  $N$ , calculated from Eqs. (7) or (8) ( $\phi$  - means Golden Mean)

$N$	$D_H^{C(N)}$	$[1/D_H^{C(0)}]^{N-1}$
0	$\phi^0$	$\phi$
1	$\phi^1$	1
2	$\phi^2$	$\phi^{-1}$
3	$\phi^3$	$\phi^{-2}$
4	$\phi^4$	$\phi^{-3}$
5	$\phi^5$	$\phi^{-4}$
$\vdots$	$\vdots$	$\vdots$
$\infty$	$\phi^\infty$	0

### **3. NEUROMOLECULAR COMPUTING**

From Table 1 we see that dimension  $N = 1$  looks like a "ghost" in this system. This one-dimensional entity ("string"), with unit sphere equal 2, is a dimension independent from all others, because there is no complement in the system through  $N = 0$ . According to this model its complement may be only  $N = 0$ , and/or through  $N = 0$ , all other dimensions.

#### **3.1 DNA as one-dimensional information entity**

DNA is composed of the so called nucleotides. One nucleotide is composed of three elements: a base, ribose and a phosphate group. Four types of bases may be represented: adenine, thymine, guanine and cytosine. Nucleotides are interconnected by hydrogen bonds in a specific double-helix structure [18]. From the aspect of organization of structure one such double-helix has the so-called aperiodic crystal [19]. DNA is one solution of 1-dimensional ( $N = 1$ ) crystallization. The term "aperiodic" signifies the irregular interchange of bases inside the helix while the phosphates and ribose are located on the outside making up a periodic structure. The "irregular" repetition of the bases within the helix represent properties of living beings, which has a meaning, from the information point of view, only as a code system. The genetic code from the aspect of chemistry is based on a triplet and that in variation of four bases gives of total of  $4^3 = 64$  possible codons for coding 20 amino acids.

#### **3.2 Proteins as other side of the DNA code**

The biochemical mechanism of protein synthesis is well-known. Messenger RNA (mRNA) is synthesized from one end of the DNA double helix, while the other end of the helix remains in the nucleus making possible the synthesis of another chain of DNA. The complete genetic

information is preserved and remains inside the nucleus. From mRNA through carrier RNA (tRNA) to ribosomal RNA(rRNA) there is a continual transmission of the genetic information message, making in effect proteins, the other side of the genetic code. Amino acids of proteins are organized in a chain as a 1-dimensional (1-D) "knot" entity giving 3-D structure. There are thousands of different proteins in a cell. We will here consider only two: tubulin and clathrin.

### *3.2.1 DNA 1-D replacement: From tubulin through microtubules to centrioles*

Tubulin is a type of globular (spherical) protein with about 450 amino acids. There exist  $\alpha$ ,  $\beta$  and  $\gamma$  subunits, but only  $\alpha$  and  $\beta$  make a  $\alpha-\beta$  heterodimer. Two subunits are able to bond as  $\alpha-\beta$  heterodimers with the aid of the strong GTP [20].

Tubulin subunits make one new type of organizational structure, microtubules. Observing microtubules in a slice cut width wise they usually (about 85%) consist of 13 subunits: however, under the microscope it is possible to see numbers of subunits varying from 7 to 17. Subunits possess electric dipole moments, and the Curie symmetry ( $\infty \infty /m$ ) for ideal spheres [21]. Since experimental results link tubulin and microtubules to bioinformation processes such as memory and learning [22,23], microtubules have become the subject of intensive research as bioinformation devices [24-26]. We have found that the microtubules possess two code systems,  $K_1(13,2^6,5)$  and  $K_2(24,3^4,13)$ , which may provide communication inside and outside of microtubules [2,25]. This self-assembled 3-D cylindrical structures, similar to DNA, are solutions of 1-D crystallization. Tubulin subunits are arranged in a cylinder with golden mean properties [3]. The lattice of tubulin subunits is a pattern with divergence  $(GM^+)^{-2}$ , where  $GM^+ = (\sqrt{5} + 1)/2$ . This structure forms a network in the cell which is responsible for intracellular transport, addressing, cell shape, growth form and many other dynamic activities [27,28].

Cell structure is organized from a central focal region near the nucleus called the microtubule organizing center (MTOC). The principle component of this center is the centriole, an organelle which consists of two perpendicular cylinders. Each of these cylinders, about 400 nm in length, is made up of nine microtubule triplets. The triplets are formed of one complete microtubule, with 13 protofilaments, and second and third partial microtubules with 10 protofilaments. Centrioles and MTOC play key roles in dynamic coordination of cell cytoplasm and its activities.

The centriole remains a central enigma in the cell biology and molecular biology. This enigmatic characterization may be resolved by considering centrioles as double Golden Mean devices: first through microtubules (divergence  $(GM^+)^{-2}$ ) and second through microtubule nine-fold symmetry triplets. These triplets may have both left and right orientation with the golden mean angle [3].

### 3.2.2 *Clathrin: DNA inverting itself into 0-D shell*

Clathrin is the major component of coated vesicles, important organelles for intracellular material transfer including synaptic neurotransmitter release. Based on molecular weights, isoelectric points and antigenetic determinants,  $\alpha$  and  $\beta$  tubulin subunits have been found to be associated with coated vesicles in both bovine brain and chicken liver [29]. However, there is evidence that synaptic vesicles are closely associated with microtubules, about five vesicles being radially disposed around a microtubule [27].

Clathrin is a Fullerene-like protein with a truncated icosahedron symmetry, as an object with 12 pentagons and 20 hexagons. Also, this protein may exist with 12 pentagons and as any number of hexagons. However, its process of self-assembly is by 0-D symmetry, as a process of crystallization around the point. This gives to clathrin a form of a shell (cage network); DNA so inverts part of its code into a protein with 3-D shell structure. The inside space of coated vesicles (shell) may be occupied or empty; experimentally, both situations have been observed [30].

It was shown that computing via self-assembly is possible, and the general self-assembly model of macromolecules based on quantum molecular computing has been proposed [31]. According to this model the free energy minimization of molecular computing may be used as a link between physics and biology, from an information point of view.

### 3.2.3 *Magic water clusters: $[\text{H}_2\text{O}]_{20}$ and $[\text{H}_3\text{O}]^+@[\text{H}_2\text{O}]_{20}$*

Although water has an overall neutral charge, the charges are asymmetrically distributed in space, which makes the molecule polar. The oxygen nucleus draws electrons away from the hydrogen nuclei, leaving these nuclei with a small net positive charge. On the other hand, the excess of electron density on the oxygen atom creates weakly negative regions. Because they are polarized, two or more adjacent water molecules can form a linkage known as a hydrogen bond. Molecules of water join together in a short-lived hydrogen-bonded lattice, cluster. Water itself has a slight tendency to ionize and therefore can act both as a weak acid and as a weak base. When it acts as an acid, it releases a proton to form a hydroxyl ion ( $[\text{OH}]^-$ ). When it acts as a base, it accepts a proton to form a hydronium ion ( $[\text{H}_3\text{O}]^+$ ).

The majority of water molecules identified from the X-ray electron density maps are individually bound to protein surfaces. Less frequently, but still in significant numbers, five-membered water rings and two-water five-fold symmetry clusters (hydrogen-bonds cage network) are found hydrogen-bonded to the protein surface [32,33].

The  $[\text{H}_3\text{O}]^+@[\text{H}_2\text{O}]_{20}$  or  $[+@*20]$  cluster has been detected under numerous experimental conditions [34-36]. This cluster is very stable due to the strong Coulombic interaction between the incaged  $[\text{H}_3\text{O}]^+$  ion and the surrounding 20 water molecules. This cluster has six local minima as its conformational state.

A basic cluster has a global minima  $[\text{H}_2\text{O}]_{20}$  or  $[*20]$  with pentagonal dodecahedron symmetry (one of the five Platonic solids). This yield an association energy of formation of the  $[*20]$  from 20 separated water molecules of about 117 kcal/mol. This cluster (dodecahedron) is an empty

cage with same elements of the symmetry and irreducible representations as icosahedron. ( $I_h$  symmetry group).

## 4. MODELS FOR BIOLOGICAL BASIS OF CONSCIOUSNESS

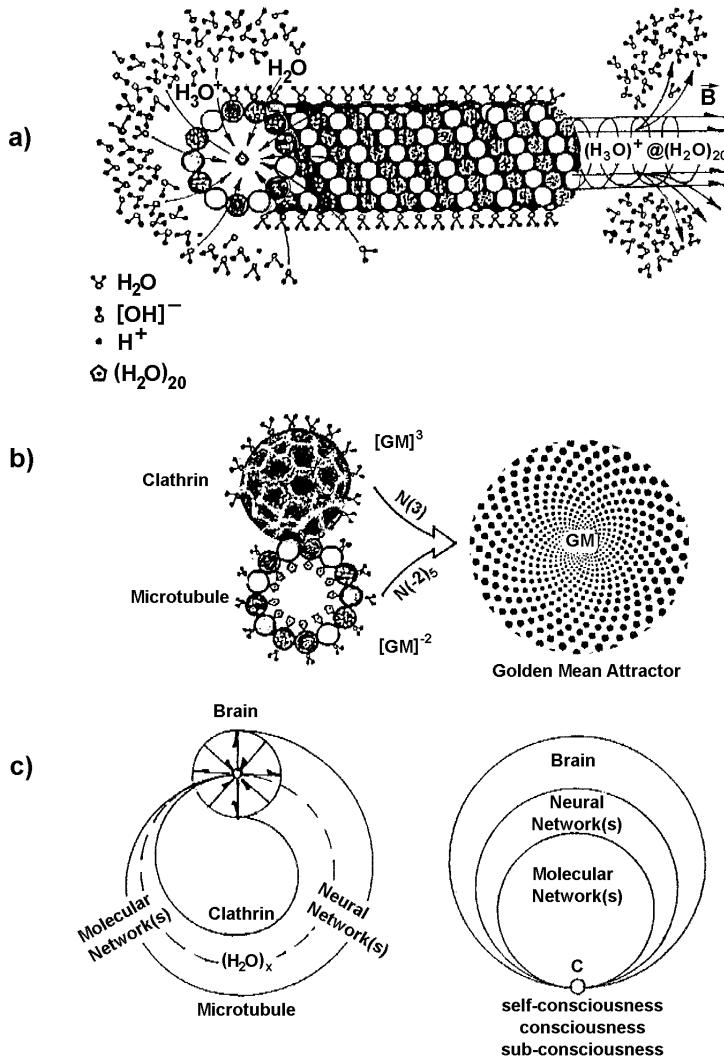
Consciousness can be viewed through: (1) DNA, as  $N = 1$  entity, information realization to become  $N = 0$  (makes own information complement in protein-water interaction), and (2) hierarchy of unity of three wave functions information processes by Golden Mean law; the first level is cytoskeleton activities (sub-consciousness), the second level is neural network dynamics (consciousness), and the third is self-control of the brain (self-consciousness).

### 4.1 Model of sub-consciousness

Microtubules are the cell cylindrical organelles with an outer diameter of about 30 nm and an inner diameter approximately 14 nm. Microtubule paracrystalline energy states are defined by microtubule Golden Mean ( $[GM^+]^2 = N[-2]$ ) lattice structure and dynamics. Basic energy fields around and inside microtubules will be defined by the same,  $[GM^+]^2$  law. According to Eqs. (6) and (8) the complementary structures of Golden Mean will interact with microtubules giving a state of  $N[0]$  with  $D_H^{C(0)}$ . This phenomenon we shall define as the excitation of molecular information (a sparkle of information quanta) or *spark* of consciousness (sub-consciousness). One of the complementary structures to microtubules is the water cluster [\*20]. The relationship between symmetry operations and energy states (like Hückel Molecular Orbital - HMO) of a water cluster [\*20] for  $T_{1g}$ ,  $T_{2g}$ ,  $T_{1u}$ , and  $T_{2u}$  (energy-irreducible representations) are  $GM^+$  and  $GM^-$ . Since information quanta were defined as  $GM^+ - GM^- = 0!$ , the water cluster [\*20] information dynamics in biological environment is an excellent candidate for the point of departure from consciousness. Because

the interior of microtubules exists as an energy field of the Golden Mean law, water molecules will be naturally (spontaneously) assembled into water clusters [\*20] inside of the tubes surface ( $20\text{H}_2\text{O} \rightarrow [\text{H}_2\text{O}]_{20}$ ), according to screw symmetry of 13 protofilaments (Fig. 2a). Inside of the water clusters [\*20] there is an empty space, and a quantum vacuum symmetry breaking approach to information processing of the wave function may be used. Inside of the microtubule and water cluster [\*20], tube currents of very low intensity (10-100 nA) may appear. Weak photo emission is generated and electromagnetic field coupling with gravitational symmetry breaking in microtubular dissipative structures may be established. There is experimental evidence of such an influence of gravity on basic cellular activities, but mechanisms involved in gravitational effects on cells are still unknown [37]. Also, there is experimental evidence that microtubules are gravity sensitive [38], while centrioles may be the gravity spatio-temporal controller of the cell.

On the quantum mechanical level of microtubule-water clusters interaction, the interior of the tube (including centrioles) may provide electromagnetic-gravity waves coupling. On that level dynamics of equilibrium gravity-electromagnetic state under environmental weak fields influence may change states of sub-consciousness and determine its activities.



**Figure 2** Concept of consciousness: a) Around microtubules are water molecules ( $\text{H}_2\text{O}$ ,  $\text{OH}^-$ ,  $\text{H}^+$ ,  $0^{-2}$ ). Most water molecules are in form  $\text{H}_2\text{O}$  on outer surface of microtubule. Because inside the microtubule exists energy field by Golden Mean properties, water molecules will spontaneously assemble in water clusters  $(\text{H}_2\text{O})_{20}$  because they are complement (dual) from structure--energy-information point of view to microtubules; b) Cross section of microtubule-water-clathrin interaction. This is a place of coherence and synchronicity of dynamic activities molecular networks, neural network(s) and brain, through wave functions by Golden Mean structure-energy-information synergy; c) Three as the one: relationship between sub-consciousness, consciousness, and self-consciousness.

## 4.2 Model of consciousness

Cytoskeletal lattices include protein polymers of microtubules, actin, intermediate filaments and more than 15 other proteins. The major neuronal architectural element is the microtubule, which interacts with clathrin on the synapses. Clathrin, as a dual form of the water cluster [\*20], interacts with microtubules from the outside of the tube by Golden Mean laws. This new phenomenon is a preamble to consciousness (Fig.2b). Microtubules, and particularly centrioles, are controllers of molecular network dynamic activities in the neuron. The cytoskeleton molecular network as a sub-neural factor of neuron networks [39,40] may play a very important role in sub-consciousness to consciousness processing integration. The parallel actions of many microtubule-clathrin interactions on synapses and dendrites by Golden Mean oscillations (molecular wave functions) are organized in many interconnecting networks, giving a new quality of information processing - consciousness.

## 4.3. Model of self-consciousness

Coherent control of quantum dynamics of microtubule-water clusters [+@\*20] inside the tube may be the basis of self-consciousness. The outer layer of the microtubule is exposed to the effect of ions from the cytoplasm, affecting the changes of mass and dipole moment of subunits. Due to the change of the dipole moment and the mass of the subunits, microtubules oscillate with the following electromagnetic and acoustic frequencies:  $f_{EM}=6\times10^{15}$  Hz and  $f_{AC}=5\times10^{10}$  Hz [41]. Ion currents of very low intensity may appear inside the tube by [+@\*20] water cluster, which is moving away from one side of the microtubule. This moving is by screw symmetry law, according to energy minimization and configuration of protofilaments. Very low concentrations of [+@\*20] ionic water cluster inside of the tube gives a relative dielectric permitittivity of  $\epsilon_r = 1+10^{-10}$ . According to a relativistic relation between the frequencies measured in the two reference frames of microtubules, inside and outside, moving away from one to another it is possible to write  $f_{in}(MT) = f_{out}(MT)\times K[\epsilon_r]$  [42]. This gives the frequency range 0.2 to 120 Hz for the different number (from 13 to 91) of ionic water clusters [+@\*20] inside the microtubule. This indicates that brainwaves (EEG) may originate from the oscillatory processes of

microtubules and ionic water clusters [+@\*20] through the collective quantum action of many ( $10^{12}$ ) of the neurons in the brain. Based on such dynamics of ion density inside microtubules different states of consciousness might be generated through different excited frequencies. EEG may be explained by deterministic chaos generally [43,44] and as deterministic randomness of Cantor-set-like structure particularly [45], whose Hausdorff dimension is Golden Mean value.

Information processes on these three levels (molecular, cellular and brain) are in quantum coherence (Fig.2c), giving us spark-excited information state (consciousness) of our inner world ( $[(-2)_5 = (GM^+)^{-2}]$  space-time structures) by Golden Mean laws. Our inner 5-dimensional (micro) world is complement with 3-dimensional outer (macro) world, we are currently conscious of.

## 5. TESTING PROPOSED MODELS OF CONSCIOUSNESS

### 5.1 Biomedical testing

General anesthesia is one of the best ways to test models of human consciousness both *in vivo* and *in vitro*. There is evidence, *in vitro*, that anesthetics act in hydrophobic pockets of bacteria and firefly luciferase proteins [46,47]. These enzymes are important *in vitro* anesthetic systems because their easily measured photoemission is inhibited by anesthetics proportional to the anesthetics' clinical potencies. In similar way it is possible *in vitro* to measure wave functions of microtubule-water clusters by STM under the influence of anesthetics. It is interesting that one of the first molecular theories of general anesthesia was proposed as the hydrate-microcrystal (pentagonal dodecahedra) of the clathrate type [48] similar to water clusters.

It was shown that halothane depolymerizes microtubules, but it occurred at higher than clinical concentrations of halothane [49]. According to our model of consciousness anesthetics act on water clusters (cage hydrogen network dynamics) and on the secondary structured dynamics of clathrin and tubulin, through hydrogen bonding. There is experimental evidence for

anesthetic effects on the protein secondary structure [50,51], while anesthetic effects on water clusters will be investigated.

## 5.2 "Artificial brain" model testing

Fullerene C-60 has the same shape and symmetry properties as clathrin in brain [3]. Its structure crystallizes about the zero-dimension ( $N = 0$ ), with a pure vacuum inside cage. Based on this molecule, with Golden Mean structure and energy properties, it will be possible to make artificial microtubule and molecular networks similar to the cytoskeleton. Also, based on Fullerene physics and chemistry it will be possible to build self-assembly systems: from artificial cells (neurons), as a simple one, through complex artificial neuron networks, to the "artificial brain" - by adopting the information physics principles from living matter.

We can then study dynamic responses of a complementary set of globally coupled quantum oscillators (biological brain-*electrical* and artificial brain-*magnetic*) with EEG randomly distributed frequencies, which will be, in the absence of external driving, able to exhibit a transition between the incoherent state and the coherent one with spontaneous synchronization (mapping of conscious contents from biological brain to "artificial brain" and vice versa). This may also solve one of the ultimate engineering goals: man-machine system interaction.

## 6. CONCLUSION

With our goal being to propose a realistic model of consciousness, we find that it is necessary to establish a new physical theory, which we named *information physics*. It was shown that biological information processing from DNA to proteins is information inversion itself. DNA is a 1-D (one-dimensional) information matrix, mapping itself into proteins as a 3-D information entities, which crystallizes around an axis - microtubule (1-D) and point - clathrin (0-D). Biological consciousness is the solution of the 1-D information entity (DNA) which invert itself into proteins (tubulins and clathrin), which interact with water clusters giving 0-D information entity - consciousness. According to this model the Hausdorff dimension of Cantor middle third set of zero-dimension is Golden Mean, as one of the major

properties of consciousness. Although we see our outer world as 3-D, according to this model it is 5-D, because mass is a manifestation of dimension  $N = 0$  in 5-D space structure. Biological information matrix, based on microtubules, is also 5-D, and there is hope that someday, someone, will be able to understand indeed the Universe and Consciousness itself.

We have considered biological consciousness through three levels: molecular (sub-consciousness), neural networks (consciousness) and brain (self-consciousness). The proposed models for testing are, first, based on molecular and cellular mechanism of general anesthesia, and second, based on mapping conscious contents from biological to "artificial brain".

**Acknowledgments:** *This research is dedicated to my parents Ljubomir and Milica Koruga. I would like to thank Richard Hofstad for technical assistance in the preparation of this paper.*

## REFERENCES

- [1] R.W.Hamming, *Coding and information theory* (Prentice-Hall, Englewood Cliffs, 1986).
- [2] D.Koruga, Neurocomputing and consciousness, *Int. J. Neural and Mass-Parallel Comp. and Inform. Syst.* 1 (1991), pp. 32-38.
- [3] D.Koruga, S.Hameroff, J.Withers, R.Loutfy and M.Sundareshan, *Fullerene C-60: History, Physics, Nanobiology, Nanotechnology* (North-Holland, Amsterdam, 1993).
- [4] N.J.A.Sloane, The packing of spheres, *Scientific American* 250 (1984), pp. 116-125.
- [5] J.Leech, and N.J.A.Sloane, Sphere packing and error-correcting codes, *Can. J. Math.* 16 (1971), pp. 657-682.
- [6] J.H.Conway and N.J.A.Sloane, *Sphere packings, lattices and groups* (Springer-Verlag, New York, 1993).
- [7] L.Nottale, Fractals and the quantum theory of spacetime, *Int. J. Mod. Phys. A* 4(19) (1989), pp. 5047-5117.
- [8] G.N.Ord, Fractal space-time: A geometric analogue of relativistic quantum mechanics, *J. Phys. A: Math. Gen.* 16 (1983), pp. 1869-1884.
- [9] O.T.Rosler, Intra-observer chaos:hidden root of quantum mechanics? *Chaos, Solitons & Fractals* 4(3) (1994), pp. 415-421.
- [10] R.S.Shaw, Strange attracters, chaotic behavior, and information flow, *Z.Naturforsch. 36A* (1981), pp. 80-105.
- [11] L.Galgani, Statistical mechanics of weakly coupled oscillators presenting stochastic thresholds, *Lett. Nuovo Cimento* 31 (1981), pp. 65-72.
- [12] O.T.Rosler, Chaos and chemistry, In: *Nonlinear Phenomena in Chemical Dynamics*, C.Vidal and A.Pacault, eds. (Springer, New York, 1981) pp. 79-87.
- [13] K.Tomita, Conjugate pair of representation in chaos and quantum mechanics, *Found. Phys.* 17 (1987), pp. 699-711.
- [14] R.D.Mauldin and S.C.Williams, Random recursive constructions: Asymptotic geometric and topological properties, *Tran. Amer. Math. Soc.* 295(1) (1986), pp. 325-347.
- [15] D.Koruga, Neurocomputing: A geometric-topological approach, In: *Theoretical aspects of neurocomputing*, E.Peliksn and M.Novak, eds. (World Scientific, Singapore, 1991), p. 1939.
- [16] M.S.El Naschie, Is quantum space a random Cantor set with a Golden Mean dimension at the core? *Chaos, Solitons and Fractals* 4(2) (1994), pp. 177-179.
- [17] H.L.Ryder, *Quantum Field Theory* (Cambridge University Press, Cambridge, 1985).
- [18] J.D.Watson and F.H.Crick Molecular structure of nucleic acids, *Nature* 171 (1953) pp.737.
- [19] E.Schrödinger, *What is Life and Matter?* (Cambridge University Press, Cambridge, 1967).
- [20] E.Mandelkov and E.-M.Mandelkov, Tubulin, microtubules and oligomers: molecular structure and implications for assembly, *Cell Movement* 2 (1989), pp. 23-45.

- [21] D.Koruga and J.Simic-Krstic, Semiconductor and crystal symmetry assessment of microtubule proteins as molecular machines, *J. Mol. Electronics* 6 (1990), pp. 167-173.
- [22] R.Mileusnic, S.P.Rose and P.Tillson, Passive avoidance learning results in region specific changes in concentration of, and incorporation into, colchicine binding proteins in the chick forebrain, *Neur. Chemistry* 34 (1980), pp. 1007-1015.
- [23] S.P.R.Rose, Early visual experience, learning and neurochemical plasticity in the rat and the chick, *Philos.Trans.R.Soc.London B* (1977), pp. 278-307.
- [24] S.R.Hameroff and R.C.Watt, Automaton model of dynamic organization in microtubules, *Ann NY Acad. Sci.* 466 (1986), pp. 949-952.
- [25] D.Koruga, Microtubular screw symmetry: Packing of spheres as a latent bioinformation code, *Ann. NY Acad. Sci.* 466 (1986), pp. 953-954.
- [26] S.R.Hameroff, J.E.Dayhoff, R.Lahoz-Belrta, S.Rasmussen, E.M.Insinna and D.Koruga, Nanoneurology and the cytoskeleton: Quantum signaling and protein conformational dynamics as cognitive substrate, In: *Rethinking neural networks: Quantum fields and biological data*, K.H.Pribram, ed. (Lawrence Erlbaum Associates Publishers, Hillsdale, 1993), pp. 317-376.
- [27] P.Dustin, *Microtubules* (Springer-Verlag, Heidelberg, 1984).
- [28] D.Soifer, ed., *Dynamics aspects of microtubule biology*, New York, *Ann. NY Acad. Sci.* Vol. 466 (1986).
- [29] W.G.Kelly, A.Passamiti, J.W.Woods, J.L.Dais, and T.F.Roth, Tubulin as a molecular component of coated vesicles, *J. Cell Bio.* 97 (1983), pp. 1191-1199.
- [30] J.Heuser and T.Kirchhausen, Deep-etch views of clathrin assemblies, *J. Ultrastructure Res.* 92 (1985), pp. 1-27.
- [31] M.Conrad, Quantum molecular computing: The self-assembly model, *Int. J. of Quantum Chemistry: Quantum Biology Symposium* 19 (1992), pp. 125-143.
- [32] M.M.Teeter, Water structure of a hydrophobic protein at atomic resolution: Pentagon rings of water molecules in crystals of crambin, *Proc. Natl. Acad. Sci. USA* 81 (1984), pp. 6014-6018.
- [33] G.A.Jeffrey and W.Saenger, *Hydrogen Bonding in Biological Structures* (Springer-Verlag, Berlin, 1991).
- [34] H.Shinohara, U.Nagashima, H.Tanaka and N.Nishi, Magic numbers for water-ammonia binary clusters: Enhanced stability of ion clathrate structures, *J.Chem. Phys.* 83(8) (1985), pp. 4183-4192.
- [35] U.Nagashima, H.Shinohara, N.Nishi and H.Tanaka, Enhanced stability of ion-clathrate structures for magic number water clusters, *J. Chem. Phys.* 84(1) (1986), pp. 209-214.
- [36] X.Yang and A.W.Jr.Castleman, Large protonated water clusters  $H^{\oplus}(H_2O)_n$  ( $1 < n > 60$ ): The production and reactivity of clathrate-like structures under thermal conditions, *J. Am. Chem. Soc.* 111 (1989), pp. 6845-6846.
- [37] A.Cogoli, A.Tschopp and P.Fuchs-Bislin, Cell sensitivity to gravity, *Science* 225 (1984), pp. 228-230.
- [38] J.Tobony and D.Job, Gravitational symmetry breaking in microtubular dissipative structures, *Proc. Natl. Acad. Sci. USA* 89 (1992), pp. 6948-6952.
- [39] D.Koruga, Molecular network as a sub-neural factor of neural network, *BioSystems* 23 (1990), pp. 297-303.

- [40] D.Koruga, M.Andjelkovic, S.Jankovic, and S.Hameroff, Cytoskeleton as feed-back control system in neuron, In: *Artificial Neural Networks 2*, I.Aleksander and J.Taylor, eds. (Elsevier Science Publishers, Amsterdam, 1992), pp. 399-402.
- [41] C.W.Smith and S.Best, *Electromagnetic Man* (J.M.Dent & Sons Ltd., New York, 1989).
- [42] D.Koruga, Neuromolecular computing, *Nanobiology* 1 (1992), pp. 5-24.
- [43] A.Babloyantz, Chaotic dynamics in brain activity, In: *Dynamics of sensory and cognitive processing by the brain*, E.Basar, ed. (Springer, Berlin, 1988), pp. 196-202.
- [44] W.J.Freeman, Simulation of chaotic EEG patterns with a dynamic model of the olfactory system, *Biol. Cybernetics* 56 (1987), pp. 139-150.
- [45] B.J.West, *Fractal Physiology and Chaos in Medicine* (World Scientific, Singapore, 1990).
- [46] N.P.Franks and W.R.Lieb, The firefly throws light on anaesthesia, *Chemistry in Britain*, October 1985, pp. 919-921.
- [47] S.Curry, W.R.Lieb, and N.P.Franks, Effects of general anesthetics on bacterial luciferase enzyme from *Vibrio harveyi*: An anesthetics target site with different sensitivity, *Biochemistry* 29 (1990), pp. 4641-4652.
- [48] L.Pauling, A molecular theory of general anesthesia, *Science* 134(3471) (1961), pp. 15-21.
- [49] A.C.Allison and J.F.Nunn, Effects of general anesthetics on microtubules, *The Lancet*, December 21 (1968), pp. 1326-1329.
- [50] A.Shibata, K.Morita, T.Yamashita, H.Kamaya, and I.Ueda, Anesthetic-Protein interaction: Effects of volatile anesthetics on the secondary structure of poly (l-lysine), *J. Pharm. Sci.* 80(11) (1991), pp. 1037-1041.
- [51] D.Koruga, J.Simic-Krstic, S.Jankovic, D.Louria, and S.Hameroff, Ethanol effects on secondary structure of firefly luciferase (to be published).

#### 4. The whole text from Reference (Rakočević, 1995)

## THE UNIVERSAL CONSCIOUSNESS AND THE UNIVERSAL CODE

Miloje M. Rakočević

Department of Chemistry, Faculty of Science, University of Niš,  
Æirila i Metodija 2, 18000 Niš, Yugoslavia

**Abstract.** There are many approaches to investigate consciousness. We propose one which utilize *logic* in the sense of *logos* (from the ancient Greek *logos* meaning *coding*, as a something that was at the beginning). For living beings their *matter logos* point of departure is the periodic system of elements because genetic code is its second hand side. In the scientific work of D.I.Mendeleev the important aspects of periodic law were examined, which the 20th century scientists have never adequately understood. Also, Darwin's diagram, in his book *Origin of Species*, represents a specific code, which is the first example of the code model and the code system in biological science. What is the point of departure of *mind logos* it is difficult to say, but on the consciousness level, after five thousand years in the development of the binary numbering system, the time has become ripe for a unification of the two different approaches in studying the universe and human beings: the global-integral method of the Ancient East, based primarily on speculation (logic of *speculative mind*), and the single-partial method of the Modern West, based primarily on experimentation (the logic of *natural mind*). Even more, today it is becoming evident that a scientific basis of human consciousness cannot be understood without such unification. Before such a unification can be done, one possible general hypothesis about the existence of a universal *Mind/Matter* code should be tested and proven. This paper is subject of that kind of consideration.

**Key words:** *logic, periodic system of elements, genetic code, microtubules, water, Yin-Yang code, consciousness*

### 1. INTRODUCTION

There are many approaches in investigating consciousness. We propose one which utilizes logic in the sense of *logos* (from the ancient Greek *logos* meaning something that was at the beginning). In spite of this seemingly very narrow approach, from the physiological-psychological aspect it is more extensive. More in accordance with Russel's [1], and Vygotskii's [2] and less with Ashby's [3] and Arbib's [4] views (Russel, p. 168: "... we are said to be 'conscious of' something; in this sense, 'consciousness' is a

relation"; p. 170: "The ... relation to an 'object', it could be said, is characteristic of every kind of consciousness", p. 173: "Nevertheless we can distinguish 'mental' events from others ..."; Vygotskii, first chapter: "... consciousness is a unity of all functions"; Arbib, p. 1: "... all the functioning of the nervous system relevant to our study is mediated solely by passage of electrical impulses by cells we call neurons"; Ashby, p. 11: "... the book deals with only one of the properties of the brain, and with property - learning - that has long been recognized to have no necessary dependence on consciousness"; p. 12: " And until such a method ... the facts of consciousness cannot be used in scientific method"). Our fundamental hypothesis in establishing our approach is that the investigation of the consciousness must always consider a whole system as a unity of physiological (including biochemical and biophysical) - psychological-logical characteristics; with two subsystems: physiological-psychological and psychological-logical. Considering this concept, together with Russel's idea about consciousness as a relation to objects in sense to be 'consciousness of' something, with Sartre's idea [5] that any consciousness is the consciousness about something, with Petronijeviæ's idea [6] that the contents of the consciousness are the notions, and finally, with Einstein's idea [7] that all notions within 'natural laws' are from the space-time nature, we will show how human consciousness in some specific ways has expressed itself in the various works by different creative investigators through different epochs. On the other hand we will show that this human consciousness is in a specific relation with a universal consciousness within universal code.

## **2. PRELIMINARIES**

More than any other scientists, Crick and Einstein are responsible for our link between the molecular basis of life and consciousness. Crick made first and important step, with arguments that consciousness is a property of molecular activities in neurons and networks of neurons in our brain [8], while a serious analysis of complete works of Einstein lead to the conclusion that when he speaks about the four-dimensional continuum of space-time, he means in fact three-four-dimensionality. This opens a possibility to speak about coding coordinates and coding spaces; by doing this, each Boole's space characterized by three-quaternitity must be taken as Boole-Einstein's space. In fact Coding Space (CS) unavoidably should to be

Coding Space-Time (CST), what is subject of the information physics as a new scientific discipline of space-time structures [9,10].

The basic parameters which determine physico-chemical characteristics of *a system* of stable chemical elements are: atomic number, number of period, number of the group and number of isotopes. Mendeleev never used word *Table*, what we usually do, to present his work, but *System*. His original *System* of elements is different from our today's *Table* of elements. Science of 20th century escapes Mendeleev "mysterious" form of his system of elements, saying that Mendeleev made some arithmetical errors (ref. [11], p. 185). Mendeleev's "errors", or our inadequate understanding of his work, is the information (coding system  $3^4 = 81$  and  $4^3 = 64$ ) approach to elements. Information approach as coding approach. The coding system  $3^4 = 81$  - because within first 84 chemical elements (from H = 1 to the Po = 84) there are exactly 81 stable elements. The coding system  $4^3 = 64$  - because 84 minus 20 "monoisotope" elements equals 64 (cf. ref. [12], ch. 27, sect. "Relations odd-even", where Gould says that all even elements to the polonium, Po = 84, have minimally two stable isotopes, except beryllium; cf. ref. [13], where Rakoèeviæ says that within chemical code there are exactly 84 elements; cf. 64 hexagrams in Fig. 5 and 81 tetragrams in Fig. 6). Mendeleev clearly and precisely gave the system of chemical elements as a four-dimensional Boolean hypercube [14]. The same approach, based on Boolean hypercube, has been used recently by Kauffman to explain the origins of order as point of departure of self-organization and selection in evolution [15].

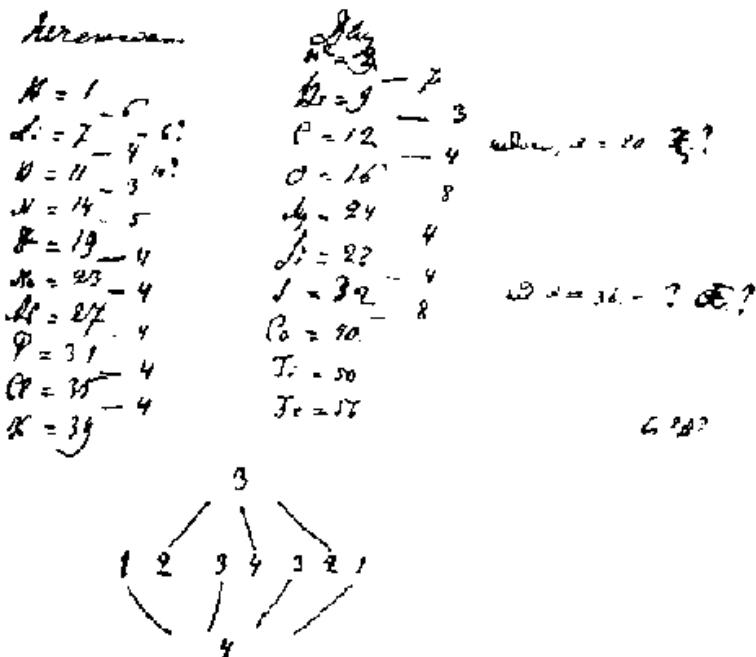


Figure 1 The universal consciousness on the universal code of the nature: the Mendeleev's system of chemical elements oddity-evenness principle; the valence trinity-quaternity system; the atomic mass distances integer system.

In the scientific work of D.I. Mendeleev, especially his original manuscript works (see ref. [11], pp. 128-129), three important aspects of periodic law were examined, which have never been adequately understood by 20<sup>th</sup> century science. These are: (1) the system relations among elements including odd-even principle, (2) spaciousness i.e. three-dimensionality (Figs. 1-2), and (3) cyclicity (photocopy XII in Kedrov: Cu, Ag, Au within first and after that within eighth group at the end, parenthetically). From these facts it follows that the third dimension of the periodic system as a "New dimension for Mendeleev" [16] is not necessary because Mendeleev was conscious of dimension 100 years ago. Also, Mendeleev was conscious of the problem of "rare earth", although there are different opinions (ref. [16], p. 13: "The two versions differ simply in their arrangements to accommodate elements such as the rare earths, but the result must be to leave many with the impression that Mendeleev had not made up his mind about something of importance"). In his long periods Table (ref. [11], p. 188) Mendeleev gave a specific position to the first element of "rare earth", i.e. lantanides (Ce) - not in the third but in fourth group; then still 13 groups

for 13 lantanides. If so, Mendeleev must have been seen conscious that the short periods Table with 8 groups of elements corresponds with cube, and long periods Table with a hypercube (8 group of non-rare earth elements plus 14 group of yes-rare earth elements with a superposition equal 14 groups; then: 14 groups plus the first group at the end, and plus the zeroth group equal 16 groups). On the other hand, Mendeleev was conscious of the fact that there must be 4 types of elements with correspondence to the square (for details see ref. [14], pp. 197-200). Bearing in mind all these Mendeleev's insights, it was possible to determine the definitive real positions of elements within 3-4 dimensional periodic system (ref. [13], pp. 19-22). Other creative investigators, primarily Darwin and Mendel [14] were conscious of square-cube-hypercube relations within natural systems.

Таблица 13

Изображение периодической таблицы элементов в системе координат  
(по Д. И. Менделееву) с пулевой группой в качестве оси абсцисс

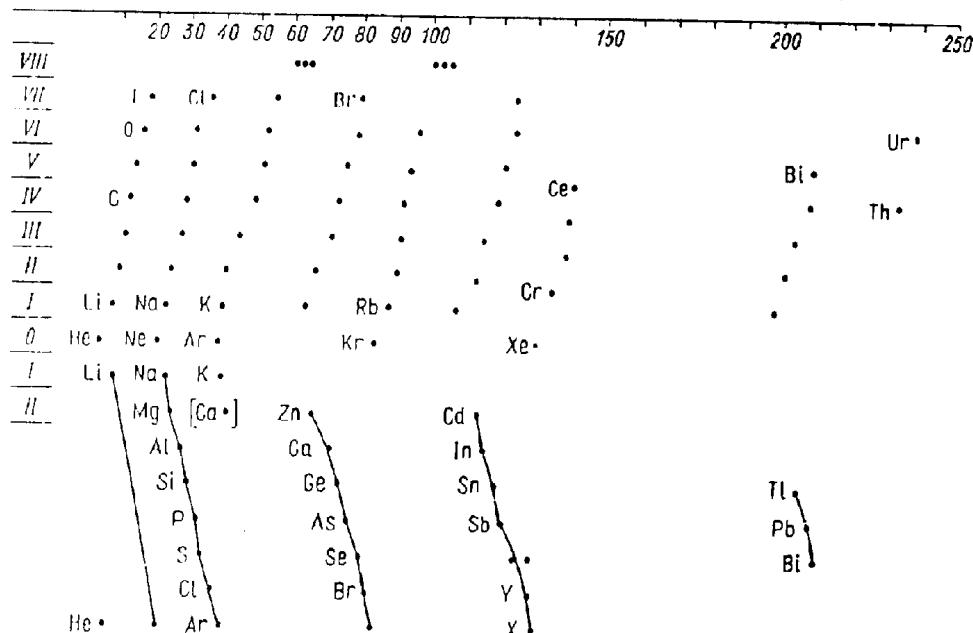


Figure 2 The universal consciousness on the universal code of the nature: the Mendeleev's three-dimensional periodic system as unity of cube-sphere; cube: through the three coordinates; sphere: the noble elements are at all three coordinates.

Darwin's diagram, binary tree, represents the first systematic information approach to the analysis of the relations between organisms [17]. This is the only diagram in his book *Origin of Species* and it represents a model of interpretations of the origin of varieties, species, genera and higher systematic categories. By its essence, his diagram represents a specific code-model and code-system as the first example in biological science. Relations of the noted elements within this code system correspond to the relations of the organisms in the natural systems. Hidden message of this diagram now is clear: if the natural systems are at the same time the coding systems, the only adequate and complete way of description and interpretation of such systems would be the creation of adequate code models with adequately corresponding relations between the elements of the one and the other model.

The main idea which is in the basis of the diagram - binary tree, is the realization of the logic of systematization and classification, separation of the parts within the whole, as well as the regularity of the hierarchy of the levels. The accordance of this logic with the model of classification of the number systems with the number basis  $N_2 = 2(2n+1)$ , where  $n = 0,1,2,3$ , is directly obvious. So, we have for  $n = 0$ ,  $N = 2$ , which corresponds to the division of the binary tree to the left tree and the right tree. This is exactly what Darwin discussed on the relations during the evolution only along two lines at the beginning of which "species (A)" and "species (I)" occur: "These two species (A) and (I), were also supposed to be very common and widely diffused species, so that they must originally have had some advantages over most of the other species of the genus". The obvious characteristic of the Darwin's diagram of the binary tree is that each transition to the next level completely follows the logic of the Gray code, since only a unit change is allowed [14].

### 3. LOGIC OF MOLECULES OF LIFE

The problem of accordance-discordance of the genetic code ( $4^3=64$ ) and its physico-chemical basis was firstly stated by Crick who demonstrated that this problem is impossible to separate from the questions related to the ratio of probabilistic and deterministic in the coding process: the fact that codons  $X_1, X_2, X_3 \dots$  are coding for amino acid  $Y$  results from numerous accidental

processes during the evolution, or here strict (deterministic) reasons could be also included [18,19].

The very approach to the three-four-dimensional system of both the chemical code (Mendeleev) and the genetic code should be integral one: it has to be emerging logic approach. We have shown [14] that atomic mass and number of isotopes represent the principal determinants of the chemical code ( $3^4 = 81$ ), while binary values of codons and amino acids represent the principal determinants of the genetic code ( $2^6 = 64$ ). (Figs. 3-4).

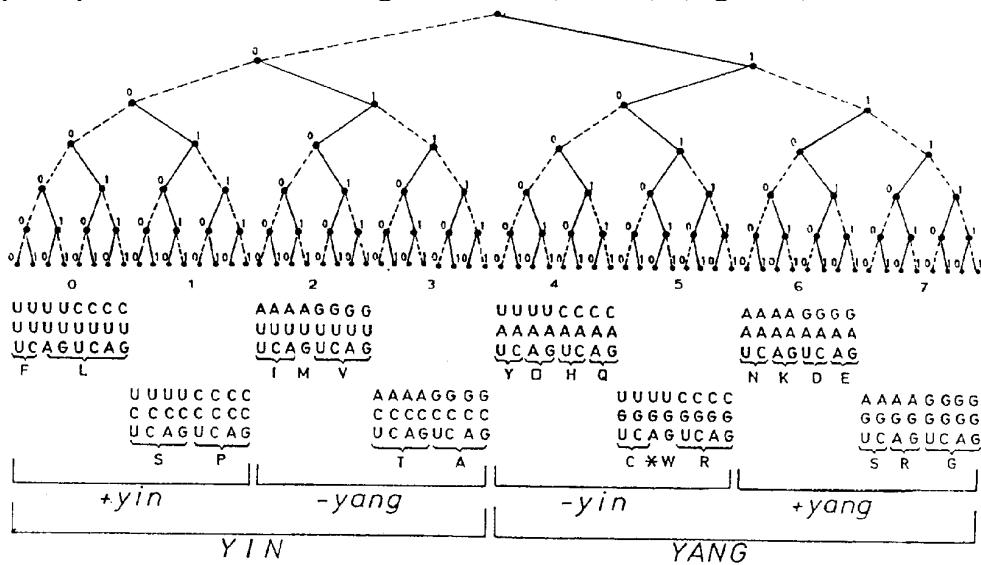


Figure 3 The universal consciousness on the universal code of the nature: the genetic code binary tree is in full accordance with the yin-yang binary tree in the oldest book - I Ching. The relations between 64 codons and 64 hexagrams: to each zero corresponds a broken line; to each one corresponds an unbroken line (see ref. [14], p. 274). This binary tree of the genetic code with the order of the eight families of codons (rosettes), which corresponds to the series of natural numbers (0-7). The four 16-codon families begin with broken lines for +Yin and/or +Yang states; full lines for -Yin and/or -Yang states, reflecting the greater or lesser influence of bases of the *Py* and/or *Pu* type.

### 3.1 The Number of Trinity-Quaternity

The basic concept from which we start is the Boolean logical square. This hidden square exists within the Gray code model of genetic code [20,21]. The Gray code model of the genetic code can be *per se* developed

in two types of the binary tree: (1) the binary tree which keeps the logic of the Gray code having characteristic that "two adjacent symbols differ at only one bit" [20], and (2) the binary tree with the logic of natural numbers series "for the numbers 0-63" [21]. With the first type of binary tree, the distances between codons are the unit Hamming distances if "measured" by weight, i.e. by norm of Boolean vector, while with the second type of binary tree the distances are also the unit ones if "measured" via the vector number.

According to the logic of the Boole's square, longitudinal diagonal of the Boole's cube has to be labeled by the following sequence of corners: 1076. The end-corner, 1776, in the diagonal is optimal only in the system of trinity-quaternity (*TQ*):  $1076 + 700 = 1776$  (trinity because it is cube; quaternity because there are four positions; optimal because the corner 7 of the three-digit-record does not change position during "transition" to four-digit-record). *TQ* system should be understood as the unity of Boole-Einstein's cube-hypercube: there are two sevens at the longitudinal diagonal as a result of permanent coping 0-7 and 7-0 (within the frame of the cube) and/or as a result of permanent motion from starting to middle point and backwards, i.e. from end-point to middle point and backwards (within the frame of the hypercube). Sevens from longitudinal diagonal are intercrossed by two sevens of the middle diagonal (within the frame of hypercube) which appear as a result of permanent coping 7-8 and 8-7. Therefore, at each moment of time, there is a system 0777-7777 in Boole-Einstein's coding space, and/or 1776-17776, if we take into account cyclicity of the system.

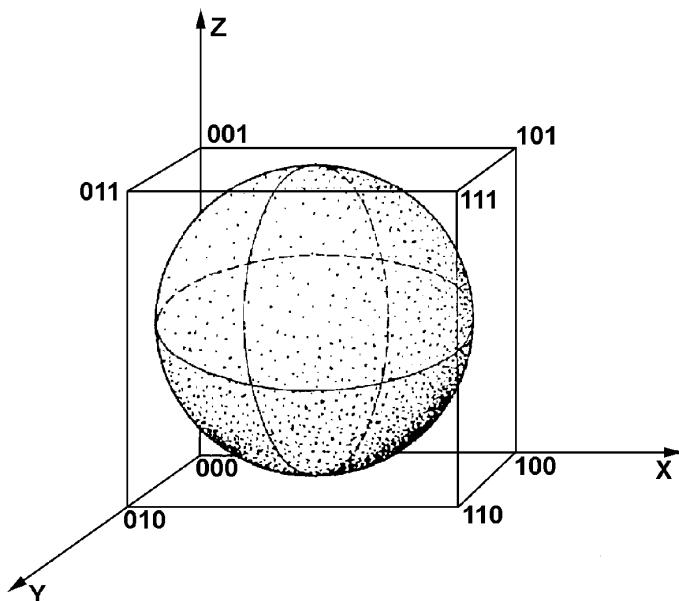


Figure 4 The universal consciousness on the universal code of nature: the **LIGHT** model of the genetic code is in full accordance with Mendeleev's cube-sphere model (Fig. 2) of the chemical code (for details see ref. [14], p. 54).

### 3.2 Genetic Code as TQ System

Analysis of the Crick's papers published immediately after 1966, upon definitive establishment of the Table of codons, demonstrates his sagacious observation of the problems imposed by Table itself, in spite of its beauty and symmetry. Determination of the number of problems and their denomination are ours, while the original statements are those of Crick.

(1) *Problem of the alphabets (problem of coding):* Four-letter language of the nucleic acids has been studied in the meantime and we know how it controls 20-letter language (amino acids) of the proteins. However, in spite of the fact that numerous problems remained unsolved, this knowledge is certain;

(2) *Problem of the neighborhood (both codons and amino acids):* Neighbor amino acids are coded for by neighbor codons;

(3) *Problem of similarity of codons:* It is sure that triplets coding for the same amino acid are most often very similar;

(4) *Problem of the position of base I, II, III in triplet*: In any case triplets with U and C at the end of the codon are coding for the same amino acid, and this is also very often the case with the triplets containing A and G as the end base;

(5) *Problem of selectivity of base I, II, III within a codon*: Amino acid is chosen mainly by the first two bases in a triplet;

(6) *Problem of a sign*: Is the allotment of a triplet to amino acids at random, or there are structural reason for this?

(7) *Problem of the meaning*: What is the sense of the synonymy through the third base, and what is the sense of the exceptions?

(8) *Problem of the form*: Form of the genetic code is established with a considerable certainty;

(9) *Problem of the essence (what is the "corner stone")*: Genetic code is an important corner stone at the long path of molecular biology and biological life;

(10) *Problem of the origin and evolution of the genetic code*: When we answer all these questions, the question of the origin of the genetic code will remain as the major problem. Is the genetic code the result of a series of evolutional coincidences? The origin of the genetic code will remain as the major problem.

Two amino acids; *serine* and *arginine* have been the main problem not only for Crick, but also for all other researchers undertaking serious studies on the essence of the genetic code. The codons coding for these two amino acids for each of them are very different and even separated in the table. This is then the reason Crick could not claim with certainty that similar codons code for similar amino acids (this should be expected on the basis of chemistry) or that neighbor codons code for the same amino acid. So, Crick could only say that they are "the most often very similar". Position of the third base within a triplet makes new problems. Coding process is not affected when pyrimidine bases (U and C) appear in the third position, i.e. the same amino acid is coded (synonymy). If we use information-topological model of the genetic code it is practically possible to solve all ten Crick's problems of genetic code: four-letter language is at the same time the language and the chemical essence; similar codons indeed code for similar amino acid and again without an exception; neighbor codons code for the same amino acid and again without an exception. In our information geometry approach [14] genetic code is completely characterized by entity of *TQ* system. Two pyrimidine (*Py*) and two purine (*Pu*) bases are inevitably expressed in the coding space as the system 3+1 (three with oxo-

group and one without it, or three with amino group and one without it). Therefore the number of codons in the table of the genetic code by positions in four groups should read from an aspect of the main coding position. In spite of this distinction 3+1 only on the basis of numbers presented in this manner it is impossible to understand possible physico-chemical meaning of *TQ* system in the genetic code without an analysis of the internal structure. Each position consisting of 16 units can be taken as position with the structure 8:8 what makes sense from a physico-chemical aspect, since 16 families of codons  $(1 \cdot 16) \cdot 4 = 64$  could be understood also as a system of  $(2 \cdot 8) \cdot 4 = 64$  codons. In one family of higher order (eight-membered one) there can be 8 codons with *Py* base in the first position, and/or 8 codons with *Pu* base in the first position. However, there is a question whether classifications such as (8-1):(8+1), (8-2):(8+2), or some others make sense? We will put forward the *hypothesis* that the classification (8-1): (8+1), so 7:9 makes sense also from an aspect of strictly determined physico-chemical parameters such as hydropathy (H) or polarity (P), as well as from an aspect of the principal parameter of the binary value. (For details see ref. [14], pp. 253 and 255-260). It was proposed the existence of binary values for the entities of two pyrimidine and two purine bases: *U(00)*, *C(01)*, *G(10)*, *A(11)* [22]. Swanson made the same assumption, but with a significant difference: *U(00)*, *C(01)*, *A(10)*, *G(11)* [20]. She showed that the binary record of a codon must begin with the second (middle) and not the first base, as was proposed by Schonberger. When we used our *TQ* coding system we satisfied all Crick's genetic code problems and find similarity with quantization the magnetic quantum number (-3,-2,-1,0,+1,+2,+3). This indicate the basic biomolecular information processing associate with quantum field trough coding. Our results of genetic code are summarized in ref. [14]. We can say that genetic code is preamble of biological consciousness which arise in proteins-water interaction and through activities of cell (molecular networks), body network, neural networks and brain, that lead human beings to be conscious [8,10].

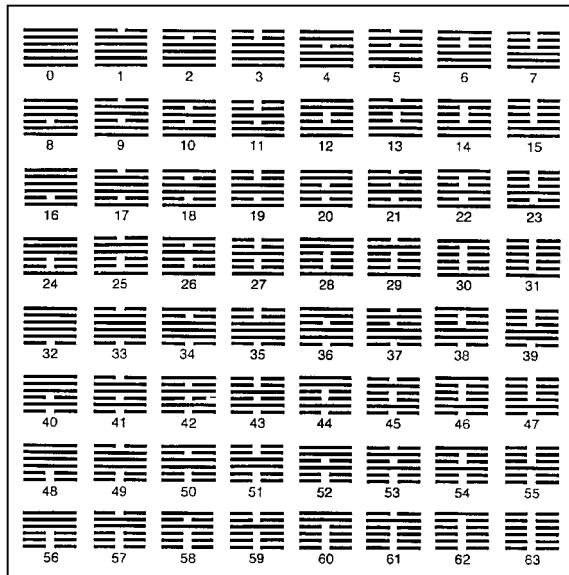
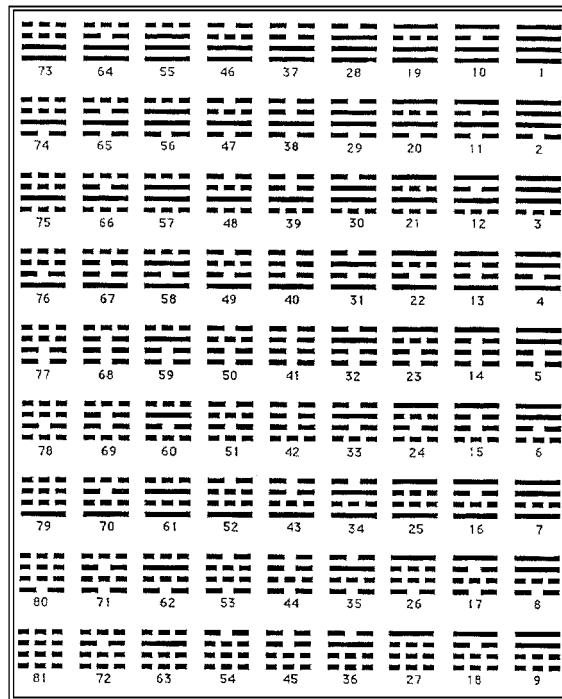


Figure 5 Speculative system of ancient China based on hexagonal arrangement: It consists of 64 hexagrams ( $2^6 = 64$ ), which display every possible combination of archetypal human situations – along with thousands of variations caused by changing lines. Notice that an inverse countdown is possible. In such a case the 63rd number is zeroth. If so, then  $2 \cdot 6$  and  $2^6$  are the numbers of lines and hexagrams respectively.

### 3.3 Microtubules Coding System

Microtubules coding system was identified by Koruga [23]. There is microtubule coding system of two codes;  $K_1[13,2^6,5]$  and  $K_2[24,3^4,13]$ . First code,  $K_1$ , is result of tubulin subunits packing in protofilaments by screw symmetry. This code has 64 codewords, length 13 and distance 5 (the best known binary error-correcting code). Second code,  $K_2$ , is result of interaction of 24 tubulin subunits and high molecular weight MAP (microtubule-associated proteins). This code has 81 codewords, length 24 and distance 13 (the best efficient code for information transmission) (cf. these 64 and 81 codewords with analogous "codewords" within chemical code, mentioned in Preliminaries).



*Figure 6 Speculative system of ancient China based on tetragram arrangement: A tetragram is constructed from four stacked lines of three types (solid, broken, and twice-broken). There are 81 possible combinations of these three types of lines ( $3^4=81$ ). It is no accident that ancient Chinese book *Tao Te Ching* has eighty-one chapters, what is also a significant number to those Chinese philosophers who treasured the symmetry of numbers. Notice that an inverse countdown is possible. In such a case the 81st number is zeroth. If so, then 3·4 and  $3^4$  are the numbers of lines and tetragrams, respectively.*

### 3.4 Biological Water Mystery

The essential role of water has been recognized in all studies of biological processes, but it is a paradox that we know very little about order and properties of "biological water". Water seems to be the fastest solvent, because simulations predict and experiments verified femtosecond dynamics of water [24]. It is well known that water molecules may be organized in different ways but one of the most promising is the clusters organization. Water clusters may exist with 10 to 1000 water molecules. An

approach of water cluster cellular automata (WCCA) may be the right way to solve the problem of its mysterious role in biological information processes. If we look at cell from the system theory approach, biological water seems to be "*intelligent solvent*".

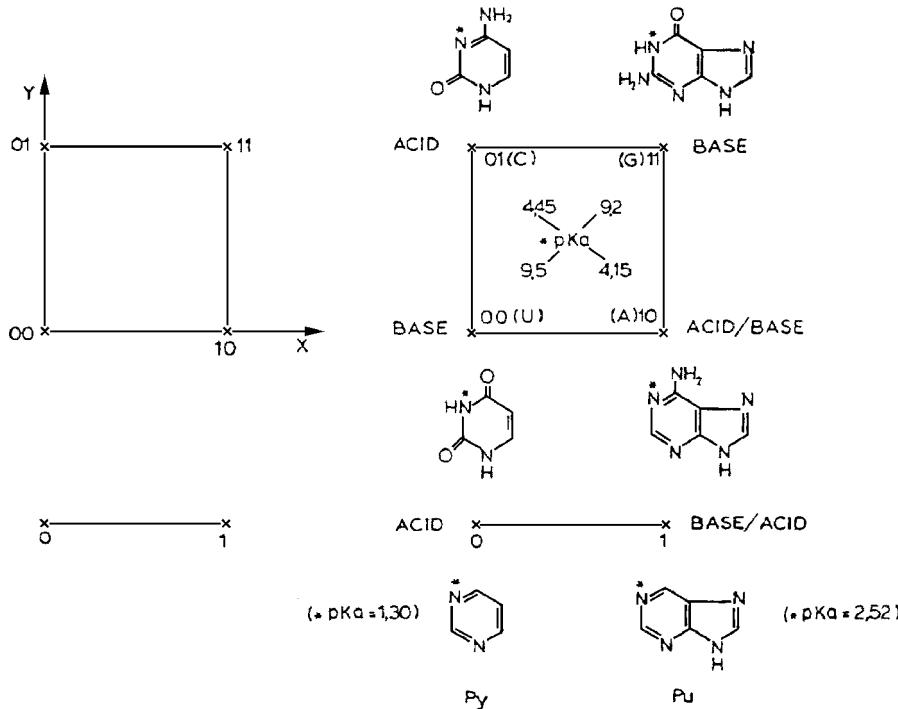


Figure 7 The universal consciousness on the universal code of Nature: the Boolean logical square of the genetic code is in full accordance with the Aristotelian (Boolean) logical square of four entities (cf. Fig. 8) (For details see ref. [14], p. 8).

#### 4. LOGIC OF SPECULATIVE MIND

The logic of *natural mind* is primary result of human brain adaptation through its interaction with environment (Nature), while logic of human *speculative mind* is primary result of human field based mind interaction with *Mind Itself*. The best example of speculative mind related to natural mind is Chinese concept of *Yin-Yang*, while Chinese concept of *Dao* may be related to *Mind Itself*. Schonberger was the first to point the possibility of

making the *I Ching* and the genetic code conform to the same model by using binary records [22].

#### 4.1 Dao and Mind Itself

From a scientific point of view we do not know yet what *Mind Itself* is, but if we identified *Mind Itself* with *Dao* we can learn that

*The Dao that can be expressed  
Is not the Dao of Absolute.  
The name that can be named  
Is not the name of the Absolute.*

*The Dao is empty and yet useful;  
Somehow it never fills up.  
So profound!  
It resembles the source of All Things.*

#### 4.2 Yin-Yang and Natural Mind

We shall demonstrate the underlying meaning of the link, coherence and interdependence of the natural code (*natural mind*) and the *I Ching* code (*speculative mind*). We shall show that there is a complete and perfect correspondence between the *Yin-Yang* entities in the *I Ching* code ( $2^6 = 64$ ) and the pyrimidine-purine entities in the genetic code. Our starting point has been Stent's discovering that *Yang* (the male or light principle) is identified with the purine bases and *Yin* (the female or dark principle) with pyrimidine bases [25]. It was known in the ancient China that the *Yin-Yang* entities may be extended, so there can be +*Yin* (Great *Yin*) and -*Yin* (Lesser *Yin*) or +*Yang* (Great *Yang*) and -*Yang* (Lesser *Yang*), what lead us to new conclusions, which can be summed up in Figs. 3-8 and the following points:

(1) Boole's logical square lies at the heart of the *I Ching* as well as of all natural codes. The logical square of the four entities of the *I Ching* should be turned  $180^0$  to make it correspond to the logic square of the four elements known to Aristotle: *Air* and *Fire* associated with *Yang*, and *Earth* and *Water* with *Yin* (Fig. 8).

(2) There is complete congruence and correspondence between the six-bit binary records for the 64 codons and the binary records for the 64 hexagrams in the *I Ching* (Fig. 5).

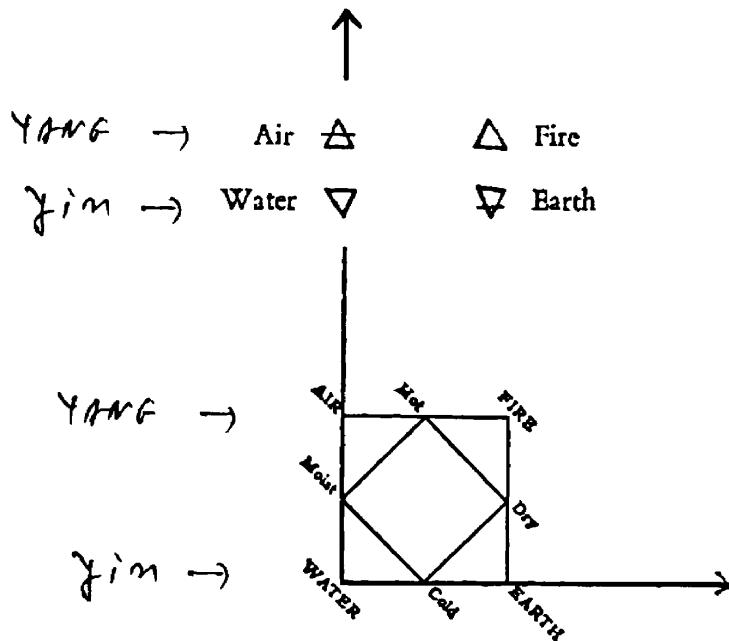


Figure 8 The universal consciousness on the universal code of Nature: the Alchemists' (Boolean) logical square follows from Aristotelian square; both are in accordance with the yin-yang system and with the genetic code logical square (cf. Fig. 7); further, they are in accordance with the fundamental particles square: neutrino (00), electron (01), quark down (10), quark up (11). For the (Boolean) logical square regarding the fundamental particles see ref. [14], p. 283.

(3) The binary tree of the *I Ching* should be turned  $180^0$  for it to fit exactly over the binary tree of the genetic code; then both binary trees correspond to Farey's binary tree, which determines the quasiperiodical transition to deterministic chaos (cf. ref. [14], p. 280).

(4) The eight trigrams in the *I Ching* are analogous to the eight rosettes, i.e. eight families of codons on the binary tree of the genetic code (Fig. 3).

However, the system: +Yang, -Yang, +Yin, -Yin may be presented as  $3^4 = 81$  (Fig. 6). This indicates that *Yin-Yang* coding system is the same as microtubules coding system.

## 5. FINAL COMMENTS

In fact the three quoted aspects of periodic law are of great importance today, when it has been demonstrated that the system entity, spatiality, periodicity and cyclicity are the most important characteristics of the genetic code (Figs. 3-4). From these figures it is clear why it is sensible to speak about the chemical code in connection with the genetic code. There are certain aspects of correspondence and coherence of the two codes: (1) within the genetic code there are exactly 61 amino acid (stable aggregation) meaning codon situations, plus 3 breaks in amino acid meaning (3 "stop" codons), plus 20 non-codon situations (20 protein amino acids); (2) within the chemical code there are exactly 61 situations (in the form of stable aggregations) which have multi-isotope meaning, plus 3 breaks in stable isotope (3 "stop" situations: Tc, Pm, Po), plus 20 non-(stable) isotope situations (20 "mono isotope" elements). Thus altogether there are 84 entities within both the genetic, and chemical codes. This, is the very topic: the chemical code, built on the very principles mentioned and in complete accordance with the genetic code. Such a surprisingly simple model at the same time represents the Logical-Informational and Geometrical-Homeomorphous-Topological (LIGHT) system of the Boolean cube-hypercube with an inscribed sphere-hypersphere (Fig. 4). The 8 vertices of the cube in Fig. 4 correspond with the 8 rosettes on the binary tree in Fig. 3 (8 families of codons); the 16 vertices of imagined hypercube correspond with 16 families of codons on the binary tree.

d	c	b	a	e	f
				.	1
1	0	<b>1</b>	<b>00001</b>		
1	0	<b>1</b>	<b>00002</b>	2	9
1	1	0	00004	1	1
1	0	<b>1</b>	<b>00008</b>	2	12
1	0	<b>1</b>	<b>00016</b>		
1	1	0	00032	2	<b>3</b>
1	1	0	00064		
1	0	<b>1</b>	<b>00128</b>	1	M
1	1	0	00256	2	<b>3</b>
0	1	0	00512		
1	0	<b>1</b>	<b>01024</b>	2	12
1	0	<b>1</b>	<b>02048</b>		
0	1	0	04096	1	1
0	0	<b>1</b>	<b>08192</b>	2	9
0	0	<b>1</b>	<b>16384</b>		
				.	1

Figure 9 The universal consciousness on the universal code of Nature: the Homer's and Njegoš's (Boolean) space sequence,  $N=2^n$ . (a) The binary sequence whose sum is  $2^{15}$ .  $1=32767$ ; (b) Homer's choice:  $27803_{10} = 66233_8 = 110110010011011_2$  (the number of verses for *Iliad* plus *Odyssey*); (c) Njegoš's choice:  $4964_{10} = 11544_8 = 001001101100100_2$  (*The Mountain Wreath*: printed version 2819 verses plus 318 person-scenes, plus 116 pages for printing = 3253; manuscript version 1528 verses plus 150 person-scenes plus 033 pages = 1711; all together - the total spaces of *The Wreath*:  $3253 + 1711 = 4964$ ); Homer's plus Njegoš's system:  $66233_8 + 11544_8 = 77777_8$ ; (d) Homer's choice:  $3583_{10} = 6777_8 = 000110111111111_2$ . The number 3583 represents the difference of *Iliad* and *Odyssey*:  $15693 - 12110 = 3583$ . The relation between two numbers  $77777_8$  and  $6777_8$  was given through a logic program: to exclude first position, and then - to write the result ( $7777_8$ ); after that: to exclude first unit in the first position, and, then to write the result ( $6777_8$ ). The choice logic for the number  $3583_{10} = 6777_8$  is as follows. From the total sequence ( $2^{15} - 1$ ) to exclude all the situations that contain the whole third perfect number 496; (e) The number of Homer's yes-choice and non-choice situations; (f) The (in literary science) known composition

sequence of *Iliad*: from the middle point Mission to Achilles) 1 day full, 9 empty of events etc.

The basic (main) relations, determinants and invariants within the binary tree, i.e. within the system of cube-hypercube are the relations of a Boolean logical square. There are 4 types of molecules within the genetic code: Uracil ( $U = 00$ ) with number 0 of Boolean vector, Cytosine ( $C = 01$ ) with number 1, Adenine ( $A = 10$ ) with number 2 and guanine ( $G = 11$ ) with number 3 of Boolean vector (Fig. 7). Also, there are 4 types of chemical elements:  $s(00)$ ,  $p(01)$ ,  $d(10)$  and  $f(11)$  in relations of the Boolean logical square; plus 4 types of fundamental particles: neutrino ( $00$ ), electron ( $01$ ), quark down ( $10$ ) and quark up ( $11$ ). Fig. 8 illustrates the consciousness about logical square relations within the Universe of Aristotle and Alchemists. The periodicity and cyclicity within the genetic, as well as within the chemical code, are in accordance with periodicity and cyclicity of the natural number systems with the base  $N_1 = 2^n$  ( $n = 1, \dots, 6$ ) and  $N_2 = 2(2n+1) = 4n+2$  ( $n = 0, \dots, 5$ ). The relations within these mathematical number systems lead to the Golden Mean, as one of the most important Laws in Nature (for details, see ref. [14]). It could be said that these number systems are the natural number systems. And then a new surprise: human consciousness in some specific way expresses itself through masterpieces such as those written by Goethe, Shakespeare, Tolstoy, Njegoš. We find their compositions were written according to the same Law as possessed by the chemical and genetic codes - the Golden Mean. Fig. 9 shows how Homer and Njegoš generated their works from the binary sequence  $N_1 = 2^n$ .

From this discussion it follows: it makes sense to give some separate hypotheses for further investigation. For example: (1) Human consciousness as a specific brain-computer code must be determined by Boolean spaces; (2) Human consciousness in the form of human language must be determined by Boolean logical square (Fig. 10); (3) Human consciousness as logical reason (syllogism etc.) must also be determined by Boolean logical square (Fig. 11).

## PRINCIPES DE PHONOLOGIE

On obtient ainsi le schéma des variations possibles :

	I	II	III	IV
a	Expiration	Expiration	Expiration	Expiration
b	Art. bucc.	Art. bucc.	Art. bucc.	Art. bucc.
c	[ ]	~~~~~	[ ]	.....
d	[ ]	[ ]	.....	.....

La colonne I désigne les sons *sourds*. II les sons *sonores*, III les sons sourds nasalisés, IV les sons sonores nasalisés.

Mais une inconnue subsiste : la nature de l'articulation buccale ; il importe donc d'en déterminer les variétés possibles.

Figure 10 The universal consciousness on the universal code of Nature: De Saussure's sound system of natural language can be seen as a specific Boolean logical square: I(00), III(01), II(10), IV(11) ... (cf. De Saussure's natural language designation system ref. [26], p. 70 and R. Swanson's genetic code designation system ref. [14], p. 10).

## 6. CONCLUSIONS

The periodicity and cyclicity within the periodic system of elements, genetic code, microtubule code and *Yin-Yang* system, are in accordance with periodicity and cyclicity of the natural number systems with the base  $N_1 = 2^n$  ( $n = 1, \dots, 6$ ) and  $N_2 = 2(2n+1)$  ( $n = 0, \dots, 5$ ). The relation within these mathematical number systems is the *Golden Mean*.

To understand the biophysical mechanism of information processes the main investigation should be done in both fields water coding system(s) and water-biomolecules interaction. There is a strong indication that water code(s) should be given by natural systems,  $N_1$  and/or  $N_2$ .

Our main hypothesis in establishing our approach is that investigation of consciousness must always consider the whole system as a unity of *mind/matter*. We have shown that *speculative mind* (*Yin-Yang*) and *matter* (periodic system of elements and genetic code) have the same coding system. The link between *Mind* (as quantum field entity of empty space - *pure vacuum*) and *Matter* (as mass organized entities) are microtubules and

their interaction with water. We believe, based on our knowledge, that human consciousness as mind/matter unity arises from this interaction. Considering this concept, together with Russel's idea about consciousness as a relation to an object in the sense of being conscious of something, we believe that human consciousness in some specific ways has expressed human being itself in the various works by different creative investigators through different epochs.

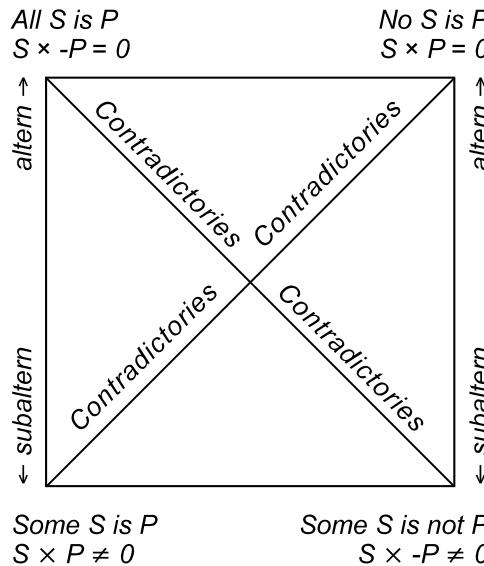


Figure 11 The universal consciousness on the universal code of Nature: Aristotle's syllogism "square of opposition" can be seen as specific Boolean logical square: No S is P (00), Some S is not P (01), Some S is P (10), All S is P (11) ... (cf. ref. [27], p. 341).

All the physiological (including biochemical and biophysical) processes of the human organism form the basis of human psychological and logical activities and all manifestations of consciousness. From the standpoint to be "conscious of" something, in this article we have shown that there exists a Universal consciousness about the universal code of Nature. This universal code as the basis of separate natural codes (chemical, genetic codes etc.) must be determined by Boolean spaces. Thus follows the hypothesis that human consciousness must also be determined by Boolean logical spaces.

This Boolean logical spaces concept of the universal consciousness must be provided by further investigation from different aspects. For example,

from the aspect of an existing or not existing accordance with the quantum physics concept [28] and information physics [9].

**Acknowledgments:** *This research is dedicated to soul of Petar Petroviæ Njegoš (1813-1851), who wrote the poem "Light of the Microcosm" by emerging logic and Golden Mean laws.*

## REFERENCES

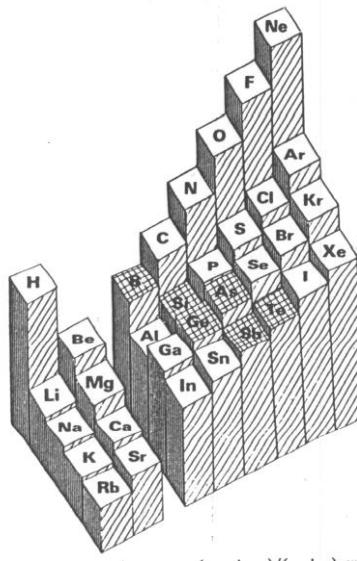
- [1] B.Russel, *An Outline of Philosophy* (Unwin, London, 1979).
- [2] L.S.Vygotskii, *Myshlenie i rech* (Moskva, 1956).
- [3] W.R.Ashby, *Design for a Brain* (Chapman & Hall, 1960).
- [4] M.A.Arbib, *Brains, Machines and Mathematics* (McGraw-Hill, New York, 1964).
- [5] J.-P.Sartre, *L'imaginaire* (Galimard, Paris, 1940).
- [6] B.Petronijeviæ, *Principien der Metaphysik* (Carl Winter, Heidelberg, 1904).
- [7] A.Einstein, *Sobranie nauchnih trudov I-IV* (Nauka, Moskva, 1965-1967).
- [8] F.Crick, *The Astonishing Hypothesis: The Scientific Search for the Soul* (Charles Scribner's Sons, New York, 1994).
- [9] D.L.Koruga, Neurocomputing and consciousness, *Neural Networks World* 1 (1991), pp. 32-38.
- [10] D.L.Koruga, Information physics: In search of scientific basis of consciousness, in Hameroff, Kaszniak, Scott, eds., *Toward a Scientific Basis for Consciousness* (MIT Press, Cambridge, 1995).
- [11] B.M.Kedrov, *Prognozi D.I.Mendeleeva v atomistike: neizvestnie elementi* (Atomizdat, Moskva, 1977).
- [12] E.S.Gould, *Inorganic Reactions and Structure* (Holt, New York, 1962).
- [13] M.M.Rakoèeviæ, The coherence of the chemical and genetic code, *Zbornik radova (Compendium) of the Faculty of Science Nis* 2 (1991), pp. 1-29.
- [14] M.M.Rakoèeviæ, *Logic of the Genetic Code* (Nauèna knjiga, Belgrade, 1994).
- [15] S.A.Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford Univ. Press, New York, 1993).
- [16] J.Maddox, New dimension for Mendeleev, *Nature* 356 (1992), p. 13.
- [17] Ch.Darwin, *The Origin of Species* (John Murray, London, 1859).
- [18] F.H.C.Crick, The genetic code yesterday, today and tomorrow, *Cold Spring Harbor Symposia on Quantitative Biology* 31 (1966), pp. 3-9.
- [19] F.H.C.Crick, The origin of genetic code. *J. Mol. Biol.* 3 (1968), pp. 67-379.
- [20] R.Swanson, A unifying concept for the amino acid code, *Bull. Math. Biol.* 46 (1984), pp. 187-203.
- [21] M.Rakoèeviæ, Information-topological concept of the amino acid code, *Bull. Fac. Sci. Nis* 1 (1990), pp. 3-23.
- [22] M.Schonberger, *The I Ching and Genetic Code*. (ASI, New York, 1980).

- [23] D.L.Koruga, Microtubular screw symmetry: Packing of spheres as a latent bioinformation code, *Ann. NY Acad. Sci.* 466 (1986), pp. 953-955.
- [24] R.Jimenez, G.R.Fleming, P.V.Kumar, and M.Maroncelli, Femtosecond solvation dynamics of water, *Nature* 369 (1994), pp. 471-473.
- [25] S.G.Stent, *The Coming of the Golden Age* (Freeman, New York, 1969).
- [26] F.Saussure de, *Cours de Linguistique Generale* (Payot, Paris, 1985).
- [27] S.K.Langer, *An Introduction to Symbolic Logic* (Dover, New York, 1953).
- [28] E.J.Squires, Quantum theory and the relation between the conscious mind and the physical world, *Synthese* 97 (1993), pp. 109-123.

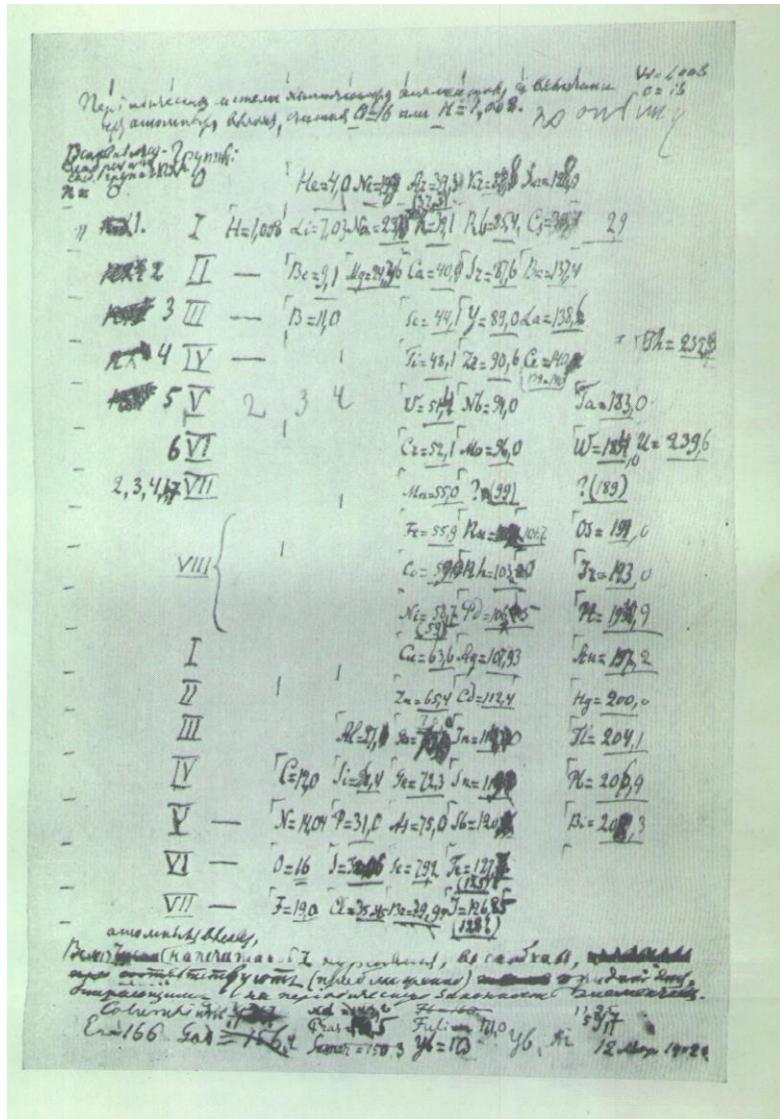
## **EPILOGUE**

## The Position Bioelements and Biomolecules Systems

The Mendeleev's periodic system is ultimately a position system (POSSYS): the system of the positions (SYSPOS) of the chemical elements within itself. Furthermore, ultimately a three-dimensional (more exactly, three-four-dimensional) position system, a LIGHT cube, i.e. a LIGHT cube-hypercube. However, if that is a three-dimensional periodic system, then that is no in the sense of Leland C. Allen's model (Allen, 1989, 1992) as here in Fig. 12, but in the sense of original Mendeleev's model as here in Fig. 13 in relation to Fig. 2 and Fig. 4 on the page 152 and 156 respectively (cf. Tables 18, 19 and 20) [Allen, 1989, p 9003: "It is argued that electronegativity is the third dimension of the Periodic Table"; Allen, 1992, p 1510: "It is clear that something is missing ... Configuration energy (CE), the average one-electron valence shell energy of a ground-state free atom, is the missing third dimension"]. Accordingly to Mendeleev, the same groups of the elements are at all three coordinates (dimensions) at the same time (Notice in Fig. 13, in relation to Fig 2 p 152, that noble gases are at the end in all three cases). That comes from the fact that the groups of elements must be at the vertices of the LIGHT cube.



**Fig. 12.** Electronegativity as a "third dimension" of the periodic system (After: Allen, 1989, p 9004).



**Fig. 13.** Mendeleev's three-dimensional periodic system as a position system (POSSYS), i.e. a system of the positions (SYSPOS) of the element groups on the eight vertices (0-7) of a LIGHT cube, in relation to Fig. 2, p 152 and to Fig. 4, p 156. A half volumen of the cube is realized with a line 2-3-4-5 and a whole volumen of the cube is realized with a line 2-3-4-6-7. By this notice a translation process: from the starting LIGHT-cube vector (0, 1), through the vectors (2, 3) and (4, 5) to the vector (6, 7), i.e. from the (000, 001), through (010, 011) and (100, 101) to the (110, 111). By this, the starting vector must be a basic fractal motive within LIGHT cube as a fractal system.

The physical, chemical and biological (physiological etc) properties of essential bioelements correspond to their positions within Mendeleev's periodic system in the manner as it is shown in Appendix 5 (pp 68-77). On the other hand, the physical, chemical and physiological properties of amino acid molecules must correspond to their positions within the PECO-DTAAS in Survey 7', p 102 and/or within the ECO-DTAAS in Survey 15, p 120; analogously, in Survey 8' and Survey 16, pp 104 and 120 respectively (cf. two classes of precursors in Table 13, p 106: I. *class*: less complex and II. *class*: more complex precursors, with two classes of enzymes in Figures 11-12, pp 112-113: I. *class*: less specific and II. *class*: more specific enzymes). The grounds and arguments are as follows.

He	Ne	Ar	Kr	Xe
Li	Na	K	Rb	Cs
Be	Mg	Ca 40	<del>Al = 87</del>	Ba 137
B	Al	In 70	<del>Li 114</del>	Sc 204
C	Si	Ge	In	Al
N	P	As	Se	Bi
O	S	Se	Te	
F	Cl	Br	I.	
30   37   77				

**Fig. 14.** Two Mendeleev's arithmetical "errors": 30/37/77 instead 30/27/67 (Fpr details see: Rakoèeviæ, 1994, pp 197-200).

In Appendix 9, p 118 are shown the relations between two systems, PECOS and ECOS, which represent a union, a specific POSSYS, i.e. SYSPoS of amino acid biomolecules. By this, the balance, except on the zig-zag lines, exists still through the divisions within two halves of the system. So, within *first half* as well as within the *second half* the PECOS and ECOS differ only for 01 atom and exactly for 10 nucleons; AN(**fh**): PECOS - ECOS = 94 - 93 and AN(**sh**): ECOS - PECOS = 143 - 142; NN(**fh**): ECOS - PECOS = 482 - 472 and NN(**sh**): PECOS - ECOS = 925 - 915.

**Note 1.** It is a noteworthy fact that one can compare these 01 - 10 situations with 01 - 10 Mendeleev's arithmetical "errors" in Fig. 14, in relation to Fig. 1 p 151. The possible questions are: Why Mendeleev wrote the atom masses in mentioned Fig. 1 only as integers, and why he made two errors for 10 in Fig. 14 - two errors in only three counting cases?! On the other hand, why and how Homer wrote the *Iliad* and *Odyssey* with validity of "minimum change relation" and/or "first - second" relation in a strict correspondence to the structure of the genetic code; moreover, how Njegoš wrote *Mountain Wreath* in a strict correspondence to the *Iliad* and *Odyssey*, and how both, Homer and Njegoš, wrote their works accordingly to a LIGHT cube and/or hypercube (see Survey 17 in relation to Fig. 9 p 165)? Abuot Njegoš's works see in Prvulovich, 1984.

**Note 2.** Notice that 3 & 6 distinctions within PECOS and ECOS correspond to the first perfect number (6); that 118 is the second perfect number (28) in module 9 ( $118 = 28 + 90$ ) and that 694 is an inversion of the third perfect number (496). Notice also that if amino acids placed on the full line, as well as on the dotted line, within ECOS (Survey 15), possess  $90 + 28$  atoms both, then amino acids handled by class I synthetases (bold positions) posses 28 atoms more whereas amino acids handled by class II synthetases (other positions) 28 atoms less. By this, NN in bold positions, together with its inversion permutation equals:  $861 + 168 = 1029$ ; in other positions:  $536 + 635 = 1171$ . If so, then the sum of tow last results is  $1029 + 1171 = 220 \times 10$  ten times the first frendly number, and the difference is  $1171 - 1029 = 284 : 2$  a half of the second frendly number.

**Note 3.** The eight triplets within PECOS/ECOS correspond to eight vertices within LIGHT cube: I - VIII to 0 - 7, i.e. to 000 - 111. Two and two zig-zag lines correspond too... If, on the fist (full) line of PECOS/ECOS there are  $694 + 0 = 694$  nucleons, then on the second (dotted) line there are  $694 + 9 = 703$  and vice versa. The result 703 corresponds to the second perfect number in a special manner. Namely, 703 is 19th case within the Shcherbak's modular table, Table 1 (Shcherbak, 1994, p 476):  $19 \times 037 = 703$ . On the other hand, the numbers 19 and 37 are the first neighbours to the 28 within the same congruent class in module 9: 19-28-37 in module 9 as 1-1-1. Notice that the sum of the LIGHT cube vertices equals exactly 28 ( $0 + 1 + \dots + 7 = 28$ ), what is the second perfect number.

## Survey 17 The verses number relations within *Iliad* and *Odyssey*.

<b>ODISSEY</b>	O)	<b>12110</b>				
			<i>139</i>			
			<i>3</i>			
<i>EVEN</i>	(I & O) - 1 =	<b>13503</b>		<i>109</i>		
				<i>5</i>		
			298		998	
<i>FIRST</i>	(I & O) - 0 =	<b>13801</b>		97		998
			201		0	0
<i>SECOND</i>	(I & O) + 0 =	<b>14002</b>		97		998
			298		998	
<i>ODD</i>	(I & O) + 1 =	<b>14300</b>		<i>109</i>		
				<i>5</i>		
			139			
			<i>3</i>			
<b>ILIAD</b>	(I	<b>15693</b>				
		<b>83409</b>			9959	
		<b>x 21</b>			<i>x 01</i>	
		<b>1751589</b>			9959	
						<b>1741630</b>

The both epics have 24 cantos each: 12 the first, 12 the second; 12 even and 12 odd. The final result 1741630 is the same as within the genetic code: "The number of all distributions in the set of 64 codons is 1741630" (Alvager et al., 1989, p 191; cf. Fig. 9, p 165 in this study). Notice that  $1751589 = 63 \times 27803$  (*Iliad plus Odyssey*); the 63 is the last point on the binary tree of the genetic code in Fig. 1, p 4 & Fig. 3, p 154 in this study; the 21 is the last point on the Floor-table of genetic code (Table 4.1, p 56 in Rakoëeviae, 1994).

The further relations, the balanced proportionalities of AN and NN within two systems, from the aspect to be first 3 or second 3 pairs, first 6 or second 6 members of the pairs, are presented in solutions 60-63 for AN and 60'-63' for NN:

$$\text{PECOS } (15 + 27) + (38 + 41) = 118 + 3 \quad (60)$$

$$\text{ECOS } (25 + 37) + (29 + 30) = 118 + 3 \quad (61)$$

$$\text{ECOS } (14 + 17) + (46 + 38) = 118 - 3 \quad (62)$$

$$\text{PECOS } (20 + 32) + (30 + 33) = 118 - 3 \quad (63)$$

$$\text{PECOS } (059 + 115) + (245 + 278) = 694 + 3 \quad (60')$$

$$\text{ECOS } (133 + 189) + (187 + 189) = 694 + 4 \quad (61')$$

$$\text{ECOS } (073 + 087) + (294 + 245) = 694 + 5 \quad (62')$$

$$\text{PECOS } (121 + 177) + (189 + 213) = 694 + 6 \quad (63')$$

As we can see the divisions into 3 & 3 pairs and 6 & 6 members of the pairs within 2 & 2 systems are followed by the distinctions for 3 & 3 atoms and 3 - 6 nucleons in relations to two zig-zag lines (full and dotted); by this “the minimum change relation” is valid for 2 & 2 systems: for 0 in AN and for 1 in NN. [Hint: If  $694 + 0 = 694$  is on the *first*, then  $694 + 9 = 703$  is on the *second* line; thus, the results in solutions 60' - 63' in relation to it are: 703 - 6, 703 - 5, 703 - 4, 703 - 3. From both lines follows an inverse symmetrical pattern: +3, +4, +5, +6 / -6, -5, -4, -3].

Within ECO-DTAAS in Survey 15 (analogously in Survey 16 too) exists a **hidden** ECO-DTAAS, a HECO-DTAAS. To discover this system one must go in several steps. First of all, two singlet lines are visible with the naked eye: I. G, S, T, P, A, H, F, D, N, K and II. V, C, M, I, L, W, Y, E, Q, R. In next step through a pairing process (each to each - first to first, second to second, etc.), the HECO-DTAAS is discovered: G-V, S-C / T-M, P-I / A-L, H-W, F-Y / D-E, N-Q, K-R, when our previous said rule comes to be self-evident: the first (less complex) amino acids are handled by class II (more specific) synthetases, whereas the second (more complex) amino acids are handled by class I (less specific) synthetases [Hint: the coplexity is related only within the pair, going from one pair to another].

For better understanding an additional discussion is needed. So, in our previous work we have shown that four stereochemical types of 20 protein amino acids exist through two classes: I. glycine type - alanine type and II. valine type - proline type. On the other hand we have also shown that the pair valine - isoleucine is a *link pair* between cyclic and non-cyclic as well as between *inner* and *outer* system within DTAAS. Now, we see that this pair makes also a link between two classes of four stereochemical types. If so, then it is a sense for a G-V pair because a similar inductive effect - electron density (IE-ED) side chain influence to  $\text{C}^\alpha$  atom exists (a symmetrical H atom in comparison with a symmetrical isopropyl group). Further, if the linkages V-I and V-P are the facts (in the discussed sense), then a P-I linkage must also exist!

And now, the last step in additional discussion. Originally, from the aspect “to be a system”, there are not 10, but 9 amino acid pairs built from only four bioelements: C, N, O, H, all four in a strong neighbourhood positions within Mendeleev’s periodic system (O and H are the neighbours through a diagonal link if H is within the 7<sup>th</sup> group). An additional 10<sup>th</sup> pair

with the 5<sup>th</sup> atom within itself will not break the original system if, and only if, two next conditions are realized: 1. the additional 5<sup>th</sup> bioelement must be the first neighbour to some from previous four bioelements (for example sulphur as a “vertical” neighbour to oxygen); if it not, then its first neighbour (if not sulphur, then selenium); 2. the “coming” additional pair can go into the original system if, and only if, a new additional pairing process is possible: minimum one pair must exist, within the original system, which can give, through the said pairing process, two new pairs; for example, the pair S-T together with the “coming” pair C-M gives the pairs: S-C and T-M.

I hope that with this additional discussion it is clear that the HECO- DTAAS is a reality. If so, then my hypothesis, given in the title of Appendix 9, has obtained a proof. This proof, on the other hand, can help in future new researches to obtain a valid answer for two Wetzel’s questions: “... the synthetases coevolved with the genetic code”, or “the synthetases evolved in the context of an already-established genetic code - a code which developed earlier in an RNA world ?” (Wetzel, 1995, p 545); for De Duve’s hypothesis of “the existence of a second genetic code, imprinted into the structure of aminoacyl-tRNA synthetases” (De Duve, 1988, p 117), and for hypothesis of several other authors of the existence of “an operational RNA code for amino acids” (Schimmel, 1995; Hipps et al., 1995); also, a valid answer to the question in how manner “the common origin of any pair of tRNAs for two different amino acids precedes the age of the present universal genetic code” (Jukes, 1995, p 537); finally, to obtain or not a satisfying proof for the hypothesis after that “the most primitive code is ... a GC code: GG coding for glycine, CC coding for proline, GC coding for alanine, CG coding for ‘arginine’”, in which case “the aminoacyl-tRNA synthetase was a copper-montmorillonite” (Hartman, 1995, p 541).

In all mentioned cases, following my prediction (*Prediction 17*), based on the characteristics of PECOs, ECOs, HECOs and PRACENZ system, the said future researches will ultimately show that genetic code was “from the beginning” a strictly arranged regular whole system with 4 *pu-pyr* bases and 20 amino acids, as a condition for the origin of the life; that genetic code was in the very first step of life such as it is today, such generated, not degenerated in any evolution process. On the other words, the future researches will show that it is no a sense to speak about the evolution of the genetic code, but only about the evolution of macromolecules and living beings. The future researches will also show that physical, chemical and biological (physiological etc.) properties of amino acid molecules must

correspond to their positions within PECOS and ECOS (*Prediction 20*); and that, on the other hand, PECOS and ECOS must be the *repères* (*Prediction 19*) for all other biomolecules systems (BIMOSs).\*

---

\* Remind that glycine side chain, corresponding to a free hydrogen atom, is the simplest possible *étau* of IE-EI for all non - cyclic molecules. On the other hand, proline side chain, corresponding to a free cyclopropane molecule, is the simplest possible *étau* of IE-EI for all cyclic molecules.

**Table 18. The periodic system of chemical elements (short periods model)**

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 <sub>1</sub>	He <sub>2</sub>							
2	2	a	Li <sub>3</sub>	Be <sub>4</sub>	B <sub>5</sub>	C <sub>6</sub>	N <sub>7</sub>	O <sub>8</sub>	F <sub>9</sub>	Ne <sub>10</sub>								
3	3	a	Na <sub>11</sub>	Mg <sub>12</sub>	Al <sub>13</sub>	Si <sub>14</sub>	P <sub>15</sub>	S <sub>16</sub>	Cl <sub>17</sub>	Ar <sub>18</sub>								
4	4	a	K <sub>19</sub>	Ca <sub>20</sub>	b	Sc <sub>21</sub>	Ti <sub>22</sub>	V <sub>23</sub>	Cr <sub>24</sub>	Mn <sub>25</sub>	Fe <sub>26</sub>	Co <sub>27</sub>	Ni <sub>28</sub>					
5	b	Cu <sub>29</sub>	Zn <sub>30</sub>	a	Ga <sub>31</sub>	Ge <sub>32</sub>	As <sub>33</sub>	Se <sub>34</sub>	Br <sub>35</sub>	Kr <sub>36</sub>								
6	5	a	Rb <sub>37</sub>	Sr <sub>38</sub>	b	Y <sub>39</sub>	Zr <sub>40</sub>	Nb <sub>41</sub>	Mo <sub>42</sub>	Tc <sub>43</sub>	Ru <sub>44</sub>	Rh <sub>45</sub>	Pd <sub>46</sub>					
7	b	Ag <sub>47</sub>	Cd <sub>48</sub>	a	In <sub>49</sub>	Sn <sub>50</sub>	Sb <sub>51</sub>	Te <sub>52</sub>	I <sub>53</sub>	Xe <sub>54</sub>								
8	a	Cs <sub>55</sub>	Ba <sub>56</sub>	b	La <sub>57</sub>	c	Ce <sub>58</sub>	Pr <sub>59</sub>	Nd <sub>60</sub>	Pm <sub>61</sub>	Sm <sub>62</sub>	Eu <sub>63</sub>	Gd <sub>64</sub>	Tb <sub>65</sub>	Dy <sub>66</sub>	Ho <sub>67</sub>	Er <sub>68</sub>	
9	6	c	Tm <sub>69</sub>	Yb <sub>70</sub>	c	Lu <sub>71</sub>	b	Hf <sub>72</sub>	Ta <sub>73</sub>	W <sub>74</sub>	Re <sub>75</sub>	Os <sub>76</sub>	Ir <sub>77</sub>	Pt <sub>78</sub>				
10	b	Au <sub>79</sub>	Hg <sub>80</sub>	a	Tl <sub>81</sub>	a	Pb <sub>82</sub>	Bi <sub>83</sub>	Po <sub>84</sub>	At <sub>85</sub>	Rn <sub>86</sub>							
	a	Fr <sub>87</sub>	Md <sub>101</sub>	Ra <sub>88</sub>	b	Ac <sub>89</sub>	c	Th <sub>90</sub>	Pa <sub>91</sub>	U <sub>92</sub>	Np <sub>93</sub>	Pu <sub>94</sub>	Am <sub>95</sub>	Cm <sub>96</sub>	Bk <sub>97</sub>	Cf <sub>98</sub>	Es <sub>99</sub>	Fm <sub>100</sub>
	c			No <sub>102</sub>	a	Lr <sub>103</sub>	a	Ku <sub>104</sub>	Ns <sub>105</sub>	106	107	108	109	110				
	b																	

The groups 0-14 equals 15 groups; the 1st group must appear still as a 15th group within a **LIGHT** hypercube. On the left: 1-10 the rows and 1-7 the periods. (After: Rakoèeviæ, 1991, p 20).

**Table 19. The periodic system of chemical elements (long periods model).**

The designations as in Table 18 (After: Rakoëviæ, 1991, p 21).

**Table 20. The periodic system of chemical elements (“cylinder” model)**

DIADS														MONAD													
I							TRIADS							II							II						
I			II		III		I			II		III		I			II		III		I			II		III	
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	Ae	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				

Within the brackets it is the number of the natural nuclides. For Polonium, Po, only five nuclides can be in this “system game”. Notice the law for TRIADS:  $Si(3)+P(1) = S(4)$  etc.(the difference is zero); then:  $La(2)+Ce(4) = Ba(7)-1$  (the difference is one); then:  $Sm(7)+Eu(2) = Gd(7)+2$  or  $Gd(7)+Eu(2) = Sm(7)+2$  (the difference is two).

## LITERATURE

Allen, L. C. (1989). Electronegativity is the Average One-Electron Energy of the Valence-Shell Electrons in Ground-State Free Atoms. *J. Am. Chem. Soc.* 111, 9003-9014.

Allen, L. C. (1992). Extension and Completion of the Periodic Table. *J. Am. Chem. Soc.* 114, 1510-1511.

Alvager, T. et. al. (1989). On the Information Content of the Genetic Code. *Bio Systems* 22, 189-196.

Attardi, G. (1985). Animal Mitochondrial DNA: an Extreme Example of Genetic Economy. *Int. Rev. Cytol.* 93, 93-145.

Caspari, W. E. (1968). The Genetic Code After the Excitement. *Advances in Genetics*, Vol. 14, 325-390.

Champe, P.C. and R. A. Harvey. (1994). Biochemistry. Philadelphia: *Lippin Cott.*

Crick, C. H. F. (1966a). The Genetic Code Yesterday, Today and Tomorrow. *Cold Spring Harbor Symposia on Quantitative Biology*. 31, 3-9.

Crick, C. H. F. (1966b). Codon - Anticodon Pairing: The Wobble Hypothesis. *J. Mol. Biol.* 19, 548-555.

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

Dayhoff, M. O. (1969). Computer Analysis of Protein Evolution. *Scientific American* 221 (1), 86-96.

Dayhoff, M. O. (1972-1978). Atlas of Protein Sequence and Structure, Vol. 5 and Supplements 1, 2 and 3. Washington, DC: *National Biomedical Research Foundation*.

Dayhoff, M. O., Orcutt, B. C. (1985). Scoring Matrices. PIR Report Document Data Protein Identification Resource. *National Biomedical Research Fundation*, Washington.

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

Dickerson, R. E. (1972). The Structure and History of an Ancient Protein. *Scientific American* 22, 189.

Dickerson, R. E. (1978). Chemical Evolution and the Origin of Life. *Scientific American*, Vol. 239, No. 3, 70-86.

Doolittle, R. F. (1981). Similar Amino Acid Sequences: Chance or Common Ancestry. *Science*, 214, 149-157.

Doolittle, R. F. (1985). Proteins Are the Molecules Encoded by Genes. *Scientific American*, Vol. 253, No. 4, 74-85.

Doolittle, R. F. and J. Kyte. (1982). A Simple Method for Displaying the Hydropathic Character of a Protein. *J. Mol. Biol.* 157, 105-132.

Douglas, E. B. et al. (1994). Concepts and Models of Inorganic Chemistry. New York: *John Wiley*.

Dubinin, N. P. (1985). Genetics (in Russian). Kishinev: *Shtiintsa*.

Dugas, H. and Ch. Penney. (1981). Bioorganic Chemistry - a Chemical Approach to Enzyme Action. New York, Berlin: *Springer-Verlag*.

Eck, R. V. (1963). Genetic Code: Emergence of a Symmetrical Pattern. *Science* 140, 477-480.

Eichhorn, G. (1993). Conformational Change Induced by Metal Ions Through Coordination. *Coordination Chemistry Review* 128, 167-173.

Eigen, M. and P. Schuster. (1979). The Hypercycle, a Principle of Natural Self-Organization. Berlin: *Springer-Verlag*.

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

Falconer, K. (1990). Fractal Geometry. *John Wiley*. New York.

Fitch, W. M., Margoliash, E. (1967). Construction of Phylogenetic Trees. *Science*, 155, 279-284.

Frömmel, C. and G. H. Holzhütter. (1985). An Estimate on the Effect of Point Mutation and Natural Selection on the Rate of Amino Acid Replacement in Proteins. *J. Mol. Evol.* 21, 233-257.

Frieden, E. (1972). The Chemical Elements of Life. *Scientific American*, 227, 52-60.

Gamow, G. (1954). Possible Relation Between DNA and Protein Structure. *Nature* 113, 318.

Gatlin, L. L. (1972). Information Theory and the Living Sistem. *Columbia University*, New York.

Gavrilov, P. G. and A. A. Sapozhenko. (1989). Selected Problems in Discrete Mathematics. Moscow: *Mir*.

Gilyarov, M. S. editor (1989). Biologicheskii encyclopedicheskii slovar. *Sovetskaya encyclopediya*, Moscow.

Grafstein, D. (1983). Stereochemical Origins of the Genetic Code. *J. Theor. Biol.* 105, 157-174.

Hartman, H. (1995). Speculations on the Origins of the Genetic Code. *J. Mol. Evol.* 40, 541-544.

Helene, C. (1987). The Structure of DNA in: *Genetika i naslednost*. *Mir*, Moscow.

Hilton, P. and J. Pedersen. (1989). Duality and the Descartes Deficiency. *Computers Math. Applic.* 17, 73-88.

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

Horowitz, S. and A. M. Gorovsky. (1985). An Unusual Genetic Code in Nuclear Genes of Tetrahymena. *Proc. Natl. Acad. Sci. USA* 82 (8), 2452-5.

Jatsimirskii B. K. (1976). Vvedenie v bioneorganicheskuy himiy. *Naukova dumka*, Kiev.

Johannsen, W. (1909). Elemente der exakten Erblichkeitslehre. *Gustav Fischer*, Jena.

Johannsen, W. (1913). Elemente der exakten Erblichkeitslehre. Second Edition. *Gustav Fischer*, Jena.

Jukes, T. H. (1963). The Genetic Code. *Amer. Scientist* 51, 127-284.

Jukes, T. H. (1966). Molecules and Evolution. New York, London: *Columbia University Press*.

Jukes, T. H. (1973). Possibilities for the Evolution of the Genetic Code from a Preceding Form. *Nature*. 246, 22-27.

Jukes, T. H. (1983). Mitochondrial Codes and Evolution (News). *Nature*. 301, (5895), 19-20.

Jukes, T. H. (1995). A Comparison of Mitochondrial tRNAs in Five Vertebrates. *J. Mol. Evol.* 40, 537-540.

Kauffman, S. A. (1993). The Origins of order: Self-organization and selection in Evolution. *Oxford University Press*. New York.

Kimura, M. (1968). Evolutionary Rate at the Molecular Level. *Nature*. 217, 624.

King, J. L. and T. H. Jukes. (1969). Non-Darwinian Evolution. *Science*, Vol. 164, 788-797

Konopka, A. K. and V. Brendel. (1985). The Missense Errors in Protein Can be Controlled by Selective Synonymous Codon Usage at the Level of Transcription. *Biochimie*. 67, 469-473.

Koruga, D. (1995). Information Physics: In Search of a Scientific Basis of Consciousness, in: Consciousness - A Scientific Challenge of the 21<sup>st</sup> Century. *ECPD*, Belgrade.

Koruga, D. L. (1992). Neuromolecular Computing. *Nanobiology*, 1, 5-24.

Koruga, D. L. et al. (1993). Fullerene C<sub>60</sub>-History, *Physics, Nanobiology, Nanotechnology*. North Holland, Amsterdam.

Koruga, D.L. (1991). Neurocomputing and Consciousness. *International Journal on Neural and Mass-Parallel Computing and Information Systems*. 1, 32-38

Kuchino, Y. et al. (1985). Tetrahymena Thermophila Glutamine tRNA and Its Gene that Corresponds to UAA Termination Codon. *Proc. Natl. Acad. Sci. USA* 82, 4758-4762.

Lagerkvist, U. (1978). Two Out of Three: an Alternative Method for Codon Readings. *Proc. Natl. Acad. USA* 75, 1759-1762.

Lagerkvist, U. et al. (1981). Codon Reading and Translational Error. *The Journal of Biological Chemistry* 256, 2635-2643.

Lehninger, A. L. (1982). Principles of Biochemistry. *The Johns Hopkins University School of Medicine*, Baltimore.

Lenissen, M. A., De Jong, W. W. (1986). Phylogenetic trees constructed from hydrophobicity values of protein sequences. *J. theor. Biol* 119, 187-196.

Lewin, B. (1987). Genes. Third Edition, New York: *John Wiley & Sons Inc.*

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

Martynenko, I. L., Spitsyn, I. V. (1983). Methodological Aspects of the Course in Inorganic Chemistry. *Mir*, Moscow.

Mathematical encyclopedic dictionary. (1988). Matematicheskii encyclopedicheskii slovar edited by Yu. B. Prohorov. *Sovetskaya encyclopediya*, Moscow.

McClain, H. W. and J. G. Seidman (1987). Genetic Conversion of G.C Base-Pairs in a Transfer RNA. *J. Molec. Biol.* 197, 605-608.

Mendel, G. (1866). Versuche über Pflanzenhybriden. *Verhandl. Des Naturvorschenden Vereins in Brunn*, IV.

Moore, G. A. (1993). The Limit of the Golden Numbers is 3/2. *The Fibonacci Quarterly*, 31.4, 354-364.

Moore, G. A. (1994). The Limit of the Golden Numbers is 3/2. *The Fibonacci Quartely*. June-Juy, 211-217.

Orgel, L. E. (1986). RNA Catalysis and the Origin of Life. *J. Theor. Biol.* 123, 127-149.

Osawa, S. et al. (1992). Recent Evidence for Evolution of the Genetic Code. *Microbiological Reviews* 56, 229-264.

Pflug, D. A. (1984). Early Geological Record and the Origin of Life. *Naturwissenschaften*. 71, 63-68.

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

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

Porschke, D. (1985). Differential Effect of Amino Acid Residues on the Stability of Double Helices Formed from Polyribonucleotides and Its Possible Relation to the Evolution of the Genetic Code. *J. Mol. Evol.* 21 (2), 192-198.

Prat, A. et al. (1986). Nucleotide Sequence of the Paramecium Primaurelia. *J. Mol. Biol.* 189, 47-60.

Prvulovich, Ž. R. (1984). Prince-Bishop Njegosh's Religious phylosophy. Birmingham: ŽRP.

Rackovsky, S. and H. A. Scheraga. (1977). Hydrophobicity, Hydrophilicity and the Radial and Orientational Distributions of Residues in Native Proteins. *Proc. Natl. Acad. Sci. USA* ju, 5248-5251.

Rakoèeviæ, M. M. (1988). Genes, Molecules, Language (in Serbian with an English Language Supplement). Belgrade: *Nauèna knjiga*.

Rakoèeviæ, M. M. (1990). Information - Topological Concept of the Amino Acid Code. *Compendium of the Faculty of Science in Niš* 1, 3-23.

Rakoèeviæ, M. M. (1991). The Coherence of the Chemical and Genetic Code. *Compendium of the Faculty of Science in Niš* 2, 1-29.

Rakoèeviæ, M. M. (1994). Logic of the Genetic Code (in English). Belgrade: *Nauèna knjiga*.

Rakoèeviæ, M. and Jokiæ, A. (1996). Four Stereochemical Types of Protein Amino Acids: Synchronic Determination with Chemical Characteristics, Atom and Nucleon Number. *J Theor. Bil.*

Reuben, J. and F. E. Polk. (1980). Nucleotide - Amino Acid Interactions and Their Relation to the Genetic Code. *J. Molec. Evol.* 15, 103-112.

Rowe, G. W. and H. E. L. Trainor. (1983b). On the Informational Content of Viral DNA. *J. Theor. Biol.* 101, 15-170.

Rumer, Yu. B. (1966). About the Codon's Systematization in the Genetic Code (in Russian). *Proc. Acad. Sci. USSR: Doklady*. 167, 1393.

Rumer, Yu. B. (1966). About the Codon's Systematization in the Genetic Code (in Russian). *Proc. Acad. Sci. U.S.S.R. Doklady*. 167, 1393.

Saenger, W. (1984). Principles of Nucleic Acid Structure. New York, Berlin: *Springer-Verlag*.

Sanger, F. et al. (1981). Sequence and Organization of the Human Mitochondrial Genom. *Nature* 290, 457-465.

Schimmel, P. and Ya. M. Hou. (1988). A Simple Structural Feature Is a Major Determinant of the Identity of a Transfer RNA. *Nature* 333, 140-145.

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. (1989). Rumer's Rule and Transformation in the Context of the Cooperative Symmetry of the Genetic Code. *J. Theor. Biol.* 139, 241-401.

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.

Shulz G. E., Schirmer R. H. (1979). Principles of protein Structure. New York.

Stakhov, A. P. (1989). The Golden Section in the Measurement Theory. *Computers Mth. Applic.*, 17, 613-638.

Swanson, R. (1984). A Unifying Concept for the Amino Acid Code. *Bull. Math. Biol.* 46 (2), 187-207.

Taylor, F. J. R., Coates, D. (1989). The Code within Codons. *BioSystems* 22, 177-187.

Taylor, R. W. (1986). The Classification of Amino Acid Conservation. *J. Theor. Biol.* 119, 205-218.

Topal, D. M. and J. R. Fresco. (1976a). Complementary Base Pairing and the Origin of Substitution Mutations. *Nature* 263, 285-289.

Topal, D. M. and J. R. Fresco. (1976b). Base Pairing and Fidelity in Codon-Anticodon Interaction. *Nature* 263, 289-294.

Van Nostrand's (1983). *Scientific Encyclopedia*. 6<sup>th</sup>. Ed. New York.

Verkhovod, A.B. (1994). Alphanumerical Divisions of the Universal Genetic Code: New Divisions Reveal New Balances. *J. Theor. Biol.* 170, 327-330.

Voet, D. and J. G. Voet, (1990). Biochemistry, New York: *John Wiley & Sons*.

Watson, J. D. and F. H. C. Crick. (1953a). Molecular Structure of Nucleic Acids. A Structure of Deoxyribose Nucleic Acid. *Nature* 171, 737-738.

Watson, J. D. and F. H. C. Crick. (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

Yamao, F. et al. (1985). UGA Is Read as Tryptophan in Mycoplasma Capricolum. *Proc. Natl. Acad. Sci. USA* 82, 2306-2309.

-

# Contents

1. INTRODUCTION .....	1
2. THE HYPOTHETICAL FRAMEWORKS .....	3
3. GENETIC CODE AS A BOOLEAN SPACE .....	5
3.1. Mathematical and Physical Models.....	6
3.2. Boolean Square Relations .....	8
3.3. The Alphabet and the Words.....	11
4. GENETIC CODE NUCLEON NUMBER .....	14
4.1. Union of Chemistry, Physics and Boolean Arithmetic .....	14
4.2. Amino Acid Systematization .....	19
5. HIERARCHY OF BOOLEAN SPACES .....	20
6. PARAMETRIC RELATIONS .....	22
6.1. The Three Rings .....	22
6.2. Codon-Anticodon and Codon-Amino Acid Relations .....	30
6.3. Boolean Spaces Through Hydropathy and Cloisteredness .....	35
7. FINAL COMMENTS .....	39
 Generation of Cyclic Amino Acid System.....	49
Arithmetical regularities within the tRNA molecules .....	53
Perfect and friendly numbers .....	60
About the number 037 .....	65
Gold and Golden Section as Determinants of Bioelements Choice .....	68
Four Stereochemical Types of Protein Amino Acids:	
Synchronical Determination with Chemical Characteristics, Atom and Nucleon Number (I) .....	78
Fractal Structure of Amino Acid (Genetic) Code .....	82
Four Stereochemical Types Of Protein Amino Acids:	
Synchronical Determination With Chemical Characteristics, Atom and Nucleon Number (II) .....	101
Four Stereochemical Types of Protein Amino Acids in Correspondence with Two Classes of Enzymes .....	119
 A D D E N D A .....	122
THE LIMIT OF THE GOLDEN NUMBERS $\frac{1}{2}$ .....	123
Sixty-four Triplets and 20 Canonical Amino Acids of the Genetic Code: the Arithmetical Regularities. Part II.....	124
INFORMATION PHYSICS: IN SEARCH OF A SCIENTIFIC BASIS	

OF CONSCIOUSNESS.....	126
1. INTRODUCTION.....	127
2. FUNDAMENTALS OF INFORMATION PHYSICS.....	128
2.1 Information theories and physics .....	128
2.2 Unit spheres .....	129
2.3 Unit spheres packing and coding .....	132
2.4 Information Physics.....	133
3. NEUROMOLECULAR COMPUTING .....	135
3.1 DNA as one-dimensional information entity .....	135
3.2 Proteins as other side of the DNA code .....	135
4. MODELS FOR BIOLOGICAL BASIS OF CONSCIOUSNESS	139
4.1 Model of sub-consciousness.....	139
4.2 Model of consciousness .....	142
4.3. Model of self-consciousness .....	142
5. TESTING PROPOSED MODELS OF CONSCIOUSNESS .....	143
5.1 Biomedical testing.....	143
5.2 "Artificial brain" model testing .....	144
6. CONCLUSION .....	144
REFERENCES .....	146
 THE UNIVERSAL CONSCIOUSNESS AND THE UNIVERSAL CODE.....	149
1. INTRODUCTION.....	149
2. PRELIMINARIES .....	150
3. LOGIC OF MOLECULES OF LIFE .....	154
4. LOGIC OF SPECULATIVE MIND .....	162
5. FINAL COMMENTS .....	165
6. CONCLUSIONS .....	168
REFERENCES .....	170
 EPILOGUE .....	172
The Position Bioelements and Biomolecules Systems .....	173

On the basis of what we have presented, on the basis of the discussion given in every one of the six chapters, as well as all integral discussion, the inevitable conclusion is that all the working hypotheses given in chapter 2 have been proved. The general hypothesis, according to which the Boolean spaces are actually the main determinants and invariants of the genetic code, has been therefore proven. The Boolean spaces have been shown to be reality-models! From this it further follows that it makes no sense to talk about the neutrality of mutations, or about a non-Darwinian evolution...

On the other hand, the genetic code, being redundant, did not become degenerate in the process of evolution, but was generated in origin as such. It is today as it was in the beginning and it will remain so in the future anywhere in the universe, because that follows from the positions of the bioelements in the periodic system of elements;; bioelements - being the constituents of the genetic code. Accordingly, it makes no sense to talk about the evolution of the genetic code, but it does make sense to talk about the evolution of the macromolecules, that is, the evolution of life which came into being on the basis of just such a code - a universal genetic code.

How has the genetic code become "from the beginning" that what it was: why with those bases and that exact number of bases; why with those amino acids and why with that exact number of amino acids; with exactly 3 "stop" codons in the alphabet which functions on the level of words, with exactly one termination situation in the alphabet which functions on the level of letters? All this follows from a strict determination by nucleon number presented in chapter 4 and a strict determination based on physical and chemical parameters presented in chapter 6. The genetic code, therefore, must be universal for life which could exist anywhere in the universe (whit this the answer to question 1 from the Introduction has been given). But it not universal in the sense that there is one-meaning correspondence between the words of one alphabet and the letters of the other alphabet in all cases. On the contrary, for the largest number of cases the correspondence is really one of one-meaning (strong), but in a number of cases there must be a deviation from one-meaning, and that by two possible levels: middle and weak.