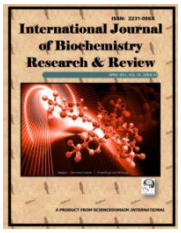


International Journal of Biochemistry Research & Review
1(3): 82-107, 2011

SCIENCEDOMAIN *international*
www.sciencedomain.org

Algorithm of Insulin Human P01308 - Position 102 - Determinante 2029

Lutvo Kurić[1*]

[1]*Kalinska 7/6 72290, Novi Travnik, Bosnia and Herzegovina.*

Research Article

Received 11th July 2011
Accepted 7th August 2011
Online Ready 28th August 2011

ABSTRACT

The modern science mainly treats the biochemical basis of sequencing in bio-macromolecules and processes in medicine and biochemistry. One can ask weather the language of biochemistry is the adequate scientific language to explain the phenomenon in that science. Is there maybe some other language, out of biochemistry, that determines how the biochemical processes will function and what the structure and organization of life systems will be? The research results provide some answers to these questions. They reveal to us that the process of sequencing in bio-macromolecules is conditioned and determined not only through biochemical, but also through cybernetic and information principles. Many studies have indicated that analysis of protein sequence codes and various sequence-based prediction approaches, such as predicting drug-target interaction networks (He et al., 2010), predicting functions of proteins (Hu et al., 2011; Kannan et al., 2008), analysis and prediction of the metabolic stability of proteins (Huang et al., 2010), predicting the network of substrate-enzyme-product triads (Chen et al., 2010), membrane protein type prediction (Cai and Chou, 2006; Cai et al., 2003; Cai et al., 2004), protein structural class prediction (Cai et al., 2006; Ding et al., 2007), protein secondary structure prediction (Chen et al., 2009; Ding et al., 2009b), enzyme family class prediction (Cai et al., 2005; Ding et al., 2009a; Wang et al., 2010), identifying cyclin proteins (Mohabatkar, 2010), protein subcellular location prediction (Chou and Shen, 2010a; Chou and Shen, 2010b; Kandaswamy et al., 2010; Liu et al., 2010), among many others as summarized in a recent review (Chou, 2011), can timely provide very useful information and insights for both basic research and drug design and hence are widely welcome by science community. The present study is attempted to develop a novel sequence-based method for studying insulin in hopes that it may become a useful tool in the relevant areas.

**Corresponding author: Email: lutvokuric@yahoo.com;*

Keywords: Human insulin; algorithm; insulin model; insulin code; determinante of insulin;

1. INTRODUCTION

The biologic role of any given protein in essential life processes, eg, insulin, depends on the positioning of its component amino acids, and is understood by the “positioning of letters forming words“. Each of these words has its biochemical base. If this base is expressed by corresponding discrete numbers, it can be seen that any given base has its own program, along with its own unique cybernetics and information characteristics.

Indeed, the sequencing of the molecule is determined not only by distin biochemical features, but also by cybernetic and information principles. For this reason, research in this field deals more with the quantitative rather than qualitative characteristcs of genetic information and its biochemical basis. For the purposes of this paper, specific physical and chemical factors have been selected in order to express the genetic information for insulin. Numerical values are them assigned to these factors, enabling them to be measured. In this way it is possible to determine oif a connection really exists between the quantitative ratios in the process of transfer of genetic information and the qualitative appearance of the insulin molecule. To select these factors, preference is given to classical physical and chemical parameters, including the number of atoms in the relevant amino acids, their analog values, the position in these amino acids in the peptide chain, and their frenquencies. There is a arge numbers of these parameters, and each of their gives important genetic information. Going through this process, it becomes clear that there is a mathematical relationship between quantitative ratios and the qualitative appearance of the biochemical “genetic processes“ and that there is a measurement method that can be used to describe the biochemistry of insulin.

2. MATERIALS AND METHODS

The biologic role of any given protein in essential life processes, e.g., insulin, depends on the positioning of its component amino acids, and is understood by the „positioning of letters forming words“. Each of these words has its biochemical base. If this base is expressed by corresponding discrete numbers, it can be seen that any given base has its own program, along with its own unique cybernetics and information characteristics. Indeed, the sequencing of the molecule is determined not only by distin biochemical features, but also by cybernetic and information principles. For this reason, research in this field deals more with the quantitative rather than qualitative characteristcs of genetic information and its biochemical basis. For the purposes of this paper, specific physical and chemical factors have been selected in order to express the genetic information for insulin. Numerical values are them assigned to these factors, enabling them to be measured. In this way it is possible to determine oif a connection really exists between the quantitative ratios in the process of transfer of genetic information and the qualitative appearance of the insulin molecule. To select these factors, preference is given to classical physical and chemical parameters, including the number of atoms in the relevant amino acids, their analog values, the position in these amino acids in the peptide chain, and their frenquencies. There is an arge numbers of these parameters, and each of their gives important genetic information. Going through this process, it becomes clear that there is a mathematical relationship between quantitative

ratios and the qualitative appearance of the biochemical „genetic processes“ and that there is a measurement method that can be used to describe the biochemistry of insulin.

Insulin can be represented by two different forms, ie, a discrete form and a sequential form. In the discrete form, a molecule of insulin is represented by a set of discrete codes or a multiple dimension vector. In the sequential form, an insulin molecule is representing by a series of amino acids according to the order of their position in the sequence lenght 110 AA. Therefore, the sequential form can naturally reflect all the information about the sequence order and lenght of an insulin molecule. The key issue is whether we can develop a different discrete method of representing an insulin molecule that will allow accommodation of partial, if not all sequence order information? Because a protein sequence is usually represented by a series of amino acids should be assigned to these codes in order to optimally convert the sequence order information into a series of numbers for the discrete form representation?

3. RESULTS

The matrix mechanism of Insulin, the evolution of biomacromolecules and, especially, the biochemical evolution of Insulin language, has been analyzed by the application of cybernetic methods, information theory and system theory, respectively. The primary structure of a molecule of Insulin is the exact specification of its atomic composition and the chemical bonds connecting those atoms.

Fig. 1. P01308 (INS_HUMAN) lenght 110 AA: Sequence length 110 AA

M	A	L	W	M	R	L	L	P	L	L	A	L	L	A	L	W	G	P	D	P	A
A	A	F	V	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G
E	R	G	F	F	Y	T	P	K	T	R	R	E	A	E	D	L	Q	V	G	Q	V
E	L	G	G	G	P	G	A	G	S	L	Q	P	L	A	L	E	G	S	L	Q	K
R	G	I	V	E	Q	C	C	T	S	I	C	S	L	Y	Q	L	E	N	Y	C	N

Aforementioned aminoacids are positioned from number 1 to 110. This positioning is of the key importance for understanding of programmatic, cybernetic and information principles in this protein. The scientific key for interpretation of bio chemical processes is the same for insulin and as well as for the other proteins and other sequences in biochemistry. The first aminoacid in this example has 20 atoms, the second one 13, the third one 22, etc. Why do they have exactly this many atoms? It is because there are many codes in the molecule of insulin, analogue codes and other coded features. In fact, there is a program-cybernetic algorithm in which it is „recorded“ that the firs amino acid has to have 20 atoms, th3, the third one 22, etc. The first amino acid has its own biochemistry, the second and the third one also. The conclusion here has to be that there is a concrete relationship between quantitative ratios in the process of transfer of genetic information and qualitative appearance, i.e., the characteristics of organisms.

3.1 Algorithm 1

We shall now give some mathematical evidences that will prove that in the biochemistry of insulin in there really is programmatic and cybernetic algorithm in which it is „recorded“, in the language of mathematics, how the molecule will be built and what will be the quantitative characteristics of the given genetic information.

Atomic progression

$$[AC1 + (AC1+ AC2) + (AC1+ AC2+ AC3)..., + (AC1+ AC2+ AC3..., + ACR)] = \mathbf{S};$$
$$AC1 = APa1;$$
$$(AC1+ AC2) = APa2 ;$$

$$(AC1+ AC2+ AC3) = APa3;$$
$$(AC1+ AC2+ AC3..., + AC306) = APaR;$$
APa1,2,3,n = Atomic progression of amino acids 1,2,3,n

Progressions can be: macro and micro, even and odd, primary and secondary, analogue, negative, positive, etc.

Cybernetic, information and system characteristics of biochemistry of insulin can be researched also using frequencies (macro and micro), primary and secondary values, standard deviations, analogue values, even and odd values, determinants, bio codes, etc.

Within the digital pictures in biochemistry, the physical and chemical parameters are in a strict compliance with programmatic, cybernetic and information principles. Each bar in the protein chain attracts only the corresponding aminoacid, and only the relevant aminoacid can be positioned at certain place in the chain. Each peptide chain can have the exact number of aminoacids necessary to meet the strictly determined mathematical conditioning. It can have as many atoms as necessary to meet the mathematical balance of the biochemical phenomenon at certain mathematical level, etc. The digital language of biochemistry has a countless number of codes and analogue codes, as well as other information content. These pictures enable us to realize the very essence of functioning of biochemical processes. There are some examples, which are shown in Table 1.

When evaluating progressions, one has to take into account the fact that there are macro and micro progressions, odd and even, primary and secondary, analogue, etc. Progressions have their category (odd and even, primary and secondary, analogue, etc.) All these progressions are in correlation with each other.

Establishing of numeric values of amino acids needs to be done through use of strictly determined criterion from the theory of systems and also from cybernetics whish, in this example, is the number of atoms in amino acids. That is only one dimension of the digital image of insulin. There are many other dimensions as well as digital images. Each of these dimensions and images has its corresponding progression. With some dimensions, one has to use some other parameters from the theory of systems and cybernetics (frequency, standard deviation, various codes and analogue codes, analogue values, primary and secondary values, odd-even relation, and many others), and not progression.

Regardless of the fact whether there is a typical correlation between parameters or not, their effect in the process of evolution can be followed through use of adequate methodology. Examples have been shown in Table 2.

Table 1. Atomic progression APa and APb (Amino acid Leu – position from3 to 105 AA)

	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Number of atoms	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22
Position AA	3	7	8	10	11	13	14	16	30	35	39	41	61	68	77	80	82	86	102	105
APa	55	150	172	211	233	268	290	325	572	652	725	771	1155	1284	1400	1459	1494	1559	1850	1916
APb	1974	1879	1857	1818	1796	1761	1739	1704	1457	1377	1304	1258	874	745	629	570	535	470	179	113
AP(a,b)	2029	2029	2029	2029	2029	2029	2029	2029	2029	2029	2029	2029	2029	2029	2029	2029	2029	2029	2029	2029

Table 2. Determinant 2029 APa and APb (Amino acid Leu – position 102)

Progression APa (From 1795 to 1198)

	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	102	102	102	102	102	102	102	102	102	102
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
APa	1850	1850	1850	1850	1850	1850	1850	1850	1850	1850
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
	1795	1700	1678	1639	1617	1582	1560	1525	1278	1198
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
APa	55	150	172	211	233	268	290	325	572	652
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	3	7	8	10	11	13	14	16	30	35

Progression APa (From 1125 to 0)

	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22
	102	102	102	102	102	102	102	102	102
	↓	↓	↓	↓	↓	↓	↓	↓	↓
APa	1850	1850	1850	1850	1850	1850	1850	1850	1850
	↓	↓	↓	↓	↓	↓	↓	↓	↓
	1125	1079	695	566	450	391	356	291	0
	↑	↑	↑	↑	↑	↑	↑	↑	↑
APa	725	771	1155	1284	1400	1459	1494	1559	1850
	↑	↑	↑	↑	↑	↑	↑	↑	↑
	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22
	3	7	8	10	11	13	14	16	30

Progression APb (From -1795 to -1198)

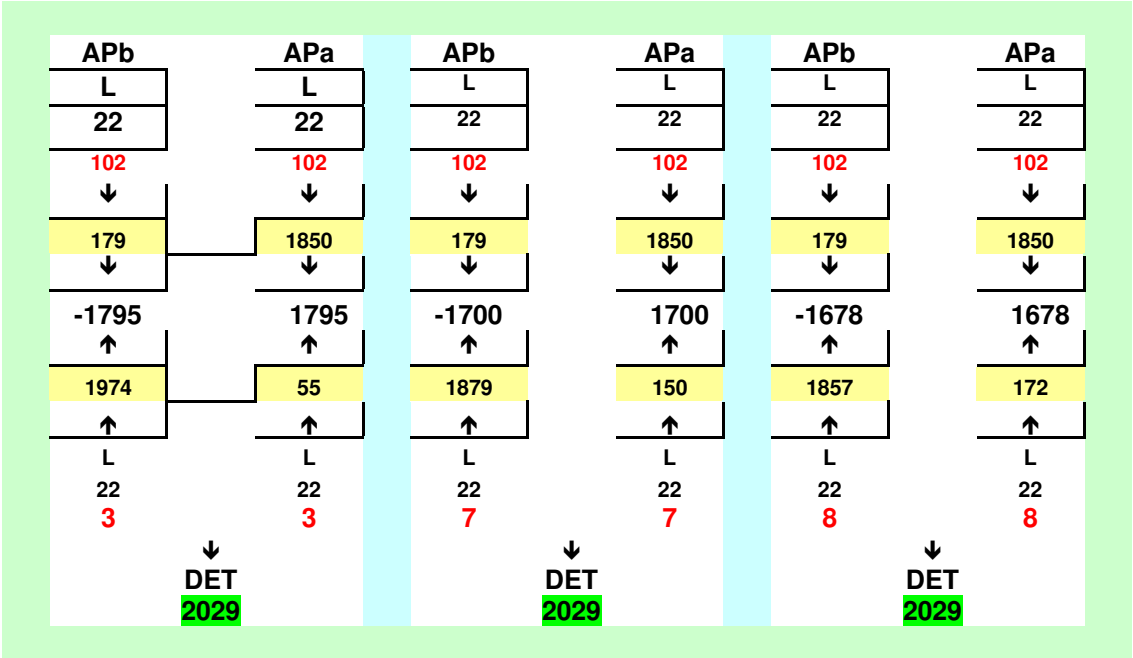
	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	102	102	102	102	102	102	102	102	102	102
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
APb	179	179	179	179	179	179	179	179	179	179
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
	-1795	-1700	-1678	-1639	-1617	-1582	-1560	-1525	-1278	-1198
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
APb	1974	1879	1857	1818	1796	1761	1739	1704	1457	1377
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	3	7	8	10	11	13	14	16	30	35

Progression APb (From -1125 to 0)

	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22
	102	102	102	102	102	102	102	102	102
	↓	↓	↓	↓	↓	↓	↓	↓	↓
APb	179	179	179	179	179	179	179	179	179
	↓	↓	↓	↓	↓	↓	↓	↓	↓
	-1125	-1079	-695	-566	-450	-391	-356	-291	0
	↑	↑	↑	↑	↑	↑	↑	↑	↑
APb	1304	1258	874	745	629	570	535	470	179
	↑	↑	↑	↑	↑	↑	↑	↑	↑
	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22
	3	7	8	10	11	13	14	16	30

Atomic progressions in correlation with each other result in the progression difference, and the progression difference result in the determinante 2029. We could say that the determinante 2029 connects all the progressions into the progression matrix Apa and Apb. That code connects the progressions with the number of atoms in insulin. This can be observed at the next example:

Table 3. Atomic progression APa and APb (Amino acid Leu – *Progression differences*)



Determinants 2x2

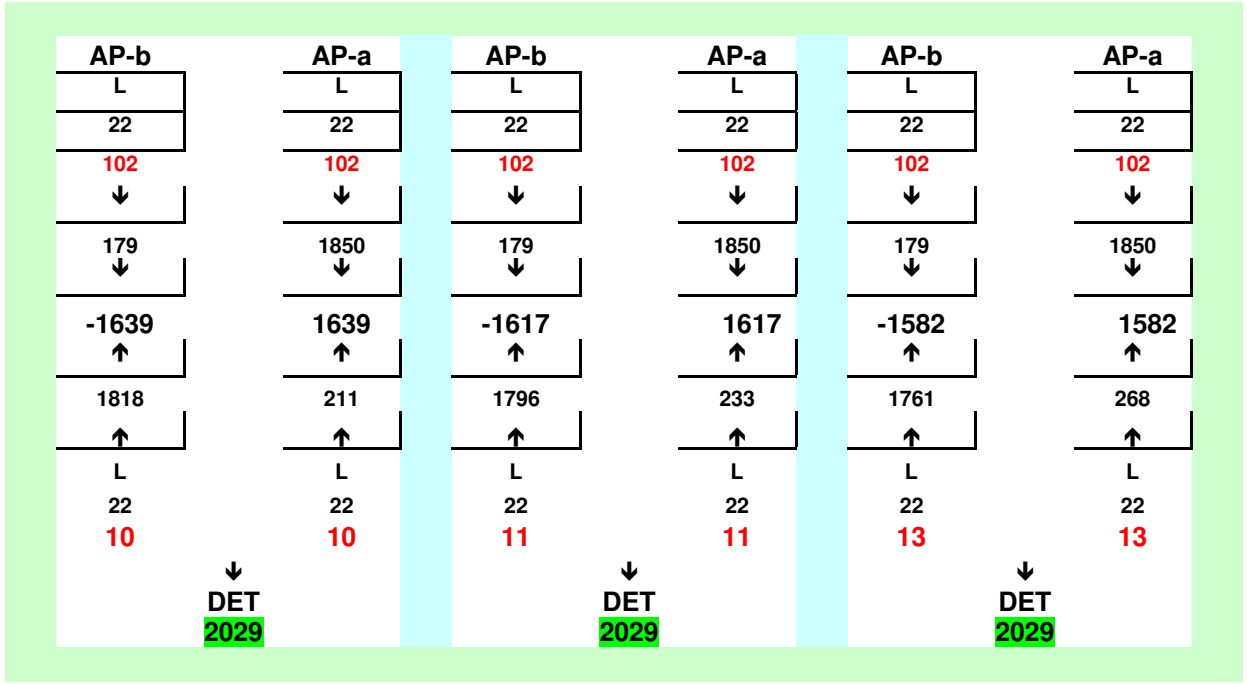
1850	55	>	1850	
179	1974		179	1974

Determinants 2x2

1850	150	>	1850	
179	1879		179	1879

Determinants 2x2

1850	172	>	1850	
179	1857		179	1857



Determinants 2x2

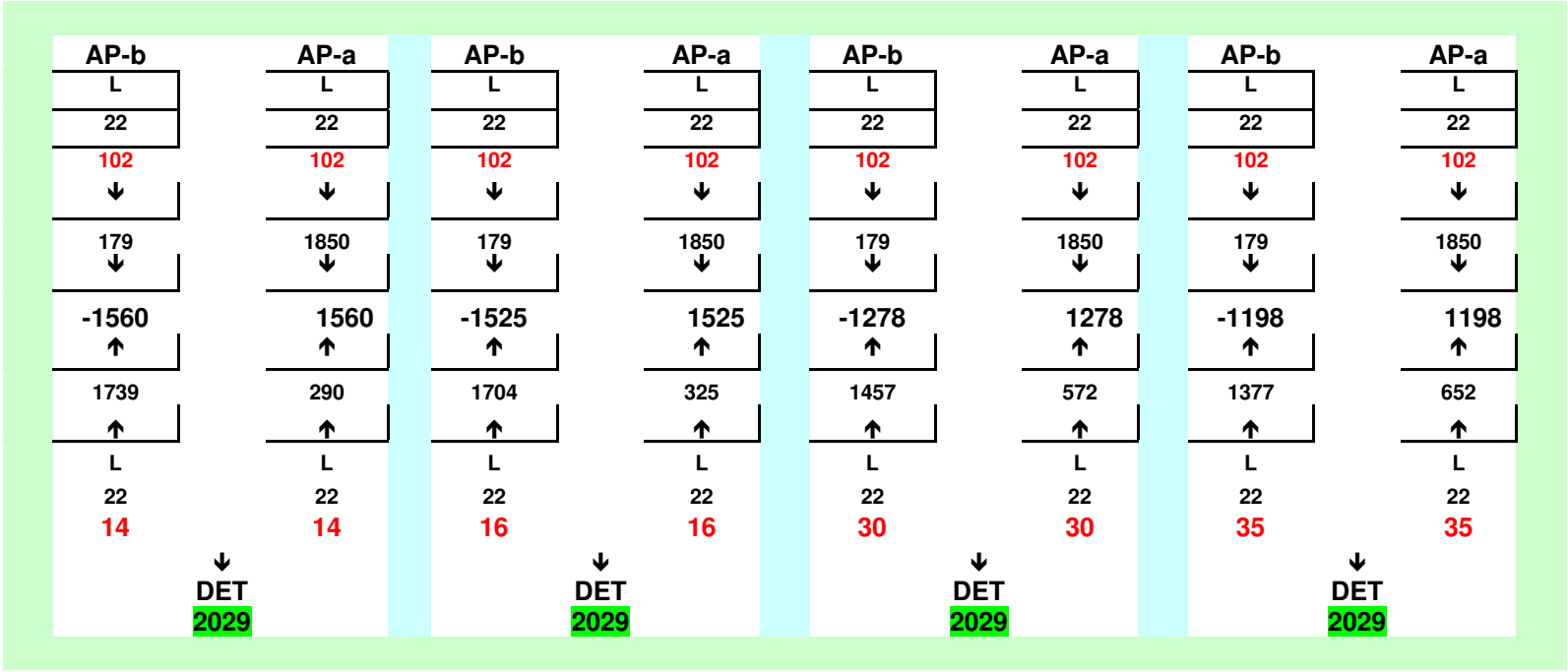
1850 211 > (1639 x 2029)
179 1818

Determinants 2x2

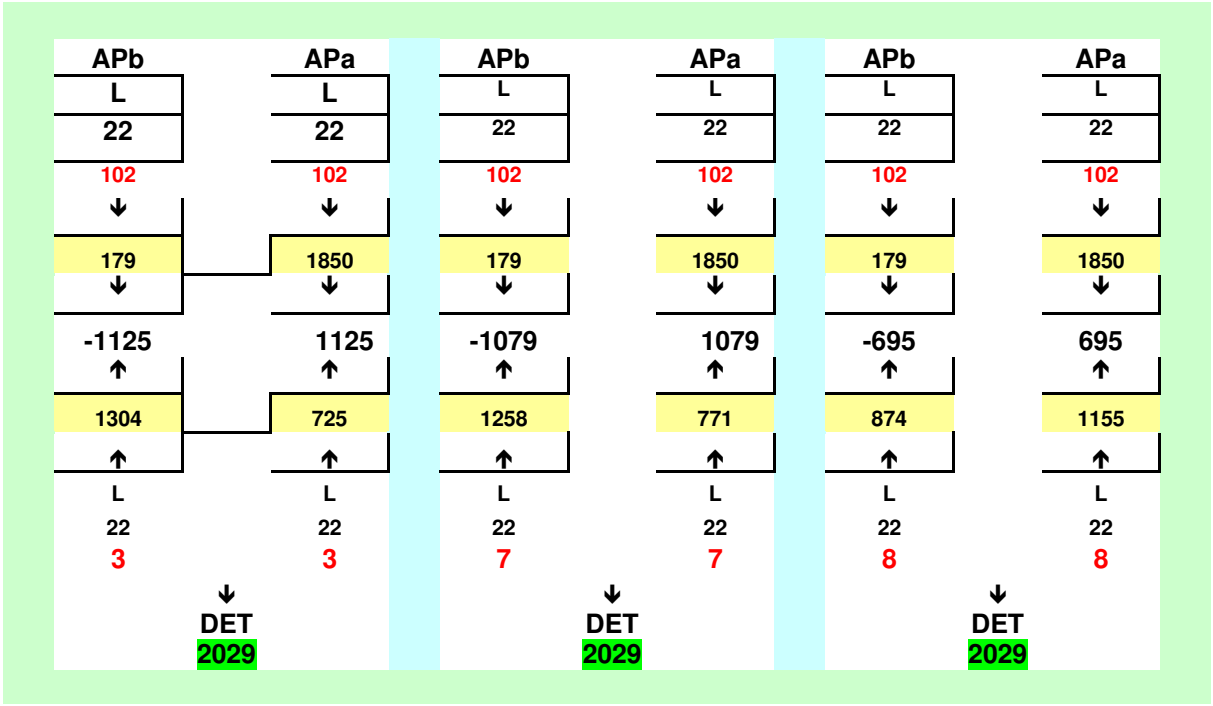
1850 233 > (1617 x 2029)
179 1796

Determinants 2x2

1850 268 > (1582 x 2029)
179 1761


AP-b
AP-a
AP-b
AP-a
AP-b
AP-a
AP-b
AP-a
L
22
102
179
-1560
1739
1560
290
-1525
1704
1525
325
-1278
1457
1278
572
-1198
1377
1198
652
14
16
30
35
DET
2029

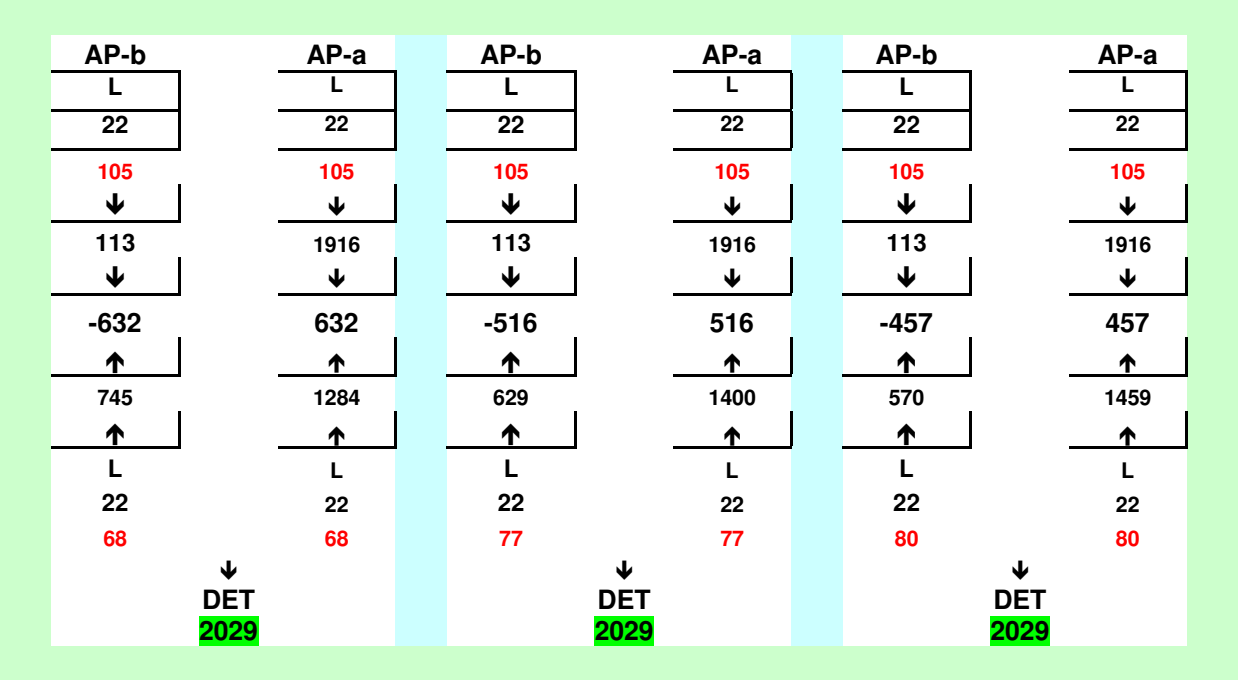
Determinants 2x2
1850 290 > (1560 x 2029)
179 1739
Determinants 2x2
1850 325 > (1525 x 2029)
179 1704
Determinants 2x2
1850 572 > (1278 x 2029)
179 1457
Determinants 2x2
1850 652 > (1198 x 2029)
179 1377



Determinants 2x2
1850 725 > (1125x 2029)
179 1304

Determinants 2x2
1850 771 > (1079x 2029)
179 1258

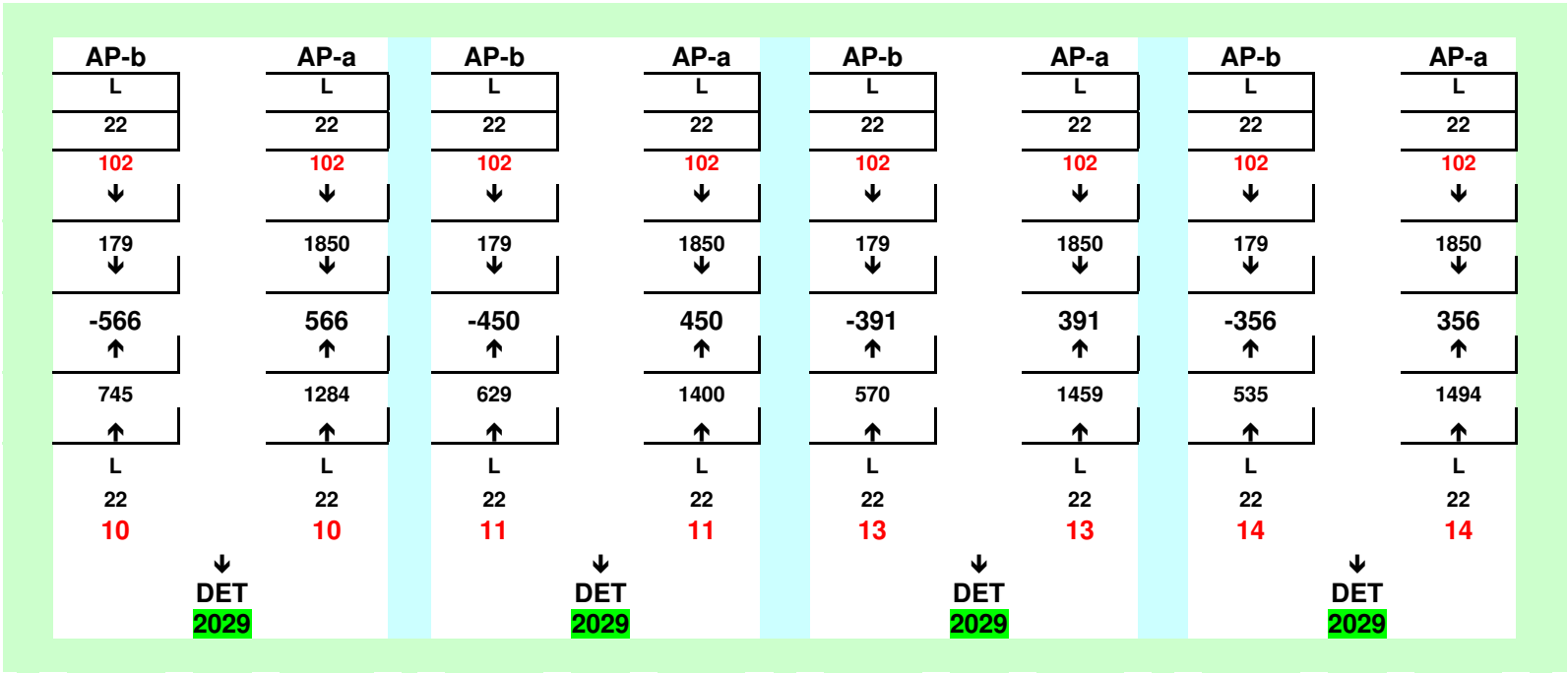
Determinants 2x2
1850 1155 > (695x 2029)
179 874



Determinants 2x2
1916 1284 > (632 x 2029)
113 745

Determinants 2x2
1916 1400 > (516 x 2029)
113 629

Determinants 2x2
1916 1459 > (457 x 2029)
113 570



Determinants 2x2
1850 1284 > (566 x 2029)
179 745
Determinants 2x2
1850 1400 > (450 x 2029)
179 629
Determinants 2x2
1850 1459 > (391 x 2029)
179 570
Determinants 2x2
1850 1494 > (356 x 2029)
179 535

Table 4. Determinante 2029 APa and APb (Amino acid Leu – position 86)

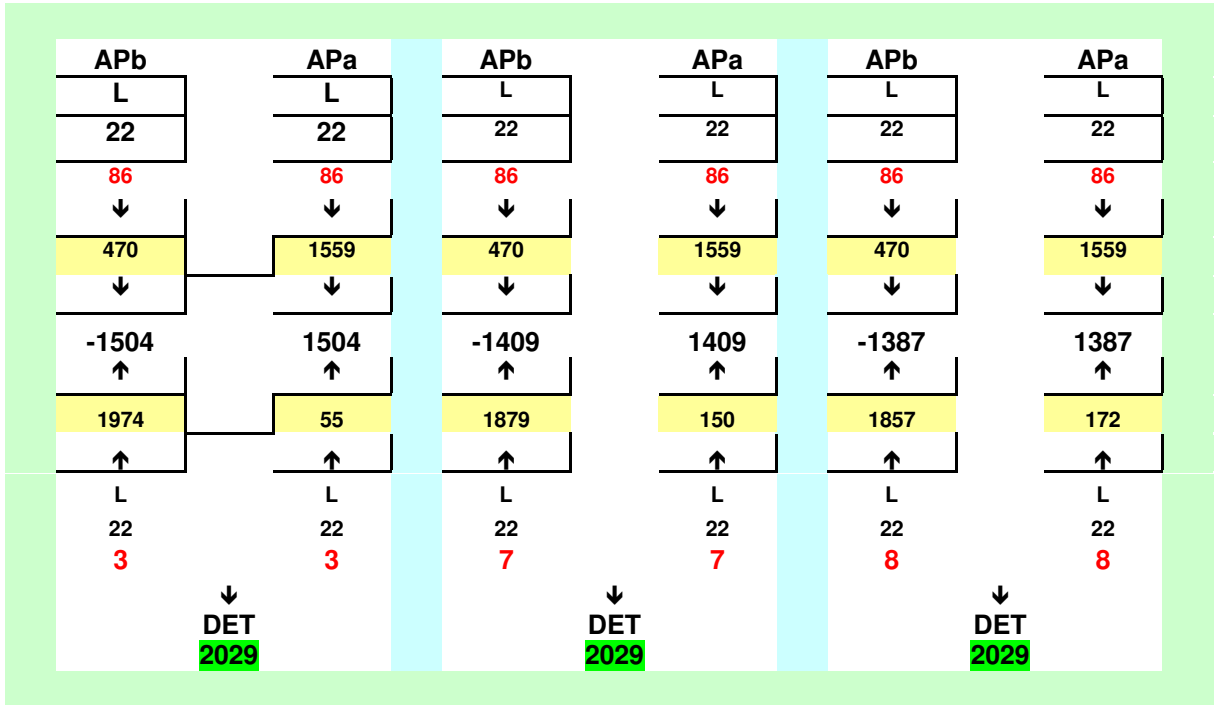
Progression APa (From 1504 to 907)

	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	86	86	86	86	86	86	86	86	86	86
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
APa	1559	1559	1559	1559	1559	1559	1559	1559	1559	1559
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
	1504	1409	1387	1348	1326	1291	1269	1234	987	907
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
APa	55	150	172	211	233	268	290	325	572	652
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	3	7	8	10	11	13	14	16	30	35

Progression APb (From 1504 to -907)

	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	86	86	86	86	86	86	86	86	86	86
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
APb	470	470	470	470	470	470	470	470	470	470
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
	-1504	-1409	-1387	-1348	-1326	-1291	-1269	-1234	-987	-907
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
APb	1974	1879	1857	1818	1796	1761	1739	1704	1457	1377
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	3	7	8	10	11	13	14	16	30	35

Table 5. Atomic progression APa and APb (Amino acid Leu – *Progression differences*)



Determinants 2x2

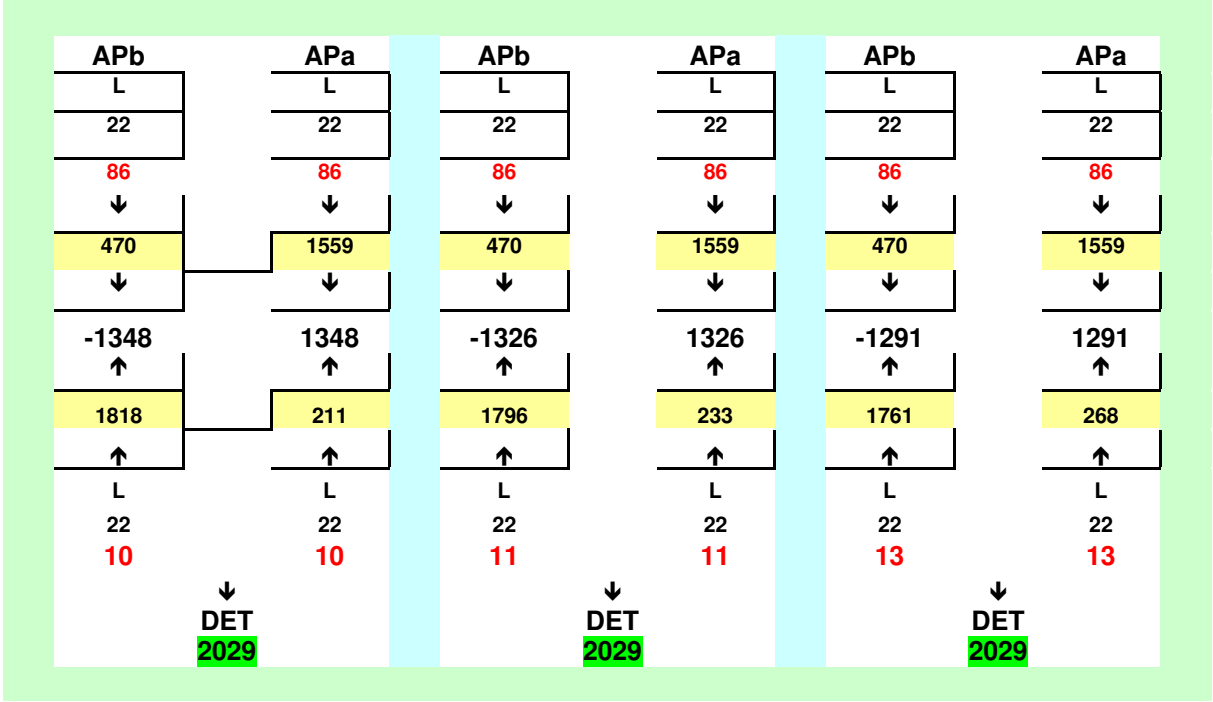
1559	55	>	(1504 x 2029)
470	1974		

Determinants 2x2

1559	150	>	(1409 x 2029)
470	1879		

Determinants 2x2

1559	172	>	(1387 x 2029)
470	1857		



Determinants 2x2

1559	211	>	(1348 x 2029)
470	1818		

Determinants 2x2

1559	233	>	(1326 x 2029)
470	1796		

Determinants 2x2

1559	268	>	(1291 x 2029)
470	1761		

Table 6. Determinante 2029 APa and APb (Amino acid Leu – position 82)

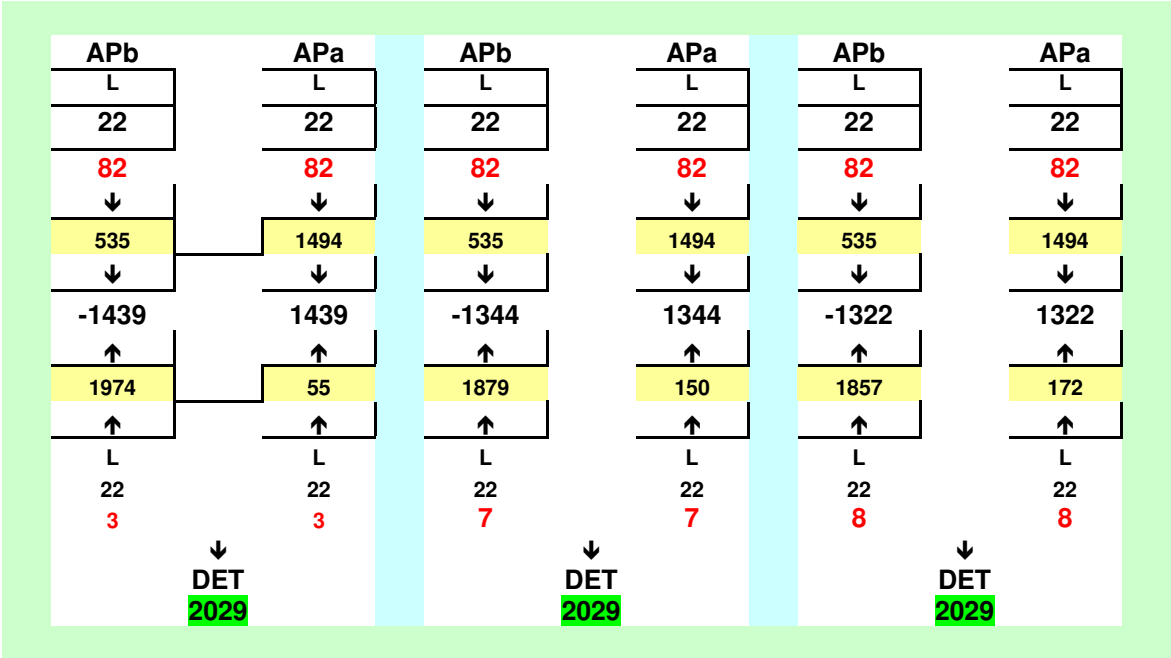
Progression APa (From 1439 to 842)

	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	82	82	82	82	82	82	82	82	82	82
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
APa	1494	1494	1494	1494	1494	1494	1494	1494	1494	1494
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
	1439	1344	1322	1283	1261	1226	1204	1169	922	842
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
APa	55	150	172	211	233	268	290	325	572	652
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	3	7	8	10	11	13	14	16	30	35

Progression APa (From -1439 to -842)

	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	82	82	82	82	82	82	82	82	82	82
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
AP-b	535	535	535	535	535	535	535	535	535	535
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
	-1439	-1344	-1322	-1283	-1261	-1226	-1204	-1169	-922	-842
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
AP-b	1974	1879	1857	1818	1796	1761	1739	1704	1457	1377
	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	L	L	L	L	L	L	L	L	L	L
	22	22	22	22	22	22	22	22	22	22
	3	7	8	10	11	13	14	16	30	35

Table 7. Atomic progression APa and APb (Amino acid Leu – *Progression differences*)



Determinants 2x2

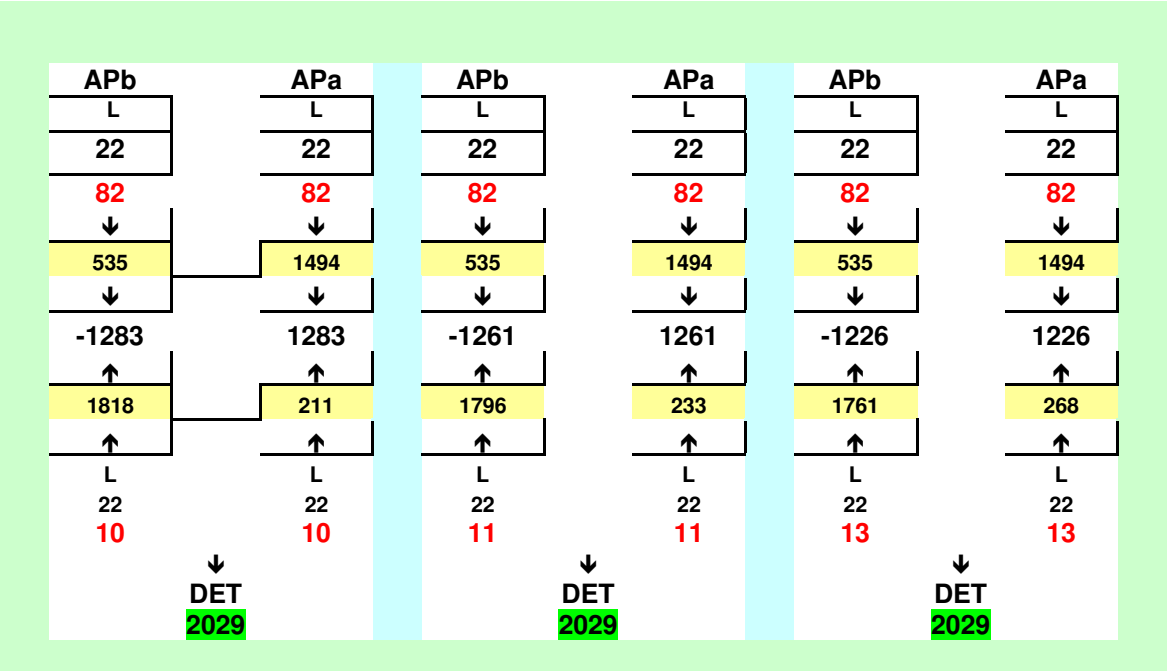
1494	55	>	(1439 x 2029)
535	1974		

Determinants 2x2

1494	150	>	(1344 x 2029)
535	1879		

Determinants 2x2

1494	172	>	(1322 x 2029)
535	1857		



Determinants 2x2

1494 211 > (1283 x 2029)
535 1818

Determinants 2x2

1494 233 > (1261 x 2029)
535 1796

Determinants 2x2

1494 268 > (1226 x 2029)
535 1761

In these examples the determinant 2029 interlinks all the APa and APb progression.

As we can see, when system and information procedures and methods are applied in research of biochemistry of insulin, using the progression matrixes, it is possible to analyze effects of classical and information parameters in the evolution process of that protein.

Therefore, different progression values developed into harmonized progression differences and harmonized sums of these progressions. It is our opinion the said harmonized progression difference proves that there really is the genetic language that can be described through use of theory of systems and cybernetics, and that functions in accordance with specific rules.

As we can see, there really is a direct mathematical connection between the progression matrix and number of atoms in insulin.

Biological specificum of proteins depends on the order of amino acids in their molecules. If this order is changed, their biological specificum will change as well. The basic parameters that determine the change in the status of system of bio-synthesis matrix of macro molecules are the system entropy, volume of transferred information and degree of probability that the genetic information will be transferred in full. System means any group of objects (elements), their inter-relations (number of atoms, atomic number, atomic weight, co-valent radius, molar rotation, electrical negativity, etc.), as well as the relations among its attributes. In this example we have analyzed only the number of atoms of the said amino acids. This number of atoms has been analyzed from one perspective only. A similar analysis can be done using some other parameters and observing these parameters from other perspectives. This creates the need for research of other cyber, information and system characteristics of biochemical basis of genetic processes of insulin. The results of that research will hopefully be published in the time to come.

The result of the research that we have carried out clearly shows that there is a matrix code in insulin. It also shows that the coding system within the amino acidic language gives a full information, not only for the amino acid „record“, but also for its structure, configuration and its various shapes. In the following text we shall discuss the issue of the existence of the insulin code, and also the issue of coding of individual structural levels in this protein.

In the previous examples we translated the physical and chemical parameters from the language of biochemistry into the digital language of programmatic, cybernetic and information principles. This we did by using the adequate mathematical algorithms. By using chemical-information procedures, we calculated the numerical value for the information content of molecules. What we got this way is the digital picture of the phenomenon of biochemistry. These digital pictures reveal to us a whole new dimension of this science. They reveal to us that the biochemical process is strictly conditioned and determined by programmatic, cybernetic and information principles.

From the previous examples we can see that this protein really has its quantitative characteristics. It can be concluded that there is a connection between quantitative characteristics in the process of transfer of genetic information and the qualitative appearance of given genetic processes.

4. DISCUSSION

The results of our research show that the processes of sequencing the molecules are conditioned and arranged not only with chemical and biochemical lawfulness, but also with

program, cybernetic and informational lawfulness too. At the first stage of our research we replaced nucleotides from the Amino Acid Code Matrix with numbers of the atoms and atomic numbers in those nucleotides. Translation of the biochemical language of these amino acids into a digital language may be very useful for developing new methods of predicting protein sub-cellular localization, membrane protein type, protein structure secondary prediction or any other protein attributes.

The success of human genome project has generated deluge of sequence information. The explosion of biological data has challenged scientists to accelerate the speed for their analysis. Nowadays, protein sequences are generally stored in the computer database system in the form of long character strings. It would act like a snail's pace for human beings to read these sequences with the naked eyes (Xiao and Chou, 2007). Also, it is very hard to extract any key features by directly reading these long character strings. However, if they can be converted to some signal process, many important features can be automatically manifested and easily studied by means of the existing tools of information theory (Xiao and Chou, 2007). The novel approach as presented here may help improve this kind of situation.

5. CONCLUSIONS AND PERSPECTIVES

The process of sequencing in bio-macromolecules is conditioned and determined not only through biochemical, but also through cybernetic and information principles. The digital pictures of biochemistry provide us with cybernetic and information interpretation of the scientific facts. Now we have the exact scientific proofs that there is a genetic language that can be described by the theory of systems and cybernetics, and which functions in accordance with certain principles.

6. DISCLOSURE

The author reports no conflict of interest in this research.

REFERENCES

Cai, Y.D. and Chou, K.C. (2006). Predicting membrane protein type by functional Domain composition and pseudo amino acid composition. J Theor Biol., 238, 395-400.

Cai, Y.D., Zhou, G.P. and Chou, K.C. (2003). Support vector machines for predicting membrane protein types by using functional domain composition. Biophys J., 84, 3257-3263.

Cai, Y.D., Zhou, G.P., Chou, K.C. (2005). Predicting enzyme family classes by hybridizing gene product composition and pseudo-amino acid composition. J Theor Biol., 234, 145-149.

Cai, Y.D., Feng, K.Y., Lu, W.C., Chou, K.C. (2006). Using LogitBoost classifier to predict protein structural classes. J Theor Biol., 238, 172-176.

Cai, Y.D., Pong-Wong, R., Feng, K., Jen, J.C.H., Chou, K.C. (2004). Application of SVM to predict membrane protein types. J Theor Biol., 226, 373-376.

Chen, C., Chen, L., Zou, X., Cai, P. (2009). Prediction of protein secondary structure content by using the concept of Chou's pseudo amino acid composition and support vector machine. Protein & Peptide Letters, 16, 27-31.

Chen, L., Feng, K.Y., Cai, Y.D., Chou, K.C., Li, H.P. (2010). Predicting the network of substrate-enzyme-product triads by combining compound similarity and functional domain composition. BMC Bioinformatics, 11, 293.

Chou, K.C. (2011). Some remarks on protein attribute prediction and pseudo amino acid composition (50th Anniversary Year Review). J Theor Biol., 273, 236-247.

Chou, K.C. and Shen, H.B. (2010a). Cell-PLoc 2.0: An improved package of web-servers for predicting subcellular localization of proteins in various organisms. Natural Sci., 2, 1090-1103 (openly accessible at http://www.scirp.org/journal/NS/).

Chou, K.C., Shen, H.B. (2010b). Plant-mPLoc: A Top-Down Strategy to Augment the Power for Predicting Plant Protein Subcellular Localization. PLoS ONE., 5, e11335.

Ding, H., Luo, L., Lin, H. (2009a). Prediction of cell wall lytic enzymes using Chou's amphiphilic pseudo amino acid composition. Protein & Peptide Letters., 16, 351-355.

Ding, Y.S., Zhang, T.L., Chou, K.C. (2007). Prediction of protein structure classes with pseudo amino acid composition and fuzzy support vector machine network. Protein & Peptide Letters, 14, 811-815.

Ding, Y.S., Zhang, T.L., Gu, Q., Zhao, P.Y., Chou, K.C. (2009b). Using maximum entropy model to predict protein secondary structure with single sequence. Protein & Peptide Letters, 16, 552-560.

He, Z.S., Zhang, J., Shi, X.H., Hu, L.L., Kong, X.G., Cai, Y.D., Chou, K.C. (2010). Predicting drug-target interaction networks based on functional groups and biological features. PLoS ONE., 5, e9603.

Hu, L., Huang, T., Shi, X., Lu, W.C., Cai, Y.D., Chou, K.C. (2011). Predicting functions of proteins in mouse based on weighted protein-protein interaction network and protein hybrid properties PLoS ONE., 6, e14556.

Huang, T., Shi, X.H., Wang, P., He, Z., Feng, K.Y., Hu, L., Kong, X., Li, Y.X., Cai, Y.D. and Chou, K.C. (2010). Analysis and prediction of the metabolic stability of proteins based on their sequential features, subcellular locations and interaction networks PLoS ONE., 5, e10972.

Kandaswamy, K.K., Pugalenthi, G., Moller, S., Hartmann, E., Kalies, K.U., Suganthan, P.N., Martinetz, T. (2010). Prediction of Apoptosis Protein Locations with Genetic Algorithms and Support Vector Machines through a New Mode of Pseudo Amino Acid Composition. Protein and Peptide Letters, 17, 1473-1479.

Kannan, S., Hauth, A.M., Burger, G. (2008). Function prediction of hypothetical proteins without sequence similarity to proteins of known function. Protein & Peptide Letters, 15, 1107-1116.

Kurić, L. (2007). The digital language of amino acids. Amino Acids, 653-661.

Kurić, L. (2009). The Atomic Genetic Code. J. Comput. Sci. Biol., 2, 101-116.

Kurić, L. (1986). Mesure complexe des caracteristiques dynamiques de series temporelles. Journal de la Societe de statistique de Paris- tome., 127, No 2.

Kurić, L. (2010). The Insulin Bio Code - Zero Frenquencies. GJMR, 10(1), 1.

Kurić, L. (2010). Molecular biocoding of insulin, Advances and Applications in Bioinformatics and Chemistry, 45 – 58.

Kurić, L. (2010). The Insulin Bio Code – Prima sequences, GJMR, 1(1), 1.

Kurić, L. (2010). ATOMIC HEMOGLOBIN CODE**,** GJMR GJMR, 10(2).

Kurić, L. (2010). Language of Insulin Decoded: Discrete code 1128, IJPBS Journal.

Kurić, L. (2010). Measures of Bio Insulin Frequencies. IJCSET, 1(4).

Kurić, L. (2010). The Insulin Bio Code - Standard Deviation, International Journal of Scientific and Engineering Research (IJSER) ISSN 2229-5518.

Kurić, L. (2011). Molecular biocoding of insulin – amino acid Gly. International Journal of Scientific and Engineering Research (IJSER).

Kurić, L. (2011). Algorithm and computational complexity of Insulin. International Journal of Computer Technology and Application, 2(2), 216-241.
Kurić, L. (2011). Algorithm and computational complexity of Insulin – amino acid Asn, International Journal of Computer Technology and Application IJCTA, 2(2), 258-273.
Kurić, L. (2011). Algorithm of Insulin – Amino acid Gln. IJCSET.
Kurić, L. (2011). Algorithm of Insulin – Amino acid Thr. International Research Journal of Biochemistry and Bioinform. IRJBB, April 2011.
Kurić, L. (2011). Algorithm of Chemistry – discret codes 19 and 7. Journal of Computer Science and Information Technologies, IJCSIT, April 2011.
Kurić, L. (2011). Algorithm of Insulin – Amino acid Thr, International Research Journal of Biochemistry and Bioinformatics (IRJBB), April issue
Kurić, L. (2011). Algorithm and computational complexity of biochemistry, Journal of Chemical Engineering and Material Science, JCEMS, 2 (5).
Kurić, L. (2011). Algorithm of Insulin Human P01308–discrete code 2029. International Journal of Scientific and Engineering Research (IJSER), 2.
Kurić, L. (2011). Language of Insulin Decoded: Discret code 1128. International Journal of Pharma and Bio Sciences, 2(2), 258-273, ISSN: 2229-6093.
Kurić, L. (2011). Molecular biocoding of insulin - amino acid Ser. Journal of Bioengineering & Biomedical Science, July 26, 2011.
Liu, T., Zheng, X., Wang, C., Wang, J. (2010). Prediction of Subcellular Location of Apoptosis Proteins using Pseudo Amino Acid Composition: An Approach from Auto Covariance Transformation. Protein & Peptide Letters., 17, 1263-9.
Mohabatkar, H. (2010). Prediction of cyclin proteins using Chou's pseudo amino acid composition. Protein & Peptide Letters, 17, 1207-1214.
Wang, Y.C., Wang, X.B., Yang, Z.X., Deng, N.Y. (2010). Prediction of enzyme subfamily class via pseudo amino acid composition by incorporating the conjoint triad feature. Protein & Peptide Letters, 17, 1441-1449.
Xiao, X., Chou, K.C. (2007). Digital coding of amino acids based on hydrophobic index. Protein Peptide Lett., 14, 871-875.