Symmetry and Asymmetry of The Ultimate Genetic Code Matrix

Jean-Yves BOULAY

Jean-Yves BOULAY independent researcher – FRANCE – jean-yvesboulay@orange.fr ORCID: 0000-0001-5636-2375
https://www.researchgate.net/profile/Jean-Yves-Boulay

Abstract: *From the table of the standard genetic code at 64 codons, 61 amino acids and 3 Stop signals, It is invested a simplified matrix of the genetic code comprising only twice twenty entities: the twenty canonical proteinogenic amino acids associated with only twenty DNA biplets. By this process, a singular organization of this matrix is highlighted in both symmetrical and asymmetrical arrangements. Also, the configurations of this matrix are organized in coincidence with the alphanumeric system introduced in a recent published article by the author. This compressed version of the genetic code is called "the ultimate genetic code matrix."*

Keywords: genetic code, symmetry, amino acids, fractal, nucleobases, set theory, molecular biology, number theory.

1. Introduction

Today, it is now firmly established that living matter is organized via a called “universal” genetic code and that this genetic code encodes only, and very precisely, twenty proteinogenic amino acids. This number is not arbitrary, it is equal to 5*x*. More precisely this number of 20 entities is equal to 3*x* + 2*x* entities with a value of *x* equal to 4.

As we will demonstrate over the different parts of this study of the genetic code, it turns out that different components that make up the set of twenty (5*x*) proteinogenic amino acids are also to 5*x* (so 3*x* + 2*x*) entities. Also, it is the same considering this time different coding characteristics (arrangement and physico-chemical properties of DNA nucleobases) linked to these twenty amino acids. We will also demonstrate that these coding characteristics, organized around values equal to 5*x*, are intimately related to the physico-chemical properties of the twenty encoded amino acids.

The genetic code consists of sixty-four coding entities and sixty-four coded entities. Thus, sixty-four triplets of DNA nucleobases encode sixty-one amino acids and three stop signals. These different values are therefore not equal to 5*x*.

We propose here to construct and study a compressed matrix of the genetic code. To do this, we will therefore match only a single codon to each of the twenty proteinogenic amino acids and ignore the nonsense information (stop signal).

To this end, we will isolate just the first codon corresponding to each AA* in a defined order from the entire array of the universal genetic code and ignore the third base of the DNA triplet, transforming it into a two-nucleobase biplet. We thus obtain a matrix of the genetic code of 5*x* coding entities associated with 5*x* coded entities, i.e. 20 biplets and 20 codons. We call this matrix "the ultimate genetic code."

Having thus isolated these two times twenty entities, we observe that they are organized symmetrically and asymmetrically according to their own physico-chemical characteristics within this ultimate matrix. This is observed in many mapped geometric configurations.

In preview published paper “*Numbering of the twenty proteinogenic amino acids and new alphanumerical nomenclature proposal to them”* [1], we introduced a concept of numbering the twenty proteinogenic amino acids and we have demonstrated that a large number of different amino acid attributes arrange themselves numerically in exact 3/2 value ratios according to this numbering system. Recalling here this new concept of amino acid numbering, we will also demonstrate that the concept of the ultimate genetic code coincides with this AA numbering.

This present article is therefore a continuation and a complement to that [1] introducing the concept of AA numbering.

** To simplify, in some parts of text and tables, AA (or AAs) is used to replace amino acid appellation.*

2. Matrix of the universal genetic code

Before introducing the concept of the ultimate genetic code, we recall here the structure of the genetic code called as universal. So a universal mechanism recognized as the main biological system coding the twenty proteinogenic amino acids (themselves standard) that the living matter widely uses.

Figure 1 illustrates this universal genetic code where 64 triplets (codons) encode 61 amino acids and three stop signals. This genetic code is presented here in a matrix of 4 by 4 boxes where each of the twenty proteinogenic amino acids can be present several times. In each of these sixteen boxes, there are four DNA triplets with the first two identical nucleobases. In this matrix, the four different bases are introduced in the order A-G-T-C. The full importance of this classification will be gradually introduced in the next demonstrations.

This matrix is the introductory support for this paper investigating the genetic code from a innovative angle of study.

AAA	Lys	GAA	Glu	TAA	-	CAA	Gln
AAG	Lys	GAG	Glu	TAG	-	CAG	Gln
AAT	Asn	GAT	Asp	TAT	Tyr	CAT	His
AAC	Asn	GAC	Asp	TAC	Tyr	CAC	His
AGA	Arg	GGA	Gly	TGA	-	CGA	Arg
AGG	Arg	GGG	Gly	TGG	Trp	CGG	Arg
AGT	Ser	GGT	Gly	TGT	Cys	CGT	Arg
AGC	Ser	GGC	Gly	TGC	Cys	CGC	Arg
ATA	Ile	GTA	Val	TTA	Leu	CTA	Leu
ATG	Met	GTG	Val	TTG	Leu	CTG	Leu
ATT	Ile	GTT	Val	TTT	Phe	CTT	Leu
ATC	Ile	GTC	Val	TTC	Phe	CTC	Leu
ACA	Thr	GCA	Ala	TCA	Ser	CCA	Pro
ACG	Thr	GCG	Ala	TCG	Ser	CCG	Pro
ACT	Thr	GCT	Ala	TCT	Ser	CCT	Pro
ACC	Thr	GCC	Ala	TCC	Ser	CCC	Pro

Figure 1: Matrix of the universal genetic code (in base order A ➞ G ➞T ➞ C).

3. The ultimate genetic code table

From the matrix of the universal genetic code (Figure 1), we are now constructing what we will call "the ultimate genetic code".

3.1. The ultimate genetic code table construction

From the matrix of 64 double entities of the universal genetic code, we will isolate just twenty double entities corresponding quite simply to the twenty proteinogenic amino acids.

To this purpose, we use a mechanism for filling in the initial table of the universal genetic code by retaining only one triplet for each of the twenty amino acids. Also we do not consider stop signals (nonsense).

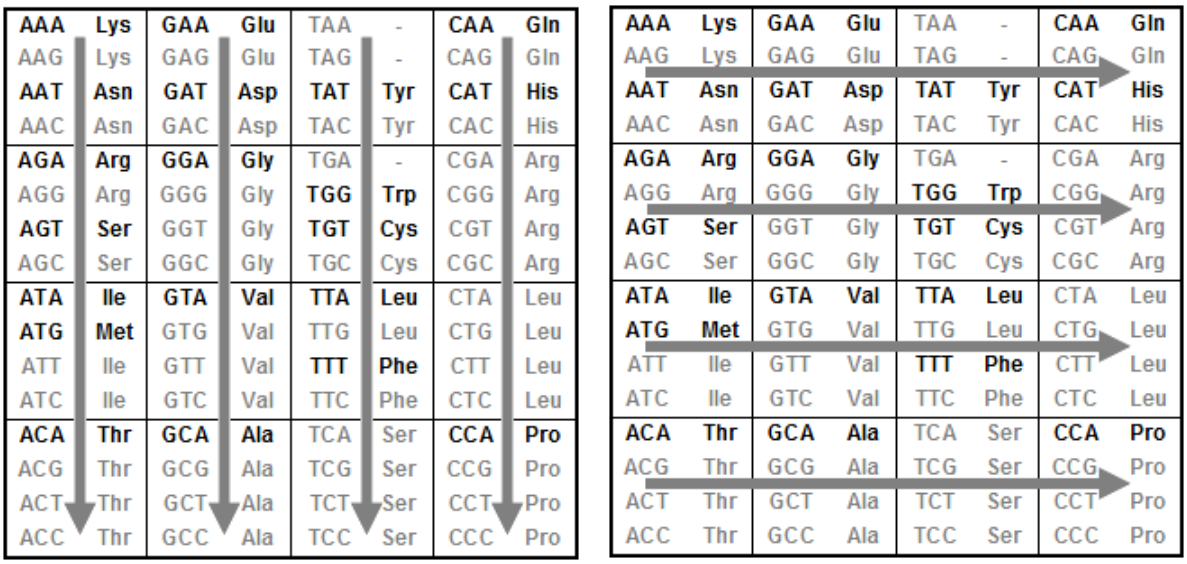


Figure 2: Vertical or horizontal filler mechanism generating the ultimate genetic code (in base order A ➞ G ➞T ➞ C).

The mechanism for filling the genetic code table can be done line by line as described in Figure 2 but also column by column while always respecting the order of the bases A-G-T-C. Also, this filling can be done in a more subtle way by alternating between a line, a diagonal and a column in two possible directions as described in Figure 3.

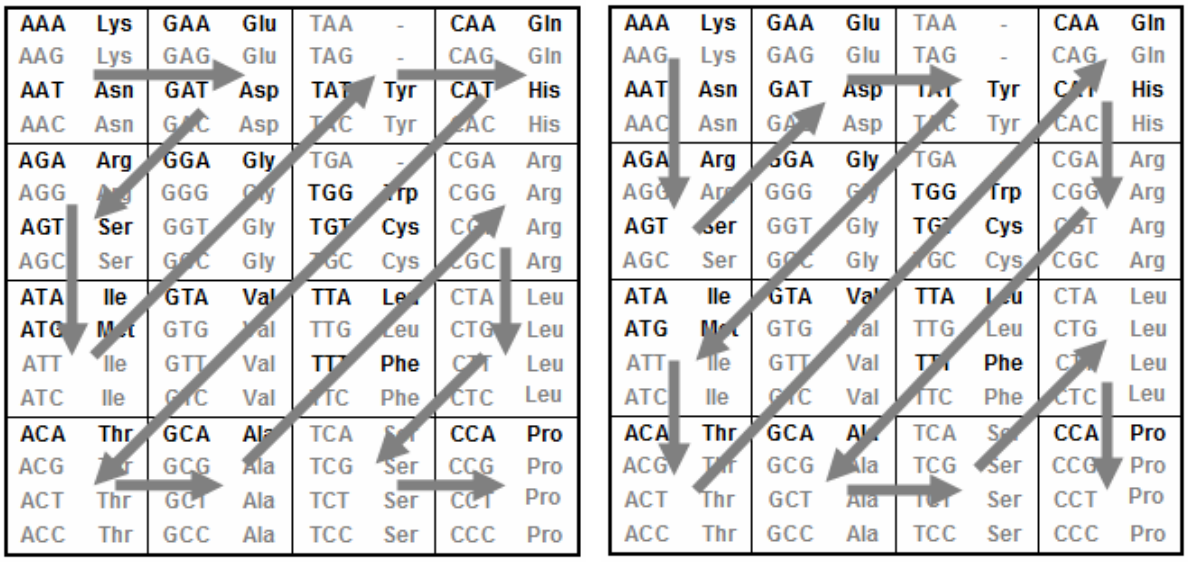


Figure 3: Alternative filler mechanism generating the ultimate genetic code (in base order A ➞ G ➞T ➞ C).

This filling order generating the ultimate genetic code is related to the degree of affinity of the bases between them: in relation to Adenine, the first base considered in the table of the genetic code (first row, first column), the second base , Guanine, is the other base of the same nature (A and G both being purines), the third, Thymine, is the complementary base (of the double DNA chain) to Adenine and the fourth and last base, Cytosine, is the opposite base: neither of the same nature nor complementary to Adenine.

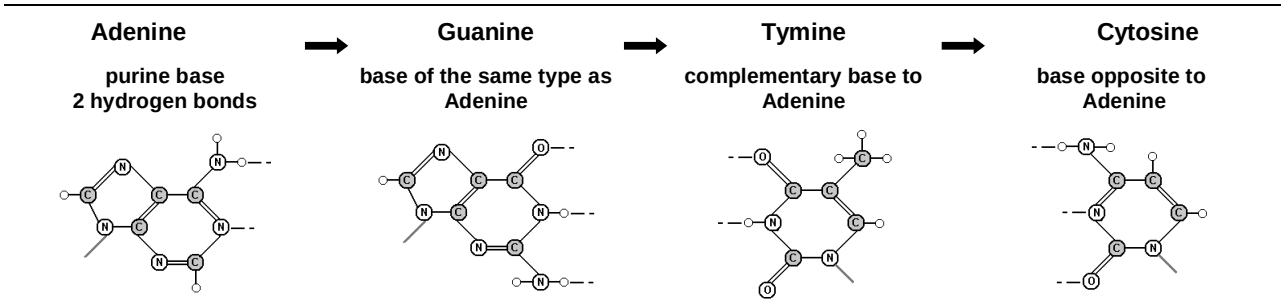


Figure 4: Classification of the four DNA bases in order A→G→T→C according to their degree of affinity.

This filling rule has an influence on the selected position of three amino acids which are normally present in two boxes of the complete table of the genetic code: SER, ARG and LEU (see Figure 1). Not demonstrating it here, the rule of filling in the order of DNA bases A-G-T-C is the only one which gives a perfect symmetrical and asymmetrical distribution of the twenty amino acids in the table of the ultimate genetic code. This base order is in fact the only one where all of the multiple remarkable phenomena presented in this unprecedented study of the genetic code are manifested.

3.2. The ultimate genetic code matrix

These different paths, always respecting the order of the bases A-G-T-C and progressively isolating, one after the other, twenty codons and twenty coded ones, produce the same final table of the ultimate genetic code as presented in Figure 5: twenty unique codons associated with twenty unique proteinogenic coded amino acids.

Generation of the ultimate genetic code

AAA	Lys	GAA	Glu	TAA	-	CAA	Gln
AAG	Lys	GAG	Glu	TAG	-	CAG	Gln
AAT	Asn	GAT	Asp	TAT	Tyr	CAT	His
AAC	Asn	GAC	Asp	TAC	Tyr	CAC	His
AGA	Arg	GGA	Gly	TGA	-	CGA	Arg
AGG	Arg	GGG	Gly	TGG	Trp	CGG	Arg
AGT	Ser	GGT	Gly	TGT	Cys	CGT	Arg
AGC	Ser	GGC	Gly	TGC	Cys	CGC	Arg
ATA	Ile	GTA	Val	TTA	Leu	CTA	Leu
ATG	Met	GTG	Val	TTG	Leu	CTG	Leu
ATT	Ile	GTT	Val	TTT	Phe	CTT	Leu
ATC	Ile	GTC	Val	TTC	Phe	CTC	Leu
ACA	Thr	GCA	Ala	TCA	Ser	CCA	Pro
ACG	Thr	GCG	Ala	TCG	Ser	CCG	Pro
ACT	Thr	GCT	Ala	TCT	Ser	CCT	Pro
ACC	Thr	GCC	Ala	TCC	Ser	CCC	Pro

The ultimate genetic code:
20 codons - 20 coded AAs

AA	Lys	GA	Glu	TA	Tyr	CA	Gln
AA	Asn	GA	Asp			CA	His
AG	Arg	GG	Gly	TG	Trp	-	
AG	Ser			TG	Cys		
AT	Ile	GT	Val	TT	Leu	-	
AT	Met			TT	Phe		
AC	Thr	GC	Ala	-		CC	Pro

Figure 5: Generating, from the table of the universal genetic code, of the table of the ultimate genetic code. See Figures 2 and 3 also.

In this ultimate genetic code, it is no longer considered the third coding base but only the first two, i.e., for each of the twenty AAs, a biplet of two nucleobases. The fact that some of these biplets encode two different amino acids is not necessarily a problem in what we present. On the contrary this amplifies the concept of the ultimate genetic code.

Thus, the study presented in this paper concerns only this matrix of the genetic code reduced to twenty double entities: twenty coding biplets and twenty coded amino acids.

4. Symmetric and asymmetric distribution of codons and coded AAs

4.1. Characteristics of DNA bases

Conventionally, the twenty codons of the ultimate genetic code consist only of two bases (nucleobases) of DNA. These two DNA bases can be Adenine, Thymine, Guanine and Cytosine. These four bases are either purines or pyrimidines. Also, these four nucleobases have either two hydrogen bonds (or bridges) or three hydrogen bonds.

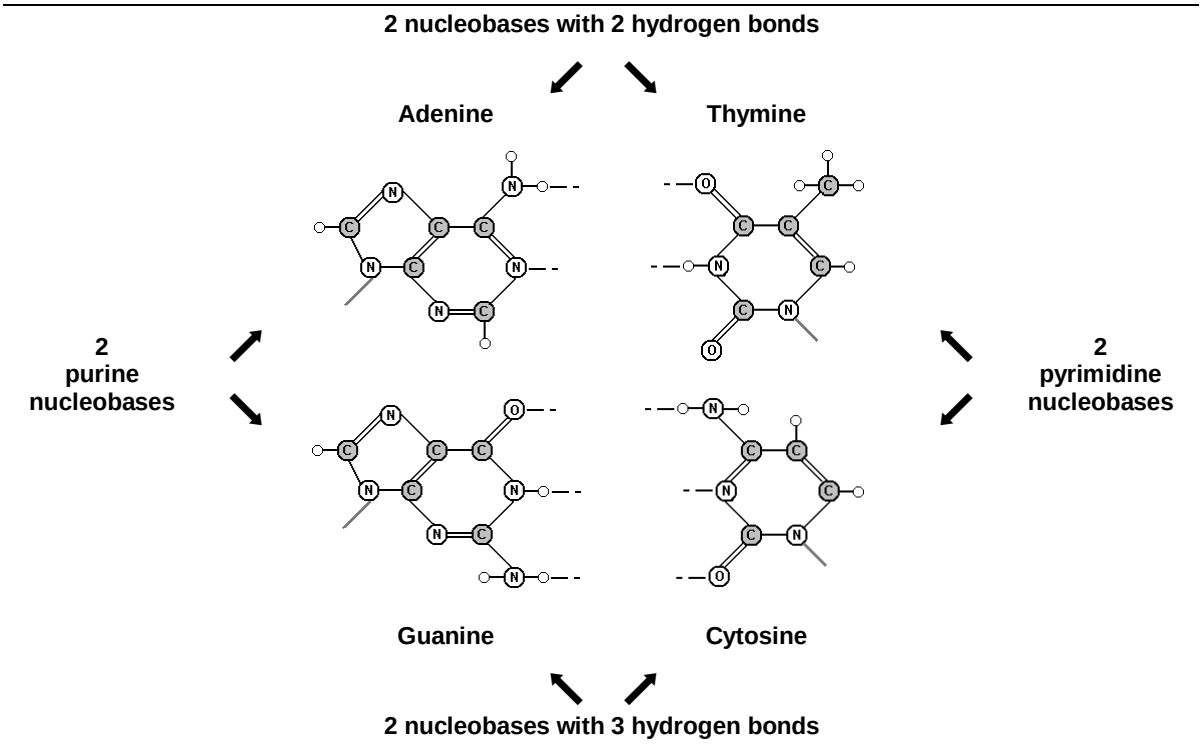


Figure 6: The four coding entities working in DNA and their specific crossed characteristics.

These two characteristics are dissociated and therefore, as shown in Figure 6, just exist:

- a purine with two hydrogen bonds: Adenine,

- a purine with three hydrogen bonds: Guanine,

- a pyrimidine with two hydrogen bonds: Thymine,

- a pyrimidine with three hydrogen bonds: Cytosine.

These two double entangled characteristics are the source of the many phenomena of symmetry and asymmetry within the matrix of the ultimate genetic code.

4.2. Symmetric and asymmetric distribution of DNA bases

In the table of the ultimate genetic code (revealed in Figure 5 of the previous chapter), each column (boxes with the same first DNA base) and each corresponding row (boxes with the same second DNA base) both have the same number of codons (and coded AAs). Thus, the number of DNA bases is identical for sets of codons having the same first DNA base or the same second DNA base.

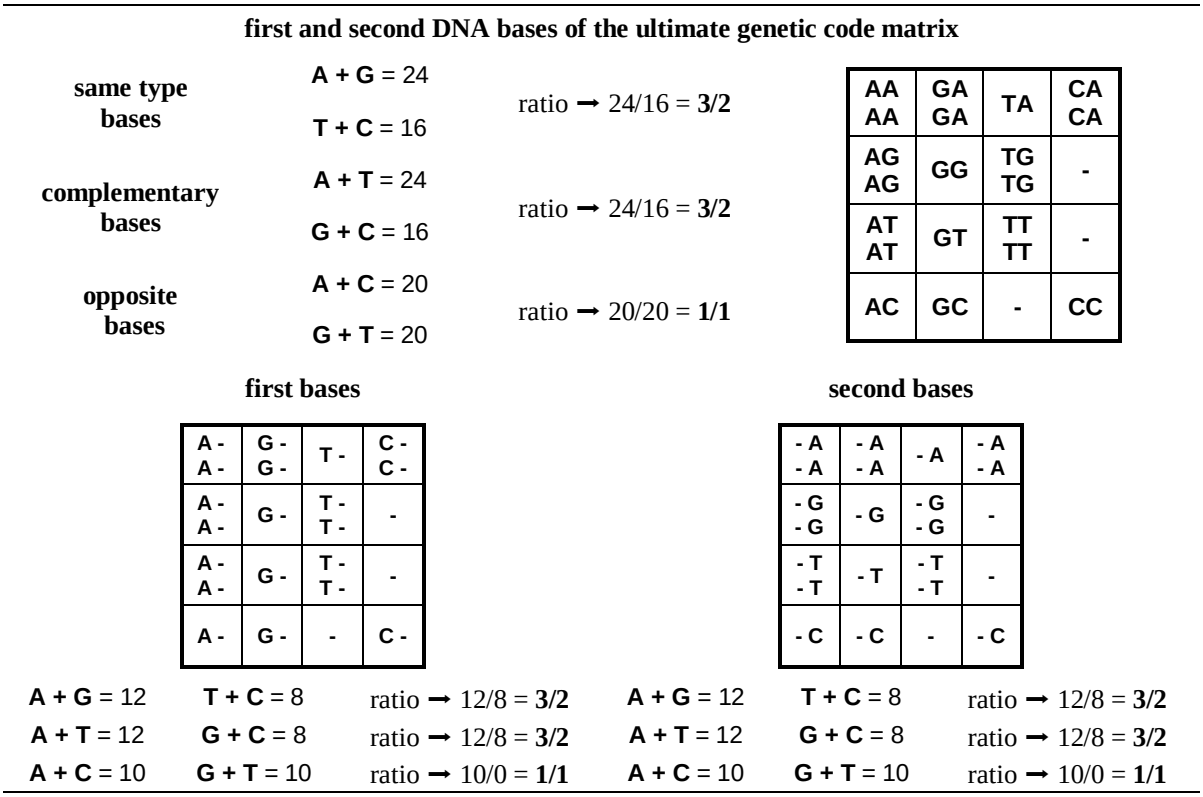


Figure 7: Depending on their respective nature, asymmetrical (3/2 ratio) and symmetrical (1/1 ratio) distribution of DNA bases in the table of the ultimate genetic code. See Figures 5 and 6 also.

As show in Figure 7, the distribution of DNA bases in the table of the ultimate genetic code is such that it generates either a 3/2 ratio or a 1/1 ratio depending on the nature of the bases considered. Thus, there are twenty-four purines (A and G) and sixteen pyrimidines (T and C): ratio 3/2. There are twenty-four complementary bases A and T and sixteen complementary bases G and C: ratio 3/2. Also there are twenty opposite bases (neither of the same nature nor complementary) A and C and twenty opposite bases G and T: ratio 1/1.

This distribution remains proportionally identical by separately distinguishing all of the first bases from all of the second bases. This observation is of great importance; it will be the source of the main arithmetic phenomena presented in this study of the genetic code.

4.3. Symmetric and asymmetric distribution of amino acids

The classification of the twenty amino acids according to the different possible characteristics of the respective DNA coding bases always gives the same series of four numbers: 7-5-5-3. Depending on the nature of the first coding base which is either A, G, T or C there are respectively, in the table of the ultimate genetic code, 7-5-5-3 coded amino acids. And depending on the

nature of the second coding base which is either A, G, T or C there are also respectively, in the table of the ultimate genetic code, 7-5-5-3 coded amino acids.

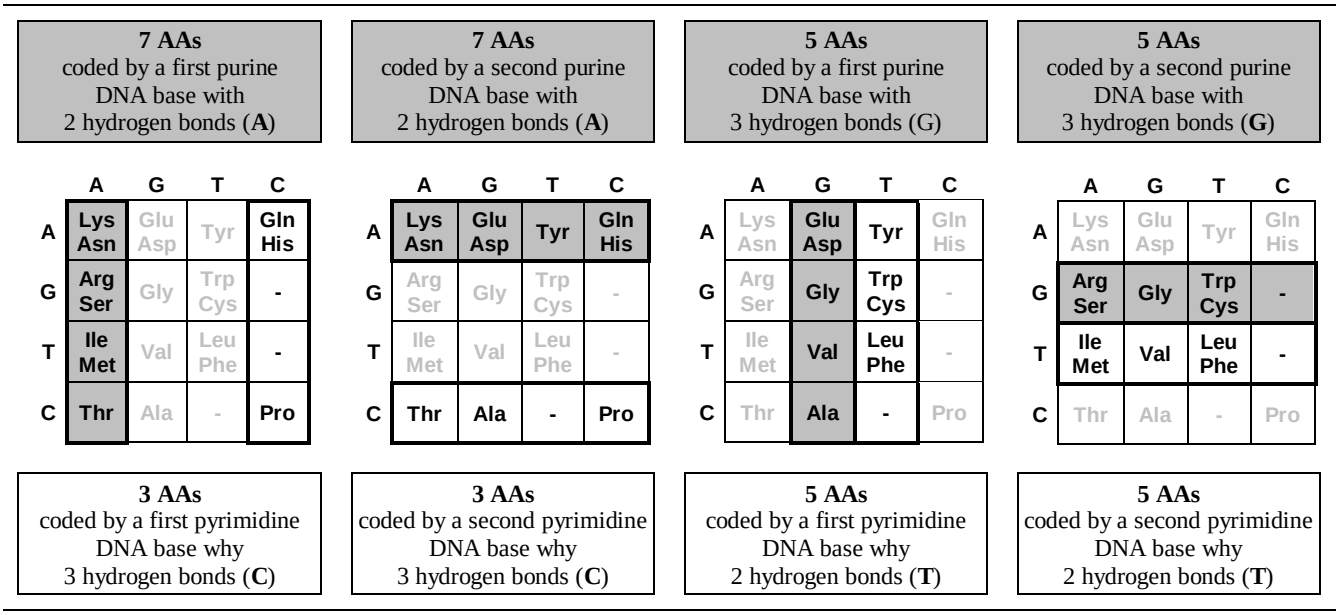


Figure 8a: Distribution of amino acids according to the nature of the first or second coding DNA base.

Depending on the nature of the two coding DNA bases which are either purines (A and G) or pyrimidines (T and C), there are respectively, in the table of the ultimate genetic code, 7-5-5-3 coded amino acids . And depending on the number of hydrogen bonds of the two coding bases which have either two bonds (A and T) or three bonds (G and C) there are also respectively, in the table of the ultimate genetic code, 7-5-5-3 coded amino acids.

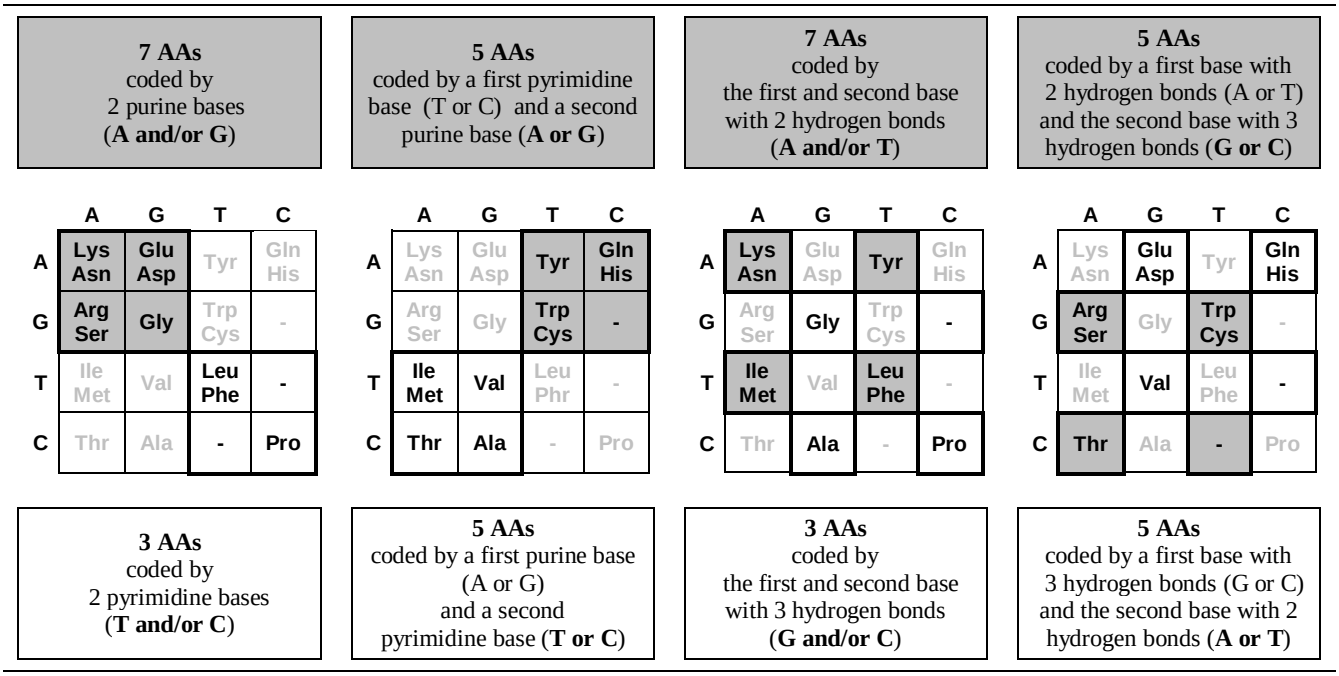


Figure 8b: Distribution of amino acids according to the nature of the first or second coding DNA base.

Thus, as shown in Figure 9, according to the different characteristics of the DNA coding bases described above and whether we consider either the first base, the second base or all of the two bases, the numbers of amino acids always remain the same.

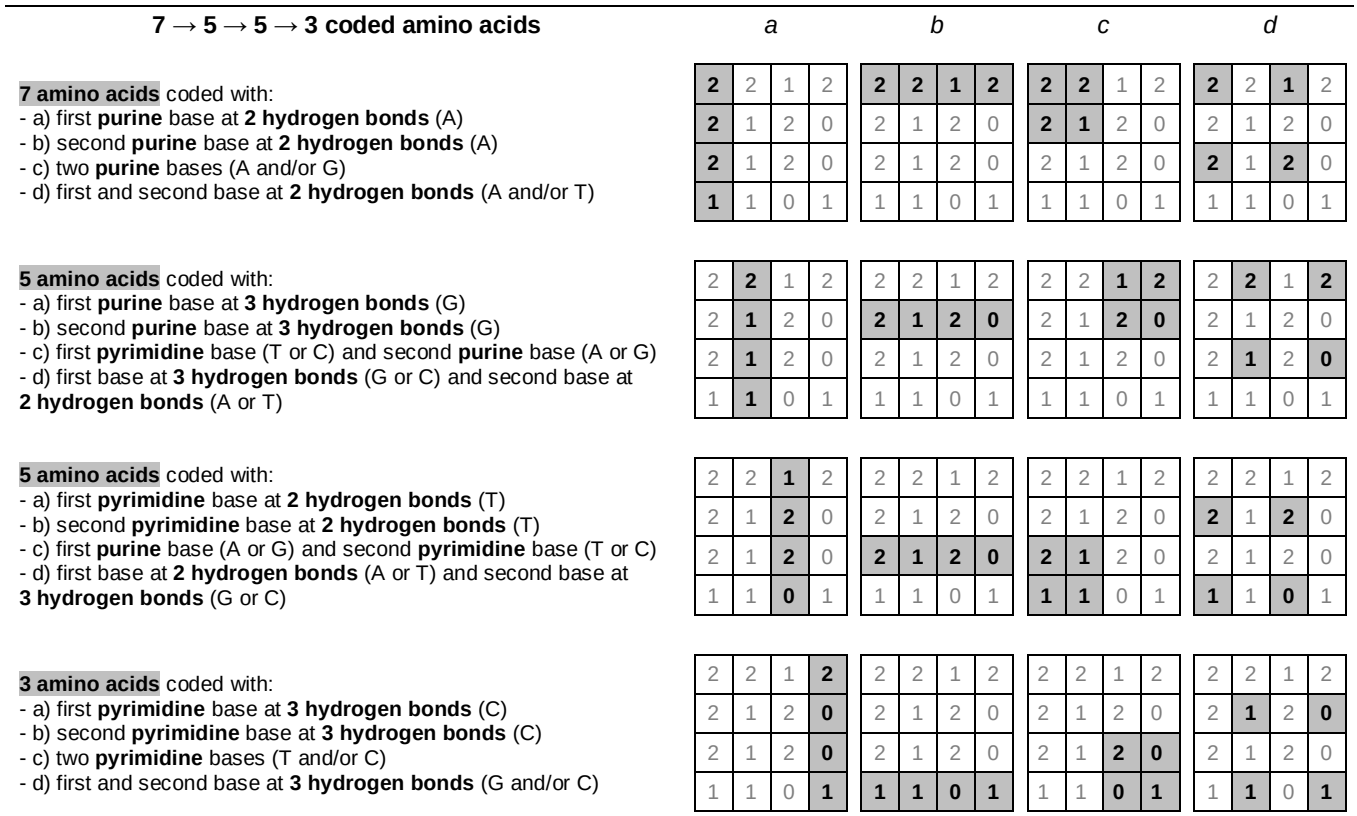


Figure 9: Configurations to **7 → 5 → 5 → 3 coded amino acids** according to the type of the two respective DNA bases in the matrix of the ultimate genetic code. See Figures 7 and 8.

4.4. Symmetrical and asymmetrical configurations

According to the previously defined criteria, the numbers of amino acids are geometrically distributed in the order 7 → 5 → 5 → 3 in the table of the ultimate genetic code. Also, these geometric configurations (reflections of the physical characteristics of the DNA bases), are organized symmetrically to form zones always having a double total of ten versus ten amino acids in a ratio of 1/1. These same configurations are organized asymmetrically to form areas always having a total of twelve versus eight amino acids in a ratio of 3/2.

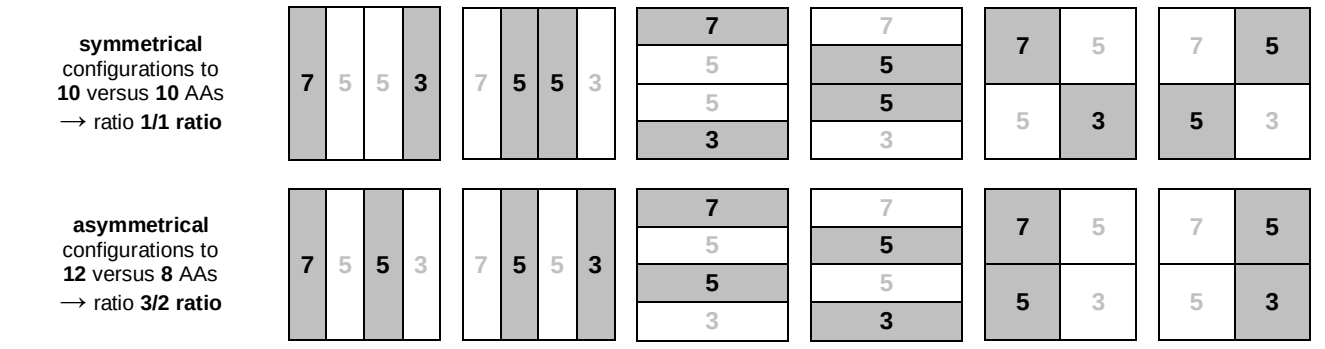


Figure 10: Symmetrical (10 versus 10 AAs) and asymmetrical (12 versus 8 AAs) configurations in the matrix of the ultimate genetic code. See Figure 9 also.

4.4.1. Symmetrical sub-configurations

Also, in the table of the ultimate genetic code, each of the row, column or square configurations (made up of four boxes) containing the same number of coded amino acids (7 → 5 → 5 → 3) can be separated into two sub-configurations two semi-rows, two semi-columns or two semi-squares (made up of two boxes) which also have the same number of amino acids depending on their positioning in the matrix.

Thus, as it appears in Figure 11, the configurations of 7 entities are split into two sub-configurations of always 4 and 3 entities. Those of 5 entities all split into sub-configurations of 3 and 2 entities and those of 3 entities always split into sub-configurations of 2 and 1 entities.

Overall, in the matrix of the ultimate genetic code of sixteen boxes, the number of amino acids in these sub-configurations is then always in order (geometric from A to C):

4 → 3 → 3 → 2 → 3 → 2 → 2 → 1

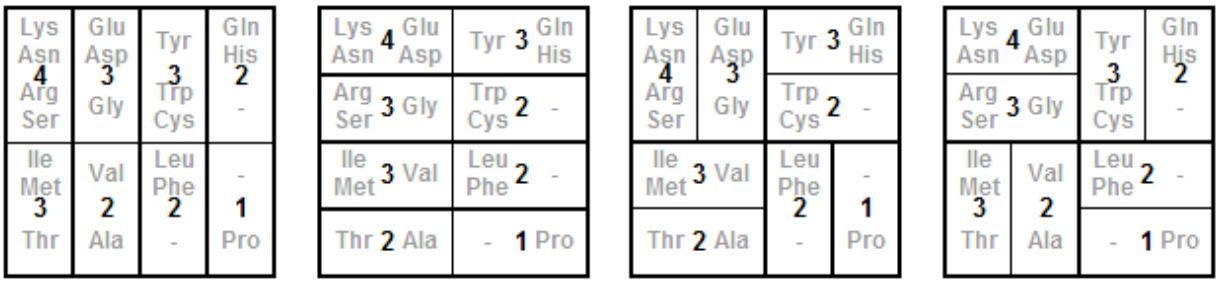


Figure 11: From configurations to 7 → 5 → 5 → 3 AAs, sub-configurations to 4 → 3 → 3 → 2 → 3 → 2 → 2 → 1 amino acids in the ultimate genetic code matrix. See Figure 9 and 10 also.

It is interesting to note that the two middle configurations are identical with five and five amino acids then, in more detail, three and two then again three and two amino acids. This can be explained by the fact that, for square configurations, it is not possible to distinguish the order of these configurations as it is for row and column configurations.

Indeed, by following the order of the bases A→G→T→C, there is only one possible path to number the order of the row or column configurations but two possible paths to number the order of the square configurations (see Figures 7 and 8, very explicit about this phenomenon). The organization of the ultimate genetic code is such that it overcomes this constraint by this arrangement in a double identical configuration in a median position.

Also, as visible in Figure 12, the symmetrical arrangement by four of the eight sub-configurations of two boxes generates, in the table of the ultimate genetic code, areas geometrically mirroring each other and whose number of coded amino acids (and respective codons) is always equal to ten.

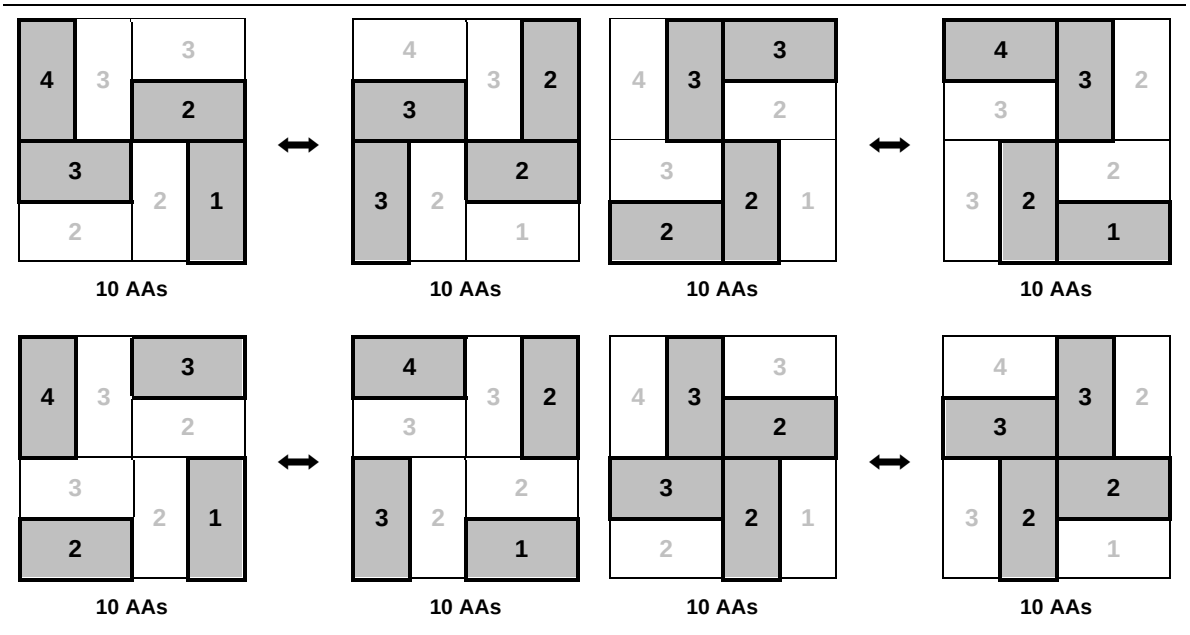


Figure 12: Symmetric mirror area configurations of always 10 and 10 amino acids in the matrix of the ultimate genetic code. See Figure 11 also.

We have just demonstrated that the distribution of the twenty proteinogenic amino acids in the matrix of the ultimate genetic code is in no way random. Indeed, depending on the respective nature of the two bases (DNA biplets) which encode them, the AAs form different groups of entities opposing each other symmetrically or asymmetrically in ratios of the exact value 3/2 or 1/1. We will now demonstrate that this ultimate matrix, emanating from the standard genetic code, is organized in a very sophisticated way in what we call the ultimate 3/2 ratio.

5. The genetic code ultimate ratio of 3 to 2

The mirror configurations of eight boxes (Figure 12 of the previous Chapter) therefore all have the same number of coded amino acids, i.e. always composed of ten amino acids. This phenomenon is due to the particular distribution of amino acids in each of the sixteen boxes of the table of the ultimate genetic code. This singular distribution of the entities of the genetic code matrix falls into what we call "the ultimate ratio 3 to 2".

5.1. Associations of the bases inside DNA biplets

Depending on their distinctive characteristics (see Figure 6 of Chapter 4.1), the four DNA bases can associate in the twenty codons of the ultimate genetic code (two-base biplets) in four different ways as illustrated Figure 13.

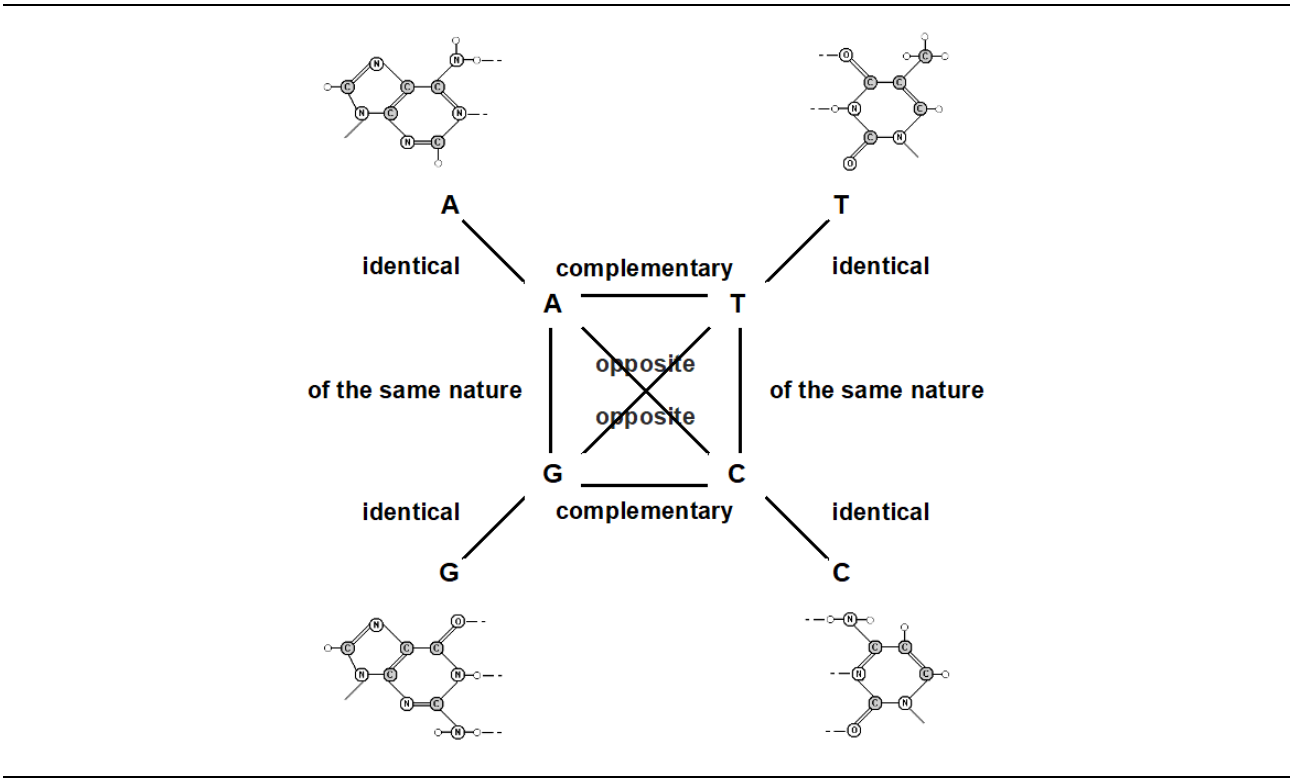


Figure 13: All possible associations of the four DNA bases inside biplets of the ultimate genetic code.
See Figure 4 and 5 also.

Thus, for each amino acid, the two respective DNA coding bases of the ultimate genetic code are, one in relation to the other:

- either identical,
- either complementary: same number of hydrogen bonds (2 or 3 bridges),
- either of the same nature:2 purines or 2 pyrimidines,
- or opposite: neither identical nor complementary nor of the same nature.

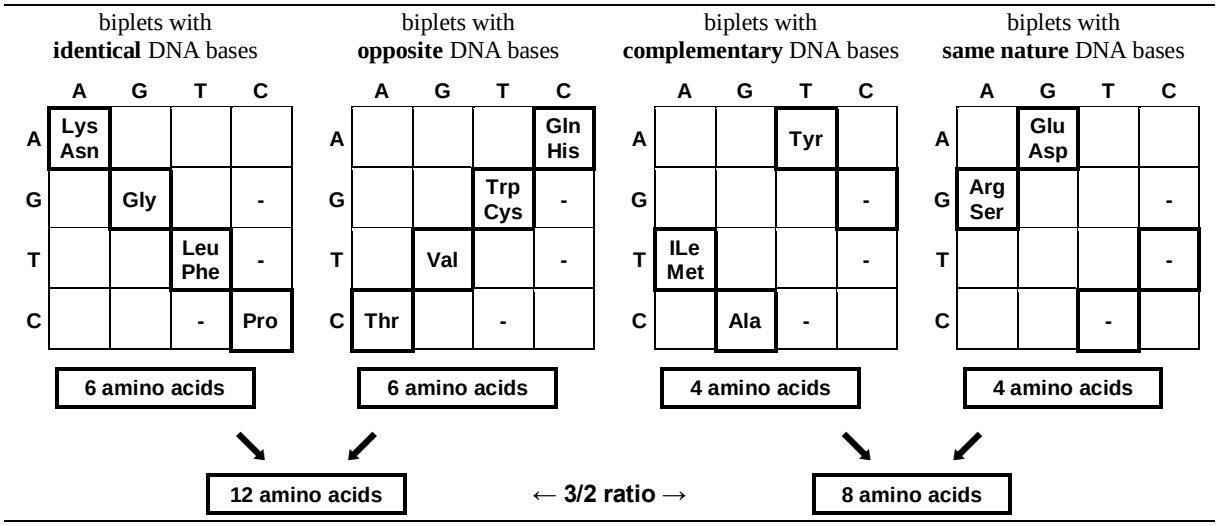


Figure 14: Related to the nature of the codons of the ultimate genetic code, the four amino acid configurations generating the ultimate 3/2 ratio.

Figure 14 highlights that, in the ultimate genetic code matrix, the quantities of coded amino acids are very dependent on the respective nature of the nucleobases pair (biplet) which encode each of the twenty proteinogenic amino acids.

In fact, the number of amino acids coded with two identical bases and the number of amino acids coded with two opposite bases are the same: six amino acids. Also, the number of amino acids coded with two complementary bases and the number of amino acids coded with two same nature bases are also identical: four amino acids. These double configurations oppose each other in a duality whose ratio is equal to 3/2 (12 versus 8 amino acids).

These four configurations, directly dependent on the respective nature of the two nucleobases encoding each of the twenty proteinogenic amino acids, are the very source of what we call the ultimate 3/2 ratio, a ratio which constitutes the main object of this study of the genetic code.

5.2. Source of the ultimate ratio 3 to 2

The synthesis of previous investigations on the respective nature of the nucleobases coding for the twenty amino acids and on the distribution of these twenty amino acids in the matrix of the ultimate genetic code reveals the remarkable phenomenon of the organization of this ultimate genetic code in the ratio 3/ 2 ultimate.

5.2.1. Pairs of symmetrical boxes

Appellation, "ultimate 3/2 ratio", as well as the appellation "ultimate genetic code", find all their legitimacy when we consider the singular distribution of amino acids (in the concept of the ultimate genetic code) into always three entities in the four pairs of coding and symmetrically opposed boxes with identical or opposite nucleobases and the equally singular distribution of amino acids in always two entities in the four pairs of coding and opposite boxes with nucleobases of the same nature or complementary. Figure 15 clearly highlights this singular phenomenon.

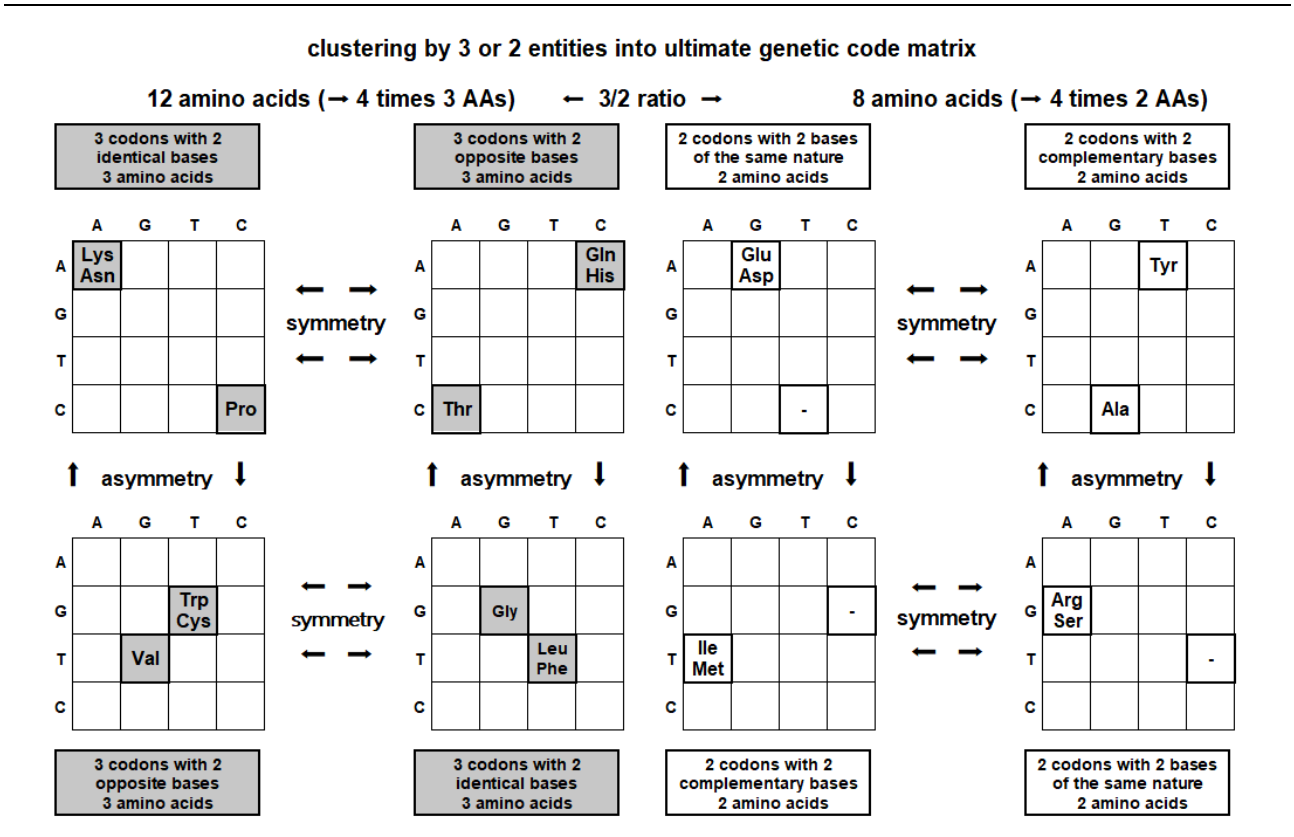


Figure 15: The eight source configurations that are the origin of the ultimate 3/2 ratio. See also figure 14.

Thus, in the table of the ultimate genetic code as introduced in Figure 5 in chapter 3.2, the association of a coding box with the opposite coding one always generates a set of three amino acids or always a set of two amino acids according to the respective nature of the two coding bases of the codon (biplet) considered.

As illustrated in Figure 15, These configurations are therefore formed by the grouping of symmetrically opposite boxes in the table of the ultimate genetic code. Two symmetrically opposite boxes always contain either a total of three amino acids or a total of two amino acids. The symmetrically opposite boxes are in relation to the nature of the two bases of the codons or biplets of the ultimate genetic code.

Thus, for example, the GA codon is symmetrically opposed to the TC codon because G and T (first bases) and A and C (second bases) are bases opposite each other: they are neither identical, nor of the same nature, nor complementary. The association of two boxes according to this criterion always gives three amino acids (and three codons) if the codons are made up of two identical bases or two opposite bases. By this same mechanism, association always gives two amino acids (and two codons) if the codons are made up of two bases of the same nature or of two complementary bases.

5.3. The ultimate ratio 3 to 2

The final association of the symmetrical and asymmetrical groups of two boxes containing either three or two amino acids form two groupings (or coding areas) of eight boxes containing respectively a total of twelve and eight amino acids (as well as twelve and eight codons respectively) . These two groupings are therefore opposed in a duality whose ratio is equal to 3/2. This phenomenon (this ratio) is thus conventionally called "The ultimate 3/2 ratio".

Figure 16 graphically introduces the concept of the ultimate 3/2 ratio into the entire table of the ultimate genetic code.

The ultimate ratio 3 to 2

12 codons
at 2 opposite bases or at 2 identical bases
12 amino acids

AA Lys AA Asn	GA Glu GA Asp	TA Tyr	CA Gln CA His
AG Arg AG Ser	GG Gly	TG Trp TG Cys	*CG* -
AT Ile AT Met	GT Val	TT Leu TT Phe	*CT* -
AC Thr	GC Ala	*TC* -	CC Pro

8 codons
at 2 complementary bases or at 2 same nature bases
8 amino acids

Figure 16: The ultimate 3/2 ratio: 12 amino acids and 12 codons with 2 opposite or identical bases versus 8 amino acids and 8 codons with 2 complementary or similar bases.

This ratio, “the ultimate 3/2 ratio” can therefore be schematized by this twin image Figure 17, source of the multiple investigations of this singular study of the genetic code. This mapping therefore represents the two zones of the ultimate 3/2 ratio within the matrix of sixteen boxes of the ultimate genetic code.

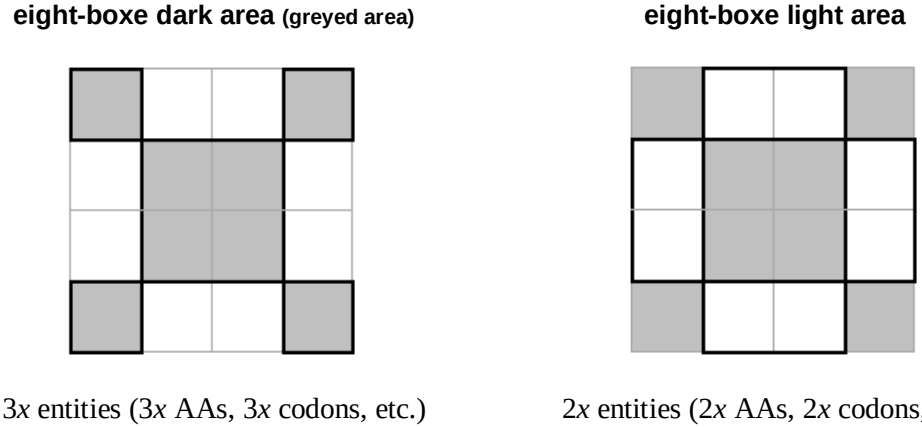


Figure 17: Mapping of the ultimate 3/2 ratio schematizing the table of the ultimate genetic code: conventionally dark zone (greyed) made up to 3*x* entities, conventionally clear area consisting of the same multiple to 2*x* entities.

Wherever the ultimate 3/2 ratio is highlighted, this form of table is now used to describe the multitude of phenomena presented.

Note: the allocation of colours (greyed and light) is purely random and only responds to a certain concern for the aesthetics of the presentations. However, by convention, these two areas are definitely named "dark (or grey) area" and "light area" of the ultimate 3/2 ratio.

6. Numbering of the twenty proteinogenic amino acids

In preview published paper "*Numbering of the twenty proteinogenic amino acids and new alphanumerical nomenclature proposal to them"* [1], we have demonstrated that a large number of different amino acid attributes arrange themselves numerically in exact 3/2 value ratios according to a numbering system of the twenty proteinogenic amino acids. We will then show that this numbering is closely related to the configuration of the matrix of the ultimate genetic code by transcending the ultimate 3/2 ratio.

From a subtle numbering of the 64 codons of the universal genetic code, we propose a numbering (from 0 to 19) of the twenty amino acids. These two numbering systems, including the first proposed by Professor Sergey Petoukhov [3], are very directly dependent on the physico-chemical properties of the four nucleobases that make up DNA. They are therefore very legitimate to be used for the study of the genetic code mechanism. When we number the twenty amino acids, which are, very importantly, 5*x* in number, then we classify them into two symmetrical sets of 12 (or 3*x*) and 8 (or 2*x*) entities.

6.1. Petoukhov's numbering of the 64 genetic code codons

In his investigations of the genetic code [3] Sergey Petoukhov assigns a number from 0 to 63 to each of the sixty-four codons. This Petoukhov numbering is directly dependent on the physico-chemical properties of the four DNA coding bases. Using a very sophisticated method, Sergey Petoukhov manages to classify the full sixty-four codons set using a binary language (or alphabet, we invite the reader to consult the full article by Sergey Petoukhov [3]). Depending on whether each nucleobase can undergo deamination or not, Sergey Petoukhov assigns them either the value 1 or the value 0 as illustrated Figure 18. Also, depending on whether each nucleobase can undergo depurination or not, Sergey Petoukhov assigns them either the value 0 or the value 1.

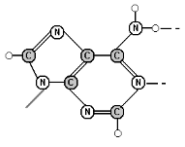
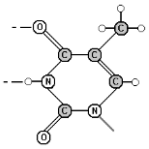
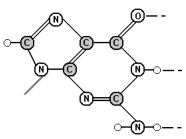
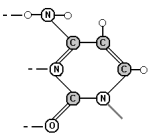
	Adenine	**Thymine**	**Guanine**	**Cytosine**
nucleobases				
Possible deamination: yes = 1 no = 0	**1**	**0**	**0**	**1**
Possible depurination: yes = 0 no = 1	**0**	**1**	**0**	**1**

Figure 18: Method of assigning a double binary value to the four DNA nucleobases according to Sergey Petoukhov [3].

This double criterion makes it possible, for each codon, to create a six-digit binary number by juxtaposition of two three-digit numbers as described in Figure 19.

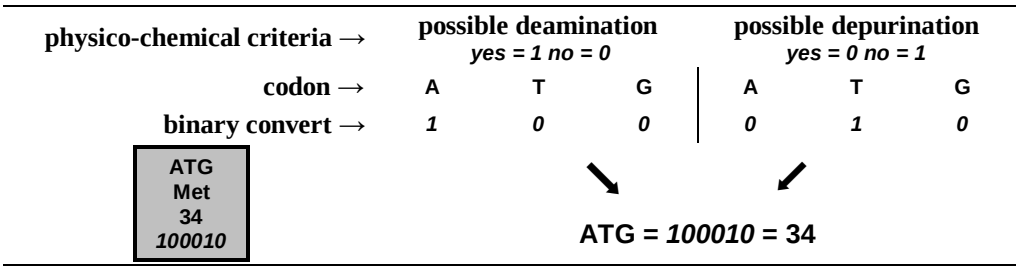


Figure 19: Method of assigning a number to codons according to Sergey Petoukhov.
See Figures 18 and 20 also.

Sergey Petoukhov then classifies very subtly in superimposed squares of 4, 16 and 64 boxes the 64 codons and numbers them in the order of the bases G→T→A→C for the first, second and third bases. In this numbering imagined by Sergey Petoukhov, the GGG codon thus bears the number 0 (binary 000000) and the CCC codon the number 63 (binary 111111). Figure 20 illustrates this complete numbering of the 64 genetic code codons set.

We would like to support here the importance that Sergey Petoukhov [3] gives to these criteria of deamination and depurination. It is in fact largely through these criteria that the four nucleobases can "mutate" within DNA chains, also allowing the evolution of living organisms.

	111	*110*	*101*	*100*	*011*	*010*	*001*	*000*
111	CCC Pro 63 111111	CCA Pro 62 111110	CAC His 61 111101	CAA Gln 60 111100	ACC Thr 59 111011	ACA Thr 58 111010	AAC Asn 57 111001	AAA Lys 56 111000
110	CCT Pro 55 110111	**CCG Pro 54 110110**	**CAT His 53 110101**	**CAG Gln 52 110100**	ACT Thr 51 110011	**ACG Thr 50 110010**	**AAT Asn 49 110001**	**AAG Lys 48 110000**
101	CTC Leu 47 101111	CTA Leu 46 101110	CGC Arg 45 101101	CGA Arg 44 101100	ATC Ile 43 101011	ATA Ile 42 101010	AGC Ser 41 101001	AGA Arg 40 101000
100	CTT Leu 39 100111	CTG Leu 38 100110	CGT Arg 37 100101	CGG Arg 36 100100	**ATT Ile 35 100011**	**ATG Met 34 100010**	AGT Ser 33 100001	**AGG Arg 32 100000**
011	TCC Ser 31 011111	TCA Ser 30 011110	TAC Tyr 29 011101	TAA Stop 28 011100	GCC Ala 27 011011	GCA Ala 26 011010	GAC Asp 25 011001	GAA Glu 24 011000
010	TCT Ser 23 010111	**TCG Ser 22 010110**	**TAT Tyr 21 010101**	TAG Stop 20 010100	GCT Ala 19 010011	**GCG Ala 18 010010**	**GAT Asp 17 010001**	**GAG Glu 16 010000**
001	TTC Phe 15 001111	TTA Leu 14 001110	TGC Cys 13 001101	TGA Stop 12 001100	GTC Val 11 001011	GTA Val 10 001010	GGC Gly 9 001001	GGA Gly 8 001000
000	**TTT Phe 7 000111**	**TTG Leu 6 000110**	**TGT Cys 5 000101**	**TGG Trp 4 000100**	GTT Val 3 000011	**GTG Val 2 000010**	GGT Gly 1 000001	**GGG Gly 0 000000**

Figure 20: Numbering of the 64 codons according to Sergey Petoukhov genetic code investigations [1] and distinction (grey areas) of the first appearance of each of the 20 coded amino acids. See Figures 18 and 19 also.

6.2. Numbering of the twenty proteinogenic amino acids

From this numbering system, in order to assign a number to each of the twenty proteinogenic amino acids, the most logical procedure is therefore proposed here, which is to follow the order of appearance of the amino acids according to this numbering of the codons (from 0 to 63) of the table by Sergey Petoukhov (Figure 20).

Only one number assigning to amino acids

Pro	Pro	His	Gln	Thr	Thr	Asn	Lys
Pro	**19 Pro**	**18 His**	**17 Gln**	Thr	**16 Thr**	**15 Asn**	**14 Lys**
Leu	Leu	Arg	Arg	Ile	Ile	Ser	Arg
Leu	Leu	Arg	Arg	**13 Ile**	**12 Met**	Ser	**11 Arg**
Ser	Ser	Tyr	*Stop*	Ala	Ala	Asp	Glu
Ser	**10 Ser**	**9 Tyr**	*Stop*	Ala	**8 Ala**	**7 Asp**	**6 Glu**
Phe	Leu	Cys	*Stop*	Val	Val	Gly	Gly
5 Phe	**4 Leu**	**3 Cys**	**2 Trp**	Val	**1 Val**	Gly	**0 Gly**

Figure 21: Assigning a single only one number to each of 20 proteinogenic amino acids in the table of the complete genetic code. See Figure 20 also.

By this process, it is thus assigned (Figure 21) number 0 to Glycine, number 1 to Valine and to Proline, the last amino acid to appear according to this order of numbering of the sixty-four genetic code codons, 19 as number.

Also, as we give the explanation in appendix A6, we now designate the twenty AAs by a 5-character alphanumeric symbol.

6.3. Symmetrical break-up of the 20 AAs in ratio 3 to 2

Now that we have determined a numbering of amino acids by assigning them a unique and personal number, we propose to isolate these twenty entities in two sets of unequal size. We therefore distinguish, in Figure 22, a first set of 12 entities then a second set of 8 other entities. As illustrated in Figure22, these two sets then oppose each other in a ratio of value 3/2.

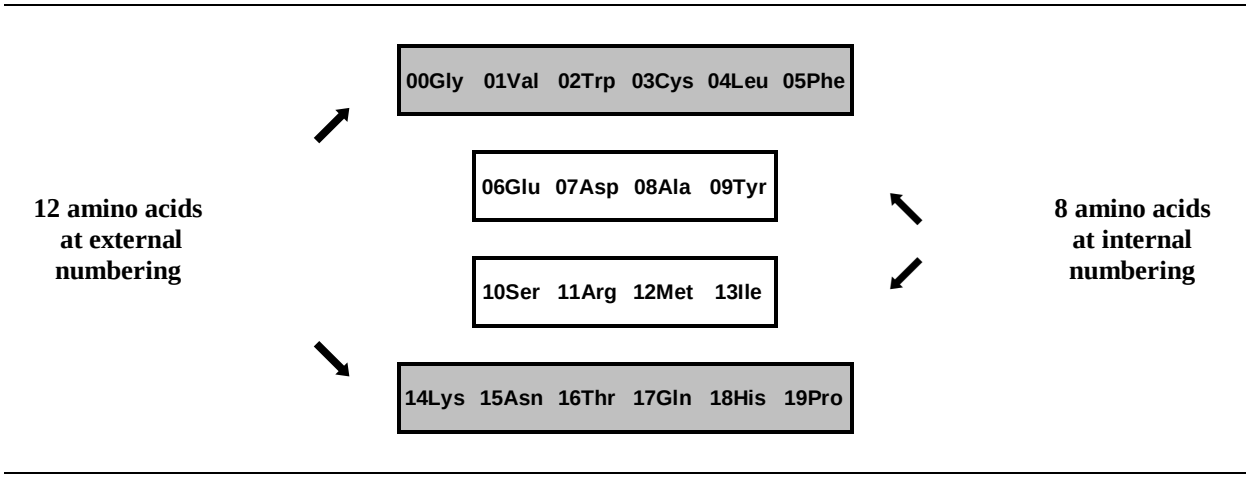


Figure 22: Conventional representation of 20 proteinogenic amino acids numbering in symmetry graphics. Since them numbering, symmetrical break-up of the 20 AAs into two sets of 2 times 6 versus 2 times 4 entities. See Figure 21 also.

Using symmetry graphics, many arithmetic phenomena presented in this paper will be presented in the way illustrated in Figure 22. Thereby, each of the 20 amino acids is symmetrically positioned to the one of opposite numbering in relation to the numbering order of these 20 AAs: *00Gly* versus *19Pro*, *01Val* versus *18His*, etc.

Also, we therefore isolate two numbering zones:

- an area called “external” with inside the first six and last six numbered AAs
- an area called “internal” with inside the two times four centrally numbered AAs.

7. Amino acid numbering and ultimate genetic code matrix

By introducing the concept of ultimate genetic code, we were able to conclude that the matrix of this optimized code is organized into two areas of 12 versus 8 entities. What we call the ultimate 3/2 ratio. Using the concept of numbering the twenty proteinogenic AAs, we have also defined two areas of 12 versus 8 entities.

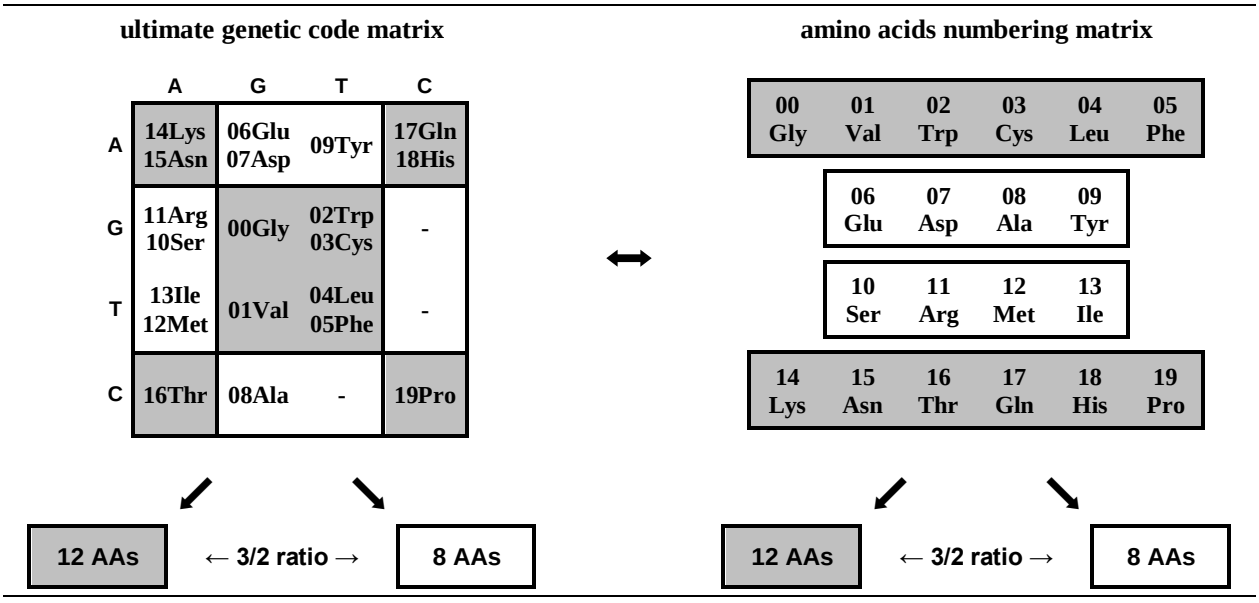


Figure 23: The two convergent matrix of the ultimate genetic code. See Figures 16 and 22.

As we can discover in Figure 23, it turns out that these two double areas are coincident. In fact, the 12 amino acids with two identical bases or two opposite bases are entirely those with external numbering. The 8 AAs with two bases of the same nature or with complementary bases are entirely those with internal numbering. This phenomenon amplifies these two twin concepts which we merge into the genetic concept of ultimate ratio 3 to 2.

Figure 24 illustrates an example of application of the concept of ultimate 3/2 ratio by the study of the prime attributes of the twenty proteinogenic amino acids that we will develop in the next chapter.

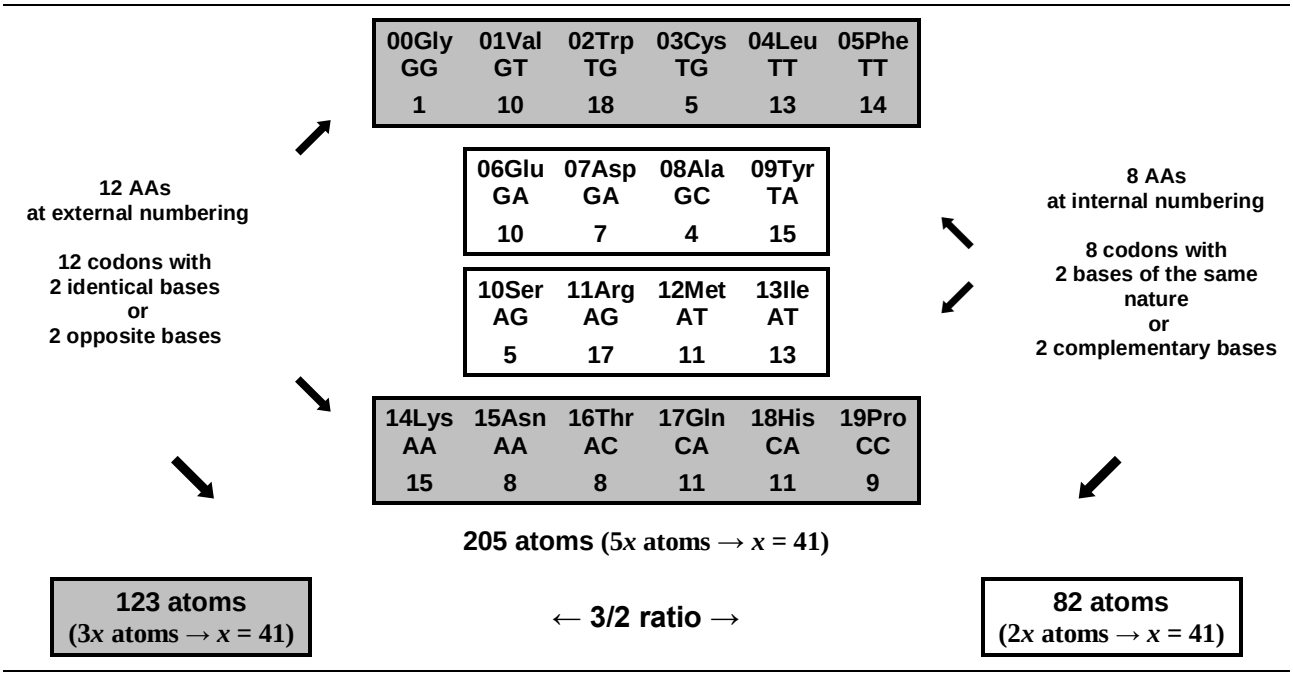


Figure 24: Ultimate genetic code concept and AA numbering concept superposition. See Figure 23 and 25 also.

We present here the very significant number of atoms contained in each of the radicals of the twenty AAs. The total number of atoms contained inside the radicals of the twenty amino acids is equal to 205. This number is therefore equal to 5*x* entities with $x = 41$. Also, as shown in Figure 24 (and as listed Figure 26), there are 123 atoms (3*x* atoms → $x = 41$) in the 12 external AAs set and there are 82 atoms (2*x* atoms → $x = 41$) in the 8 AAs of the internal group. Thus, these two sets are opposed in an exact ratio of 3/2 as value.

By superposition of the previously defined concepts, these two sets are also those of the AAs at (ultimate) codon with 2 identical bases or 2 opposite bases and those of the AAs at (ultimate) codon with 2 bases of the same nature or 2 complementary bases. Thus, as we will also demonstrate in the next chapter, the prime attributes of the twenty proteinogenic amino acids are also organized in perfect ratios of value 3/2 in agreement with their numbering and their ultimate coding.

8. Amino acid attributes and ultimate ratio 3 to 2

The phenomena presented here come from the published article *"Numbering of the twenty proteinogenic amino acids"* [1], we have demonstrated that a large number of different amino acid attributes arrange themselves numerically in exact 3/2 value ratios according to their numbering system.

8.1 Prime amino acid attributes

The following table (Figure 25) lists the 20 amino acids in mapping of the ultimate genetic code matrix previously introduced. In this study of the genetic code, it is considered the amino acids in their isolated and saturated state, therefore (important point of view) in non-ionized states.

This schematic representation (inspired from Petoukhov paper [3]) of the 20 proteinogenic amino acids makes it possible to highlight several of their physicochemical characteristics studied here in this paper.

This table therefore highlights for each AA (and not restrictively) :

- alphanumeric appellation,
- DNA biplet,
- number of atoms,
- type of atom,
- number of CH_2 groups.

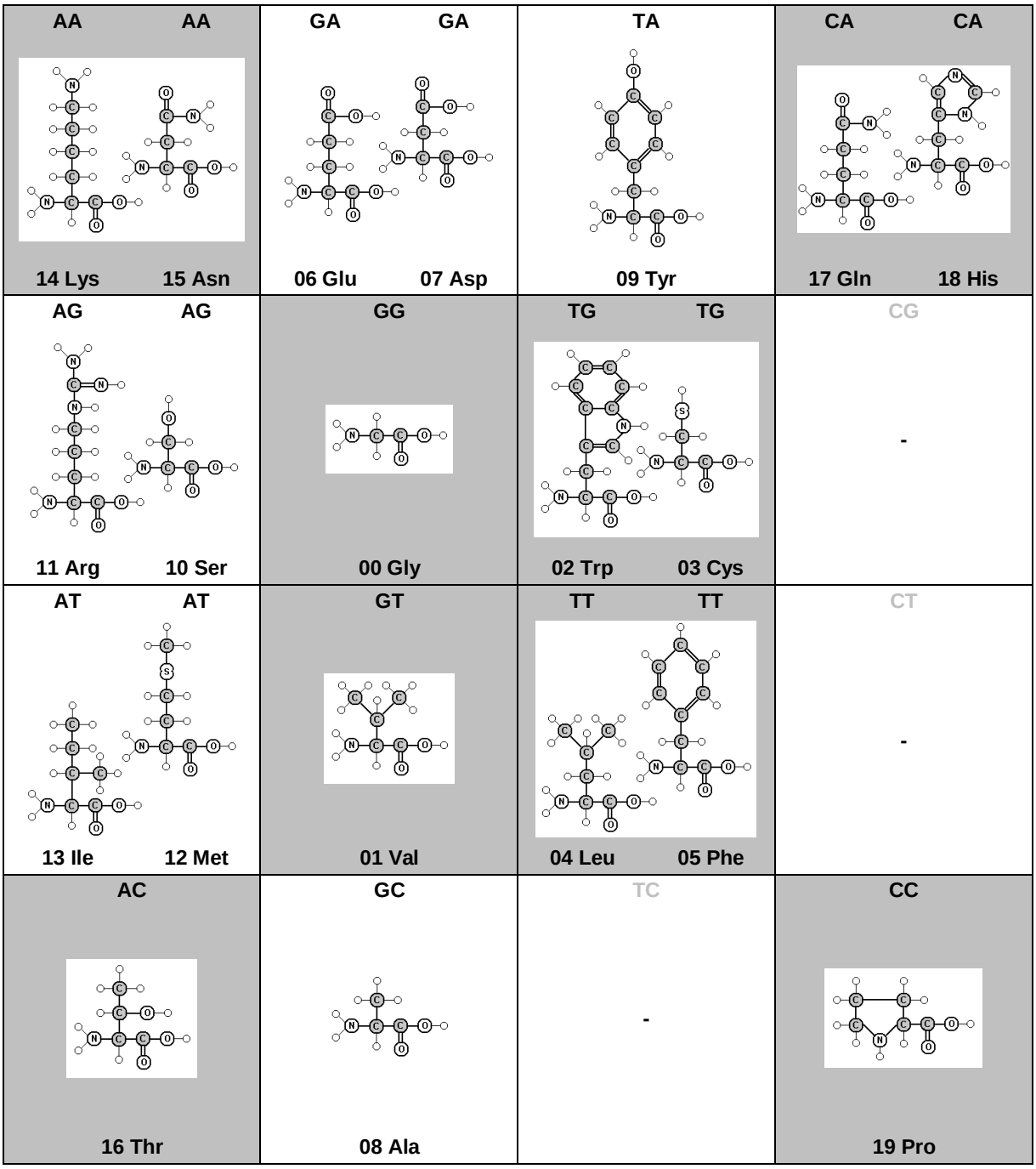


Figure 25: Table of the ultimate genetic code: codon (biplet), name and number (see Chapter 7) structure of the amino acid. Distinction of the two areas (greyed and light) of the ultimate 3/2 ratio.

The following table in Figure 26 condenses the main values relating to the twenty proteinogenic amino acids. This table lists absolute (integer) values which are therefore not suggestive and subject to debate. For each amino acid, it is so listed, its numbering, its atom number inside its radical, detail of this number with atoms at even or odd electron shells, its number of CH_2 groups (methylene bridges directly connected to alpha carbon).

It is also listed, for each of twenty proteinogenic amino acids, its rank of OMH hydrophobicity index and its number of codons in its largest respective codon set with two first same DNA bases.

A brief overview of this table already demonstrates the powerful tendency of the mechanism of the genetic code to organize itself into perfect arithmetic ratios of 3/2 values.

We study here from the outset, both the anatomy of amino acids (physical structure), their chemical properties (hydrophobicity) and the coding genetic structure associated with them. This voluntary approach in order to suggest the inter connectivity of the phenomena studied.

AA	*Molecular formula**	***a***	***b***	***c***	***d***	***e***	***f***	***g***
Gly	H	0	1	0	1	0	15	4
Val	C3H7	1	10	3	7	0	6	4
Trp	C9H8N	2	18	10	8	1	7	1
Cys	CH3S	3	5	1	4	1	8	2
Leu	C4H9	4	13	4	9	1	4	4
Phe	C7H7	5	14	7	7	1	1	2
Glu	C3H5O2	6	10	5	5	2	19	2
Asp	C2H3O2	7	7	4	3	1	20	2
Ala	CH3	8	4	1	3	0	10	4
Tyr	C7H7O	9	15	8	7	1	2	2
Ser	CH3O	10	5	2	3	1	12	4
Arg	C4H10N3	11	17	7	10	3	13	4
Met	C3H7S	12	11	3	8	2	5	1
Ile	C4H9	13	13	4	9	0	3	3
Lys	C4H10N	14	15	5	10	4	16	2
Asn	C2H4NO	15	8	4	4	1	18	2
Thr	C2H5O	16	8	3	5	0	9	4
Gln	C3H6NO	17	11	5	6	2	17	2
His	C4H5N2	18	11	6	5	1	14	2
Pro	C3H6	19	9	3	6	3	11	4
Cumulated values		**190**	**205**	**85**	**120**	**25**	**210**	**55**
12 AAs at codons with 2 identical bases or 2 opposite bases		**114**	**123**	**51**	**72**	**15**	**126**	**33**
8 AAs at codons with 2 bases of the same nature or 2 complementary bases		**76**	**82**	**34**	**48**	**10**	**84**	**22**
	ratio →	**3/2**	**3/2**	**3/2**	**3/2**	**3/2**	**3/2**	**3/2**

- ***a*** numbering of the twenty amino acids
- ***b*** number of atoms in the radical
- ***c*** number of atoms with an even number of electron shells (C, N and O)
- ***d*** number of atoms with a odd number of electron shells (H and S)
- ***e*** number of CH_2 groups (methylene bridge directly connected to alpha carbon)
- ***f*** rank of OMH hydrophobicity index: rank from the highest index to the lowest index**
- ***g*** number of codons in largest codon sets with 2 first same DNA bases

Figure 26: Some prime amino acid attributes. See Figure 25 also. *Radical only and in non-ionized state.** Exact OMH index values are listed in table Figure 44.

8.2. Atom number

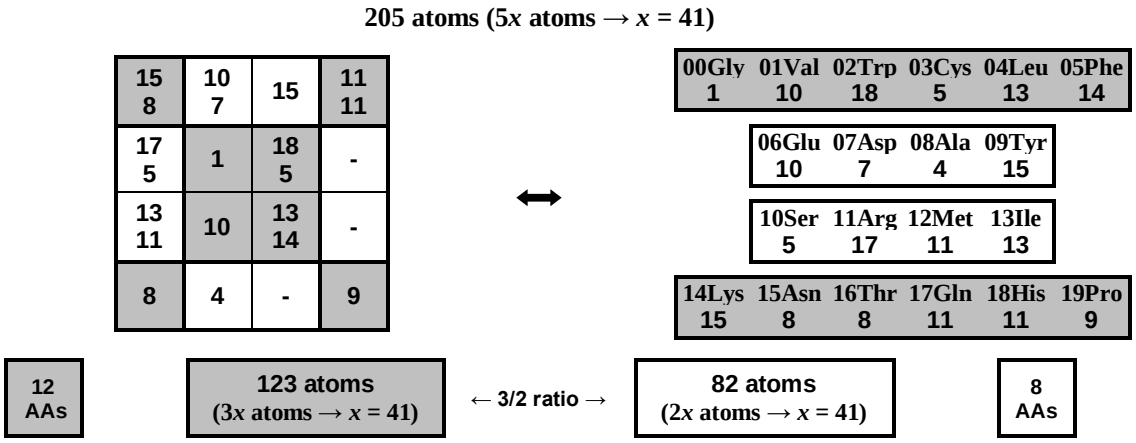


Figure 27: Atom number counting in radical of AAs: organization in the 3/2 ultimate ratio depending of AA numbering and ultimate genetic code matrix.

We recall here, as we illustrated by significant example in Figure 24 of the previous chapter, that the total number of atoms contained inside the radicals of the twenty amino acids is equal to 205. This number is therefore equal to 5*x* entities with *x* = 41.

Also, there are 123 atoms (3*x* atoms → *x* = 41) in the 12-AA dark area set and there are 82 atoms (2*x* atoms → *x* = 41) in the 8-AA light area set. Thus, these two sets are opposed in an exact ratio of 3/2 as value.

Figure 27 illustrates this distribution of atoms in the two twin matrices previously introduced: that of the ultimate genetic code and that of the numbering of the twenty proteinogenic AAs. We will then use the same representation to introduce the different studied entities of amino acids which are listed in Figure 26.

8.3. Atom number and quantum shells

The 3/2 ratio operates simultaneously within the genetic code in different aspects. Thus, the differentiation of two categories of atoms, respectively with an even number of electron shells and with a odd number of electron shells, oppose them in 3/2 ratios.

Indeed, the amino acids are (only) made up of Carbon, Nitrogen and Oxygen, three atoms with two electron shells and of Hydrogen and Sulphur, two atoms with one and three electron shells. Figure 28 illustrates this prime opposition of genetic code constituents in 3/2 values ratio.

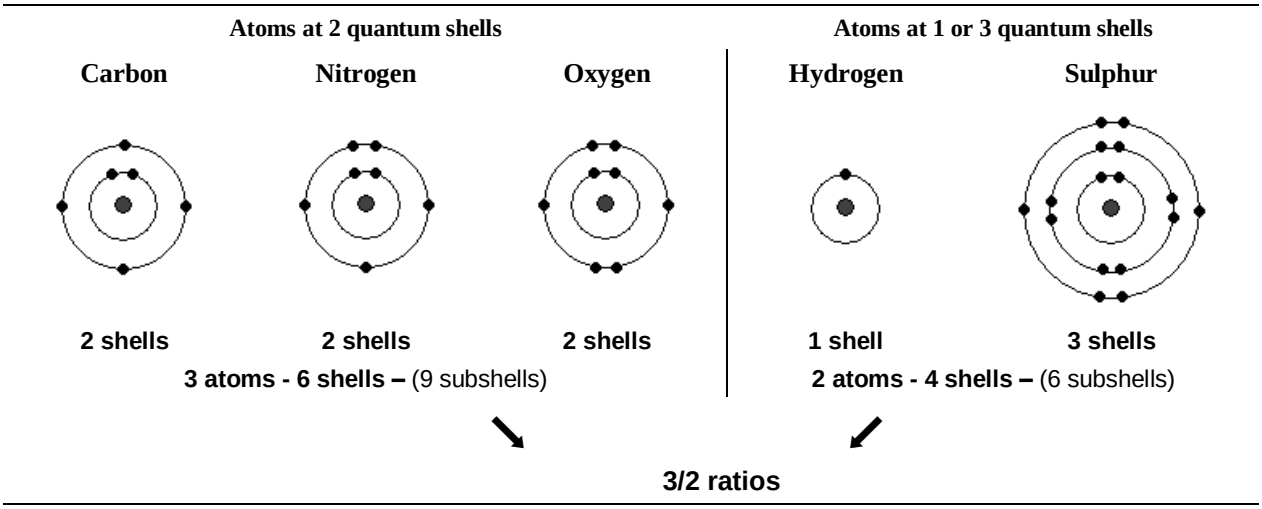


Figure 28: Differentiation of the 5 atoms constituting the 20 proteinogenic amino acids into 2 sets of 3 and 2 atoms according to the parity of their number of electron shells.

Figure 28 illustrates just a few oppositions between these two sets of three and two atoms. In a previous paper [4] we demonstrated that a very large number of their physico-chemical and so even quantum characteristics also oppose each other in exact ratios of value 3/2.

The imposing table Figure A1, presented in the appendix of this present paper, lists all these observations on these two sets of three and two atoms, chemical elements, components of the genetic code always opposing each other in various ratios of value 3/2.

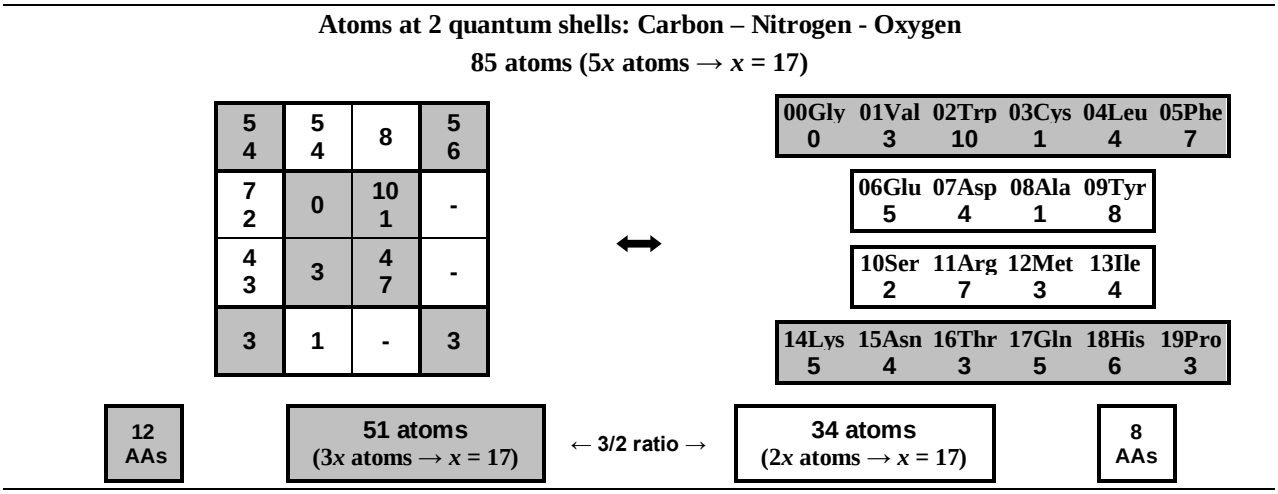


Figure 29: Counting of atoms at 2 quantum shells in radical of AAs: organization in the 3/2 ultimate ratio depending of AA numbering and ultimate genetic code matrix.

As it is visible in Figure 29 then the one 30, the separate counting of these two categories of atoms, which grouped together gave (Figure 28) already an opposition of the values in a 3/2 ratio according to the twin concept of numbering of the 20 amino acids and of ultimate genetic code, continues to generate the same exact ratios of 3/2 values.

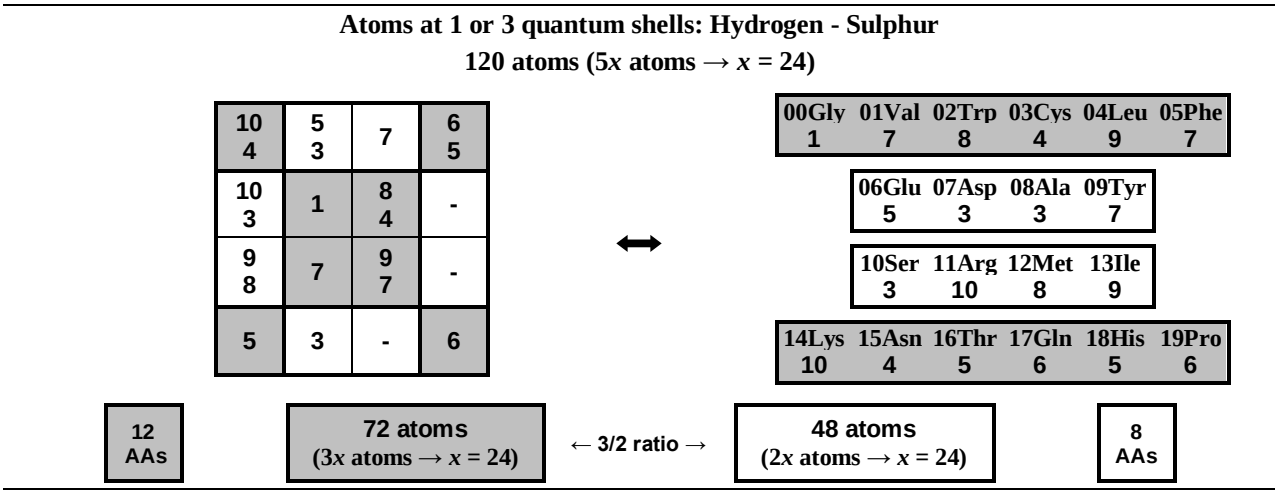


Figure 30: Counting of atoms at 1 or 3 quantum shells in radical of AAs: organization in the 3/2 ultimate ratio depending of AA numbering and ultimate genetic code matrix. See Figure 25 and 26.

8.4. CH_2 groups

Many of the twenty amino acids contains CH_2 groups (methylene bridges) in their radical. All these methylene bridges are located just after the alpha carbon either directly connected to it, alone or in a chain. There is however an exception to this in Isoleucine where the CH_2 group is not connected directly to the alpha carbon. Figure 31 illustrates these two configuration types.

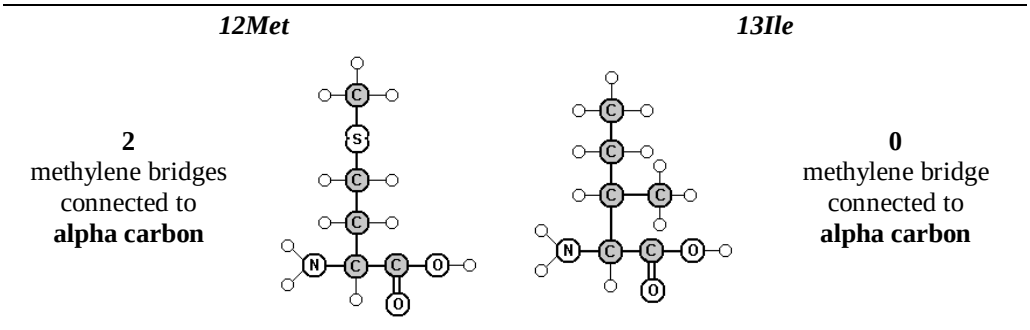


Figure 31: CH_2 groups (methylene bridges) differentiation. See figure 25 also.

There are therefore 26 CH_2 groups in all of the twenty amino acid radicals but just 25 directly connected (alone or in a chain) to the alpha carbon. Thus, this number is equal to 5*x* entities.

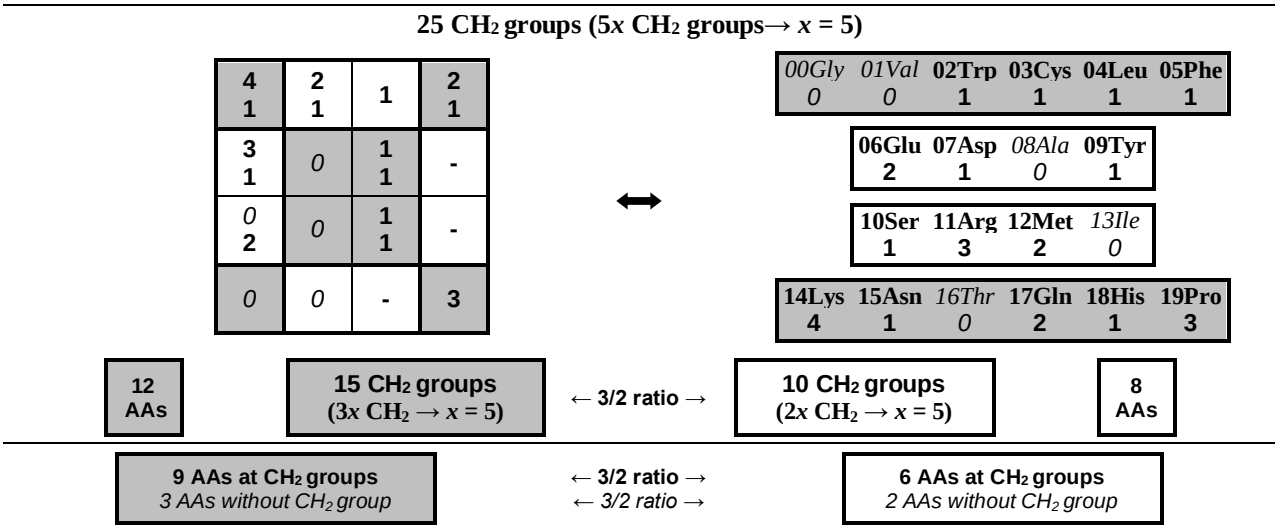


Figure 32: Counting of CH_2 groups in radical of AAs: organization in the 3/2 ultimate ratio depending of AA numbering and ultimate genetic code matrix. See Figure 25 and 26.

Also, as it appears in Figure 32, it turns out that there are 15 CH_2 groups (3*x* groups → *x* = 5) in the 12-AA dark area set and there are 10 CH_2 groups (2*x* groups → *x* = 5) in the 8-AA light area set. Thus, these two sets are opposed in an exact ratio of 3/2 as value.

*In fact, 15 AAs (so 5*x* AAs) contains CH_2 groups. We observe that these 15 AAs are also distributed in the ultimate 3/2 ratio with 9 located in the dark area versus 6 in the light area. The five AAs without group are in fact of the same 3/2 configuration.

8.5. OMH index rank

According to the exact values of the OMH scale shown Figure 44 in Chapter 11, we created (*f* in Figure 26) an index rank scale ranging from 1 (largest index) to 20 (lowest index) for the twenty amino acids.

8.5.1. Full OMH index rank

The cumulative value of these ranks gives a value of 210 (5*x* ranks→ *x* = 42) including 126 in the 12-AA dark area set and a value of 84 in the 8-AA light area set as this is illustrated Figure 33. So these two sets oppose in a 3/2 ratio.

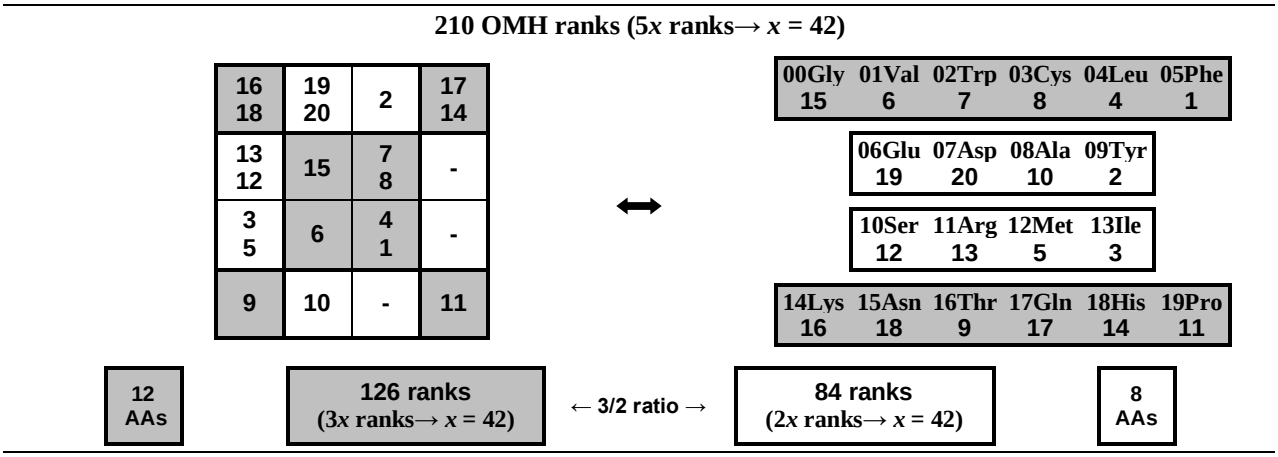


Figure 33: OMH index ranks distribution: organization in the 3/2 ultimate ratio depending of AA numbering and ultimate genetic code matrix.

The OMH index [5] is universally recognized in the study of the twenty proteinogenic amino acids and it is highly unlikely that this perfect arithmetic arrangement is so by pure chance. What emerges from the next demonstration will reinforce this point of view.

8.5.2. OMH index ranks parity

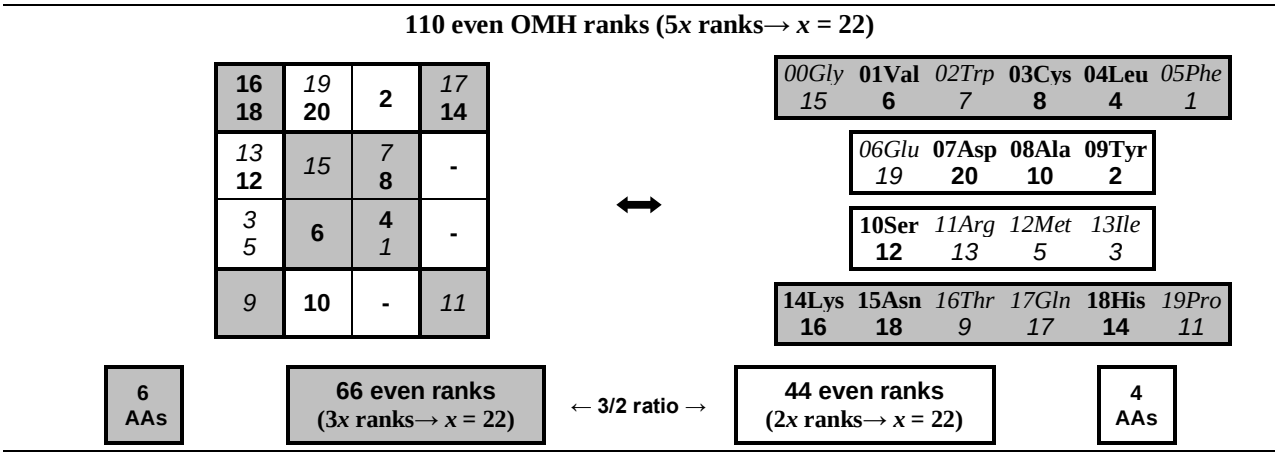


Figure 34: OMH index ranks distribution: organization in the 3/2 ultimate ratio depending of AA numbering and ultimate genetic code matrix.

Although the distribution of the different OMH index ranks (Figure 33) seems random within the two defined AAs sets of dark area and of light area, distribution of the even and odd isolated values continues to generate a perfect 3/2 ratio between these two sets as demonstrated in Figures 34 and 35. Also, independently of the consideration of the accumulation of values, six of the AAs at OMH even rank are in dark area compared to four in light area, i.e. a distribution in a 3/2 ratio. The ten other AAs of OMH odd rank are distributed in this same ratio.

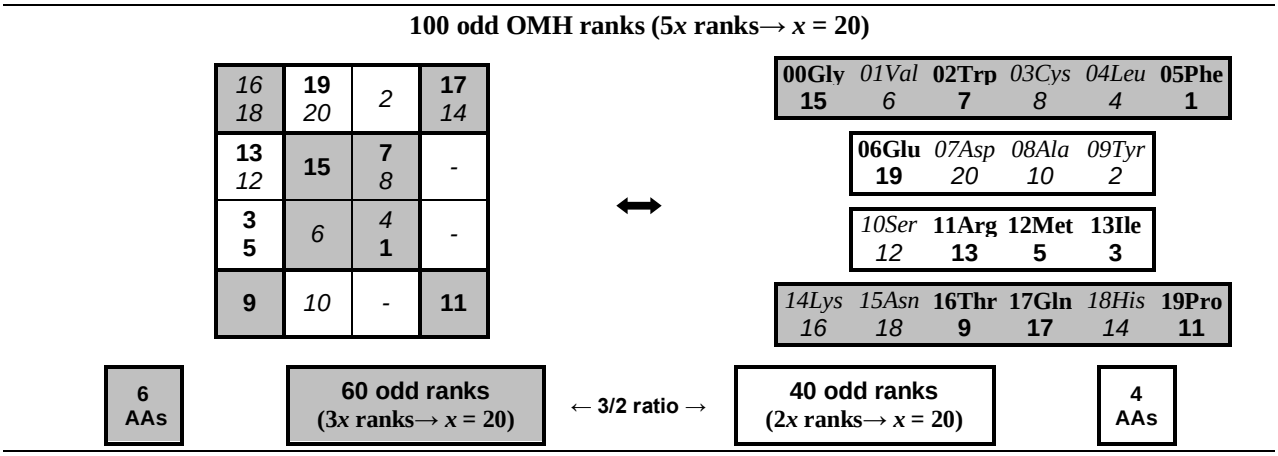


Figure 35: OMH index ranks distribution: organization in the 3/2 ultimate ratio depending of AA numbering and ultimate genetic code matrix.

8.5.3. OMH rank parity and ultimate genetic code source

In Chapter 5 and illustrated Figure 14, we have demonstrated that the two areas defined as dark and light are assemblies of symmetrical groups of coding boxes with the same reciprocal characteristics. Here we return to this illustration, Figure 36, showing that the amino acids are distributed in equal quantities according to the parity of their OMH rank in these subsets.

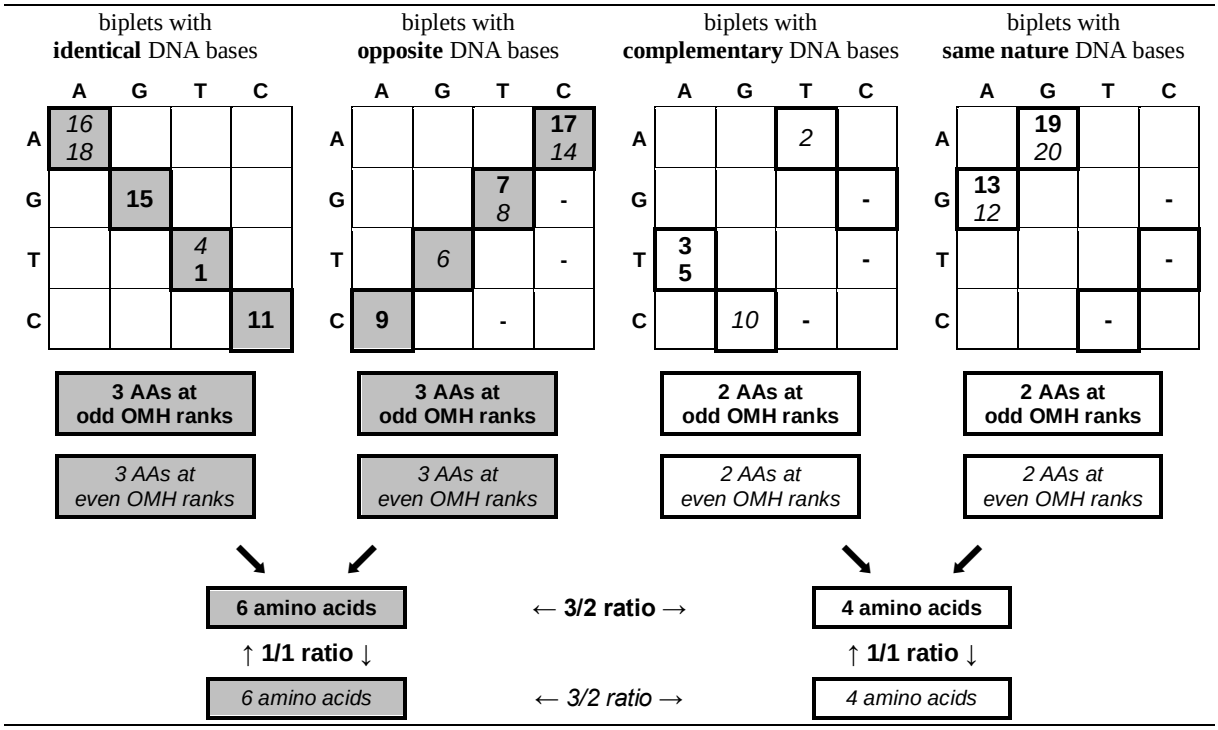


Figure 36: Related to the nature of the codons of the ultimate genetic code, the four amino acid configurations generating the ultimate 3/2 ratio and equal distribution of AAs at odd OMH rank and at even OMK rank.

This is just one of the very unusual phenomena linking the OMH scale and the ultimate codification (their respective biplets) of the twenty amino acids within the ultimate matrix of two previously defined areas. We will demonstrate this in Chapter 11 which follows.

8.6. Greater number of codons with the first two identical nucleobases

Due to the structural mechanics of the genetic code, i.e. the association of three out of four possible nucleobases to form a coding signal, 64 codons are necessary for the encoding of 20 amino acids. It turns out that each amino acid is associated with a seemingly random number of codons.

In fact, AAs are usually encoded with codons at the same first two identical nucleobases. In this universal genetic code, each amino acid is associated with a set of codons with two identical first bases varying from 1 to 4 codons.

Three amino acids are encoded with more than one of these sets at the same first nucleobases. Arginine, Leucine and Serine encoded by these 4-codon sets are also encoded with 2-codon sets.

Complete genetic code:
64 codons to 20 AAs and 1 stop signal

AAA	Lys	GAA	Glu	TAA	-	CAA	Gln
AAG	Lys	GAG	Glu	TAG	-	CAG	Gln
AAT	Asn	GAT	Asp	TAT	Tyr	CAT	His
AAC	Asn	GAC	Asp	TAC	Tyr	CAC	His
AGA	Arg	GGA	Gly	TGA	-	CGA	Arg
AGG	Arg	GGG	Gly	TGG	Trp	CGG	Arg
AGT	Ser	GGT	Gly	TGT	Cys	CGT	Arg
AGC	Ser	GGC	Gly	TGC	Cys	CGC	Arg
ATA	Ile	GTA	Val	TTA	Leu	CTA	Leu
ATG	Met	GTG	Val	TTG	Leu	CTG	Leu
ATT	Ile	GTT	Val	TTT	Phe	CTT	Leu
ATC	Ile	GTC	Val	TTC	Phe	CTC	Leu
ACA	Thr	GCA	Ala	TCA	Ser	CCA	Pro
ACG	Thr	GCG	Ala	TCG	Ser	CCG	Pro
ACT	Thr	GCT	Ala	TCT	Ser	CCT	Pro
ACC	Thr	GCC	Ala	TCC	Ser	CCC	Pro

Lighter genetic code:
55 codons to 20 AAs → 5*x* codons to 5*x'* AAs

AAA	Lys	GAA	Glu	*TAA*	*-*	CAA	Gln
AAG	Lys	GAG	Glu	*TAG*	*-*	CAG	Gln
AAT	Asn	GAT	Asp	TAT	Tyr	CAT	His
AAC	Asn	GAC	Asp	TAC	Tyr	CAC	His
AGA	*Arg*	GGA	Gly	*TGA*	*-*	CGA	Arg
AGG	*Arg*	GGG	Gly	TGG	Trp	CGG	Arg
AGT	*Ser*	GGT	Gly	TGT	Cys	CGT	Arg
AGC	*Ser*	GGC	Gly	TGC	Cys	CGC	Arg
ATA	Ile	GTA	Val	*TTA*	*Leu*	CTA	Leu
ATG	Met	GTG	Val	*TTG*	*Leu*	CTG	Leu
ATT	Ile	GTT	Val	TTT	Phe	CTT	Leu
ATC	Ile	GTC	Val	TTC	Phe	CTC	Leu
ACA	Thr	GCA	Ala	TCA	Ser	CCA	Pro
ACG	Thr	GCG	Ala	TCG	Ser	CCG	Pro
ACT	Thr	GCT	Ala	TCT	Ser	CCT	Pro
ACC	Thr	GCC	Ala	TCC	Ser	CCC	Pro

Figure 37: Complete genetic code (64 codons) and lighter genetic code (55 codons) with only, for each AA, consideration of the largest codons sets with the first two identical nucleobases.

We therefore propose, right part Figure 37, to consider only a reduced genetic code of these smaller sets since for these three AAs, a set of larger codons also encodes them. We therefore subtract from the initial 64 codons these three times two codons and we do not consider the three "Stop" codons either. Thus, we lighten the initial genetic code of 9 codons and therefore keep only 55 codons for the coding of the 20 amino acids.

This residual number of 55 codons is therefore equal to 5*x* entities. The final count of these 55 residual codons in the two predefined sets of 12 and 8 AAs respectively qualified as external and internal (dark and light matrix areas) generates here also an exact ratio of value 3/2 with 33 versus 22 counted codons as this is illustrated Figure 38.

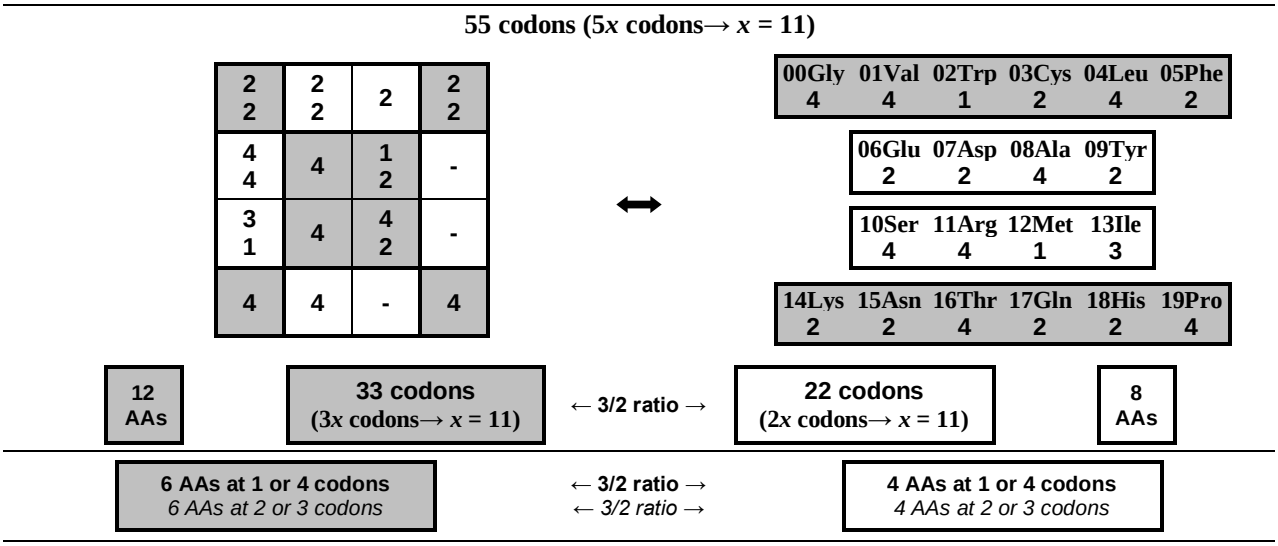


Figure 38: Codon number in largest codon sets with 2 first same DNA bases: organization in the 3/2 ultimate ratio depending of AA numbering and ultimate genetic code matrix. See Figures 37 and 39 also.

8.6.1. Greater number of codons and ultimate genetic code source

Also, independently of the consideration of the accumulation of values, It turns out that 10 AAs have 1 or 4 codons (complete lighter code) and 10 others have 2 or 3 codons. Six of the AAs at 1 or 4 codons are in dark area compared to four in light area, i.e. a distribution in a 3/2 ratio. The ten other AAs at 2 or 3 codons are distributed in this same ratio.

Finally, Figure 39, exactly as about the parity of OMH ranks, it turns out that in symmetrical groups of coding boxes with the same reciprocal characteristics, the ten amino acids at 1 and 4 codons and the ten at 2 or 3 codons are distributed in equal quantities according to this criteria in these four sub areas.

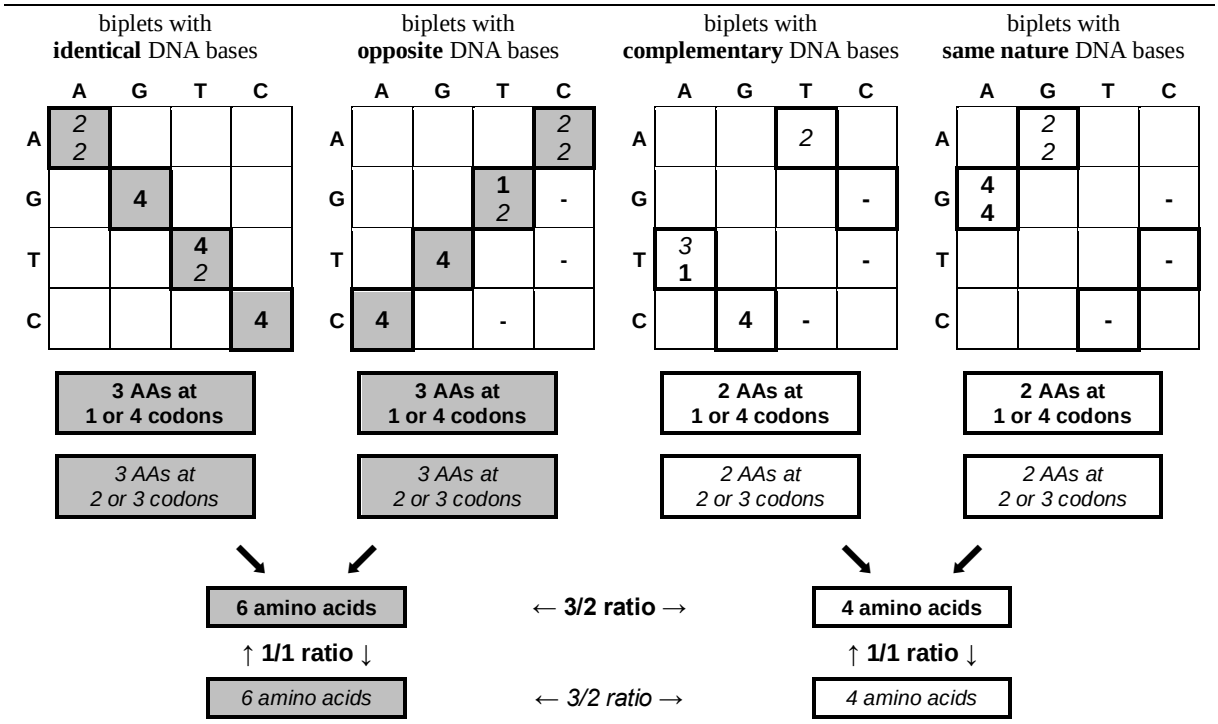


Figure 39: Related to the nature of the codons of the ultimate genetic code, the four amino acid configurations generating the ultimate 3/2 ratio and equal distribution of AAs at 1 or 4 codons and at 2 or 3 codons from complet lighter genetic code. See Figure 37 and 38 also.

8.7. Synthesis of the distribution of the AA prime attributes

Figure 40 summarizes the main attributes of the twenty amino acids that we just studied. It should therefore be noted that there are always 5*x* in number and that they are also always divided into 3*x* entities in the set of 12 AAs with external numbering (dark matrix area) and into 2*x* entities in that of 8 AAs with internal numbering (light matrix area).

Genetic code entities	Total Entities number	12 external numbering amino acids	8 internal numbering amino acids
20 amino acids	**20** **$5x \rightarrow x = 4$**	**12** **$3x \rightarrow x = 4$**	**8** **$2x \rightarrow x = 4$**
205 radical atoms	**205** **$5x \rightarrow x = 41$**	**123** **$3x \rightarrow x = 41$**	**82** **$2x \rightarrow x = 41$**
85 radical atoms at even number of electron shells (C - N - O)	**85** **$5x \rightarrow x = 17$**	**51** **$3x \rightarrow x = 17$**	**34** **$2x \rightarrow x = 17$**
120 radical atoms at odd number of electron shells (H - S)	**120** **$5x \rightarrow x = 24$**	**72** **$3x \rightarrow x = 24$**	**48** **$2x \rightarrow x = 24$**
25 CH_2 groups (methylene bridges connected to alpha carbon)	**25** **$5x \rightarrow x = 5$**	**15** **$3x \rightarrow x = 5$**	**10** **$2x \rightarrow x = 5$**
210 ranks of OMH hydrophobicity index (from 1 to 20)	**210** **$5x \rightarrow x = 42$**	**126** **$3x \rightarrow x = 42$**	**84** **$2x \rightarrow x = 42$**
110 even ranks of OMH hydrophobicity index (from 2 to 20)	**110** **$5x \rightarrow x = 22$**	**66** **$3x \rightarrow x = 22$**	**44** **$2x \rightarrow x = 22$**
100 odd ranks of OMH hydrophobicity index (from 1 to 19)	**100** **$5x \rightarrow x = 20$**	**60** **$3x \rightarrow x = 20$**	**40** **$2x \rightarrow x = 20$**
55 codons (largest codon sets with 2 first same nucleobases)	**55** **$5x \rightarrow x = 11$**	**33** **$3x \rightarrow x = 11$**	**22** **$2x \rightarrow x = 11$**

Figure 40: Depending on the two defined areas of the ultimate genetic code matrix, synthesis of the distribution of the prime attributes (to 5*x* in number) related to the 20 proteinogenic amino acids in exact 3/2 ratios.

In view of this first synthesis, and because they concern very different aspects, it seems very unlikely that all these physico-arithmetic arrangements are so by chance. The results of the next more subtle investigations will reinforce this hypothesis.

9. Atom number and symmetry of the ultimate genetic code matrix

In accordance with their respective number of atoms, the twenty proteinogenic amino acids can be separated into two sets:

- those whose radical contains 1 to 10 atoms,

- those whose radical contains more than 10 atoms.

It turns out that each of these two sets is made up of exactly 10 AAs and these two groups also correspond to the ten AAs at lowest and the ten at highest van der Waals volume.

Also, as previously illustrated Figure 27 in Chapter 8.2, these two sets of 10 AAs distribute in a 3/2 ratio according of the two areas of the ultimate genetic code matrix with for each set, 6 AAs in the dark area versus 4 in the light area.

9.1. Atom number and eight symmetrical areas

In chapter 4.3 and illustrated Figures 8a and 8b we identified symmetrical areas in the matrix of the ultimate genetic code that always contain a total of ten amino acids (and of course 10 associated biplets). These different coding areas therefore correspond to the different physicochemical attributes of the DNA bases, which legitimizes their consideration.

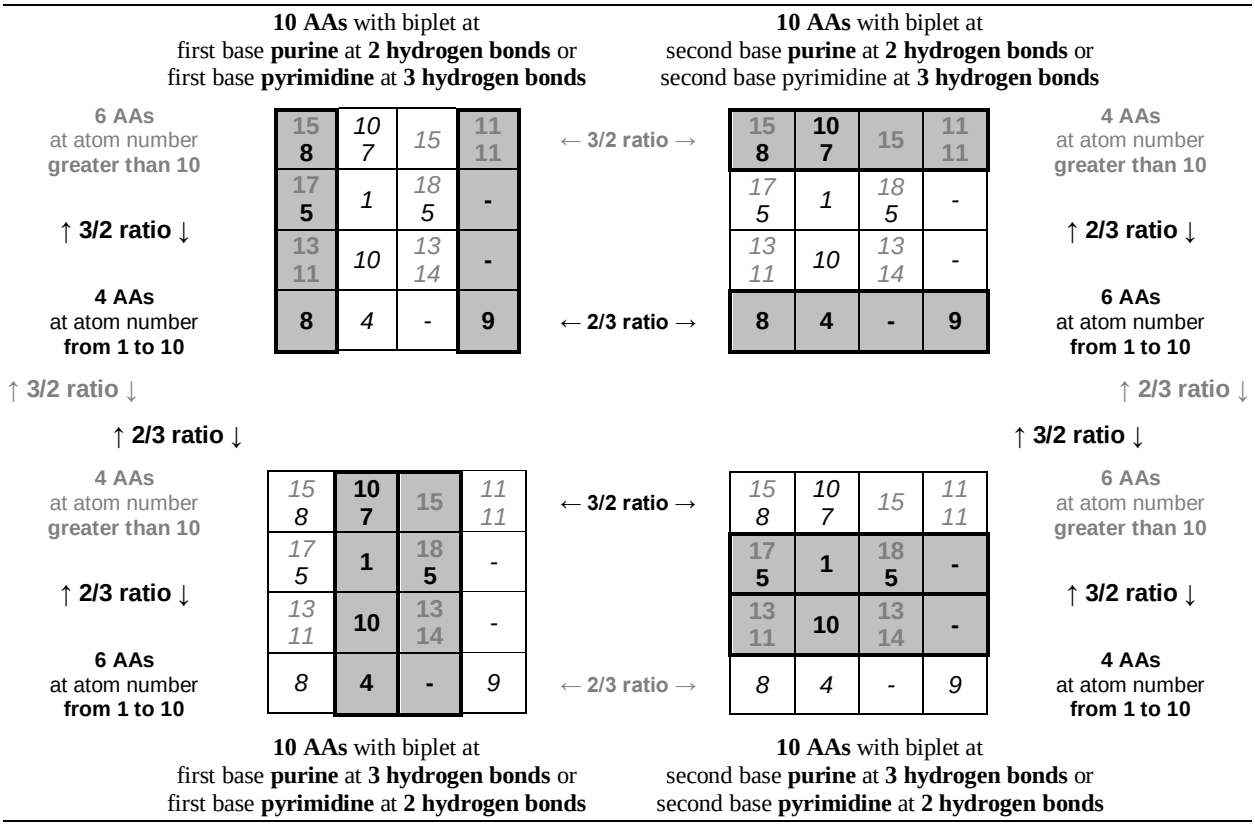


Figure 41: According to their number of atoms (radical only) distribution of amino acids in the symmetrical areas of the matrix of the ultimate genetic code always in an opposition of 3*x* versus 2*x* entities. See Figure A8 also.

Figure 41 then the following Figure 42, illustrate a perfect opposition in always a ratio of value 3/2 of the sets of amino acids with a number of atoms between 1 and 10 and those with a number of atoms greater than 10. This by considering all the symmetric areas of the ultimate matrix of the genetic code which total 10 AAs.

Thus, considering the matrix of the ultimate genetic code, in each of these eight configurations of ten amino acids, six are with the lowest number of atoms and four with the greatest number, or vice versa. From a more physical aspect, in these eight configurations linked to the genetic coding, there are therefore always either six amino acids with the smallest van der Walls volume and four with the largest volume, or the reverse.

At the end of the appendix, Chapter A5, is presented in Figure A8 a recombination of Figure 41 where other symmetrical configurations of 10 versus 10 AAs present the same singular characteristics in accordance with the number of atoms of the amino acids.

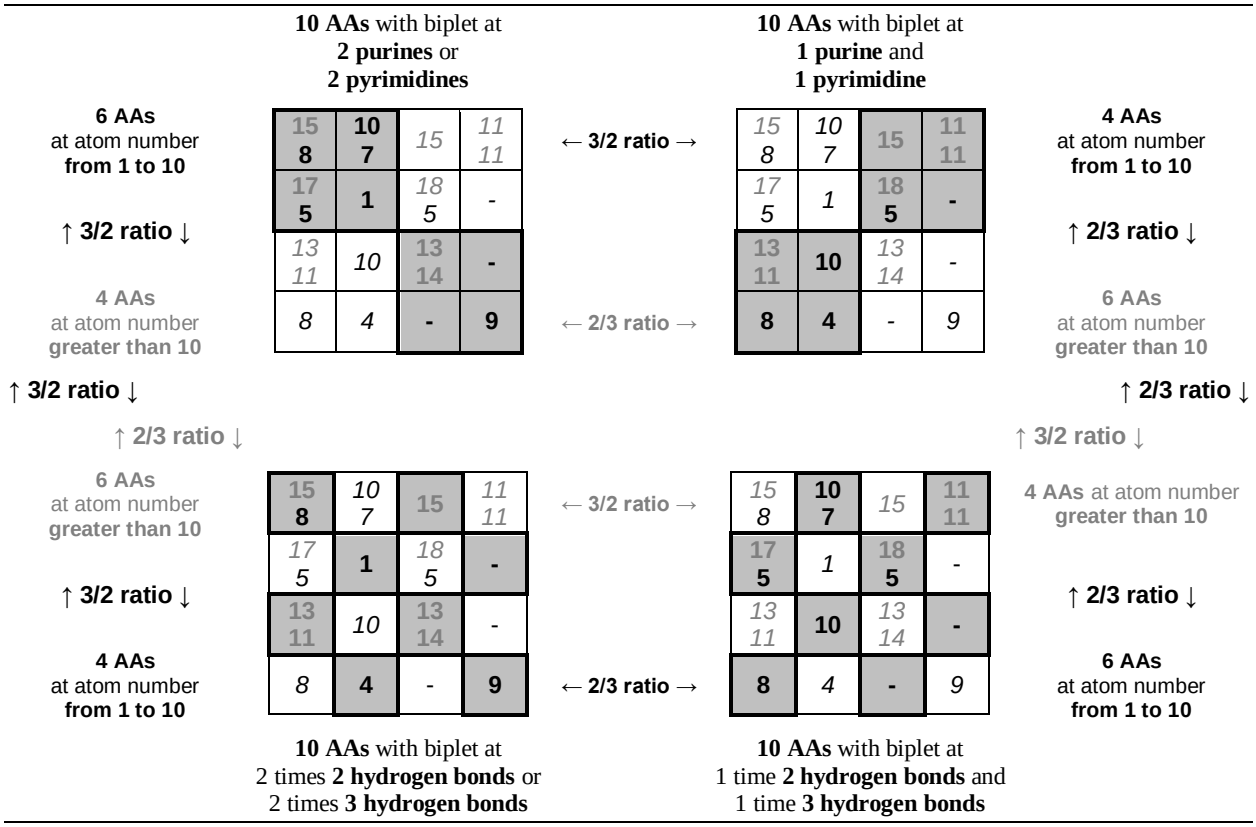


Figure 42: According to their number of atoms (radical only), distribution of amino acids in the symmetrical areas of the matrix of the ultimate genetic code always in an opposition of 3*x* versus 2*x* entities.

9.2. Atom number and four symmetric or asymmetric areas

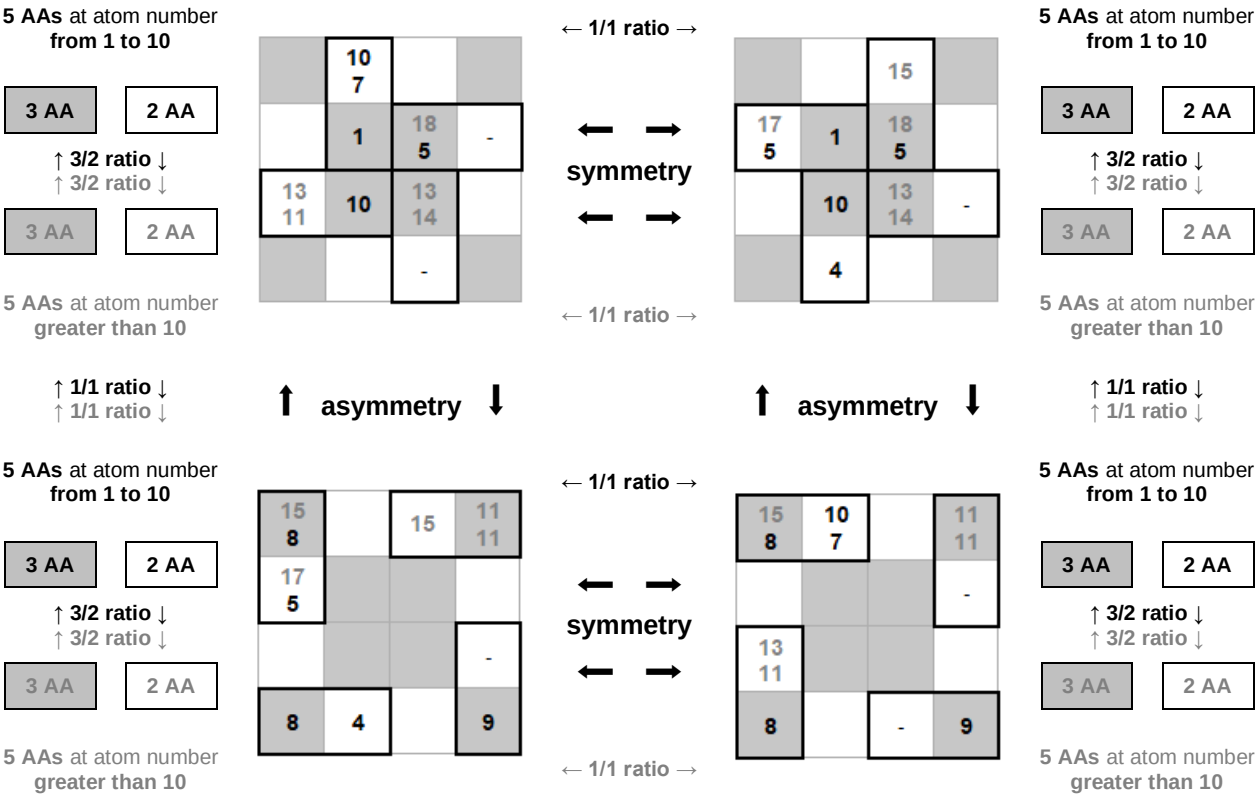


Figure 43: According to their number of atoms (radical only), distribution of amino acids in the symmetrical areas of the matrix of the ultimate genetic code always in an global opposition of 5*x* versus 5*x* entities then of 3*x* versus 2*x* entities according to the two defined dark or light areas. See Figure 11 also.

In Figure 11 Chapter 4, within the ultimate matrix of the genetic code, we identified eight double areas with intriguing arithmetic properties. These areas of two adjacent boxes are all made up of from 4 to 1 entities, i.e. containing from 4 to 1 amino acids (and from 4 to 1 corresponding biplets). We demonstrated that their symmetrical associations always grouped together 5 AAs.

Figure 43 clearly illustrates how the differentiation of the 10 smallest AAs from the 10 largest (i.e. from those with 10 atoms maximum to those with more than 10 atoms) is intimately, and of very entangled way, in relation to the particular configuration of the matrix of the genetic code that we have called ultimate.

Thus, in these four geometrically coding zones, all made up of 4 dark boxes and 4 light boxes, there are always 5 AAs with the highest number of atoms and 5 AAs with the lowest number of atoms. Also, each time, 3 of these 5 AAs are located in the dark area compared to 2 in the light area.

10. OMH hydrophobicity index ranks and ultimate genetic code matrix

According to the exact values of the OMH scale shown in the left part Figure 44, we created an index rank scale ranging from 1 (largest index) to 20 (lowest index) for the twenty amino acids.

The cumulative value of these ranks gives a amount of 126 for the dark area set of AAs and 84 for the light one as this is illustrated in right part Figure 44.

Amino acid	OMH scale	rank of OMH hydrophobicity index*
00Gly	-0.67	**15**
01Val	0.91	**6**
02Trp	0.5	**7**
03Cys	0.17	**8**
04Leu	1.22	**4**
05Phe	1.92	**1**
06Glu	-1.22	**19**
07Asp	-1.31	**20**
08Ala	-0.4	**10**
09Tyr	1.67	**2**
10Ser	-0.55	**12**
11Arg	-0.59	**13**
12Met	1.02	**5**
13Ile	1.25	**3**
14Lys	-0.67	**16**
15Asn	-0.92	**18**
16Thr	-0.28	**9**
17Gln	-0.91	**17**
18His	-0.64	**14**
19Pro	-0.49	**11**

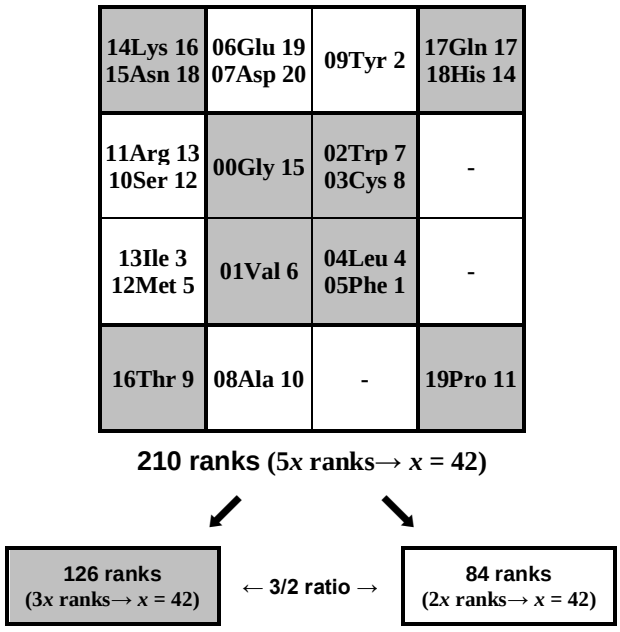


Figure 44: OMH index ranks and these distribution in exact 3/2 ratio into the two sets of AAs: 12-AAs dark area and 8-AAs light area. * Rank from the highest index to the lowest index.

10.1. Total entanglement of OMH index and ultimate genetic code matrix

In previous Chapter 8.5, we have demonstrated the importance of differentiating the twenty OMH index ranks according to their parity.

We will now reveal an entanglement between this concept of index rank parity and the numbering parity of the twenty proteinogenic amino acids. All this super entangled with the order of magnitude of these ranks and the two predefined areas of the matrix of the ultimate genetic code. We recall that all these concepts depend on physicochemical properties either linked to the AAs or to the associated DNA bases.

What follows (like the whole of this paper) therefore has nothing to do with any exercise in numerology although we are handling numbers here as an entity of study.

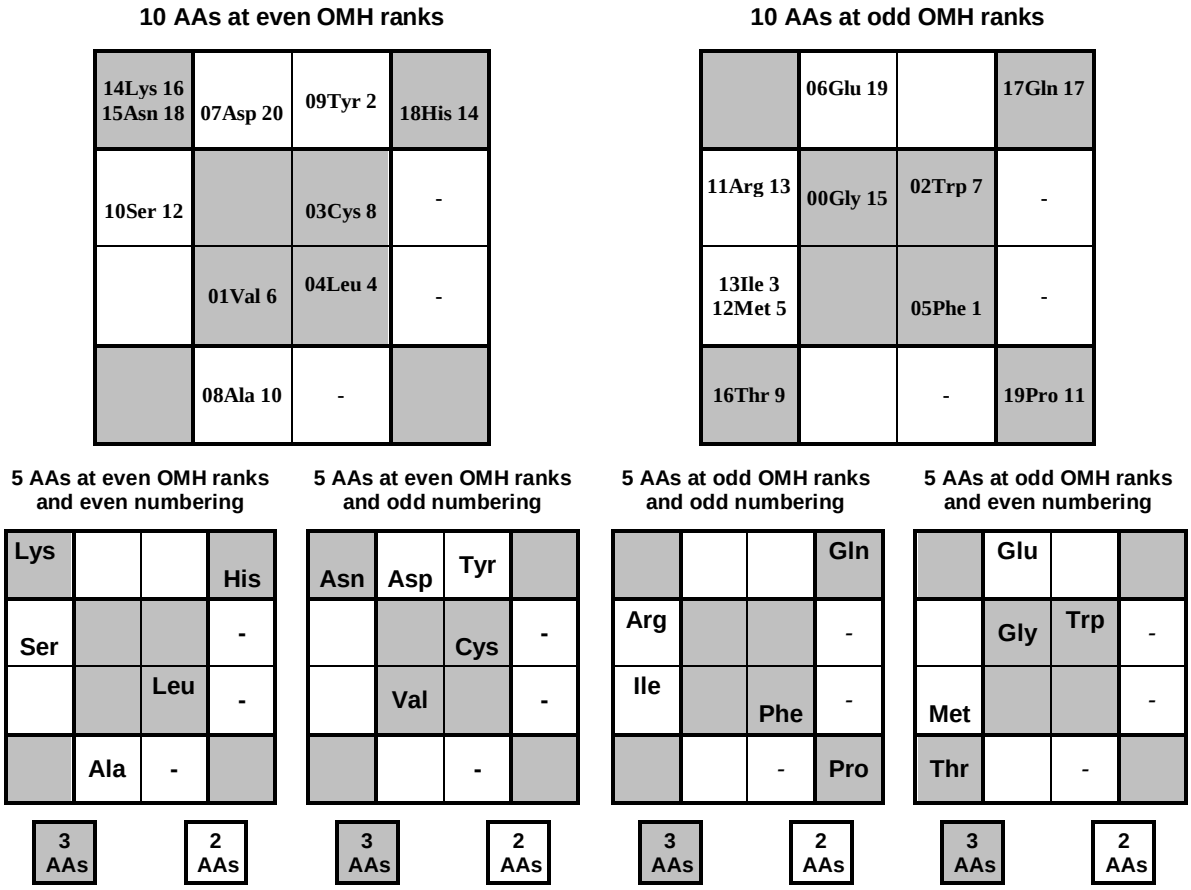


Figure 45: Entanglement of OMH rank parity, AA numbering parity and ultimate genetic code matrix concept. See Figure 46 also.

As illustrated in Figure 45, it turns out that among the 10 AAs with even rank of OMH index there are 5 AAs with even numbering and 5 others with odd numbering. The same phenomenon operates for the 10-AA set with odd OMH index rank. Also each of these four subsets of 5 entities formed according to this double parity criterion is organized in an ultimate 3/2 ratio with always 3 dark area AAS versus 2 light area AAs.

10.2. Total fractal organization of the ultimate genetic code

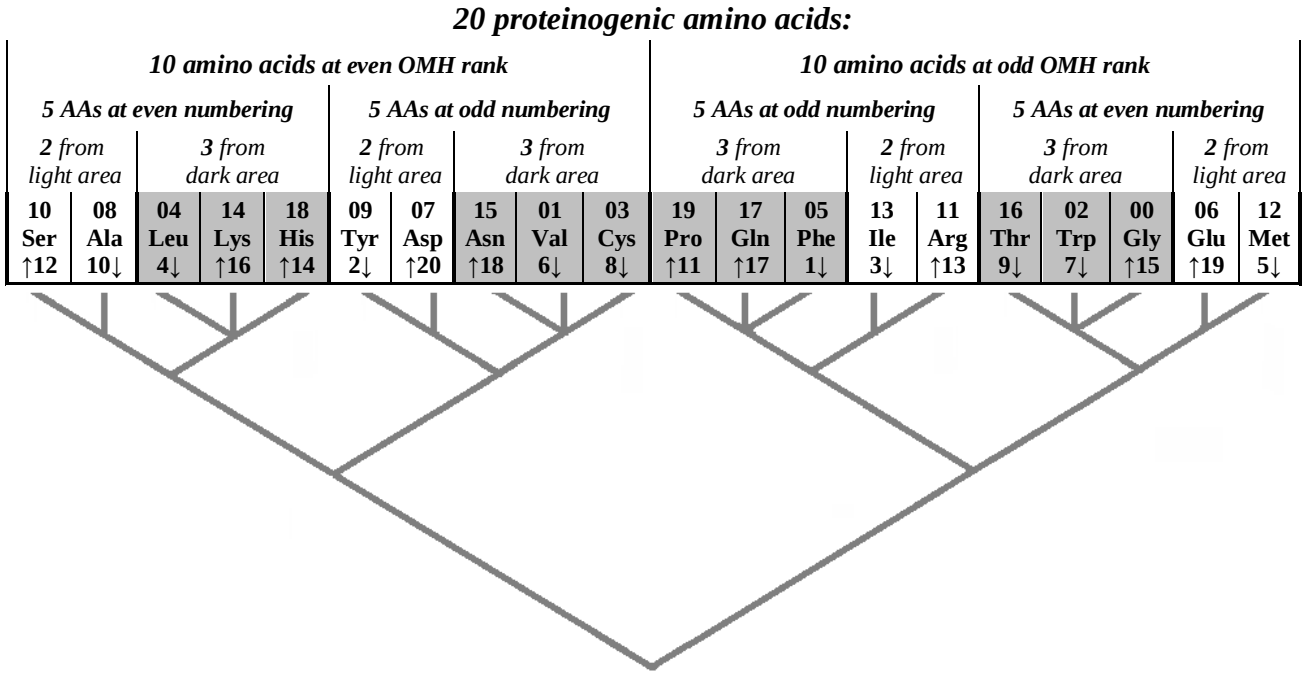


Figure 46: Symmetric perfect fractal distribution of amino acids according to four criteria: OMH rank parity, AA numbering parity, AA genetic coding and OMH rank amplitude (↓↑). See Figures 45 and 47 also.

Figure 46, and complementary Figure 47, illustrate more subtly the total entanglement of phenomena which are organized according to a perfect fractal concept.

This symmetric fractal representation makes better appear how we go from 20 entities to the final ratio 3/2. Indeed, from the twenty entities of the genetic code that are the proteinogenic amino acids, two sets of 10 entities can be isolated according to different physico-chemical criteria: OMH rank parity, AA numbering parity*, OMH rank amplitude and nucleobase character of respective proteinogenic amino acids in the matrix of the ultimate genetic code.

Depending to them numbering parity, these two sets can each be split into two subsets of 5 AAs. Then each of these subsets can be separated into sets with ultimate numbers of 3 and 2 entities. Finally, according to codifying criteria, within these subsets of 2 or 3 entities, there is always 1 amino acid with the highest OMH rank and 1 with the lowest rank in the light area subsets and 1 versus 2 of the highest or lowest rank in the dark area subsets.

*We recall, as explained in the Chapter 6, that the numbering of the twenty amino acids is highly dependent on the physico-chemical properties of the four DNA bases.

Figure 47, focus on the lower end of this fractal mapping highlights clarifies how the ultimate genetic code is subtly organized.

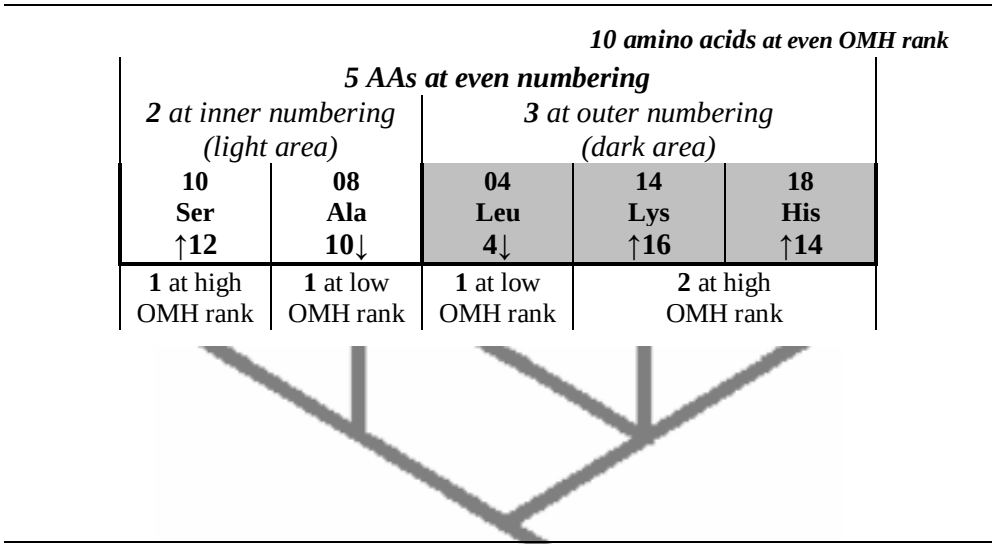


Figure 47: Focus on the final fractal distribution of the 20 AAs described in mapping Figure 46.

Thus, overall, according to four depths of physicochemical criteria, 12 amino acids have a unique range of properties that they do not share with any other AA and 8 other AAs have a unique range of properties shared with just one other AA.

As an example, Serine (*10Ser*) from light area, it's the only one at high OMH rank, at inner numbering, at even AA numbering and at even OMH rank. Lysine and Histidine (*14Lys* and *18His*), from dark area, are the only two at high OMH rank, at outer numbering, at even AA numbering and at even OMH rank. These two sets of 12 and 8 amino acids oppose in a perfect ratio of 3 to 2 in a intricate fractal configuration.

11. The 3/2 ratio and genetic code organization

As a preliminary conclusion, it seems essential to us to speak about the importance of the arithmetic ratio of value 3/2 in the organization of the genetic code.

The ultimate genetic code concept and the one of numbering of the twenty proteinogenic amino acids is not the only concepts to generate singular arithmetic phenomena opposing the entities of the genetic code in various ratios of value 3/2. In a preview paper *"Genetic code, quantum physics and the 3/2 ratio"* [4], we have revealed in great detail, a multitude of arithmetic arrangements of the components of the genetic code in this 3/2 ratio.

For example, we are drawing attention to the fact that Glycine, which is simply like an amino acid base, has all these various components at *5x* in number (10 atoms, 40 protons, 75 nucleons, etc.) and that these can be opposed in 3*x* and 2*x* in number. The same phenomena are also observed in the composition of the five atoms constituting the twenty proteinogenic amino acids (Hydrogen, Carbon, Nitrogen, Oxygen and Sulphur) which can also be opposed in various ratios of 3/2 values. Finally, depending on whether or not they are organic, the first ten chemical elements also oppose their nuclear charge number (atomic number) in a ratio of value 3/2. These many observations confirm the main idea of this article that the genetic code, confused AAs and nucleobases, is arithmetically organized according with the ratio 3/2. Some of these singular phenomena are more fully illustrated in the appendix at the end of this paper.

Also, various other genetic code investigations from many authors are in relationship with the subject of this paper especially about ratio 3 to 2, symmetry, listing of proteinogenic amino acids or more generally connections between number theory and the genetic code. As example and not limited to, some of these investigations are listed in references [6 to 14].

From this reference list, we draw particular attention to two of our previous preprints. In *"Preproinsulin molecule and numbering of the twenty proteinogenic amino acids"* [6], we demonstrate that the amino acid sequence of the 110-amino acid preproinsulin, the initial product of the translation of insulin mRNA, is in close dependence with the twice concept of the numbering of the twenty proteinogenic amino acids and the ultimate genetic code. In fact, It turns out that the orders of occurrence of the various preproinsulin amino acids, both direct and inverse sequence, are organized in numerous ratios of exact value 3/2. This, according to the amino acid numbering concept. The degree of abundance of these amino acids in this initial single-chain molecule reveals same numerical rational phenomena.

In *"Amino acid numbering, ultimate numbers and the 3/2 ratio"* [7], we demonstrate relationships between the mechanism of the genetic code, of the field of Biology, and the number theory field, of which more precisely the notion of *ultimate number* [15], one of the four classes of Mathematics entities proposed to constitute the set of whole numbers. These connections are revealed in an physico-arithmetic organization of the genetic code in various ratios of 3/2 value as global configurations.

12. Discusions and conclusions

By the construction of a matrix of the genetic code comprising only twenty double entities, namely simply the twenty proteinogenic amino acids each associated with a single biplet, we revealed a coding structure presenting both a symmetrical and asymmetrical character. Within this matrix of the genetic code that we call "ultimate", we have identified two sets of amino acids opposing each other in a perfect ratio of value 3 to 2.

These two sets of 12 and 8 AA which correspond to the criteria coding them, correspond very precisely to the two sets identified according to their numbering, a concept previously introduced in the published paper *"Numbering of the twenty proteinogenic amino acids and new alphanumerical nomenclature proposal to them"* [1]. This initial concept is also closely dependent on the DNA coding of the twenty amino acids and therefore has every legitimacy of consideration.

The very sophisticated, very complex but also very harmonious mapping of the different attributes of the AAs and associated biplet, within this matrix of the ultimate genetic code, suggests an impossible chance evolution of the genetic code. On the contrary, it appears that the genetic code is subject to strong physico-numerical constraints. As the periodic table of elements is entirely built on number entities, it is the same with regard to the structure of the genetic code.

Thus, the properties of the twenty amino acids, quantified by numbers and the properties of the four nucleobases assembled in triplets as well in biplets and also number quantified, are totally interdependent in an entangled whole. Indeed, it is always in a global way that the different phenomena of symmetry and asymmetry are observed in the matrix of the ultimate genetic code.

The new investigations presented here reinforce our conviction to propose a new aphanumerical nomenclature for the twenty canonical proteinogenic amino acids.

We have therefore firmly established, in many aspects, that the various characteristics of the twenty proteinogenic amino acids are closely linked to their numbering, which itself depends on their DNA codification. Also this numbering is in coincidence with special arrangements of the twenty AAs in the ultimate genetic code matrix. This is why we suggest here, to enrich the current nomenclature applied to these twenty entities, the creation of new standardized alphanumeric symbols making it possible to identify these twenty proteinogenic amino acids.

References

1. Jean-Yves Boulay. Numbering of the twenty proteinogenic amino acids and new alphanumerical nomenclature proposal to them. Symmetry: Culture and Science Volume 34, Number 1, pages 061-086. 2023. https://doi.org/10.26830/symmetry_2023_1_061.

2. Jean-Yves Boulay. Numbering of the twenty proteinogenic amino acids: 3/2 ratios inside the genetic code. 2022. https://www.researchgate.net/publication/363952852.

3. S.V. Petoukhov. Genetic Code and the Ancient Chinese Book Of Changes. Symmetry: Culture and Science Vol. 10, Nos. 3-4, p. 211-226. 1999. https://www.researchgate.net/publication/285143425 .

4. Jean-Yves Boulay. Genetic code, quantum physics and the 3/2 ratio. 2020. https://www.researchgate.net/publication/343064455.

5. Sweet R.M., Eisenberg D. Optimized matching hydrophobicity (OMH). J. Mol. Biol. 171:479-488. 1983; https://doi.org/10.1016/0022-2836(83)90041-4.

6. Jean-Yves Boulay. Preproinsulin molecule and numbering of the twenty proteinogenic amino acids. 2023. https://www.researchgate.net/publication/369880921.

7. Jean-Yves Boulay. Amino acid numbering, ultimate numbers and the 3/2 ratio. 2022. https://www.researchgate.net/publication/364374047

8. Ming Liu. Biosynthesis, structure, and folding of the insulin precursor protein. Volume20, IssueS. 2Supplement: Update on Islet Hormone Production: A Tribute to Donald Steiner. Proceedings of the 19th Servier - IGIS Symposium, St Jean Cap Ferrat, France, 22 – 25 March 2018. https://doi.org/10.1111/dom.13378.

9. S.V. Petoukhov.The Bi-periodic Table of Genetic Code and Number of Protons, Foreword of K. V. Frolov, Moscow, 258. 2001.

10. Then, A., Mácha, K., Ibrahim, B. et al. A novel method for achieving an optimal classification of the proteinogenic amino acids. Sci Rep 10, 15321, 2020; https://doi.org/10.1038/s41598-020-72174-5.

11. Wohlin A. Numerical analysis of 3/2-relations in the genetic code and correlations with the basic series of integers 5-0. Biomed Genet Genomics 1, 2016; https://doi.org/10.15761/BGG.1000118.

12. Petoukhov S., He M. Symmetrical Analysis Techniques for Genetic Systems and Bioinformatics: Advanced Patterns and Applications. IGI Global, Hershey, USA, p. 271, 2009; https://doi.org/10.4018/978-1-60566-124-7.

13. Darvas G., Koblyakov A.A., Petoukhov S.V., Stepanyan I.V. Symmetries in molecular-genetic systems and musical harmony. Symmetry Culture and Science, vol. 23, №3-4, p. 343, 2012; http://symmetry.hu/scs_online/SCS_23_3-4.pdf.

14. Petoukhov S.V. Genetic code, musical harmony, stochastic resonance and the Ancient Chinese book of I-Ching. Editor-in-Chief Solar G. p. 160-180, 2022, . Published by: The First Clinic of Acupuncture and Natural Medicine of G.Solar, Ltd, Samorin, Slovak Republic. ISBN 978-80-974284-1-9. EAN 9788097428419. This publication was created thanks to support from the European Union Erasmus program, project number 2020-1-SK01-KA202-078222; https://www.acuclinic.eu/ecompendium/.

15. Jean-Yves Boulay. The ultimate numbers and the 3/2 ratio. Just two primary sets of whole numbers. 2023. https://www.researchgate.net/publication/339943634

APPENDIX

Some of phenomena presented here are taken from the author's previous paper: *Jean-Yves Boulay. Genetic code, quantum physics and the 3/2 ratio. 2020.* [4].

This appendix is of a global theme and sometimes without direct references to the different chapters introduced previously. Its important relationship, however, is the organization of the genetic code around a numerical constraint revealed in an exact ratio of value 3 to 2.

A1. Anatomy of Glycine as 3/2 ratio

A1.1. Glycine as glycined base

Within the mechanism of the genetic code and therefore among the twenty amino acids, Glycine is distinguished by its absence of radical. Its radical is reduced to a simple hydrogen atom which in a way simply closes the "base" structure common to each amino acid. The quantum study of this *glycined base*, identifying with Glycine, reveals singular arithmetic arrangements of its different components.

A1.2. Modules of Petoukhov

The notion of modules is an original system proposed by Sergey Petoukhov [3 and 8] to describe the structure of biological molecules. According to this genetic code researcher, in organic chemistry, module is a group formed of just one non-hydrogen atom with possibly its satellite hydrogen atoms attached. Also, Sergey Petoukhov considers Sulphur as constituted in a twice module.

A1.3. Detailed structure of Glycine

Figure A1 describes the structure of Glycine (or saturated base called *glycined base*) according to many criteria including its chemical composition, modular, but also atomic. It turns out that Glycine consists of 40 protons, either 5x protons or (3 + 2)x protons.

This glycined base also consists of 5 groups or modules, i.e. (3 + 2)x chemical groups. In Glycine, the number of protons is therefore an exact multiple of 8 (5 times 8 protons) and it turns out that the average number of protons per chemical group (or Petoukhov module) is therefore 8. For two groups (CH_2 and O), the amount of protons is exactly 8 whereas for the other three groups, these proton amounts are 9 or 6 ($NH_2 \rightarrow 9$, $OH \rightarrow 9$ and $C \rightarrow 6$).

The differentiation of these two types of modules, made up or not made up of 8 protons reveals a multitude of oppositions of the different natures of the components of Glycine (glycined base) in always an arithmetical ratio of 3/2 value.

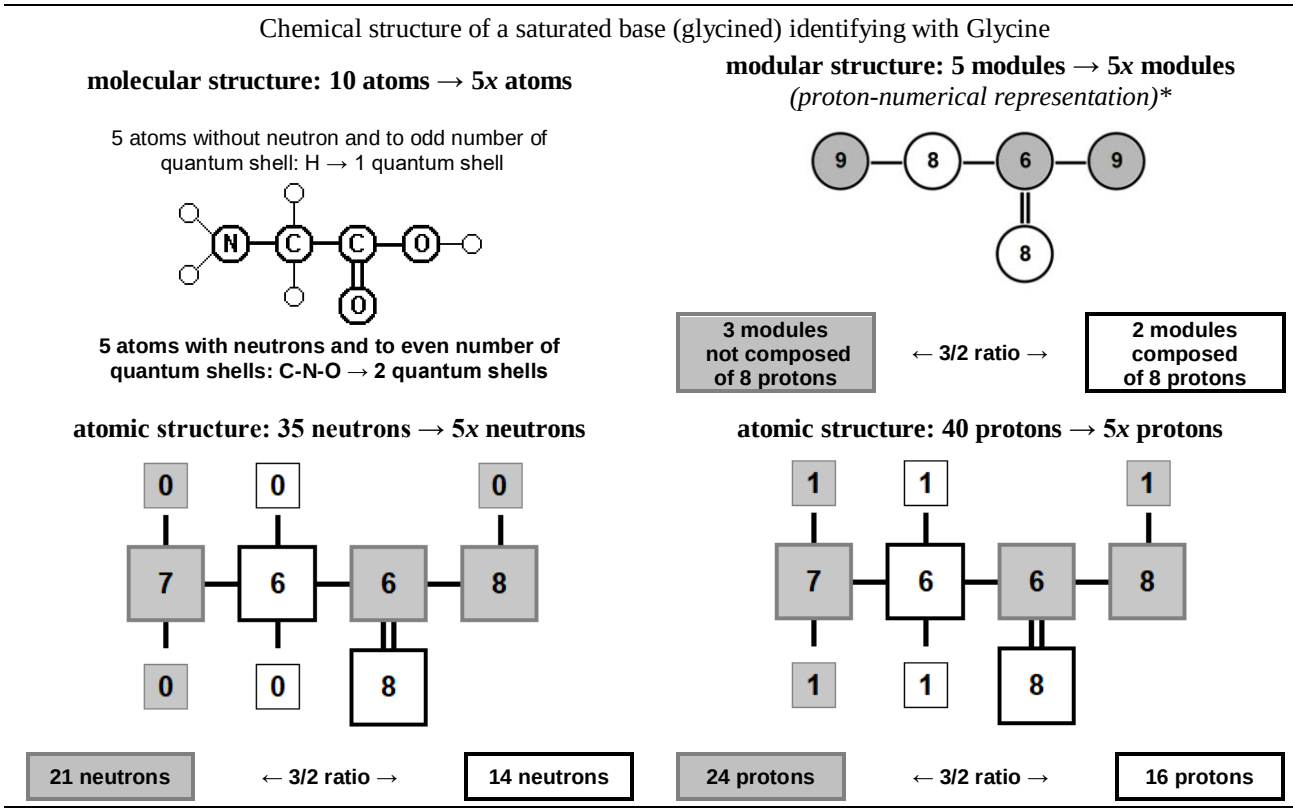


Figure A1: Chemical, modular and atomic structure of a saturated base identified with the amino acid Glycine: 5 modules, 10 atoms, 40 protons and 35 neutrons. See also Figure A2. * inspired representation from Sergey Petoukhov [3 and 9].

Glycine is made up of a multitude of entities whose numbers are all multiples of five. Thus the glycined base consists of five modules, two times five atoms, five of which have one electron shell (H) and five at two shells (C, N and O). Also Glycine consists of 5 times 15 nucleons (75) including 5 times 7 (35) neutrons and 5 times 8 (40) protons. Valences of these different components are also in numbers which are equal to 5*x* entities.

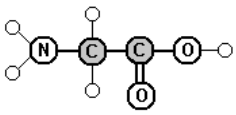
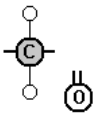
Glycine entities	Total entities number 	Entities account in 3 no 8-proton modules 	Entities account in 2 8-proton modules 
5 modules	**5** **5*x* → *x* = 1**	**3** **3*x* → *x* = 1**	**2** **2*x* → *x* = 1**
10 atoms	**10** **5*x* → *x* = 2**	**6** **3*x* → *x* = 4**	**4** **2*x* → *x* = 4**
5 non-hydrogen atoms (at even number quantum shells)	**5** **5*x* → *x* = 1**	**3** **3*x* → *x* = 1**	**2** **2*x* → *x* = 1**
5 hydrogen atoms (at odd number quantum shells)	**5** **5*x* → *x* = 1**	**3** **3*x* → *x* = 1**	**2** **2*x* → *x* = 1**
75 nucleons	**75** **5*x* → *x* = 15**	**45** **3*x* → *x* = 15**	**30** **2*x* → *x* = 15**
40 protons	**40** **5*x* → *x* = 8**	**24** **3*x* → *x* = 8**	**16** **2*x* → *x* = 8**
35 neutrons	**35** **5*x* → *x* = 7**	**21** **3*x* → *x* = 7**	**14** **2*x* → *x* = 7**
20 valences (cumulated by atom)	**20** **5*x* → *x* = 4**	**12** **3*x* → *x* = 4**	**8** **2*x* → *x* = 4**
15 valences in non-hydrogen atoms	**15** **5*x* → *x* = 3**	**9** **3*x* → *x* = 3**	**6** **2*x* → *x* = 3**
5 valences in hydrogen atoms	**5** **5*x* → *x* = 1**	**3** **3*x* → *x* = 1**	**2** **2*x* → *x* = 1**

Figure A2: Distribution of the prime attributes (to 5*x* in number) of Glycine. Arrangement in 3/2 ratios according to module proton number which can be equal to 8 or not to 8.

Also, it therefore appears, Figures A1 and A2, that the different constituents of Glycine, always 5*x* in number, are always at 3 same *x* entities in the set of three modules (chemical groups) with number of protons not equal to 8 and always of amount at 2 same *x* entities in the set of two modules whose number of protons is equal to 8.

A2. Five living matter atoms

Proteinogenic amino acids (and nucleotides) are just constituted by arrangements of five different atoms. The opposition of the values of Carbon, Nitrogen and Oxygen to those of Hydrogen and Sulphur (Phosphorus for nucleotides in DNA), always generates an arithmetic ratio of value 3/2 according to multiple criteria studied.

A2.1. Quantum anatomy of the five living matter atoms

The table in Figure A3 lists the impressive series of quantum situations in which this remarkable duality takes place between sets of 3*x* entities versus 2*x* entities. Thus, the ratio for the numbers of electron subshells (1*s*, 2*s*, 2*p*, 3*s*, 3*p*) is 3/2. It is still 3/2 if we detail the subshells of those where the quantum number $l = 0$ of those where the quantum number $l = 1$. Also, the ratio for the numbers of orbitals is 3/2. It is still on 3/2 if we detail the orbitals of those where the quantum number $m = 0$, of those where the quantum number $m = -1$ and those where the quantum number $m = 1$.

This ratio is always 3/2 if we detail the orbitals of those where the quantum number $l = 0$ of those where the quantum number $l = 1$. Also, the maximum number of electrons that can orbit inside all of the electronic shells of these two groups of atoms is still in a ratio of 3/2: thirty electrons can orbit inside the electronic shells of Carbon, Nitrogen and Oxygen versus twenty on the electron shells of Hydrogen and Sulphur (Phosphorus for DNA bases). For this last criterion, the distinction of the electrons which can orbit either on the first internal shell (2 electrons for each of the five atoms) or on the set of the other (external) shells always opposes the different values in ratios 3/2: 6 versus 4 electrons for the inner shell and 24 versus 16 for the other shells.

Quantum criteria:	Atoms to **even number** of electron quantum shells				Atoms to **odd number** of electron quantum shells	
Number of atoms	**Carbon** 1	**Nitrogen** 1	**Oxygen** 1		**Hydrogen** 1	**Sulphur*** 1
	3 atoms			**← 3/2 ratio →**	**2 atoms**	
Number of electron shells (*K*, *L*, *M*)	Carbon 2	Nitrogen 2	Oxygen 2		Hydrogen 1	Sulphur* 3
	6 electron shells			**← 3/2 ratio →**	**4 electron shells**	
Number of subshells (1*s*, 2*s*, 2*p*, 3*s*, 3*p*)	Carbon 3	Nitrogen 3	Oxygen 3		Hydrogen 1	Sulphur* 5
	9 subshells			**← 3/2 ratio →**	**6 subshells**	
Number of subshells where the quantum number $l = 0$ where the quantum number $l = 1$	Carbon 2 1	Nytrogen 2 1	Oxygen 2 1	**← 3/2 ratio →** **← 3/2 ratio →**	Hydrogen 1 0	Sulphur* 3 2
	6 subshells where $l = 0$ **3 subshells** where $l = 1$				**4 subshells** where $l = 0$ **2 subshells** where $l = 1$	
Maximum number of orbitals	Carbon 5	Nitrogen 5	Oxygen 5		Hydrogen 1	Sulphur* 9
	15 orbitals			**← 3/2 ratio →**	**10 orbitals**	
Number of orbitals where the quantum number $m = 0$ where the quantum number $m = -1$ where the quantum number $m = 1$	Carbon 3 1 1	Nitrogen 3 1 1	Oxygen 3 1 1		Hydrogen 1 0 0	Sulphur* 5 2 2
	9 orbitals where $m = 0$ **3 orbitals** where $m = -1$ **3 orbitals** where $m = +1$			**← 3/2 ratio →** **← 3/2 ratio →** **← 3/2 ratio →**	**6 orbitals** where $m = 0$ **2 orbitals** where $m = -1$ **2 orbitals** where $m = +1$	
number of orbitals where the quantum number $l = 0$ where the quantum number $l = 1$	Carbon 2 3	Nitrogen 2 3	Oxygen 2 3		Hydrogen 1 0	Sulphur* 3 6
	6 orbitals where $l = 0$ **9 orbitals** where $l = 1$			**← 3/2 ratio →** **← 3/2 ratio →**	**4 orbitals** where $l = 0$ **6 orbitals** where $l = 1$	
Maximum number of electrons orbiting on quantum shells of which the first shell (internal) of which the outer shell (s)	Carbon 10 2 8	Nitrogen 10 2 8	Oxygen 10 2 8		Hydrogen 2 2 -	Sulphur* 18 2 8+8
	30 electrons **6 electrons** **24 electrons**			**← 3/2 ratio →** **← 3/2 ratio →** **← 3/2 ratio →**	**20 electrons** **4 electrons** **16 electrons**	

Fig. A3 3/2 ratio of the electron shells and subshells, orbitals and maximum numbers of electrons according to the parity of the number of electron shells of the five atoms constituting the twenty amino acids (* Or Phosphorus for DNA). Other 3/2 ratios generated in relation to the values of the different quantum numbers of the electrons.

Thus, fourteen different quantum criteria oppose, in a duality of ratio 3/2, the five atoms constituting the twenty amino acids (and also constituting the four DNA nucleotides with the Phosphorus in place of Sulphur). The fact that the genetic code is organized only with these five different atoms in this duality is therefore not random. The perfect complementarity of the quantum characteristics of Hydrogen and Sulphur (Phosphorus in DNA) is particularly remarkable.

These last two atoms have indeed very different quantum characteristics (in contrast to Carbon, Nitrogen and Oxygen with common characteristics) which however complement each other perfectly to always oppose in a 3/2 ratio to three other atoms, constituents of amino acids (and DNA bases). For example, Sulphur has a maximum number of nine orbitals versus only one for Hydrogen. These two very different values nevertheless complement each other (10 orbitals) to oppose in a duality of ratio 3/2 to the three times five quantum orbitals of Carbon, Nitrogen and Oxygen (15 orbitals).

Thus, the 3/2 ratio is revealed at the bottomest of the subatomic structure of the constituents of the twenty amino acids that are on the one hand the three atoms of Carbon, Nitrogen and Oxygen and on the other hand the two atoms of Hydrogen and Sulphur. It is therefore remarkable to note that these same phenomena are found in DNA, another mechanical component of the genetic code, where the quantum properties of the Phosphorus mimic those of Sulphur.

A2.2. Five living matter atoms and remarkable identity

Thus, these various ratios opposing the subshells and shells and transversely, the two categories of atoms previously defined according to the parity of their number of quantum shells, are organized in the remarkable identity $(a + b)^2 = a^2 + 2ab + b^2$ where a and b have the respective values 3 and 2.

Figure A4 explains this arithmetic organization operating in the quantum structure of the five elements working within the genetic code.

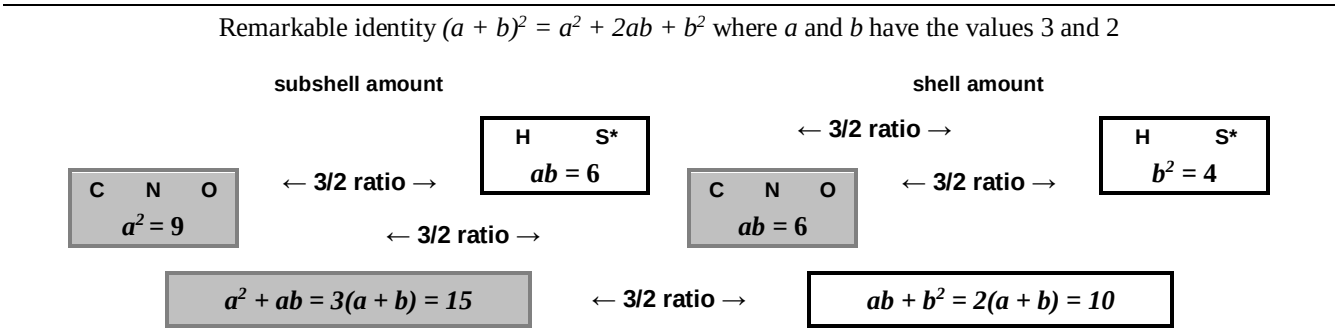


Figure A4: Remarkable identity revealed in the count of subshells and quantum shells of the five elements H, C, N, O and S (*P in DNA). See Figure A3.

Thus, the quantity of subshells in C, N and O corresponds to the value a^2 of the remarkable identity and the quantity of subshells in H and S corresponds to the value ab. The quantity of quantum shells in C, N and O also corresponds to the value ab and that in H and S corresponds to the value b^2. These different values therefore transcend into these equal ratios:

$$(a^2/ab) = (ab/b^2) = (a^2+ab)/(ab+b^2)$$

$$(3^2/6) = (6/2^2) = (3^2+6)/(6+2^2)$$

$$(9/6) = (6/4) = (15)/(10)$$

In a similar fashion, this remarkable identity therefore also operates in the counts of electrons according to their azimuthal quantum number and according to their magnetic number. In these electron counts, the values are just double and, for a and b at the root values 3 and 2, the respective and transcendent values are equal to:

$$2a^2 \rightarrow 2ab \rightarrow 2ab \rightarrow 2b^2$$
$$18 \rightarrow 12 \rightarrow 12 \rightarrow 8$$

A3. Ten first atoms and living matter

It turns out that four of the six organic chemical elements are among the first classified ten elements. Thus, among these 10 (i.e. 5x elements) first elements, in a ratio of value 3/2, six are not organic and four participate in the organization of living matter by being present in the twenty proteinogenic amino acids (and also in nucleotides).

The cumulative value of the atomic numbers, also called nuclear charge numbers, of the first ten elements is mathematically equal to $5x$ with $x = 11$, i.e. a cumulative charge equal to 55.

The first ten chemical elements

6 non-organic chemical elements						← 3/2 ratio →	4 organic chemical elements			
Helium	**Lithium**	**Beryllium**	**Boron**	**Fluorine**	**Neon**		**Hydrogen**	**Carbon**	**Nitrogen**	**Oxygen**
2	**3**	**4**	**5**	**9**	**10**		**1**	**6**	**7**	**8**
33 cumulated atomic number (33 protons)						← 3/2 ratio →	**22 cumulated atomic number** (22 protons)			

Figure A5: Opposition of the 6 non-organic chemical elements and 4 organic chemical elements about their respective cumulated atomic number.

It is found, there again, that the cumulated value of the nuclear charges of the six inorganic elements opposes that cumulated of the four organic chemical elements in a ratio of exact value 3/2. Indeed, as shown in Figure A5, the six inorganic elements total 33 nuclear charges (33 protons) and the four organic elements which are Hydrogen, Carbon, Nitrogen and Oxygen total 22 nuclear charges (22 protons). Since all the other phenomena presented previously, it seems very unlikely that this ratio will appear there also by simple chance.

A4. Alphanumeric symbol of the 20 proteinogenic amino acids

We have therefore firmly established, in many aspects, that the various characteristics of the twenty proteinogenic amino acids are closely linked to their numbering, which itself depends on their DNA codification as proposed at the beginning of the article. This is why we suggest here, to enrich the current nomenclature applied to these twenty entities, the creation of new standardized alphanumeric symbols making it possible to identify these twenty proteinogenic amino acids.

In first preview papers [1 and 2], we initially numbered these entities from 0 to 19 and we attached their respective number to their three-letter alphabetic symbol. For example, we have described Valine as *1Val* and Arginine as *11Arg* so by respectively four and five characters.

For the sake of standardization (and even formalization), we propose, as illustrated Figures A6 and A7, to describe all the twenty AAs with five characters, two of which are numeric and three alphabetical. So we add the number symbol 0 (zero) to the first ten AAs numbered from 0 to 9. By this, for each AA, we therefore propose a unified symbol of 2 digits + 3 letters.

The 20 proteinogenic amino acids conventional nomenclature:			Alphanumeric symbol proposal
Trivial name	*symbol*	*one letter symbol*	
Glycine	Gly	G	**00Gly**
Valine	Val	V	**01Val**
Tryptophan	Trp	W	**02Trp**
Cysteine	Cys	C	**03Cys**
Leucine	Leu	L	**04Leu**
Phenylalanine	Phe	F	**05Phe**
Glutamic acid	Glu	E	**06Glu**
Aspartic acid	Asp	D	**07Asp**
Alanine	Ala	A	**08Ala**
Tyrosine	Tyr	Y	**09Tyr**
Serine	Ser	S	**10Ser**
Arginine	Arg	R	**11Arg**
Methionine	Met	M	**12Met**
Isoleucine	Ile	I	**13Ile**
Lysine	Lys	K	**14Lys**
Asparagine	Asn	N	**15Asn**
Threonine	Thr	T	**16Thr**
Glutamine	Gln	Q	**17Gln**
Histidine	His	H	**18His**
Proline	Pro	P	**19Pro**

Figure A6: Conventional nomenclature and alphanumeric symbol proposal to the twenty proteinogenic amino acids into 5 characters: 2 digits + 3 letters.

The table in Figure A6 therefore lists all of the 20 proteinogenic amino acids involved in the mechanism of the universal genetic code. It is therefore described, from the conventional nomenclature, the trivial name, the symbol in 3 letters and the one letter symbol. To this is added, for each AA, its alphanumeric symbol of 5 characters that we propose as a new standardized and official nomenclature.

Conventional nomenclature			Proposed alphanumeric symbol **into 5 characters**
Trivial name	➞	***Glycine***	
Symbol (three letters)	➞	***Gly***	➞ ***00Gly***
One letter symbol	➞	***G***	

Figure A7: Conventional nomenclature and alphanumeric symbol proposal to proteinogenic amino acids into 5 characters: 2 digits + 3 letters. Here Glycine as example. See Figure A6 also.

A5. Other singular configurations

Figure A8 suggests a recombination of the arrangements presented in Figure 41 Chapter 9. In these other symmetrical areas of eight boxes, there are also always 10 AAs and these still oppose each other in a value ratio of 3/2 depending on their number of atoms.

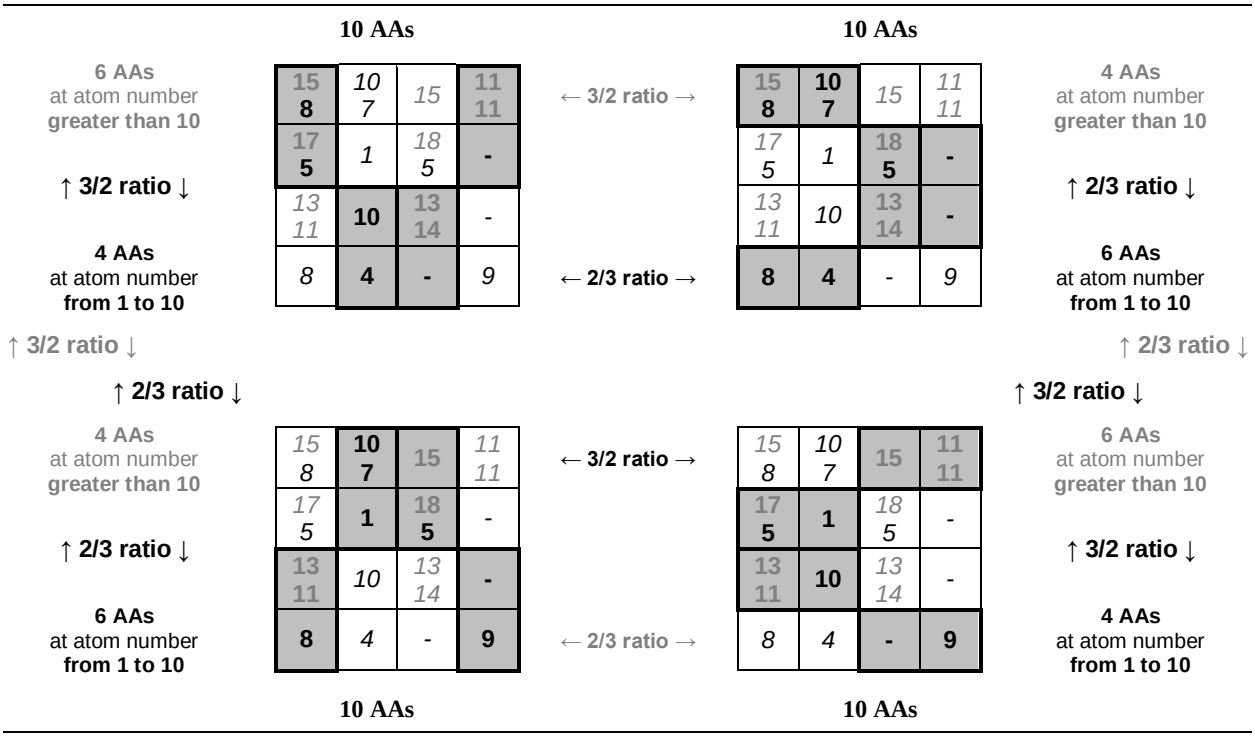


Figure A8: According to their number of atoms (radical only) distribution of amino acids in other symmetrical areas of the matrix of the ultimate genetic code always in an opposition of 3*x* versus 2*x* entities. See Figure 41 to comparison.

A6. Requests about scientific circles

If you are convinced of the relevance of creating a new alphanumeric nomenclature to the twenty proteinogenic amino acids, thank you, according to your influential reputation, for helping the author by working in this way with qualified world institutions.

Jean-Yves BOULAY independent researcher – FRANCE – jean-yvesboulay@orange.fr ORCID: 0000-0001-5636-2375
https://www.researchgate.net/profile/Jean-Yves-Boulay