Prediction of Physical-Chemical Properties of Amino Acids from Genetic Code

L. FRAPPAT^{1,3}, A. SCIARRINO² and P. SORBA¹

¹*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*

²*Dipartimento di Scienze Fisiche, Università di Napoli 'Federico II', and I.N.F.N., Sezione di Napoli, Complesso Universitario di Monte S. Angelo, Via Cintia, I-80126 Napoli, Italy*

³*Member of Institut Universitaire de France*

Abstract. Using the crystal basis model of the genetic code, a set of relations between the physical-chemical properties of the amino acids are derived and compared with the experimental data. A prediction for the not yet measured thermodynamical parameters of three amino acids is done.

Key words: PACS number: 87.10.+e, 02.10.-v

1. Introduction

It is a known observation [1] that a relationship exists between the codons and the physical-chemical properties of the coded amino acids. The observed pattern is read either as a relic of some kind of interaction between the amino acids and the nucleotides at an early stage of evolution or as the existence of a mechanism relating the properties of codons with those of amino acids. It is also observed that the relationship depends essentially on the nature of the *second* nucleotide in the codons and it holds when the second nucleotide is A, U, C, not when it is G. To our knowledge neither the anomalous behaviour of G nor the existence of a closest relationship between some of the amino acids is understood. In this paper we provide an explanation of both these facts in the framework of the crystal basis model of genetic code we have discussed in [2]. After recalling briefly the model in Sect. 2, we derive in Sect. 3 a set of relations between the physical-chemical properties of the amino acids, based on the content of the irreducible representations of the dinucleotides and of the codons coding for the amino acids. Finally we compare our predictions with the experimental data.

2. The Model

The four nucleotides A, C, G, U are considered as basic states of the $(\frac{1}{2}, \frac{1}{2})$ irreducible representation of the quantum enveloping algebra $\mathcal{U}_q[sl(2) \oplus sl(2)]$ in the limit $q \rightarrow 0$, with the following assignment of quantum numbers:

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

In this framework, the codons are built as composite states of the nucleotide states by tensoring three such $(\frac{1}{2}, \frac{1}{2})$ representations. The crystal basis, which exists in the limit $q \rightarrow 0$ of the quantum enveloping algebra $\mathcal{U}_q(G)$ for any semi-simple Lie algebra G , is the only way to provide such composite states as *pure* states, and to ensure the existence of an *order* in the tensoring procedure. However, in a multiplet of codons relative to a specific amino acid, the first two constituents of a codon are ‘relatively stable’, the degeneracy being mainly generated by the third nucleotide. Considering the tensor product $(\frac{1}{2}, \frac{1}{2})^{\otimes 2}$ and using Kashiwara’s theorem [3], we get (the subscripts H ($:=$ horizontal) and V ($:=$ vertical) specify the two $sl(2)$):

$$\begin{array}{ccccc} \rightarrow & sl(2)_H & (0, 0) & (CA) & (1, 0) \quad (CG \quad UG \quad UA) \\ \downarrow & & & & \\ & sl(2)_V & (0, 1) & \begin{pmatrix} CU \\ GU \\ GA \end{pmatrix} & (1, 1) \quad \begin{pmatrix} CC & UC & UU \\ GC & AC & AU \\ GG & AG & AA \end{pmatrix} \end{array} \quad (2)$$

The dinucleotide states formed by the first two nucleotides in a codon can be put in correspondence with multiplets of codons relative to an amino acid [2]. We define in our model the ‘charge’ Q of a dinucleotide state by $Q = J_{3,H} + \frac{1}{4} C_V (J_{3,V} + 1) - \frac{1}{4}$ (the values of $J_{3,H}$ and $J_{3,V}$ are directly read in eq. (2) and $C_V = J_V (J_V + 1)$). The dinucleotide states are then split into two octets with respect to the charge Q : the eight *strong* dinucleotides CC, GC, CG, GG, CU, GU, UC, AC associated to quadruplets of codons satisfy $Q > 0$, while the eight *weak* dinucleotides AA, AU, UA, UU, UG, AG, CA, GA associated to doublets or singlets of codons satisfy $Q < 0$. Remark that by the change $C \leftrightarrow A$ and $U \leftrightarrow G$, equivalent to the change of the sign of $J_{3,\alpha}$ or to reflexion with respect to the diagonals of the eq. (1), the 8 strong dinucleotides are transformed into weak ones and vice-versa. A first attempt to differentiate between strong and weak dinucleotides is contained in [4].

3. Relationship between the Physical-Chemical Properties of Amino Acids

We assume that some physical-chemical property of a given amino acid are related to the nature of the codons, in particular they depend on the following mathematical features, written in hierarchical order:

1. the irreducible representation of the dinucleotide formed by the first two nucleotides;
2. the sign of the charge Q on the dinucleotide state;

3. the value of the third component of $J_{3,V}$ inside a fixed irreducible representations for the dinucleotides;
4. the upper label(s) of the codon irreducible representation(s);

Not all the physical-chemical properties are supposed to follow the scheme above; some of them are essentially given by the specific chemical structure of the amino acid itself. In the following, we analyze the physical-chemical properties of the amino acids in the light of the dinucleotide content of the irreducible representations of eq. (2).

– Representation (0,0): the codons of the form CAN (N = C, U, G, A) all belong to the irreducible representation $(\frac{1}{2}, \frac{1}{2})^4$ and code for His and Gln, both being coded by doublets and differing by the value of $J_{3,V}$. Then we expect that the physical-chemical properties of His and Gln are very close.

– Representation (1,0): we analyze the codons CG ($Q > 0$), UG, UA (both $Q < 0$). The codons CGS (S = C, G), resp. CGW (W = U, A), belonging to irreducible representation $(3/2, 1/2)^2$, resp. $(1/2, 1/2)^2$, all code for Arg, so we do not have any relation. The codons UGS, resp. UGW, belonging to irreducible representation $(3/2, 1/2)^2$, resp. $(1/2, 1/2)^2$, code for Cys and Trp, resp. the other Cys and Ter. So we expect some affinity between the physical-chemical properties of Cys and Trp, not very strong indeed as the former is encoded by a doublet and the latter by a singlet. The codons UAN, belonging to the irreducible representation $(3/2, 1/2)^2$, code for the Tyr and Ter. So we expect some affinity between the amino acids coded by UGN and UAN, in particular between Cys and Tyr both being coded by doublets.

– Representation (0,1): we analyze the codons CU, GU (both $Q > 0$) and GA ($Q < 0$). The codons CUY and GUY (Y = C, U), resp. CUR and GUR (R = G, A), belonging to irreducible representation $(1/2, 3/2)^2$, resp. $(1/2, 1/2)^3$, code for Leu and Val. Therefore we do not have any relation between amino acids coded by the same dinucleotide, but we expect that the physical-chemical properties of Leu and Val are close since CU and GU both belong to the same irreducible representation and are both strong. The codons GAN belong to the irreducible representation $(1/2, 3/2)^2$ and they code Asp and Glu (both doublets). Then we expect the physical-chemical properties of Asp and Glu to be very close.

– Representation (1,1): the dinucleotide irreducible representation (1, 1) contains five states with $Q > 0$ (CC, UC, GC, AC, GG). The codons CCN and UCN (resp. GCN and ACN) belong to four different irreducible representations and code for Pro and Ser (resp. Ala and Thr). We expect a strong affinity between the physical-chemical properties of Pro and Ser on the one hand and between the physical-chemical properties of Ala and Thr on the other hand. The codons GGN belong to two different irreducible representations and code for Gly, so we expect an affinity of physical-chemical properties of Gly with those of Pro, Ser, Ala, Thr. Now let us look at the four states with $Q < 0$ (UU, AU, AG, AA). The codons UUN belong to two different irreducible representations and code for Leu, the doublet

subpart of the sextet, and for Phe (doublet). An affinity is expected between the physical-chemical properties of these two amino acids. The codons AUN belong to two different irreducible representations and code Ile (triplet) and Met (singlet) and, in fact, the values of physical-chemical properties of these two amino acids are not very different. The codons AGN belong to two different irreducible representations and code for Ser and Arg, the doublet subpart of the sextet, so an affinity between the physical-chemical properties of these codons is expected. The codons AAN belong to the same irreducible representation ($3/2, 3/2$) and code for Asn and Lys, so the values of the physical-chemical properties of these amino acids should be close.

Note that for the three sextets (Arg, Leu, Ser) the quartet (doublet) subpart is coded by a codon with a strong (weak) dinucleotide.

4. Discussion

We have compared our theoretical predictions with 10 physical-chemical properties:

1. the Chou-Fasman conformational parameters [5] P_α , P_β and P_τ which gives a measure of the probability of the amino acids to form respectively a helix, a sheet and a turn. The sum $P_\alpha + P_\beta$ appears more appropriate to characterize the generic structure forming potential and the difference $P_\alpha - P_\beta$ the helix forming potential, this quantity depending more on the particular amino acid [6]. So we compare with $P_\alpha + P_\beta$ and P_τ ;
2. the Grantham polarity P_G [7];
3. the relative hydrophilicity R_f as computed in [8];
4. the thermodynamic activation parameters at 298 K: ΔH (enthalpy, in kJ/mol), ΔG (free energy, in kJ/mol) and ΔS (entropy, in J/mole/K) as obtained in [9];
5. the negative of the logarithm of the dissociation constants at 298 K: pK_a for the α -COOH group and pK_b for the α -NH₃⁺ group [10];
6. the isoelectric point pI [10], i.e. the pH value at which no electrophoresis occurs.

The comparison between the theoretical relations and the experimental values shows, see Tables I and III ($\cong$ means strong affinity, $\approx$ affinity, $\sim$ weak affinity):

1. His $\cong$ Gln – The agreement, except for pI , is very good.
2. Asp $\cong$ Glu – The agreement, except for P_τ , is very good.
3. Asn $\cong$ Lys $\sim$ Arg, Ser – The agreement, except for pI and P_τ is very good.
The comparison with the values of physical-chemical properties of Ser and Arg is satisfactory.
4. Cys $\cong$ Tyr $\approx$ Trp – Except for R_f , the agreement between the first two amino acids is very good, while with Trp is satisfactory.
5. Leu $\cong$ Val – The agreement is very good.
6. Pro $\cong$ Ser $\approx$ Gly – The agreement is very good, except for $P_\alpha + P_\beta$ and ΔH , and with Gly more than satisfactory.

7. Ala $\cong$ Thr $\approx$ Gly, Pro, Ser – The agreement is very good between the first two amino acids except for P_r and satisfactory with the others except for the conformational parameters.
8. Ile $\cong$ Met $\approx$ Phe – The agreement is very good between the first two amino acids and satisfactory with Phe.

In order to have a more quantitative evaluation of the agreement between the data and the theoretical model, we compute in 3 different cases, the mean value $\bar{x}$ and the standard deviation σ_x of the whole population (i.e. the twenty amino-acids) for each generic physical-chemical property x :

1. considering all the amino acids, i.e. summing over the whole set of amino acids ($n = 20$);
2. considering the amino acids coded by the codons with the same second nucleotide ($n = 4$ for C, $n = 5$ for U and G, $n = 7$ for A);
3. considering the couples ($n = 2$) of amino acids given by our 8 relations.

The results are reported in Tables I to III. As an estimate of the accuracy of our predictions, we compute the sum of the adimensional quantities given for any physical-chemical property by the ratios $\eta = \sigma_x/\bar{x}$ of the standard deviation to the average value, in the three different cases considered above.

all				1.82			
A	1.94	C	1.13	G	1.82	U	0.70
Gln/His	Asp/Glu	Asn/Lys	Cys/Tyr	Pro/Ser	Ala/Thr	Ile/Met	Leu/Val
0.35	0.63	0.72	0.80	0.77	0.48	0.39	0.38

$$\text{Values of } \eta = \sum \sigma_x/\bar{x}$$

We remark in the above table that the characterization of the physical-chemical properties from the nature of the second nucleotide is, except for U, indeed not really discriminatory. On the contrary, the relations derived by our model reduces dramatically the values of η .

From Table I we note that the values of the thermodynamical quantities for some amino acids are not yet measured. Our relations predict that the values of ΔH and $-\Delta S$ for His should be around 60 kJ/mol and 120 kJ/mol/K. For the values of Asp and Glu coded by the unique weak nucleotide GA in the irreducible representation (0,1), we do not have any available data to compare with. However, making the further assumption, supported by an inspection of the available data, that the values of the thermodynamical quantities depend slightly on the nature of the charge, we predict that the values of these quantities for Asp and Glu should not be very dissimilar from the values of the analogous quantities for Leu and Val, coded by the strong dinucleotide CU and GU in the irreducible representation (0,1). So we predict that for Asp and Glu, one should find $\Delta H \approx 60$ kJ/mol, $-\Delta S \approx 135$ kJ/mol/K and $\Delta G \approx 100$ kJ/mol.

Table I. Table of amino-acid properties

Amino acid	$P_\alpha + P_\beta$	P_τ	P_G	R_f	ΔH	$-\Delta S$	ΔG	pK_a	pK_b	pI
Ala	2.25	0.66	8.09	0.89	50.7	147.8	94.66	2.34	9.69	6.00
Arg	1.91	0.95	10.50	0.88	54.8	143.2	97.5	2.17	9.04	11.15
Asn	1.60	1.56	11.5	0.89	55	135.8	95.51	2.02	8.80	5.41
Asp	1.55	1.46	13	0.87	—	—	—	1.88	9.60	2.77
Cys	1.89	1.19	5.5	0.85	43.9	163.2	92.89	1.96	10.28	5.02
Gln	2.21	0.98	10.5	0.82	61.1	121.8	97.30	2.17	9.13	5.65
Glu	1.88	0.74	12.2	0.84	—	—	—	2.19	9.67	3.22
Gly	1.32	1.56	9	0.92	49.0	140.6	91.08	2.34	9.60	5.97
His	1.87	0.95	10.4	0.83	—	—	94.59	1.82	9.17	7.47
Ile	2.68	0.47	5.2	0.76	57.3	152.3	102.6	2.36	9.60	5.94
Leu	2.51	0.59	4.9	0.73	59.4	128.1	97.70	2.36	9.60	5.98
Lys	1.90	1.01	11.3	0.97	57.8	135.6	98.16	2.18	8.95	9.59
Met	2.50	0.60	5.7	0.74	58.2	128.5	96.65	2.28	9.21	5.74
Phe	2.51	0.60	5.2	0.52	48.3	156.1	94.76	1.83	9.13	5.48
Pro	1.12	1.52	8	0.82	50.7	164.5	99.49	1.99	10.60	6.30
Ser	1.52	1.43	9.19	0.96	36.0	180	89.98	2.21	9.15	5.68
Thr	2.02	0.96	8.59	0.92	53.4	136.0	93.98	2.09	9.10	5.64
Trp	2.45	0.96	5.4	0.2	51.1	151.5	96.25	2.83	9.39	5.89
Tyr	2.16	1.14	6.2	0.49	54.4	139.0	95.68	2.20	9.11	5.66
Val	2.76	0.50	5.9	0.85	60.2	139.0	101.7	2.32	9.62	5.96
$\bar{x}$	2.03	0.99	8.31	0.79	53.02	144.88	96.14	2.18	9.42	6.03
σ_x	0.45	0.36	2.61	0.18	6.23	15.57	3.16	0.23	0.43	1.76
$\sigma_x/\bar{x}$	22 %	36 %	31 %	23 %	12 %	10 %	3 %	10 %	5 %	29 %

5. Conclusion

In conclusion, the values of physical-chemical properties show, with a few exceptions, a pattern of correlations which is expected from the assumptions of the crystal basis model. The remarked property that the amino acids coded by codons whose second nucleotide is G do not share similarity in the physical-chemical properties with other amino acids does find an explanation in the model, as it is immediate to verify that there are no two states with G in second position which share simultaneously the properties of belonging to the same irreducible representation and being characterized by the same value of Q .

It has been suggested that the physico-chemical properties of the amino acids have played a fundamental role in the evolution and the organization of the genetic code. In this framework, a measure of the relative distances between amino acids

Table II. Physical-chemical properties of amino-acids coded by codons with same second nucleotide (written at the left)

		$P_\alpha + P_\beta$	P_τ	P_G	R_f	ΔH	$-\Delta S$	ΔG	pK_a	pK_b	pI
A	$\bar{x}$	1.88	1.12	10.73	0.82	57.08	133.05	96.25	2.07	9.20	5.68
	σ_x	0.23	0.27	2.03	0.14	2.65	6.63	1.29	0.15	0.30	2.17
	$\sigma_x/\bar{x}$	12 %	24 %	19 %	17 %	5 %	5 %	1 %	7 %	3 %	38 %
C	$\bar{x}$	1.73	1.14	8.47	0.90	47.70	157.08	94.53	2.16	9.64	5.90
	σ_x	0.44	0.35	0.47	0.05	6.84	16.66	3.38	0.13	0.60	0.28
	$\sigma_x/\bar{x}$	25 %	31 %	6 %	6 %	14 %	11 %	4 %	6 %	6 %	5 %
G	$\bar{x}$	1.82	1.22	7.92	0.76	46.96	155.70	93.54	2.30	9.49	6.74
	σ_x	0.39	0.25	2.08	0.28	6.52	14.48	2.90	0.29	0.44	2.23
	$\sigma_x/\bar{x}$	21 %	20 %	26 %	37 %	14 %	9 %	3 %	13 %	5 %	33 %
U	$\bar{x}$	2.59	0.55	5.38	0.72	56.68	140.80	98.68	2.23	9.43	5.82
	σ_x	0.11	0.06	0.37	0.11	4.31	11.68	3.00	0.20	0.22	0.19
	$\sigma_x/\bar{x}$	4 %	10 %	7 %	15 %	8 %	8 %	3 %	9 %	2 %	3 %

has been defined, see [11] and references therein. The picture which emerges in [11] has some striking point of agreement with the predictions of our model. So it is worthwhile to shortly analyze in more detail a few points which are related to the changes of the encoding correspondence in the evolution process. In particular in the coevolution theory of Wong [12] it is argued that, in an early primordial stage, codons encoding Gln (resp. Asn) belonged to the multiplet encoding Glu (resp. Asp). In Table IV we report the physical-chemical properties of these four amino acids together with our estimator η in order to compare their affinity in the framework of our approach. The computed value of η is low, confirming indeed the affinity between the couples of amino acids. It should also be remarked that the affinity is even more evident if the comparison is limited to the subset of the first four physical-chemical quantities of the Table IV and to some other quantities not reported in the tables (molecular weight, surface and volume), which play a more fundamental role in the evolution process. As the four amino acids are encoded by codons with a nucleotide A in second position, it is interesting to make a somewhat more detailed analysis of the physical-chemical properties of the whole set of amino acids sharing this feature. In Table V we report the value of the estimator η for the pairwise comparison between the seven amino acids encoded by codons with A in second position. Looking at Table V we see that the lowest value of η are in agreement with the prediction of our approach and of the coevolution theory. There a few more couples of amino-acids exhibiting a low value of η , in particular the couples: His/Asn ($\eta = 0.64$), His/Lys ($\eta = 0.38$), Gln/Asn ($\eta = 0.55$), Gln/Lys ($\eta = 0.48$), which have the peculiar feature that the dinucleotide (CA) of the first amino acid is a singlet, while the dinucleotide (AA) of the second one is the lowest weight state of the irreducible representation (1,1). In conclusion the pairwise comparison of the physical-chemical properties of the amino acids encoded by codons

Table III. Table of amino-acid doublet properties

		$P_\alpha + P_\beta$	P_τ	P_G	R_f	ΔH	$-\Delta S$	ΔG	pK_a	pK_b	pI
Gln/His	$\bar{x}$	2.04	0.97	10.45	0.83	61.10	121.80	95.95	2.00	9.15	6.56
	σ_x	0.17	0.01	0.05	0.01	—	—	1.36	0.17	0.02	0.91
	$\sigma_x/\bar{x}$	8 %	2 %	0 %	1 %	—	—	1 %	9 %	0 %	14 %
Asp/Glu	$\bar{x}$	1.72	1.10	12.60	0.86				2.04	9.64	3.00
	σ_x	0.17	0.36	0.40	0.02				0.16	0.04	0.22
	$\sigma_x/\bar{x}$	10 %	33 %	3 %	2 %				8 %	0 %	8 %
Asn/Lys	$\bar{x}$	1.75	1.29	11.40	0.93	56.40	135.70	96.84	2.10	8.88	7.50
	σ_x	0.15	0.28	0.10	0.04	1.40	0.10	1.32	0.08	0.08	2.09
	$\sigma_x/\bar{x}$	9 %	21 %	1 %	4 %	2 %	0 %	1 %	4 %	1 %	28 %
Cys/Tyr	$\bar{x}$	2.03	1.17	5.85	0.67	49.15	151.10	94.29	2.08	9.70	5.34
	σ_x	0.14	0.02	0.35	0.18	5.25	12.10	1.40	0.12	0.58	0.32
	$\sigma_x/\bar{x}$	7 %	2 %	6 %	27 %	11 %	8 %	1 %	6 %	6 %	6 %
Pro/Ser	$\bar{x}$	1.32	1.48	8.60	0.89	43.35	172.25	94.74	2.10	9.88	5.99
	σ_x	0.20	0.04	0.60	0.07	7.35	7.75	4.76	0.11	0.73	0.31
	$\sigma_x/\bar{x}$	15 %	3 %	7 %	8 %	17 %	4 %	5 %	5 %	7 %	5 %
Ala/Thr	$\bar{x}$	2.14	0.81	8.34	0.91	52.05	141.9	94.32	2.22	9.40	5.82
	σ_x	0.12	0.15	0.25	0.02	1.35	5.90	0.34	0.13	0.29	0.18
	$\sigma_x/\bar{x}$	5 %	19 %	3 %	2 %	3 %	4 %	0 %	6 %	3 %	3 %
Ile/Met	$\bar{x}$	2.59	0.54	5.45	0.75	57.75	140.40	99.63	2.32	9.41	5.84
	σ_x	0.09	0.07	0.25	0.01	0.45	11.90	2.97	0.04	0.19	0.10
	$\sigma_x/\bar{x}$	3 %	12 %	5 %	1 %	1 %	8 %	3 %	2 %	2 %	2 %
Leu/Val	$\bar{x}$	2.64	0.55	5.40	0.79	59.80	133.55	99.70	2.34	9.61	5.97
	σ_x	0.12	0.05	0.50	0.06	0.40	5.45	2.00	0.02	0.01	0.01
	$\sigma_x/\bar{x}$	5 %	8 %	9 %	8 %	1 %	4 %	2 %	1 %	0 %	0 %

Table IV. Table of amino-acid doublet properties: Gln/Glu and Asn/Asp

		$P_\alpha + P_\beta$	P_τ	P_G	R_f	pK_a	pK_b	pI
Gln/Glu	$\bar{x}$	2.05	0.86	11.35	0.83	2.18	9.40	4.44
	σ_x	0.17	0.12	0.85	0.01	0.01	0.27	1.22
	$\sigma_x/\bar{x}$	8 %	14 %	7 %	1 %	0 %	3 %	27 %
Asn/Asp	$\bar{x}$	1.58	1.51	12.25	0.88	1.95	9.20	4.09
	σ_x	0.02	0.05	0.75	0.01	0.07	0.40	1.21
	$\sigma_x/\bar{x}$	2 %	3 %	6 %	1 %	4 %	4 %	32 %

Table V. Values of the estimators η and η' for amino-acids with nucleotide A in central position. η is defined as in section 4 with the data for $P_\alpha + P_\beta$, P_τ , P_G , R_f , pK_a , pK_b and pI . η' is defined similarly by adding the data for the refractivity, the bulkiness, the molecular weight, the surface and volume data.

	η	η'		η	η'		η	η'		η	η'
Asn/Asp	0.52	0.67	Asn/Gln	0.55	0.98	Asn/Glu	0.84	1.23	Asn/His	0.64	1.25
Asn/Lys	0.68	1.37	Asn/Tyr	0.98	2.16	Asp/Gln	0.95	1.52	Asp/Glu	0.63	1.17
Asp/His	0.94	1.69	Asp/Lys	1.07	1.90	Asp/Tyr	1.37	2.69	Gln/Glu	0.61	0.70
Gln/His	0.34	0.58	Gln/Lys	0.48	0.75	Gln/Tyr	0.61	1.39	Glu/His	0.73	0.94
Glu/Lys	0.81	1.11	Glu/Tyr	1.18	1.99	His/Lys	0.38	0.56	His/Tyr	0.91	1.52
Lys/Tyr	1.02	1.54									

with A in second position gives a value of the affinity estimator η ranging between 0.34 and 1.37. The low values of η (< 0.7) appear in couples which are related by close irreducible representations in our scheme, or by closeness in an early time, in agreement with Wong coevolution theory or are of the form CAN and AAN. So we argue that the assignment of codons in multiplets, corresponding to the irreducible representations of our model, may have a deep connection with the evolutionary organization. This is a fascinating and interesting subject which will be discussed in more detail in the next future.

Acknowledgements

We are grateful to M.L. Chiusano for pointing us the interest to analyze the physical-chemical properties in the light of the crystal basis model. Partially supported by MURST (Italy) and MAE (France) in the framework of french-italian collaboration Galileo. The authors would like to thank the referee for very useful comments about the problem of evolution.

References

1. Sjostrom, M. and Wold, S.: A Multivariate Study of the Relationship Between the Genetic Code and the Physical-Chemical Properties of Amino Acids, *J. Mol. Evol.* **22** (1985), 272.
2. Frappat, L., Sciarrino, A. and Sorba, P.: A Crystal Base for the Genetic Code, *Phys. Lett. A* **250** (1998), 214.
3. Kashiwara, M.: Crystalizing the q -Analogue of Universal Enveloping Algebras, *Commun. Math. Phys.* **133** (1990), 249.
4. Konopel'chenko, B.G. and Rumer, Yu.B.: Classification of Codons in the Genetic Code, Translated from *Doklady Akademi Nauk SSSR* **223**, N.2, (1975), 471.
5. Chou, P.Y. and Fasman, G.D.: Conformational Parameters for Amino Acids in Helical, β -Sheet, and Random Coil Regions Calculated from Proteins, *Biochemistry* **13** (1974), 211.

6. Siemion, I.Z. and Siemion, P.J.: The Informational Context of the Third Base in Amino Acid Codons, *BioSystems* **33** (1994), 139.
7. Grantham, R.: Amino Acid Difference Formula to Help Explain Protein Evolution, *Science* **185** (1974), 862.
8. Weber, A.L. and Lacey, J.C.: Genetic Code Correlations: Amino Acids and Their Anticodon Nucleotides, *J. Mol. Evol.* **11** (1978), 199.
9. Siemion, I.Z. and Stefanowicz, P.: Periodical Change of Amino Acid Reactivity within the Genetic Code, *BioSystems* **27** (1992), 77.
10. Lide, D.R., ed.: *Handbook of Chemistry and Physics*, 73rd ed., CRC Press, Boca Raton, Florida, USA, 1992.
11. Di Giulio, M.: Some Aspects of the Organization and Evolution of the Genetic Code, *J. Mol. Evol.* **29** (1989), 191.
12. Wong, J.T.Z.: A Coevolution Theory of the Genetic Code, *Proc. Natl. Acad. Sci. USA* **72** (1975), 1909.