Article

Groups of Symmetries of the Two Classes of Synthetases in the Four-Dimensional Hypercubes of the Extended Code Type II

Marco V. José [*] , Eberto R. Morgado , Juan R. Bobadilla

Posted Date: 29 August 2023

doi: 10.20944/preprints202308.1951.v1

Keywords: Aminoacyl-tRNA synthetases; Standard Genetic Code; Symmetry groups; Origin and evolution of the genetic code; Group Theory

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2023 doi:10.20944/preprints202308.1951.v1

Article

Groups of Symmetries of the Two Classes of Synthetases in the Four-Dimensional Hypercubes of the Extended Code Type II

Marco V. José [1,*], Eberto R. Morgado [2] and Juan R. Bobadilla [1]

[1] Theoretical Biology Group, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Apdo, Postal 70228, Ciudad de México 04510, México. jromanb@unam.mx (JRB)
[2] Facultad de Matemática, Física y Computación, Universidad Central "Marta Abreu" de Las Villas, Santa Clara, Cuba, ebertorodobaldo@gmail.com (ERM)
* Correspondence: marcojose@biomedicas.unam.mx (MVJ)

Abstract: Aminoacyl-tRNA synthetases (aaRSs) originated from an ancestral bidirectional gene (mirror symmetry) and through the evolution of the genetic code the twenty aaRSs exhibit a symmetrical distribution in a 6-dimensional hypercube of the Standard Genetic Code. In this work, we assume a primeval RNY code, and the Extended Genetic RNA code type II which includes codons of the type YNY, YNR and RNR. Each of the 4 subsets of codons can be represented in a 4-dimensional hypercube. Altogether these 4 subcodes constitute the 6-dimensional representation of the SGC. We identify the aaRSs' symmetry groups in each of these hypercubes. We show that each of the four hypercubes contains the following sets of symmetries for the two known Classes of synthetases: RNY: dihedral group D_4; YNY: binary group $\square_2$; YNR: amplified octahedral group $\mathrm{O_h}$; and RNR: binary group $\square_2$. We demonstrate that for each hypercube, the group of symmetries of Class 1 is the same as the group of symmetries of Class 2. The biological implications of these findings are discussed.

Keywords: aminoacyl-tRNA synthetases; standard genetic code; symmetry groups; origin and evolution of the genetic code; group theory

---

1. Introduction

Aminoacyl-tRNA synthetases (aaRSs) are key players of the genetic code in all living beings. AaRS attaches an amino acid to the cognate tRNA, and the aminoacyl-tRNA is then used for translation upon binding to mRNA according to the codon–anticodon interaction on the ribosome. The Standard Genetic Code (SGC) is the mapping of 61 codons or triplets to 20 canonical amino acids. There are 20 aaRSs, one for each of the 20 standard amino acids. AaRSs are divided into two mutually exclusive Classes, 1 and 2, based on their structural, functional, and evolutionary relatedness [1–8]. Each aaRS falls into either Class 1 or Class 2, except for lysyl-tRNA synthetase (LysRS), which has a representative in both classes. The correspondence between amino acids and synthetases is one to one, that is, bijective. For every triplet, or codon there is a synthetase, associated to the amino acid it specifies. In a previous work [9], we outlined the group of symmetries of both Classes of synthetases in each of the 4-dimensional hypercubes of the so-called Extended Code of type II [10,11]. Herein, we rigorously determine the symmetries of aaRSs in the 4 subcodes RNY, YNY, YNR, and RNR. The article is organized as follows. First, we provide the basic mathematics needed to understand the derivation of symmetries of the SGC. In Appendix A we present formal mathematical definitions and concepts. Second, we determine the symmetries in each of the four subcodes. We show that the group of symmetries of Class 1 is the same as the group of symmetries of Class 2. We also analyze the 5-dimensional hypercubes NNY and NNR as obtained by the union of RNA with YNY, and YNR with

RNR, respectively. The 4 subcodes constitute the whole SGC. Finally, we discuss the biological implications of the results.

1.1. Basic Mathematics of the SGC

The cartesian product $N^3 = N \times N \times N$, being $N = \{C, U, A, G\}$ the set of the four RNA-nucleotides, C=Cytosine, U=Uracil, A=Adenine, G=Guanine, is the set of the 64 triplets (X, Y, Z), where $X, Y, Z \in N$. The standard genetic code (SGC) may be seen as a surjective function $f : N^3 \rightarrow A_S = A \bigcup \{S\}$, being A the set of the 20 known amino acids, and S the stop signal, which means the instruction of finalizing the process of synthesis of a protein.

For every $a \in A$, its preimage set $f^{-1}\{a\} = \{(X, Y, Z) | f(X, Y, Z) = a\}$, is the set of triplets, also called codons, that encode the amino acid, or stop signal, $a \in A_S$. Since the decipherment of the genetic code [12] (Marshall W. Nirenberg, Har Gobind Khorana and Robert W. Holley, 1967), it is known that the number of coding triplets for any $a \in A_S$ is one of the numbers 1,2,3,4, or 6.

In **Table 1** we show the table of the genetic code, highlighting the 4 subcodes RNY, YNY, YNR and RNR. The table also indicates the subcodes NNY and NNR.

Table 1. Genetic code: Triplets and amino acids in each of the 4-, 5-, and 6-dimensional hypercubes. aaRS Class 1, red *; aaRS Class 2, blue **.

5D NNY		5D NNR	
4D RNY	4D YNY	4D YNR	4D RNR
ACC	CCC	CCA	ACA
Thr**	Pro**	Pro**	Thr**
ACU	CCU	CCG	ACG
Thr**	Pro**	Pro**	Thr**
AUC	CUC	CUA	AUA
Ile*	Leu*	Leu*	Ile*
AUU	CUU	CUG	AUG
Ile*	Leu*	Leu*	Met*
AAC	CAC	CAA	AAA
Asn**	His**	Gln*	Lys*
AAU	CAU	CAG	AAG
Asn**	His**	Gln*	Lys*
AGC	CGC	CGA	AGA
Ser**	Arg*	Arg*	Arg*
AGU	CGU	CGG	AGG
Ser**	Arg*	Arg*	Arg*
GCC	UCC	UCA	GCA
Ala**	Ser**	Ser**	Ala**
GCU	UCU	UCG	GCG
Ala**	Ser**	Ser**	Ala**
GUC	UUC	UUA	GUA
Val*	Phe**	Leu*	Val*
GUU	UUU	UUG	GUG
Val*	Phe**	Leu*	Val*
GAC	UAC	UAA	GAA
Asp**	Tyr*	Stop	Glu*
GAU	UAU	UAG	GAG
Asp**	Tyr*	Stop	Glu*
GGC	UGC	UGA	GGA
Gly**	Cys*	Stop	Gly**
GGU	UGU	UGG	GGG
Gly**	Cys*	Trp*	Gly**

1.2. Arithmetization of the genetic code

The bijective correspondence $C \leftrightarrow 00,\ U \leftrightarrow 01,\ A \leftrightarrow 10, G \leftrightarrow 11,$ between the set N and the set $\mathbb{Z}_2^2 = \mathbb{Z}_2 \times \mathbb{Z}_2$, being $\mathbb{Z}_2 = \{0,1\}$, induces a bijective correspondence between the set NNN of the 64 triplets and the set $\mathbb{Z}_2^6$ of all the sextuples of zeros and ones.

Binary operations: In the binary set $\mathbb{Z}_2 = \{0,1\}$, the two binary operations $\oplus_2, \otimes_2$, the so-called addition and product, module 2, are defined. Their Cayley Tables (**Table 2**) are:

Table 2. The Cayley tables of sum and multiplication of the binary set $\mathbb{Z}_2$..

$\oplus_2$	0	1
0	0	1
1	1	0

$\otimes_2$	0	1
0	0	0
1	0	1

They define, in the set $\mathbb{Z}_2 = \{0,1\}$, the algebraic structure of a field, or commutative division ring. This is the smallest possible field, with only two elements: the neutral of the addition and the neutral of the product, respectively.

The addition $\oplus_2$ is extended, component-wise, to the set $\mathbb{Z}_2^6$, which, with the natural definition of the product of any scalar $\alpha \in \mathbb{Z}_2$, ,by a sextuple $(a_1, a_2, a_3, a_4, a_5, a_6)$, becomes a 6-dimensional vector space over the binary field $(\mathbb{Z}_2, \oplus_2, \otimes_2)$. This vector space is the so-called 6-dimensional binary hypercube.

The bijection between $\mathbb{Z}_2^6$ and the set NNN of the 64 triplets, induces in it the algebraic structure of a 6-dimensional vector space, whose canonical basis is the system: (ACC, UCC, CAC, CUC, CCA, CCU). The addition of nucleotides is shown in the following **Table 3**:

Table 3. Sum module 2 of nucleotide bases.

$\oplus_2$	C	U	A	G
C	C	U	A	G
U	U	U	G	A
A	A	G	C	U
G	G	A	U	C

It is seen that the nucleotide C, cytosine, is the neutral element of the group. This group is isomorphic to de group $(\mathbb{Z}_2, \mathbb{Z}_2, \oplus)$, which is known as the Four Klein Group. It is an Abelian group of order 4, where each element is its own inverse. It is, in the Felix Klein list of finite groups, the fourth, and the first that is not cyclic.

Mutations in the triplets of the genetic code: A mutation in a triplet XYZ is the substitution of any of its components, by some nucleotide. For example, GUC→GAC, where U is replaced by A.

Obviously, a mutation in a triplet may produce the change of the amino acid it encodes. In our example, valine, that is encoded by GUC, is converted into aspartic acid, encoded by GAC. Algebraically, a mutation in a triplet is performed by the addition of triplets of the canonical basis. In our example, the mutation GUC→GAC is obtained by the addition of the triplet CGC= CAC+CUC.

Classification of the nucleotides

There is a partition $\{\mathrm{Y},\mathrm{R}\}$ of the set N, being $\mathrm{Y}=\{\mathrm{C},\mathrm{U}\}$ pyrimidines, and $\mathrm{R}=\{\mathrm{A},\mathrm{G}\}$ purines. The condition of being a pyrimidine or a purine is called the chemical type of a nucleotide.

Transitions and transversion: A mutation is called a transition if it does not change the chemical type, and it is called a transversion if the chemical type is changed. It is easy to notice that transitions are produced by additions of pyrimidines and transversions by addition of purines.

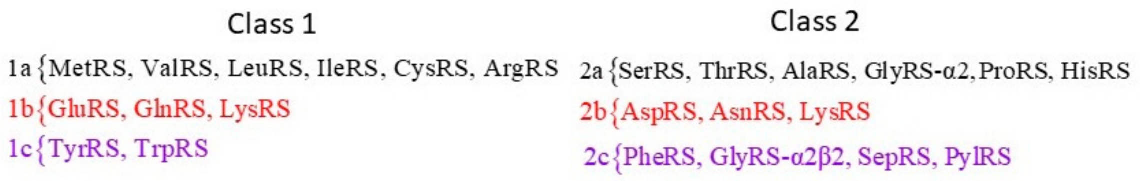
The primaeval genetic code. It is assumed that there was, in the RNA world, a primitive, or primaeval genetic code, of only sixteen triplets and eight amino acids: the set $\mathrm{RNY}=\{(\mathrm{X},\mathrm{Y},\mathrm{Z})\}$, where X is a purine, Y is any of them, and Z is a pyrimidine. It is a 4-dimensional hypercube, a translation of the 4-dimensional subspace YNY, derived from RNY for transversion in the first nucleotide. By transversions in the first or the third component, the hypercubes $\mathrm{YNY}, \mathrm{YNR}$ and RNR are derived from the primaeval RNY. These four sets are pairwise disjoint, and they cover the whole set NNN. Hence, the set of sets $\{\mathrm{YNY}, \mathrm{RNY}, \mathrm{YNR}, \mathrm{RNR}\}$ is a partition of NNN. The hypercube YNY determines, with the addition $\oplus_2$, a subgroup of the additive group $(\mathrm{NNN},\oplus_2)$, being the other $\mathrm{RNY}, \mathrm{YNR}$ and RNR its cosets. With the addition and the product of scalars by vectors, YNY determines a 4-dimensional vector subspace, while $\mathrm{RNY}, \mathrm{YNR}$ and RNR are its affine subspaces, that is, translations of it, namely:

$$\mathrm{RNY} = \mathrm{YNY} + \{\mathrm{ACC}\}$$
$$\mathrm{YNR} = \mathrm{YNY} + \{\mathrm{CCA}\}$$
$$\mathrm{RNR} = \mathrm{YNY} + \{\mathrm{ACA}\}$$

The Classes of Synthetases

The synthetases are enzymes that regulate the selection of amino acids that shall be charged to tRNA molecules. A little more than 20 aaRSs are found in modern organisms. They are classified into two groups, Class 1, and Class 2, each having three subclasses (a, b, and c) based on similarity in sequences and structures [13,14]. The classification is shown in **Table 4**:

Table 4. Classes and subclasses of aminoacyl tRNA synthetases.



Class 1	Class 2
1a {MetRS, ValRS, LeuRS, IleRS, CysRS, ArgRS	2a {SerRS, ThrRS, AlaRS, GlyRS-α2, ProRS, HisRS
1b {GluRS, GlnRS, LysRS	2b {AspRS, AsnRS, LysRS
1c {TyrRS, TrpRS	2c {PheRS, GlyRS-α2β2, SepRS, PylRS

In general, aaRS consists of a catalytic domain, an anticodon-binding domain, and often an editing domain. Each class harbors class-specific characteristic motifs and structural topology in its catalytic domains [3].

According to the RO model [15–18] the table of the genetic code can be divided into the sub-codes NAN, NGN, NUN, NCN. We have also shown that there exists an automorphism *F* of the cube defined also piecewise, which transforms that division into the sub-codes RNR, YNR, RNY, YNY, respectively, which is precisely our algebraic model [19].

3. Results

3.1. Group of symmetries of the classes of synthetases in the hypercube RNY.

The members of Class 2 are in the vertexes of a square, contained in the set $[-1,1]^4$, the convex closure of the set $\{-1,1\}^4$, taken, by a suitable change of coordinates, rotations, and translations of axis, as a representation of the hypercube RNY. See **Figure 1**:

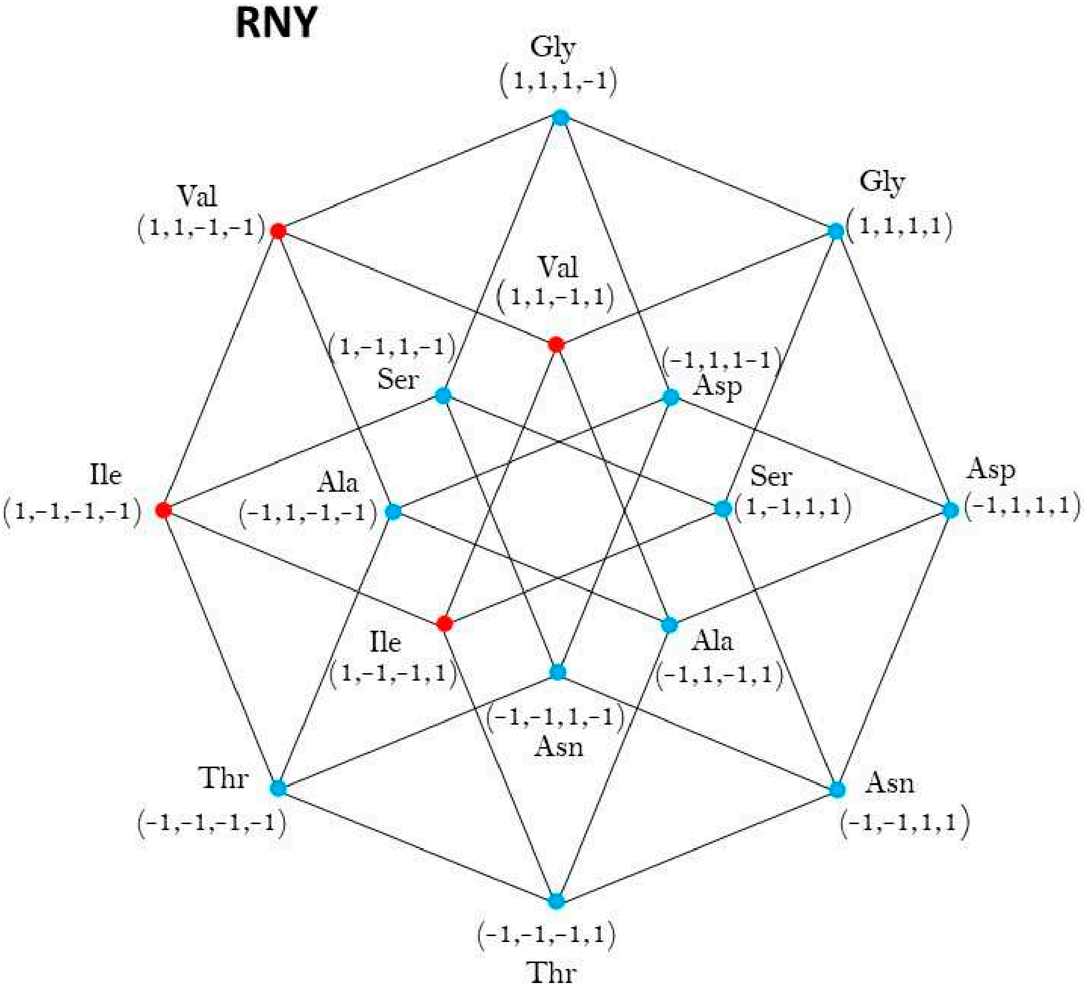


Figure 1. aaRSs distribution in code RNY. Class 1 (red, 2 amino acids) and Class 2 (blue, 6 amino acids). Graphic representation of the subsets RYY and RRY. First 4-dimensional hypercube of the RNY code: $\mathrm{RNY} = \mathrm{RYY} \bigcup \mathrm{RYY}$..

The quadruplets that belong to Class 2 of synthetases (colored in black), are in the set V_2 of vertexes of a square S, one of the 24 faces of the hypercube. V_2 is a subset of the set $\{-1,1\}^4$ of the sixteen vertexes of the hypercube RNY. The set of vertexes of that square is:

$$V_2 = \{v_1, v_2, v_3, v_4\},$$

where

$v_1 = (1,-1,-1,-1)$, $v_2 = (1,-1,-1,1)$, $v_3 = (1,1,-1,1)$, and $v_4 = (1,1,-1,1)$, black colored.

doi:10.20944/preprints202308.1951.v1

The center of this square is the point $c_2 = \frac{v_1 + v_2 + v_3 + v_4}{4} = (1, 0, -1, 0)$.

The members of Class 1 are in the complementary set of vertexes V_1, that is, the other 12 vertexes, red colored, out of the four vertexes of the square S.

$$V_1 = (u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}) \text{ where}$$

$$u_1 = (-1,-1,-1,-1), u_2 = (-1,-1,-1,1), u_3 = (-1,1,-1,1), u_4 = (-1,1,-1,-1),$$

$$u_5 = (-1,-1,1,-1), u_6 = (-1,-1,1,1), u_7 = (-1,1,1,1), u_8 = (-1,1,1,-1),$$

$$u_9 = (1,-1,1,-1), u_{10} = (1,-1,1,1), u_{11} = (1,1,1,1), u_{12} = (1,1,1,-1)$$

Now we can state the following:

3.2. Theorem I. The group of symmetries of Class 1 are the same as that of Class 2.

Proof: As a symmetry of the Class 1 is a bijective isometrical function f from RNY onto itself, it preserves the set $\text{Class } 1 \in RNY$. It also preserves the square S. that contains the quadruples of Class 2, that is, the vertexes of the square S. It means that both classes have the same group of symmetries. That is so, because the binary set of sets $\{\text{Class } 1, \text{Class } 2\}$ is a partition of the finite set $\{-1,1\}^4$ of the 16 vertexes of the hypercube RNY.

Observation: A similar theorem takes place in any of the other hypercubes of the Extended code type II, and in the whole 6-dimensional hypercubes NNN. In the case of the 4-dimensional YNR, the 5-dimensional NNR and the whole 6-dimensional NNN, it is valid if, for methodological reasons, we assume the stop signal as it would be an amino acid of Class 1.

Let us consider the orthogonal matrix $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$, which represents a rotation of angle $\frac{3\pi}{2}$ in the plane $\langle e_2, e_4 \rangle$. It operates over the set of vertexes of the square S in the following way:

$$A: \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} \to \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \to \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \to \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \to \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$$

This means that A is an isometry of the square S.

Let us now consider the orthogonal matrix $B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$, which represents a reflection through the 3-dimensional hyperplane $\langle e_1, e_2, e_3 \rangle$, that reverses the fourth axis □ e_4. It operates over the set of vertexes of the square S in the following way:

$$\mathrm{B}\!:\begin{pmatrix}1\\-1\\-1\\-1\end{pmatrix}\rightarrow\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}\rightarrow\begin{pmatrix}10\\-1\\-1\\-1\end{pmatrix}$$

$$\mathrm{B}:\begin{pmatrix}1\\1\\-1\\1\end{pmatrix}\rightarrow\begin{pmatrix}1\\1\\-1\\-1\end{pmatrix}\rightarrow\begin{pmatrix}1\\1\\-1\\1\end{pmatrix}.$$

This means that B is an isometry of the square S. It is not difficult to verify the following equalities, that occur between the matrixes A and B: $\mathrm{A}^4=\mathrm{B}^2=\mathrm{I}; \mathrm{BA}=\mathrm{A}^3\mathrm{B}.$

They are the defining relations of the dihedral group D_4, the group of all the symmetries of a square.

Now, we will see which is the action of the dihedral group D_4, generated by the matrixes A and B, over the set V_1 of the 12 quadruplets that represent the synthetases of Class 1.

The action of A over the set V_1 is,

$$\mathrm{A}:\mathrm{u}_1=\begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix}\rightarrow\mathrm{u}_2=\begin{pmatrix}-1\\-1\\-1\\1\end{pmatrix}\rightarrow\mathrm{u}_3=\begin{pmatrix}-1\\1\\-1\\1\end{pmatrix}\rightarrow\mathrm{u}_4=\begin{pmatrix}-1\\1\\-1\\-1\end{pmatrix}\rightarrow\mathrm{u}_1=\begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix}$$

$$\mathrm{u}_5=\begin{pmatrix}-1\\-1\\1\\-1\end{pmatrix}\rightarrow\mathrm{u}_6=\begin{pmatrix}-1\\-1\\1\\1\end{pmatrix}\rightarrow\mathrm{u}_7=\begin{pmatrix}-1\\1\\1\\1\end{pmatrix}\rightarrow\mathrm{u}_8=\begin{pmatrix}-1\\1\\1\\-1\end{pmatrix}\rightarrow\mathrm{u}_5=\begin{pmatrix}-1\\-1\\1\\-1\end{pmatrix}$$

$$\mathrm{u}_9=\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}\rightarrow\mathrm{u}_{10}=\begin{pmatrix}1\\-1\\1\\1\end{pmatrix}\rightarrow\mathrm{u}_{11}=\begin{pmatrix}1\\1\\1\\1\end{pmatrix}\rightarrow\mathrm{u}_{12}=\begin{pmatrix}1\\1\\1\\-1\end{pmatrix}\rightarrow\mathrm{u}_9=\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}$$

It is seen that A is an isometry of the set of synthetases of Class 1.

The action of B over the set V_1 is:

$$\mathrm{B}:\mathrm{u}_1=\begin{pmatrix}-1\\-1\\-1\\-1\end{pmatrix}\leftrightarrow\mathrm{u}_2=\begin{pmatrix}-1\\-1\\-1\\1\end{pmatrix}\quad\mathrm{u}_3=\begin{pmatrix}-1\\1\\-1\\1\end{pmatrix}\leftrightarrow\mathrm{u}_4=\begin{pmatrix}-1\\1\\-1\\-1\end{pmatrix}$$

doi:10.20944/preprints202308.1951.v1

$$\mathrm{u}_5=\begin{pmatrix}-1\\-1\\1\\-1\end{pmatrix}\leftrightarrow \mathrm{u}_6=\begin{pmatrix}-1\\-1\\1\\1\end{pmatrix}\quad \mathrm{u}_7=\begin{pmatrix}-1\\1\\1\\1\end{pmatrix}\leftrightarrow \mathrm{u}_8=\begin{pmatrix}-1\\1\\1\\-1\end{pmatrix}$$

$$\mathrm{u}_9=\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}\leftrightarrow \mathrm{u}_{10}=\begin{pmatrix}1\\-1\\1\\1\end{pmatrix}\quad \mathrm{u}_{11}=\begin{pmatrix}1\\1\\1\\1\end{pmatrix}\leftrightarrow \mathrm{u}_{12}=\begin{pmatrix}1\\1\\1\\-1\end{pmatrix}$$

Note that B is an isometry of the set of synthetases of Class 1.

We see that B is a symmetry of the set V_1 vertexes of the Class 1 of synthetases. As A and B are symmetries of the Class 1 of synthetases this result confirms the assertion of Theorem 1, where it was proved that boss classes have the same group of symmetries.

The main result of Part I; We have proved that the group of symmetries of both classes of synthetases, in the hypercube RNY, is the dihedral group D_4.

3.3. *Group of symmetries of the classes of synthetases in the hypercube* YNY

In this case the members of each class are in eight vertexes, out of the sixteen vertexes of the hypercube YNY. They are represented by the sixteen quadruples of the set $\{-1,1\}^4$ of ones and minus one, being the hypercube $[-1,1]^4$, the convex closure of the set $\{-1,1\}^4$, The set $[-1,1]^4$ is taken, by a suitable change of coordinates, rotations, and translations of axis, as a representation of the hypercube YNY. The center of this hypercube is the origin of coordinates $\mathrm{O}=(0,0,0,0)$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 August 2023 doi:10.20944/preprints202308.1951.v1

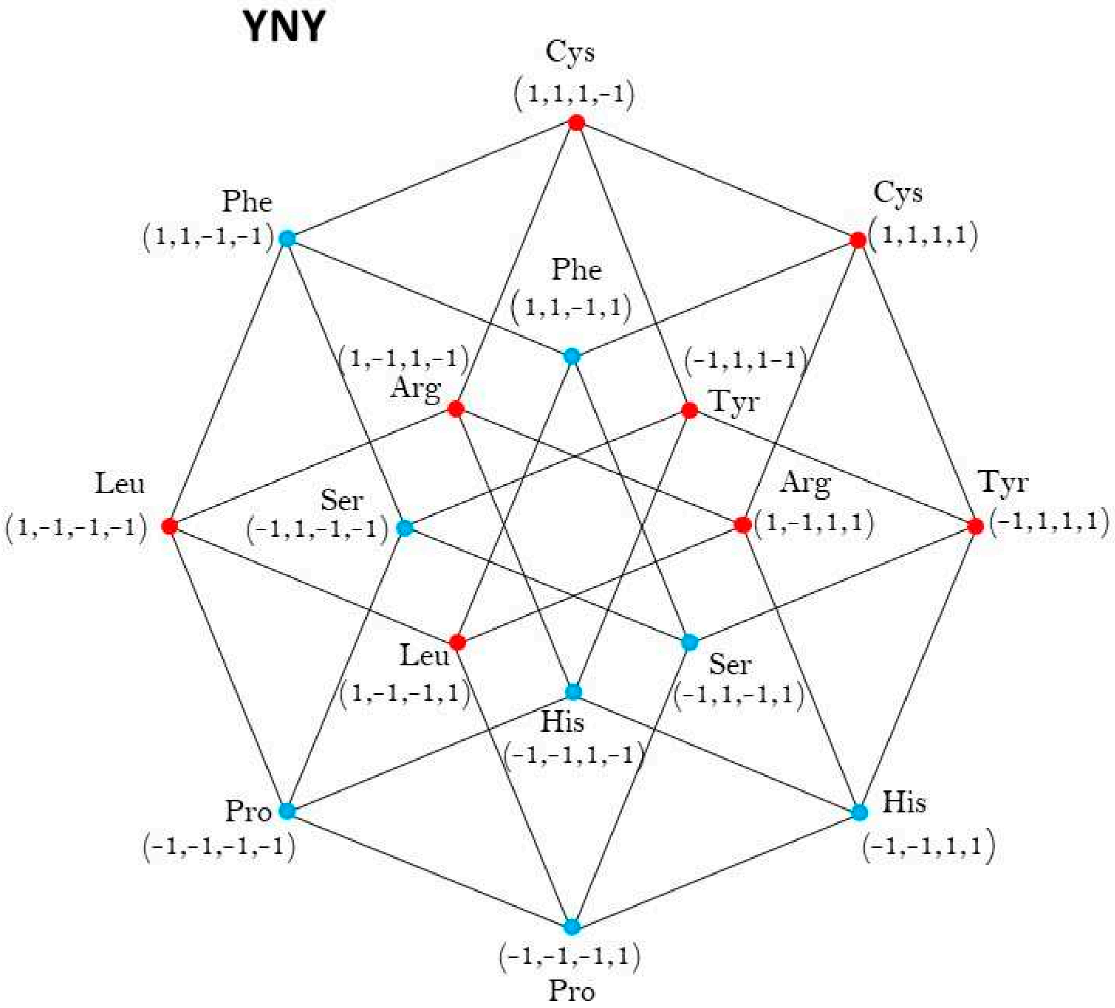


Figure 2. aaRSs distribution in the YNY code. Class 1 (red, 4 amino acids) and Class 2 (blue, 4 amino acids). Graphic representation of the subsets YYY and YRY. Second 4-dimensional hypercube of the Extended RNA-code type 2: $\mathrm{YNY} = \mathrm{YYY} \bigcup \mathrm{YRY}$..

The set of vertexes of the Class 1, colored in red, is the set:
$V_1 = \{(1,1,-1,-1),(1,1,-1,1),(-1,1,-1,-1),(-1,1,-1,1),(-1,-1,-1,-1),(-1,-1,-1,1),(-1,-1,1,-1),(-1,-1,1,1)\}.$

The members of Class 2 are in the complementary set of vertexes, that is, the other eight vertexes, colored in black, out of the eight of Class 1.

This is the set:

$$V_2 = \{(1,-1,-1,-1),(1,-1,-1,1),(1,-1,1,-1),(1,-1,1,1),(1,1,1,-1),(1,1,1,1),(-1,1,1,-1),(-1,1,1,1)\}.$$

Let us consider the affine transformation: $T_{(1/2,\ 1/2,\ 1/2,\ ?} \circ (\text{絀})$, that converts each -1 into 0 and 1 into itself. It represents a change of coordinates, that converts the hypercube $[-1,1]^4$ into its isometric $[0,1]^4$, whose center is the point (1/2, 1/2, 1/2, 1/2). The sets V_1 and V_2 are now:

$$V_1 = \{(1,1,0,0),(1,1,0,1),(0,1,0,0),(0,1,0,1),(0,0,0,0),(0,0,0,1),(0,0,1,0),(0,0,1,1)\}.$$

$$V_2 = \{(1,0,0,0),(1,0,0,1),(1,0,1,0),(1,0,1,1),(1,1,1,0),(1,1,1,1),(0,1,1,0),(0,1,1,1)\}.$$

Let us consider the orthogonal matrix $A=\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, which represents a rotation of angle $\pi = 180^\circ$ around the vectorial plane generated by the vectors e_4 and $e_1 + e_3 + e_4 = (1,0,1,1)$, which fixes the point e_4 changes the sign the vector e_2 and interchanges the vectors e_1 and e_3.

Let us now see what the action of the rotation A over the set V_1 is:

$$A: v_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \rightarrow v_7 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

$$A: v_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \rightarrow v_8 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

$$A: v_3 = \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \rightarrow v_5 = \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$$

$$A: v_4 = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \rightarrow v_6 = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

$$A: v_5 = \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} \rightarrow v_3 = \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

$$A: v_6 = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \rightarrow v_4 = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

$$A : v_7 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \to v_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

$$A : v_8 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \to v_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

It is seen that the rotation A is a symmetry of the set V_1.

Next, we will see the action of A over the set of vertexes V_2:

$$A : u_1 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} \to u_7 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

$$A : u_2 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \to u_8 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$A : u_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \to u_5 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

$$A : u_4 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \to u_6 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$A : u_5 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \to u_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

$$A : u_6 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \to u_4 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

$$\mathrm{A} : \mathrm{u}_7 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \rightarrow \mathrm{u}_1 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$$

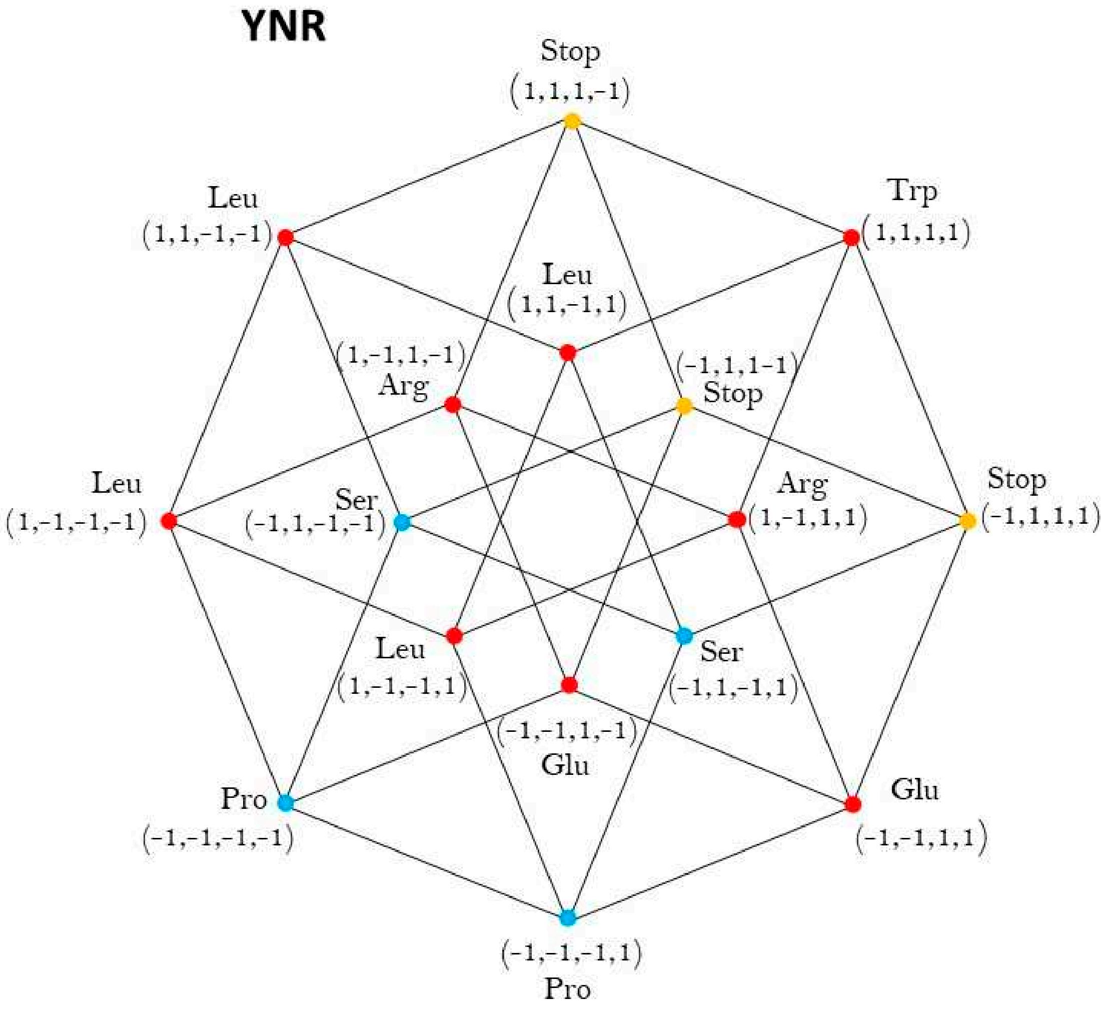
$$\mathrm{A} : \mathrm{u}_8 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \rightarrow \mathrm{u}_2 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

Notice that rotation A is a symmetry of the set V_2.

The rotation A of angle $\pi = 180^{\circ}$ is an element of order two. Then, the group of symmetries of the classes of synthetases in the hypercube YNY contains the group generated by the rotation A, which is isomorphic to the binary group $(\mathbb{Z}_2, \oplus_2)$.

The isometric transformation A is the only symmetry of the sets V_1 and V_2 in the hypercube YNY. Then, the group of symmetries of both classes of synthetases is the binary group $(\mathbb{Z}_2, \oplus_2)$.

3.4. Group of symmetries of the synthetases in the hypercube YNR



doi:10.20944/preprints202308.1951.v1

Figure 3. aaRSs distribution in the YNR code. Class 1 (red, 3 amino acids) and Class 2 (blue, 3 amino acids and yellow, the stop signals). Graphical representation of the subsets YYR and $\mathrm{YRR}.$ Third 4-dimensional hypercube of the Extended $\mathrm{RNA}-$ code type II: $\mathrm{YNR}=\mathrm{YYR}\bigcup \mathrm{YRR}.$.

For methodological reasons we have assumed the three stop codons as members of Class 1. In this case the members of each Class are in the eight vertexes of a 3-dimensional cube. Both Classes coincide with the cubes YNU and YNG. The vertexes of YNR are represented by the sixteen quadruples of the set $\{-1,1\}^4$ of ones and minus ones, being the hypercube $[-1,1]^4$, the convex closure of the set $\{-1,1\}^4$, The set $[-1,1]^4$ is taken, by a suitable change of coordinates, rotation, and translations of axis, as a representation of the hypercube YNR. The center of this hypercube is the origin of coordinates $\mathrm{O}=(0,0,0,0)$.

The set of vertexes of the cube, associated to Class I, colored in red, is the set:

$$\mathrm{V}_1=\{(-1,-1,-1,-1),(-1,-1,-1,1),(-1,1,-1,1),(-1,1,-1,-1),(1,-1,-1,-1),(1,-1,-1,1),(1,1,-1,1),(1,1,-1,-1)\}.$$

The center of this cube is the point $-\mathrm{e}_1=(0,0,-1,0),$ and the set of vertexes of the Class 2, colored in black, is the set:

$$\mathrm{V}_2=\{(-1,-1,1,-1),(-1,-1,1,1),(-1,1,1,1),(-1,1,1,-1),(1,-1,1,-1),(1,-1,1,1),(1,1,1,1),(1,1,1,-1)\}.$$

whose center is the point $\mathrm{e}_3=(0,0,1,0)$.

Then, we have proved that the group of symmetries of both classes of synthetases in the hypercube YNR is the extended octahedral group $(\mathrm{O}_h,\circ)$.

Let us consider the translation $\mathrm{T}_{\mathrm{e}3}$ associated to the point $\mathrm{e}_3=(0,0,1,0)$. Applying it to the set V_1 we obtain:

$$\mathrm{T}_{\mathrm{e}3}(\mathrm{V}_1)=\mathrm{U}_1=\{(-1,-1,0,-1),(-1,-1,0,1),(-1,1,0,1),(-1,1,0,-1),(1,-1,0,-1),(1,-1,0,1),(1,1,0,1),(1,1,0,-1)\},$$

which is the set of vertexes of another cube, with center in the origin $\mathrm{O}=(0,0,0,0)$ of coordinates.

Analogously, applying to the set V_1 the translation $\mathrm{T}_{-\mathrm{e}_3}$ we obtain:

$$\mathrm{T}_{-\mathrm{e}_3}(\mathrm{V}_2)=\mathrm{U}_2=\{(-1,-1,0,-1),(-1,-1,0,1),(-1,1,0,1)(-1,1,0,-1),(1,-1,0,-1),(1,-1,0,1),(1,1,0,1),(1,1,0,1)\},$$

which is the same set U_1, of vertexes of a cube with center in the origin $\mathrm{O}=(0,0,0,0)$ of coordinates.

Let us now consider the orthogonal matrices:

$$\mathrm{A}=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix},\quad \mathrm{B}=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}$$

$$\mathrm{C}=\begin{pmatrix}1&0&0&0\\0&0&0&1\\0&1&0&0\\0&0&1&0\end{pmatrix},\quad \mathrm{D}=\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&-1\end{pmatrix}$$

which represent rotations, A, B and D of angles π radians and C of angle $\frac{2\pi}{3}$ radians. They satisfy the relations: $\mathrm{A}^2 = \mathrm{B}^2 = \mathrm{C}^3 = \mathrm{D}^2 = \mathrm{I}$,

$$\mathrm{BA} = \mathrm{AB},\ \mathrm{CA} = \mathrm{BC},\ \mathrm{CB} = \mathrm{ABC},\ \mathrm{DA} = \mathrm{BD},\ \mathrm{DB} = \mathrm{AD},\ \mathrm{DC} = \mathrm{AC}^2\mathrm{D}.$$

that are the defining relations of the octahedral group (O, o), which is the group of the rotational symmetries of a cube or regular hexahedron. It is also the group of rotational symmetries of a regular octahedron, which is the dual figure of the regular hexahedron. Adding to the set of rotational matrixes $\mathrm{A}, \mathrm{B}, \mathrm{C}$ and D, the matrix

$$\mathrm{R} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

which represents a reflection through the three-dimensional subspace XYZ, generated by the unitary vectors $\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3$, that reverses the fourth axis □ e_4, the amplified octahedral group $(\mathrm{O_h}, \mathrm{o})$ is obtained, where $\mathrm{O_h} = \mathrm{O} \bigcup (\mathrm{O} \circ \mathrm{R})$.

We invite the lector to verify that these rotations, and the reflection R, are symmetries of the sets V_1 and V_2 of the vertexes of the Class 1 and the Class 2, respectively.

From this it follows that the group of symmetries of classes of synthetases in the hypercube YNR is the extended octahedral group $(\mathrm{O_h}, \mathrm{o})$.

3.5. Group of symmetries of the synthetases in the hypercube RNR

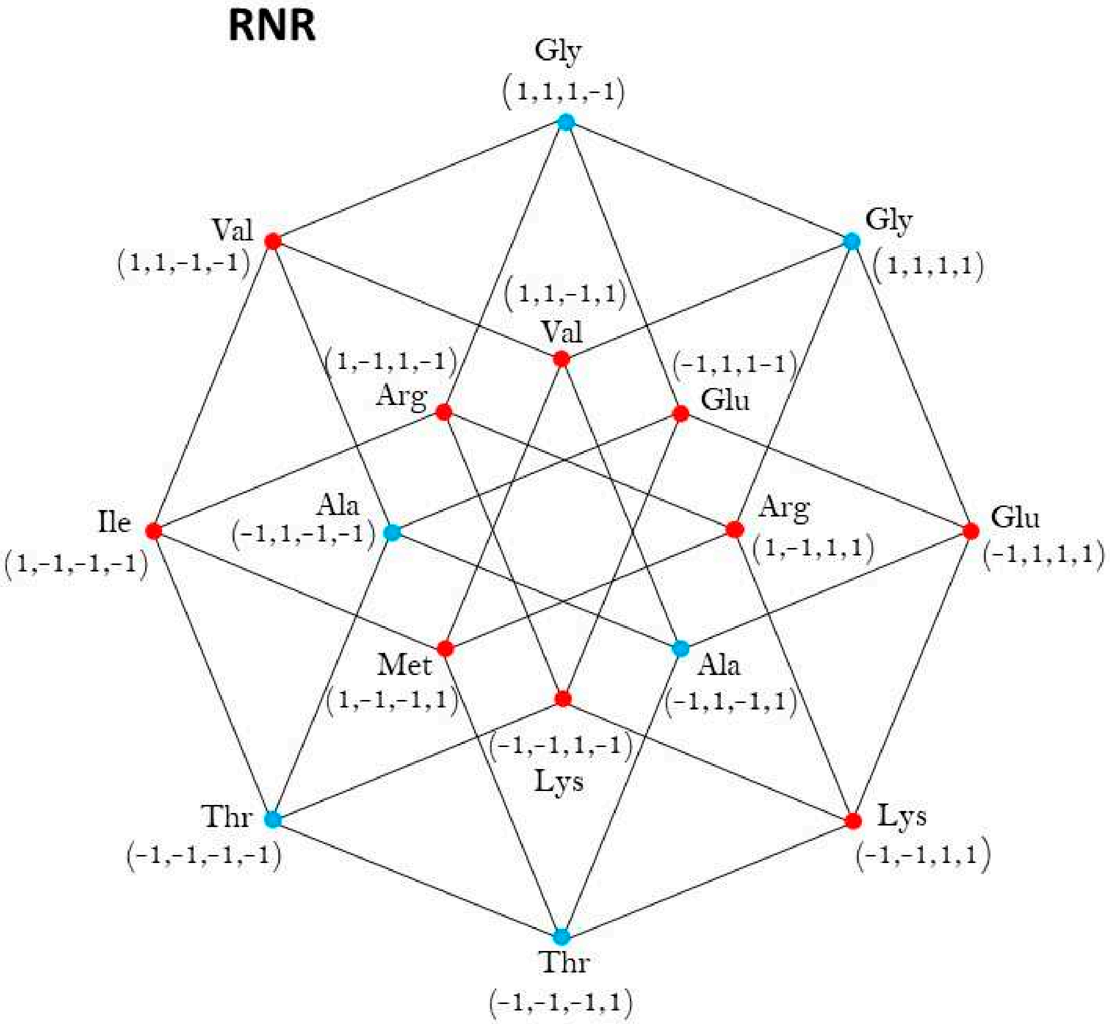


Figure 4. aaRSs distribution in the RNR code. Class I (red, 3 amino acids) and Class 2 (blue, 5 amino acids). Graphic representation of the subsets RYR and RRR. Fourth 4-dimensional hypercube of the Extended RNA-code type II: $RNR = RYR \bigcup RRR$..

In this case the vertexes of Class 1, colored in red, are the set of six vectors:

$V_1 = \{v_1 = (-1,-1,-1,-1),\ v_2 = (-1,-1,-1,1),\ v_3 = (-1,1,-1,1),\ v_4 = (-1,1,-1,-1),$ $v_5 = (,1,1,-1),\ v_6 = (1,1,1,1)\}$. which is the union of the square of vertexes V_1, V_2, V_3, V_4 with the edge of vertexes V_5, V_6.

And the set of vertexes of Class 2, colored in black, are the set of the ten vertexes:

$$V_2 = \left\{ \begin{array}{l} u_1 = (1,-1,-1,-1),\ u_2 = (1,-1,-1,1),\ u_3 = (1,1,-1,1),\ u_4 = (1,1,-1,-1), \\ u_5 = (-1,-1,1,-1),\ u_6 = (-1,-1,1,1),\ u_7 = (-1,1,1.1),\ u_8 = (-1,1,1,-1), \\ u_9 = (1,-1,1,-1),\ u_{10} = (1,-1,1,1) \end{array} \right\}$$

which is he union of the squares of vertexes u_1, u_2, u_3, u_4 and u_5, u_6, u_7, u_8 with the edge of vertexes u_9, u_{10}.

Now, consider the matrix $R=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$, which represents the same reflection of **Part** III, through the hyperplane XYZ, that reverses the axis □ e_4. This reflection changes the sign of the fourth component of each quadruple and it easy to notice that it produces the following interchange of vectors in the Class I of synthetases:

$$v_1 \leftrightarrow v_2, v_3 \leftrightarrow v_4, v_5 \leftrightarrow v_6$$

Hence, R is a symmetry of the Class I of synthetases in the hypercube RNR.

By Theorem 1 it should also be a symmetry of the set V_2 of vertexes of Class 2. But it is straightforward to verify that it produces the interchange of vectors:

$$u_1 \leftrightarrow u_2, u_3 \leftrightarrow u_4, u_5 \leftrightarrow u_6, u_7 \leftrightarrow u_8, u_9 \leftrightarrow u_{10}.$$

Then, R is also a symmetry of Class 2.

It is not difficult to notice that the reflection R is the only symmetry of both Classes 1 and 2. As it is an element of order two, the group $(\{R, I\} \circ)$, isomorphic to the Abelian group (儈$_2$,$\oplus_2$), is, as in the case of Part II, the group of symmetries of the classes of synthetases in the hypercube RNR.

3.6. Conclusions

It has been proved that the groups of symmetries of both classes of synthetases in each of the four hypercubes are the following (**Table 5)**:

Table 5. Summary of symmetry groups of aaRSs for each subcode.

Hypercube	**Group of symmetries**
RNY	D_4
YNY	□ $_2$
YNR	O_h
RNR	□ $_2$

3.7. The Genetic 5-dimensional hypercube

The 5-dimensional hypercube NNY is the disjoint union of the 4-dimensional hypercubes YNY and RNY. YNY is the subspace generated by the four unitary vectors UCC, CAC, CUC, CCU and RNY is its affine subspace $YNY + \{ACC\}$. Here the sextuples of zeros and ones have been replaced by triplets of the letters C, U, A, G, with the following equivalence: $C = 00, U = 01, A = 10, G = 11$. Y is the binary set $\{C, U\}$, R is the also binary $\{A, G\}$, and N is the set $\{C, U, A, G\} = \{Y, R\}$.

The four triplets UCC, CAC, CUC, CCU are associated to the canonical unitary vectors $e_2, e_3, e_4 . e_6$. In **Figure 5** they are vectors initiated in the point CCC, colored in red, that generate the hypercube YNY. The triplet ACC, associated to the unitary vector e_1, also initiated in

CCC, and colored in blue, completes the canonical basis of the 5-dimensional hypercube NNY. Obviously, the translation associated to convert YNY into its affine subspace RNY, colored in blue, and the union of both YNY with RNY, is the 5-dimensional hypercube NNY. The vertexes of NNY are the triplets or codons that end in a pyrimidine.

It can be seen that the translation T_e, associated to the unitary vector e_5, represented by the triplet CCA, converts the 5-dimensional subspace NNY into its 5-dimensional affine subspace NNR, whose triplets are those that end in a purine (**Figure A1**). It is the disjoint union of the 4-dimensional hypercubes YNR and RNR. It is clear that the 6-dimensional hypercube NNN is the union of the 5-dimensional hypercube NNY with its affine subspace, also a 5-dimensional hypercube, $\mathrm{NNR} = \mathrm{NNY} + \{\mathrm{CCA}\}$. Finally, we have, that NNN is the disjoint union $\mathrm{RNY} \bigcup \mathrm{YNY} \bigcup \mathrm{YNR} \bigcup \mathrm{RNR}$ of the four 4-dimensional hypercubes, derived from the primaeval RNY by transversions in the first or the third nucleotide.

Note: The subspace (hypercube) YNY, is the solution set of the homogeneous system of two liner equations $x_1 = 0, x_5 = 0,$ with the six unknowns $x_1, x_2, x_3, x_4, x_5 . x_6$. Its affine subspace (hypercube) RNY, is the solution set of the non-homogeneous system $x_1 = 0, x_5 = 0,$ also of rank 2, with the same six unknowns.

The 5D hypercube NNR (see **Figure A1**) is obtained by the translation T_{CCA}, that converts every triplet that ends in pyrimidine into other that ends in purine, obtaining the hypercube NNR. It represents the disjoint union of the subcodes YNR and RNR. Note that this hypercube contains the stop codons $\mathrm{UGA}, \mathrm{UAG}$ and UAA. The disjoint union of NNY and NNR is equal to the 6D hypercube NNN that represents the entire SGC. For the 6D genetic code each triplet is mapped into sextuples of o´s and 1's. This hypercube is the set NNN isomorphic to $\square_2^6$, which is the binary vector space over the binary field $\square_2$,

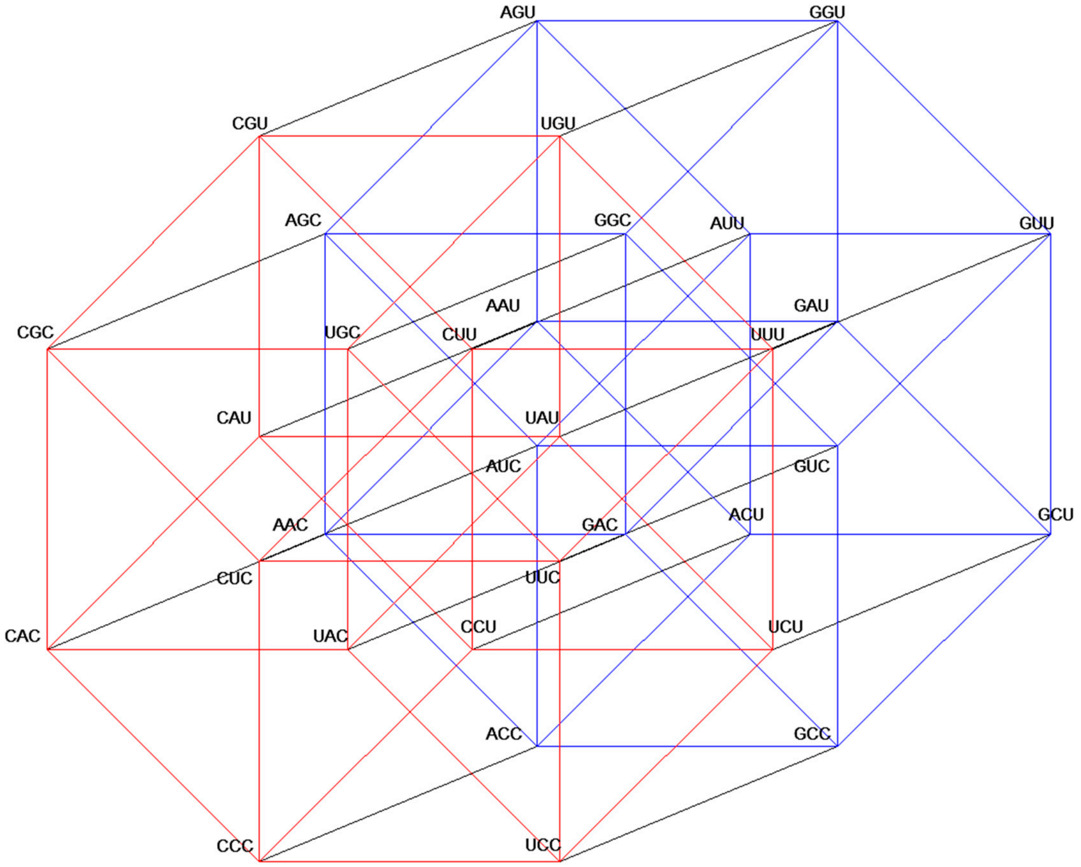


Figure 5. The five-dimensional hypercube of triplets NNY, which is the union of the 4-dimensional hypercubes RNY and YNY. YNY is obtained from the RNY by means of transversion of each nucleotide at the left of the triplet. YNY is the subspace generated by the four unitary vectors CAC, CUC, CCU, UCC, and RNY is its affine subspace $YNY+\{ACC\}$. The translation that represents these transversions is situated at an angle of 22.5 degrees with respect to the horizontal axis generated by the canonical vector UCC. At this angle none of the vertexes are overlapped.

The edges of the hypercube

Two vertexes are said to be neighbors if the distance between them is equal to 1. It means that they differ in only one of their n components. An edge of the hypercube is the segment that joins two neighbor vertexes. This segment has a length equal to 1. For every vertex there are n neighbors. As there are 2^n vertexes, the product $2^n \times n$ would count the total number of edges. But as every edge has two extremes, it would be counted twice. Then, the actual number of edges is equal to $\frac{2^n \times n}{2} = 2^{n-1} \times n$. For example, in the ordinary cube, 3-dimensional, the number of edges is $2^2 \times 3 = 12$ and, in the 4-dimensional, this number is $2^3 \times 4 = 32$.

The faces of the hypercube

Any face of the hypercube is a plane square determined by four vertexes and four edges, in such a way that two edges with a common vertex are orthogonal. If we fix n-2 components with values zeros and ones, and leave free the other two, the four obtained vertexes are the vertexes of a plane square, then, of a face of the hypercube. It may be done in $\binom{n}{n-2} \times 2^{n-2}$ ways. Hence, the total

doi:10.20944/preprints202308.1951.v1

number of different faces is precisely $\binom{n}{n-2}\times 2^{n-2}$. (The combinatorial number $\binom{n}{k}=\frac{n!}{k!(n-k)!}$ is the number of subsets of k elements of a set of n elements, being $0\le k\le n$). For example, for n=3 the number of faces is $3\times 2 = 6$, and for n=4 it is 6x4=24.

The hypercube of dimension 6 NNN.

It is the disjoint union of the 5-dimensional NNY and NNR. NNY is the subspace generated by the five unitary vectors CAC, CUC, CCU, UCC, ACC and NNR is its affine subspace $\mathrm{NNY}+\{\mathrm{CCA}\}$. The vector space NNN is generated by the six unitary vectors CAC, CUC, CCU, UCC, ACC, CCA. They conform the canonical basis of $\square_2^6=\mathrm{NNN}$. Conversion of the hypercube $[0,1]^n$ in other whose components are ones and minus ones. Multiplying by 2 all the vectors of the space, the set of the 2^n vertexes is converted into the set $\{0,2\}^n$, of vertexes of the hypercube $[0,2]^n$ whose edges have length 2. Next, performing the substraction of the n-tuple $E = (1,1,\ldots 1)$ the hypercube $[0,2]^n$ is converted into $[-1,1]^n$, which is another hypercube, also with edges of length 2, whose set of vertexes is the set $\{-1,1\}^n$ of the 2^n n-tuples of ones and minus ones.

Note: The n-tuple E is the opposite vertex, in the hypercube $[0,1]^n$, of the null vector $O=(0,0,\ldots 0)$. The distance between them is the real number $\sqrt{n}=|E|$, which is the diameter of that hypercube. The diameter of a hypercube is the diameter of the hypersphere that circumscribes it-

Actually, the hypercube $[-1,1]^n$ is the image of $[0,1]^n$ under the affine transformation $T_{-E}\circ 2I$ being I the identical automorphism. The new hypercube has its center in the origin $O=(0,0,\ldots 0)=(T_{-E}\circ 2I)\left(\frac{1}{2},\frac{1}{2},\ldots\frac{1}{2}\right)$ of coordinates, and its radius is $\sqrt{n}$. The affine transformation $T_{-E}\circ 2I$ converts every 0 into -1, and the number 1 into itself. The linear transformation $2I$ is a homothetic transformation of ratio 2, that duplicate the size of every n-dimensional Figure. Its only fixed point is the origin O of coordinates. The affine transformation $T_{-E}\circ 2I$ is also a homothety of ratio 2, whose only fixed point is the point E.

A metric in the hypercube $\square_2^6$.

The Hamming Norm and the Hamming Distance defined in $\square_2^6$.

The Hamming Norm: The Hamming Norm $|u|$ of a sextuple $u=(a_1,a_2,a_3,a_4,a_5,a_6)$ is defined as the sum $\sum a_i$, in $\square$, of its components a_i. It is equal to the number of ones it has. The only unitary sextuples, that is, of norm equal to 1, are those of the canonical basis.

The Hamming distance. The Hamming distance $H(u,v)$ between two sextuples $u=(a_1,a_2,a_3,a_4,a_5,a_6)$ and $v=(b_1,b_2,b_3,b_4,b_5,b_6)$ is defined as the norm $|u\oplus_2 v|$ of their module 2 addition. It is equal to the number of places where they differ. It is not difficult to prove that the distance so defined is actually a distance, in the mathematical sense. It means that it has the following properties:

D1) For all u,v, $H(u,v)\ge 0$, $H(u,v)=0\Leftrightarrow u=v$ (positivity of the distance between two different sextuples u and v), and nullity of the distance of a sextuple to itself.

D2) For all u,v, $H(u,v)=H(v,u)$ (Symmetry)

D3) For u,v,w, $H(u,w)\le H(u,v)+H(v,w)$ (Triangle inequality)

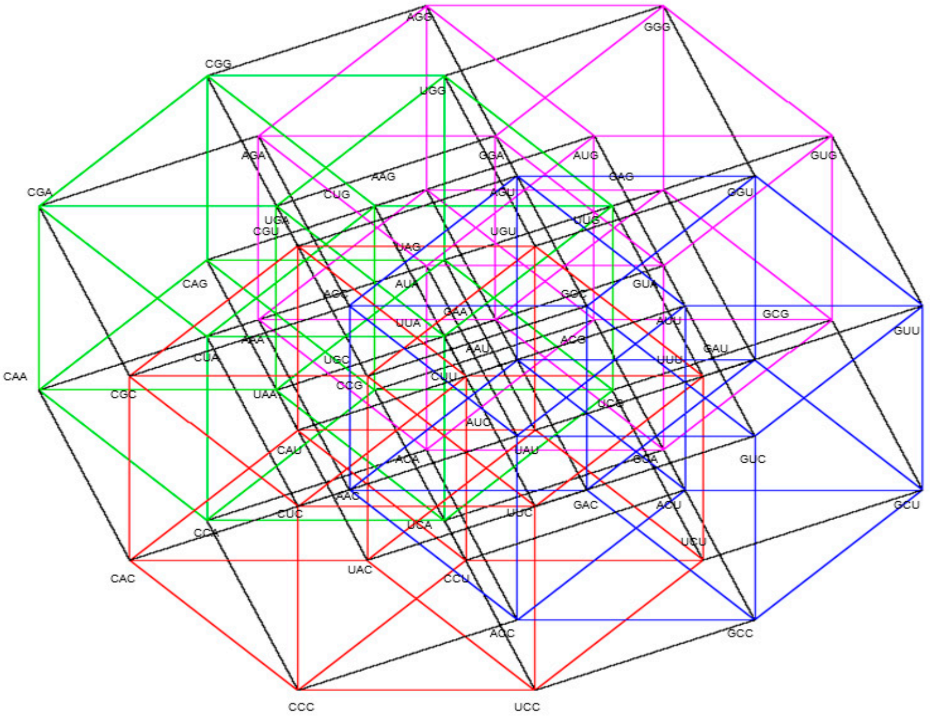


Figure 6. The 6D representation of the Standard Genetic Code.

Discussion

We have determined the type of symmetries of aaRSs in each of the 4 subcodes of the Extended RNA code type II. We used the 4D representation of each subcode. The RNY subcode has the dihedral group, with 8 symmetries, 4 rotations and 4 reflections. The YNY subcode has the $\square_2$ symmetry, with 2 symmetries; The YNR subcode exhibits an octahedral amplified symmetry, with 48 elements, with 24 rotations and 24 roto-reflections (the octahedral classic group contains only 24 rotations)**;** The RNR subcode display the symmetry of the binary group $\square_2$. In each subcode, symmetrical distributions of both Classes of aaRSs were found. Indeed, we proved that for each hypercube, the group of symmetries of Class 1 is the same as the group of symmetries of Class 2. This theorem holds for both 5D- hypercubes and for the whole 6D representation of the SGC.

The consideration of stop codons in the determination of the type of symmetry of the subcode is not fortuitous. Mitochondrial codes present variations principally in the codons for the stop signals and unicoded amino acids. The mitochondrial genetic codes of yeast, invertebrates, and vertebrates present variations principally in the codons for the stop signals and unicoded amino acids https://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi?chapter=tgencodes#SG24 (accessed on July 31 2023).The stop codons are tricodonic in SGC, tetracodonic in vertebrate mitochondria, and dicodonic in invertebrate and yeast mitochondria.

In computer science, a byte is an octet of 0′s and 1′s, where each bit represents 0 or 1. Hence, in the genetic code a byte would correspond to a sextuple of 0′s and 1′s, where each of them represents a triplet or codon of the nucleotides $\{C, U, A, G\}$. The presence of stop codons converts the genetic code into an algorithm that carries out the protein synthesis. This means that the whole process of protein synthesis is carried out by a Turing machine, i.e., by a recursive function. Unlike the Turing machine the genetic code has the additional property of heritability. In a forthcoming work we will develop these concepts.

Our work can be regarded as a possible pathway of the distribution of the 2 Classes of aRSs during the formation of the SGC. According to the model of Rodin-Ohno {15 -18], there was a single gene encoding for 2 synthetases: They proposed a single anti-parallel complementary strands of a single base-paired nucleic acid molecule. The Rodin–Ohno model divides the table of the genetic code into two classes of aminoacyl-tRNA synthetases (Classes 1 and 2) with recognition from the minor or major groove sides of the tRNA acceptor stem [15–18]. According to the table of the genetic code, the RO model is almost symmetric. It turns out that the RO model is symmetric in a six-dimensional (6D) hypercube (José et al. 2017). Conversely, using the same automorphisms, the RO-model can lead to the SGC. Class aaRS 1 (Class aaRS 2) can be converted into Class aaRS II (Class aaRS I) by means of isometric algebraic functions [19]. Notably, the 6D algebraic model is compatible with both the SGC (based upon the primeval RNY code) and the RO model [19].

Author Contributions: Conceptualization, MVJ and ERM; methodology, MVJ and ERM; formal analysis, ERM and MVJ.; investigation, MVJ and ERM; writing—original draft preparation, ERM and MVJ; writing—review and editing, MVJ and ERM; visualization, JRB.; Figures JRB; funding acquisition, MVJ. All authors have read and agreed to the published version of the manuscript.

Funding: M.V.J. was funded by Dirección General de Asuntos del Personal Académico (DGAPA), Universidad Nacional Autónoma de México, UNAM (PAPIIT-IN224015).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Mathematical background

The Euclidean n-dimensional $\mathbb{R}$ – space, $(\mathbb{R}^n,+,*)$ being $(\mathbb{R},+,x)$, the field of real numbers, has a basis $(e_1,e_2,\ldots,e_n)$, the so-called canonical basis, where, for each $i\in\{1,2,\ldots n\}$, $e_i=(0,0,\ldots 1,\ldots,0)$, with 1 in the place i, and 0 in the remainders.

A remarkable observation: The n-dimensional vector space $(\mathbb{R}^n,+,*)$ is a generalization of the ordinary 3-dimensional $\mathbb{R}$ – space, $(\mathbb{R}^3,+,*)$, where every triplet (x_1,x_2,x_3) represents a point P of the space, endowed with a system of axis X, Y and Z, pairwise orthogonal or perpendicular, intercepted in a common point denoted as O and called the origin of coordinates, represented by the triplet (0,0,0). Usually, every point P is identified with the triplet (x_1,x_2,x_3) that represents it, as they were the same object. So, the equality $P=(x_1,x_2,x_3)$ is admitted. The axis are right lines defined as $X=\mathbb{R}e_1=\{(x_1,0,0|x_1\in\mathbb{R}\}$, $Y=\mathbb{R}e_2=\{(0,x_2,0|x_2\in\mathbb{R}\}$, $Z=\mathbb{R}e_3=\{(0,0,x_3|x_3\in\mathbb{R}\}$. The addition of two points $P=(x_1,x_2,x_3)$ and $Q=(y_1,y_2,y_3)$ is defined as $P+Q=(x_1+y_1,x_2+y_2,x_3+y_3)$. This operation, so defined, is associative, has a neutral element, O, and for every point $P=(x_1,x_2,x_3)$ its additive inverse $-P=(-x_1,-x_2,-x_3)$ exists, such that $P+(-P)=O$. Then, the ordered pair $(\mathbb{R}^3,+)$ is a group. As the addition + is also commutative, it is an Abelian group. The product of a real number $\alpha\in\mathbb{R}$ by a point $P=(x_1,x_2,x_3)$ is defined as $\alpha*P=(\alpha x_1,\alpha x_2,\alpha x_3)$. This external operation, of a real number by a point, has the following algebraic properties:

1. $\alpha*(\beta*P)=(\alpha\times\beta)*P$ for $\alpha,\beta\in\mathbb{R},P\in\mathbb{R}^3$ (Mixed associativity).
2. $1*P=P$, for every $P\in\mathbb{R}^3$ (Existence of an external neutral element).
3. $\alpha*(P+Q)=\alpha*P+\alpha*Q$, for all $\alpha\in\mathbb{R},P,Q\in\mathbb{R}^3$ (Distributivity of the product with respect to the addition of points, also called distributivity at the left).
4. $(\alpha+\beta)*P=\alpha*P+\beta*P$ for all $\alpha,\beta\in\mathbb{R},P\in\mathbb{R}^3$ (Distributivity of the product with respect to the addition of numbers, also called distributivity at the right.)

The condition of Abelian group for the addition and the four properties of the external operation, completes the structure of a vector space over the field $(\mathbb{R},+,\times)$ of the real numbers, in the set $\mathbb{R}^3$ of all the points of the space.

The concept of vector: Given two different points $P,Q\in\mathbb{R}^3$ we call the vector of origin P and extreme Q to the oriented segment of right line that joins P with Q. It is denoted as $\overrightarrow{PQ}$. We define the addition of two vectors $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ with common origin P as the vector $\overrightarrow{PS}$ such that $S=R+Q-P$. If the points R,Q and P are not collinear the four points are the vertexes of a plane parallelogram. It means that the lines $r(P,Q)$ and $r(R,S)$ are parallel. If $P=O$, the null point $\{0,0,0\}$, the function: $f:P\rightarrow\overrightarrow{PO}$, where for $P=O$, $\overrightarrow{OO}=0$ is defined as a vector whose initial and final points coincide, is a bijective function between the set $\mathbb{R}^3$ of all the triplets or points of the space, and the set V of all the vectors with origin O. This correspondence induces over the set V the operations $+$ and $*$ of addition and product of numbers by vectors and V is endowed the structure of a vector space over the field $\{\mathbb{R},+,\times\}$ of the real numbers.

Euclidean norm

The Euclidean norm, module, or length, of a vector $v=(x_1,x_2,\dots x_n)$ is defined as $|v|=\sqrt{x_1{}^2+x_2{}^2+\dots x_n{}^2}$, the non-negative square root of the sum of the squares of its n components.

Main properties of the Euclidean norm

N1) For every v, $|v|\geq 0$, and it is =0 if, and only if, v=o, the null vector $O=(0,0,0\dots0)$, (Positivity of the norms for all the non-null vectors)

N2) For $v\in\mathbb{R}^n$ and $\alpha\in\mathbb{R}$, $|\alpha*v|=|\alpha||v|$. (The norm $|\alpha|$, module, or absolute value, of the real number $\alpha\neq0$ is the positive of the couple $\{\alpha,-\alpha\}$, and, for $\alpha=0$, $|\alpha|=0$.

From N2 we have that $|-v|=|v|$ for every $v\in\mathbb{R}^n$.

N3) For all $u,v\in\mathbb{R}^n$, $|u+v|\leq|u|+|v|$ (Triangle inequality).

The Euclidean Distance between vectors

For $u,v\in\mathbb{R}^n$, the Euclidean distance $d(u,v)$ between their extremes, is defined as $|u\text{-}v|$.

Main properties of the Euclidean distance

D1) For every v, $d(v,O)=|v|\geq0$, being O the null vector $O=(0,0\dots0)$, (Positivity of the distance to the origin for all the final points of the non-null vectors).

D2) For all $u,v\in\mathbb{R}^n$, $d(u,v)=d(v,u)$ (Symmetry).

D3) For all $u,v\in\mathbb{R}^n$, $d(u,w)\leq d(u,v)+d(v,w)$ (Triangle inequality).

With this distance the set $\mathbb{R}^n$, of vectors, is a metric space and the vector space $(\mathbb{R}^n,+,*)$ is called the n-dimensional Euclidean Vector Space.

Linear transformations

A function $f:\mathbb{R}^n\rightarrow\mathbb{R}^n$ is called a linear transformation, or linear endomorphism, if it has the following property:

For all $u,v\in\mathbb{R}^n$ and $\alpha,\beta\in\mathbb{R}$, the equality $f(\alpha*u+\beta*v)=\alpha*f(u)+\beta*f(v)$ holds.

Note: The latter condition is equivalent to the two conditions:

L1) $f(u+v)=f(u)+f(v)$, for all $u,v\in\mathbb{R}^n$. (This means that f is a group homomorphism of the additive group $(\mathbb{R}^n,+)$ onto itself).

L2) $f(\alpha * u)=\alpha * f(u)$ for all $u\in\mathbb{R}^n$ and $\alpha\in\mathbb{R}$.

If the function f is bijective, the linear transformation is a linear isomorphism, and it is called a linear automorphism of the vector space.

Matrix representation of a linear transformation

For every endomorphism f there is a unique square matrix $A=\left(a_{ij}\right)$, such that for every j, $f(e_j)=(a_{1j},a_{2j},\ldots a_{nj})$, being $\begin{pmatrix} a_{1j}\\ a_{2j}\\ ..\\ a_{nj}\end{pmatrix}$ the j-th column of A. Then, A is a matrix whose columns are the column matrixes associated of the vector images of the canonical unitary vectors. It is called the matrix of f with respect to the canonical basis. It will be denoted as $A=[f]$. In the case of an automorphism the matrix A is a non-singular or invertible matrix, hence, its determinant is different from 0.

For every vector $v=(x_1,x_2,\ldots x_n)\in\mathbb{R}^n$ the column matrix $[V]=\begin{bmatrix} x_1\\ x_2\\ .\\ .\\ x_n\end{bmatrix}$ is taken as the matrix representation of V. It is well known that the matrix representation of the image vector $f(V)$ is the matrix product $[f][V]$. Conversely, every square $n\times n$ matrix A defines an endomorphism f, whose matrix representation $[f]$ is A.

Translations

For any vector u the translation defined by u, is the function $T_u:\mathbb{R}^n\to\mathbb{R}^n$, such that, for all v, $T_u(v)=v+u$. For every u, T_u is a bijective function that preserves the norm of each vector and the distance between different vectors of the space. Hence, it is an isometry of the Euclidean vector space. For two translations T_u and T_v the composition $T_u\circ T_v$ is equal to the translations T_{u+v}.

The correspondence $u\leftrightarrow T_u$ between $\mathbb{R}^n$ and the set T, of all the translations, is a bijective function, that induces over T the structure of a n-dimensional vector space, isomorphic to the Euclidean n-space $(\mathbb{R}^n,+,*)$.

The concept of affine transformation

An affine transformation is a function that is a composition $T_u\circ f$ of a linear endomorphism f, followed by a translation T_u. It acts over every vector v as $\left(T_u\circ f\right)\left(v\right)=f(v)+u$. The affine transformation $T_u\circ f$ is bijective if, and only if, its linear part f is an automorphism. The set of all the affine transformations of the vector space $(\mathbb{R}^n,+,*)$. will be denoted as $A_{ff}(\mathbb{R}^n)$. The composition $\left(T_u\circ f\right)\circ\left(T_v\circ g\right)$ of two affine transformations $T_u\circ f$ and $T_v\circ g$ acts over a vector W in the form:

$$w\mapsto(T_u\circ f)(g(w)+v)=f(g(w)+v)+u=(f\circ g)(w)+f(v)+u=(T_{f(v)+u}\circ(f\circ g))(w)$$

Then, the composition is the affine transformation $T_{f(v)+u}\circ(f\circ g)$. Hence, we have the equality $(T_u\circ f)\circ(T_v\circ g)=T_{f(v)+u}\circ(f\circ g)$

We see that the composition of two affine transformations is an affine transformation. Then, the composition is an inner operation in the set $A_{ff}(\mathbb{R}^n)$. As the operation is associative and its neutral element, the identity function I, is the affine transformation $T_o\circ I$, the ordered pair $(A_{ff}(\mathbb{R}^n),\circ)$ is a monoid, or semigroup with neutral element.

Matrix representation

The matrix representation of the affine transformation $T_u\circ f$ is given by the composition $T_{[u]}\circ A$, being A the matrix of f, which acts over the matrix $[v]$ of a vector v as $[v]\to A[v]+[u]$.

The conjugated of a translation by a linear automorphism

The conjugated $f\circ T_V\circ f^{-1}$ acts over any vector w as $w\to\left(f\circ T_v\circ f^{-1}\right)(w)=f\left(f^{-1}(w)+v\right)=w+f(v)=T_{(v)}(w)$. Then $f\circ T_v\circ f^{-1}=T_{f(v)}$. The latter means that the group of translations $(T,\circ)$ is a normal subgroup of the group $\left(A\left(\mathbb{R}^n\right),\circ\right)$ being $A\left(\mathbb{R}^n\right)=T\circ Aut\left(\mathbb{R}^n\right)$ the set of all the invertible affine transformations, and $\left(Aut\left(\mathbb{R}^n\right),\circ\right)$ the group of all the linear automorphisms of the space $(\mathbb{R}^n,+,*)$. The group $(A(\mathbb{R}^n),\circ)$, that is called the affine group of the space $(\mathbb{R}^n,+,*)$, is a subgroup of the group of all the bijective transformation of the set $\mathbb{R}^n$. It contains as a normal subgroup, the group $(T,\circ)$ of all the translations. In fact, it is the semidirect product of the group $(T,\circ)$ with the group $(Aut(\mathbb{R}^n),\circ)$ of all the linear automorphisms. (A group $(G,*)$ is semidirect product of its subgroups $(K,*)$ and $(H,*)$ if $G=K*H$ and $(K,*)$ is a normal subgroup of $(G,*)$). The semidirect product will be denoted as $(K\lhd H,*)$. Then, $(A(\mathbb{R}^n),\circ)=(T\lhd Aut(\mathbb{R}^n),\circ)$ The affine group $(A(\mathbb{R}^n),\circ)$ is the group of invertible elements of the monoid $(A_{ff}(\mathbb{R}^n),\circ)$.

The n-dimensional binary hypercube.

The binary vector space $\left(\mathbb{Z}_2^n,\oplus_2,*\right)$, vector space over the binary field $(\mathbb{Z}_2,\oplus_2,\otimes_2)$, is also called the n-dimensional binary hypercube. Its set of vertexes coincides with the set $\{0,1\}^n$ of the 2^n n-tuples of zeros and ones. This set is a subset of the set $\mathbb{Z}^n$ of all the n-tuples of the whole numbers. The set $\mathbb{Z}^n$ determines the Abelian group $(\mathbb{Z}^n,+)$, subgroup of the additive group $(\mathbb{R}^n,+)$ of the n-dimensional Euclidean n-space $(\mathbb{R}^n,+,*)$.

The n-dimensional hypercube, whose vertexes are the elements of the set $\{0,1\}^n$ is defined as the convex closure $[0,1]^n$ of the set $\{0,1\}^n$. Here, $[0,1]$ denotes the closed interval of the real numbers between 0 and 1. The convex closure of a set is the minimal convex set that contains it.

A convex set is a set $C\subseteq\mathbb{R}^n$ that, for two different elements a and b of C, the segment that joins them is entirely contained in C.

The center of the hypercube $[0,1]^n$ is the point $C=\left\{\frac{1}{2}.\frac{1}{2},\ldots,\frac{1}{2}\right\}$ obtained by the sum of the 2^n vertexes divided by 2^n. That is so because, for every place i of the n-tuples, 2^{n-1} of them, that is, the half, have 1, and the other 2^{n-1} have 0. Then, for each place i, the quotient is $\frac{2^{n-1}}{2^n}=\frac{1}{2}$.

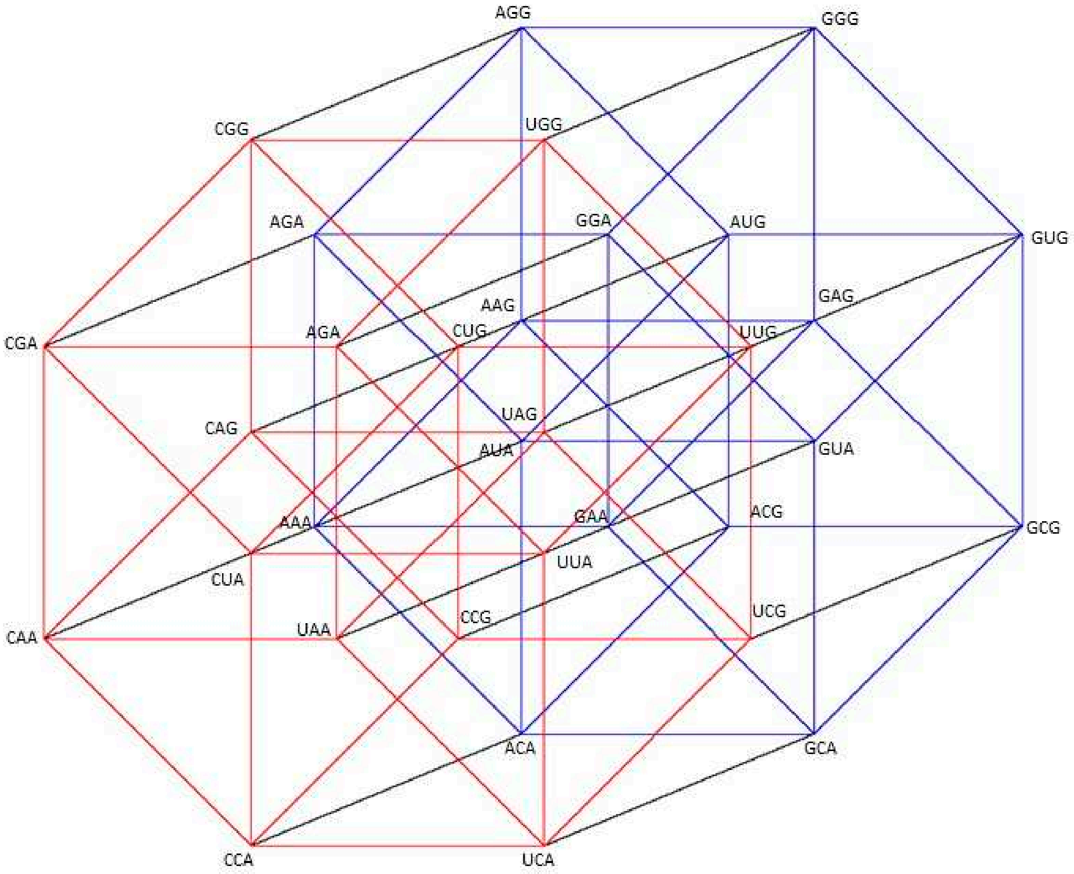


Figure A1. Five-dimensional hypercube NNR **Figure 5.** The five-dimensional hypercube of triplets NNR, which is the union of the 4-dimensional hypercubes YNR and RNR..

References

1. Ibba M, SÖll D (2000) Aminoacyl-tRNA synthesis. *Annu Rev Biochem* 69:617–650. https:// doi. org/ 10. 1146/ annur ev. bioch em. 69.1.617
2. Cusack S, Berthet-Colominas C, Hartlein M, Nassar N, Leberman R (1990) A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 A. *Nature* 347:249–255. https:// doi. org/ 10. 1038/ 34724 9a0
3. Eriani G, Delarue M, Poch O, Gangloff J, Moras D (1990) Partition of aminoacyl-tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* 347:203–206. https://doi. org/ 10. 1038/ 34720 3a0
4. Burbaum JJ, Schimmel P (1991) Structural relationships and the classification of aminoacyl-tRNA synthetases. *J Biol Chem* 266:16965–16968
5. Ribas de Pouplana L, Schimmel P (2001a) Two classes of tRNA synthetases suggested by sterically compatible dockings on tRNA acceptor stem. *Cell* 104:191–193
6. Ribas De Pouplana L, Schimmel P (2001b) Operational RNA code for amino acids in relation to genetic code in evolution. *J Biol Chem* 276:6881–6884. https:// doi. org/ 10. 1074/ jbc. R0000 32200

7. Ribas de Pouplana L (2020) The evolution of aminoacyl-tRNA synthetases: From dawn to LUCA. In Enzymes 48, 11–37, Academic Press
8. Schimmel P, Giege R, Moras D, Yokoyama S (1993) An operational RNA code for amino acids and possible relation to genetic code. *Proc Nat Acad Sci USA* 90(19):8763–8768. https:// doi. org/ 10.1073/ pnas. 90. 19. 8763
9. José M.V., Román J.R., Zamudio G.S., Farías S.T. Symmetrical Distributions of Aminoacyl-tRNA Synthetases during the evolution of the Genetic Code. *Theory in Biosciences* (2023). https://doi.org/10.1007/s12064-023-00394-0.
10. José MV, Morgado ER, Govezensky T. (2007) An extended RNA code and its relationship to the standard genetic code: an algebraic and geometrical approach. *Bull. Math. Biol.* (2007). 69:215–243. https:// doi.org/ 10. 1007/ s11538- 006- 9119-3
11. José M.V., Morgado E.R., Govezensky T. Genetic Hotels for the Standard Genetic Code: Evolutionary Analysis based upon Novel Three-dimensional Algebraic Models. *Bull. Math. Biol.* (2011). **73**: 1443-1476.
12. Nirenberg MW, Matthaei JH. (1961) The dependance of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl Acad. Sci. USA* **47**, 1588–1602. (doi:10.1073/pnas.47.10.1588)
13. Nagel GM, Doolittle RF (1995) Phylogenetic analysis of the aminoacyl-tRNA synthetases. *J Mol Evol* 40:487–498. https:// doi. org/10. 1007/ BF001 66617
14. Woese CR, Olsen GJ, Ibba M, SÖll D (2000) Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol Mol Biol Rev* 64:202–236. https:// doi. org/ 10. 1128/ MMBR. 64.1.202- 236. 2000
15. Rodin SN, Ohno S (1995) Two types of aminoacyl-tRNA synthetases originally encoded by complementary strands of the same nucleic acid. *Orig Life Evol Biosph* 25:565–589. https:// doi. org/ 10. 1007/BF015 82025
16. Rodin SN, Rodin AS (2008) On the origin of the genetic code: signatures of its primordial complementarity in tRNAs and Aminoacyl tRNA synthetases. *Heredity* 100:341–355
17. Rodin AS, Rodin SN, Carter CW (2009) On primordial sense–antisense coding. *J Mol Evol* 69(5):555–567
18. Carter CW Jr et al (2014) The Rodin-Ohno hypothesis that two enzyme superfamilies descended from one ancestral gene: an unlikely scenario for the origins of translation that will not be dismissed. *Biol Direct* 9:11
19. José MV, Zamudio GS, Morgado ER (2017) A unified model of the standard genetic code. *R Soc Open Sci*. 4:160908. https:// doi. org/10. 1098/ rsos1 60908