Journal of Theoretical Biology 369 (2015) 95–109

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Numeral series hidden in the distribution of atomic mass of amino acids to codon domains in the genetic code

Åsa Wohlin *,[1]

Möregatan 10, 118 27 Stockholm, Sweden

HIGHLIGHTS

- A 24-codon table for 20 amino acids (aa) shows masses of aa as codon domains.
- 12 aa with mixed codons form a 2–3-dimensional table with high mass regularity.
- A number series 5 – 0, exp. 2/3, × 100, gives stepwise aa masses of codon domains.
- The number series behind the periodic system, × 16, is shown as a related series.
- Transcriptions between number-bases 10, 8, 6 connect codons with aa mass domains.

ARTICLE INFO

Article history:
Received 29 September 2014
Received in revised form
13 December 2014
Accepted 13 January 2015
Available online 24 January 2015

Keywords:
Codon assignment
Atomic mass analysis
Exponent 2/3
Periodic system
Number-base transformation

ABSTRACT

The distribution of codons in the nearly universal genetic code is a long discussed issue. At the atomic level, the numeral series $2x^2$ ($x=5-0$) lies behind electron shells and orbitals. Numeral series appear in formulas for spectral lines of hydrogen. The question here was if some similar scheme could be found in the genetic code. A table of 24 codons was constructed (synonyms counted as one) for 20 amino acids, four of which have two different codons. An atomic mass analysis was performed, built on common isotopes. It was found that a numeral series 5 to 0 with exponent 2/3 times 10^2 revealed detailed congruency with codon-grouped amino acid side-chains, simultaneously with the division on atom kinds, further with main 3rd base groups, backbone chains and with codon-grouped amino acids in relation to their origin from glycolysis or the citrate cycle. Hence, it is proposed that this series in a dynamic way may have guided the selection of amino acids into codon domains. Series with simpler exponents also showed noteworthy correlations with the atomic mass distribution on main codon domains; especially the $2x^2$-series times a factor 16 appeared as a conceivable underlying level, both for the atomic mass and charge distribution. Furthermore, it was found that atomic mass transformations between numeral systems, possibly interpretable as dimension degree steps, connected the atomic mass of codon bases with codon-grouped amino acids and with the exponent 2/3-series in several astonishing ways. Thus, it is suggested that they may be part of a deeper reference system.

1. Introduction

1.1. Some general views

Some rough relations are certainly found between the distribution of codons and biochemical properties of the coded amino acids (aa). However, there are hundreds of contributors with different aspects and proposals in this field, and even if a few theories have been dominating, there is still no consensus on how to explain the genetic code, which leaves the field open for further proposals.

A deeper understanding of the coding system as such seems also needed, besides its similarity with a language. One aspect on the relations between the proteins and the coding bases could possibly be illustrated by two different ways to read a hexagonal pattern (Fig. 1): i.e., as composed of centers with three radii in plane projection, like tetrahedrons of aa, or as hexagon rings like coding bases. The bases are essentially made up of aa and each center is defined as such by three hexagon rings, to compare with the codon triplets.

Some general assumptions about the emergence of cells are proposed here, with the aim of connecting more physical associations to the arithmetical findings in this study.

* Tel.: +46 8 6424231.
E-mail address: a.wohlin@u5d.net
[1] Individual research.

http://dx.doi.org/10.1016/j.jtbi.2015.01.013

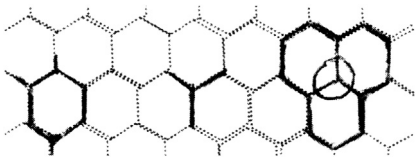

Fig. 1. Two ways of reading a hexagonal pattern, as central points with three radii or as hexagonal rings. The centers are defined by three rings in a similar way as aa are defined by codons.

The enormous complexity of a cell seems only comprehensible as a result of *internal differentiations* inside a unit, somehow defined as such a unit through complementary structures and opposite forces between a center and a semipermeable anticenter.

All forces (recognized or not) on the physical level should naturally be expected to appear in some form on the biological level, not only the electromagnetic one. The hydrophobic bonds of cell membranes, for instance, could be thought of as an appearance of the atomic strong force on the higher molecular level.

Regarding positive–negative charge, a cell represent an inversion in the center–anticenter relation of an atom. Such inversions or conjugates of higher dimensions may be part of the transition of forces between the atomic and the biological level. (An arithmetical inversion of the proton–electron mass quotient as an eventual root for the atomic masses in the genetic code is touched upon in Section 5, paragraphs 3 and 4, of this study.)

Furthermore, it may be time to consider the *concept of dimensions* not only as a number of "independent variables", but rather as deeply interdependent dimensional polarities in dynamic processes. Such a consideration implies a view where structural 3-dimensional aspects and all different kinds of biochemical gradients are integrated into a dynamic multidimensional system.

The genetic code has been described as an *information system*. This concept may sometimes give too narrow associations and should surely be taken in its widest sense to include many different kinds of mutual references within a cell, including complementarity in structures, polarity of charge, inversions, resonances, conjugates, the relations between interdependent dimension degrees (d-degrees) and so on.

The *relevance of mass* in general or number of nucleons especially (the main issue in this research) seems to have been neglected under long periods or even rejected in the discussion of the genetic code, mainly because unusual but stable isotopes do not seem to change the processes in a cell, e.g. the recognition of aa by tRNAs. Yet, several authors did early pay attention to it as an essential feature of the code, then in terms of moleular weight, size or volumes of aa, among others (Di Giulio, 1989). See further in Section 7.

It is reasonable to expect that all properties of atoms were essential at the emergence of the genetic code. If we can assume that mass is a property of higher *d*-degree than charge and as such represents an underlying level, it might imply that mass was decisive for elementary structures, while the more superficial level of charge, expressed in electron shells, becomes the relevant level in processes, in metabolism as characterized by more of released kinetic energy.

To search for *a guiding principle* behind the life building structures should be equally reasonable as it was when the patterns on the atomic level were found. Why numeral series are still regarded as special exceptions in biological systems is unclear, when there are such series on the atomic level. One established example is the $2x^2$-chain behind the periodic system, $50-32-18-8-2-0$; another is the formula for spectral lines of hydrogen: $1/\lambda = \mathrm{RH}\ [1/m^2 - 1/n^2]$, where, in the Balmer series, for instance, $m=2$ and $n=3, 4, 5,\ldots$ Those series appear as just the way of Nature to organize itself.

Among all the *published* studies and theories about the genetic code, it seems as relatively few have dealt with number regularities: e.g. Shcherbak (1993, 1994, 2003), Rakočević (2004), and Downes and Richardson (2002). Recently, Perez (2010) showed that the human genome as a whole single strand was of a fractal type with regard to frequency of codons. Indeed, it has long been known that features of Fibonacci number series and the golden section appear in Nature. (Such series showed up also in the atomic mass analysis of the genetic code presented below, Section 3.6.3.)

1.2. Background

The actual background behind this research was a very elementary 5-dimensional model or conceptual structure, suggested by the author for interpretations in theoretical physics (unpublished theory). It is only mentioned here because some of the arithmetical patterns detected in the present study reflect features of this model. Briefly, the model, with some redefinitions of the dimension concept and of the 4th dimension, implies a development from a 5-dimensional "entirety" through steps of polarizations towards lower degrees $5\rightarrow4\rightarrow3\ldots$ with debranched degrees translated to external motions (ultimately to pure kinetic energy in the last step $1\rightarrow0/\infty$, the "*d*-degree of motions") — or, in a closed system, meeting "the other way around" (Fig. 2).

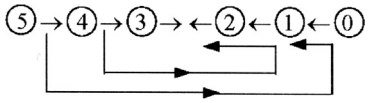

Fig. 2. A 5-dimensional chain. Debranched degrees from higher steps outwards rebuilt into structures "the other way around".

The figure, referred to in the text below, illustrates in other words a two-way directed process of disintegration and synthesis, implying a series of quantized steps between qualitatively different states, simultaneously expressed as polarizations into complementary partial structures, where conjugates may be one example.

The thought was that such a model, if of any value, should also appear as a pattern in some form on superposed levels of higher complexity, not least in a development towards life and the genetic code.

2. A 24-codon table and a numeral series

This study concerns eventual patterns in the distribution of atomic mass (nucleon number) of aa on codons and complementary codon types or domains. It is a general condition in this study that most common isotopes are used.

This section 2 presents a 24-codon table, leading to a division of aa into those with mixed and non-mixed codons; then correlations between atomic mass of codon-grouped aa with a numeral series. Section 3 concerns the same series related to other principles of distribution than codons, Section 4 the relation to the origin of aa from glycolysis and the citrate cycle. In Section 5 series with other exponents are studied, and Section 6 deals with transformations of atomic mass numbers between number-base systems, connecting the codon bases with aa and with the numeral series in this Section 2.

2.1. Mass of aa with mixed and non-mixed codons

The research started with making up a table of 24 codons (see online Appendix Table 4) for the 20 "classical" aa, four of which have two different codons. Hence, the counting implied that where U or C in 3rd position, like A or G in 3rd position, makes no difference, the codons were regarded as one and the same, as when the 3rd base makes no difference at all. The four aa with two codons are Arg

CG+AG−A/G, Ser UC+AG−U/C, Leu CU+UU−A/G and Ile AU−U/C+AUA, which differ only in 3rd base *type*.

Sums of A-and Z-numbers of the side-chains (R) of the coded aa from online appendix Table 4 are shown in Fig. 3.

G1, 5 aa	105 Z, 191 A.	— | —	411 A, 224 Z,	6 aa, G2
C1, 5 aa	195 Z, 353 A	— | —	133 A, 75 Z,	4 aa, C2
G1+C1	**300 Z, 544** A	— | —	**544 A, 299 Z,**	**G2+C2**
U1, 6 aa	250 Z, 463 A	— | —	437 A, 247 Z,	7 aa, U2
A1, 8 aa	278 Z, 497 A	— | —	523 A, 282 Z,	7 aa, A2
U1+A1	**528 Z, 960** A	— | —	**960 A, 529 Z,**	**U2+A2**

Fig. 3. A- and Z-sums of the codon-grouped side-chains of 24 amino acids in 1st and 2nd base order.

For a footnote on counting with charged basic and uncharged acidic aa, (Karlson 1974), see online appendix Table 4.

Unless stated otherwise, all numbers in this study refer to atomic mass (A-number) and to the R-chains of the aa.

Fig. 3 shows that the sums of aa in the G+C- and U+A-coded groups become the same in the 1st and 2nd base order. Hence, aa with what here is called mixed codons (1st base G or C with 2nd base A or U and with 1st base U or A with 2nd base G or C) that "change positions" between base-pair groups from 1st to 2nd base order have the same sum. The mixed codons represent 12 aa with the atomic mass=770 (Table 1), the other, non-mixed codons represent 12 aa with the atomic mass=734 (Table 2). They will here be referred to as the 12-group 770 and the 12-group 734.

Table 1
Group of 12 amino acids with mixed codons divided into 4 groups of 3 after the first base in the codons. Total atomic mass 770.

GA	Glu	CA	His	UG	Trp	AG	Arg	→ **385**
GA	Asp	CA	Gln	UG	Cys	AG	Ser	→ **209**
GU	Val	CU	Leu	UC	Ser	AC	Thr	→ **176**
175		**210**		**208**		**177**		
385				**385**				**770**

The remarkable regularities (more about these in Sectiom 2.8) in the 2-dimensional Table 1 (-/+1), led to the further investigations. The N-Z-division (see Section 2.5) in this table makes it almost 3- dimensional. Another annotation already here:

In this Table 1 the atomic mass sum of aa in rows 1 and 2 with differentiating 3rd base in codons becomes 594, $2 \times 11 \times 3^3$, and the so-called "fourfold degenerated" aa in row $3=176=2 \times 11 \times 2^3$. This is one reason why codons with indifferent 3rd base here preferably are regarded as "2-base-coded". About the numbers 385, 27 and 8, see further Section 5.4.

Table 2
Group of 12 amino acids with non-mixed codons divided in 4 groups after the first base in the corresponding codons. Total atomic mass 734.

414						320		Sum
GG	Gly	CC	Pro	UU	Leu	AA	Lys	**322**
				UU	Phe	AA	Asn	
GC	Ala	CG	Arg	UA	Tyr	AU	Ile1	**412**
						AU	Met	
						AU	Ile2	
16		143		255		320		
159				**575**				**734**

(The tables may have some similarity with a table by Rumer (1966) but does not concern the same views or symmetries and was simply derived from the ordinary table of 64 codons.)

Based on the 1st and 2nd bases there are 4 types of codons with 6 aa in each group, here referred to as Form-codons GA, AG, UC, CU and Cross-codons GU, UG, CA, AC, these two in Table 1, RNA-codons GC, CG, UA, AU and Pair-codons GG, CC, UU, AA, these two in Table 2.

Abbreviations used below: G1, C1, U1, A1 refer to the total atomic mass of aa with G, C, U or A respectively as 1st base in their codons, and G2, C2, U2, A2 refer to the total atomic mass of aa with G, C, U or A as 2nd base in their codons. G+C and U+A refer to the corresponding atomic mass sums of aa, which became equal in 1st and 2nd base order.

2.2. The numeral series $x^{2/3}$ $(x=5-0) \times 10^2$

It was found that a numeral series $x=5-0$ with exponent $2/3 \times 10^2$, here called the exponent series (ES) or the ES-chain, with abbreviated numbers=292−252−208−159−100 (Fig. 4) highly correlated with codon-grouped aa.

$5^{2/3}$ — $4^{2/3}$ — $3^{2/3}$ — $2^{2/3}$ — $1^{2/3}$ — 0, x 10^2
= 292.4. 252.0. 208.0. 158.7 100
292 252 208 159 100 → abridged
752 ↓ **259**
544 **x 2 = 416**
2 x 752 = 1504, total sum of 24 aa R
544 = G+C-coded aa = 752 - 208
544 + 416 = U+A-coded aa = 752 + 208

Fig. 4. The "exponent series" (ES). The first three numbers in the series, sum 752, times 2, give the total sum 1504 of the 24 aa R. The G+C- and U+A-groups 544 and 960 respectively are given as 752 minus/plus the 3rd number 208.

The limitation to first three numbers in the series do not exclude the lower ones; the ES-chain implies operations minus/plus lower numbers and intervals for the derivation of codon-grouped aa in agreement with the feature shown in Fig. 2.

2.3. G+C-coded aa group 544

Individual G- and C-groups were obtained from the -/+ last two numbers in the ES-chain:

G1=292–100, -1* C2 = 292 − 159 = 133
C1=252+100, +1, G2=252+159=411

*101=Arg, CG, as if displaced from a GG-codon? Arg gets its end-group from the G-base (and forms creatine phosphate with Gly GG).

2.4. The two 12-groups 770 and 734

The halves of these 12-groups, 385 and 367, are shown in Fig. 5.

367 **2 x 367 = 734**
/ \
292 — 252 — 208 — 159 — 100 — 0
544 <——|——>|
385 **2 x 385 = 770**
2 (544 - 159) = 770, Cross- and Form-coded aa
2 (208 + 159) = 734, Pair- and RNA-coded aa

Fig. 5. The two 12-groups of amino acids in the ES-chain halved: 385 an interval, 367 a sum.

2.5. Pyrimidine and purine aa groups

These aa-groups were obtained in a similar way from the numbers 544 and 208:

C2+U2=2(544−159−100)=570
G2+A2=2(208+159+100)=934, 467×2
C1+U1=1/2×544+544=816
G1+A1=1/2×544+416=688

2.6. U- and A-coded aa groups

U- and A-groups in 2nd base order with sum 544+2×208 were obtained through −/+ the interval 208−100, −1.

U2=544−108=436, +1.
A2=416+108, −1.
(Note 107=Tyr UA, deriving from Phe, UU-coded.)

(Cf. A1=544−47, U1=416+47, and 47=Cys UG; as from Met AUG? 47 is not an interval in the ES-chain but a factor in the total 1504.) We have also:

U1 = 463 = G1 + 1/2 x 544
A1 = 497 = C1 − 1/2 x 544 + 416

Individual codon-groups differ from 1st to 2nd base order with a number 59, the interval 159−100 in the series, −/+1:

G1–58=C2, C1+58=G2
U1+60=A2, A1−60=U2

A summary of correlations of codon-grouped atomic mass sums of aa with numbers from the ES-chain is given in Table 3.

Numbers 292−252−208−159−100 are from here on often designated 5' - 4' - 3' - 2' - 1'.

2.7. Individual groups of aa in the 12-group 770

Some atomic mass sums of the individual mixed codon-groups are shown in Fig. 6.

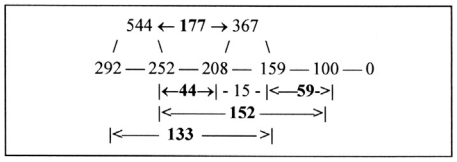

Fig. 6. Individual codon-groups of aa in 12-group 770.

177=UG Trp+Cys.
208−177=31=UC, Ser1
133−1=GA, Glu+Asp=AG, Arg2+Ser2
152+1=CA, Gln+His
44 −/+1=GU and AC, Val and Thr
2×44=UC+CU, Ser+Leu

The numbers 177 and 208 are given in another way in Table 3, rows 13 and 14. Note the interval 133=Asp with its backbone chain (B) included and R=59 (159−100). Asp is often seen as typical for an aa. So is Ala, R 15, an interval of 2nd order, 44–59, bridging over the middle step. (Cf. additions to the citrate cycle in Section 4.1 below.)

The numbers 177 and 367 represent a 3rd division of 544 after 292+252 and 336+208. With −/+ the interval 44, they give C2=177−44=133 and G2=367+44=411.

2.8. The 12-group 770, number 77 and factor 11

77 is the whole ES-chain divided in step 4' − 3' = 544 − 467 (208+159+100). The factor is remarkably distributed 2×77 and 3×77 in Table 1:

Row 1:…G1+C1=2×77, U1+A1=3×77
Row 2+3: G1+C1+3x77, U1+A1=2x77
231 is divided 130−101 −/+1 ∗ in groups 3×77.
154 is divided 77 −/+4 on G1−C1, −/+1 on U1−A1

Table 3
Summary of correlations of codon-grouped atomic mass of aa with the ES-chain.

	Codons for aa, sums R		=ES-chain	Follows from column 4	Number divisions
1	G1+C1+U1+A1	1504	2 (5′+4′+3′)	=G2+C2+U2+A2	Total =1504
2	G1+C1	544	5′+4′	U1+A1=5′+4′+2×3′	1504 in 544–960
3	G2+C2	544	5′+4′	U2+A2=5′+4′+2×3′	1st=2nd base order
4	G1	191	5′−1′−**1**	C1=4′+1′ +**1**	544 in 191–353
5	C2	133	5′−2′	G2=4′+2′	544 in 133–411
6	C2+U2	570	2 (5′+4′−2′−1′)	G2+A2=2 (3′+2′+1′)[a]	1504 in 570–934
7	C1+U1	816	5′+4+1/2(5′+4′)	G1+A1=2×3′+1/2(5′+4′)[a]	1504 in 816–688
8	U2[b]	437	5′+4′−(3′−1′)+**1**	A2=2×3′+(3′−1′)−**1**	960 in 437–523
9	U1[b]	463	4′+3′+**3**	A1=3′+5′−**3**	960 in 463–497
Mixed codons, 12 aa					
10	Total	770	2(5′−2′)	2×3′+2′=non-mixed codons	1504 in 770–734
11	GA+GU+CA+CU	385	5′+4′−2′	5′+4′−2″=UG+UC+AG+AC	770 in 385–385
12	GA+AG-+UC+CU	352	4′+1′	2×2′+1′=UG+GU+CA+AC	770 in 352–418[c]
13	GA+GU	175	1/2(4′+1′)−**1**	1/2(4′+1′)+**1**=AG+AC	352 in 175–177
14	CA+CU	210	1/2(2×2′+1′)+**1**	1/2(2×2′+1′)−**1**=UG+UC	418 in 210–208
Non-mixed codons, 12 aa					
15	GG+GC+CC+CG	159	2′	2×3′+2′=UU+UA+AU+AA	734 in 159–575

Bold figures −/+1, −/+3=deviations from exact correlations.

[a] Single U1, A1, U2 and A2 groups may be indirectly derived exactly from correlations in rows 6 and 7.
[b] More direct derivations of U2 and A2, U1 and A1.
[c] Division between aa with Form-codons and Cross-codons respectively.

*The 11-factor in this group is hypothetically a result of two-way direction in steps of the basic chain 5−0 read as two-digit numbers.

77=43+34: 3×43=*129*, 3×34=*102*

This two-way direction also influences the codons with regard to the 1st and 2nd bases: Adding aa with mirror-codons in this 12-group 770 gives sums that all include factor 11: UG+GU=220, GA+AG=264, CA+AC=198 and CU+UC=88.

"Mistakes" at transcription such as CU for Ser (normally a Leu codon) might depend on this two-way direction of the underlying structure.

2.9. The 12-group 734

This 12-group of aa implies a less regular division of atomic mass on the codons; however, the main groups correlate clearly with the ES-chain:

U+A-coded aa=575=2×208+159

G+C-coded aa=159

The G–C-group 159 is divided in accordance with the intervals in the ES-chain −/+1:

Ala GC+Gly GG+Pro CC=59−1

Arg CG=100+1

2.10. Number of aa

The codon groups correlate in number of aa with first three numbers of the underlying elementary chain 5–0:

5+4+3=12, ×2=24

G1+C1: 5+5 aa, U2+A2: 7+7 aa

A1=4×2, U1=3×2

G2=3×2, C2=2×2

Cf. flower plants with 5-, 4- or 3-merous plans (increasingly asymmetric), plans that certainly are encoded in the genes.

With ES-numbers designated 5', 4', 3', 2', 1' there is a relation with the main groups of aa through −/+ interval 5′−4′=40:

2×5′=584, −40=544=G+C, 10 aa

2×(4′+3′)=920, +40=960=U+A, 14 aa

In the U2- and A2-groups a division was found between 4 aa with 1st base G or A, 3 aa with 1st base U or C. The division of the 5 aa in G1 and C1 showed a similar pattern, where 3 aa have 2nd base G or A and 2 aa have 2nd base U or C.

2.11. 3rd base groups and number 292

There are 8 aa in each 3rd base group, 8 with 3rd base A/G and either A or G, 8 with U/C and 8 with indifferent 3rd base (4 of these in each 12-group).

Sums of aa in 1st base order with codons differentiated in the 3rd base were found to be 2×292 in both the purine and pyrimidine groups (+1 in C1+U1).

In G1+A1=2×292=584

In C1+U1=2×292=584, +1

In G2+A2=2×416, C2+U2=336+1.

The "2-base-coded" aa=544−208=336, −1.

2.12. The parents of the RNA-bases=292

The sum of hypoxanthine 136 and orotate 156, the parents of the purine and pyrimidine bases in the codons, was 5′=292. They are naturally connected with the differentiations in the 3rd base and to the first base in codons through the anticodons of tRNAs.

Curiously, when the sum 292 was distributed to the following two numbers in the ES-chain, it gave, times 2, directly the main pyrimidine and purine groups of aa:

2 (252+Orotate 156)=816=U1+C1

2 (208+Hypoxanthine 136)=688=G1+A1

3. The ES-series in distribution of atomic mass on bases other than codons

3.1. Backbone chains (B-chains)

B-chains of aa, although seemingly unrelated to the distribution of codons, appear in the ES-chain in groups of six (Fig. 7).

292 - 252 - 208 - 159 - 100 ← Condensation
544 <——————> 100 = **444** = 6 B-chains à 74 A*
208 <——————>100 = **108** = minus 6 × 18 (H_2O)
544 <——>208 = **336** = 6 B-chains bound

Fig. 7. Backbone chains in groups of six, unbound and bound *(More correct 443. Cf. Arg charged 101).

Dividing the interval 444 into two parts around the middle number, 544→367 and 367→100, gives numbers 177=6×29.5 and 267=6×44.5, corresponding to parts of an unbound B-chain (H)H_2N–CH– and –COO(H), which supports the identification of B-chains in the ES-series.

3.2. Atomic mass division on atom kinds

It was found that when dividing the total atomic mass of aa R on carbon atoms C and other atoms (N, O, S, H), the sums became the same as between U+A- and G+C-groups:

C-atoms=960, ∼U+A

N+O+S+H=544, ∼G+C

The sum 544 of N+O+S+H was divided as in the ES-series:

292 on G1+A1

252 on U1+C1

H=252−100, N+O+S=292+100.

These relations to the ES-chain are a further indication that a numeral scheme could have guided the organization as along different coordinate axes and on different levels.

The atomic mass of C-atoms is divided on the two 12-groups of aa in accordance with the ES-chain:

770-group: 444=544−100

734-group: 516=416+100

Note the step to number 3′ of the ES-chain in the 12-group 734, here as in Fig. 5 (Section 2.4). The division on atom kinds shows a

special regularity in aa grouped according to the keto-/amino polarity and in 2nd base order.

G2+U2: C=576=3 × 192. Rest=272
C2+A2: C=384=2 × 192. Rest=272

3.3. Number of C-atoms in R-chains

One basis for grouping the aa is the number of C-atoms in their R-chains. The atomic mass of aa with 4 C-atoms in the R-chains was 2 × 5′=584. Here, the aa are grouped after two times the ES-numbers 5′, 4′, 3′:

584=group 4C, 8 aa
504=groups 7C+3C, +0C (Gly), 8 aa
416=groups 2C+1C, +Trp 9C, 8 aa

Phe and Tyr are constructed through 3C+4C, while Trp gets 9C by 3+4+5C-atoms with reductions. Adding Trp to group 1C may seem odd but consider that Ser 1C takes part in the construction of Trp, Trp shares codon UG with Cys 1C, and Trp can break down to Ala 1C.

Approximate doublings in the number of C-atoms were also noted, from G1 9 to C1 18C, from G1+C1 27C to U1+A1 53C. (C2+G2 26C, U2+A2 54C.)

3.4. End-groups of R-chains

The sum of aa with only CH*x* or H=420, the sum of aa with OH*x*, without Gln and Asn=468. The sum of all aa with N in the R-chains=616.

888=2(544−100), sum of first two groups
616=2(208+100), the N-contenting group

3.5. N–Z-relations

The numbers of N and Z in codon groups 544 and 960 differed only by one unit from 1st to 2nd base order (Fig. 3). It was observed that the division in N and Z equals the division in codon type groups −/+1 in the 12-group 770; N=351=form-coded aa 352−1, Z=419=cross-coded aa 418+1. In the 12-group 734 a similar pattern was found within the U+A-group 575: N=256=U1 255+1, Z=319=A1 320−1.

H-atoms in the 24 R-chains seemed to be distributed in accordance with codons: H-atoms=152=8 × 19 were closely divided in 1/8 and 2/8 on G1 and C1, 2/8 and 3/8 on U1 and A1, only exchanged within the groups in 2nd base order.

Separating differences in N and Z between codon domains in 1st and 2nd base order produced numbers of the ES-series as a border (Fig. 8), +1 in C- and A-groups.

4′+3′=460, 3′+5′=500:

G1: 191 —> + 101 N —> **292** —> + 119 Z —> 411 = G2
C1: 353 —> - 100 N —> **253** —> - 120 Z —> 133 = C2
U2: 437 —> + 23 N —> **460** —> + 3 Z —> 463 = U1
A2: 523 —> - 22 N —> **501** —> - 4 Z —> 497 = A1

Fig. 8. Displacements in N and Z from 1st to 2nd base order (G–C) and from the opposite direction (U–A).

The interval in the middle step 3′−2′ of the ES-chain=49. ES-numbers 5′ and 4′ in the G–C-group and 1/2 × 960 in the U-A-group −/+ this interval gave closely the division in N and Z in 1st base order:

G1+C1:N 244=292−49, +1
Z 300=252+49, −1
U1 + A1: N 432 = 480 − 49, +1
Z 528=480+49, −1.

3.6. Geometrical aspects

3.6.1. An inverted exponent 3/2

Two first intervals in the ES-chain with an inverted exponent seems to imply a kind of feedback system, establishing close to the ES-numbers:

292−252=40; $40^{3/2}$=253.
252−208=44; $44^{3/2}$=292.
$84^{3/2}$= 770.; 2 x $42^{3/2}$=544.4.

3.6.2. Relations of the two 12-groups of aa

Halves of the 12-groups 770 and 734 transform approximately into one another through $\sqrt{2}$ and inversions (/\). (259=2′+1′ in the ES-chain):

$385 \times \sqrt{2}=544.5,\ \times\frac{1}{2}=272.2,\ /\backslash=367.3 \times 10^{-5}$
$367.3 \times \sqrt{2}=519.5,\ \times\frac{1}{2}=259.7,\ /\backslash=385 \times 10^{-5}$

Hence, relations as sides and diagonals in squares.

3.6.3. Inversions among codon bases

Relations through inversions are found also between the bases as if they expressed one kind of references. The central contribution to purines is Gly 75 (R+B), and the main contribution to pyrimidines is Asp 133 (R+B). These numbers are inversions of each other as 3/4 and 4/3 and are intervals in the ES-chain: 75=292−367 and 133=292−159; sum 208=3′. Sum of the three bases U+C+T=349 are inverted=286.5. × 10^{-5}. 286 is the sum of A+G. Another detected inversion could describe the polarity within base pairs. Atomic mass of two G-bases=302, of two U-bases 224:

$302^{2/3}$, /\ × $10^4 \approx 2 \times 111.1$, ∼2C,
$224^{2/3}$, /\ × $10^4 \approx 2 \times 135.6$, ∼2A.

3.6.4. The golden section

Starting from 2′=158.74 and multiplying with the number of the golden section 1.6180. in steps (below represented by arrows) gave the following number series, which here are more abbreviated in higher numbers:

2′:158.74→256.8*→415.6→672.4→1088→1760→2848.

*208+interval 49=257.

First four numbers=1504, the total of R-chains; sum of 2nd, 3rd and 4th number=1344, the 24 B-chains bound=2 × 672; 1088=2 × 544, and 2848 is the sum of 24 aa R+B bound, in the next step leading to $\sim 8 \times 24^2$, i.e., 8 times the number of aa squared. Other similar steps:

3′: 208→336.5→544.5. ∼G+C
3′+2′: 367→593.8→960.8. ∼U+A

3.6.5. Structures of side-chains

The R-chains of aa may roughly be divided in (a) ring-closed (the aromatic, U1-coded, +Pro, His, C1-coded), (b) “straight”, and

(c) aa with branched end-groups (with atoms C–C, O–O, O–N or N–N), G2- and A2-coded:

(a) ring-closed=451=5′+2′
(b) "straight"=318=2 × 2′; (a)+(b)=*770*, −1
(c) branched:=2(3′+2′)=734, +1

Cf. sums of the two 12-goups, yet not the same groups of aa.

4. Origin of the amino acids from glycolysis and the citrate cycle

4.1. Halves of aa from each process?

One of the first and best known theories concerning the distribution of codons is the "coevolution theory" proposed by Wong (1975), which essentially presumes a stepwise evolution of the code into domains and subdomains simultaneously with the process for aa synthesis.

Here, the research started from the fact that all aa with U as 1st and/or 2nd base in their codons derive from stations in the glycolysis – in the aa syntheses of today. A connection is suggested with the fact that UTP-(-UDP-UMD) is the coenzyme associated with the binding and break-down of glycogen. The total sum of these aa R=752, i.e., exactly half the total sum of R 1504=5′ +4′+3′ in the ES-chain.

An equal division of aa (+/−1=Gly) outside/inside mitochondria membrane can be presumed under certain conditions, i.e., if the GC-coded Ala is taken as derived from oxaloacetate (while it also may derive from pyruvate), and the AG-coded Ser2 likewise, as if derived via the outer "loop" from the citrate cycle via homoserine to 3P-glycerate. Then the two 12-groups 770 and 734 become divided between glycolysis and the citrate cycle as shown in Fig. 9.

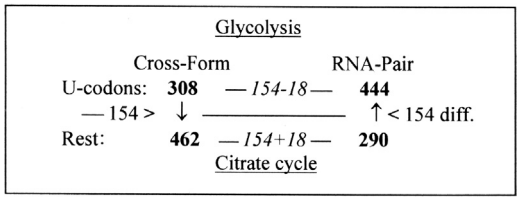

Fig. 9. Codon type groups divided based on their origin from glycolysis or the citrate cycle.

U-codons gets divided in 544−100=444 and 208+100=308. The rest gets divided in 292−2 and 252+208=460, +2.

The distribution confirms the relevance of number 77 (see Section 2.8)=the interval step 4′−3′=544–467 in division of the whole ES-chain:

385−77=308; 367+77=444
385+77=462; 367−77=290

The numbers +/−18 (Fig. 9) and +/−2 are atomic mass numbers of water and 2H, the processing of which are central themes in glycolysis and the citrate cycle.

Conversions between number base systems (nb-*x*) suggests that the groups (−/+1) in Fig. 9 "refer" to each other through transformations from nb-10 to nb-8: 308−1 in nb-10=462+1 in nb-8, and 290+1 in nb-10=444−1 in nb-8 (see further Section 7).

Dividing aa in codon-groups G+C and U+A with origins in glycolysis or the citrate cycle respectively (Fig. 10), gave atomic mass sums that agreed with simple divisions in the ES-chain.

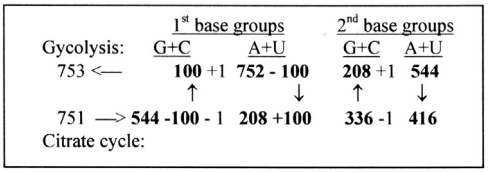

Fig. 10. Divisions of atomic mass of aa based on their origins in glycolysis or the citrate cycle.

(Displacements from 1st to 2nd base order here are all −/+ the interval 108, and also equivalent to (∼) 6 H_2O, the central theme in the processes and=3′−1′ in the ES-chain.)

It was also observed that the two main contributions into the citrate cycle that are CO_2, 44A, and acetyl(∼CoA)+OH, 59 (or 60) A, appear as intervals in the ES-chain: 44 in step 4′→3′ and 59 in the *corresponding* step 2′←1′ with views from the background model in Fig. 2.

Sums of aa may be compared with sums of stations in glycolysis and the citrate cycle: (a) bound and charged, (b) with +1 for charges and bonds to sulfur- and phosphorus enzymes:

Glycolysis, eight 3C-stations: (a) 752, (b) 766.

Citrate c., ten stations[a]: (a) 2 × 738, (b) 2 × 748.

[a]Malate, oxaloacetate, citrate, cis-aconitate, isocitrate, oxalosuccinate, α-ketoglutarate, succinyl-(CoA), succinate, fumarate.

In the citrate cycle 6 stations (one with 5C, 5 with 4C) sum up to 748 when counting in the way of (b) above; then the sum of 4 stations with 6C is also=748. If an addition of 4 × 2H released in the cycle should be quite correct, this sums up to 1504, 2 × 752, the same sum as only the R-chains of aa.

That several individual atomic mass numbers of stations appear among aa is not surprising; however, that the sums of stations add up to approximate the sum of aa R is not a matter of course. It might indicate a close relation with the genetic code when the processes developed and influence from a common number scheme. Many ES-numbers and intervals appear in the stations, as 192 (citrate), 100 (succinyl∼), 146 (α-ketoglutarate, $\frac{1}{2} \times 292$), and 133 ± 1 (malate–oxaloacetate). However, 192 is not an individual aa number but appears in groups of aa, e.g., 7 × 192=24 B-chains bound; 5 × 192=960, the whole U+A-group; 3 × 192= 576=U+A-group +1 in 12-group 734 and 2 × 192=384, the halves of 12-group 770−1.

4.2. Codons and anticodons

Most aa with U in 1st and/or 2nd position of codons from *glycolysis* have anticodons that code for aa from the citrate cycle; Ile2 (AUA) and Tyr are exceptions. (The only aa from the *citrate cycle*, with the assumptions above, that have no anticodon for aa from glycolysis was Ser2, AG-U/C.) The atomic masses for aa from glycolysis and those from the citrate cycle with anticodons was compared and showed number divisions from the ES-chain.

U-codons with 2nd base G or C, (UG+UC), with aa sum=208, gave anticodons for Gly, Ala, Thr, Pro, Arg1 and Arg2; with Ser2 (assumed above) included=aa sum 336; together 544. U-codons with 2nd base U or A, with the aa sum 544 when Tyr and Ile2 were included, gave anticodons for Lys, Gln, Glu, Asp, Asn and His=aa sum 416.

These number divisions might give additional support for the thought that half the codons originate from tRNAs as from the opposite strand of DNA.

5. ES-chain compared with other series

5.1. Simple quotients of 24 and the ES-chain

It should first be observed that the quotients 7/24, 6/24 and $5/24 \times 10^3 = 291.7.$, 250 and 208.3. approximate numbers 5′, 4′, 3′ in the ES-chain and times 2 give 1500, close to the total of aa R 1504.

The parts 14, 12 and 10 are equal to numbers of electrons in filled shells of the *f*, *s*+*d* and *d* orbitals of the periodic system.

5.2. The $2x^2$-series (x = 5-0) behind the periodic system

5.2.1. A factor 16 times the $2x^2$-chain

Numbers in the $2x^2$-series ($x = 5-0$) represent additions to whole shells in the periodic system: 50 – 32 – 18 – 8 – 2 – 0, and intervals the number of electrons in orbitals for whole shells. (Orbital 18 and number 50 not found in Nature.)

Life might be regarded as a continuation of "fusion", i.e., of the carbon–nitrogen cycle in the sun, outsourced to a planet, while transformed to the molecular level, with internal relations of mainly C, N, O and H inverted to outer bonds.

The ES-chain is the cubic root out of the $2x^2$-chain halved × 100 and was therefore compared with this series. A simple first observation was that 32 and 18 times 30 gives the U+A-coded aa 960 and the G+C-coded aa – 4, 540; 30 = 2 times the series 5 to 0.

Factor 16 times numbers in the series (Fig. 11) was found to give the main codon groups. (Codon groups of aa in 1st base order are all divisible by 16 −/+1.)

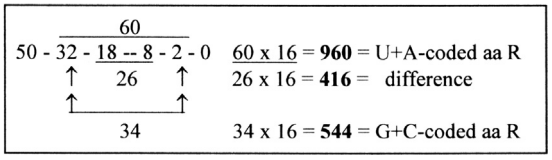

Fig. 11. Numbers 32 to 2 in the $2x^2$-series multiplied by 16, giving sums of codon-grouped aa.

The number 544 of G+C-coded aa may represent a deeper level, representing a first "polarization" of 5 to 4+1, as the number comes "first" in the ES-chain. Cf. Fig. 2 when it concerns synthesis, i.e., the opposite direction.

Many other atomic mass sums of codon-groups of aa are given −/+1 (Fig. 12).

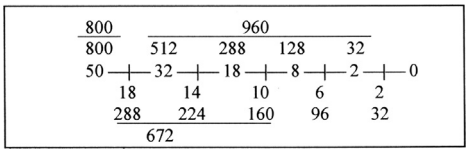

Fig. 12. Numbers for 16 times the $2x^2$-chain.

The 12-group 734 and its codon domains were derived from middle numbers in the ES-chain. So too in this chain:

$2 \times 288 = 576 = U1 + Al + 1,$
$U1 = 288 - 32, -1, A1 = 288 + 32.$
$288 - 128 = 160 = G1 + C1\ 159, +1$

The 12-group 770 halved = 385, −1 = 800 − 416 or 512 − 128, an interval here as in the ES-chain. 12 B-chains bound = 672, the sum of 3 intervals.

Here as in the ES-chain the 3rd number times 2 in the chain belongs to the U+A-group. (Cf. that these +8 and +18 to whole shells in the $2x^2$-chain are repeated in the periodic system; the *d*-orbitals of these shells filled later than the *s*-orbital of next shell).

The factor 16 times $2x^2$ may of course be replaced by 32 (2^5) times x^2. (First three numbers 25, 16, 9 represent a Pythagoras' triangle. They are also found in the intervals of DNA-bases, T to A 9, A to G 16 and G to T 25.) The squared integers give a reason to remind of Rydberg's formula for wavelengths of spectral lines of hydrogen with intervals between such inverted numbers. (Other interpretations of a factor 16 are not excluded.)

5.2.2. The total Z-sum of unbound aa

Total Z of the 24 unbound aa was found to be $1760 = 16 \times 110$, and 110 is the sum of the $2x^2$-series.

Counting with 40 Z in each B-chain, as in Gly, implies taking 24H from R-chains "back" to the central C-atom of B-chains, plus the 4H lacking in B-chains of Arg 1,2, Lys and Pro. The Z-sum of B-chains become 960 and Z-sum of R-chains 800. (960 divided into 20 B-chains = 800Z, +4 = 128 + 32Z).

With this way of counting, the Z-sum of U1+A1 becomes = 512 (528Z − 14, –2H for Arg AG and Lys AA), and the Z-sum of G1+C1 becomes 288 (300Z − 10 − 2H for Arg CG and Pro CC). The agreement of these numbers with the division of 800 in Fig. 12 above indicates further the impact of the $2x^2$-chain as a possible underlying or preceding level in the evolution of the code.

The Z-distribution on atom kinds becomes equal to numbers in Fig. 12; in B-chains C-atoms = 288, N+H-atoms = 288, O-atoms = 384; in R-chains C-atoms = 512 − 32 and N+O+H+S = 288 + 32 (~2S), a relation 3 to 2×160, 480 − 320.

These findings could raise the question as to whether B-chains, like L-waves of dipoles, preceded the evolution of R-chains in a step to proteins as "T-waves".

5.2.3. The number 26

The sum of the middle of the $2x^2$-chain = 18 + 8 gave squared, × 2, the atomic mass of R-chains without H-atoms, = 1352. The total atomic mass 3276 of 24 aa unbound = 26×126.

5.2.4. Sum of all integers 1–110

The series 1–110 = $55 \times 111 = 6105$. The inversion of this number happens to give half the atomic mass 3276 of 24 unbound aa, this as periodic decimals, 0.000163800163800…

$1/55 - 1/555 = 1638.00 \times 10^{-5}$ as period
$1/55 \times 1/555 = 3276.00 \times 10^{-8}$ as period

(1/55 = 18 as period, atomic mass of H_2O, 1/555 = 180 as period, atomic mass of fructose; both essential molecules in the origin of aa.)

5.2.5. Selenocysteine

Numbers 34 and 60 in Fig. 11, × 16, gave the G+C- and U+A-groups of aa. The sum 94 is the R-chain of *selenocysteine* (with selenium 34 *Z*), which is called the 21st aa, encoded as it is in a special way. It seems to express the unity of main codon domains, to compare with Cys with half the atomic mass, which binds protein chains on a higher level.

5.2.6. Mass divisions hidden in a shared field?

Disintegration of atomic mass seems reflected in the total of electrons or protons. 1760 was the total Z of 24 unbound aa. The number of C, N, O and S atoms in these aa R+B is 224 (16 times the orbital number 14 in step 4 − 3). Reducing 1760 with 224 in steps (as a kind of activation in suppressing of deeper orbital levels)

with one electron in each step, until those of the C-atoms are zero, gives six "phases" (Fig. 13):

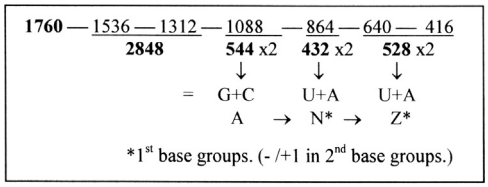

Fig. 13. Reducing total Z=e 1760 (R+B) with 224 electrons per step gives six "phases" that correspond to atomic mass division on codon groups of aa and/or on atom kinds.

2848 is the atomic mass sum of the 24 bound aa R+B; the first number in the series, *1536*, is the sum of all 128 C-atoms, the second number *1312* the sum of all other atoms. The whole interval 1760−416=*1344* is the atomic mass of 24 bound B-chains, 4×56 in each step.

Following "phases"=2(*544, 432, 320, 208*) gave the R-chains (note doubled) of G+C, then N and Z of U1+A1. This finding may indicate that atomic masses somehow are collectively expressed in/on the levels of electron shells; e.g., in K-shells the total R+B bound and in L-shells the R-chains or the equal distribution on substituents and carbon.

(With the same operation on R- and B-chains separately, Z 828 and Z 932, first two phases in R-chains, −104 per step, gave the atomic mass sum of all bound B-chains, and first two phases in B-chains, −120 per step, gave the atomic mass sum of all R-chains.)

5.3. Half orbitals as a superposed chain

5.3.1. Wavy readings of 6- and 2-digit numbers

Half orbitals as elementary integers may be seen as a superposed chain (Fig. 14) to the elementary 5−4−3−2−1 series. Read as triplets it gives 975+531=1506=24R-chains+2.

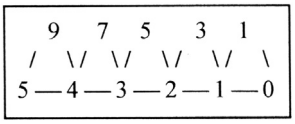

Fig. 14. Halved orbital numbers as a superposed series.

(The triplets 975 and 531 are roughly the division of 24 aa on Class I and II of aminoacyl tRNA synthetases: 975+2, 531−4.)

A wavy reading of 6-digit numbers from lower 5 to upper 5 outwards and the opposite way back, taking the square root out of these numbers, was found to give the two 12-groups −/+1:

594735→√=771.2.∼12-group 770 +1
537495→√=733.1.∼12-group 734 − 1

The same operation from lower 4 to upper 3 gives the sum of 20 aa without the double set of 4 aa; outwards Z −1≈688, inwards N+1≈570, (also=G1+A1 and C2+U2 among the 24 aa).

Reading 2-digit numbers in the same wavy way gives:

(a) from 5: 59+94+47+73=273; ×2=546
(b) from 4: 47+73+35+52=207, ×2=414

Cf. 544=G+C, and U+A=544+416.
273 and 207 are also the mass of π^+-mesons and μ-leptons in units of electrons. Protons disintegrate into π-mesons and these into μ-leptons at proton–antiproton collisions. In proteins aa appear as the "elementary particles" and "carriers of forces" on the higher level of biology.

5.3.2. Has 544 its root in the e/p-quotient?

The number 544 might perhaps be rooted in the p/e quotient as an inversion of this. The quotient protons p^+ to electrons e^- is very close to the quotient of triplets in the superposed, half orbital series, divided in step 3–2, $\times 10^3$:

$975/531 \times 10^3 = 1836.158 \sim p/e$

The inverse $e/p = 544.6. \times 10^{-6}$

Cf. that a cell could be thought of as inversion of an atom with the nuclear force inverted to hydrophobic bonds between fatty acids and the shared lack to full electron shells among atoms transformed to inner covalent bonds in proteins, which are responsible for most of the relative negative inner charge of a cell.

5.4. The x^3 series, 125−64−27−8−1−0

5.4.1. Numbers 544, and 12-group 2×385

In the middle step 3−2 of this chain, the factors in Table 1 (Section 2.1) are found, 27 and 8, and in the columns the 11-factor divided in 5 and 6: 5×35=175, the G1-column, 6×35=210, the C1-column.

The square root of 8/27 approximates to the e/p-quotient $\times 10^3$ above.

It is noteworthy that inverted squares of 27 and 8 ($\times 10^3$) give numbers close to the division of 544 on the G1- and C1-groups of aa:

$\sqrt{27/8} \rightarrow /\backslash, \times 10^3 = 544.33$
$3^{3/2}$, /\, $\times 10^3 = 192.45. \sim G1+1$
$2^{3/2}$, /\, $\times 10^3 = 353.55. \sim C1$

The sum 546, ×6, is 3276, the total of 24 aa unbound. (A "condensed" elementary chain 5–4–3–2–1∼543+3∼546.)

An association may be allowed to numbers 27 and 8 in particle physics (Gell-Mann and Ne'eman 1964); this mentioned with no closer knowledge in that field.

The sum 27+8=35, the interval 19, and 27 disintegrated into interval 19 and 8, times 11=209 and 176, sums −/+1 in Table 1, the 12-group 770, 385×2. (The number of H-atoms in R.-chains was 19×8.)

35 divided 3 to 2 into 21 and 14, ×11 gave the sums 231 and 154 of aa groups within Table 1. Dividing the middle interval 19 in two numbers with difference 9, as 32−23, gives the numbers 5 and 14. 5×27=135, the atomic mass of the A-base, 14×8 gives 112, the atomic mass of the U-base, their difference=23.

The first number 125, ×6, is 750, and −/+ 6(27+8),=−/+ 210, gives 540 and 960, approximating the ES-chain, and 6 times the intermediate number 64=385−1. (The whole series=15^2 is the square of the sum of the basic series 5−0 times 100.)

5.4.2. Number of codons

There are 27 out of 64 possible codons if 3 stop codons are included, disintegrating into 3×8 codons for aa when grouped after 3rd base, and the stop codons. (Total number of codons for aa with all synonyms included=61 and might perhaps derive from the first interval 125–64 in this x^3-chain.)

There are a total of 5 bases, T, U, C, G, A; 4 inwards to DNA, 4 outwards in RNA, becoming 3 or 2 (called "degenerated") in codons and 1 in nucleotides in their function as coenzymes. This may be a way to see the series 5 to 1 that also locates codon triplets to the middle step.

5.5. The simple series x^1 read as triplets

This series read as triplets gives a chain that closely correspond to main codon groups of aa.

543 - 432 - 321 - 210 = 543 + 963

543 + .1 = G + C; 963 - 3 = U + A

Or 2(543 + 210) = 2 × 753 = total aa R + 2.

The series can be interpreted as steps of disintegration from total atomic mass of G+C to neutrons to polarizations into charge (from bottom upwards):

210−1=Z in UG+UC+AC+AG *

321−2=Z in UU+UA+AU+AA;…Z-sum 528

432=N in U1+A1

543+1=A in G1+C1

*G and C return in the last step as 2nd bases.

Doubled, this series is nearly the same as later steps in the chain in Section 5.2.6 (Fig. 13) from 1760 to 416 with reduction of 224 per step.

A similar division of number 753-1 also appears in groups of aa from both glycolysis and the citrate cycle with regard to 3rd base:

3rd base A/G (and A or G)=319=321−2

3rd base U/C+ "2-base-coded"=433=432+1

(If the "2-base-coded" aa are added to the A/G-group, these groups sum up to=975 −2; the U/C-group is 531; i.e., a division close to the superposed chain in Section 5.3.)

The steps 111 in the triplet series are 3/2 × 74; 74 the usual atomic mass of B-chains unbound. It could perhaps be a reason why the protein synthesis proceeds. Cf. the interval 5′ to 3′ in the ES-chain=84=3/2 × 56, the bound B-chains. (Cf. for number 111 Shcherbak, 1994).

5.6. The x^4 series ($x=5-0$)

This series 625−256−81−16−1 gave main codon groups in a way that is reminiscent of the ES-chain, where -/+ 3rd number in the chain gave those groups:

625−81=544=G+C.

625+256+81=962=U+A, +2

Minus/plus the number 81 gave the 4 codon-groups of aa in 1st base order from ES-numbers 544 halved and 416:

1/2 × 544−81=G1, 1/2 × 544+81=C1

544−81=U1, 416+81=A1

81 (= His R) is also=H_2PO_3∼ . Adding +16+1 in the x^4-chain gives 98=H_3PO_4.

6. Transformations between number base systems

6.1. Codon bases transformed to amino acids

All geometrical dimensions should surely be expected as developed in a cell simultaneously, and they are here assumed as interdependent through transformations into one another. One simple example is the geometries of proteins, forming linear threads, sheets and globular structures.

One thought in this research was that different dimensional degrees could be associated with different number-base systems (nb-*x*), as nb-10, nb-8, nb-6, for *x*=5, 4, 3. A first test on the atomic masses of codon bases showed the astonishing "connections" with main groups of aa (Fig. 15). Many other such relations were discovered with codon-grouped aa and with the ES-series. Some of the main findings are presented below.

	nb-10	nb-8	24 aa R
G:	151	——> 227	
C:	111	——> 157...Sum **384***, x 2 = 768, ∼ **770**	
U:	112	——> 160	
A:	135	——> 207...Sum **367,** x 2 = **734**	
Σ	509		
x 2 = 1018		——> **1772 = 24 B-chains unbound**	

* G+C transformed together = 386 in nb-8.

Fig. 15. Two sets of RNA-bases in nb-10, transformed separately, give codon-groups of aa R in nb-8. Transformed together (sum 1018) they give exactly the sum of 24 B-chains unbound, 1772.

Derived numbers in lower nb-*x* systems are regarded as "references" to other atomic mass sums in nb-10 and may therefore be repeatedly transformed. All operations of numbers in lower nb-*x* systems are performed in nb-10.

Indexes for *x* in nb-*x* are often used and numbers in nb-8 and nb-6 are often rewritten with digits from nb-10. One of many open questions is how to interpret nb-16.

Two sets of the G+C-bases in nb-8=768 give a total sum 3276 in nb-6.

768-**10**→3276-**6**, the 24 aa R+B unbound in nb-10.

The sum of the four bases in nb-8 were=752 −/+1,∼sum of 5′+4′+3′ in the ES-chain:

752-**10**→2848-**6**, the 24 aa R+B bound in nb-10

Half 752=376 as a number in nb-16:

376-**16**→886-**10**=12 B-chains unbound

6.2. From base pairs to sets of nucleotides

Two sets of base-pairs in nb-8 gave further transformed two sets of the four whole nucleotides in chain binding uncharged (G, C, U, A=345, 305, 306, 329; T=320): 2(G-8+C-8)=768 gives *RNA* and 2(U-**8**+A-**8**)=734 gives *DNA*. Here, an arrow represents a step nb-10 to nb-8:

768→1400|1400→2570=2 × 1285 (RNA)

734→1336|1336→2470=2 × 1235 (DNA)

6.3. Total sum of aa unbound

Number 3276-**10**, the total of unbound aa R+B=CCC in nb-16, which could be written 12.12.12 as illustrating atomic mass of C-atoms in displaced positions, guided by a factor 16 (oxygen?).

CC in nb-16=204 in nb-10, the atomic mass of Trp R+B, the heaviest aa.

6.4. Bases in nb-8 in the ES-chain

Sum of G+U+C in nb-8=544, A-base 208–1 (Fig. 16). U-**8** 160 and C-**8** 157 approximates to the number 2′=159 in the ES-series, together 317. In nb-10 number 385 is the interval 544 to 159. Here G-**8** becomes the same interval to the sum of U-**8**+C-**8**. Cf. that the G-base can bind to both.

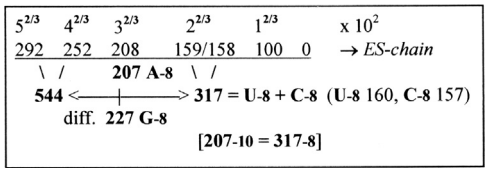

Fig. 16. Numbers of the bases transformed to nb-8 in the ES-chain.

6.5. Generation of the 12-groups from 385

From numbers 177 (=544–367) and 208 with the sum 385 in the ES-chain the 12-groups 770 and 734 were generated through repeated transformations nb-10→nb-8, indexes here not marked:

177→261→385, ×2=770
208→320, rewritten 318→476→734

The interval 177 to 385 above =208. Interval 318 to 734=416 gets divided in the two steps 158 and 258, ~numbers 159 and 159+100 (−1) in the ES-chain.

6.6. Five times numbers in the ES-chain:

Number 544 may be divided in 292 and 252 or in 336 (544−208) and 208. It was found that 5 times these ES-numbers in nb-8 (Fig. 17) gave the division on codon groups of aa in nb-10. Remarkably, the sum 10×544 in nb-8 gave the total 2848 in nb-10, sum of the 24 aa R+B bound.

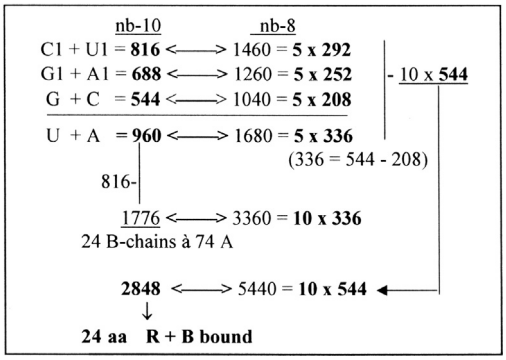

Fig. 17. Five times the ES-parts 292 and 252 of 544 from nb-8 to nb-10 give the codon groups U1+C1 and G1+A1, and five times the secondary ES-parts 336 and 208 of 544 from nb-8 to nb-10 give G+C and U+A. Ten times 544 in nb-8 give in nb-10 the sum of 24 aa bound.

In addition, 5(208+292)=2500. As a number in nb-8 it gives 1344 in nb-10, the sum of the 24 B-chains bound.

6.7. H-atoms in R-chains

The sum of H-atoms were 152=252−100 in the ES-chain, an interval that can be divided in 44 = step 4′−3′ and 108, step 3′−1′. Transforming these parts from nb-16 to nb-6 gives the atomic mass of H-atoms in R and the atomic mass of N=676=26^2 (also *Z* without H).

44-**16**→152-**6**, ~H in nb-10; (+108)
108-**16**→676-**6** (rewritten)=N in nb-10

6.8. B-chains

The first number 292 in the ES-chain transformed from nb-10 to nb-8=444, ~6 B-chains à 74A.

74 in nb-10=112 in nb-8, i.e., the sum of 2 bound B-chains in nb-10. This could possibly be a factor behind repeated peptide bonds?

6.9. Number of aa

The number of "classical" aa is 20. Why have 4 of these aa two codons? One aspect in this context could be that 20 in nb-10 = 24 in nb-8.

6.10. P-contenting nucleotides and coenzymes

AUG codes for Met, which starts the protein synthesis in most cells. The three nucleotides A, U and G transformed from nb-10 to nb-8 was found to give the sum 1504 of the 24 aa (Fig. 18).

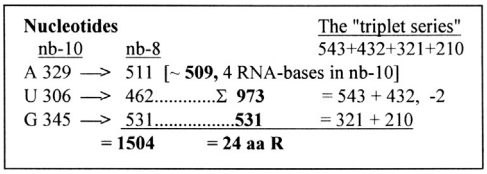

Fig. 18. AUG-nucleotides transformed to the sum of all 24 aa R.

P-groups H_3PO_4= 98 and HPO_3^-=80. Transformations of these P-group numbers gave one to three P-groups +ribose when bound to bases or in chains (e.g. 3×98+150, −4×18=372), as if it was an operator in the formation of nucleotides:

80-**10**→98-**8**=−+H_2O
98-**16**→372-**6**=P~P~P-ribose~
80-**16**→292-**6**=P~P-ribose~
80-**10**→212-**6**=P~ribose~

Rewriting the numbers 372 to 412 and 292 to 332 in nb-6 gives +80. The sum 412+332=744, equivalent to the atomic mass of NADP.

6.11. Fatty acids

Fatty acids with their zigzag lines could be a 3rd way to read the hexagonal pattern in Fig. 1. Cf. that P-lipids form hexagons at certain temperatures. Fatty acids are as essential to life as the genetic code. Two of the most common fatty acids when transformed to nb-6 give 3 times the numbers 367 and 385, +/-1, a relation to R-chains of the 24 aa in nb-8=3/2:

$C_{16}H_{32}O_2$: 256-**10**→1104-**6**=**3×368**
$C_{18}H_{36}O_2$: 284-**10**→1152-**6**=**3×384**

(For more material, see www.u5d.net/Genetic-code/index.html.)

7. Discussion

The correlations between the "classical" genetic code and the ES-chain seem too many to ignore. The frozen hazard theory (Crick, 1968) and the theory of evolution towards an error minimizing system concerns the distribution of *codons* in the genetic code. Yet, it is the choice of encoded *aa* that in the first place seems to be a random one. Why, for instance, exactly ten oxygen in R-chains of aa and twelve nitrogen, even if both atom types surely are needed besides the hydrophobic ones with only CH*x*? The choice indicates that another factor, like for instance the

nucleon number or chemical energy of aa in relation to RNA-strands may have been decisive.

The fact that such correlations here were detected for seemingly *independent* properties, as for both the atomic mass distribution on codon domains and on atom kinds, supports in itself the general suggestion that the code was built on an arithmetical scheme. Other such examples were found in groups of B-chains, in parents to the bases and in origin of aa from stations in glycolysis and the citrate cycle.

The ES-chain (and more elementary versions of such a series) seems to operate along different coordinate axes and levels. The 2-dimensional Table 1 for mixed codons is not the only one of that kind. (Counting number of atoms in 48 coding bases, 1st and 2nd position, C+N atoms equal exactly the number of atoms in the A+G-bases and O+H equals exactly the number of atoms in the U+C-bases. The total is precisely 2/3 of the whole ES-chain.)

Many different and intricate relations between the $2x^2$-series and the genetic code were detected. In that series behind the periodic system, times a factor 16, both atomic mass and charge numbers for main codon domains of aa could be read. The numbers of aa in different codon groups correspond to orbital numbers in the periodic system. It demonstrates a shared serial character and may be one example of the physical connection of the code to underlying atomic levels.

A main objection against the hypothesis here is that currently no accepted mechanism is known that could explain a construction of molecules along a numeral series. Yet, the "key and lock" hypotheses and the vague term "affinity" are used in explanations but hardly understood, and numeral series are established facts in the mentioned examples, the formulas for spectral lines of hydrogen and the periodic system. When there are organizing principles at the atomic level, why could not similar principles appear also at the molecular level?

Both the mentioned atomic schemes may have served as intermediaries between the arithmetical and biochemical levels. (For example, the quotients between the spectral lines of hydrogen (a) 0.21, (b) 0.1875, and (c) 0.13889 in the Balmer series are: a/b $\times 10^2 = 112 =$ atomic mass of the U-base, b/c $\times 10^2 = 135 =$ atomic mass of the A-base and a/c $\times 10^2 = 151.2$, $\approx$ atomic mass of the G-base 151.)

It may be added that scientists usually first detect patterns and regularities, then invent the cause, as gravity in macrocosm (or bending of space-time). Spin as a quantum number of elementary particles was invented by Pauli to explain his "exclusion principle" for pairs of electrons. New mathematics is also invented when needed.

Furthermore, it may be questioned in which sense the periodic system is "explained" in terms of cause and effect, with the "octet rule" not well understood, although assuredly tested.

About evolution, a numeral series as a "cause" to the actual code would not necessary represent a very first state, later hid by growing complexity and processes driven by the electromagnetic force. It should be remembered that the periodic system did not "exist" as a recognizable pattern in the first materialized Universe and have appeared first after many evolutionary steps of fusion in stars to all the known elements. What may be regarded as a cause in direction of disintegration may appear last in synthesizing direction, (cf. Fig. 2, Section 1.2).

Evolution of the code might be illustrated by the approximate numbers for main domains in the simpler series (Section 5): a stepwise evolution of the code towards the middle step when exponents are assumed as expressions for dimensions.

$x^4 \rightarrow (3 \text{ or } 6)x^3 \rightarrow x^{2/3} \leftarrow (2)x^2 \leftarrow x^1$

A similar pattern of two-way direction can be recognized on the higher level of protein synthesis, where mRNA moves outwards from DNA to ribosomes, and tRNAs as individualized units, come as from the opposite strand of DNA to meet mRNA "the other way around".

"Evolution" in more biological terms is an almost inevitable concept since Darwin. Yet, the question should be raised, at least once, if there really have been an evolution of the genetic code in the sense of gradual synthesis or incorporation of new aa; this as long as nothing like a cell has been found that can manage without the 20 "classical" ones. (With modern technology it will perhaps be possible to test this, depriving a cell its possibility to synthesize a certain aa or to get it from the environment?)

The natural thought that the number of aa developed from simpler to more complex ones in the code could be a false analogy with chemical elements. It contradicts the fact that continual breaking-down and synthesis characterizes the life of a cell. Such a double-direction as inwards–outwards in geometrical terms should have been fundamental also to separate and define a unit as such and further its main structural polarity in "circular" membranes and a "radial" protein skeleton. The opposite directions are most clearly expressed in the hydrophobic–hydrophilic polarity among aa. (Cf. gravity as an answer to Big Bang!)

It would still be possible to imagine that the first G1-coded aa as Gly, Ala, Val and Asp, those which most easily get synthesized in the Miller-type of experiments, got connected with RNA in an *open* environment. The G1-coded aa come "first" also in the ES-chain, however, from the highest number in the chain.

With the background model here in mind (Section 1.2 and Fig. 2) the directions of "first" and "last" become indeed ambiguous.

Gradual changes of the coding system itself, inside a cell, is something else. About mass as an essential property, Di Giulio (1989) examined relationships between codons when a single base in the triplets was thought to be exchanged by mutations or translational errors. He found in a cluster analysis of these codon relations high correlations with only two physico-chemical properties, chiefly with polarity but also with molecular *volume*, both statistically significant. Apart from the G2-group of aa his clusters correlated also with 2nd base groups. He had the interesting idea that it could be the deviations from the mean value of mass of aa that had importance sooner than the mass as such and found it in an analysis with a few exceptions highly significant. This thought certainly includes the concepts *double-direction* and *polarization.*

(His exceptions were Arg, Cys and Tyr, whose mass numbers 101, 47 and 107 somehow happen to be the deviations from ES-numbers in aa-groups G1-C1, A1-U1 and U2-A2; cf. Section 2.2).

About distances of aa masses from a mean value, it may be added that counting on 24 aa, the 14 lightest aa below the mean value (here nearly 63) has the atomic mass 600, the 10 above the mean value 904, close to a 2/3 relation.

When Di Giulio analyzed the correlation (negative) between number of synonymous codons for aa and molecular weight, excluding Arg, he found it highly significant ($p < 0.001$). It seems rather natural, since 6 of the 8 aa with indifferent 3rd base, equivalent with many synonyms, here shortened the "2B"-coded, are the lightest ones. This invites the thought that an early code of RNA could have had only 2 bases, and also the idea of an evolution towards heavier aa. Di Giulio discusses such thoughts mostly through references to others, not excluding any of their aspects.

Here only two first annotations to this discussion, more below. If some aa originated from tRNA, none of the A2-coded aa or Ile2 could have been derived from existing "2B"-codons.

A mentioned alternative is that early tRNAs (or tRNA syntethases?) were less specified and could have accepted whole groups of aa. This would agree with the "domain" concept (Wong, 1975) and also with the ES-chain. However, it should be asked which character in that case that may have defined such groups? Codons? Degree of polarity? Size? Structure? Does the answer necessary lie in the "key and lock" conception?

Take for instance the geometrical structure, more difficult to give a measure (Section 3.6.4). It connect branching in a wider sense with the middle step in the ES-chain and atomic masses with the hydrophobic–hydrophilic polarity that the 2nd base in codons since long has been associated with. This structural division implies another way of grouping aa, as along another coordinate axis. When Di Giulio (in this early article) says that a 2-dimensional representation (of polarity and molecular weight) is sufficient to describe the organization of the code, it can really be questioned.

Another property was suggested by Taylor and Coates (1989), the pathways from the five stations in glycolysis and the citrate cycle along which the aa now in general are seen constructed. They suggest a scheme where the 3rd base in the codons gives information about mass, the 2nd base about polarity and the 1st base, apart from G1, information about pathways for the origin of aa. As for the relevance of mass, they explicitly refer to the fact that 6 out of 8 aa with no information in the 3rd base make up the start of the weight series of aa.

Their scheme fits nicely with the underlying dimensional views in this article and the background model, roughly: pathways as 1-dimensional and charge, expressed in polarity, as a 2-dimensional property in relation to mass as 3-dimensional when analyzed as one aspect on volumes.

Unfortunately, some of their pathways to get the scheme are hard to accept. With reference to the map by Nicholson (1976) there are more straight ways from pyruvate to the U2-coded Val *and Ile* and via acetyl-Coa to the CU-and UU-coded Leu (on pathways to fatty acids), with all bases represented in the 1st position of their codons; hydrophobicity is clearly the shared property. (Met is also more naturally derived from homocysteine on a way from Cys.)

However, the authors emphasize an essential thing: "There are multiple levels at which selective forces operate, and an explanation at one level does not eliminate those operating at other levels." This should imply that we can expect the deeper property of mass stretching along the whole codon for an aa – and surely along all codons as in this research, and the next higher level of charge (polar–non polar aa) acting also at the level above it.

Furthermore, these levels should have their roots in still deeper levels of physics and under that level pure math, levels where we could expect more purely quantitative, numerical and geometrical aspects on the code and the unification of codons into the whole coding system.

Taylor and Coates did not take up a definite position about evolution of the code from an eventual earlier "2B"-code to the actual triplets, only talks about a "gradual use" of the 3rd base for the property of mass. They exclude the "extras" as the AG-codons for Arg and Ser from their scheme, assuming them as "later" developed.

Hasegawa and Miyata (1980) did not exclude these "extras" for Arg, Ser and Leu when they proposed an earlier, more symmetric table of codons than the actual one. Their table, with rows and columns in the order C–G–U–A, gets the form of what they call two big "L" (if turned, one for the "2B-coded" upside down, one for the divided squares for aa with differentiating 3rd base turned left–right, hence, complementary along two axes; added here).

The left "L" with 8 "2B"-coded aa includes the 6 lightest aa, and the authors emphasize this correlation with molecular weight as surely not accidental.

About this "2B"-group of codons the authors refer to others, pointing to the stronger bonds at translation with G and C in first two positions (or at least in one of these), which had been proposed to be a reason for a "2-out of 3" reading. It should be noted here that the relation between 3 hydrogen bonds in the G–C-pair and 2 in the U–A-pair also implies a step of *polarization* 2→1 of attached atoms to the rings: both oxygen (O) and nitrogen (N) in the G- and C-bases, only O in the U-base, only N in the A-base.

For the right "L" in the table, the authors only count on a halving of the squares through the 3rd base Y (U/C) or R (A/G), referring to the "wobbling rule" of Crick 1966. (About "wobbling", cf. perhaps Fig. 16, Section 6.) However, both the parents of the bases, orotate as well as hypoxanthine, seem to have had a big numeral impact on the grouping of aa on codons in 1st base order (Section 2.12).

The authors get the biggest problem for their more symmetric table with the AUA-coded Ile2, irregularly invading an eventual, earlier Met-codon AUR found in some mitochondria. They have excluded it and include instead one stop-codon UAR to get a nice, regular 8-8-8-codon table.

This AUA-codon may also seem a bit odd in the 24-codon table for only aa that was the start of this research. Yet, it exists, also in primitive cells. Did it perhaps become essential as the anticodon to Tyr UAU to stabilize the process where the atomic mass of Tyr (3′ - 1′,-1) expresses the mass deviations of U2- and A2-groups of aa (437 and 523) from main number domains in the ES-chain, 544 and 416 (Section 2.6)?

Another guess may be an origin in the superposed chain of halved orbital numbers 9-7-5-3-1 (Section 5.3): atomic masses of Met 75 and Ile 57 mirror one another, and we have the sum 575 (read→←?) in this U+A-group of aa with non-mixed codons. AUA could be a special code for expressing this double-direction?

A third function could be to stress the extra derivation loop from Thr to Ile as a kind of "feed back" from the citrate cycle to glycolysis?

In all three guesses it would in different ways be a sign of double-direction.

With the aspects in this research, does anything indicate an evolution from an earlier simpler "2B-code" to the triplet one, or that the construction of mostly lighter and simpler aa preceded the heavier ones? One example could eventually be the 11-factor among mixed codons in sums of aa of the type UG+GU, CA+AC (Section 2.8), where only the first two bases in the UG- and CA-codons are regarded. Another could be to apply a view of polarizations: 64→32 (the "2B-codons")→16 (U/C-codons), and the resting 16 divided in 10A/G-codons, 6 divided in 3 A *or* G as the 3rd base and 3 stop-codons. However, this view departs from 64, the possible *triplets*.

Following two arithmetical observations of structural relations in the code contradict such a unidirectional view.

In the ES-chain the first three numbers and intervals are 292–(40)–252–(44)–208.

$4 \times 292 = 2 \times 584$, the sums of aa with differentiated 3rd base in the groups A1+G1 and with +1 in the groups U1+C1 (Section 2.11).

4×40, the first interval, 160, $-1 = 159$, the aa with "2B"-codons of only G and C.

$4 \times 44 = 176$ is the sum of "2B"-coded aa among mixed codons.

Hence, the "2B"-coded aa could be derived as intervals from first two steps, perhaps to see as debranched and meeting 'the other way around' in accordance with Fig. 2 (Section 1.2). There is the relation between the first and last base through the translational process.

The other observation concerns the $2x^2$-chain behind the periodic system in Section 5.2:

2 out of 14 aa in the U+A-group are "2B"-coded, 6 out of 10 in the G+C-group, so in both the 1st and 2nd base order. A similarity with the numbers of *s-p-d-f*-orbitals 2–6–10–14 appears (cf. Fig. 2, Section 1.2).

If we now build "Pascal's triangles" or two-peak number pyramids on the $2x^2$-chain 2–8–18–32–50 as level 0 and add the numbers on level 1, 2 and 3, we get **158** on the side of the left pyramid built on 2-to-32 and **574** as total of the right 8-to-50

pyramid; interval 416. It is the division in "2B"- and "3B"-coded aa among *non-mixed* codons (Table 2, Section 2.1), −1 in each number.

With the same two-peak pyramids built on the *cumulative* series 2–10–28–60–110, the left peak, built on numbers 2-to-60, is **176**, the right peak, built on numbers 10-to-110, is **384** (interval 208). It is the numbers for row 3 ("2B"-coded aa) and row 1 of aa with *mixed codons*, −1 in interval (row 2) and row 1 (Table 1, Section 2.1).

If these correlations reveal a truth about the genetic code on a mathematical level, we can note the displacement of only one serial step between the two-peak pyramids, reminding of the K- and L-shells and the "octet rule" in chemistry, but they are deeply involved in one another. Numbers up to 32 are involved also in the sum 158 above, which simultaneously is the sum of level 1 in the right 574 pyramid. (What the first pair of pyramids have in common is 152 = 4′ - 1′ in the ES-chain, equal to the number of H-atoms in aa side-chains; in the second pyramids 252=4′.)

A conclusion would be that the very first "2B"-coded aa perhaps may have initiated the process but at the same time draw out the scheme for the whole code, developed along an axis vertical to the usually assumed serial development of the periodic system.

About the questioned relevance of atomic mass, just a couple of additional remarks. Much research has been focused on the stability of the code. One reason for its attributed stability could reasonably be the use of most common isotopes. Recently, Sassi et al. (2014) reported that carbon-14 as an important β-emitter is a source for non-canonical bases in DNA.

Another argument is that the aa Ile sometimes get mistaken for Leu by tRNAs in spite of different structures but with the same atomic mass and atoms. This suggests that these properties are essential. Mass has also been used as one factor to predict the destination and functions of proteins (Chou, 2005) with obviously good results.

The existence of Englert–Higgs mass field has been accepted by the detection of Higgs boson. However, it still does not explain the masses of individual elementary particles. Cf. aa as such particles on the cellular level. There is always much more to be discovered.

How should the striking relations in transformations between nb-systems be interpreted? A number 770 in one language, nb-10, may refer to atomic mass of the 12-group of aa R with mixed codons, in another language, nb-8, to the added atomic mass of two base pairs. (10 US-dollars are not the same as 10 crowns, but the two currencies may be bound to each other.)

The bases are essentially built by aa. In this sense they represent the "same" thing expressed in different languages, applying the suggested view on their relation in Fig. 1.

The base structures imply a step of sp-hybridization from tetrahedrons of aa, a dimensional step 4→3 or 3→2 in a simpler view. This may be one reason for a nearly unavoidable association of different nb-systems with d-degrees.

That 24 aa R+B unbound were obtained from 16 bases (Fig. 15, Section 6) also shows on a relation 2 to 3 in number. (In nb-2 the number 16=24 rewritten. About lower nb-systems, the series 50–40–30–20–10 read in nb-10, 8, 6, 4 and 2 respectively gives the $2x^2$-chain for the periodic system, 50–32–18–8–2.)

It is especially noteworthy that in both operations in Figs. 15 and 17 the transformation of *summed* numbers give the B-chains too, only the B-chains unbound in Fig. 15, the bound B-chains included in Fig. 17. This finding indicates that, among other things, the polarization of aa in R- and B-chains may be seen as a secondary step in disintegrating direction to a right-angled relation.

The main domains of aa in Fig. 17 transformed from ES-numbers and especially from number 544 not only supports the relevance of the ES-chain but also the suggestion (Section 5.3.2) that this number (=5′+4′ in the ES-series) has a deep physical root, eventually in the e^-/p^+ quotient. How to join that thought with the ES-chain? A partial answer could be that 544 in nb-6=208 in nb-10=3′ in that chain.

Thus, it is the hypothesis here that such transformations between nb-systems reveal a hidden reference system in the code and its processes. The abundance of possibilities is of course a problem and even more so when rewritings are included.

Naturally, the findings here do not exclude other presented patterns and theories. The attractive stereochemical theory (Yarus, 1998) reminds of the illustration in Fig. 1, which suggests that such a kind of close affinity between aa and codons could have a deeper foundation, before time factors evolved that transformed structure units into processes. The stereochemical theory is compatible with other proposals (Sukhodolets, 1989) that a whole unit of aa and RNA preceded the disintegration, implicit also in the numeral series in this study.

With the opposite synthesizing aspect, one proposal concerns how aa could have been created on RNA (Copley et al., 2005). Assuming such a view, in spite of first arguments above, some numeral similarities between their Fig. 1 and the ES-chain may be noted: a P-P-ribose group = 5′, binding two P-ribose groups=2 (4′+3′) of a dinucleotide, and the base pair bound circa 2′+1′, G+C or A+T. (The bonds, −18 and −36, within the P-ribose-parts and to the bases here neglected.) It should imply a growth of aa between ribose and the bases as in a middle step in the ES-chain.

If all began with RNA-strands capturing or creating aa, or aa "came first", creating RNA bases, seems to be a very open question hitherto.

The proposed ES-chain here is in several ways consistent with the coevolution theory (Wong, 1975, 2005). There is the connection with the biochemical origins of aa from glycolysis and the citrate cycle and the view of codon domains as totals, differentiated in following steps, although here related to the atomic masses. The G1-coded aa "arrive first" in the ES-chain too, 5 out of the 7 aa assumed as primary in that theory, GG–GC–GU–GA–GA besides Ser UC and Phe UU. However, are really the origins of aa today in agreement with the metabolic processes in anaerobic, very early bacteria?

The numeral approach in itself in this research seems justified also by plants with their 5-, 4- and 3-merous plans, where the 5- and 4-merous plants are dicotyledons, the 3-merous monocotyledons derived from "halved" seeds.

Finally, about the ES-exponent 2/3, it may be mentioned that the Cheops pyramid has the height of ~1/2 × 5′ and the side of its base ~1/2 (4′+3′) in meters. (The Pharaoh measure "ell" is said to have been about half a meter.) Thus, the first three ES-numbers with exponent 2/3 seems to have been known a very long time ago.

Conflict of interest

The author declares no conflict of interest.

Appendix. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jtbi.2015.01.013.

References

Chou, KC, 2005. Using amphiphilic pseudo amino acid composition to predict enzyme subfamily classes. Bioinformatics 21.1, 10–19.
Copley, SD, Smith, E, Morowitz, HJ, 2005. A mechanism for the association of amino acids with their codons and the origin of the genetic code. Proc. Natl. Acad. Sci. 102, 4442–4447.
Crick, FH, 1966. Codon-anticodon pairing: the wobble hypothesis. J. Mol. Biol. 19, 548–555.
Di Giulio, M, 1989. Some aspects of the organization and evolution of the genetic code. J. Mol. Evol. 29, 191–201.
Downes, AM, Richardson, J, 2002. Relationships between genomic base content and distribution of mass in coded proteins. J. Mol. Evol. 55, 476–490.
Gell-Mann, M., Ne'eman, Y., 1964. The Eightfold Way. W. A. Benjamin, Inc., New York, Amsterdam p. 15, 29, 85.

Hasegawa, M, Miyata, T, 1980. On the antisymmetry of the amino acid code table. Orig. Life 10, 265–270.
Karlson, P, 1974. Biokemi (*Biochemistry*).. LiberLäromedel, Lund.
Nicholson, D.E, 1976. Metabolic Pathways. Koch-Light Laboratories Ltd, Coinbrook, Bucks, England.
Perez, JC, 2010. Codon populations in single-stranded whole human genome DNA are fractal and fine-tuned by the Golden Ratio 1.618. Interdiscip. Sci. 2 (3), 228–240.
Rakočević, MM, 2004. A harmonic structure of the genetic code. J. Theor. Biol. 229, 221–234.
Rumer, Yu.B., 1966. On systematizing codons in the genetic code. Dokl. Akad. Nauk SSSR 167 (6), 1394.
Sassi, M, Carter, DJ, Uberuaga, BP, Stanek, CR, Marks, NA, 2014. Carbon-14 decay as a source of non-canonical bases in DNA. Biochim. Biophys. Acta , http://dx.doi.org/10.1016/j.bbagen.2013.10.003.
Shcherbak, VI, 1993. Twenty canonical amino acids of the genetic code: the arithmetical regularities. Part I. Theor. Biol. 162 (3), 399–401.
Shcherbak, VI, 1994. Sixty-four triplets and 20 canonical amino acids of the genetic code: the arithmetical regularities. Part II. J. Theor. Biol. 166 (4), 475–477.
Shcherbak, VI, 2003. Arithmetic inside the universal genetic code. Biosystems 70 (3), 187–209.
Sukhodolets, VV, 1989. The genetic code as a clue to understanding of molecular evolution. J. Theor. Biol. 141 (3), 379–389.
Taylor, F.J.R, Coates, D, 1989. The code within the codons. BioSystems 22, 177–187.
Wong, J.T.F., 1975. A co-evolution theory of the genetic code. Proc. Natl. Acad. Sci. USA 72 (5), 1909–1912.
Wong, JT, 2005. Coevolution theory of the genetic code at age thirty. Bioessays 27, 416–425.
Yarus, M, 1998. Amino acids as RNA ligands: a direct-RNA-template theory for the code's origin. J. Mol. Evol. 47, 109–117.