CODON-ANTICODON INTERACTION and GENETIC CODE EVOLUTION

A.Sciarrino and P.Sorba

Plan:

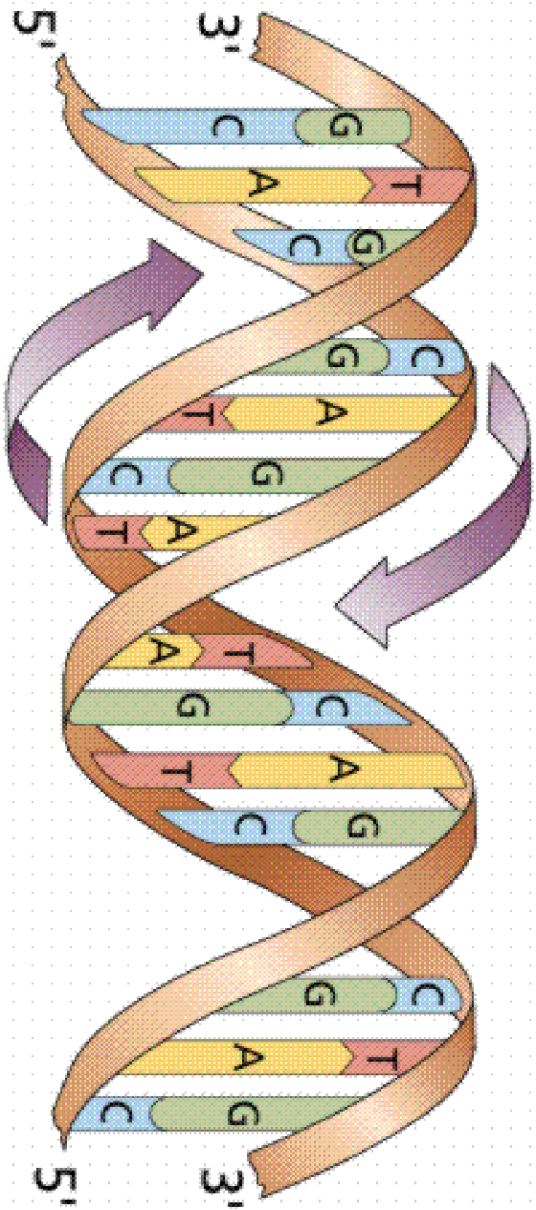
1. *Genetic code : a brief survey*
2. *Crystal Basis Model of the Genetic Code*
3. *Codon-anticodon Interaction*
4. *Evolution of the Genetic code*

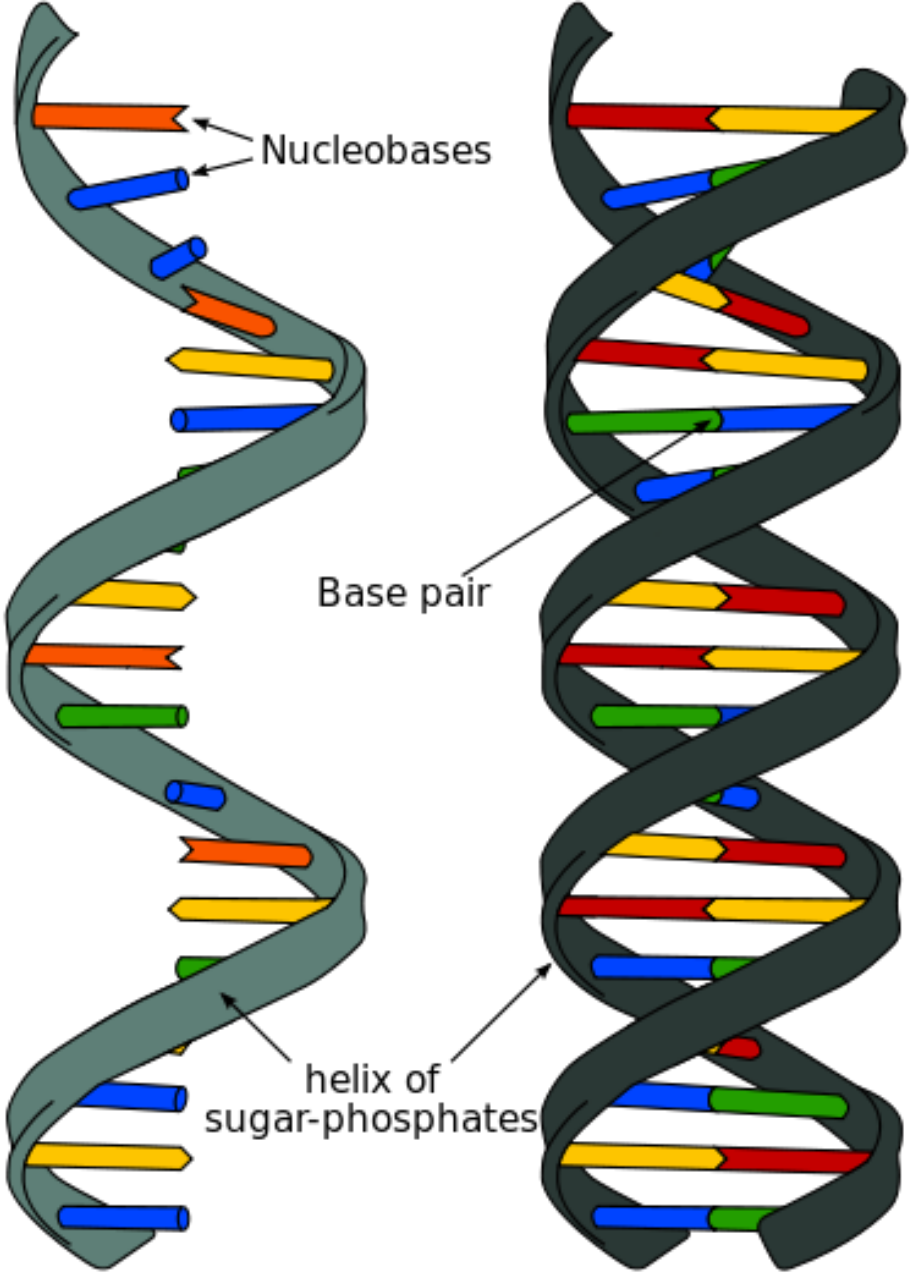
Biosystems 107 (2011) ,113-119; 111 (2013), 175-180

paul.sorba@lapth.cnrs.fr

1- Genetic Code: a brief survey

The transmission of information from DNA (formed by 4 bases **C,G,A,T**) (nucleotides) to protein building is a complex process of transcription and translation.

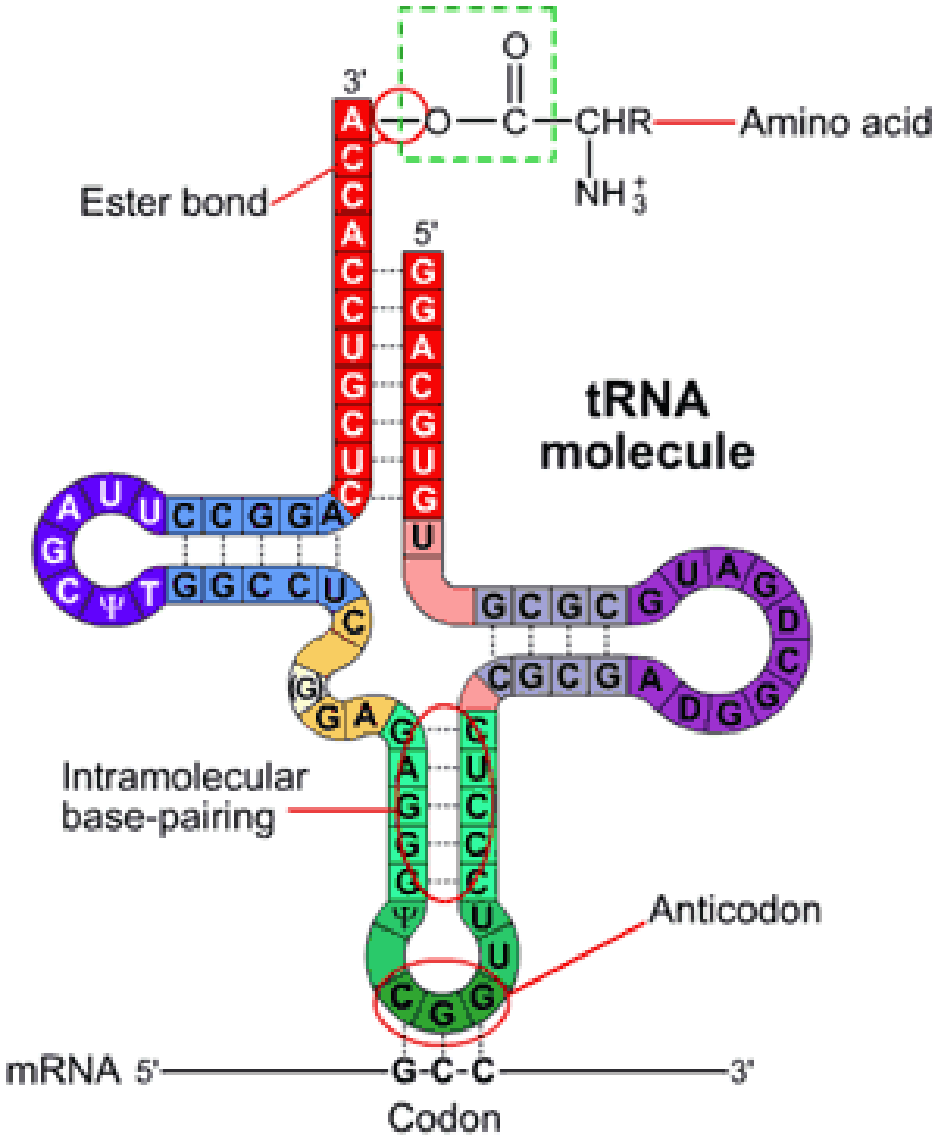




RNA
Ribonucleic acid

DNA
Deoxyribonucleic acid

From mRNA to tRNA and Amino acid



The Genetic Code

Triple of **4** **nucleotides**
(adenine (A), guanine (G), uracile (U), cytosine (C))
➔ **64 codons**

Number of amino-acids = 20

GENETIC CODE is degenerate

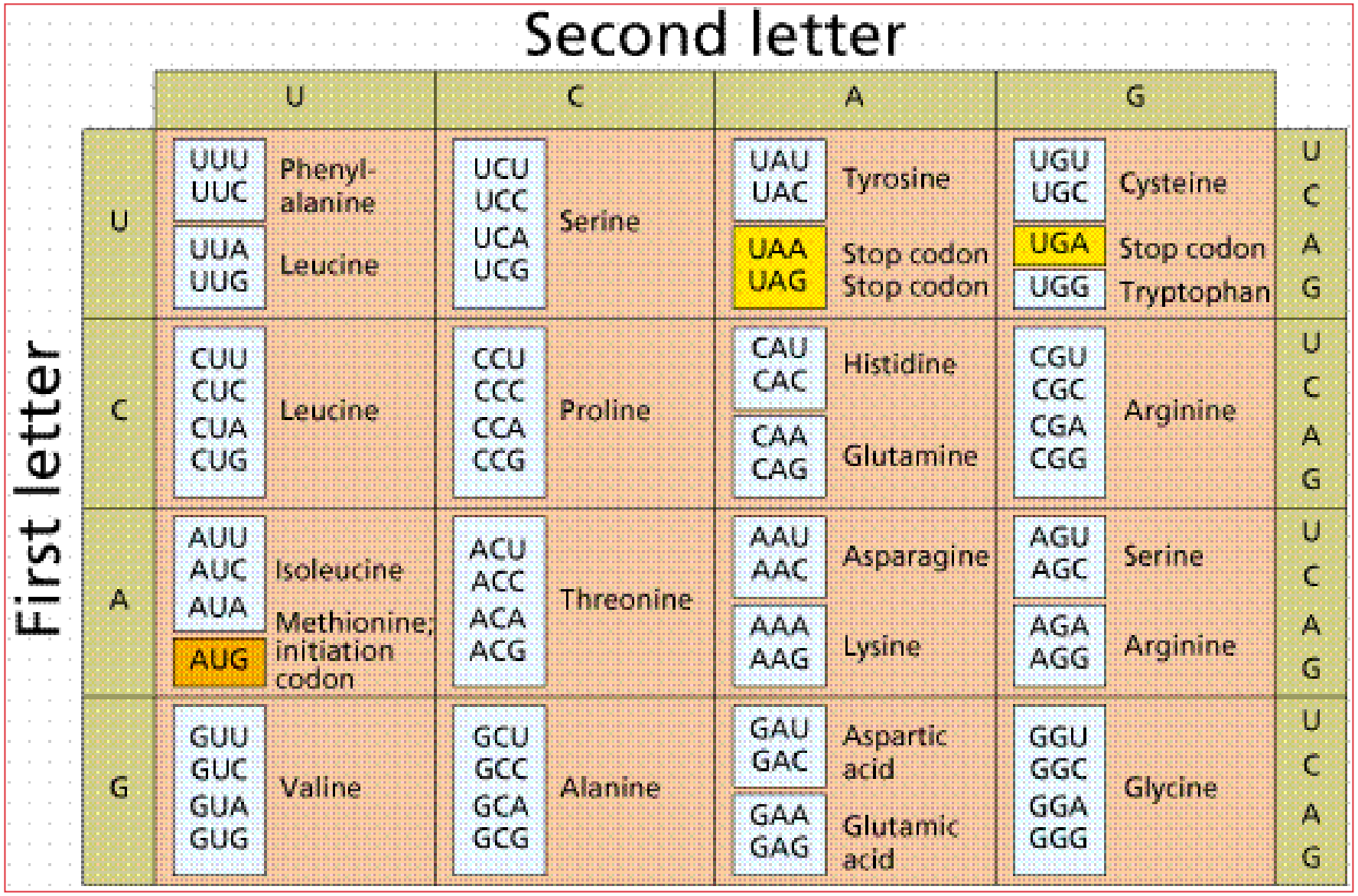
Standard code

61 codons code for 20 amino-acids + 3 Stop

The correspondence between codons and amino-acids is by multiplets (of synonymous codons)

n. 3 sextets
n. 5 quartets
n. 2 triplets
n. 9 doublets
n. 2 singlets

Standard Genetic Code



2- Crystal Basis Model of G.C.

4 bases : purines : (A, G) *and* bases
pyrimidines : $(C, T/U)$ complementarity

$\Rightarrow (\frac{1}{2}, \frac{1}{2})$ representation of $SU(2) \times SU(2)$.

$$
\begin{array}{ccc}
 & \overset{su(2)_H}{\longleftrightarrow} & \\
C \equiv (+,+) & & U \equiv (-,+) \\
su(2)_V \updownarrow & & \updownarrow su(2)_V \\
G \equiv (+,-) & \longleftrightarrow & A \equiv (-,-) \\
 & \underset{su(2)_H}{} &
\end{array}
$$

Analogy	between	quark (q)	and	baryon ($3q$)
	and	base (b)	and	codon ($3b$).

But :

$$|p> \sim |uud> + |udu> + |duu> \text{ (implicit spin structure)}$$

while:

$$UAG \neq AUG \text{ in codons.} \quad \textit{no mixing}$$

$\Longrightarrow$ Limit of the quantum (deformed) algebra
$\mathcal{U}_q[sl(2) \oplus sl(2)]$ when $q \to 0$

(remind: $q \to 1$ usual $\mathcal{U}[sl(2) \oplus sl(2)]$).
Then:

tensorial product of representations
=
"pure" states of constituent states
in **crystal bases**.

$U_q(sl(2))$

$$[J_z, J_\pm] = \pm J_\pm$$

$$[J_+, J_-] = \frac{q^{J_z} - q^{-J_z}}{q^{1/2} - q^{-1/2}}$$

$q = 1 \Rightarrow sl(2)$

$q = 0 \Rightarrow$ Crystal Basis

In our model: a codon $\in (\frac{1}{2},\frac{1}{2}) \otimes (\frac{1}{2},\frac{1}{2}) \otimes (\frac{1}{2},\frac{1}{2})$.

product of two representations

$$(\tfrac{1}{2},\tfrac{1}{2}) \otimes (\tfrac{1}{2},\tfrac{1}{2}) = (1,1) \oplus (1,0) \oplus (0,1) \oplus (0,0)$$

$\rightarrow$ $su(2)_H$ $(J_H, J_V) = (0,0)$ (CA) $(1,0)$ $(CG \quad UG \quad UA)$

$\downarrow$

$su(2)_V$ $(0,1)$ $\begin{pmatrix} CU \\ GU \\ GA \end{pmatrix}$ $(1,1)$ $\begin{pmatrix} CC & UC & UU \\ GC & AC & AU \\ GG & AG & AA \end{pmatrix}$

Property:

$$Q = T_{3H} + \frac{1}{4} C_V (T_{3V} + \Omega - \frac{1}{4})$$

quadruplets (as well as those in sextets) s.t. $J_{3,H} > 0$ or $J_{3,H} = 0$ and $J_{3,V} \geq 0$, $J_V \neq 0$

doublets (and others: triplet, singlets) s.t. $J_{3,H} < 0$ or $J_{3,H} = 0$ and $J_{3,V} < 0$ or $J_V = 0$

product of three representations = codons

$$(\tfrac{1}{2},\tfrac{1}{2})\otimes(\tfrac{1}{2},\tfrac{1}{2})\otimes(\tfrac{1}{2},\tfrac{1}{2}) = (\tfrac{3}{2},\tfrac{3}{2})\oplus 2\,(\tfrac{3}{2},\tfrac{1}{2})\oplus 2\,(\tfrac{1}{2},\tfrac{3}{2})\oplus 4\,(\tfrac{1}{2},\tfrac{1}{2})$$

$$(\tfrac{3}{2},\tfrac{3}{2}) \equiv \begin{pmatrix} CCC & UCC & UUC & UUU \\ GCC & ACC & AUC & AUU \\ GGC & AGC & AAC & AAU \\ GGG & AGG & AAG & AAA \end{pmatrix}$$

$$(\tfrac{3}{2},\tfrac{1}{2}) \equiv \begin{pmatrix} CCG & UCG & UUG & UUA \\ GCG & ACG & AUG & AUA \end{pmatrix}$$

$$(\tfrac{3}{2},\tfrac{1}{2})' \equiv \begin{pmatrix} CGC & UGC & UAC & UAU \\ CGG & UGG & UAG & UAA \end{pmatrix}$$

$$(\tfrac{1}{2},\tfrac{3}{2}) \equiv \begin{pmatrix} CCU & UCU \\ GCU & ACU \\ GGU & AGU \\ GGA & AGA \end{pmatrix} \qquad (\tfrac{1}{2},\tfrac{3}{2})' \equiv \begin{pmatrix} CUC & CUU \\ GUC & GUU \\ GAC & GAU \\ GAG & GAA \end{pmatrix}$$

$$(\tfrac{1}{2},\tfrac{1}{2}) \equiv \begin{pmatrix} CCA & UCA \\ GCA & ACA \end{pmatrix} \qquad (\tfrac{1}{2},\tfrac{1}{2})' \equiv \begin{pmatrix} CGU & UGU \\ CGA & UGA \end{pmatrix}$$

$$(\tfrac{1}{2},\tfrac{1}{2})'' \equiv \begin{pmatrix} CUG & CUA \\ GUG & GUA \end{pmatrix} \qquad (\tfrac{1}{2},\tfrac{1}{2})''' \equiv \begin{pmatrix} CAC & CAU \\ CAG & CAA \end{pmatrix}$$

Cod.	a.a.	J_H	J_V	Cod.	a.a.	J_H	J_V
CCC	Pro	3/2	3/2	UCC	Ser	3/2	3/2
CCU	Pro	(1/2	3/2)[1]	UCU	Ser	(1/2	3/2)[1]
CCG	Pro	(3/2	1/2)[1]	UCG	Ser	(3/2	1/2)[1]
CCA	Pro	(1/2	1/2)[1]	UCA	Ser	(1/2	1/2)[1]
CUC	Leu	(1/2	3/2)[2]	UUC	Phe	3/2	3/2
CUU	Leu	(1/2	3/2)[2]	UUU	Phe	3/2	3/2
CUG	Leu	(1/2	1/2)[3]	UUG	Leu	(3/2	1/2)[1]
CUA	Leu	(1/2	1/2)[3]	UUA	Leu	(3/2	1/2)[1]
CGC	Arg	(3/2	1/2)[2]	UGC	Cys	(3/2	1/2)[2]
CGU	Arg	(1/2	1/2)[2]	UGU	Cys	(1/2	1/2)[2]
CGG	Arg	(3/2	1/2)[2]	UGG	Trp	(3/2	1/2)[2]
CGA	Arg	(1/2	1/2)[2]	UGA	Trp	(1/2	1/2)[2]
CAC	His	(1/2	1/2)[4]	UAC	Tyr	(3/2	1/2)[2]
CAU	His	(1/2	1/2)[4]	UAU	Tyr	(3/2	1/2)[2]
CAG	Gln	(1/2	1/2)[4]	UAG	Ter	(3/2	1/2)[2]
CAA	Gln	(1/2	1/2)[4]	UAA	Ter	(3/2	1/2)[2]
GCC	Ala	3/2	3/2	ACC	Thr	3/2	3/2
GCU	Ala	(1/2	3/2)[1]	ACU	Thr	(1/2	3/2)[1]
GUG	Ala	(3/2	1/2)[1]	ACG	Thr	(3/2	1/2)[1]
GCA	Ala	(1/2	1/2)[1]	ACA	Thr	(1/2	1/2)[1]
GUC	Val	(1/2	3/2)[2]	AUC	Ile	3/2	3/2
GUU	Val	(1/2	3/2)[2]	AUU	Ile	3/2	3/2
GUG	Val	(1/2	1/2)[3]	AUG	Met	(3/2	1/2)[1]
GUA	Val	(1/2	1/2)[3]	AUA	Ile	(3/2	1/2)[1]
GGC	Gly	3/2	3/2	AGC	Ser	3/2	3/2
GGU	Gly	(1/2	3/2)[1]	AGU	Ser	(1/2	3/2)[1]
GGG	Gly	3/2	3/2	AGG	Arg	3/2	3/2
GGA	Gly	(1/2	3/2)[1]	AGA	Arg	(1/2	3/2)[1]
GAC	Asp	(1/2	3/2)[2]	AAC	Asn	3/2	3/2
GAU	Asp	(1/2	3/2)[2]	AAU	Asn	3/2	3/2
GAG	Glu	(1/2	3/2)[2]	AAG	Lys	3/2	3/2
GAA	Glu	(1/2	3/2)[2]	AAA	Lys	3/2	3/2

- ***Applications*** *of this model provided in a series of papers (L.Frappat, A.Sciarrino and P.S.) in the years 1998-2005. Among them:*
- *Study of codon usage probabilities, elaboration of sum rules.*
- *Relations between physico-chemical properties of amino-acids (a.a.) and predictions.*
- *More mathematical aspects: operator relating a.a. and codons for any known genetic code; attempts to describe mutations,etc.*

3- Codon-anticodon Interaction

Position of the problem:

codon: XYZ ---- anti-codon: Z'Y'X'

with nucleotids Z', Y', X' associated to Z, Y, X

- In tRNA process, codon –anticodon pairing does not follow the usual Watson-Crick pattern (i.e. pairing C --- G , U --- A).

- This leads Crick (1966) to propose the wobble hypothesis :

A specified anti-codon can recognize more than one codon differing only in the third nucleotide.

i.e. standard pairing for X--X' and Y—Y' while Z' may pair to different Z.

- Two main hypotheses proposed in this context:
  1- for doublets, the first nucleotide Z' in anticodon should have G (resp. U) to read for codon with Y (resp. R) in third position Z.
  2- the chosen anticodon is the one with first position nucleotide pairing the (third position of the) most abundant codon among synonymous codons.

(Y= C, U pyrimidine, R= G,A purine)

Considering the Mitochondrial Code:

there are: 2 sextets,
6 quadruplets
12 doublets of codons specifying the 20 amino-acids.

So, a minimum number of 22 anticodons is needed.

And this appears to be the case in mitochondria of animals (*Sprinzl et al., 1998*)

Data seem to confirm the empirical rule just above.

It is this set of data that we will consider now in the framework of the Crystal Basis Model.

The Minimum Principle

Consider the operator:

$$T\,(\text{anticodon, codon}) = c_H \vec{J}^{c}_{H} . \vec{J}^{a}_{H} + c_V \vec{J}^{c}_{V} . \vec{J}^{a}_{V}$$

where : $\vec{J}^{c} . \vec{J}^{a} = \frac{1}{2} \{ (\vec{J}^{c} + \vec{J}^{a})^2 - (\vec{J}^{c})^2 - (\vec{J}^{a})^2 \}$

and: $\vec{J} = (J_1, J_2, J_3)$ generators of $su(2)_{\substack{H \\ V}}$ group.

Then define:

- for quadruplets: taking as an example Val (GUN; N= C,U,G,A) and as a possible anticodon CAC:

$$T\text{aver.}\,(\text{CAC, Val}) = \sum_{N} P^{val}_{N} . T\,(\text{CAC, GUN})$$

with : $P^{q}_{C} + P^{q}_{U} + P^{q}_{G} + P^{q}_{A} = 1$

- for doublets : taking as an ex. Asp (GAC, GAU) and as a possible anticodon CUC:

$$T_{aver.}(CUC, Asp) = \sum_{Y}^{Asp} P_Y . T(CUC, GUY)$$

with : $P_Y^d = P_C^d + P_U^d = 1$ (and $P_R^d = P_G^d + P_A^d = 1$)

Question:

Can we determine c_H and c_V such that for each given quadruplet (or doublet) of codons, the anticodon minimizing $T_{aver.}$ is the one given by the data ?

Let us remind that the possible anticodons to the codon XY N is N'Y'X' with X -- X' and Y --Y' related by the "usual pairing" (i.e. C -- G, U -- A) and N' is any nucleotide C,G,U, A).

In Vertebral Mitochondrial Code:

most used anticodons for mitochondria of animals (*Sprinzl et al, 1998*)

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	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{1}{2}$	$\frac{3}{2}$	
CUU	L	$(\frac{1}{2}\ \frac{3}{2})^2$		$-\frac{1}{2}$	$\frac{3}{2}$		UUU	F	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{3}{2}$	$\frac{3}{2}$	*GAA*
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*

Results:

- for quadruplets: choose simply $c_H > 0$ and $c_V < 0$ to be in accordance with data.

-for doublets: choose $c_V > 0$
and the sign of c_H such that:

c_H is $>$ for the doublets : UUY, UAY, AUY, AAY
CAR, UGR, AGR, GAR

c_H is $<$ for the other doublets: UUR, UAR, AUR, AAR
CAY, UGY, AGY, GAY

(c_H of opposite sign for two doublets with same dinucleotide but ending with a purine or a pyrimidine).

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	

Conclusion

- Anticodons minimizing the conjectured operator T_{aver} in very good agreement with the observed ones for mitochondria of animals.
- Results depending only of the sign of two coupling constants.
- One may expect a more complicated pattern in the general case, following the biological species. Might happen that the « universal » feature of C_V and C_H should be released and T expressions modified.
- Then the crystal basis offers a lot of possibilities, for ex. adjunction of a term of « spin-spin » interaction of the type:

$$g_H \, J^{c}_{H,3} \cdot J^{a}_{H,3} + g_V \, J^{c}_{V,3} \cdot J^{a}_{V,3}$$

4- Evolution of the Genetic Code

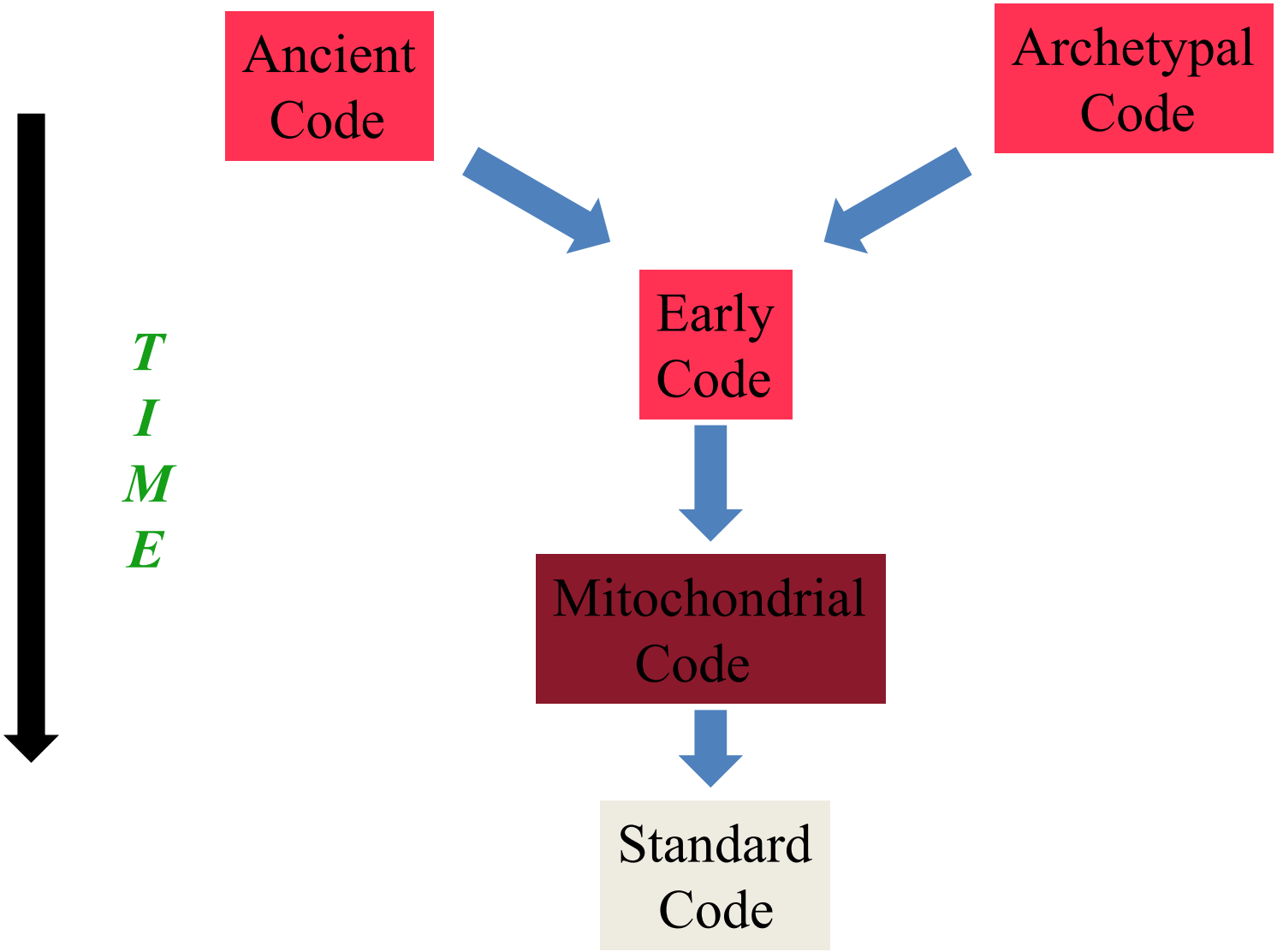
The genetic code has undergone an evolutionary process

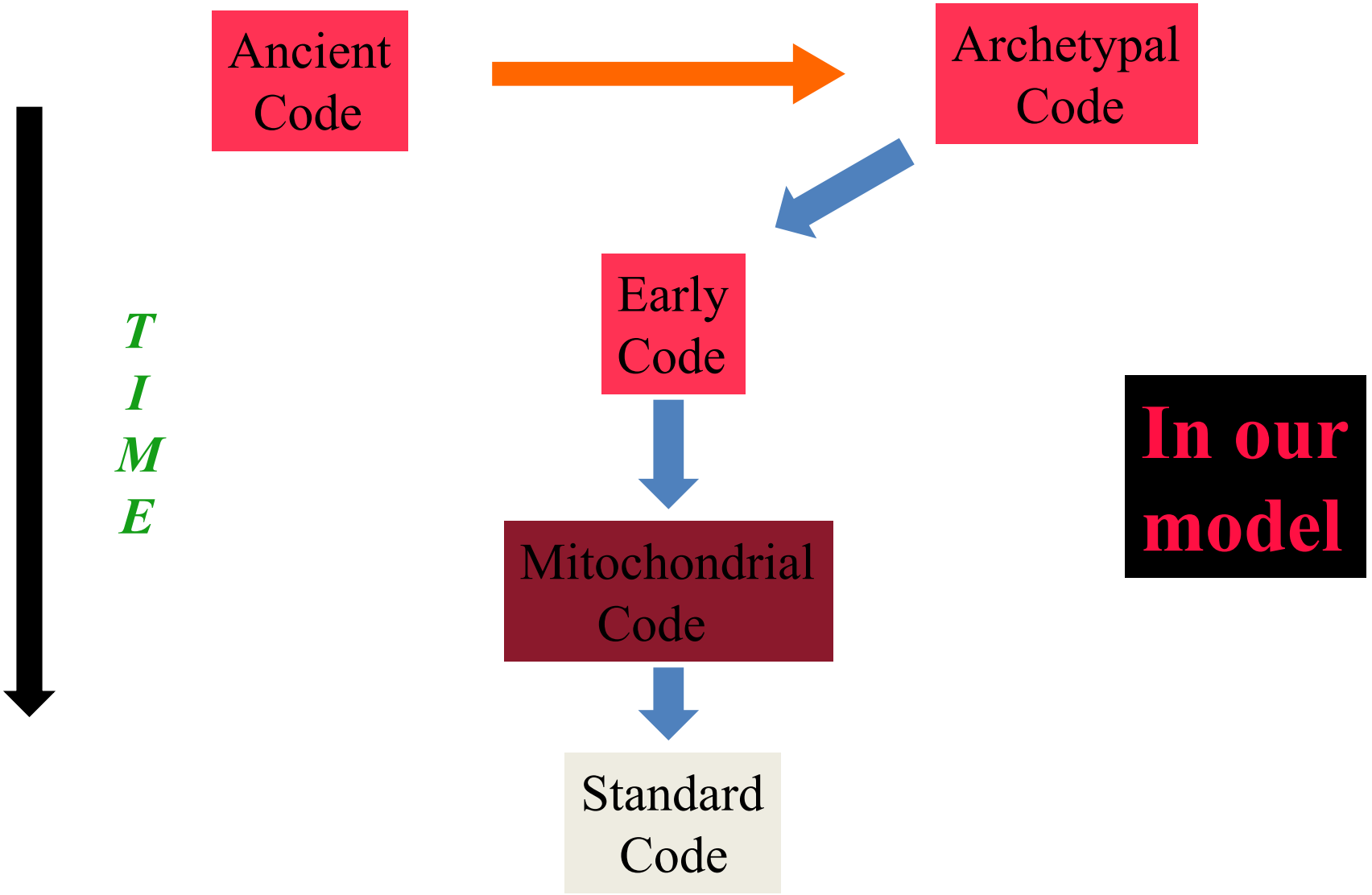
- An evolutionary theory is the « codon capture theory »
(review Ohama et al. (2008)):

. number of the encoded amino acids is kept constant,
equal to 20.

. the coding codons change, key role in this process being played by the anticodon.

In this context, 3 main codes: the Ancient and Early Codes, and an alternative: the Archetypal Code.


Ancient Code
Archetypal Code
Early Code
Mitochondrial Code
Standard Code
TIME


Ancient Code
Archetypal Code
Early Code
Mitochondrial Code
Standard Code
TIME
In our model

A comparison and a proposition

Differences:

- In Ancient code, 20 codons and 20 anti-codons.
- In Early code, all codons involved and 20 anti-codons.
- In Archetypal code, all codons but 16 anticodons .

A proposal: **Minimisation Conjecture** in following way:

- **T (**anticodon-codon**)** in Ancient Code :
with Watson-Crick pairing between nucleotids (i.e. C-G , U-A)

-**T aver(**anticodon-codon**)** in Archetypal and Early Codes
with « wobble » mechanism.

Ancient code

in bold the differences with Ohama(2008)

a.a	codon	anticodon
Pro	CCG	CGG
Leu	**CUA**	**UAG**
Arg	CGG	CCG
Ala	GCG	CGC
Val	**GUA**	**UAC**
Gly	GGG	CCC
Ser	UCG	CGA
Thr	ACG	CGU
His	CAC	GUG
Gln	CAG	CUG
Phe	UUC	GAA
Cys	**UGU**	**ACA**
Trp	**UGA**	**UCA**
Tyr	UAC	GUA
Asp	GAC	GUC
Glu	GAG	CUC
Ile	AUC	GAU
Met	AUG	CAU
Asn	AAC	GUU
Lys	AAG	CUU

Table 2: The couple of codon-anticodon which minimizes the operator $\mathcal{T}$ with $c_V < 0$ and $c_H < 0$ for the strong dinucleotides (first 8 rows) and c_V underdetermined $c_H > 0$ for the weak dinucleotides in the Ancient Genetic Code.

Archetypal code

a.a	codon	anticodon	sign c_H	sign c_V	note
Pro	CCN	UGG	+	-	
Leu	CUN	UAG	+	-	
Arg	CGN	UCG	+	-	
Ala	GCN	UGC	+	-	
Val	GUN	UAC	+	-	
Gly	GGN	UCC	+	-	
Ser	UCN	UGA	+	-	
Thr	ACN	UGU	+	-	
His/Gln	CAN	UUG	+	-	$P_S > 1/4$
Phe/Leu'	UUN	UAA	-	-	
Cys/Trp	UGN	UCA	+	+	
Tyr	UAY	GUA	+	+	
Asp/Glu	GAN	UUC	+	-	$P_S > 1/4$
Ile/Met	AUN	UAU	-	-	$P_Y > 1/8$
Asn/Lys	AAN	UUU	-	-	
Ser'/Arg'	AGN	UCU	+	-	

Table 3: Sign of coupling constants minimizing the operator $\mathcal{T}$, averaged over the codons, for any amino acid encoded in the Archetypal Genetic Code. We denote by a prime the a.a. encoded by the sub-part of the sextet corresponding to a doublet.

Early Code

a.a	codon	anticodon	sign c_H	sign c_V	note
Pro	CCN	UGG	+	-	
Leu	CUN	UAG	+	-	$P_S > 1/4$
Arg	CGN	UCG	+	-	
Ala	GCN	UGC	+	-	
Val	GUN	UAC	+	-	$P_S > 1/4$
Gly	GGN	UCC	+	-	
Ser	UCN	UGA	+	-	
Thr	ACN	UGU	+	-	$P_Y > 1/8$
His	CAY	GUG	+	+	$P_C > 3/8$
Gln	CAR	UUG	-	und.	$P_G < 1/4$
Phe	UUY	GAA	+	+	
Leu'	UUR	UAA	-	und.	
Cys	UGY	GCA	-	+	
Trp	UGR	**UCA**	+	und.	
Tyr	UAY	GUA	+	+	
Asp	GAY	GUC	+	+	$P_C > 1/4$
Glu	GAR	UUC	-	und.	$P_G < 1/4$
Ile	AUY	GAU	+	+	
Met	AUR	UAU	-	und.	
Asn	AAY	GUU	+	+	
Lys	AAR	UUU	-	und.	
Ser'	AGY	GCU	-	+	
Arg'	AGR	UCU	+	und.	

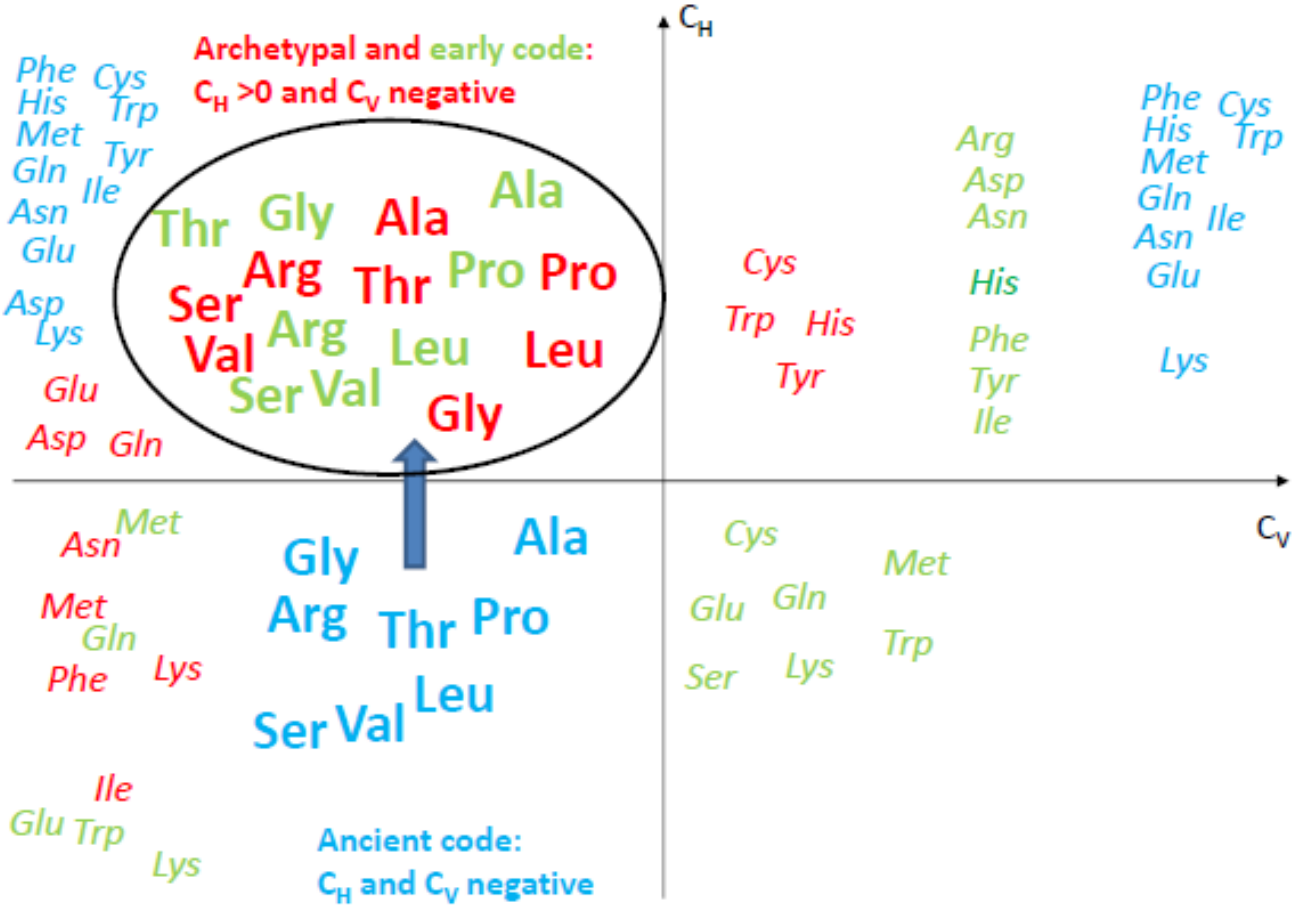
Table 4: Sign of coupling constants minimizing the operator $\mathcal{T}$, averaged over the codons, for any amino acid encoded in the Early Genetic Code.

Remark:

« evolution » of the signs of the Constants:

example of Asn/Lys

$$c_H^{AAN} > 0 \implies c_H^{AAN} < 0 \implies \begin{cases} c_H^{AAY} > 0 \\ \\ c_H^{AAR} < 0 \end{cases}$$



Strong di nucleotides evolution of the constant: bold

Weak di nucleotide constant evolution: italic

Issues with a.a. Leu, Ser and Arg.

Blue text: ancient code
Red text: archetypal code
And early code (strong dinucleotide)
Green text: early code for weak nucleotide

Conclusion

- *The « constants » C_V and C_H deserve a more precise study. Their branching points would correspond to the advent of different Genetic Codes, with the standard G.C. emerging as the one exhibiting selective advantages.*
- *Could we verify that the existing G.C., that is the branching point which has survived, satifies the required optimality conditions?*

Summary

- *The evolution of the Genetic Code is analysed from the view-point of the **codon-anticodon interaction**.*
- *Imposing a **Minimum principle for the Interaction**, we deermine, in the framework of the **Crystal Basis Model**, the structure of anticodons in the Ancient, Archetypal and Early Genetic Codes, that are all reconciled in an unique frame.*
- *The above obtained results, joined to previous ones, encourage us to introduce the notion of **« BIO-SPIN »** related to the Uq(SU(2)+SU(2)) group of our model.*