A Minimum Principle in Codon-Anticodon Interaction

A. Sciarrino[a,b,*], P. Sorba[c]

[a] *Dipartimento di Scienze Fisiche, Università di Napoli "Federico II"*
Complesso Universitario di Monte S. Angelo
Via Cinthia, I-80126 Napoli, Italy
[b] *I.N.F.N., Sezione di Napol*
sciarrino@na.infn.it

[c] *Laboratoire d'Annecy-le-Vieux de Physique Théorique LAPTH*
CNRS, UMR 5108, associée à l'Université de Savoie
BP 110, F-74941 Annecy-le-Vieux Cedex, France
sorba@lapp.in2p3.fr

Abstract

Imposing a minimum principle in the framework of the so called crystal basis model of the genetic code, we determine the structure of the minimum set of anticodons which allows the translational-transcription for animal mitochondrial code. The results are in very good agreement with the observed anticodons.

Keywords: genetic code, codon, anticodon

*Corresponding author: Tel +39/081676807 – Fax +39/081676346
Email addresses: sciarrino@na.infn.it (A. Sciarrino), sorba@lapp.in2p3.fr (P. Sorba)

1. Introduction

The translational-transcription process from DNA to proteins is a very complex process carried on in several steps. A key step is the translation from coding sequences of nucleotides in mRNA to the proteins chaines. In this process a role is played by the tRNA in which a triplet of nucleotides (anticodon) pairs to the triplet of nucleotides (codon) reading the genetic information. Since there are 60 codons (in mitochondrial code) specifying amino acids, the cell should contain 60 different tRNA molecules, each with a different anticodon in order to have a pairing codon anticodon following the usual Watson-Crick pattern, i.d. the pairing respectively between the nucleotides C and G, and U and A. Actually, however, the number of observed anticodons is less than 60. This implies that an anticodon may pair to more than one codon. Already in the middle of the sixties, it was realized that the pairing anticodon-codon does not follow the standard rule and Crick (Crick , 1966) proposed, on the basis of the base-pair stereochemistry, the "wobble hypothesis". According to this hypothesis a single tRNA type, with a a specified anticodon, is able to recognize two or more codons in particular differing only in the third nucleotide, i.e only the first two nucleotides of a codon triplet in mRNA have the standard precise pairing with the bases of the tRNA anticodon while the first nucleotide in the anticodon may pair to more than a nucleotide in the third position of the codon.

This rule has been subsequently widely confirmed and extended, with a better understanding of the chemical nucleotide modifications, for a review see (Agris , 2004). Since the years seventies the questions were raised (Jukes , 1977): how many anticodons do we need? which anticodons do manifest?

In order to explain which anticodon do manifest two main hypothesis have been advanced:

1. The conventional wobble versatility hypothesis assumes that the the first position of anticodon should have G (U) to read for codon with Y (respectively R) in third position.
2. The codon adaptation hypothesis states that the first position of anticodon should pair the most abondant codon in the family of synonymous codons.

For a comparison and discussion of the two hypothesis in fungal mitochondrial genomes and for marine bivalve mitochondrial genomes, see (Carullo and Xia , 2000) and (Hong Yu and Qi li , 2011).

In order to have a correct translation process between codons and amino-acids in the mitochondrial code we need a minimum number of 22 anticodons. In fact, in this code, the 20 amino-acids (a.a) are encoded by 2 sextets, 6 quadruplets and 12 doublets of codons. Considering a sextet as the sum of a quadruplet and a doublet, we need to dispose at least of 22 anti-codons, of which 8 should "read" the quadruplets and 14 the doublets. Indeed this seems to happen for the mitochondria of animals (Sprinzl et al. , 1998; Higgs et al. , 2003; Wilhelm and Nikolajewa , 2004; Nikolajewa et al. , 2006; Nikolaeva and Wilhelm , 2005). The data seem to confirm the empirical rule that the most used anticodons have as second and third nucleotide, respectively, the complementary to the first and second nucleotide of the codons, while the first nucleotide is U for the anticodons pairing the quadruplets, G and U for the anticodons pairing, respectively, the doublets ending with a pyrimidine and with a purine, with exception of Met.

The aim of this paper is to propose a mathematical approach in the framework of the "crystal basis model" model of the genetic code (Frappat et al , 1998), to determine which anticodon is chosen to translate the genetic information stored into the quadruplets and the doublets of codons. More generally, the idea is to require the minimization of a suitable operator or function, mathematically expressed in terms of the quantities defined in the model, to explain why and which anticodon is used to "read" more than a codon[1].

Let us very quickly recall the main ideas of the model introduced in (Frappat et al , 1998), for a review and some applications see (Frappat et al , 2001). In that paper we have proposed a mathematical framework in which the codons appear as composite states of nucleotides. The four nucleotides being assigned to the fundamental irreducible representation (irrep.) of the quantum group $\mathcal{U}_q(su(2) \oplus su(2))$ in the limit $q \to 0$, the codons are obtained as tensor product of nucleotides. Indeed, the properties of quantum group representations in the limit $q \to 0$, or crystal basis, are crucial to take into account the fact that a codon is an ordered triple of nucleotides. The nucleotide content of the $(\frac{1}{2}, \frac{1}{2})$ (fundamental) representation of $\mathcal{U}_{q\to 0}(su(2) \oplus su(2))$, i.e. the eigenvalues of $J_{H,3}$, $J_{V,3}$, is chosen as follows:

$$ \mathrm{C} \equiv (+\frac{1}{2}, +\frac{1}{2}) \qquad \mathrm{U} \equiv (-\frac{1}{2}, +\frac{1}{2}) \qquad \mathrm{G} \equiv (+\frac{1}{2}, -\frac{1}{2}) \qquad \mathrm{A} \equiv (-\frac{1}{2}, -\frac{1}{2}) \qquad (1) $$

where the first $su(2)$ - denoted $su(2)_H$- corresponds to the distinction between the purine bases A,G and the pyrimidine ones C,U and the second one - denoted $su(2)_V$ - corresponds to the complementarity rule C/G and U/A, Thus to represent a codon, we have to perform the tensor product of three $(\frac{1}{2}, \frac{1}{2})$ or fundamental representations of $\mathcal{U}_{q\to 0}(su(2) \oplus su(2))$ and we get the results, reported in Table 2, where we have also written the observed anticodon for the mitochondria of animals taken from (Sprinzl et al. , 1998). Really in the present paper we use the minimum principle in a reduced form as we are only interested to find the composition of the minimum number of anticodons. However in the last section we hint at some more general application of our schema.

The paper is organised as follows: in Sec. 2 we present the minimum principle, in Sec. 3 we apply the principle to the mitochondrial code for animals and we compare our theoretical results with the data of (Sprinzl et al. , 1998). In the final Section we give a short summary and some highlights on future developments and applications.

2. The "minimum" principle

Given a codon[2] XYZ ($X, Y, Z \in \{C, A, G, U\}$) we conjecture that an anticodon $X^a Y^a Z^a$, where $Y^a Z^a = Y_c X_c$, N_c denoting the nucleotide complementary to the nucleotide N according to the Watson-Crick pairing rule[3], pairs to the codon XYZ, i.e. it is most used to "read" the codon XYZ

---

[1]We do not discuss here the chemical modified structure of the nucleotides, e.g see (Agris , 2004).

[2]In the paper we use the notation $N = C, A, G, U$.; $R = G, A$. (*purine*); $Y = C, U$. (*pyrimidine*).

[3]This property is observed to be verified in most, but not in all, the observed cases. To simplify we shall assume it.

if it minimizes the operator $\mathcal{T}$, explicitly written in eq.(2) and computed between the "states", which can be read from Table 2, describing the codon and anticodon in the "crystal basis model". We write both codons (c) and anticodons (a) in 5" → 3" direction. As an anticodon is antiparallel to codon, the 1st nucleotide (respectively the 3rd nucleotide) of the anticodon is paired to the 3rd (respectively the 1st) nucleotide of the codon.

$$\mathcal{T} = 8c_H \, \vec{J}_H^c \cdot \vec{J}_H^a + 8c_V \, \vec{J}_V^c \cdot \vec{J}_V^a \tag{2}$$

where:

- $c_H.c_V$ are constants depending on the "biological species" and weakly depending on the encoded a.a., as we will later specify.

- J_H^c, J_V^c (resp. J_H^a, J_V^a) are the labels of $U_{q\to 0}(su(2)_H \oplus su(2)_V)$ specifying the state (Frappat et al , 1998) describing the codon XYZ (resp. the anticodon NY_cX_c pairing the codon XYZ).

- $\vec{J}_\alpha^c \cdot \vec{J}_\alpha^a$ $(\alpha = H, V)$ should be read as

$$\vec{J}_\alpha^c \cdot \vec{J}_\alpha^a = \frac{1}{2} \left\{ \left(\vec{J}_\alpha^c \oplus \vec{J}_\alpha^{\,a} \right)^2 - (\vec{J}_\alpha^c)^2 - (\vec{J}_\alpha^a)^2 \right\} \tag{3}$$

and $\vec{J}_\alpha^c \oplus \vec{J}_\alpha^a \equiv \vec{J}_\alpha^T$ stands for the irreducible representation which the codon-anticodon state under consideration belongs to, the tensor product of $\vec{J}_\alpha^c$ and $\vec{J}_\alpha^a$ being performed according to the rule of (Kashiwara , 1990), choosing the codon as first vector and the anticodon as second vector. Note that $\vec{J}_\alpha^2$ should be read as the Casimir operator whose eigenvalues are given by $J_\alpha(J_\alpha + 1)$.

For example the value of $\mathcal{T}$ between the anticodon UUU and the codon AAC is, using Table 4:

$$< UUU|\mathcal{T}|AAC >= -6\,c_H + 18\,c_V \tag{4}$$

As we are interested in finding the composition of the 22 anticodons, minimun number to ensure a faithful translation, we shall assume that the used anticodon for each quartet and each doublet is the one which minimizes the averaged value of the operator given in eq.(2), the average being performed over the 4 (2) codons for quadruplets (doublets), see next section. Indeed it is well known that synonymous codons are not used with equal frequency. Therefore, in finding the structure of the anticodons in the minimum set, it appears reasonable that the codon usage probability plays a role in the determination of the chosen anticodon. If the codon XYZ is used more frequently than the codon XYZ', the codon XYZ should give an impact larger than the codon XYZ' in determining the choice of the anticodon in the $\mathcal{T}$. Therefore, $\mathcal{T}_{av}$ appears more appropriate than $\mathcal{T}$ for our purpose.

3. Structure of the minimum number of anticodons

According to our conjecture on the existence of a minimum principle we determine, for each quadruplet (q) and each doublet (d), the anticodon which minimizes the averaged value $\mathcal{T}_{av}$ of the operator $\mathcal{T}$ (see below). We analyse separately the case of quadruplets and doublets.

3

3.1. Quadruplets

Let us give an example of what we mean by averaged value of $\mathcal{T}$. For example let us consider the anticodon CAC for the a.a. Val, we have to compute

$$
\begin{aligned}
\mathcal{T}_{av}(CAC, Val) &= \sum_N P_N^q < CAC|\mathcal{T}|GUN > \\
&= P_C^q < CAC|\mathcal{T}|GUC > + P_U^q < CAC|\mathcal{T}|GUU > \\
&\quad + P_G^q < CAC|\mathcal{T}|GUG > + P_A^q < CAC|\mathcal{T}|GUA > \\
&= 2(P_C^q + P_U^q + P_G^q + P_A^q)\, c_H + (6P_C^q + 6P_U^q + 2P_G^q + 2P_A^q)\, c_V \\
&= 2\, c_H + [6P_Y^q + 2(1 - P_Y^q)]\, c_V
\end{aligned}
\tag{5}
$$

In the computation we have to take into account the codon usage frequency or relative percentage of the appearance of each codon in the quadruplet and we have denoted with P_N^q the codon usage frequency for codon ending with N. Really we need to introduce the following four positive frequencies P_Y^q P_R^q, P_S^q, P_W^q, with the normalization condition:

$$
P_Y^q + P_R^q = P_S^q + P_W^q = 1
\tag{6}
$$

where, respectively, P_Y^q, P_R^q, P_S^q and P_W^q denote the relative usage frequency of the codons ending with nucleotides C,U (pyrimidine), G,A (purine), C,G and U,A. From Table 3 we can compute the value which we report in Table 5.

3.2. Doublets

In the computation we have to take into account the codon usage frequency in the doublet. Now we need to introduce the following four positive frequencies P_C^d, P_U^d, P_G^d, P_A^d, with the normalization condition

$$
P_C^d + P_U^d = P_G^d + P_A^d = 1
\tag{7}
$$

As example let us compute the averaged value of $\mathcal{T}$ for Asp. Let us consider the anticodon CUC we have to compute

$$
\begin{aligned}
\mathcal{T}_{av}(CUC, Asp) &= \sum_Y P_Y^d < CUC|\mathcal{T}|GAY >= P_C^d < CUC|\mathcal{T}|GAC > + P_U^d < CUC|\mathcal{T}|GAU > \\
&= 2\, c_H + 18\, c_V
\end{aligned}
\tag{8}
$$

From Table 4, we can compute the values, which we report in Table 6.

Let us remark that:

- for all a.a. the contribution of $su_V(2)$ verifies the same property than for the quadruplets and, moreover, is not depending on the codon usage;

- for 4 a.a. the contribution of $su_V(2)$ is the same for all anticodon.

4

From the above remarks we easily realize that the case of doublets is more complicated than the one of the quadruplets. In some sense the contribution of $su_V(2)$ plays a role only in establishing the most preferred anticodon. Moreover, as we do not want a priori to exclude any anticodon, we have to face the possibility that an anticodon can be chosen to read for more than one doublet. In order to avoid this problem, in contradiction with the requirement of a faithful translation process, we make the following choice:

1. the sign of the constant c_H for the doublets ending with a purine is the opposite of the sign of the doublets ending with a pyrimidine with the same dinucleotide (if it does exist)[4].
2. the sign of c_H for the 8 weak dinucleotides encoding doublets is positive for the following 4 doublets UUY, UAY, AUY, AAY and negative for the remaining 4, i.e. CAY, UGY, AGY, GAY.

and fix the following procedure, while considering doublets with the same dinucleotide:

1. first we select, among the four possible anticodons, the one giving the lowest value for $\mathcal{T}$ averaged on the two codons of each doublet and assign this anticodon to the corresponding doublet.
2. then the anticodon reading the second doublet is chosen between the two ones containing as a first nucleotide a purine, resp. a pyrimidine, if the first nucleotide of the anticodon already determined for the first doublet is a pyrimidine (resp. a purine).

As an illustration, we take the case of the Cys and Trp amino acids. The anticodon GCA can minimize both of them, but more Cys (due to the value -6 for the c_V coefficient) than Trp (with value 2 for the same c_V coefficient). Thus GCA will be taken as the anticodon relative to Cys, while the choice for the Trp anticodon will be made between UCA and CCA, that is the two candidates with a pyrimidine as a first nucleotide, the anticodon GCA starting with a purine.

Let us remark that, even if the above assumptions seem rather ad hoc, indeed a general symmetric pattern shows up: for half of a.a. c_H is positive and for the other half is negative; the first set of 4 dinucleotides involves only 'weak" nucleotides, the second one a "strong" nucleotide; the dinucleotides XY and YX correspond to the same sign.

3.3. Discussion

For all quadruplets, from Table 3, we remark that for $c_H > 0$ and $c_V < 0$ the anticodon minimizing the average value of $\mathcal{T}$ has the composition UX_cY_c. For Leu, Val and Thr a very weak condition for codon usage frequency has to be satisfied, i.e for the first two a.a. $P_S^q > 0,25$ and for the last one $P_Y^q > 0,125$[5]. The results are in agreement with the observed anticodons, see (Sprinzl et al. , 1998) and Table 2.

---

[4]We call dinucleotide the first two nucleotids of the codon.
[5] The constraint on the codon usage frequency can be released, imposing a suitable condition between c_H and c_V.

For doublets, we remark that, with the choice of sign of c_H above specified and $c_V > 0$ for all a.a., the anticodons minimizing the average value of $\mathcal{T}$ are in agreement with the observed anticodon, see (Sprinzl et al. , 1998) and Table 2. We summarize in Table 1 the results for the doublets.

a.a	sign c_H	anticodon	note
His	-	GUG	$P_C^d > 0,25$
Gln	+	UUG	$P_G^d > 0,25$
Phe	-	GAA	
Leu	+	UAA	
Cys	+	GCA	
Trp	-	UCA	
Tyr	-	GUA	
Ser	+	GCU	
Asp	+	GUC	$P_C^d > 0,25$
Glu	-	UUC	$P_G^d > 0,25$
Ile	+	GAU	
Met	-	CAU	
Asn	-	GUU	
Lys	+	UUU	

Table 1: Anticodon minimizing the operator $\mathcal{T}$, averaged over the two codons, for any amino acid encoded by a doublet, specifying the sign of c_H.

Let us remark that we find that for Met the anticodon is not UAU, as it should be expected from the empirical rule above quoted, but CAU which seems in agreement with the data, see (Sprinzl et al. , 1998).

4. Conclusions

We have found that the anticodons minimizing the conjectured operator $\mathcal{T}$ given in eq.(2), averaged over the concerned multiplets, are in very good agreement, the results depending only on the signs of the two coupling constants, with the observed ones, even if we have made comparison with a limited database.

The fact that the crystal basis model is able to explain, in a relatively simple way, the observed anticodon-codon pairing which has its roots on the stereochemical properties of nucleotides (Lim and Curran , 2001) strongly suggests that our modeling is able to incorporate some crucial features of the complex physico-chemical structure of the genetic code.

It is rather clear that the operator $\mathcal{T}$ can be looked at as a codon-anticodon interaction operator. In this spirit, given a codon, the selected anti-codon for fixing the corresponding amino-acid, is determined as the one which minimizes the interaction. It is hard at this point to be more specific on the physico-chemical aspects of this quantity.

Indeed, the analysis in (Percudani and Ottonello , 1999) suggests that the use of the wobble behaviour is also dictated from the requirement of the optimization of translational efficiency and in (Lehmann , 2000) the free-energy change of anticodon-codon interaction has been put into relation with the dissociation time, depending on the molecular structure of the encoded amino-acid, while in (Lim and Curran , 2001) emphasis has been put in the hydrogen and ionic bondings.

It is intriguing that complex behaviour involving thermodynamical considerations as well as evolutionary effects can be cast in a single simple mininum principle.

It might be interesting to note, among the different previously obtained applications of the crystal basis model, that this model has previously allowed to establish a pattern of correlations between the physico-chemical properties of the amino-acids and the assignment of the corresponding coding codons in the model (Frappat et al , 2002). Incidentally let us remark that the model explains the symmetry codon anticodon remarked in (Wilhelm and Nikolajewa , 2004). Let us stress that our modeling has a very peculiar feature which makes it very different from the standard 4-letter alphabet, used to identify the nucleotides, as well as with the usual modeling of nucleotide chain as spin chain. Indeed the identification of the nucleotides with the fundamental irrep. of $\mathcal{U}_q(su(2)_H \oplus su(2)_V)$ introduces a sort of double "bio-spin", which allows the description of any ordered sequence of n nucleotides as as state of an irrep. and allows to describe interactions using the standard powerful mathematical language used in physical spin models.

In the present paper we have faced the problem to find the structure of the mimimum set of anticodons and, then, we have used a very simple form for the operator $\mathcal{T}$, with the main aim to present a simple, mathematical modeling of the extremely complex codon-anticodon interaction. We have not at all discussed the possible appearance of any other anticodon, which should require a more quantitative discussion. For such analysis, as well as for the eukaryotic code, the situation may be different and more than an anticodon may pair to a quartet. For this future aim we have here reported Tables 3 and 4.

The pattern, which in the general case may show up, is undoubtedly more complicated, depending on the biological species and on the concerned biosynthesis process, but it is natural to argue that the usage of anticodons exhibits the general feature to assure an "efficient" translation process by a number of anticodons, minimum with respect to the involved constraints. A more refined and quantitative analysis, as well as comparison with other organisms, which should require more data, depends on the value of these constants. Very likely, it might happen that the assumption of the "universal" feature of c_H and c_V should be released and that the expression of the operator eq.(2) should be modified, for example by adding a term of "spin-spin" interaction of the type

$$4g_H J^c_{H,3} J^a_{H,3} + 4g_V J^c_{V,3} J^a_{V,3} \qquad (9)$$

where the values of $J_{H,3}$ and $J_{V,3}$, both for codons and anticodons, can be read out from their nucleotide composition, see Table 2. However the pattern which shows up in Tables 3 and 4, with the values of the coefficients equal in pairs, strongly suggests that the minimum number of anticodons should be 32 (3 for the sextets, 2 for quadruplets and triplet and 1 for doublets and singlets).

codon	a.a.	J_H	J_V	$J_{3,H}$	$J_{3,V}$	anticodon	codon	a.a.	J_H	J_V	$J_{3,H}$	$J_{3,V}$	anticodon
CCC	P	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$		UCC	S	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	
CCU	P	$(\frac{1}{2}\ \frac{3}{2})^1$		$\frac{1}{2}$	$\frac{3}{2}$		UCU	S	$(\frac{1}{2}\ \frac{3}{2})^1$		$-\frac{1}{2}$	$\frac{3}{2}$	
CCG	P	$(\frac{3}{2}\ \frac{1}{2})^1$		$\frac{3}{2}$	$\frac{1}{2}$	**UGG**	UCG	S	$(\frac{3}{2}\ \frac{1}{2})^1$		$\frac{1}{2}$	$\frac{1}{2}$	**UGA**
CCA	P	$(\frac{1}{2}\ \frac{1}{2})^1$		$\frac{1}{2}$	$\frac{1}{2}$		UCA	S	$(\frac{1}{2}\ \frac{1}{2})^1$		$-\frac{1}{2}$	$\frac{1}{2}$	
CUC	L	$(\frac{1}{2}\ \frac{3}{2})^2$		$\frac{1}{2}$	$\frac{3}{2}$		UUC	F	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{1}{2}$	$\frac{3}{2}$	*GAA*
CUU	L	$(\frac{1}{2}\ \frac{3}{2})^2$		$-\frac{1}{2}$	$\frac{3}{2}$		UUU	F	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{3}{2}$	$\frac{3}{2}$	
CUG	L	$(\frac{1}{2}\ \frac{1}{2})^3$		$\frac{1}{2}$	$\frac{1}{2}$	**UAG**	UUG	L	$(\frac{3}{2}\ \frac{1}{2})^1$		$-\frac{1}{2}$	$\frac{1}{2}$	
CUA	L	$(\frac{1}{2}\ \frac{1}{2})^3$		$-\frac{1}{2}$	$\frac{1}{2}$		UUA	L	$(\frac{3}{2}\ \frac{1}{2})^1$		$-\frac{3}{2}$	$\frac{1}{2}$	*UAA*
CGC	R	$(\frac{3}{2}\ \frac{1}{2})^2$		$\frac{3}{2}$	$\frac{1}{2}$		UGC	C	$(\frac{3}{2}\ \frac{1}{2})^2$		$\frac{1}{2}$	$\frac{1}{2}$	
CGU	R	$(\frac{1}{2}\ \frac{1}{2})^2$		$\frac{1}{2}$	$\frac{1}{2}$		UGU	C	$(\frac{1}{2}\ \frac{1}{2})^2$		$-\frac{1}{2}$	$\frac{1}{2}$	*GCA*
CGG	R	$(\frac{3}{2}\ \frac{1}{2})^2$		$\frac{3}{2}$	$-\frac{1}{2}$	**UCG**	UGG	W	$(\frac{3}{2}\ \frac{1}{2})^2$		$\frac{1}{2}$	$-\frac{1}{2}$	
CGA	R	$(\frac{1}{2}\ \frac{1}{2})^2$		$\frac{1}{2}$	$-\frac{1}{2}$		UGA	W	$(\frac{1}{2}\ \frac{1}{2})^2$		$-\frac{1}{2}$	$-\frac{1}{2}$	*UCA*
CAC	H	$(\frac{1}{2}\ \frac{1}{2})^4$		$\frac{1}{2}$	$\frac{1}{2}$		UAC	Y	$(\frac{3}{2}\ \frac{1}{2})^2$		$-\frac{1}{2}$	$\frac{1}{2}$	
CAU	H	$(\frac{1}{2}\ \frac{1}{2})^4$		$-\frac{1}{2}$	$\frac{1}{2}$	*GUG*	UAU	Y	$(\frac{3}{2}\ \frac{1}{2})^2$		$-\frac{3}{2}$	$\frac{1}{2}$	*GUA*
CAG	Q	$(\frac{1}{2}\ \frac{1}{2})^4$		$\frac{1}{2}$	$-\frac{1}{2}$		UAG	**Ter**	$(\frac{3}{2}\ \frac{1}{2})^2$		$-\frac{1}{2}$	$-\frac{1}{2}$	—
CAA	Q	$(\frac{1}{2}\ \frac{1}{2})^4$		$-\frac{1}{2}$	$-\frac{1}{2}$	*UUG*	UAA	**Ter**	$(\frac{3}{2}\ \frac{1}{2})^2$		$-\frac{3}{2}$	$-\frac{1}{2}$	—
GCC	A	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{1}{2}$		ACC	T	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	
GCU	A	$(\frac{1}{2}\ \frac{3}{2})^1$		$\frac{1}{2}$	$\frac{1}{2}$		ACU	T	$(\frac{1}{2}\ \frac{3}{2})^1$		$-\frac{1}{2}$	$\frac{1}{2}$	
GCG	A	$(\frac{3}{2}\ \frac{1}{2})^1$		$\frac{3}{2}$	$-\frac{1}{2}$	**UGC**	ACG	T	$(\frac{3}{2}\ \frac{1}{2})^1$		$\frac{1}{2}$	$-\frac{1}{2}$	**UGU**
GCA	A	$(\frac{1}{2}\ \frac{1}{2})^1$		$\frac{1}{2}$	$-\frac{1}{2}$		ACA	T	$(\frac{1}{2}\ \frac{1}{2})^1$		$-\frac{1}{2}$	$-\frac{1}{2}$	
GUC	V	$(\frac{1}{2}\ \frac{3}{2})^2$		$\frac{1}{2}$	$\frac{1}{2}$		AUC	I	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	
GUU	V	$(\frac{1}{2}\ \frac{3}{2})^2$		$-\frac{1}{2}$	$\frac{1}{2}$		AUU	I	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{3}{2}$	$\frac{1}{2}$	*GAU*
GUG	V	$(\frac{1}{2}\ \frac{1}{2})^3$		$\frac{1}{2}$	$-\frac{1}{2}$	**UAC**	AUG	M	$(\frac{3}{2}\ \frac{1}{2})^1$		$-\frac{1}{2}$	$-\frac{1}{2}$	
GUA	V	$(\frac{1}{2}\ \frac{1}{2})^3$		$-\frac{1}{2}$	$-\frac{1}{2}$		AUA	M	$(\frac{3}{2}\ \frac{1}{2})^1$		$-\frac{3}{2}$	$-\frac{1}{2}$	*CAU*
GGC	G	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{1}{2}$		AGC	S	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	
GGU	G	$(\frac{1}{2}\ \frac{3}{2})^1$		$\frac{1}{2}$	$-\frac{1}{2}$		AGU	S	$(\frac{1}{2}\ \frac{3}{2})^1$		$-\frac{1}{2}$	$-\frac{1}{2}$	*GCU*
GGG	G	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{3}{2}$	**UCC**	AGG	**Ter**	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{3}{2}$	
GGA	G	$(\frac{1}{2}\ \frac{3}{2})^1$		$\frac{1}{2}$	$-\frac{3}{2}$		AGA	**Ter**	$(\frac{1}{2}\ \frac{3}{2})^1$		$-\frac{1}{2}$	$-\frac{3}{2}$	—
GAC	D	$(\frac{1}{2}\ \frac{3}{2})^2$		$\frac{1}{2}$	$-\frac{1}{2}$		AAC	N	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	
GAU	D	$(\frac{1}{2}\ \frac{3}{2})^2$		$-\frac{1}{2}$	$-\frac{1}{2}$	*GUC*	AAU	N	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{3}{2}$	$-\frac{1}{2}$	*GUU*
GAG	E	$(\frac{1}{2}\ \frac{3}{2})^2$		$\frac{1}{2}$	$-\frac{3}{2}$		AAG	K	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	
GAA	E	$(\frac{1}{2}\ \frac{3}{2})^2$		$-\frac{1}{2}$	$-\frac{3}{2}$	*UUC*	AAA	K	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{3}{2}$	$-\frac{3}{2}$	*UUU*

Table 2: The vertebral mitochondrial code. The upper label denotes different irreducible representations. We list the most used anticodons for mitochondria of animals, see (Sprinzl et al. , 1998). In bold-red (italic-blue) the anticodons reading quadruplets (resp. doublets).

a.a	codon	$K_{H,C}$	$K_{V,C}$	$K_{H,U}$	$K_{V,U}$	$K_{H,G}$	$K_{V,G}$	$K_{H,A}$	$K_{V,A}$
Pro	CCC	18	-10	-6	-10	18	-30	-6	-30
	CCU	6	-10	-10	-10	6	-30	-10	-30
	CCG	18	-6	-6	-6	18	-10	-6	-10
	CCA	6	-6	-10	-6	6	-10	-10	-10
Leu	CUC	2	-10	-10	-10	2	-30	-10	-30
	CUU	2	-10	6	-10	2	-30	6	-30
	CUG	2	-6	-10	-6	2	-10	-10	-10
	CUA	2	-6	6	-6	2	-10	6	-10
Arg	CGC	18	2	-6	2	18	-6	-6	-6
	CGU	6	2	-10	2	6	-6	-10	-6
	CGG	18	2	-6	2	18	2	-6	2
	CGA	6	2	-10	2	6	2	-10	2
Ala	GCC	18	6	-6	6	18	-22	-6	-22
	GCU	6	6	-10	6	6	-22	-10	-22
	GCG	18	2	-6	2	18	6	-6	6
	GCA	6	2	-10	2	6	6	-10	6
Gly	GGC	18	18	-6	18	18	-6	-6	-6
	GGU	6	18	-10	18	6	-6	-10	-6
	GGG	18	18	-6	18	18	18	-6	18
	GGA	6	18	-10	18	6	18	-10	18
Val	GUC	2	6	-10	6	2	-22	-10	-22
	GUU	2	6	6	6	2	-22	6	-22
	GUG	2	2	-10	2	2	6	-10	6
	GUA	2	2	6	2	2	6	6	6
Ser	UCC	6	-10	-10	-10	6	-30	-10	-30
	UCU	2	-10	2	-10	2	-30	2	-30
	UCG	6	-6	-10	-6	6	-10	-10	-10
	UCA	2	-6	2	-6	2	-10	2	-10
Thr	ACC	6	6	-10	6	6	-22	-10	-22
	ACU	2	6	2	6	2	-22	2	-22
	ACG	6	2	-10	2	6	6	-10	6
	ACA	2	2	2	2	2	6	2	6

Table 3: Values of the coefficient multiplying c_H $(K_H = 8\vec{J_H^c} \cdot \vec{J_H^a})$ and c_V $(K_V = 8\vec{J_V^c} \cdot \vec{J_V^a})$ computed from the value of the tensor product of the codon XYZ with the anticodon NY_cX_c, for the quadruplets.

a.a	codon	$K_{H,C}$	$K_{V,C}$	$K_{H,U}$	$K_{V,U}$	$K_{H,G}$	$K_{V,G}$	$K_{H,A}$	$K_{V,A}$
His	CAC	2	2	-10	2	2	-6	-10	-6
	CAU	2	2	6	2	2	-6	6	-6
Gln	CAG	2	2	-10	2	2	2	6	2
	CAA	2	2	6	2	2	2	6	2
Phe	UUC	-10	-10	-6	-10	-10	-30	-6	-30
	UUU	6	-10	18	-10	6	-30	18	-30
Leu	UUG	-10	-6	-6	-6	-10	-10	-6	-10
	UUA	6	-6	18	-6	6	-10	18	-10
Cys	UGC	6	2	-10	2	6	-6	-10	-6
	UGU	2	2	2	2	2	-6	2	-6
Trp	UGG	6	2	-10	2	6	2	-10	2
	UGA	2	2	2	2	2	2	2	2
Tyr	UAC	-10	2	-6	2	-10	-6	-6	-6
	UAU	6	2	18	2	6	-6	18	-6
Asp	GAC	2	18	-10	18	2	-6	-10	-6
	GAU	2	18	6	18	2	-6	6	-6
Glu	GAG	2	18	-10	18	2	18	-10	18
	GAA	2	18	6	18	2	18	6	18
Ile	AUC	-10	6	-6	6	-10	-22	-6	-22
	AUU	6	6	18	6	6	-22	18	-22
Met	AUG	-10	2	-6	2	-10	6	-6	6
	AUA	6	2	18	2	6	6	18	6
Ser	AGC	6	18	-10	18	6	-6	-10	-6
	AGU	2	18	2	18	2	-6	2	-6
Asn	AAC	-10	18	-6	18	-10	-6	-6	-6
	AAU	6	18	18	18	6	-6	18	-6
Lys	AAG	-10	18	-6	18	-10	18	-6	18
	AAA	6	18	18	18	6	18	18	18

Table 4: Values of the coefficient multiplying c_H ($K_H = 8\vec{J}_H^c \cdot \vec{J}_H^a$) and c_V ($K_V = 8\vec{J}_V^c \cdot \vec{J}_V^a$) computed from the value of the tensor product of the codon XYZ with the anticodon $N\bar{Y}_c\bar{X}_c$, for the doublets.

a.a	anticodon	coeff. c_H	coeff. c_V
Pro	CGG	$18P_S^q + 6(1-P_S^q)$	$-10P_Y^q - 6(1-P_Y^q)$
	UGG	$-6P_S^q - 10(1-P_S^q)$	$-10P_Y^q - 6(1-P_Y^q)$
	GGG	$18P_S^q + 6(1-P_S^q)$	$-30P_Y^q - 10(1-P_Y^q)$
	AGG	$-6P_S^q - 10(1-P_S^q$	$-30P_Y^q - 10(1-P_Y^q)$
Leu	CAG	2	$-10P_Y^q - 6(1-P_Y^q)$
	UAG	$-10P_S^q + 6(1-P_S^q)$	$-10P_Y^q - 6(1-P_Y^q)$
	GAG	2	$-30P_Y^q - 10(1-P_Y^q)$
	AAG	$-10P_S^q + 6(1-P_S^q)$	$-30P_Y^q - 10(1-P_Y^q)$
Arg	CCG	$18P_S^q + 6(1-P_S^q)$	2
	UCG	$-6P_S^q - 10(1-P_S^q)$	2
	GCG	$18P_S^q + 6(1-P_S^q)$	$-6P_Y^q + 2(1-P_Y^q)$
	ACG	$-6P_S^q - 10(1-P_S^q)$	$-6P_Y^q + 2(1-P_Y^q)$
Ala	CGC	$18P_S^q + 6(1-P_S^q)$	$6P_Y^q + 2(1-P_Y^q)$
	UGC	$-6P_S^q - 10(1-P_S^q)$	$6P_Y^q + 2(1-P_Y^q)$
	GGC	$18P_S^q + 6(1-P_S^q)$	$-22P_Y^q + 6(1-P_Y^q)$
	AGC	$-6P_S^q - 10(1-P_S^q)$	$-22P_Y^q + 6(1-P_Y^q)$
Gly	CCC	$18P_S^q + 6(1-P_S^q)$	18
	UCC	$-6P_S^q - 10(1-P_S^q)$	18
	GCC	$18P_S^q + 6(1-P_S^q)$	$6P_Y^q + 18(1-P_Y^q)$
	ACC	$-6P_S^q - 10(1-P_S^q)$	$6P_Y^q + 18(1-P_Y^q)$
Val	CAC	2	$6P_Y^q + 2(1-P_Y^q)$
	UAC	$-10P_S^q + 6(1-P_S^q)$	$6P_Y^q + 18(1-P_Y^q)$
	GAC	2	$-22P_Y^q + 6(1-P_Y^q)$
	AAC	$-10P_S^q + 6(1-P_S^q)$	$-22P_Y^q + 6(1-P_Y^q)$
Ser	CGA	$6P_S^q + 2(1-P_S^q)$	$-10P_Y^q - 6(1-P_Y^q)$
	UGA	$-10P_S^q + 2(1-P_S^q)$	$-10P_Y^q - 6(1-P_Y^q)$
	GGA	$6P_S^q + 2(1-P_S^q)$	$-30P_Y^q - 10(1-P_Y^q)$
	AGA	$-10P_S^q + 2(1-P_S^q)$	$-30P_Y^q - 10(1-P_Y^q)$
Thr	CGU	$6P_S^q + 2(1-P_S^q)$	$6P_Y^q + 2(1-P_Y^q)$
	UGU	$-10P_S^q + 2(1-P_S^q)$	$6P_Y^q + 2(1-P_Y^q)$
	GGU	$6P_S^q + 2(1-P_S^q)$	$-22P_Y^q + 6(1-P_Y^q)$
	AGU	$-10P_S^q + 2(1-P_S^q)$	$-22P_Y^q + 6(1-P_Y^q)$

Table 5: Value of the coefficients multiplying c_H and c_V in $\mathcal{T}_{av}$, computed for any anticodon and averaged over the four codons for each quadruplet.

a.a	**a**nticodon	**c**oeff. c_H	**c**oeff. c_V
His	CUG	2	2
	UUG	$-10P_C^d + 6(1\text{-}P_C^d)$	2
	GUG	2	-6
	AUG	$-10P_C^d + 6(1\text{-}P_C^d)$	-6
Gln	CUG	2	2
	UUG	$-10P_G^d + 6(1\text{-}P_G^d)$	2
	GUG	2	2
	AUG	6	2
Phe	CAA	$-10P_C^d + 6(1\text{-}P_C^d)$	-10
	UAA	$-6P_C^d + 18(1\text{-}P_C^d)$	-10
	GAA	$-10P_C^d + 6(1\text{-}P_C^d)$	-30
	AAA	$-6P_C^d + 18(1\text{-}P_C^d)$	-30
Leu	CAA	$-10P_G^d + 6(1\text{-}P_G^d)$	-6
	UAA	$-6P_G^d + 18(1\text{-}P_G^d)$	-6
	GAA	$-10P_G^d + 6(1\text{-}P_G^d)$	-10
	AAA	$-6P_G^d + 18(1\text{-}P_G^d)$	-10
Cys	CCA	$6P_C^d + 2(1\text{-}P_C^d)$	2
	UCA	$-10P_C^d + 2(1\text{-}P_C^d)$	2
	GCA	$6P_C^d + 2(1\text{-}P_C^d)$	-6
	ACA	$-10P_C^d + 2(1\text{-}P_C^d)$	-6
Trp	CCA	$6P_G^d + 2(1\text{-}P_G^d)$	2
	UCA	$-10P_G^d + 2(1\text{-}P_G^d)$	2
	GCA	$6P_G^d + 2(1\text{-}P_G^d)$	2
	ACA	$-10P_G^d + 2(1\text{-}P_G^d)$	2
Tyr	CUA	$-10P_C^d + 6(1\text{-}P_C^d)$	2
	UUA	$-6P_C^d + 18(1\text{-}P_C^d)$	2
	GUA	$-10P_C^d + 6(1\text{-}P_C^d)$	-6
	AUA	$-6P_C^d + 18(1\text{-}P_C^d)$	-6
Ser	CCU	$6P_C^d + 2(1\text{-}P_C^d)$	18
	UCU	$-10P_C^d + 2(1\text{-}P_C^d)$	18
	GCU	$6P_C^d + 2(1\text{-}P_C^d)$	-6
	ACU	$-10P_C^d + 2(1\text{-}P_C^d)$	-6
Asp	CUC	2	18
	UUC	$-10P_C^d + 6(1\text{-}P_C^d)$	18
	GUC	2	-6
	AUC	$-10P_C^d + 6(1\text{-}P_C^d)$	-6
Glu	CUC	2	18
	UUC	$-10P_G^d + 6(1\text{-}P_G^d)$	18
	GCA	2	18

Continued on next page

a.a	**a**nticodon	**c**oeff. c_H	**c**oeff. c_V
Continued from previous page			
	ACA	$-10P_G^d + 6(1\text{-}P_G^d)$	18
Ile	CAU	$-10P_C^d + 6(1\text{-}P_C^d)$	6
	UAU	$-6P_C^d + 18(1\text{-}P_C^d)$	6
	GAU	$-10P_C^d + 6(1\text{-}P_C^d$	-22
	AAU	$-6P_C^d + 18(1\text{-}P_C^d)$	-22
Met	CAU	$-10P_G^d + 6(1\text{-}P_G^d)$	2
	UAU	$-6P_G^d +186(1\text{-}P_G^d)$	2
	GAU	$-10P_G^d + 6(1\text{-}P_G^d)$	6
	AAU	$18P_G^d + 6(1\text{-}P_G^d)$	6
Asn	CUU	$-10P_C^d + 6(1\text{-}P_C^d)$	18
	UUU	$-6P_C^d + 18(1\text{-}P_C^d)$	18
	GUU	$-10P_C^d + 6(1\text{-}P_C^d)$	-6
	AUU	$-6P_C^d + 18(1\text{-}P_C^d)$	-6
Lys	CUU	$-10P_G^d + 6(1\text{-}P_G^d)$	18
	UUU	$-6P_G^d + 18(1\text{-}P_G^d)$	18
	GUU	$-10P_G^d + 6(1\text{-}P_G^d)$	18
	AUU	$-6P_G^d + 18(1\text{-}P_G^d)$	18

Table 6: Value of the coefficients multiplying c_H and c_V in $\mathcal{T}_{av}$, computed for any anticodon and averaged over the two codons for each doublet.

References

Agris P.F, 2004. *Decoding the genome: a modified view*, Nucleic Acids Research **32** 223-238.

Carullo M. and Xia X., 2000. *An Extensive Study of Mutation and Selection on Wobble Nucleotide in tRNA Anticodons in Fungal Mitochondrial Genomes*, J.Mol.Evol. **66** 484-493.

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

Frappat L , Sciarrino A. and Sorba P., 1998. *A Crystal base for the genetic code*, Phys.Lett. A, **250** 214-221.

Frappat L., Sciarrino A. and Sorba P., 2001. *Crystalizing the Genetic Code*, J.Biol.Phys. **27** 1-34.

Frappat L., Sciarrino A. and Sorba P., 2002. *Prediction of Physico-chemical Properties of amino-acids from the Genetic Code*, J.Biol.Phys. **28** 17-26.

Higgs P.G., Jameson D., Jow H. and Rattray M., 2003. *The Evolution of tRNA-L Genes in Animal Mitochondrial Genomes*, J.Mol.Evol. **57** 435-445.

Hong Yu and Qi Li, 2011. *Mutation and Selection on the Wobble tRNA Anticodons in Marine Bivalve Mitochondrial Genomes,* PLoS ONE. **6**) e1617.

Jukes T.H., 1977. *How Many Anticodons ?,* Science. **198** 319-320.

Kashiwara M., 1990. *Crystalizing the q-Analogue of Universal Enveloping Algebras,*Commun. Math. Phys. **133** 249-260.

Lehmann J., 2000. *Physico-chemical Constraints Connected with the Coding Properties of the Genetic System,* J.theor.Biol. **202** 129-114.

Lim V.I. and Curran J.S., 2001. *Analysis of codon:anticodon interactions within the ribosome provides new insights into codon reading and the genetic code structure,* RNA **7** 942-957.

Nikolaeva S.L. and Wilhelm T., 2005. *The Purine-Pyrimidine Classification Scheme Reveals New Patterns in the Genetic Code,*
http://www.imb-jena.de/∼sweta/genetic_code_and_evolution/Finalpaper.pdf

Nikolajewa S., Friedel M., Beyer A. and Wilhelm T., 2006. *The New Classification Scheme of the Genetic Code, its Early Evolution and tRNA Usage* J. Bioinformatics and Computational Biology, **4**609-620.

Percudani R. and Ottonello S., 1999. *Selection at the Wobble Position of Codons Read by the Same tRNA in Saccharomyces cerevisiae,* Mod.Biol.Evol. **16** 1752-1762.

Sprinzl M., Horn C., Brown M. loudovitch A. and Steinberg S., 1998. *Compilation of tRNA sequences and sequences of tRNA genes,* Nucleic.Acids.Researches **26** 148-153.
http://www.staff.uni-bayreuth.de/ ∼ btc914/search/index.html

Wilhelm T. and Nikolajewa S., 2004. *A New Classification Scheme of the Genetic Code,* J.Mol.Evol. **59** 598-605.