Numbering of the twenty proteinogenic amino acids
3/2 ratios inside the genetic code

Jean-Yves BOULAY

Abstract

By proposing a numbering of the twenty proteinogenic amino acids deduced from the physicochemical properties of the four coding DNA nucleobases, it is established that this amino acid number, equal to 5*x* entities, is not arbitrary. Indeed, we demonstrate that many attributes of these twenty amino acids, as a whole, are also 5*x* in number and that by isolating, since their numbering, the 3*x* peripheral amino acids from the 2*x* internal ones, these attributes are divided into ratios of 3/2 as exact value. This is verified both as the physicochemical properties of the 20 amino acids and as the coding configurations of the nucleobases, the source of this numbering.

1. Introduction

Today, it is now firmly established that living matter is organized via a so-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.

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 [1], 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. By "genetic code", we consider in this paper the totality of its components, namely simultaneously the 64 codons and the twenty encoded amino acids.

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. By doing this, then it turns out that the different attributes respective to these two groups are always also in a ratio whose value is very precisely equal to 3/2. This, both for different and numerous attributes (physico-chemical properties: number of atoms, CH_2 groups, etc.) of the twenty proteinogenic amino acids and in some aspects for the characteristics of their respective codons.

Finally, according to many scales of their physico-chemical properties (hydrophobicity, propensity to promote a α-helix, etc.) the twenty amino acids are further divided into 3/2 ratios according to their numbering. Numbering which, we remind you with insistence, is very dependent on the physico-chemical properties of the nucleobases which encode them.

2. Numbering of the twenty proteinogenic amino acids

In order to be able to number the twenty proteinogenic amino acids, we must first proceed to a numbering of the 64 codons of the universal genetic code. Also, this numbering of amino acids must depend on the physico-chemical character of the nucleobases constituting the codons. To this end, we use the very original numbering devised by Professor Sergey Petoukhov, which is based on the possible deamination and depurination of the four nucleobases.

2.1 Petoukhov’s numbering of the 64 genetic code codons

In his investigations of the genetic code [1] 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 Sergei Petoukhov [1]). Depending on whether each nucleobase can undergo deamination or not, Sergey Petoukhov assigns them either the value 1 or the value 0 (table Figure 1). Also, depending on whether each nucleobase can undergo depurination or not, Sergey Petoukhov assigns them either the value 0 or the value 1.

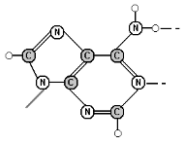
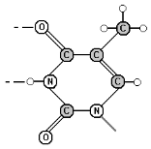
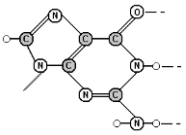
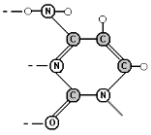
	Adenine	Thymine	Guanine	Cytosine
nucleobases				
Possible deamination: yes = 1 no = 0	**1**	**0**	**0**	**1**
Possible depurination: yes = 0 no = 1	**0**	**1**	**0**	**1**

Fig. 1 Method of assigning a double binary value to the four DNA nucleobases according to Sergey Petoukhov [1].

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 2.

physico-chemical criteria →	**possible deamination** *yes = 1 no = 0*			**possible depurination** *yes = 0 no = 1*		
codon →	**A**	**T**	**G**	**A**	**T**	**G**
binary convert →	***1***	***0***	***0***	***0***	***1***	***0***

ATG
Met
34
100010

ATG = *100010* = 34

Fig. 2 Method of assigning a number to codons according to Sergey Petoukhov. See Fig. 1 and 3 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 3 illustrates this complet numbering of the 64 genetic code codons set.

	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**

Fig. 3 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.

It is important to emphasize that, in this innovative codon arrangement, DNA triplets located on the same line all have their three bases (even though they are unique) classified in the same order according to the criterion of possible deamination and that those located on the same column all have their three bases (yet unique) classified in the same order according to the criterion of possible depurination.

2.2 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 3).

By this process, it is thus assigned (Figure 4) 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.

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**

Fig. 4 Assigning a single only one number to each of 20 proteinogenic amino acids in the table of the complete genetic code. See Fig. 3 also.

2.2.1 Symmetrical break-up of the 20 AAs in 3/2 ratio

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 a first set of 12 entities then a second set of 8 other entities. As illustrated in Figure 5, these two sets then oppose each other in a ratio of value 3/2.

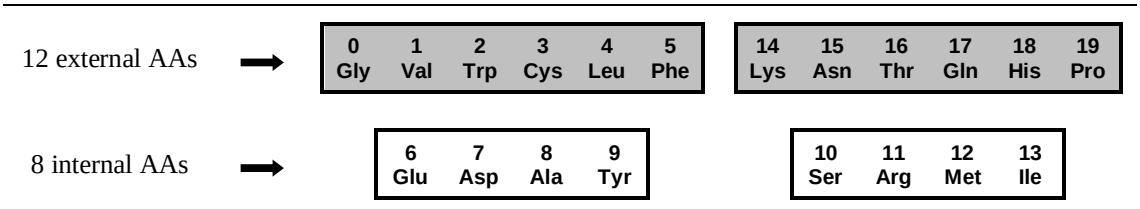


Fig. 5 Since them numbering, symmetrical break-up of the 20 AAs into two sets of 2 times 6 versus 2 times 4 entities.

2.2.2 Conventional representation of amino acids numbering.

Using symmetry graphics, many arithmetic phenomena presented in this paper will be presented in the way illustrated in Figure 6. 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*: 0Gly versus 19Pro, 1Val versus 18His, etc.

Also, we therefore isolate two numbering zones:

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

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

This is only by this way that appears many singular arithmetic arrangements about amino acids attributes. Also, in some demonstrations, we can likewise speak of grey area (grey) and light area (white) to define these two sets respectively made up of 3*x* and 2*x* entities.

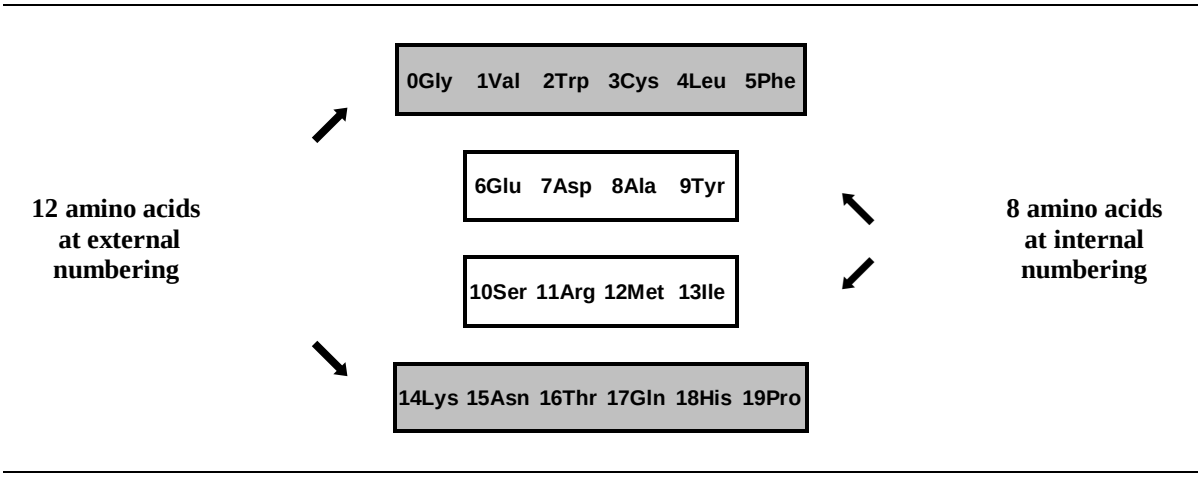


Fig. 6 Conventional representation of 20 proteinogenic amino acids numbering in symmetry graphics.

3. Depiction of the twenty proteinogenic amino acids

3.1 Schematic graphics table

In this paper, the different atoms, five in number, which make up the 20 proteinogenic amino acids, will be represented, in an unconventional way, as in the table in Figure 7.

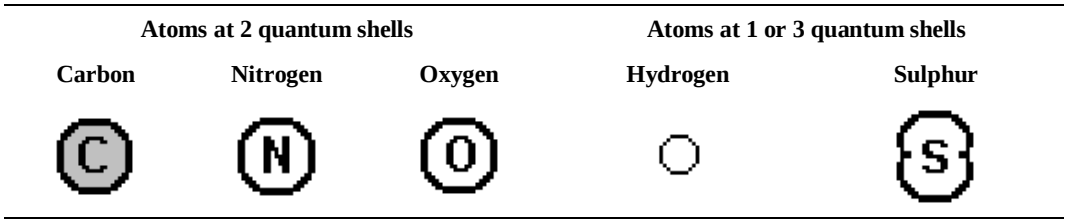


Fig. 7 Unconventional representation of the 5 atoms making up the 20 AAs. See Fig. 11 also.

This is inspired by proton-numerical module concept introduced by Sergey Petoukhov [1 and 2]. 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 Sulfur as constituted in a twice module. We will briefly discuss this concept of module in Chapter A.1 (appendix) about Glycine.

Not by chance, there are precisely five atoms that make up the 20 AAs. Also, as we will demonstrate in the next chapter, these five entities oppose each other in a ratio of 3/2 value.

The following table (Figure 8) lists the 20 amino acids in the order of their numbering just introduced in Chapter 2. 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 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,
- number of atoms,
- sort of atom,
- number of CH_2 groups
- radical symmetry (or not symmetry)

Other physical properties studied will be introduced in various tables throughout the chapters, such as hydrophobicity scales for example.

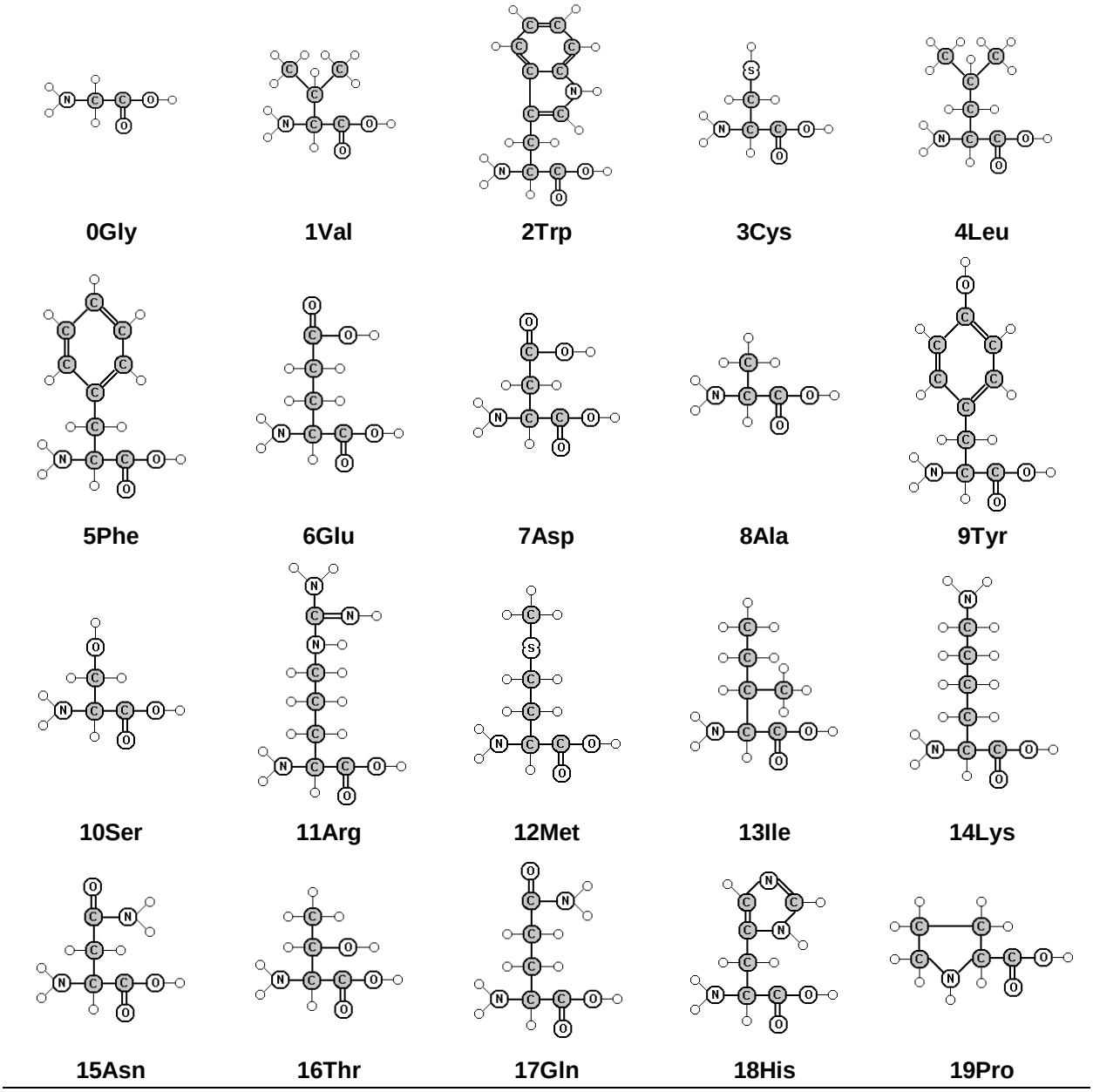


Fig. 8 Molecular description of the 20 proteinogenic AAs in the order of their numbering (in an unconventional graphic design which is inspired by S. Petoukhov papers [1-2]).

3.2 Some prime amino acid attributes.

The table in Figure 9 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. The criteria introduced here will be more widely illustrated throughout the various chapters presented.

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.

Other various aspects relating to amino acids and the genetic code will complete and flesh out this first approach.

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 external values (from 0 to 5 and from 14 to 19)		**114**	**123**	**51**	**72**	**15**	**126**	**33**
8 internal values (from 6 to 13)		**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

Fig. 9 Some prime amino acid attributes. See Figure 8 also.*Radical only and in non-ionized state.** Exact OMH index values are listed table Figure 22, Chapter 5.

4. Organization of the 20 amino acids attributes into 3/2 ratios

4.1 Atom number

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 10, there are 123 atoms ($3x$ atoms → $x = 41$) in the 12 external AAs set as defined in Chapter 2.2 and there are 82 atoms ($2x$ 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.

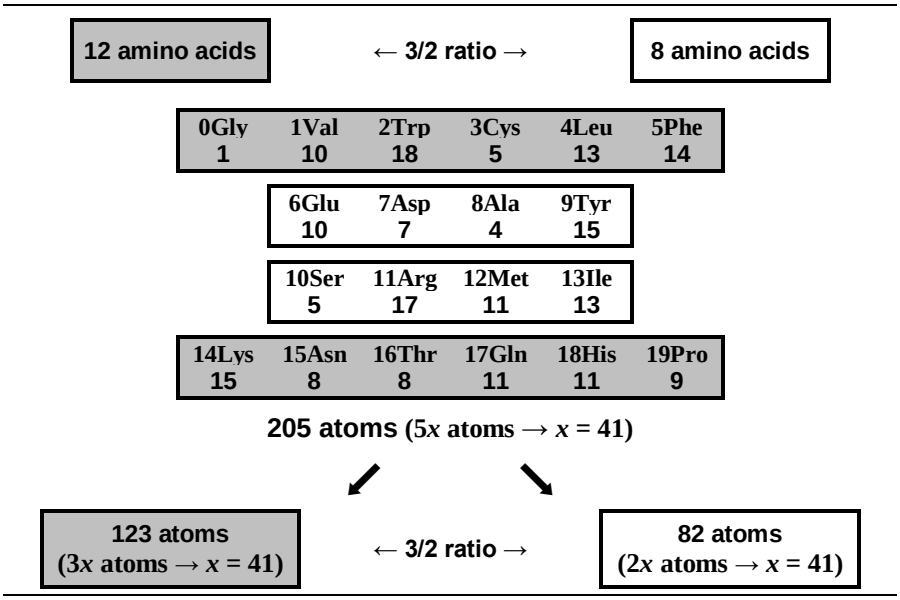


Fig. 10 Atom counting inside radical of each proteinogenic amino acid. Distribution in 3/2 ratio regarding AA numbering areas.

4.2 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 11 illustrates this prime opposition of genetic code constituents in 3/2 values ratio.

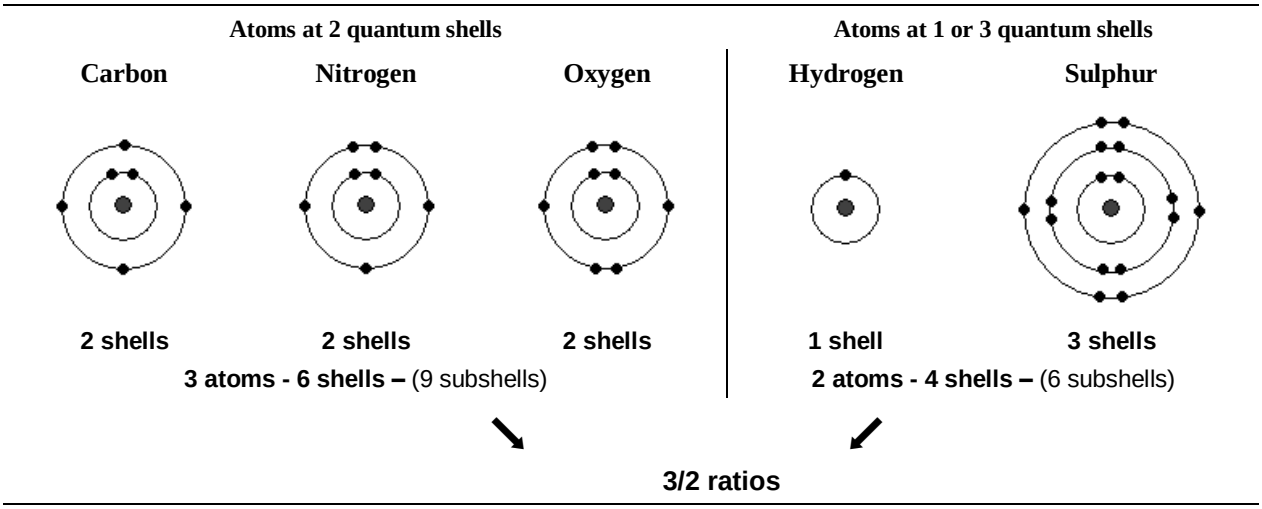


Fig. 11 Differentiation of the 5 atoms constituting the 20 proteinogenic amino acids into 2 sets of 3 and 2 atoms according to the parity of their number of electron shells.

Figure 11 illustrates just a few oppositions between these two sets of three and two atoms. In a previous paper [9] 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 A3, 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.

The separate counting of these two categories of atoms, which grouped together gave (Figure 12) already an opposition of the values in a 3/2 ratio according to the numbering of the 20 amino acids, continues to generate the same exact ratios of 3/2 values.

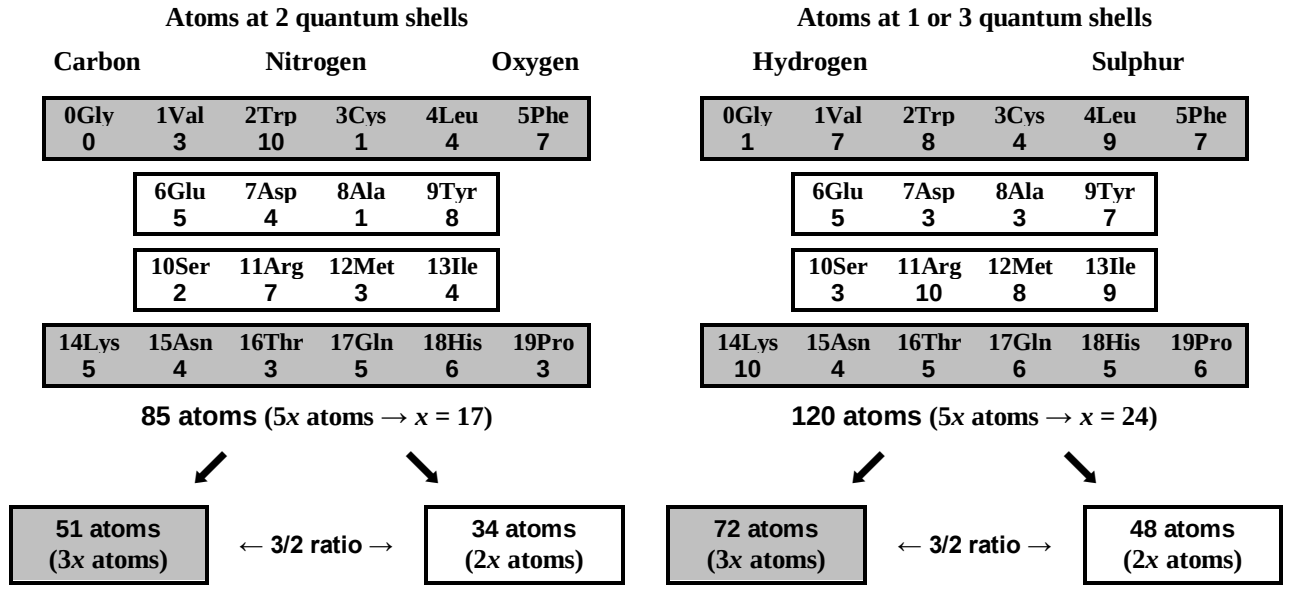


Fig. 12 Atom counting inside radical of each proteinogenic amino acid. Distribution in 3/2 ratio regarding AAs numbering and number parity of electron shells.

4.2.2 Atom number gap

Since each of the two distributions of atoms illustrated in Figure 12 generates an opposition of the values in ratio 3/2, in relative value, it is arithmetically logical that the differences in the numbers of atoms with an even and odd number of shells of electrons also generates the same ratios.

However, it is remarkable to note that these ratios, of exact value 3/2, are maintained by considering the differences of atoms in absolute values as it appears in Figure 13.

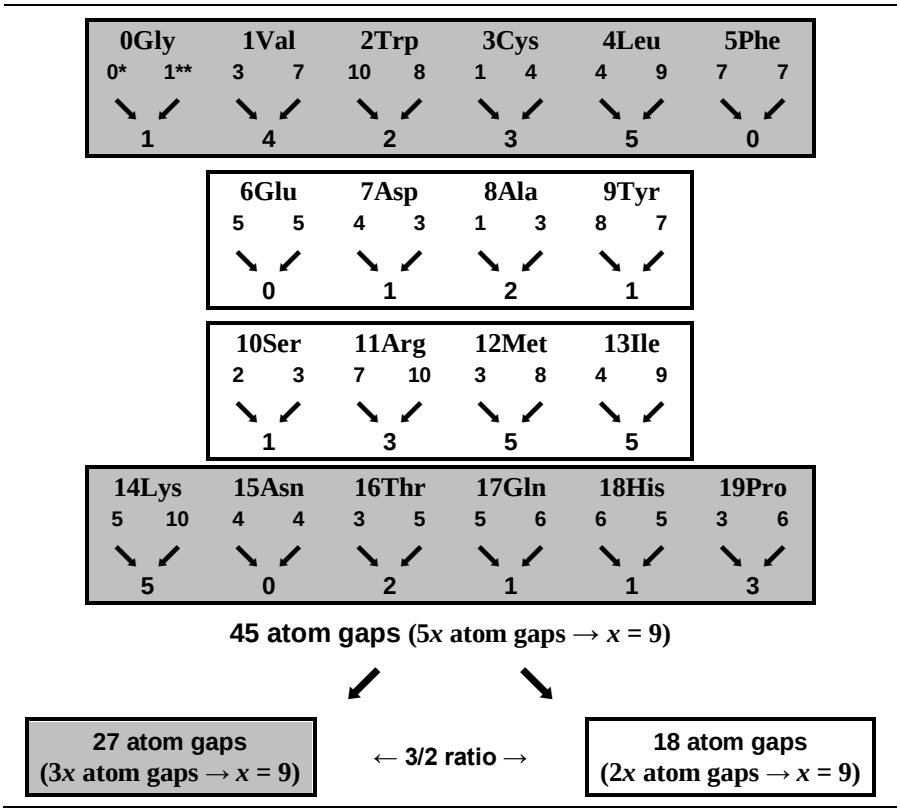


Fig. 13 In absolute values: atom gap counting regarding number parity of electron shells. * C,N and O. ** H and S. See Fig. 12 also.

4.3 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 14 illustrates these two configuration types.

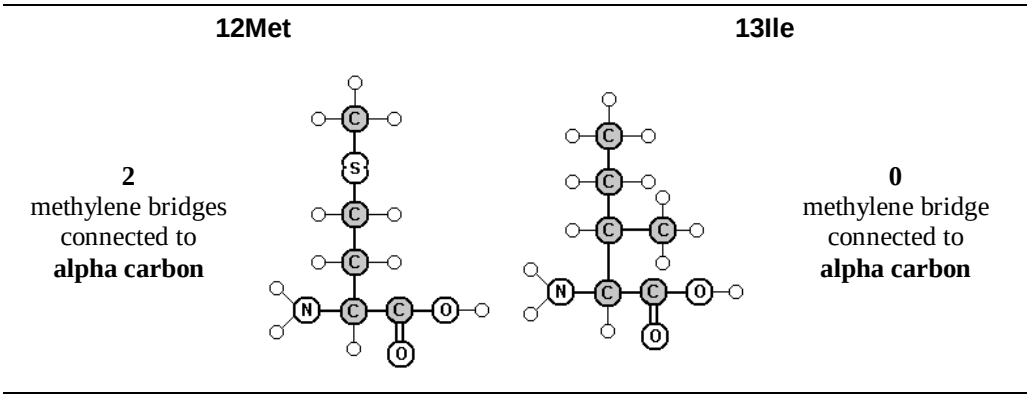


Fig. 14 CH_2 groups (methylene bridges) differentiation.

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.

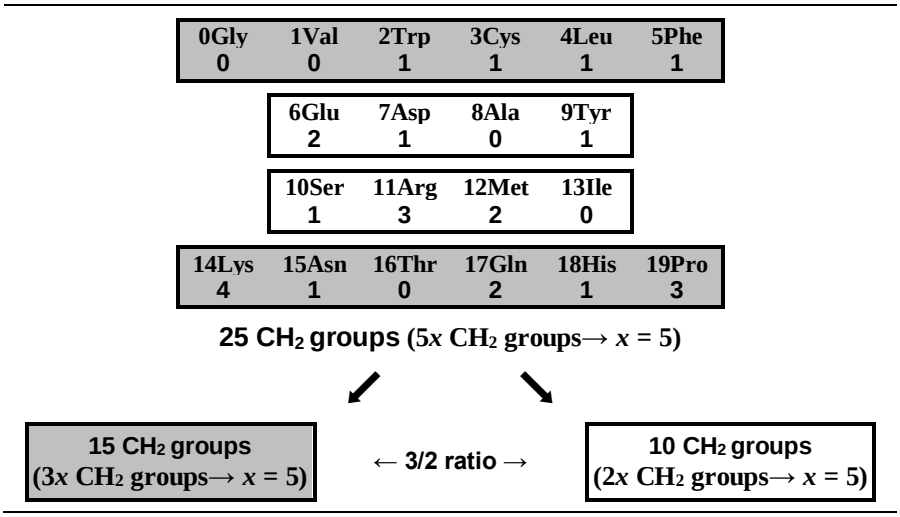


Fig. 15 CH_2 groups counting. Only methylene bridges directly connected with alpha carbon.

As it appears in Figure 15, it turns out that these 25 CH_2 groups are distributed in an exact ratio of value 3/2 between the set of 12 external AAs and the set of 8 internal AAs with respectively 15 ($3x \rightarrow \boldsymbol{x} = 5$) and 10 ($2x \rightarrow \boldsymbol{x} = 5$) entities listed.

*In fact, 15 AAs (so 5*x* AAs) contains CH_2 groups. This aspect will developed Chapter 9 with connections to amino acid coding.

4.3.1 CH_2 groups gaps and numbering

Depending on their numbering, the twenty amino acids therefore oppose their quantity of methylene bridges in a 3/2 ratio according to their membership of the external and internal groups defined in Chapter 2. This numerical differentiation is not the only one to generate this type of arithmetic phenomenon.

The difference in the number of CH_2 groups between two amino acids of opposite numbering (0 versus 19, 1 versus 18, etc.) also overall generates an opposition of values in an exact 3/2 ratio. This, once more between the two sets of external and internal entities as shown in Figure 16.

This therefore greatly reinforces the numbering proposal introduced and studied in this paper relating to the genetic code and its physico-chemical attributes (AAs, nucleobases, etc.).

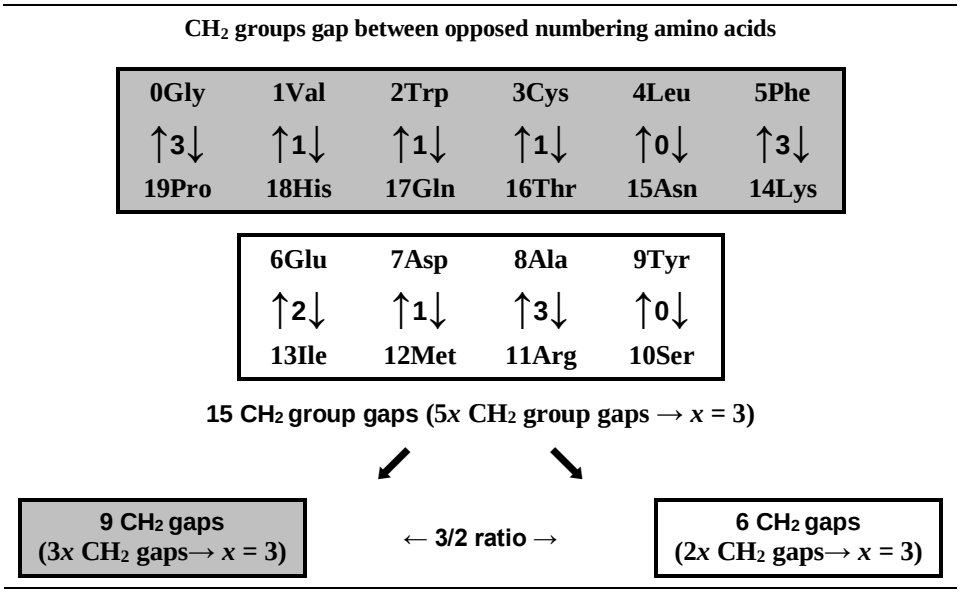


Fig. 16 CH_2 groups gaps between opposed numbering amino acids. See Fig 15 also.

Because we consider this concept as primordial in amino acid structures, others interesting arithmetic phenomena about these CH_2 groups will developed in some next chapters.

4.4 OMH index rank

According to the exact values of the OMH scale shown in Figure 22 in Chapter 5, we created (*f* in Figure 9) an index rank scale ranging from 1 (largest index) to 20 (lowest index) for the twenty amino acids.

4.4.1 Full OMH index rank

The cumulative value of these ranks gives a value of 126 for the external set of AAs and a value of 84 for the internal one as this is illustrated Figure 17.

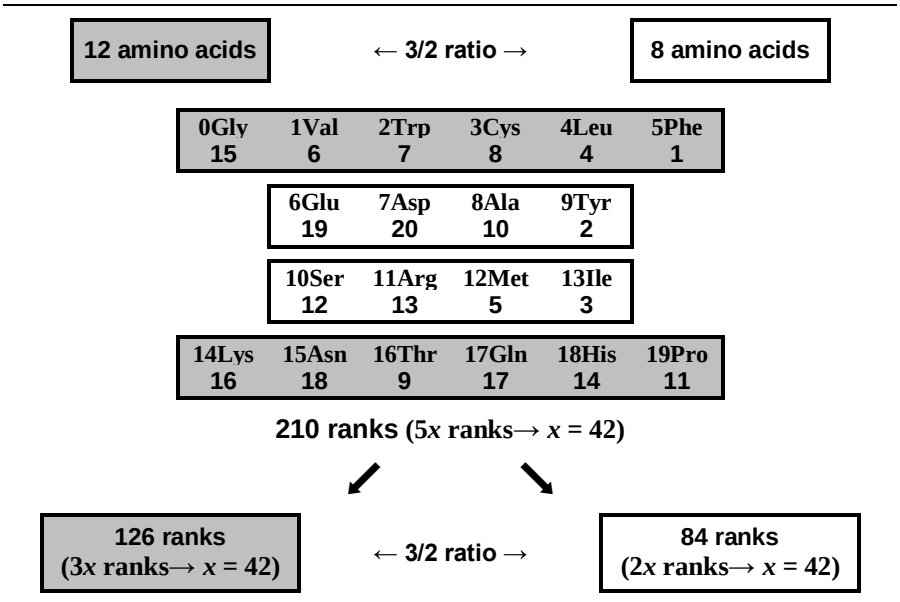


Fig. 17 OMH index ranks distribution in exact 3/2 ratio into two external and internal sets of AAs.

The OMH index [3] 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.

4.4.2 OMH index ranks parity

Although the distribution of the different OMH index ranks (Fig. 17 and 18) seems random within the two defined AAs sets of external and internal, the even and odd isolated values continue to generate a perfect 3/2 ratio between these two sets.

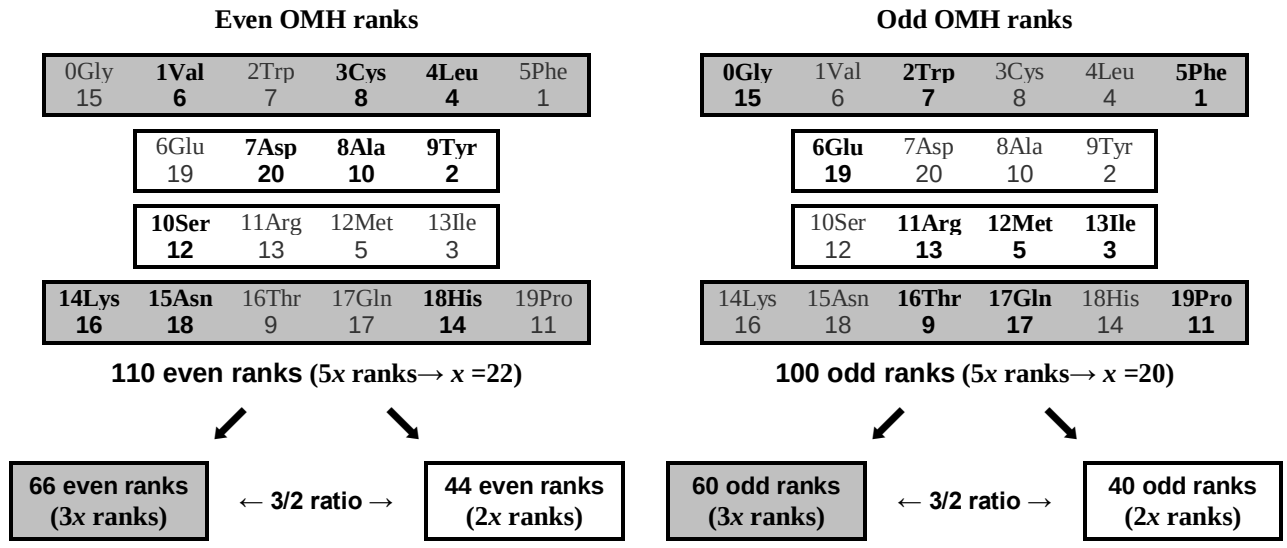


Fig. 18 According of the rank parities: OMH index ranks distribution in exact 3/2 ratio into two external and internal sets of AAs.

4.5 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, codons 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

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

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

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

Fig. 19 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, Figure 19, 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 generates here also an exact ratio of value 3/2 with 33 versus 22 counted codons as this is illustrated Figure 20.

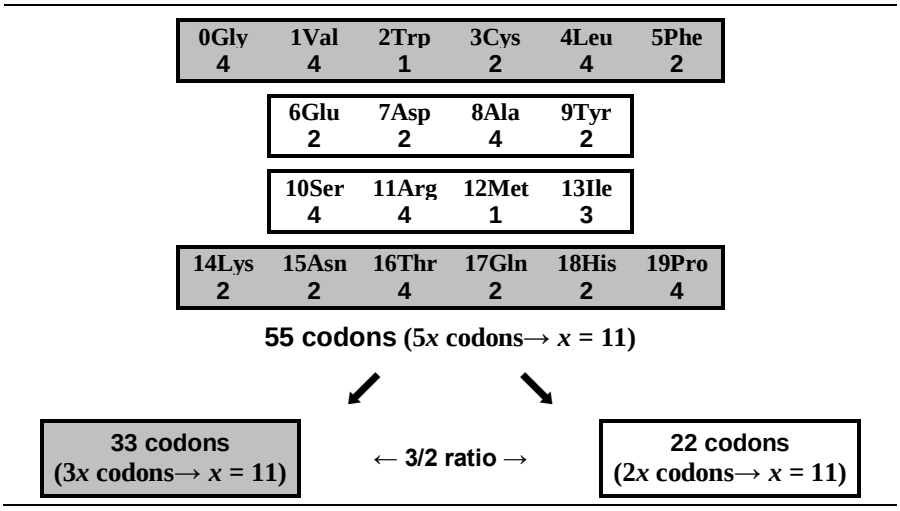


Fig. 20 Codon number in largest codon sets with 2 first same DNA bases.

4.6 First synthesis of the distribution of the AAs prime attributes

Figure 21 summarizes the main attributes of the twenty amino acids. 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 and into 2*x* entities in that of 8 AAs with internal numbering.

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** $2x \rightarrow x = 11$

Fig. 21 Depending on their numbering (external or internal): 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 investigations will reinforce this hypothesis.

5 Some prime hydrophobicity and other property scales

Here will be investigated some of the main scales of hydrophobicity and other properties of the twenty proteinogenic amino acids. It is therefore relative values that are studied, unlike the previous ones which were absolute data. In relation to this aspect, for each situation, the ten AAs with the highest indices or values will be opposed to the 10 others with the lowest level.

5.1 Scales

Table in Figure 22 lists some prime hydrophobicity and other property scales about studied amino acids. Here, investigations are on OMH scale, Eisenberg W. hydrophobicity scale, Van der Walls volume scale and three Garnier-Osguthorpe-Robson scales about of each AA its propensity to promote a structure in α-helix, β-sheet and β-bend.

Amino acid	OMH scale	Eisenberg W. hydrophobicity scale	Van der Walls volume (A° 3)	Garnier-Osguthorpe-Robson scales Propensity to promote a structure in:		
				α-helix	β-sheet	β-bend
0Gly	-0.67	0.16	48	0.56	0.92	1.64
1Val	0.91	0.54	105	0.91	1.49	0.47
2Trp	0.5	0.37	163	0.99	1.14	0.75
3Cys	0.17	0.04	86	1.11	0.74	0.80
4Leu	1.22	0.53	124	1.30	1.02	0.59
5Phe	1.92	0.61	135	1.07	1.32	0.58
6Glu	-1.22	-0.62	109	1.44	0.75	1.00
7Asp	-1.31	-0.72	91	1.04	0.72	1.41
8Ala	-0.4	0.25	67	1.29	0.90	0.78
9Tyr	1.67	0.02	141	0.72	1.25	1.05
10Ser	-0.55	-0.26	73	0.82	0.95	1.33
11Arg	-0.59	-1.80	148	0.96	0.99	0.88
12Met	1.02	0.26	124	1.47	0.97	0.39
13Ile	1.25	0.73	124	0.97	1.45	0.51
14Lys	-0.67	-1.10	135	1.23	0.77	0.96
15Asn	-0.92	-0.64	96	0.90	0.76	1.28
16Thr	-0.28	-0.18	93	0.82	1.21	1.03
17Gln	-0.91	-0.69	114	1.27	0.80	0.97
18His	-0.64	-0.40	118	1.22	1.08	0.69
19Pro	-0.49	-0.07	90	0.52	0.64	1.91

Fig. 22 Some prime hydrophobicity and other AA property scales.

5.2 Ten first and ten last amino acids

Depending on the data listed in the table of Figure 22, for each criterion, a distinction is made between the ten AAs with the highest values and 10 others with the lowest. This is highlighted in the table in Figure 23.

For all these scales, although concerning very different criteria, it is remarkable to note that in each of the two sets of 12 external and 8 internal numbered amino acids, there is always half of the entities with highest values and another half to lowest.

Thus, in each situation, among the ten AAs with highest values, six are from the external zone and four from the internal numbering zone. The ten AAs with lowest values are therefore also divided into this perfect 3/2 ratio.

AA	*a*	*a'*	*b*	*b'*	*c*	*c'*	*d*	*d'*	*e*	*e'*	*f*	*f'*
0Gly		x	x			x		x		x	x	
1Val	x		x			x		x	x			x
2Trp	x		x		x			x	x			x
3Cys	x		x			x	x			x		x
4Leu	x		x		x		x		x			x
5Phe	x		x		x		x		x			x
6Glu		x		x		x	x			x	x	
7Asp		x		x		x	x			x	x	
8Ala	x		x			x	x			x		x
9Tyr	x		x		x			x	x		x	
10Ser		x		x		x		x		x	x	
11Arg		x		x	x			x	x			x
12Met	x		x		x		x		x			x
13Ile	x		x		x			x	x			x
14Lys		x		x	x		x			x	x	
15Asn		x		x		x		x		x	x	
16Thr	x			x		x		x	x		x	
17Gln		x		x	x		x			x	x	
18His		x		x	x		x		x			x
19Pro		x		x		x		x		x	x	
12 external AAs	6	6	6	6	6	6	6	6	6	6	6	6
8 internal AAs	4	4	4	4	4	4	4	4	4	4	4	4
ratio →	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2

a 10 amino acids with highest **OMH index**
a' 10 amino acids with lowest **OMH index**
b 10 amino acids with highest **Eisenberg Weiss hydrophobicity index**
b' 10 amino acids with lowest **Eisenberg Weiss hydrophobicity index**
c 10 amino acids with largest **van der Waals radius**
c' 10 amino acids with smallest **van der Waals radius**
d 10 amino acids with a higher propensity to promote an **α-helix structure**
d' 10 amino acids with a lower propensity to promote an **α-helix structure**
e 10 amino acids with a higher propensity to promote a **β-sheet structure**
e' 10 amino acids with a lower propensity to promote a **β-sheet structure**
f 10 amino acids with a higher propensity to promote a **β-bend structure**
f' 10 amino acids with a lower propensity to promote a **β-bend structure**

Fig. 23 According to different scale data (listed Fig. 22), distribution of the ten highest and ten lowest values in perfect 3/2 ratio in the two predefined external and internal numbering zones.

5.3 OMH scale

Chapter 4.4, we have already extensively presented singular phenomena concerning the ranks of the OMH index of amino acids. It can be seen that, in addition to this, as illustrated in the table Figure 22 (*a* and *a'*), the group of 10 AAs with the highest rank like that of the group of 10 with the lowest rank are distributed in the ratio 3/2 in the two numbering zones.

5.4 Eisenberg Weiss scale and numbering concept

The Eisenberg Weiss scale, that we have taken as a reference in this paper, and which as shown in Figure 23, generates a distribution of the twenty amino acids in exact ratio 3/2 according to their index intensity and their numbering, is of good reputation in genetic code study. Therefore what will now be presented has good credibility and may be of great consideration.

Indeed, as it appears in Figure 24, the values of this scale are in close communion with the AAs numbering concept proposed in this paper. Thus, the first six externally numbered AAs all have a high Eisenberg Weiss index while the last six all have a low index. Oppositely but also alternately 2 by 2, we observe the same phenomenon concerning the set of eight AAs with internal numbering.

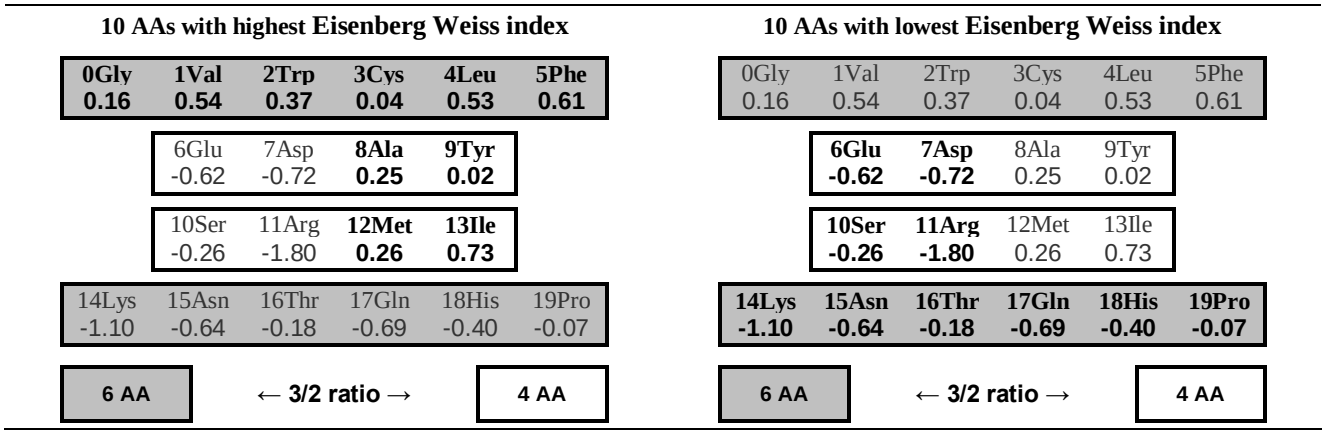


Fig. 24 Eisenberg Weiss index specific distribution in connection with AAs numbering.

In fact, depending on whether one considers pairs of amino acids in opposition of numbering or in joint numbering, the index values of the Weiss scale are distributed in a very structured way.

5.4.1 Pairs of opposite numbered AAs

As it is illustrated Figure 25, very oddly, absolutely all of the ten pairs of oppositely numbered amino acids are made up of one high-index entity and one low-Eisenberg index.

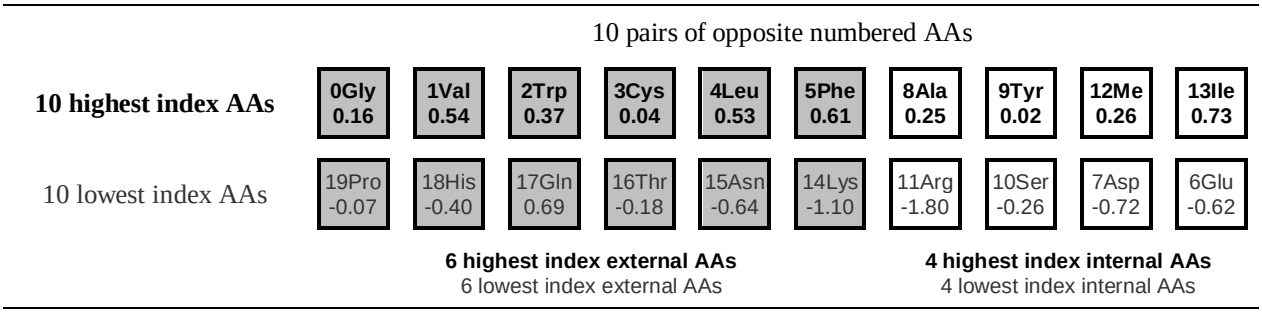


Fig. 25 Eisenberg Weiss index specific distribution in connection with opposite AAs numbering.

5.4.2 Pairs of consecutive numbered AAs

In parallel to this and as it is illustrated Figure 26, again about that same Eisenberg scale, very oddly, absolutely all of the ten pairs of consecutive numbered amino acids are make up of either two high index entities or two low index entities.

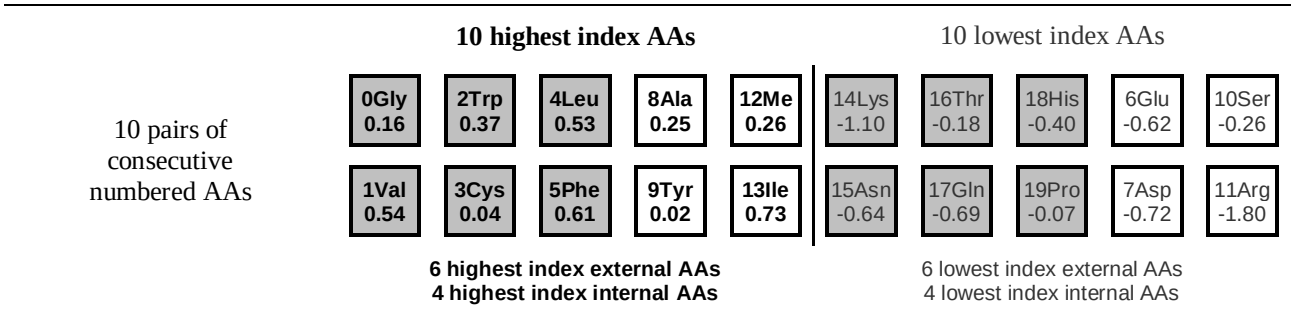


Fig. 26 Eisenberg Weiss index specific distribution in connection with consecutive AAs numbering.

It turns out, as will be demonstrated in Chapters 6 and 7, that these singular arrangements also operates in exact same way considering several other attribute scales of the twenty amino acids.

5.5 Van der Walls volume

As it is illustrated Figure 27, the 10 AAs with the largest Van der Waals volume indices and the 10 AAs with the smallest indices are distributed in 3/2 ratios in accordance with the two external and internal numbering zones.

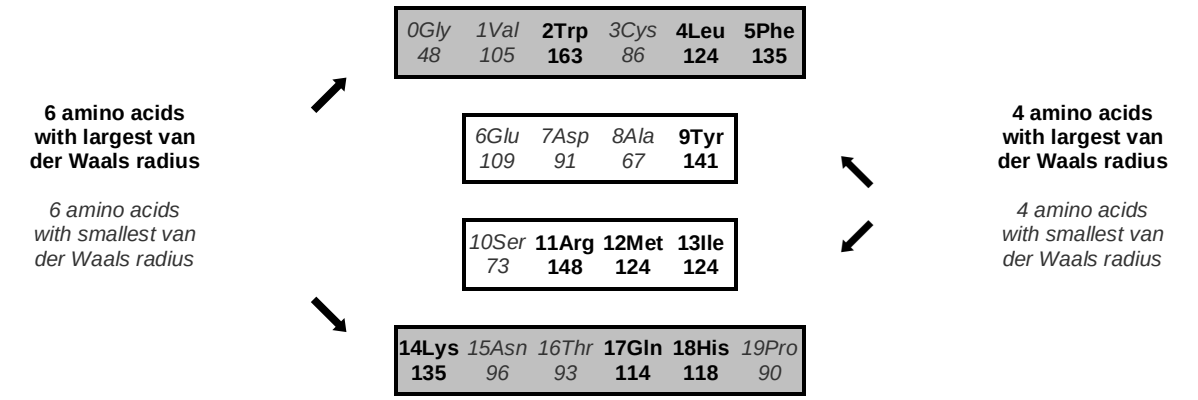


Fig. 27 Van der Waals radius index distribution in 3/2 ratio according to AAs internal and external numbering.

5.5.1 Van der Walls volume and atom number

About Van der Waals volume, the cumulative value (see Figures 22 and 27) of the radii of the 12 amino acids in the outer numbering zone is 1307 and that of the 8 AAs in the inner zone is 877. The ratio between these two combinations is very close to 3/2 since it is equal to 1.49. This inaccuracy is simply explained by the nature of the values of this scale (data in Figure 22) which are not absolute (unlike for example the number of atoms as illustrated in Chapter 4).

Also, as can be seen in Figure 28, the ten amino acids with the greatest number of atoms and the ten with the smallest number are distributed in the ratio 3/2 in the two numbering zones. It further turns out that these two groups of amino acids are the same as larger and smaller Van der Waals volume ones.

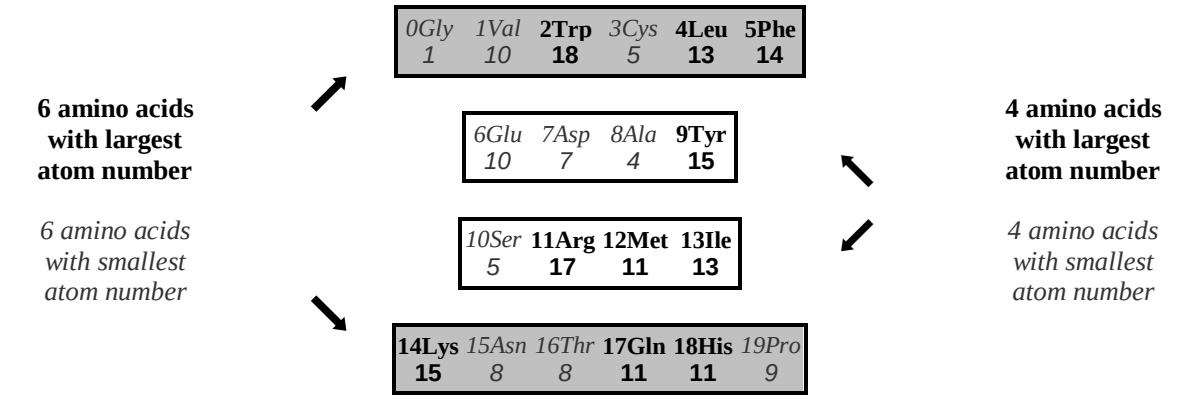


Fig. 28 Largest and smallest atom number AAs distribution in 3/2 ratio according to internal and external numbering. See Fig. 27 also.

5.6 AAs propensity to promote α-helix, β-sheet and β-bend

To the right side of table in Figure 22 groups together the probabilities for each of the twenty amino acids of being found in one of the three main structural units within a protein, i.e. within an α helix, a β sheet and a bend β.

Values below 1 indicate a tendency not to promote the type of structure (α-helix, β-sheet or β-bend) while those above 1 indicate a tendency to promote this type of structure. These values are taken from the GOR (Garnier-Osguthorpe-Robson) scales [6].

Illustrated Figure 23 and in more details Figure 29, it turns out that for each of these triple scales (α-helix, β-sheet and β-bend), by isolating the twenty amino acids into two groups of ten, one of which with the strongest indices (therefore the strongest tendencies to promote the structure) and the other with the lowest indices, the amino acids are distributed in the 3/2 ratio according the two defined numbering zones.

Thus, studying here the secondary structure of proteins and despite the different types of structure studied (α-helix, β-sheet and β-bend), in each situation, two groups of ten amino acids are distinguished in the ratio 3/2 in the two defined numbering zones, although these different pairs of two groups of ten amino acids are never completely made up of the same ten entities.

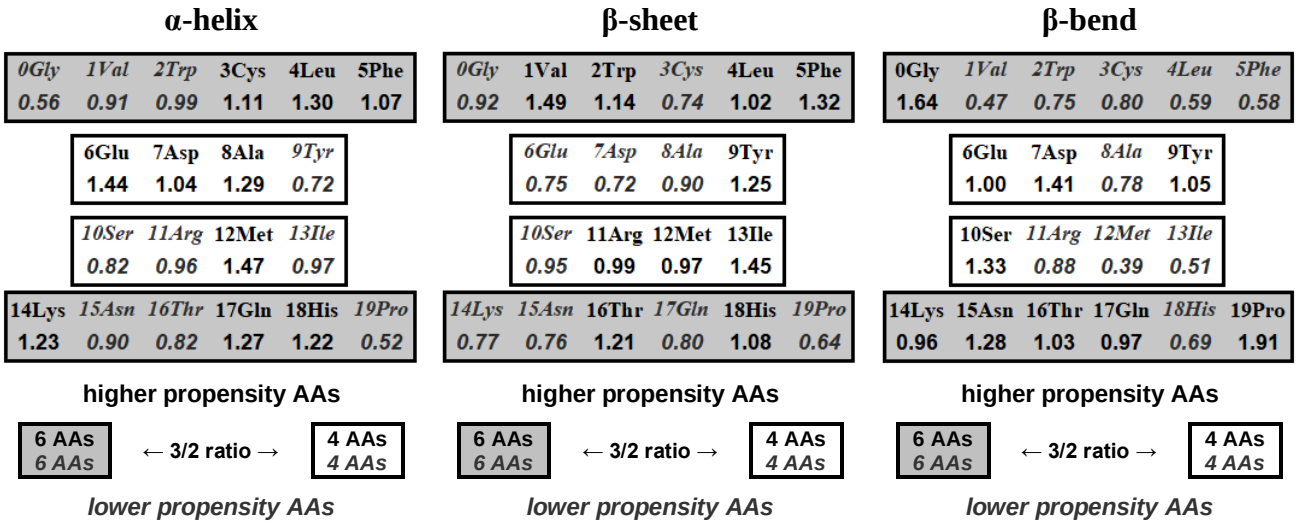


Fig. 29 Garnier-Osguthorpe-Robson scales index distribution in 3/2 ratio according to AAs internal and external numbering. See Fig. 22 and 23 also.

6 Others hydrophobicity scales

As we are going to demonstrate now, according to a very large number of hydrophobicity scales, the twenty proteinogenic amino acids are always organized in the 3/2 ratio according to the same criteria:

- the 10 with the highest index values and the 10 with the lowest value,
- the 12 with external numbering and the 8 with internal numbering.

6.1 Some various hydrophobicity scales

Here is invested a first series of credible and recognized hydrophobicity scales. This data is taken from the referencing listed at the end of the article in [7].

6.1.1 Hydrophobicity scales table 1

AA	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*
0Gly	-0.	0.48	0.01	0.74	0	0	0	0.72
1Val	4.2	1.08	0.07	-0.31	1.73	-1.5	4.7	0.86
2Trp	-0.9	0.81	-1.85	0.3	2.56	-3.4	1	0.85
3Cys	2.5	0.29	-0.24	-0.13	0.58	-1	4.1	0.91
4Leu	3.8	1.06	-0.56	-0.55	2.46	-1.8	5.7	0.85
5Phe	2.8	1.19	-1.13	-0.32	2.54	-2.5	4.4	0.88
6Glu	-3.5	-0.74	2.02	2.68	-0.34	3	-1.8	0.62
7Asp	-3.5	-0.9	1.23	3.49	-0.31	3	-3.1	0.62
8Ala	1.8	0.62	0.17	0.11	0.44	-0.5	0.2	0.74
9Tyr	-1.3	0.26	-0.94	0.68	1.63	-2.3	3.2	0.76
10Ser	-0.8	-0.18	0.13	0.84	-0.84	0.3	-0.5	0.66
11Arg	-4.5	-2.53	0.81	2.58	-2.42	3	1.4	0.64
12Met	1.9	0.64	-0.23	-0.1	1.1	-1.3	4.2	0.85
13Ile	4.5	1.38	-0.31	-0.6	2.46	-1.8	4.8	0.88
14Lys	-3.9	-1.5	0.99	2.71	-2.45	3	-3.1	0.52
15Asn	-3.5	-0.78	0.42	2.05	-1.32	0.2	-0.5	0.63
16Thr	-0.7	-0.05	0.14	0.52	-0.41	-0.4	-1.9	0.7
17Gln	-3.5	-0.85	0.58	2.36	-0.71	0.2	-2.8	0.62
18His	-3.2	-0.4	0.96	2.06	-0.01	-0.5	0.5	0.78
19Pro	-1.6	0.12	0.45	2.23	1.29	0	-2.2	0.64

a Kyte and Doolittle
b Eisenberg
c Wimley and S.H. White
d T. Hessa
e Abraham D.J., Leo A.J
f Hopp and Woods
g Cornette
h Rose

Fig. 30 Some amino acid hydrophobicity scales. See Figure 31 also. See complet references [7].

6.1.2 Hydrophobicity scales table 1 results

In the table Figure 31, for each of the different scales, the 10 amino acids with the highest indices and the 10 with the lowest are isolated by a twice column (*a* and *a'* for example to Kyte and Doolittle scale).

For all of these scales, each time it turns out that these two groups of ten amino acids are distributed perfectly in the ratio 3/2 in the two numbering zones.

AA	*a*	*a'*	*b*	*b'*	*c*	*c'*	*d*	*d'*	*e*	*e'*	*f*	*f'*	*g*	*g'*	*h*	*h'*
0Gly	x		x			x	x			x	x			x		x
1Val	x		x			x		x	x			x	x		x	
2Trp		x	x			x		x	x			x	x		x	
3Cys	x		x			x		x	x			x	x		x	
4Leu	x		x			x		x	x			x	x		x	
5Phe	x		x			x		x	x			x	x		x	
6Glu		x		x	x		x			x	x			x		x
7Asp		x		x	x		x			x	x			x		x
8Ala	x		x		x			x	x			x		x	x	
9Tyr		x	x			x		x	x			x	x		x	
10Ser	x			x		x	x			x	x			x		x
11Arg		x		x	x		x			x	x		x			x
12Met	x		x			x		x	x			x	x		x	
13Ile	x		x			x		x	x			x	x		x	
14Lys		x		x	x		x			x	x			x		x
15Asn		x		x	x		x			x	x			x		x
16Thr	x			x	x			x		x	x			x		x
17Gln		x		x	x		x			x	x			x		x
18His		x		x	x		x			x		x	x		x	
19Pro		x		x	x		x		x		x			x		x
12 external AAs	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
8 internal AAs	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
ratio →	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2

10 amino acids with highest index	Scales:	10 amino acids with lowest index
a	Kyte and Doolittle	***a'***
b	Eisenberg	***b'***
c	Wimley and S.H. White	***c'***
d	T. Hessa	***d'***
e	Abraham D.J., Leo A.J	***e'***
f	Hopp and Woods	***f'***
g	Cornette	***g'***
h	Rose	***h'***

Fig. 31 According to different scale data (listed Fig. 30), distribution of the ten highest and ten lowest values in perfect 3/2 ratios in the two predefined external and internal numbering zones.

We can notice that although different, in *b*, the other Eisenberg scale ([7] Eisenberg D., Schwarz E., Komarony M., Normalized consensus hydrophobicity scale. Wall R. J. Mol. Biol. 179: 125-142. 1984.) is organized in the same configurations that singular phenomena highlighted Chapter 5.4 about Eisenberg Weiss scale [4] and illustrated in Figures 24, 25 and 26.

Thus, in this one other scale, the first six externally numbered AAs all have a high index while the last six all have a low index. Oppositely but also alternately 2 by 2, we observe the same phenomenon concerning the set of eight AAs with internal numbering. The ten pairs of oppositely numbered amino acids are so all made up of one high-index entity and one low index. Also, the ten pairs of consecutive numbered amino acids are so always make up of either two high index entities or two low index entities.

6.2 Some others various hydrophobicity scales

Here is invested a second series of credible and recognized hydrophobicity scales. This data is taken from the referencing listed at the end of the article in [7].

6.2.1 Hydrophobicity scales table 2

AA	*i*	*j*	*k*	*l*	*m*	*n*	*o*	*p*
0Gly	0	0	4.48	0	12.43	0.501	0	0.48
1Val	1.32	1.3	7.63	1.34	15.71	0.825	1.22	1.8
2Trp	1.35	2.13	7.66	1.46	13.93	0.878	2.25	0.81
3Cys	0.76	0.25	7.93	0.84	14.63	0.68	1.54	0.29
4Leu	1.8	1.82	8.47	1.8	14.9	0.943	1.7	1.53
5Phe	1.69	2.27	9.03	1.74	14	1	1.79	1.19
6Glu	-1.95	-2.91	3.65	-0.37	11.89	0.043	-0.64	-0.74
7Asp	-2.15	-3.81	3.59	-0.51	10.85	0.028	-0.77	-0.09
8Ala	0.35	0.39	5.33	0.42	12.97	0.616	0.31	0.62
9Tyr	0.39	1.47	5.89	0.51	13.42	0.88	0.96	0.26
10Ser	-0.63	-1.24	4.09	-0.64	11.23	0.359	-0.04	-0.18
11Arg	-1.5	-3.95	4.18	-1.56	11.72	0	-1.01	-2.53
12Met	1.1	0.96	8.95	1.18	14.39	0.738	1.23	0.64
13Ile	1.83	1.82	8.83	1.81	15.67	0.943	1.8	1.38
14Lys	-1.54	-2.77	2.95	-2.03	11.36	0.283	-0.99	-1.5
15Asn	-0.99	-1.91	3.71	-1.03	11.42	0.236	-0.6	-0.78
16Thr	-0.27	-1	4.49	-0.26	11.69	0.45	0.26	-0.05
17Gln	-0.93	-1.3	3.87	-0.96	11.76	0.251	-0.22	-0.85
18His	-0.65	-0.64	5.1	-2.28	12.16	0.165	0.13	-0.4
19Pro	0.84	0.99	3.87	0.86	11.37	0.711	0.72	0.12

i	Cowan R., Whittaker R.G.	***m***	Manavalan P., Ponnuswamy P.K.
j	Roseman M.A.	***n***	Black S.D., Mould D.R.
k	Miyazawa S., Jernigen R.L	***o***	Fauchere J.-L., Pliska V.E
l	Cowan R., Whittaker R.G.	***p***	Tanford C.

Fig. 32 Some amino acid hydrophobicity scales. See Figure 33 also. See complet references [7].

6.2.2 Hydrophobicity scales table 2 results

In the table Figure 33, for each of the different scales, the 10 amino acids with the highest indices and the 10 with the lowest are isolated by columns (*i* and *i'* for example to Cowan R., Whittaker R.G. scale).

For all of these scales, each time it turns out that these two groups of ten amino acids are distributed perfectly in the ratio 3/2 in the two numbering zones.

AA	*i*	*i'*	*j*	*j'*	*k*	*k'*	*l*	*l'*	*m*	*m'*	*n*	*n'*	*o*	*o'*	*p*	*p'*
0Gly		x		x		x		x	x			x		x	x	
1Val	x		x		x		x		x		x		x		x	
2Trp	x		x		x		x		x		x		x		x	
3Cys	x		x		x		x		x		x		x		x	
4Leu	x		x		x		x		x		x		x		x	
5Phe	x		x		x		x		x		x		x		x	
6Glu		x		x		x		x		x		x		x		x
7Asp		x		x		x		x		x		x		x		x
8Ala	x		x		x		x		x		x		x		x	
9Tyr	x		x		x		x		x		x		x		x	
10Ser		x		x		x		x		x		x		x		x
11Arg		x		x		x		x		x		x		x		x
12Met	x		x		x		x		x		x		x		x	
13Ile	x		x		x		x		x		x		x		x	
14Lys		x		x		x		x		x		x		x		x
15Asn		x		x		x		x		x		x		x		x
16Thr		x		x		x		x		x		x		x		x
17Gln		x		x		x		x		x		x		x		x
18His		x		x	x			x		x		x		x		x
19Pro	x		x			x	x			x	x		x			x
12 external AAs	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
8 internal AAs	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
ratio →	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2

10 amino acids with highest index	**Scales:**	**10 amino acids with lowest index**
i	Cowan R., Whittaker R.G.	***i'***
j	Roseman M.A.	***j'***
k	Miyazawa S., Jernigen R.L	***k'***
l	Cowan R., Whittaker R.G.	***l'***
m	Manavalan P., Ponnuswamy P.K.	***m'***
n	Black S.D., Mould D.R.	***n'***
o	Fauchere J.-L., Pliska V.E	***o'***
p	Tanford C.	***p'***

Fig. 33 According to different scale data (listed Fig. 32), distribution of the ten highest and ten lowest values in perfect 3/2 ratios in the two predefined external and internal numbering zones.

We can notice that although different, in *m* and *p*, the Manavalan P., Ponnuswamy P.K. scale and the Tanford C. scale [7] are organized in the same configurations that singular phenomena highlighted Chapter 5.4 about Eisenberg Weiss scale [4] and illustrated in Figures 24, 25 and 26.

Thus, in this two ones other scales, the first six externally numbered AAs all have a high index while the last six all have a low index. Oppositely but also alternately 2 by 2, we observe the same phenomenon concerning the set of eight AAs with internal numbering. The ten pairs of oppositely numbered amino acids are so all made up of one high-index entity and one low index. Also, the ten pairs of consecutive numbered amino acids are so always make up of either two high index entities or two low index entities.

7 Some others various amino acids scales

As we are going to demonstrate now, according to many other various attributes scales, the twenty proteinogenic amino acids are always organized in the 3/2 ratio according to the same criteria:

- the 10 with the highest index values and the 10 with the lowest value,
- the 12 with external numbering and the 8 with internal numbering.

7.1 Amino acids various scales table

Here is invested a series of credible and recognized scales about many and various amino acids attributes. This data is taken from the referencing listed at the end of the article in [8].

AA	*q*	*r*	*s*	*t*	*u*	*v*	*w*	*x*
0Gly	9	0.74	-1.2	0	0	-0.07	0.57	0.79
1Val	5.9	0	3.5	2.7	13.92	-0.06	1.06	2.63
2Trp	5.4	0.13	16.3	14.9	42.53	-0.14	1.08	0.89
3Cys	5.5	2.75	-9.2	-6.8	35.77	0.21	0.7	0.91
4Leu	4.9	0	20	8.8	18.78	-0.18	1.21	1.42
5Phe	5.2	0	19.2	13.2	29.4	-0.12	1.13	1.3
6Glu	12.3	0.92	-7.1	-16.9	17.26	0.05	1.51	0.59
7Asp	13	1.38	-2.9	-8.2	12	0.05	1.01	0.5
8Ala	8.1	0	7.3	0.5	4.34	0.03	1.42	1
9Tyr	6.2	0.2	5.9	6.1	31.53	-0.04	0.69	1.08
10Ser	9.2	1.42	-4.1	1.2	6.35	0.13	0.77	0.7
11Arg	10.5	0.65	-3.6	0.8	26.66	0.06	0.98	0.68
12Met	5.7	0	5.6	4.8	21.64	0.03	1.45	1.49
13Ile	5.2	0	6.6	13.9	19.06	-0.01	1.08	2.6
14Lys	11.3	0.33	-3.7	0.1	21.29	0.1	1.14	0.59
15Asn	11.6	1.33	-5.7	0.8	13.28	0.08	0.67	0.54
16Thr	8.6	0.71	0.8	2.7	11.01	0.1	0.83	0.59
17Gln	10.5	0.89	-0.3	-4.8	17.56	0.05	1.11	0.28
18His	10.4	0.58	-2.1	-3.5	21.81	0.36	1	0.38
19Pro	8	0.39	5.1	6.1	10.93	-0.2	0.57	0.35

q	Grantham R.	***u***	Jones. D.D.
r	Grantham R.	***v***	M.J. Betts, R.B. Russell
s	Browne C.A.	***w***	Chou-Fasman
t	Meek J.L.	***x***	Lifson S., Sander C.

Fig. 34 Some amino acid various attributes scales. See Figure 35 also. See complet references [8].

Here is in more detail what these data scales correspond to:

- *q* scale is about polarity,
- *r* scale is about atomic weight ratio of hetero elements in end group to C in side chain,
- *s* scale is about retention coefficient in TFA,
- *t* scale is about retention coefficient in HPLC,
- *u* scale is about refractivity,
- *v* scale is about amino acid properties and consequences of substitutions,
- *w* scale is about conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins,
- *x* scale is about conformational preference for parallel beta strand.

7.2 Amino acids various scales table results

Thus, as illustrated Figure 35, with still other various fields of study, the attributes of the twenty amino acids are also organized in exact ratios of value 3/2 in the two predefined zones of numberings according to the size of the indices.

AA	*q*	*q'*	*r*	*r'*	*s*	*s'*	*t*	*t'*	*u*	*u'*	*v*	*v'*	*w*	*w'*	*x*	*x'*
0Gly	x		x			x		x		x		x		x	x	
1Val		x		x	x		x			x		x	x		x	
2Trp		x		x	x		x		x			x	x		x	
3Cys		x	x			x		x	x		x			x	x	
4Leu		x		x	x		x		x			x	x		x	
5Phe		x		x	x		x		x			x	x		x	
6Glu	x		x			x		x		x	x		x			x
7Asp	x		x			x		x		x	x			x		x
8Ala		x			x			x		x		x	x		x	
9Tyr		x			x		x		x			x		x	x	
10Ser	x		x			x	x			x	x			x		x
11Arg	x		x			x		x	x		x			x		x
12Met		x			x		x		x			x	x		x	
13Ile		x			x		x		x			x	x		x	
14Lys	x			x		x		x	x		x		x			x
15Asn	x		x			x		x		x	x			x		x
16Thr	x		x		x		x			x	x			x		x
17Gln	x		x			x		x		x	x		x			x
18His	x		x			x		x	x		x			x		x
19Pro		x		x	x		x			x		x		x		x
12 external AAs	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
8 internal AAs	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
ratio →	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2	3/2

10 amino acids with highest index	**Scales:**	**10 amino acids with lowest index**
q	Grantham R.	***q'***
r	Grantham R.	***r'***
s	Browne C.A.	***s'***
t	Meek J.L.	***t'***
u	Jones. D.D.	***u'***
v	M.J. Betts, R.B. Russell	***v'***
w	Chou-Fasman	***w'***
x	Lifson S., Sander C.	***x'***

Fig. 35 According to different scale data (listed Fig.34), distribution of the ten highest and ten lowest values in perfect 3/2 ratios in the two predefined external and internal numbering zones.

7.3 Singular results convergence

The first and last ten indices of the last scale presented in Figure 34, that of the Lifson and Sander scale (*x*) on the propensity of amino acids to promote a β-sheet structure, have a very particular distribution in the ratio 3/ 2. This distribution of amino acids, intimately linked to the genetic coding, is identical to that of the Eisenberg scale introduced in Chapter 5.4. As *b* scale, *n* scale and *p* scale, the first six externally numbered AAs all have a high index while the last six all have a low index. Oppositely but also alternately 2 by 2, we observe the same phenomenon concerning the set of eight AAs with internal numbering. The ten pairs of oppositely numbered amino acids are so all made up of one high-index entity and one low index. Also, the ten pairs of consecutive numbered amino acids are so always make up of either two high index entities or two low index entities.

Thus it turns out that five different scales about different attributes of the twenty proteinogenic amino acids, are organized identically in a singular phenomenon very dependent on the numbering system proposed in this paper. The result convergence of these five attribute scales greatly accredits the veracity and the great consideration that must be made of these singular arrangements in connection with the numbering system of the twenty proteinogenic amino acids.

8. Symmetry and atom count

Here is studied anatomy of the twenty amino acids and connections between radical symmetry and atom count of complet AA.

8.1 Amino acid symmetry

From a schematic point of view, we distinguish here between amino acids with symmetric radicals and those with asymmetric radicals. This symmetry concerns the way in which the molecular chain is distributed beyond the alpha carbon. The nature of the bonds (which can be single or double) is not considered but just the distribution of the groups of atoms in the radical.

As illustrated Figure 36, just ten amino acids have a symmetrical radical.

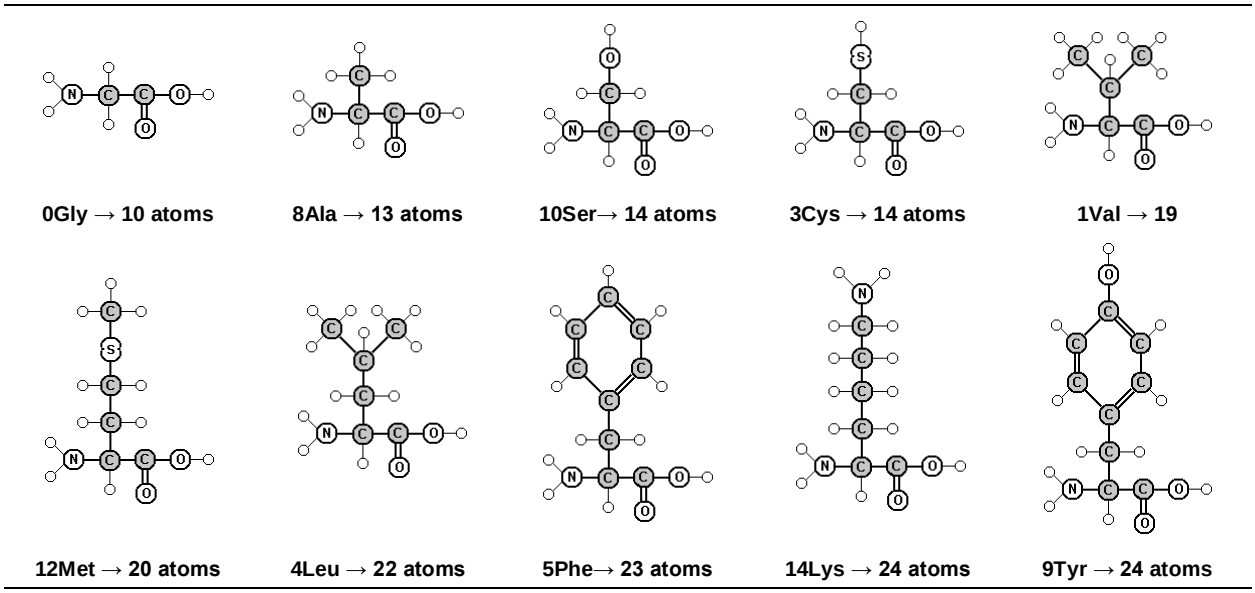


Fig. 36 The ten amino acids with symmetrical radical: 5 AAs at atom count from 10 to 19 and 5 AAs from 20 to 29.

Consecutively, Figure 37, ten amino acids do not have a symmetric radical structure.

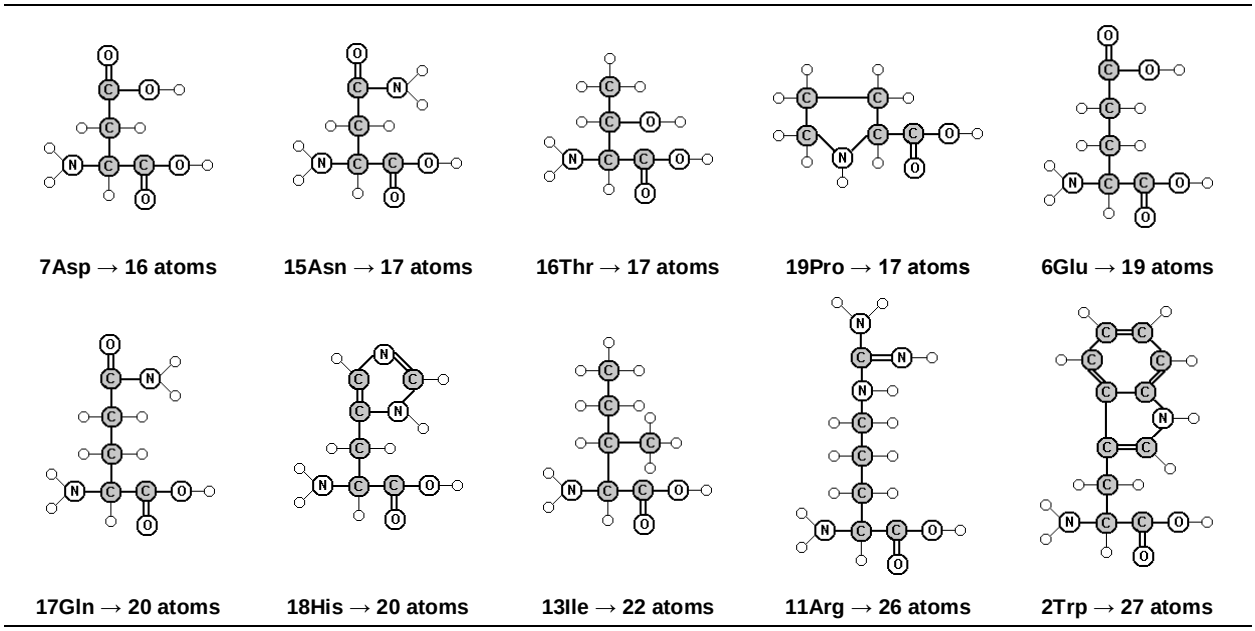


Fig. 37 The ten amino acids with asymmetrical radical: 5 AAs at atom count from 10 to 19 and 5 AAs from 20 to 29.

It turns out, Figure 38, that the distribution of these two sets of ten amino acids, distinguished by their symmetry or non-symmetry, is organized in 3/2 ratios in the two predefined numbering zones.

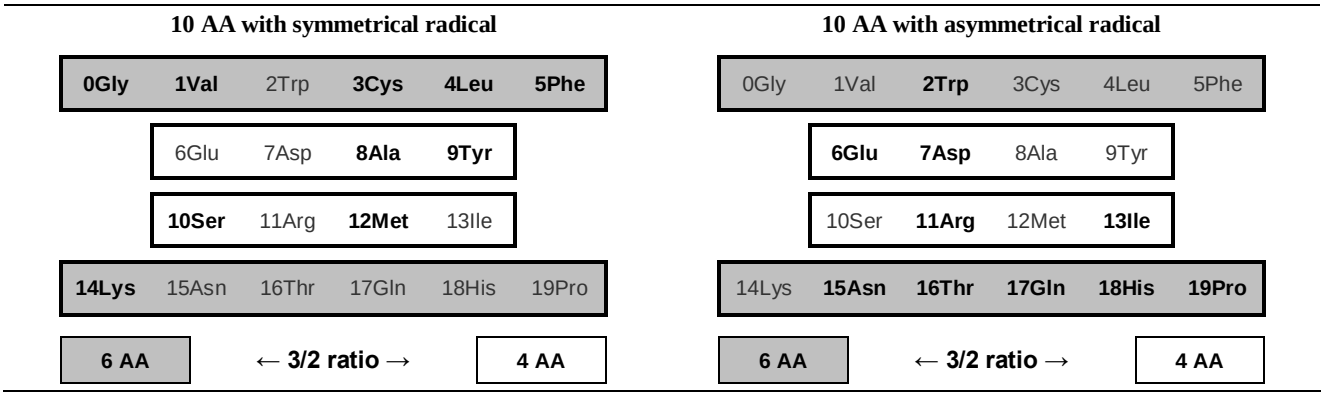


Fig. 38 Distribution of the two sets of the ten symmetrical and ten asymmetrical AAs in 3/2 ratio according to internal and external numbering. See Figures 36 and 37 also.

8.2 Atom count

We demonstrate here that the number of atoms contained in each of the twenty amino acids, in their complete version (base + radical as presented in the introduction Figure 8) is related to the decimal system.

Indeed, it turns out that the smallest amino acid, Glycine has exactly 10 atoms. Also, the ten amino acids with the smallest number of atoms have 19 as maximum 19. The other ten have a number of atoms from 20 to 29 (27 more precisely for Trp). Thus, we can say that 10 AAs have a number of atoms of ten and 10 others of two tens.

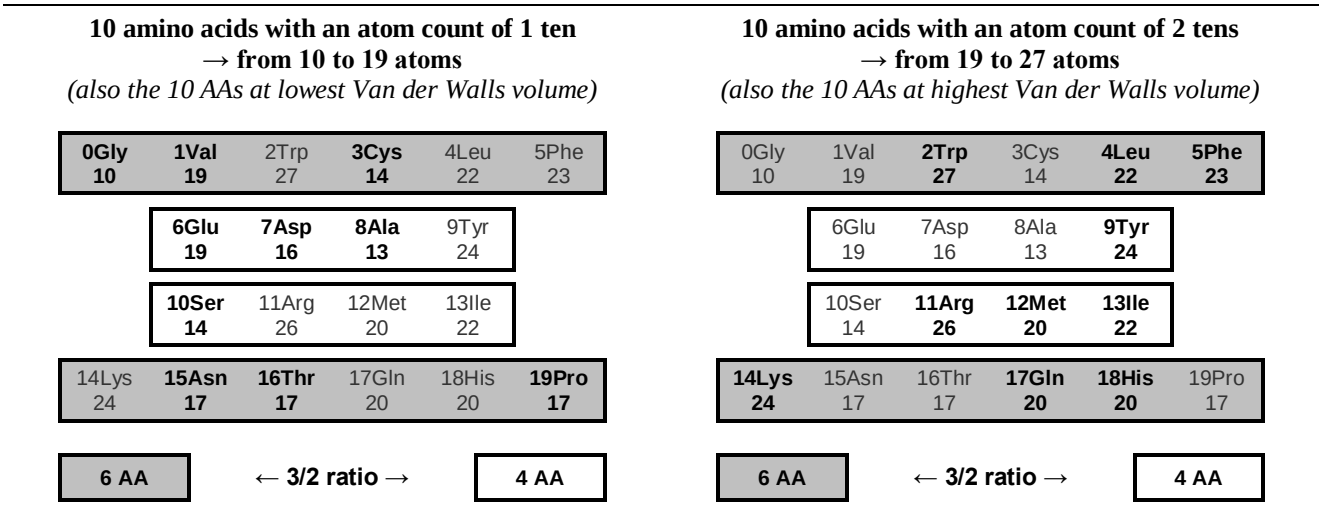


Fig. 39 Distribution of the two sets of the ten AAs at one tens number of atoms and at two tens numbers in 3/2 ratios according to internal and external numbering. See Figures 36 and 37 also. See Figures 27 and 28 also about Van der Walls volume.

These two sets of 10 AAs are distributed in perfect 3/2 ratios in accordance with the two predefined numbering zones as illustrated Figure 39. Also, these two sets are identical to those introduced Chapter 5.5 about Van der Walls volume.

8. 3 Symmetry and atom count transcendence

As it is clearly visible in Figures 36 and 37 but even more synthesized in Figure 40, it turns out that these two notions introduced here, that of symmetry of the radical (or not symmetry) and that of quantity of number of atoms transcend each other completely.

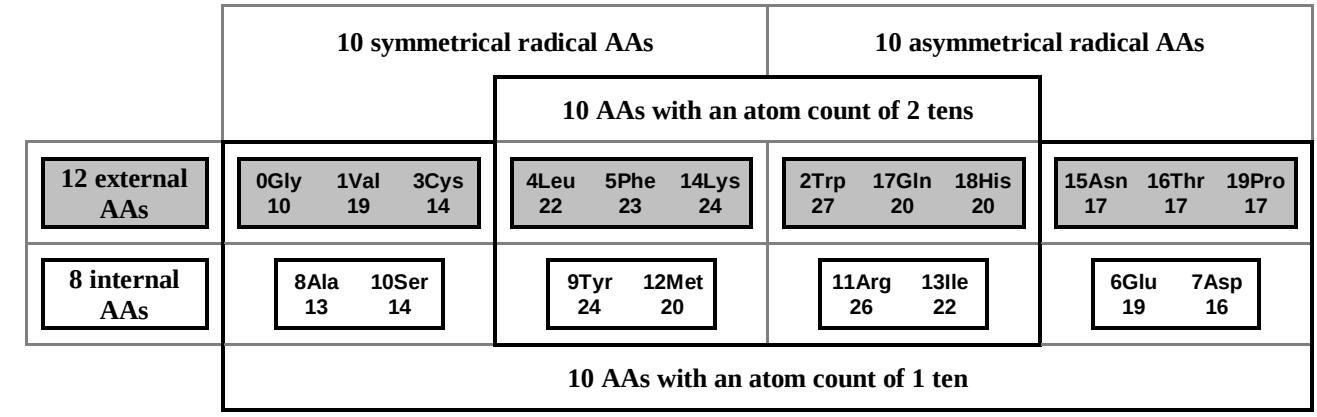


Fig. 40 Distribution of four AAs subsets in perfect 3/2 ratios according to their numbering, their radical symmetry state and their atom counting.

Thus, do we identify four subsets of five amino acids:

- 5 AAs at symmetrical radical and at atom number of one ten,
- 5 AAs at symmetrical radical and at atom number of two tens,
- 5 AAs at asymmetrical radical and at atom number of one ten,
- 5 AAs at asymmetrical radical and at atom number of two tens.

Also, systematically, in each of these four subsets, in exact 3/2 ratios, three amino acids are externally numbered and 2 AAs are internally numbered.

8.3.1 Amino acid fractal organization

This remarkable organization of the twenty proteinogenic amino acids turns out in reality to be fractal in nature, as illustrated by the graphic in Figure 41.

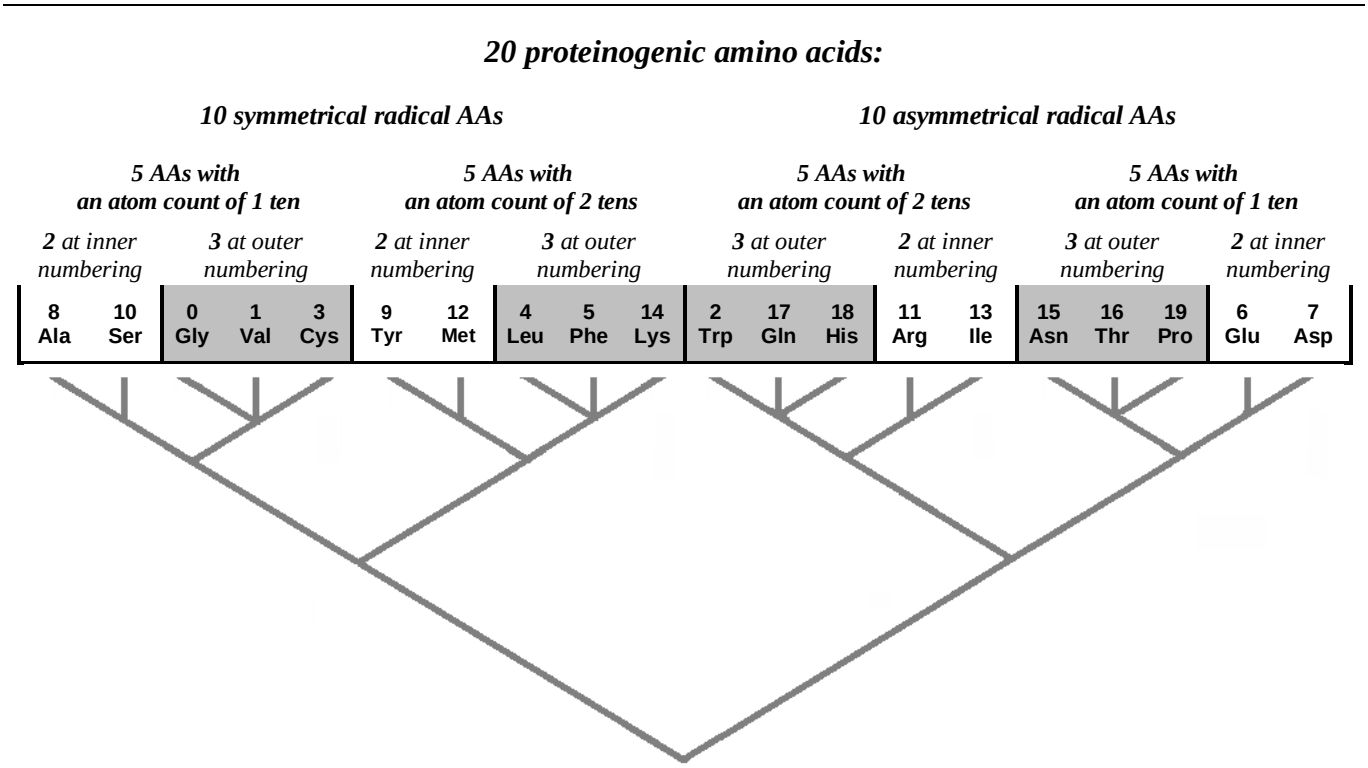


Fig. 41. Fractal distribution of amino acids according to two physico-chemical criteria and to one numbering criterion (derived from the physico-chemical properties of nucleobases). Appearance of the final 3/2 ratio in this fractal configuration of the genetic code.

This 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 physico-chemical criteria. These two sets can each be split into two subsets of 5 AAs. Finally each of these subsets can be separated into sets with ultimate numbers of 3 and 2 entities.

9 Phenomena about DNA coding

As we are now going to demonstrate, it is not only the structure or the various attributes of the twenty proteinogenic amino acids that generate singular phenomena of opposition of value in 3/2 ratios. Thus, the coding attributes of AAs also generate this same type of arithmetic phenomena.

9.1 Largest codon sets

In chapter 4.5 we demonstrated that, at most, each amino acid can be encoded with from one to four codons at two identical first nucleobases and that this generate 3/2 ratios arrangements.

It turns out that, as illustrated in Figure 42, precisely 10 AAs are coded with 1 or 4 codons as largest set (with 2 identical first bases) and 10 others with 2 or 3 codons as largest set. Also, these two sets of 10 AAs, are distributed in exact ratios of value 3/2 with for each set, 6 AAs at external numbering and 4 at internal numbering.

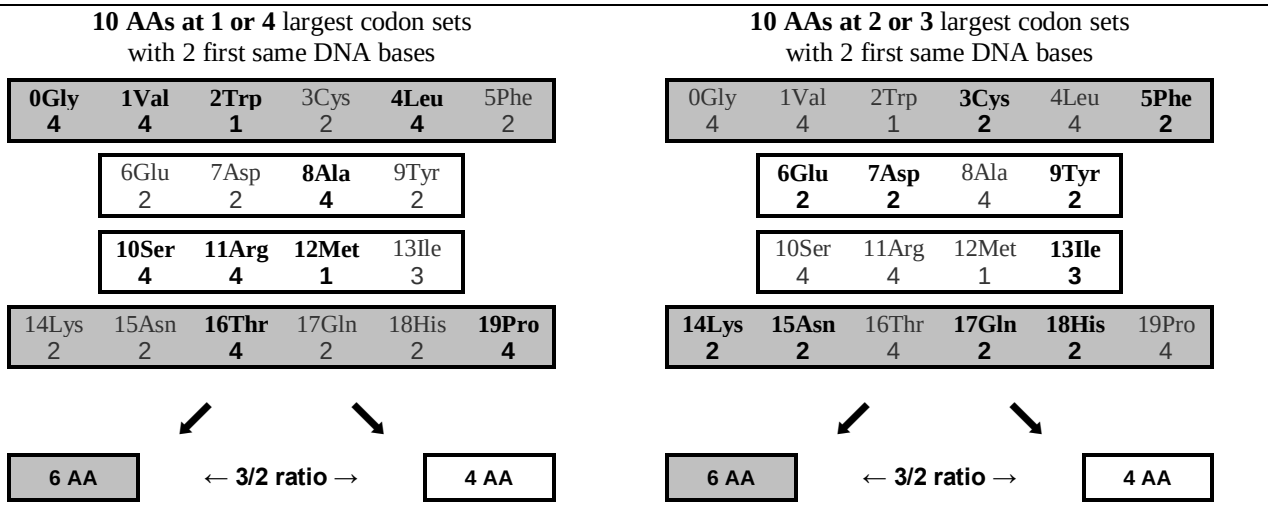


Fig. 42 AA distribution in perfect 3/2 ratios in the two predefined numbering zones according to greatest codon number set. See Fig. 19 also about lighter genetic code.

9.2 Identical or complementary nucleobases as codons

In the genetic code table of Figure 43 are identified the codons with three identical or complementary nucleobases, that is to say with either three bases A or/and T or with three bases G or/and C.

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

Fig.43 Identification of codons with only three identical or complementary nucleobases in the complet genetic code table.

9.2.1 Two sets of 10 AAs

It turns out that 16 codons have this feature but, in the end, these 16 codons just code for exactly 10 different amino acids. Anecdotally*, we also note that among these 10 AAs, in a ratio of 3/2, 6 AAs are coded with 3 G or/and C bases and 4 AAs with 3 A or/and T bases.

We therefore identify here two AA sets:

- 10 AAs with codons at only three identical or complementary nucleobases,
- 10 AAs without codons at only three identical or complementary nucleobases.

As illustrated Figure 44, these two sets of 10 AAs are distributed in perfect 3/2 ratios in the two predefined numbering zones.

* But this is perhaps not by chance and yet not discussed here.

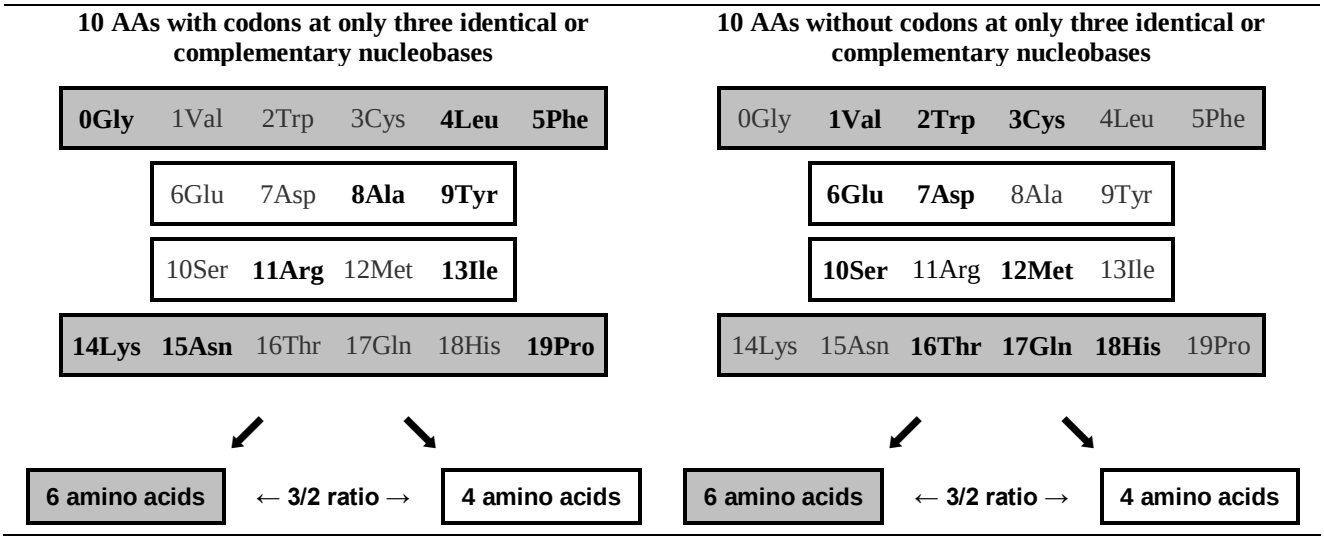


Fig. 44 Distribution in perfect 3/2 ratios in the two predefined numbering zones of the two AA sets with or without codons at only three identical or complementary nucleobases. See Figure 43 also about genetic code.

9.2.2 Identical or complementary nucleobases and largest codon sets transcendence

As it is clearly visible and synthesized in Figure 45, from the tables in Figures 42 and 44, it turns out that these two notions introduced here, that of largest codon sets with 2 first same DNA bases and that with (or without) codons at only 3 identical or complementary nucleobases transcend each other completely.

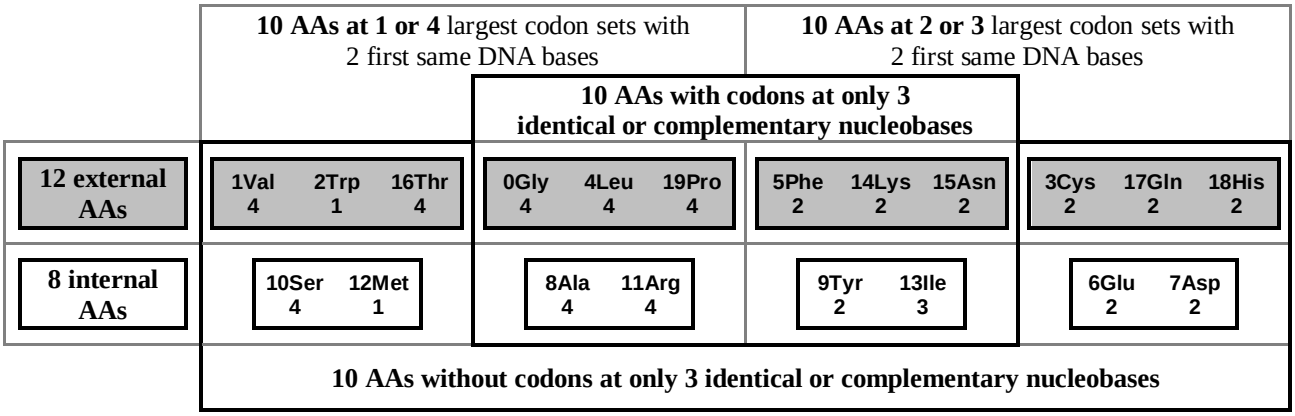


Fig. 45 Distribution of four AAs subsets in perfect 3/2 ratios according to their numbering, and to several aspects of their DNA coding. See Figures 19, 42, 43 and 44 also.

Thus, do we identify four subsets of five amino acids:

- 5 AAs at 1 or 4 largest codon sets* and at codons at only 3 identical or complementary nucleobases
- 5 AAs at 1 or 4 largest codon sets* and without codons at only 3 identical or complementary nucleobases,
- 5 AAs at 2 or 3 largest codon sets* and at codons at only 3 identical or complementary nucleobases,
- 5 AAs at 2 or 3 largest codon sets* and without codons at only 3 identical or complementary nucleobases.

Also, systematically, in each of these four subsets, in exact 3/2 ratios, three amino acids are externally numbered and 2 AAs are internally numbered.

Thus, according to these coding criteria, here again a fractal organization of the genetic code emerges. These layouts are indeed very similar to those presented in Figure 41 in Chapter 8.3.1 about the criteria of radical symmetry and number of atoms of AAs.

* with 2 first same DNA bases.

9.3 Symmetrical codons (same 1st and 3rd identical nucleobases)

Among the set of 64 codons there are symmetrical codons, i.e. codons with the same first and last nucleobase. For example, we consider the *ATA* codon to be symmetric.

9.3.1 Symmetrical codons

As it appears in the table of the genetic code in Figure 46, sixteen codons are symmetrical. also, these 16 codons code for 15 different amino acids, i.e. 5*x* AAs with *x* equal to 3.

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

Fig. 46 Identification of symmetrical codons (same 1st and 3rd identical nucleobases) in the complet genetic code table.

This consideration does not separate the 20 amino acids into two sets of equal size but nevertheless into two sets of 5*x* entities: 15 at symmetrical codons versus 5 at not symmetrical codons.

9.3.2 Symmetrical codons and numbering zones

Nevertheless, as it appears in Figure 47, these two sets of AAs of unequal size, are however also distributed in a ratio of value 3/2 in the two zones of predefined numbering.

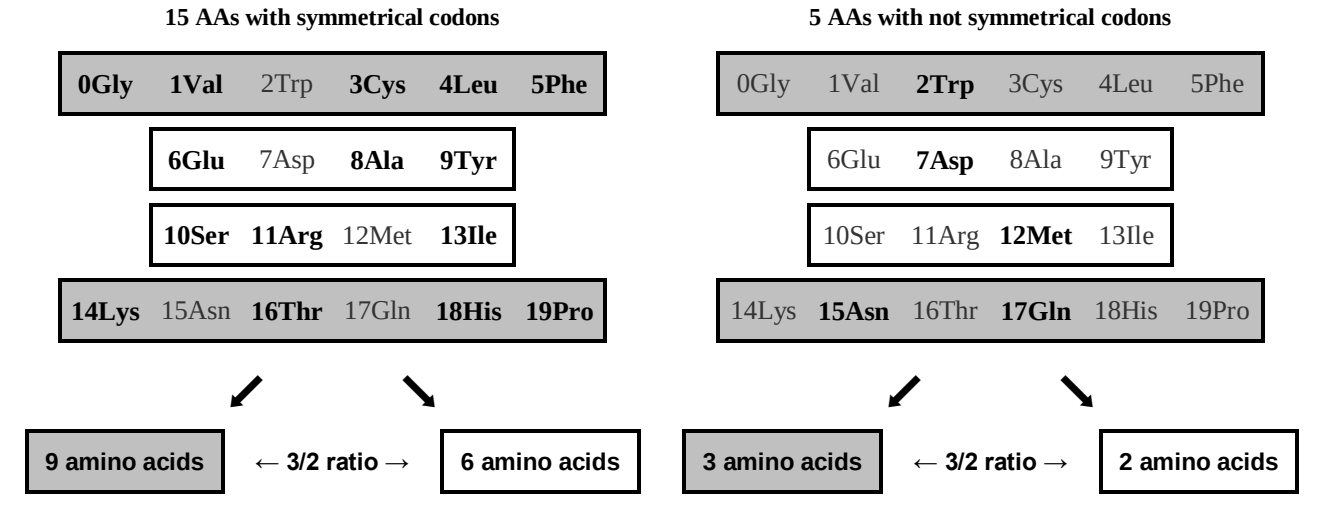


Fig. 47 Distribution in 3/2 ratios of the two AA sets at symmetrical codons (same 1st and 3rd identical nucleobases) or at not symmetrical codons according to predefined numbering zones.

9.3.3 CH_2 groups and numbering zones

The codon symmetry criterion is not the only one to separate the twenty proteinogenic amino acids into two sets of unequal size.

As shown in Figure 48, the distinction of amino acids with CH_2 group, a concept introduced in Chapter 4.3, from those without these groups, isolates the twenty AAs into two sets of the same inequalities with 15 AAs with CH_2 groups versus 5 without CH_2 group.

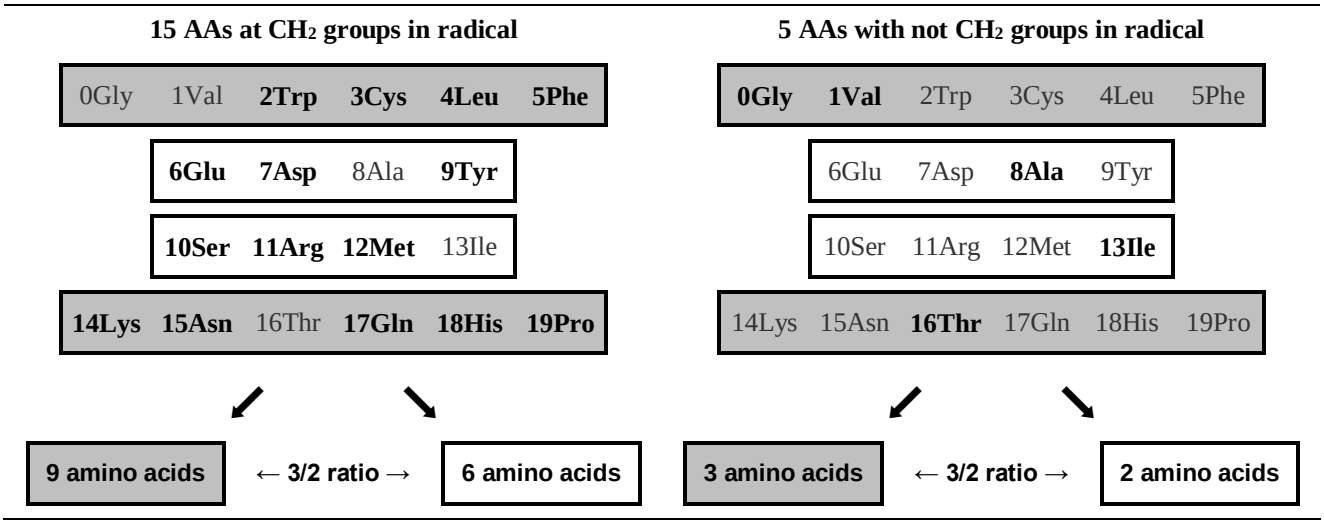


Fig. 48 Distribution in 3/2 ratios of the two AA sets with CH_2 groups in radical or without CH_2 groups according to predefined numbering zones. See Figure 7 Chapter 3.1 and Figure 14 Chapter 4.3 also.

9.3.4 Symmetrical codons and CH_2 groups transcendence

In view of the list of criteria synthesized in Figure 49, it turns out that the presence or absence of CH_2 groups in the radicals of the amino acids and the fact that they are coded or not by a symmetrical codon are in strong interactions.

AA	15 AAs with CH_2 groups	15 AAs with symmetrical codons	5 AAs without CH_2 groups	5 AAs without symmetrical codons	10 AAs with CH_2 groups and symmetrical codons	10 AAs without CH_2 groups or symmetrical codons
0Gly		x	x			x
1Val		x	x			x
2Trp	x			x		x
3Cys	x	x			x	
4Leu	x	x			x	
5Phe	x	x			x	
6Glu	x	x			x	
7Asp	x			x		x
8Ala		x	x			x
9Tyr	x	x			x	
10Ser	x	x			x	
11Arg	x	x			x	
12Met	x			x		x
13Ile		x	x			x
14Lys	x	x			x	
15Asn	x			x		x
16Thr		x	x			x
17Gln	x			x		x
18His	x	x			x	
19Pro	x	x			x	
Counting by numbering area	9	9	3	3	6	6
	6	6	2	2	4	4
ratio →	3/2	3/2	3/2	3/2	3/2	3/2

Fig. 49 Listing of amino acids according to whether or not they are composed of CH_2 groups and whether they are symmetrically coded or not.

It appears in fact that each of the five amino acids not consisting of a methylene bridge (CH_2) directly connected to the alpha carbon is systematically coded by (at least) one symmetrical codon as defined above. Also, each of the five amino acids not encoded by a symmetric DNA triplet always has at least one CH_2 group witch is connected to alpha carbon.

Thus, none of the twenty amino acids are found at the same time in one or the other of these two sets of five entities. This constitutes an important point of view which links, as a whole, the molecular structure and the coding of the twenty proteinogenic amino acids. Table in Figure 50 summarizes this symbiosis between these molecular and coding criteria about amino acids.

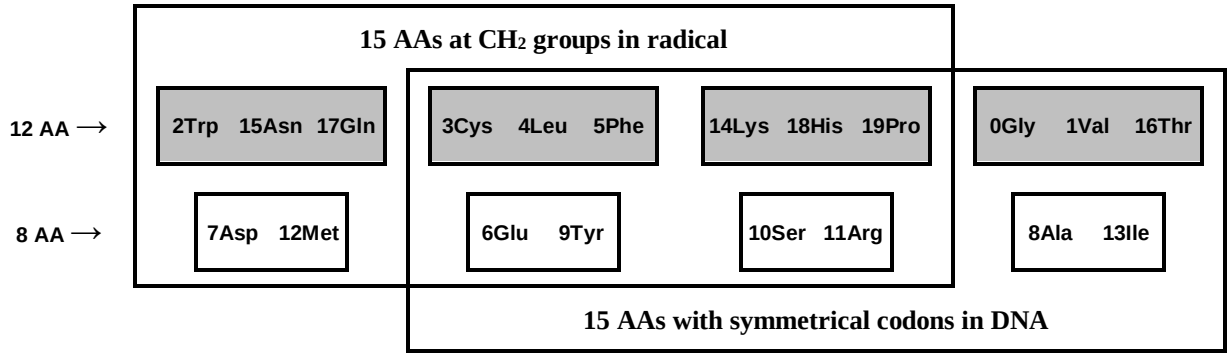


Fig. 50 Criterion symbiosis of nature coding and molecular structural and transcending also with numbering concept. See Fig. 47, Fig 48 and 49 also.

Also, of course, these different sets of amino acids always respect a distribution in 3/2 ratios according to the two numbering zones.

This has the consequence of being able to isolate two groups of ten amino acids:

- 10 AAs with symmetrical codons and at CH_2 groups,
- 10 AAs with not symmetrical codons or without CH_2 group.

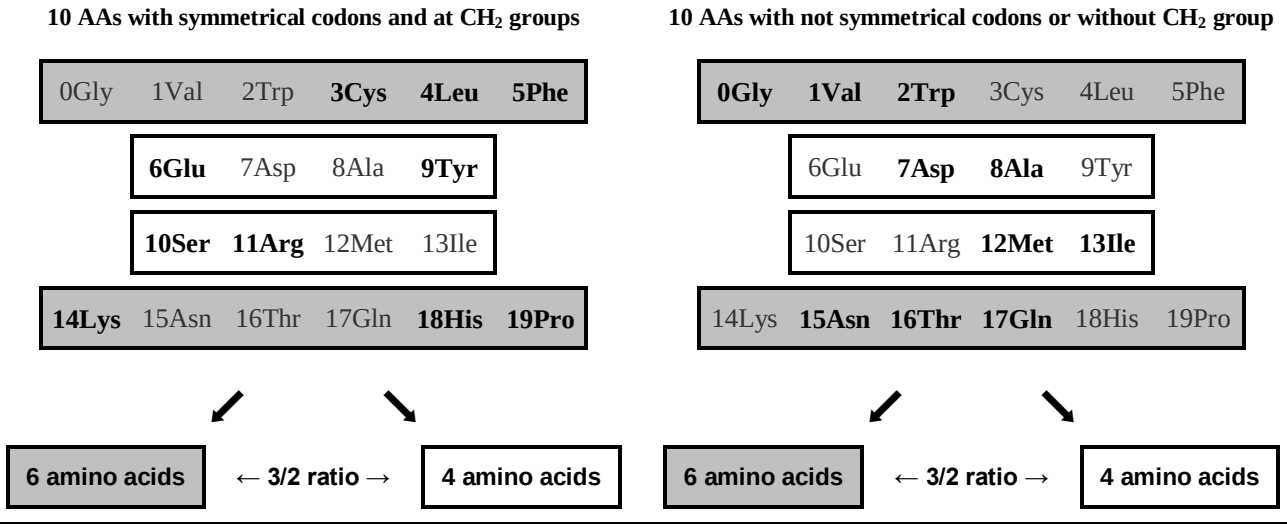


Fig. 51 The two sets of 10 AAs generated from CH_2 group concept and symmetrical coding concept.

Thus, as it is summarized in table Figure 51, there is a set of 10 entities combining the two defined criteria (molecular structure and coding) and another set of 10 amino acids individually presenting only one property.

10 Strong focus on CH_2 groups concept

We have already discussed several times the concept of possession of methylene bridges that amino acids can have. We continue our investigations here by reinforcing the idea that this notion is intimately linked to that of amino acid numbering, the main subject of this paper.

10.1 CH_2 groups number parity

In Chapter 4.3 was demonstrated that account of CH_2 groups is equal to 5*x* entities and that these were distributed in 3/2 ratios in accordance with the two predefined numbering zones. Inside the radicals of the twenty amino acids, there are from 0 to 4 of these CH_2 groups directly attached to the alpha carbon.

We now demonstrate that the parity of the number of these CH_2 groups also generates singular arithmetic phenomena in connection with the concept of numbering proposed in this paper.

As illustrated Figure 52, segregation of CH_2 groups in even or odd number inside AA radicals generate also a opposition of these entities in 3/2 ratios according to the two numbering areas. Moreover, these two sets of CH_2 groups oppose themselves in 3/2 ratio with 15 CH_2 groups in even number versus 10 in odd number.

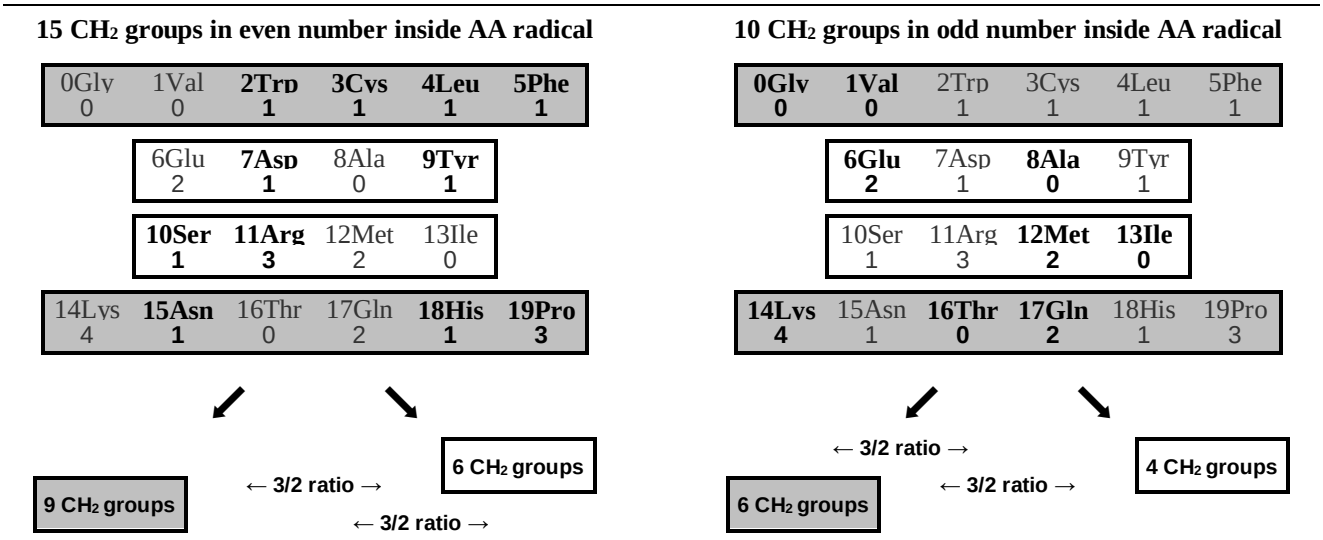


Fig.52 Distribution of the 25 CH_2 groups in various and interconnected 3/2 ratios according to the parity of their number in the radicals and of the two predefined numbering zones.

10.1.1 CH_2 Groups parity and remarkable identity

As shown in Figure 52 and much more in detail in Figure 53, the distribution and number of CH_2 groups according to the double criterion of parity and numbering introduced here is perfectly organized into the remarkable identity:

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

This, with 3 and 2 as respective values for *a* and *b*.

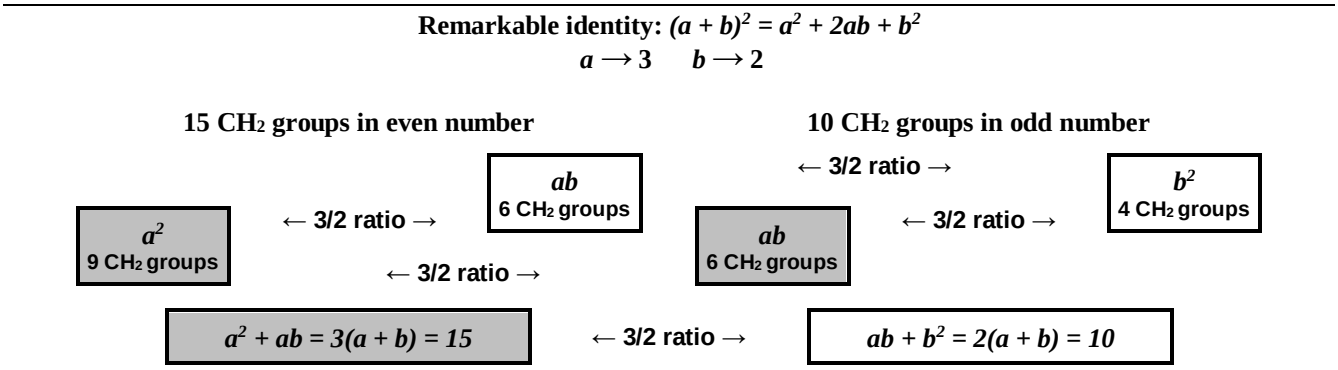


Fig. 53 Remarkable identity revealed in the count of CH_2 groups in even number or in odd number and according to numbering areas. See Figure 52 also. See Figure A4 in appendix for comparison.

Thus, the quantity of CH_2 groups in even number and in external numerated AAs corresponds to the value a^2 of the remarkable identity and the quantity of CH_2 groups in even number and in internal numerated AAs corresponds to the value *ab*.

The quantity of CH_2 groups in odd number and in external numerated AAs also corresponds to the value *ab* and that of CH_2 groups in odd number and in internal numerated AAs corresponds to the value b^2.

These different values therefore transcend into these equal ratios:

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

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

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

This perfect arithmetic arrangement of the 25 CH_2 groups variously distributed within the radicals of the twenty proteinogenic amino acids confirms the idea that this cannot be a hazard phenomenon. The following demonstration will greatly support this point of view.

Also, about quantum study of the five living matter atoms, a very similar arithmetical organization will revelled in appendix Chapter A2. for these five atoms, very singularly, shells and subshells are identically opposite.

10.2. CH_2 Groups and other numbering areas

Here, the twenty amino acids are distributed in slightly different numbering areas. However, the differentiation of these zones remains in the same general idea as those studied in the whole of this paper. Thus, Figure 54, we isolate now the six external AAs numbered from the four internal ones from 0 to 9 and likewise for the ten AAs numbered from 10 to 19.

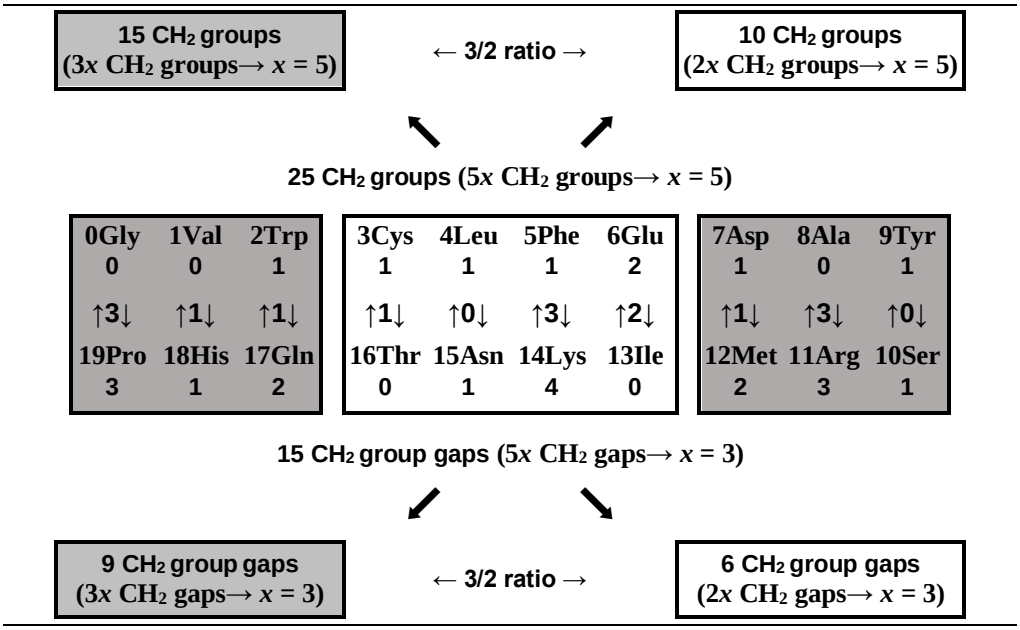


Fig. 54 CH_2 groups counting in two new symmetrical numbering zones of 12 versus 8 entities.

In this other arrangement some close to the prime one, it is remarkable to note that the 25 CH_2 groups continue to oppose each other in exact ratio of value 3/2 with 15 CH_2 counted in the outer numbering zone of 12 AAs versus 10 CH_2 in the inner zone of 8 AAs. Also as in the prime numbering arrangement (See Chapter 4.3.1), in absolute value, the difference in the number of CH_2 groups between two amino acids of opposite numbering (0 versus 19, 1 versus 18, etc.) also overall generates an opposition of values in an exact 3/2 ratio with 9 CH_2 group gaps in external area versus 6 CH_2 group gaps in internal area.

At last, as shown in Figure 55, according to the AA numbering parity and new numbering areas introduced here, the CH_2 groups distribution is also perfectly organized into the remarkable identity $(a + b)^2 = a^2 + 2ab + b^2$. This, with 3 and 2 as respective values for *a* and *b*.

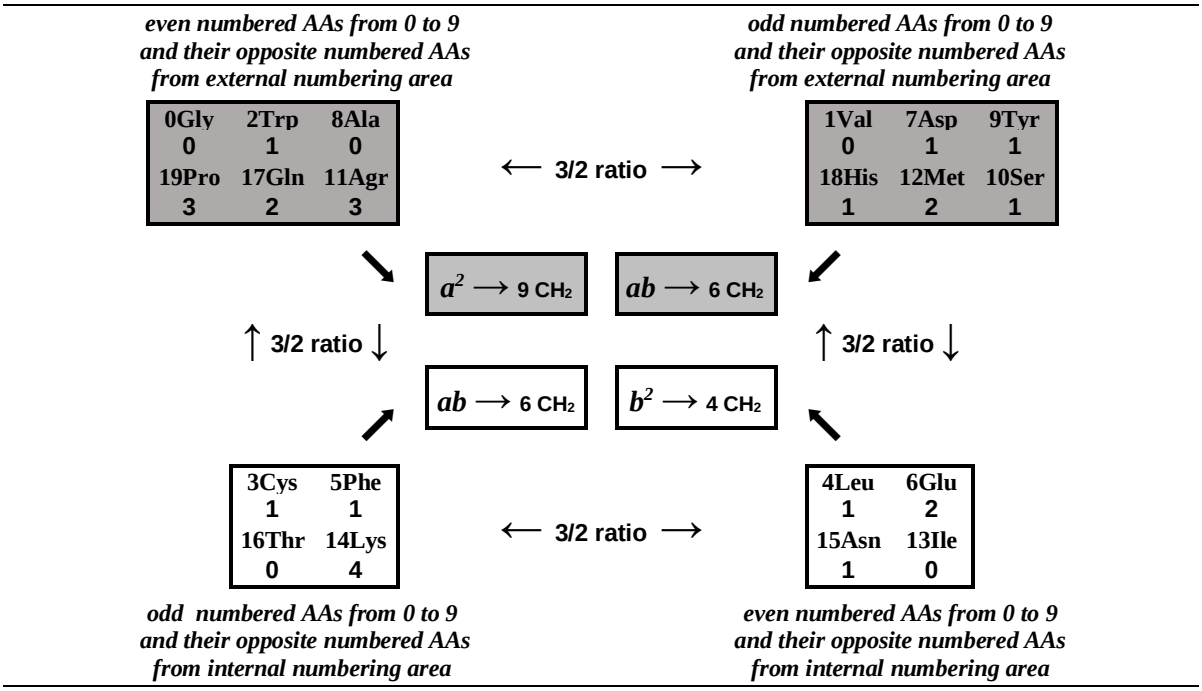


Fig. 55 Organisation in remarkable identity of 25 CH_2 groups according to the AAs numbering parity and new numbering areas.

These last singular observations close the large number of investigations of this paper made about the proposed numbering system of the twenty proteinogenic amino acids.

11. Alphanumeric symbol proposal

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 this paper, we have numbered these entities from 0 to 19 and have 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 56 and 57, 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.

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

Fig. 56 Conventional nomenclature and alphanumeric symbol proposal to proteinogenic amino acids into 5 characters: 2 digits + 3 letters. Here Glycine as example.

The table in Figure 57 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.

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**

Fig. 57 Conventional nomenclature and alphanumeric symbol proposal to the twenty proteinogenic amino acids into 5 characters: 2 digits + 3 letters.

We also propose, in a legitimate logic, that these twenty amino acids appear in the tables listing them in this order. Indeed, this sequence corresponds to certain real coding criteria and even, as demonstrated throughout this paper, to physical criteria both linked to the DNA triplets and to the specific characteristics of the amino acids.

Thus we propose to no longer present them in alphabetical order as practiced until now because this order does not correspond to any physico-chemical criteria. In their new alphanumeric order, the data concerning them will be more directly visible and usable. This, as the various tables presented in this paper have already demonstrated.

12. Other 3/2 ratios regarding genetic code organization

The numbering of the twenty proteinogenic amino acids is not the only concept to generate singular arithmetic phenomena opposing the entities of the genetic code in various ratios of value 3/2. But as this is not the main subject of this paper, these other investigations are presented in the appendix.

We are just drawing attention here to the fact that Glycine, which is simply like an amino acid base, has all these various components at *5x* in number (atoms, protons, 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.

We therefore strongly encourage the reader to consult this appendix, the observations of which confirm the main idea of this article that the genetic code, confused AAs and nucleobases, is organized arithmetically in the ratio of 3/2 value.

Also, some other genetic code investigations are in connections with the subject of this paper especially about ratio 3/2, symmetry, listing of proteinogenic amino acids or more generally connections between number theory and the genetic code. As example, some of these investigations are listed in references in [9].

13. Discussions and conclusions

The universal genetic code encodes very exactly and only twenty proteinogenic amino acids, that is 5*x* entities. These 5*x* entities are made up of only 5 different atoms, including three with two quantum shells and two with 1 or 3 quantum shells, i.e. three atoms with an even number of shells and two atoms with an odd number of quantum shells. The set of twenty amino acids operating within the universal genetic code totals 205 atoms, or 5*x* atoms. Among these 205 atoms, 85, or 5*x* atoms, have an even number of quantum shells and 120, or 5*x* atoms have an odd number of quantum shells. Also, in all of these twenty proteinogenic amino acids, there are very precisely 25 CH_2 groups (methylene bridges) linked to the alpha carbon, so again 5*x* entities.

Since their coding characteristics, that is to say the nature of the nucleobases which encode them, the twenty amino acids can be legitimately numbered from 0 to 19, i.e. twenty different numbers being assigned to them. Since this numbering, by operating a symmetrical distinction between the 3/5th amino acids with external numbering and the 2/5th with internal numbering, it turns out that the main components and other attributes of these twenty amino acids, considered as a whole, oppose in exact ratios of value 3/2 in these two zones of numbering made up of 3*x* and 2*x* amino acids.

Also it turns out that, for a very large number of attribute scales, such as the various recognized hydrophobicity scales, the distinction of the ten amino acids with the strongest indices from the ten with the weakest generates a distribution in various ratios of exact value 3/2 in these two numbering zones with always, for each set of 10 AAs, 6 AAs distributed in the external numbering zone versus 4 distributed in the internal zone.

In addition, according to other various criteria, both in terms of radical structure and coding characteristics (configuration of DNA triplets), two sets of 5*x* AAs are always distinguished, i.e. 10 versus 10 entities or sometimes 15 versus 5 AAs. These different sets of 5*x* entities are also divided into ratios of value 3/2 in the two numbering areas which are subject of this paper.

Also, some criteria transcend each other, revealing a fractal organization of the genetic code. This fractal organization, like all the other arithmetic arrangements presented in this paper, operates with the three absolute values 2, 3 and 5. These numbers are not arbitrary. They are all simply the first three prime numbers. Also the number 5 is the only prime number which is the sum of two consecutive primes: 2 and 3. And these two numbers are the only ones which are simultaneously two consecutive primes and two consecutive integers. Thus we can say that the global structure of the genetic code is directly related to the theory of numbers since it is organized with the first three and singular prime numbers that are 2, 3 and 5.

This study further reveals that the set of 64 codons, or the coding structure, and the set of 20 proteinogenic amino acids, or the coded structure, must in fact be considered as a single and unique entity that we more formally name *The Genetic Code*.

Since we have clearly demonstrated here that the structure of the twenty amino acids, in many aspects, is intimately linked to the nature of nucleobases and to the numbering of both codons and coded which emanates from them, we propose a new nomenclature official list of these twenty proteinogenic amino acids. This nomenclature, supplementing the current one (trivial name, 3-letter and 1-letter symbols) is alphanumeric in nature and is standardized in 2 digits and 3 letters as form.

It is therefore strongly suggested to list the twenty amino acids according to this new nomenclature and in numerical order (from 00Gly to 19Pro) rather than as currently in alphabetical order. This new nomenclature being related to numerous and real physico-chemical criteria (both coding and structure of AAs) must therefore be favoured in the study of the mechanism of the genetic code and of the twenty proteinogenic amino acids, a set of study subject which we call *The Genetic Code*.

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. hal-02902700v4* [10].

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 [1 and 2] 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 Sulfur 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 5*x* 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 → 9 and C → 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. As described in the author's previous paper [9], the multiplicity of protons/modules within an 8/1 ratio of amino acids is not random, but concerns exactly 50% of the twenty amino acids used in the genetic code, i.e. 10 amino acids out of 20.

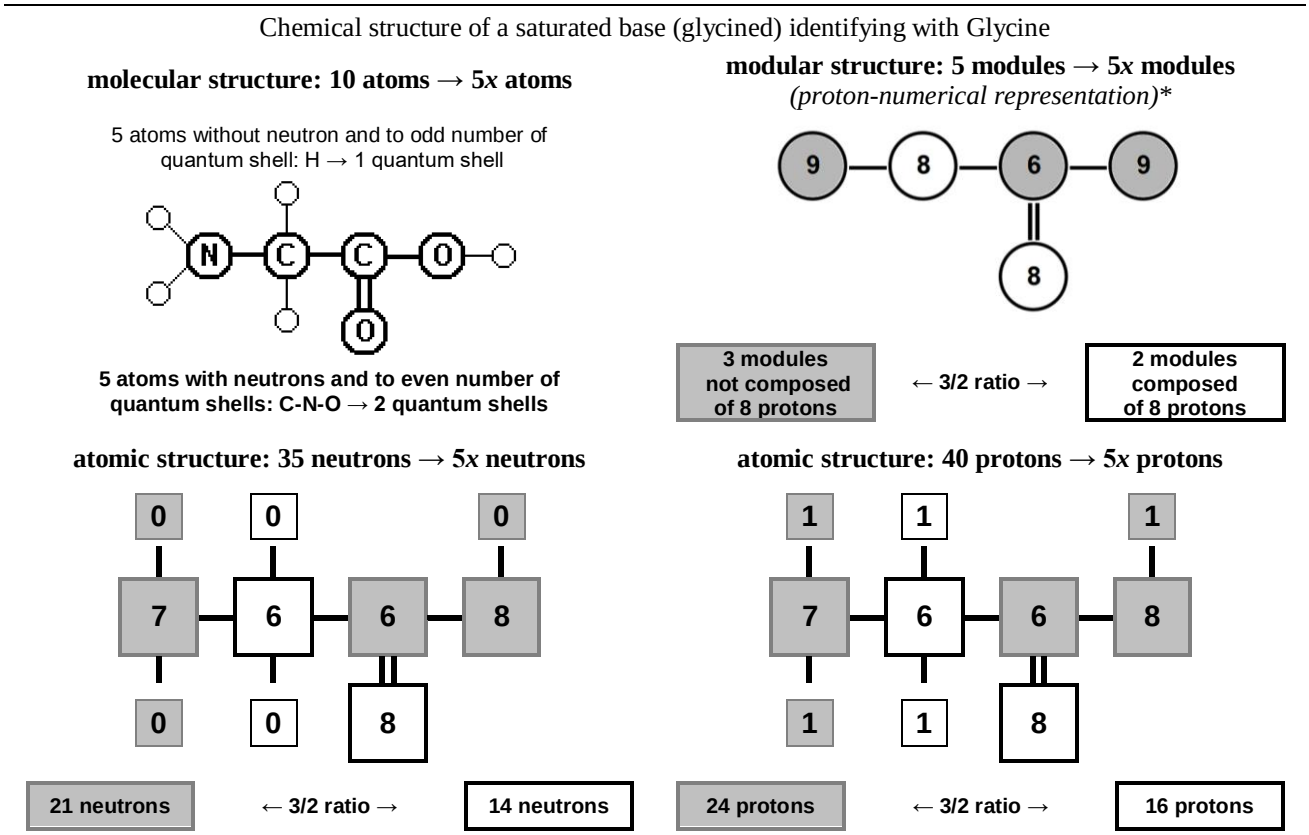


Fig. A1 Chemical, modular and atomic structure of a saturated base identified with the amino acid Glycine: 5 modules, 10 atoms, 40 protons and 35 neutrons. See also Fig. A2. * inspired representation from Sergey Petoukhov [1 and 2].

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.

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.

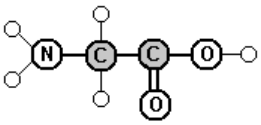
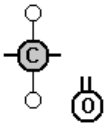
Glycine entities	Total entities number 	Entities account in 3 no 8-proton modules 	Entities account in 2 8-proton modules 
5 modules	**5** **5*x* → *x* = 1**	**3** **3*x* → *x* = 1**	**2** **2*x* → *x* = 1**
10 atoms	**10** **5*x* → *x* = 2**	**6** **3*x* → *x* = 4**	**4** **2*x* → *x* = 4**
5 non-hydrogen 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**

Fig. 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.

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 (1s, 2s, 2p, 3s, 3p) 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 (*1s*, *2s*, *2p*, *3s*, *3p*)	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		Hydrogène 1	Soufre* 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. Also, this is absolutely similar to the organization of the CH_2 groups as it was exposed Figure 53 Chapter 10.1 about the anatomy of the twenty amino acids.

Remarkable identity $(a + b)^2 = a^2 + 2ab + b^2$ where a and b have the values 3 and 2

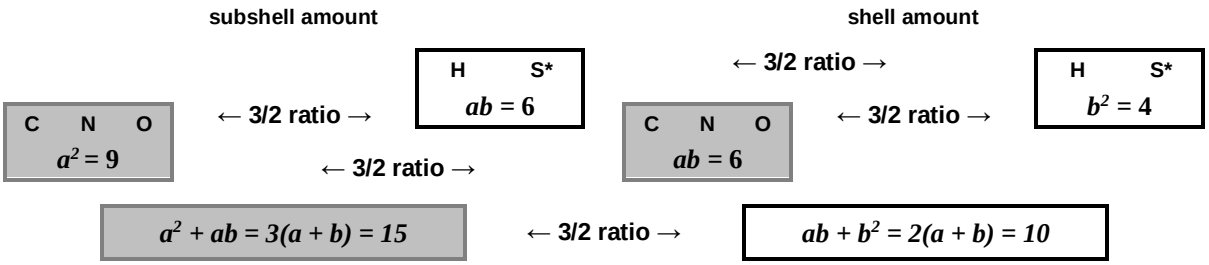


Fig. A4 Remarkable identity revealed in the count of subshells and quantum shells of the five elements H, C, N, O and S (*P in DNA). See Fig. A3. See Fig. 53 chapter 10.1 for comparison.

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

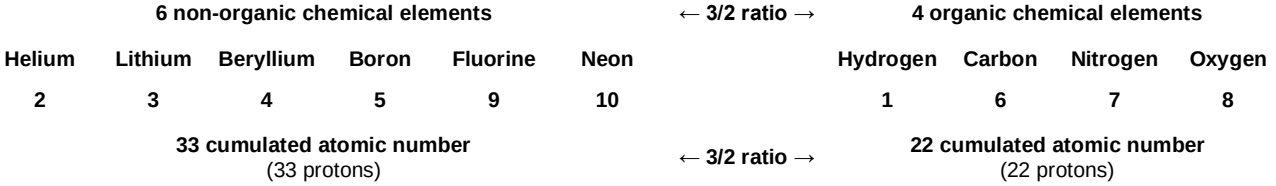


Fig. A5 Opposition of the 6 non-organic chemical elements and 4 organic chemical elements about their respective cumulated atomic number.

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.

References

1. 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.

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

3. Sweet R.M., Eisenberg D. Optimized matching hydrophobicity (OMH). J. Mol. Biol. 171:479-488. 1983.

4. D. Eisenberg, R. M. Weiss and T. C. Terwilliger. Hydrophobic Moments and Protein Structure. Faraday Symp. Chem. Soc. 17, 109-120 1982 and , Nature (London), 299,371. 1982.

5. T. Creighton. Proteins Structure and Molecular Properties. NY. W. H. Freeman & Co. 1994.

6. Garnier-Osguthorpe-Robson (GOR) Garnier J, Osguthorpe DJ, Robson B. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J Mol Biol 120:97-120. 1978.

7. Hydrophobicity scales:
(a) Hydropathicity. J. Kyte and R.F. Doolittle, "A Simple Method for Displaying the Hydropathic Character of a Protein" J Mol Biol 157:105. 1982.
(b) Eisenberg D., Schwarz E., Komarony M., Normalized consensus hydrophobicity scale. Wall R. J. Mol. Biol. 179: 125-142. 1984.
(c) W.C. Wimley and S.H. White, "Experimentally determined hydrophobicity scale for proteins at membrane interfaces" Nature Struct Biol 3:842 1996.
(d) T. Hessa, H. Kim, K. Bihlmaier, C. Lundin, J. Boekel, H. Andersson, I. Nilsson, S.H. White, and G. von Heijne, "Recognition of transmembrane helices by the endoplasmic reticulum translocon" Nature 433:377. 2005.
(e) Abraham D.J., Leo A.J. Proteins: Structure, Function and Genetics 2:130-152. 1987.
(f) Hydrophilicity. Hopp T.P., Woods K.R. Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828. 1981.
(g) J. L. Cornette; K. B. Cease; H. Margalit; J. L. Spouge; J. A. Berzofsky & C. DeLisi. Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. J Mol Biol, 195, 659-685. 1987.
(h) Rose G.D., Geselowitz A.R., Lesser G.J., Lee R.H., Zehfus M.H. Science 229: 834-838. 1985.
(i) Hydrophobicity indices at ph 7.5 determined by HPLC. Cowan R., Whittaker R.G. Peptide Research 3:75-80. 1990.
(j) Hydrophobicity (pi-r). Roseman M.A. J. Mol. Biol. 200:513-522. 1988.
(k) Hydrophobicity (contact energy derived from 3D data). Miyazawa S., Jernigen R.L. Macromolecules 18:534-552. 1985.
(l) Hydrophobicity indices at ph 3.4 determined by HPLC. Cowan R., Whittaker R.G. Peptide Research 3:75-80. 1990.
(m) Average surrounding hydrophobicity. Manavalan P., Ponnuswamy P.K. Nature 275:673-674. 1978.
(n) Hydrophobicity of physiological L-alpha amino acids. Black S.D., Mould D.R. Anal. Biochem. 193:72-82. 1991.
(o) Hydrophobicity (pi-r).Fauchere J.-L., Pliska V.E. Eur. J. Med. Chem. 18:369-375. 1983.
(p) Hydrophobicity scale (Contribution of hydrophobic interactions to the stability of the globular conformation of proteins). Tanford C. J. Am. Chem. Soc. 84:4240-4274. 1962.

8. Others scales:
(q) Polarity. Grantham R. Science 185:862-864. 1974.
(r) Atomic weight ratio of hetero elements in end group to C in side chain. Grantham R. Science 185:862-864. 1974.
(s) Retention coefficient in TFA. Browne C.A., Bennett H.P.J., Solomon S. Anal. Biochem. 124:201-208. 1982.
(t) Retention coefficient in HPLC, pH 7.4. Meek J.L. Proc. Natl. Acad. Sci. USA 77:1632-1636. 1980.
(u) Refractivity. Jones. D.D. J. Theor. Biol. 50:167-184. 1975.
(v) M.J. Betts, R.B. Russell. Amino acid properties and consequences of substitutions. In Bioinformatics for Geneticists, M.R. Barnes, I.C. Gray eds, Wiley. 2003.
(w) Chou, P.Y. and G.D. Fasman. "Conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins." Biochem. 13: 211-222. 1974
(x) Conformational preference for parallel beta strand. Lifson S., Sander C. Nature 282:109-111. 1979.

9. Others references

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.

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

Petoukhov S., He M. Symmetrical Analysis Techniques for Genetic Systems and Bioinformatics: Advanced Patterns and Applications.- IGI Global, Hershey, USA, p. 271, 2009

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

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

10. Jean-Yves Boulay. Genetic code, quantum physics and the 3/2 ratio. 2020. ⟨ hal-02902700v4⟩

Acknowledgements

We are very grateful to Sergey Petoukhov, and György Darvas for their invaluable technical and logistical assistance for the successful publication of this article.

- Sergey Petoukhov. Chief of Laboratory of biomechanical systems, Mechanical Engineering Research Institute of Russian Academy of Sciences, Moscow, Russia.

- György Darvas. Director at Symmetrion http://symmetry.hu/ (Retired from the IRO of the Hungarian Academy of Sciences and from the Faculty of Sciences of the Eötvös Loránd University Budapest)

Jean-Yves BOULAY independent researcher (without affiliation) – FRANCE –
jean-yvesboulay@orange.fr ORCID: 0000-0001-5636-2375