Symmetry: Culture and Science
Vol. 10, Nos. 3-4, p. 211-226, 1999

GENETIC CODE AND THE ANCIENT CHINESE *BOOK OF CHANGES*

Sergei V. Petoukhov

Physicist, biologist, mathematician, Doctor of Physical-Mathematical Sciences, 1988; Candidate of Biological Sciences, 1973, (b. Moscow, Russia, 1946).

Address: Department of Biomechanics, Mechanical Engineering Research Institute, Russian Academy of Sciences, M. Kharitonyevsky pereulok 4, Moscow, Russia 101830, E-mail: petoukhov@hotmail.com.

Fields of interest: biosymmetry, self-organizing systems, geometry, solitons, genetic code, biomechanics.

Awards: State Prize Award, 1986.

Publications: (1981) *Biomechanics, Bionics and Symmetry*, Moscow (in Russian); (1988*) Geometries of Living Nature and Algorithms of Self-organization*, Moscow (in Russian); (1999) *Biosolitons. The Bases of Soliton Biology*, Moscow (in Rusian); (2001) *The Bi-periodic Table of Genetic Code and Number of Protons*, Moscow (in Russian); Non-Euclidean geometries and algorithms of living bodies - *Computers & Mathematics with Applications*, (1989) 17, 4-6, 505-534; Symmetric-algorithmic properties of regular biostructures, *Symmetry: Culture and Science*, 1 (1990), 3, 295-312; Highest biosymmetries and the concept of replication morphogenesis, *Symmetry: Culture and Science*, 3 (1992), 1, 60-61.

«C'est peut-être I Ching qu'il faudrait étudier
pour saisir les relations entre hérédité et langage»
by F. Jacob, the Nobel prize winner on molecular genetics (Jacob, 1974, p. 205)

«Perhaps, for seizing of relations between genetics and language it would be necessary to study namely through the Ancient Chinese *Book of Changes*»

This article is devoted to genetic language and its parallels with symbolical system of Ancient Chinese "The Book of Changes" ("I Ching"), written a several thousand years ago according to some estimations (see for example Wilhelm, 1960).

The research is based on symmetrological analysis of system of biochemical elements used by a nature in genetic coding. The study of symmetries in molecular genetics has a traditional character and it is one of the basic methods of knowledge in this field.

For example, the analysis of symmetries has revealed the major fact of complementation of the nitrogenous bases of a code in a molecule of DNA (deoxyribonucleic acid), structure of this molecule as a regular double spiral, and also many other important things of molecular genetics.

A fact of amazing simplicity and generality of the basic principles of coding of the hereditary information in living organisms belongs to number of major scientific discoveries of mankind in past century. Now it is well-known that for realization of the genetic messages, which encodes sequences of amino acids in proteins, all kinds of organisms use the "alphabet" consisting of only four "letters" or nitrogenous bases (Figure 1): adenine (A), cytosine (C), guanine (G), thymine (T) {or uracil (U) in RNA – ribonucleic acid}. Linear sequences of these four letters on strings of molecules of heredity (DNA and RNA) contain the genetic information for protein synthesis in all living bodies - from bacteria up to a whale and from worm up to a bird and man. The given set of four letters is usually considered as the elementary alphabet of a genetic code. The modern science does not know why the alphabet of genetic language has four letters (it could have any other number of the letters in principle) and why just these four nitrogenous bases are chosen by nature as elements of the genetic alphabet from billions possible chemical connections.

The starting point of the original research, stated in the article, was the following. The author paid attention to the fact that these four nitrogenous bases represent specific poly-nuclear constructions with the special biochemical properties. The set of these four constructions is not completely heterogeneous, and it has on itself a substantial system of attributes for uniting and distinguishing of its letters (or, more precisely, the set has pairs of "attribute – anti-attribute"). This system of attributes divides the four-letter alphabet into various pairs of letters, which are equivalent from a viewpoint of one of these attributes or its absence. This situation is illustrated below for a four-letter set A, C, G, U, which is explored for discussing in molecular genetics traditionally and is appropriated for RNA.

BINARY SUB-ALPHABETS AND THREE PARALLEL LANGUAGES

Two of these letters C and U are equivalent of each other from the viewpoint of presence of an attribute of their belonging to a chemical class of pyrimidines (they have one ring in their molecular structure) and two other letters - A and G - are equivalent

each other from viewpoint of absence of this attribute (they concern to a chemical class of purines which have two rings). A relation of equivalence between letters will be marked by symbol "=" below.

The other two pairs of equivalent letters, where A=C and U=G, are formed according to another attribute, which is a property of amino-mutating of two nitrogenous bases – A and C - in RNA under action of nitrous acid HNO_2. This property is connected with detachment of amino-group NH_2 from molecules of these two bases (Wittmann 1961), where it is located in one of six tops of a molecular ring just opposite to atom of nitrogen. The other two bases U and G do not have the property of amino-mutating and do not have such a located amino-group; so they are equivalent from viewpoint of absence of this attribute.

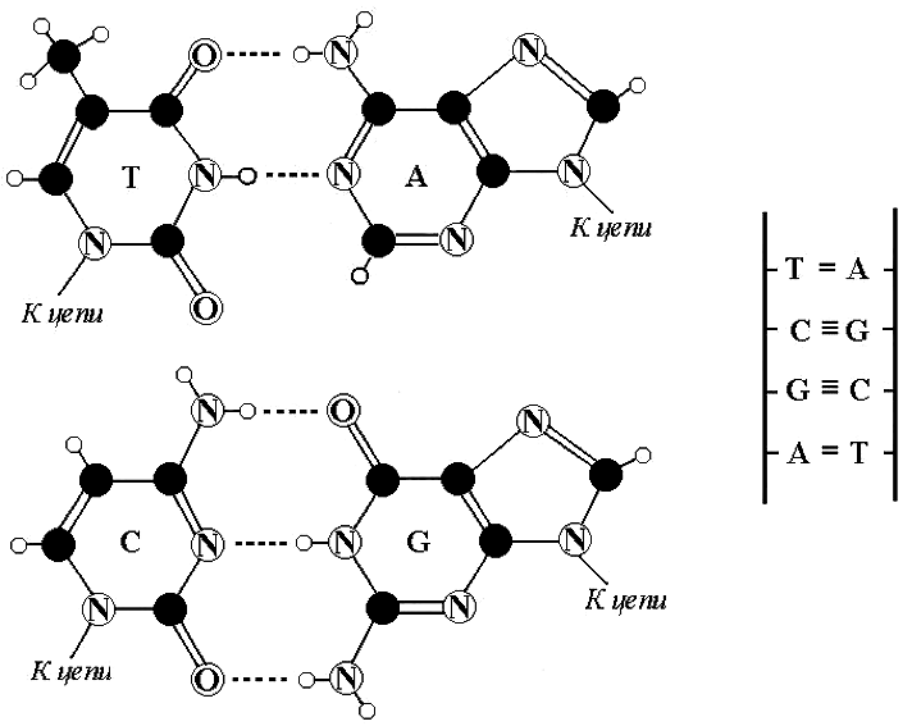


Figure 1: Complementary pairs of four nitrogenous bases in DNA: A - T (adenine and thymine), C - G (cytosine and guanine). By a dotted line are specified hydrogen bonds in these pairs. Black circles are atoms of carbon, small white circles - hydrogen, circles with the letter N - nitrogen, and circles with the letter O – oxygen.

Third kind of pairs of equivalent letters is formed on the basis of the attribute of complementation of these nitrogenous bases in molecules of nucleic acids: C=G (they form complementary pair with three hydrogen bonds between them) and A=U (they form complementary pair not with three, but with two hydrogen bonds). The symbol of binary opposition can be conditionally ascribed to each letter in the described cases: symbol “1” - in a case of presence of the considered attribute at the letter, and symbol “0” - in a case of its absence. The “elementary” four-letter alphabet of genetic code comprises three binary sub-alphabets according to three kinds of biochemical attributes (Petoukhov 2001, 2002]. Table 1 illustrates this fact (indexes at symbols “0” and “1” are identical to number of an attribute type).

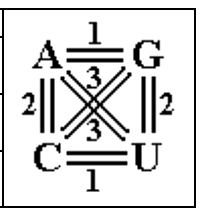
ATTRIBUTE	**G**	**A**	**U**	**C**	
1) Belonging to pyrimidine class (with one ring in the molecule)	0_1	0_1	1_1	1_1	
2) Amino-mutating (or special location of NH_2 in molecular ring)	0_2	1_2	0_2	1_2	
3) Belonging to complementary pair with three hydrogen bonds	1_3	0_3	0_3	1_3	

Table 1

So, each letter of the code alphabet has three “faces” or meanings in three binary sub-alphabets in connection with the three kinds of attributes. The four-letter alphabet of a code is curtailed into the two-letter alphabet on the basis of each kind of attributes. For example, to the first attribute we have (instead of the four-letter alphabet) the alphabet from two letters O_1 and 1_1, which the author names as “the binary sub-alphabet to a first attribute”. Accordingly, the genetic text as a sequence of four letters of genetic language is submitted by three parallel and various sequences of zero and unit. It illustrates a following example of an initial piece of a genetic sequence for insulin (Table 2).

Attribute number	Sequence of “letters” in a gene of insulin
	A U G G G C A U C G U U G A A C A G U G U **.....**
№ 1	0_1 1_1 0_1 0_1 0_1 1_1 0_1 1_1 1_1 0_1 1_1 1_1 0_1 0_1 0_1 1_1 0_1 0_1 1_1 0_1 1_1 .....
№ 2	1_2 0_2 0_2 0_2 0_2 1_2 1_2 0_2 1_2 0_2 0_2 0_2 0_2 1_2 1_2 1_2 1_2 0_2 0_2 0_2 0_2 **.....**
№ 3	0_3 0_3 1_3 1_3 1_3 1_3 0_3 0_3 1_3 1_3 0_3 0_3 1_3 0_3 0_3 1_3 0_3 1_3 0_3 1_3 0_3 **.....**

Table 2

Figuratively speaking, the genetic text appears as a bunch of parallel texts in three different languages, and genetic sequences have a property of poly-languages. This property was missed earlier from view. The sense of the genetic message, probably, is enclosed in an interlacing or simultaneous use of these multilingual texts. Each letter of the genetic language in view of its meaning in three binary sub-alphabets can be expressed by a code vector in three-dimensional Cartesian space; each coordinate of this space corresponds to one of the binary sub-alphabets.

It is well-known that a set of three letters of a genetic language forms a triplet (or a codon), which encodes one of 20 amino acids or one of punctuation marks of protein synthesis. Due to three binary sub-alphabets, any genetic triplet can be presented by a set of three code vectors or by numerical matrix (3x3). Determinants of these matrixes are equal to "0" (at 40 triplets), "+1" (at 12 triplets) and "-1" (at 12 triplets). It's obvious that various mathematical structures are existed in connection with this ensemble of the binary sub-alphabets and genetic texts from their binary elements. Formal theory of such structures composes a separate subject of mathematical research.

The ensemble of the binary sub-alphabets of genetic language is a new system, it has not been investigated in biology earlier. This ensemble has, according to data of the author, important meaning for understanding of all biological scheme of a storage and transfer of the hereditary information. Just as in physics the concept of an "elementary" particle (as particle which does not consist of parts) appears insolvent at revealing that this particle actually contains component parts, so the concept of the "elementary" alphabet of molecular genetics appears transient in light of the fact that this alphabet consists of the substantial binary sub-alphabets. Binary sub-alphabets bring together a living nature and computer technology: both fields use a principle of a storage and transfer of the information on the basis of sequences of "0" and "1".

OCTET FAMILIES

Let us consider the set $64=4^3$ genetic triplets (it is possible to make several various triplets on the basis of the four-letter alphabet). Due to attributes, which transform the four-letter alphabet of genetic language into the binary sub-alphabets, internal structure and well-ordering of the triplets set are revealed. Octet bi-periodic table of codons (Table 3), constructed by the author, demonstrates it. All triplets are renumbered by a natural manner and are located in tabular system with double – on a horizontal and

vertical – periodicity and with secretive fractal structure (described in Petoukhov /2001/ in detail).

№	111 CHYAN	110 TUI	101 LI	100 CHEN	011 HSUN	010 KAN	001 KEN	000 KUN
111 CHYAN	CCC 63	CCA 62	CAC 61	CAA 60	ACC 59	ACA 58	AAC 57	AAA 56
110 TUI	CCU 55	CCG 54	CAU 53	CAG 52	ACU 51	ACG 50	AAU 49	AAG 48
101 LI	CUC 47	CUA 46	CGC 45	CGA 44	AUC 43	AUA 42	AGC 41	AGA 40
011 HSUN	UCC 31	UCA 30	UAC 29	UAA 28	GCC 27	GCA 26	GAC 25	GAA 24
100 CHEN	CUU 39	CUG 38	CGU 37	CGG 36	AUU 35	AUG 34	AGU 33	AGG 32
010 KAN	UCU 23	UCG 22	UAU 21	UAG 20	GCU 19	GCG 18	GAU 17	GAG 16
001 KEN	UUC 15	UUA 14	UGC 13	UGA 12	GUC 11	GUA 10	GGC 9	GGA 8
000 KUN	UUU 7	UUG 6	UGU 5	UGG 4	GUU 3	GUG 2	GGU 1	GGG 0

Table 3: The bi-periodic table of genetic triplets

All 64 triplets are submitted in 64 cells of this table. Any column consists of 8 triplets equivalent of each other from the viewpoint of the binary sub-alphabet of the first attribute from Table 1. In other words, triplets of one column are indiscernible from each other in relation of positions engaged in them by pyrimidines and purines. For example, in second column of Table 3 (from the left), two pyrimidines (C or U) are on first two positions of all its triplets and one purine (A or G) is on its third position. According to Table 1, binary symbol 1_1 corresponds to pyrimidines, and binary symbol 0_1 corresponds to purines. In this connection, each column together with all its eight triplets receives an individual binary-numerical symbol from the following series of eight three-digit numbers: $0_10_10_1$, $0_10_11_1$, $0_11_10_1$, $0_11_11_1$, $1_10_10_1$, $1_10_11_1$, $1_11_10_1$, $1_11_11_1$ (in the Table 2 the indexes are omitted). While translating from binary in decimal numeration system, this series will be transformed to a series of integers 0, 1, 2, 3, 4, 5, 6, 7. In our Table 3, all columns with their families of triplets are put in descending order under these numbers. It is obvious, that in a binary system it is possible to use not only symbols of “0” and “1”, but also, for example (as it was made in Ancient Chinese

"The Book of Changes"), a symbol of broken line instead of "0" and a symbol of unbroken line instead of "1". According to Chinese manner, these lines can be put in vertical trigrams which should be read from below upwards (see Table 3). Then each from eight tabular families of triplets will receive the appropriate symbol of Ancient Chinese trigram, given in the Table 3 with its Chinese name.

These octet families of triplets are known in a history of molecular genetics due to the work (Wittmann 1961) on fragmentation of all set of triplets into eight families in connection with property of amino-mutating of nitrogenous bases in RNA, i.e., in connection with a second attribute from Table 1. Table 3, offered by us, represents an association of these Wittmann's octets in a family of eight tabular columns with their ordering according to the natural binary numeration. It is essential, that we construct Table 3 on the basis of real biochemical materials.

Each row of the Table 3 has also the binary number connected to an attribute of amino-mutating (an attribute №2 in Table 1) in triplets of a line. All eight triplets of each row are equivalent each other from the viewpoint of the binary sub-alphabet connected with this second attribute of Table 1. For example, its second row (from above) is characterized by a symbol $1_2 1_2 0_2$, i.e. all triplets of the row have identical quantity and order of the letters with these binary symbols 1_2 and 0_2. (In Table 3 the binary symbols are printed without indexes, but the symbols of rows are underlined and are printed with italics to distinguish them from symbols of columns). The sequence of rows in Table 3 is built in the decreasing order of their natural binary numbers and this order is broken only for two average rows, which are rearranged by their places in connection with biochemical features of Wittmann's octets.

In Table 3 each triplet has coordinates from a symbol of its row and symbol of its column. Both of these symbols can be written out jointly as a single whole six-digit binary number (for example, let us accept, that coordinate three-digit number of a tabular row goes in the beginning of this six-digit number, and coordinate number of a column goes in its end). Then all triplets will be renumbered with six-digit binary numbers. At translation of these binary numbers into decimal numeration, well-ordered series of numbers from 0 up to 63 is formed. All 64 triplets in the appropriate cells of Table 3 are numerated by numbers of this series. In this natural system of ordering of triplets, numerical invariant exists, which characterizes all pairs of "codon-anticodon" (codon and anticodon refer to triplets, which are located opposite each other in pairs of complementary strings of molecules DNA or RNA). This invariant is a sum of coordinate numbers of any codon and its anticodon, and this sum is equal to 63 always.

And, on the contrary, if the sum of coordinate numbers of any two triplets in Table 3 is not equal to 63, these triplets are not a pair of "codon-anticodon".

For example, codon CUA and its anticodon GAU have coordinate numbers 46 and 17 accordingly, and their sum is equal to 63. In addition to this, each codon and its anticodon are located in cells, which are inverse-symmetrical in relation of the tabular center.

THE TABLE OF HEXAGRAMS AND "THE BOOK OF CHANGES"

At use for coordinates of triplets in Table 3 not digital but Chinese trigram forms, each triplet receives an individual coordinate symbol as an Ancient Chinese hexagram: a pile of six broken and unbroken lines. This vertical set of six lines for any triplet has at its bottom the coordinate trigram of tabular row of the triplet and has on top a coordinate trigram of its tabular column. At replacement of each triplet by its coordinate hexagram, Table 3 is transformed into the table, which coincides wonderfully with the historically famous table of 64 hexagrams from "The Book of Changes" (Table 4) in the order of *Fu-Xi*. This book is a unique monument of world culture and philosophical thought, the first one in the row of sacred books of the East (see, for example, Wilhelm 1960). According to statements of Chinese sources during thousands years, this ancient table is general natural archetype. It is very interesting that the biochemical system of a genetic code is connected with this Ancient Chinese symbolical system. The detailed consideration of this analogy is carried out in the special chapter to our book (Petoukhov 2001, chapter 5), where information on other researchers of similar parallelisms is given also.

A lot of other interesting numerical and other analogies between genetic system and "The Book of Changes" have come to light during author's researches. Only a few of them will be described in this small article. Let us remember that the concept of numbers had a completely special meaning in Ancient China. One of the most important places was devoted to the four of numbers 6, 7, 8, 9. Even numbers 6 and 8 symbolized accordingly old (variable or mutable) and young (not variable) Yin, and odd numbers 9 and 7 - old (variable) and young (not variable) Yang.

Another important one was number 60. It was the basis of the universal 60-nary cyclic system, on which many Chinese systems were constructed: the 60-years' Chinese calendar, 60 tune keys of traditional Chinese music, etc. (Eremeev 1993).

TOP→ BOTTOM↓	111 CHYAN	110 TUI	101 LI	100 CHEN	011 HSUN	010 KAN	001 KEN	000 KUN
111 CHYAN	111111 63	111110 62	111101 61	111100 60	111011 59	111010 58	111001 57	111000 56
110 TUI	110111 55	110110 54	110101 53	110100 52	110011 51	110010 50	110001 49	110000 48
101 LI	101111 47	101110 46	101101 45	101100 44	101011 43	101010 42	101001 41	101000 40
100 CHEN	100111 39	100110 38	100101 37	100100 36	100011 35	100010 34	100001 33	100000 32
011 HSUN	011111 31	011110 30	011101 29	011100 28	011011 27	011010 26	011001 25	011000 24
010 KAN	010111 23	010110 22	010101 21	010100 20	010011 19	010010 18	010001 17	010000 16
001 KEN	001111 15	001110 14	001101 13	001100 12	001011 11	001010 10	001001 9	001000 8
000 KUN	000111 7	000110 6	000101 5	000100 4	000011 3	000010 2	000001 1	000000 0

Table 4: A square arrangement of hexagrams according to *Fu-Xi*'s order from "The Book of Changes" with supplement for each hexagram of its binary-numerical value and its equivalent in a decimal numeration.

These numbers are connected to Table 3 by an interesting way. Really, let's recollect a third attribute of Table 1, according to which the pair of complementary nitrogenous bases C and G has three hydrogen bonds, and other pair of complementary bases A and U has two hydrogen bonds. Let us replace each triplet in Table 3 by the sum of these numbers of its hydrogen bonds, considering C=G=3, A=U=2. For example, the triplet CAU will be replaced by number 7 (=3+2+2). In result the octet numerical matrix, submitted in the Table 5, will be produced. The sums «Σ» of number of considered

hydrogen bonds for all triplets of each row and of each column of Table 3 are represented in Table 5, too.

All rows and all columns of this matrix differ from each other by sequences of their numbers. But the sums of all hydrogen bonds in cells of each row and of each column are identical each other and are equal to that number 60, which has been considered in connection with Chinese numbers ideas already. The number 60 is used at the account of time also: 60 minutes in an hour, 60 seconds in a minute, etc.

Those numbers 6, 7, 8 and 9, which symbolized young and old Yin and Yang in Ancient China, appear in cells of the given matrix only. The certain laws are observed in their location in the Table 3. Only a few of them will be noted here.

	111 CHYAN	110 TUI	101 LI	100 CHEN	011 HSUN	010 KAN	001 KEN	000 KUN	Σ
111 CHYAN	9	8	8	7	8	7	7	6	60
110 TUI	8	9	7	8	7	8	6	7	60
101 LI	8	7	9	8	7	6	8	7	60
011 HSUN	8	7	7	6	9	8	8	7	60
100 CHEN	7	8	8	9	6	7	7	8	60
010 KAN	7	8	6	7	8	9	7	8	60
001 KEN	7	6	8	7	8	7	9	8	60
000 KUN	6	7	7	8	7	8	8	9	60
Σ	60	60	60	60	60	60	60	60	480

Table 5: A numerical matrix of the bi-periodic table of codons with a system of sums of hydrogen bonds of triplets in cells (C=G=3, A=U=2).

For additional obviousness, double frames mark tabular cells with number 9, and bold frames mark cells with number 6. These two numbers are located in cells on diagonals

of a matrix only, and numbers 7 and 8 - outside of diagonals only. The diagonals divide a matrix in four sectors (top, right, bottom and left) consisting from cells with numbers 7 and 8. Each of sectors contains in its cells the sum of numbers equal to 90, and all four sectors – the sum 4x90=360. Each of these sectors contains six pairs of mirror-symmetrically located cells with numbers 7 and 8 in each pair, giving number 15 in the sum for each pair. All four sectors contain 24 pairs, the sum of numbers in which is equal to 24x15=360. Here analogy to traditional Chinese representations about a calendar of solar year takes place: the Chinese solar year contains 360 days and four seasons with 90 days each (in more fractional Chinese view, solar year consists of 24 phases with 15 days in each) (Too 1993). These and other analogies make it very probable that the numerical matrix of the Table 5 was used in Ancient China as one of the basic classification-numerical systems, under which many representations about the organization of the whole world were built.

So called "center-summable magic squares" (3x3) with numbers in their cells are known in connection with "The Book of Changes". Each of pairs of diagonal-opposite peripheral cells of such square has an identical sum, which is equal to a number in its central cell. The genetic table 5 has the direct relation to such squares. Really, let's write eight numbers of any tabular row (and any its column) into peripheral cells of square (3x3) according to classical rule of Fu-Xi (Figure 2). Then each of pairs of its peripheral cells will have an identical sum - 15, which can be written into the central cell.

TUI	CHYAN	HSUN
LI		KAN
CHEN	KUN	KEN

Pattern of Fu-Xi

8	9	8
8	***15***	7
7	6	7

First row

9	8	7
7	***15***	8
8	7	6

Second row

7	8	7
9	***15***	6
8	7	8

Third row

Figure 2: Examples of center-summable magic squares for three upper rows of Table 5.
From the left: the pattern of Fu-Xi for square location of numbers from rows and columns of Table 5.

MOSAIC VISUALIZATION OF ANCIENT SYMBOLS

Let us consider a location of cells with figure 7 in Table 5. For this purpose these cells will be painted over with dark color, and all numbers will be cleaned in all cells with the exception of four central ones. For the better presentation of arising symbolics more dark color is used for painted cells in one half of table, than in another. The result it is illustrated in Table 6.

Number **69** is visualized in this mosaic unexpectedly. This number is constituted with blacker figure **6** from the left half of the table and with brighter figure **9** from the right half. Pictures **6** and **9** are inverse-symmetrical each other. Triplets from tabular cells, which are included in structure of the figure **6**, are anticodons in relation to triplets of those cells, which form the figure **9**. These figures are associated with numbers of hydrogen bonds 6 and 9 in the center of the table. It is interesting, that numbers 6 and 9, which are visualized in this mosaic of bi-periodic table of genetic code, were the main traditional symbols Yin and Yang in Ancient China. It testifies about deep connection of biogenetic system with symbolical system of “The Book of Changes” additionally. It is possible to believe, that so habitual form of figures 6 and 9 is not invented arbitrary by our ancestors, but is built in a nature deeply and therefrom is somehow transferred by the man into written language. These materials are connected with concept of Carl G. Jung about archetypes or instinctive figures (see for example Pauli 1952).

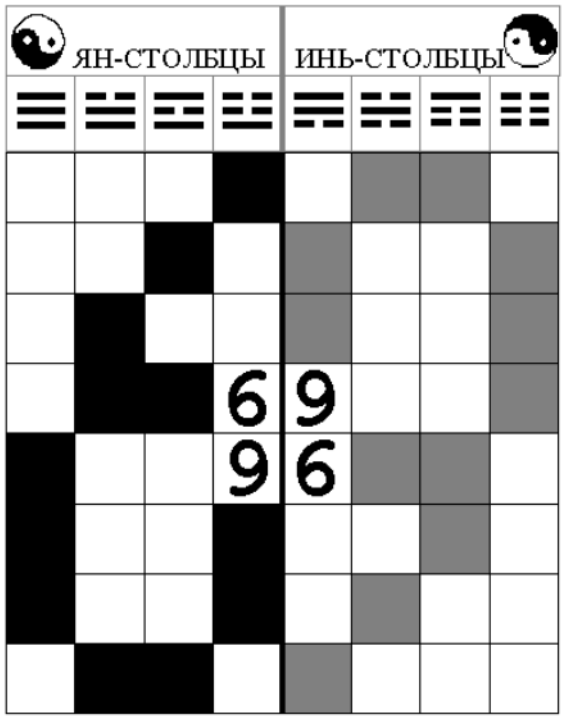


Table 6: Yin columns and Yang columns. A mosaic of cells, triplets of which have 7 hydrogen bonds, in the genetic table 5.

A similar mosaic of tabular cells with number not 7 but 8 (these cells are located in mirror-symmetrical positions concerning cells with number 7 in the Table 5) visualizes a mirror image of the figure **69**. The location of two other numbers 6 and 9 in the Table 5 is connected with a Buddhist cross (swastika) which is ancient East symbol of kind destiny and well-being. This symbol was known also as an emblem of the Christ up to the end of the Middle Ages.

A more profound consideration of the known data about genetic system from the viewpoint of the binary sub-alphabets and the bi-periodic table of genetic code (in particular, on specificity of tabular locations of 20 amino acids and of punctuation marks of protein synthesis which are encoded by 64 triplets) reveals the whole hierarchy of binary systems of Yin-Yang type at different levels of genetic system. The bi-periodic table of genetic code is connected with certain dichotomy Yin-Yang trees of binary sub-alphabet elements. It gives new opportunities for the analysis of genetic phenomena on the basis of the theory of fractals and synergetic. Symmetrological

analysis of described genetic structures in connection with ideas of unitary symmetries is very interesting also (Petoukhov 2001).

The number of variants of location of 64 triplets in octet tables is equal 64! or 10^{89} approximately. It's unimaginably huge number. The bi-periodic octet Table 3, constructed by the author in accordance with determined biochemical data, is unique variant, which reveals natural ordering in set of triplets and demonstrates new structural properties of genetic system.

NUMBER OF PROTONS

The rest of the article will be devoted to results of researches of system-organizing role of protons in genetic coding and bioorganics. The attention of the author was attracted to protons and their quantities due to following circumstances.

First, so important for genetic molecules the hydrogen bonds are based on one-proton atom of hydrogen. Secondly, carbon, nitrogen, oxygen and amino-group NH_2, which are the basic components of molecular genetic elements, have 6, 7, 8 and 9 protons in their atoms accordingly. In other words, the living substance is constructed on those nuclear formations, the numbers of protons in which are equal to Ancient Chinese numerical symbols of old and young Yin and Yang. (Amino group NH_2, presented in this list, is connected to the second attribute of Table 1 and determines of amino-mutating properties of nitrogenous bases of genetic code. Besides that, the amino-group is a foundation for function of recognition of amino acid by a ferment). Phosphorus and sulfur, which are met in molecular genetic elements rather seldom (for example, no other amino acid contains phosphorus and only two from twenty amino acids contain sulfur), contain in their atoms 15 (=7+8=6+9) and 16 (=8+8) protons accordingly.

Thirdly, the protons are the system-organizing factor in the Mendeleev's Periodic table of chemical elements. Serial numbers of all chemical elements in this Periodic table are determined only by the number of protons in atoms. The number of protons sets a value of an electrical charge of a nuclear nucleus of elements. Whether is spread not only up to the world of atoms, but also (although partially) up to the world of biological molecules, including genetic molecules? Whether has of sense to consider the named molecules how the poly-proton constructions, constructed on unified proton-numerical modules? Fourthly, in molecular biology it is well-known an important role of protons for maintenance of many biochemical processes, for example, fermentative processes,

which are realized in living bodies with such high speeds which are unattainable in conditions of modern laboratories.

How are the proton groupings submitted in elements of a genetic code? It is known that carbon, nitrogen, oxygen and hydrogen, which have numbers of protons 6, 7, 8, 1 in the atoms accordingly, dominate in structures of the majority of biological molecules or even define them completely in many cases. Let us try to visualize a structure of four nitrogenous bases submitted in Figure 1. For this purpose new – proton-numerical – symbolism will be used in habitual formal schemes of molecules. Instead of usual symbols of atoms (in a form of circles with the letters inside) new symbolism will represent each non-hydrogen atom by number of protons of its nucleus. For example, atom of nitrogen will be represented by number 7. If one or several atoms of hydrogen are joined to some non-hydrogen atom, we shall add protons of hydrogen to number of protons of non-hydrogen atom always. Such group of the non-hydrogen atom with its adjoined hydrogen atoms (i.e., "protonated" non-hydrogen atom) will be denote by sum of protons of their nucleuses. For example, amino group NH_2 will be denoted by number 9 which is sum of 7 protons of nitrogen atom and 2 protons of two hydrogen atoms. In such schemes the atoms of hydrogen are not represented independently at all, and their presence at a molecule simply increases general number of protons of the atoms, connected to them directly. The received thus proton-numerical portraits of four nitrogenous bases of genetic code are submitted in Figure 2.

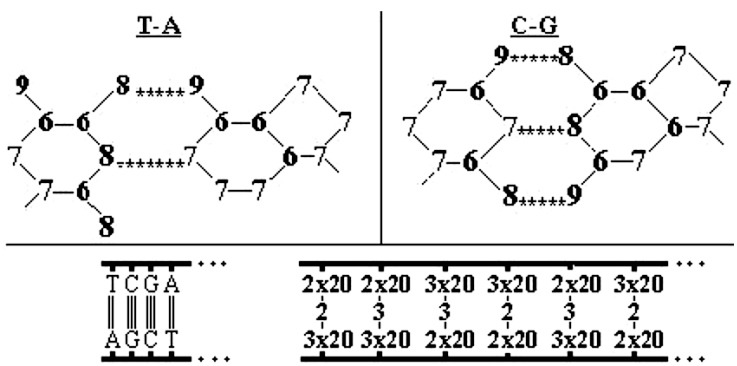


Figure 2: On top: proton-numerical representation of complementary pairs of nitrogenous bases in DNA (compare to Figure 1). At bottom: proton-numerical representation of a sequence of complementary pairs of bases in DNA as a sequence of numbers of hydrogen bonds in the given pairs (average row made up on basis of 2 and 3) and as a numerical sequence of protons at "protonated" atoms of molecules rings of these nitrogenous bases (top and bottom rows made up on basis of 2x20 and 3x20).

It is interesting that proton-numerical portraits of these code elements include only the four numbers 6, 7, 8 and 9, which are Ancient Chinese symbols of old and young Yin and Yang. The proton-numerical portraits of many other bioorganic molecules, including molecules of amino acids, DNA and ATF, give a similar picture of their connection with these four proton-numerical modules (see these portraits in Petoukhov 2001). It reveals a new - proton-numerical - area of unification in living substance.

Let us note in addition, that the sum of protons of rings of nitrogenous bases is equal to 40 at each of pyrimidine bases (C, T, U) and is equal to 60 at each of purine bases (A and G). Thus, in molecules of a heredity (DNA and RNA) each complementary pair of bases bears in its rings 100 protons (=40+60) exactly. The ratio 40:60 is equal to the ratio 2:3, which characterizes distinction of hydrogen bonds in the same complementary pairs of bases. From the viewpoint of number of hydrogen bonds in these pairs and of protons number at rings of the bases, the molecule of DNA can be presented as a stairs with steps, which are made up on basis of numbers 2, 3, 2x20, 3x20 (Figure 2, at the bottom). These and other facts force the author to suppose existence of some “arithmetic of protons” in genetic system.

Coordination of number of protons is present not only in a structure of separate molecules, but in entire system of triplets also. For example, if total number of protons of all atoms of any triplet of DNA is put into the bi-periodic table No. 3 instead of this triplet, then interesting regular numerical matrix will be generated (Petoukhov 2001, p. 143). In particular, this matrix has a numerical invariant for all 32 pairs of “codon-anticodon”: any of such pair has identical sum of protons of all its atoms and this sum is equal to 408 (in spite of the fact that number of protons at different nitrogenous bases in DNA differs: A has 70 protons, C - 58, G - 78, T - 66).

It is necessary to continue consecutive search of parallels between biogenetic system and symbol-numerical system of “The Book of Changes” which has played an extraordinarily important role in the development of science and culture in the East. The investigation results, which are described in this paper, have many applications for further analysis of hierarchic genetic system with a set of 20 amino acids; for understanding of genetic sequences; for consideration of problem of genetic code origin; for studying of biological evolution, etc. Some of them are lighted in the book (Petoukhov 2001), a summary and contents of which are at the Internet site http://members.tripod.com/vismath/sg/petoukhov.htm. On this way of investigation one can meet very unexpected problems. For example, an interesting formal analogy exists between described concept of binary sub-alphabets and the theory of quarks in physics

of elementary particles (this analogy permits to introduce a new notion "genetic alphabetical quarks" or "genoquarks" for developing of corresponding theory of genetic structures). Described investigation of genetic code gives an opportunity to borrow wonderful ideas of modern group theory physics into biology, in particularly, for modeling of some aspects of genetic structures on the base of theory of unitary symmetries. The author hopes to publish other thematic results in this symmetrological journal later.

ACKNOWLEDGMENT

I am grateful to Academician K. V. Frolov for permanent support in the researches.

REFERENCES

Eremeev, V. E. (1993) *Design of Anthropocosmos*, Moscow, 153 (in Russian).

Jacob, F. (1974) Le modele linquistique en biologie, *Critique, Mars, tome* XXX, № 322, 306-315.

Pauli, W. (1952) Der Einfluss archetypisher Vorstellungen auf die Bildung naturwissenschaftlicher Theorien bei Kepler, *Natureklarung und Psyche*, Berlin, S.109-194.

Petoukhov, S.V. (2001) *The Bi-periodic Table of Genetic Code and Number of Protons*, Foreword of K. V. Frolov, Moscow, 258 (in Russian).

Petoukhov, S. V. (2002) Binary sub-alphabets of genetic language and problem of unification bases of biological languages, *IX International Conference "Mathematics, computer, education"*, Russia, Dubna, January 28-31, 191 (in Russian).

Too, L. (1993) *Applied Pa-Kua and Lo Shu Feng Shui*, Hong–Kong, Oriental Publication, 253.

Wilhelm, H. (1960) *Change: Eight Lectures on the I Ching*, N.Y.

Wittmann, H.G. (1961) Ansatze zur Entschlusselung des genetishen Codes. – *Die Naturwissenschaften*, B.48, 24, S. 55.