

Discrete dynamic system oriented on the formation of prebiotic dipeptides from Rode's experiment

Carlos Polanco¹✉, José Lino Samaniego², Thomas Buhse³ and Jorge Alberto Castañón González²

¹Departamento de Matemáticas, Facultad de Ciencias, Universidad Nacional Autónoma de México, México; ²Facultad de Ciencias de la Salud, Universidad Anáhuac, Col. Lomas Anáhuac Huixquilucan Estado de México, México; ³Centro de Investigaciones Químicas, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, México

This work attempts to rationalize the possible prebiotic profile of the first dipeptides of about 4 billion years ago based on a computational discrete dynamic system that uses the final yields of the dipeptides obtained in Rode's experiments of salt-induced peptide formation (Rode *et al.*, 1999, *Peptides* 20: 773–786). The system built a prebiotic scenario that allowed us to observe that (i) the primordial peptide generation was strongly affected by the abundances of the amino acid monomers, (ii) small variations in the concentration of the monomers have almost no effect on the final distribution pattern of the dipeptides and (iii) the most plausible chemical reaction of prebiotic peptide bond formation can be linked to Rode's hypothesis of a salt-induced scenario. The results of our computational simulations were related to former simulations of the Miller, and Fox & Harada experiments on amino acid monomer and oligomer generation, respectively, offering additional information to our approach.

Key words: origins of life, biogenesis, dipeptides, salt-induced peptide formation

Received: 08 November, 2013; **revised:** 12 July, 2014; **accepted:** 17 October, 2014; **available on-line:** 16 December, 2014

INTRODUCTION

If we consider the geochemical conditions that supposedly prevailed on earth 4 billion years ago (Vogel, 1998), it seems that peptides had a greater chance to be formed than any other bio-molecules. One plausible chemical scenario for their generation is the salt-induced peptide formation (SIPF) proposed by Rode and co-workers (Schwendinger & Rode, 1989, Schwendinger & Rode, 1992, Plankensteiner *et al.*, 2005; Reiner *et al.*, 2006; Fraser *et al.*, 2011) involving high concentrations of NaCl subjected to wetting/drying cycles (Saetia *et al.*, 1993) and acting as a condensation reagent for the peptide bond formation. The SIPF hypothesis is supported by the estimated oxygen content of the secondary primitive earth atmosphere, which allowed the oxidation of Cu(I) to Cu(II) (Cloud, 1973; Ochiai, 1978) that is considered as a fundamental condition for the characterization of the amino acid side chain electronegativities (Schwendinger & Rode, 1992; Rode, 1999). An acidic pH and temperatures between 80 and 100°C must have prevailed in the cooling process of the earth after the

formation of the first hydrosphere as well as regular drying/wetting and day/night cycles, heavy rainfalls, tidal fluctuations and various atmospheric processes.

Under such scenario, laboratory experiments indicated formation of peptides from binary mixtures of amino acids. It turned out that some amino acids promote the formation of homo-dipeptides and others of hetero-dipeptides. In this context, Rode carried out a systematic study of those amino acids that played a major role in the formation of dipeptides and observed their generation under SIPF conditions. His pioneering work yielded a detailed quantitative description of 81 dipeptides formed by the combination of 9 amino acids: Gly, Ala, His, Asp, Glu, Lys, Pro, Val, and Leu (Rode, 1999). The obtained concentrations of the 81 dipeptides are called here the Rode profile Table 1 (Table 6; Rode, 1999). Such profile is very useful since it can be considered as a quantifiable precedent of the relative composition profile of the starting amino acid monomers. Like in the case of the amino acid distribution in the Murchison meteorite (Wolman *et al.*, 1972), the Rode profile, i.e. the measure of the final composition of the dipeptides in Rode's experiments, can serve as a valuable information enabling us to build a mathematical-computational model about the assumed prebiotic peptide formation.

The present work focuses on the possible prebiotic peptide profile formed 4 billion years ago by using the information of the Rode profile through computational simulation and by comparing this profile with our former studies (Polanco *et al.*, 2013; Polanco *et al.*, 2013a) on the Miller-type generation of amino acid monomers (Miller, 1953) as well as with the experiments by Fox & Harada (1960) on the generation of the so-called "proteinoids". In particular, we simulated in three computational scenarios the hypothetical peptide building (i) resulting from the Miller experiments on the lightning-induced amino acid generation by using the experimentally observed monomer abundances, (ii) considering the initial conditions of the Fox & Harada experiments as well as (iii) reproducing the Rode profile taking into account the starting mixtures of the Rode experiments. The latter allowed us to perform extrapolations of the future and past states of peptide building under those salt-induced conditions, i.e. the hypothetical building of longer peptides than dimers and the inverse process, respectively.

Our computational model intends to recreate the prebiotic scenario from a discrete dynamics system that

✉ e-mail: polanco@unam.mx

satisfies the Markov conjecture. This computational scheme allows multiple variables to delimit the model affecting neither the complexity nor the required processing time, due to the assumption of the Markov property (Isaacson & Madsen, 1976). A Markov process is a stochastic system in which the occurrence of a future state depends on the immediately previous state and only on that previous state. Thus the set of random variables $\{X_n\}$ in a process is said to have the Markov property if it is verified that $P\{X_n=j_n | X_{n-1}=j_{n-1}, X_{n-2}=j_{n-2}, \dots, X_0=j_0\} = P\{X_n=j_n | X_{n-1}=j_{n-1}\}$. Roughly speaking, the Markov property is satisfied if the future location of the object in study depends on its present state and not on its past state.

From this Markov process three relevant results can be identified: (1) The Rode profile enabled us to build up a past-future profile of the prebiotic composition very accurately and with a minimal number of amino acids. (2) The profile of the final composition from the Miller experiment on amino acid monomer formation and those of Fox & Harada, and Rode on the amino acid oligomerization converges into a single profile despite significantly different numbers and proportions of the involved amino acids as well as the circumstance that the Rode approach results in peptide bond formation and the Fox & Harada does not and (3). The polarity bias in the amino acids does not seem to affect the composition of the prebiotic peptides constructed this way. The comparison of the three experimental approaches was performed by constructing a polarity matrix for each one of them. The polarity matrix plays a fundamental role in the polarity index method that we have been using as a versatile fingerprint to identify the main pathogenic role of antimicrobial peptides (Polanco *et al.*, 2012).

The dipeptide formation was considered in the spirit of our former toy model simulations (Polanco *et al.*, 2013), i.e. without taking into account the thermodynamic details of a particular chemical process.

MATERIAL AND METHODS

This work is essentially a comparative study of the abiogenetic experiments by Miller, Fox & Harada, and Rode. The first two experimental approaches have already been computationally modeled by us considering the polarity as a bias (Polanco *et al.*, 2013 and 2013a). The computational platform was designed to simulate the evolutionary process of polymerization based on

the abundance and concentration of the amino acids to project the future trend of the dipeptide formation. Subsequently, a polarity matrix was calculated for each set of the obtained dipeptides. This polarity matrix was then linearized and its geometric representation as smooth curves was used to compare the future trend of the dipeptides.

To carry out the computational modeling, the same polar classification was considered for the amino acids in each experiment, i.e. the four polar groups P+, P-, N and NP as well as the same proportion and number of amino acids. In the Miller and Fox & Harada experiments particularly, the number of amino acids considered was like that prevailing in the prebiotic 4 billion years ago. To recreate the Rode experiments, the experimental final composition of the dipeptides (matrix A^1 ; Section Discrete dynamic system) was not used. Instead, we estimated the dipeptide formation starting from a prebiotic scenario based on the amino acid monomer abundances as predicted by the Miller and Fox & Harada experiments. To achieve this, each of the 81 dipeptide proportions were extrapolated first forward (expressed in matrix A^6 ; Section Discrete dynamic system) and then, from the construction of analytic functions, backwards (expressed in the $B^{0.9997}$ matrix; Section Construction of the B^6 matrix). These functions were verified with the dipeptide proportions in the Rode experiments (Section Proximity between two matrices). The $B^{0.9997}$ matrix was verified by measuring its proximity to the Rode matrix A^1 (Section Past-future profile). Then the matrices $B^{0.9997}$ and A^1 were iterated to obtain the B^6 matrix, representing the distant future of the $B^{0.9997}$, A^6 , and A^1 matrix. Afterwards the proximity between the $B^{0.9997}$ and A^1 was verified. In this way we built a broader past-future scenario than defined by the A^1 matrix from Rode's experiments. Finally, with the $B^{0.9997}$ matrix, the abiogenetic laboratory experiments were computationally modeled.

Then with the restrictions of abundance, polarity and number of amino acids, each of the abiogenetics experiments were computationally evolved by enabling and disabling the polarity bias and in all six cases the polarity matrices were calculated. Finally, the polarity matrices, expressed as smooth curves, were compared with and without the polarity bias. In both comparisons the consolidated set of genes from Delaye *et al.* (2005) of three microorganisms was included, representing the closest experimental precedent of the evolutionary trend.

Table 1. A^1 matrix profile.

$\downarrow(i,j)\rightarrow$	Asp	Glu	Gly	Pro	Lys	His	Ala	Leu	Val
Asp	0.3800	0.1700	0.7300	0.0001	0.4200	0.0001	0.1900	0.1000	0.1500
Glu	0.2300	0.7900	0.3000	0.0001	0.0100	0.0001	0.0100	0.0100	0.0000
Gly	0.2400	0.5100	6.5700	0.5400	0.9400	0.5400	0.9100	2.9400	2.0500
Pro	0.0001	0.0001	0.9600	0.0000	0.0001	0.0001	1.5500	0.8600	0.1200
Lys	1.0600	0.0100	0.3000	0.0001	0.2600	0.0001	0.2800	0.0001	0.0100
His	0.0001	0.0001	0.2500	0.0001	0.0001	0.3300	0.3200	0.1900	1.3300
Ala	0.3700	0.2500	1.2000	0.2500	0.6400	0.8500	1.8600	0.2000	1.1700
Leu	0.1100	0.0000	0.5800	0.0100	0.0001	0.2000	0.2700	0.3000	0.3800
Val	0.0000	0.0100	0.8900	0.0100	0.0100	0.5200	0.2600	0.1600	0.9600

Initial amino acid concentrations allowing dipeptide formation (in mM) (Table 6; Rode, 1999), where $(i,j) = i-j$ linkage yields in the ij amino acids. (na): Linkages not investigated yet, the value is 0.0001 instead of zero, (data supplied by Rode). (nf): Linkages analyzed but not found, the value should be 0.0000 dipeptide. (tr): Linkages found with traces but not measurable, we used 0.0100 (data by Rode).

Table 2. A^6 matrix profile.

$\downarrow(i,j) \rightarrow$	Asp	Glu	Gly	Pro	Lys	His	Ala	Leu	Val
Asp	1498.1847	1683.0925	19797.2656	1561.2090	2998.8906	2772.1150	4597.2891	8448.6709	7919.3940
Glu	616.6678	693.5171	8157.6558	643.2009	1235.0631	1140.2960	1891.3710	3482.2202	3259.9736
Gly	12848.5840	14438.4785	169870.0000	13395.6680	25727.8184	23786.9004	39442.5078	72495.7109	67955.4922
Pro	2501.3716	2804.3193	32959.9727	2600.1177	4998.9639	4631.6011	7679.3623	14058.0732	13211.7168
Lys	871.6351	977.9959	11496.3125	906.7371	1742.7059	1611.7126	2673.3232	4904.9390	4601.7808
His	977.5129	1097.0634	12902.7900	1017.7202	1955.7428	1812.7819	3004.0120	5503.9214	5172.0156
Ala	3557.3606	3988.1301	46873.2695	3697.6851	7109.2676	6587.9189	10922.1113	19991.8398	18791.0098
Leu	1291.2758	1450.2263	17058.6523	1345.3429	2584.5276	2391.4329	3964.8489	7278.8799	6828.7632
Val	1948.6941	2189.3479	25757.8184	2031.3108	3901.7222	3610.0811	5984.6724	10991.4307	10309.9434

Future trend of peptide linkage composition (in mM) once the A^1 matrix was iterated six times (Section 2.1). (**na**): Linkages not investigated yet. (**nf**): Linkages analyzed but not found. (**tr**): Linkages found with traces.

Discrete dynamic system

The typical Rode experiment (Rode, 1999) consisted of the 9 amino acids Asp, Glu, Gly, Pro, Lys, His, Ala, Leu and Val. The computer simulation of the prebiotic scenario considered a discrete dynamic system (Thom, 1975), which can be written as a matrix equation of the form: $A^k = A^1, \dots, A^1$, k times. The A^1 matrix represented the final abundance of the experimentally formed dipeptides. The (i,j) element of the A^k matrix is read as the yield of i - j peptide linkage of the i, j amino acids in time k . The notation we used in this paper to refer to an (i,j) element from the k -th matrix was $A_{(i,j)}^k$.

Construction of the future. The sequence of the Rode system started with the A^1 matrix (Table 1) (Table 6; Rode, 1999), that multiplied by itself $A^1 A^1$ produced the A^2 matrix, i.e. $A^2 = A^1 A^1$. Since the system represented the transformation occurring to the A^1 matrix in time k , then the continuous iterations of the A^1 matrix took us to the future state of the A^1 matrix. This procedure induced a succession of $A^1, A^2, A^3, \dots, A^k, A^{k+1}, \dots$ matrices, in which the left-end element corresponded to the past state of the system and the right-end element to the future state of the system. It is important to note that the discrete dynamic system from a present state intends to build a future state, but a present state could not be used to build a past state. The matrix representing the future of the A^1 matrix was then set to 6 iterations and it was called A^6 (Table 2).

Construction of the past. In order to know the past of the A^1 matrix (Table 1) (Rode, 1999) it was necessary to know the information of the future trend of it, that is A^1, A^2, \dots, A^6 . As each of the 81 elements of these matrices represented the final measure of a dipeptide, then to obtain a B matrix that represented the past of the A^1 matrix we designed 81 sixth degree polynomials, to act as a predictor function, to be used later to extrapolate the values in time to represent the past of A^1 matrix. Here, there is an example to clarify this procedure. If we look for the past of the composition Asp-Asp = $(1,1)$ (element located in line 1 column 1 of A matrix), we take all values corresponding to element $(1,1)$ of the A^1, A^2, \dots, A^6 matrices i.e. $A_{(1,1)}^1 = 0.3800$ (Table 1), $A_{(1,1)}^2 = 0.8852$ (data not shown in Tables) successively until $A_{(1,1)}^6 = 1498.1847$ (Table 2). This induces the succession of points: $(x_k, y_{(i,j)_k}) = (1, 0.3800), (2, 0.8852), \dots, (6, 1498.1847)$. With points $(x_k, y_{(i,j)_k})$ we build the polynomial $P(x) = 0.80772x^6 - 10.74002x^5 + 52.65633x^4 - 108.34695x^3 + 58.31035x^2 + 76.21210x - 68.51954$, using the least-squares method. Finally we evaluate in this

polynomial, all values less than $(1, 0.3800)$ to extrapolate the succession to the past. Following this example let us take the value $(1, 0.9999)$, the polynomial evaluated at this point is $P(0.9999) = 0.3770$, then $B^{0.9999}$ matrix in its element $(1,1)$ has the value 0.3770, i.e. $B_{(1,1)}^{0.9999} = 0.3770$. With this procedure, points $\{0.9999, 0.9998 \text{ and } 0.9997\}$ (Table 3) are evaluated generating the $B^{0.9999}, B^{0.9998}, B^{0.9997}$ matrices that represent the remote past of A^1 , with the $B^{0.9997}$ matrix particularly representing the most remote past of A^1 matrix.

Construction of the B^6 matrix

The B^6 matrix (Table 4) was built by multiplying the $B^{0.9997}$ matrix for 6 iterations, i.e. $B^6 = B^{0.9997} B^{0.9997} \dots B^{0.9997}$. Just as the A^6 matrix represented the future of the A^1 matrix (Table 2), the B^6 matrix represented the future of the $B^{0.9997}$ matrix. Note that the $B^{0.9997}$ matrix was built by polynomial extrapolation (Section Discrete dynamic system) and not as the result of experimental inspection.

Proximity between two matrices

The distance between two matrices for each of the 81 elements was calculated through the metric $|A(i,j) - B(i,j)| / |A(i,j)|$, where $(|x - y| / |x|)$ was the absolute value between elements x and y respect to x ; $(i,j) = (1,1), \dots, (9,9)$.

Proximity of the A^6 and B^6 matrices

The A^6 and B^6 matrices represented the most distant trend to the future of the peptide linkage composition. The first one corresponded to the experimental data and the second one was the result of the discrete dynamics system. The verification of these matrices was regarding the proximity between their respective elements (Table 5).

Proximity of the A^1 and $B^{0.9997}$ matrices

The A^1 and $B^{0.9997}$ matrices represented the trend of the most distant past of the peptide linkage composition, the first one corresponded to Rode's experiment and the second one was the result of polynomial extrapolation (Table 6).

The Rode approach

The $B^{0.9997}$ matrix represented, by polynomial extrapolation, the remote past of the A^1 matrix (Section Discrete dynamic system), and for us it was a measure

of the abundance of the 81 different interactions from Rode's experiment forming the dipeptides taking 9 amino acids from that remote past. With a computer program already used before to recreate prebiotic scenarios (Polanco *et al.*, 2013), we generated a set of 3000 short peptides. The model used two factors: abundance and polarity. As a bias for the abundance we used the inverse relative abundance represented by the $B^{0.9997}$ matrix (Table 7) and for the polarity we used two inverse polarity distributions in which one induced a bias (Table 8-A) and one without bias (Table 8-B).

The Miller approach

The hypothetical peptide generation based on Miller's experimental results was computed by considering a group of 21 amino acids, where only 11 of them (Gly,

Ala, Val, Leu, Ile, Pro, Asp, Glu, Ser, Thr and Lys) are currently identified as basic amino acids, while the others (numbered from 0 to 9): α -Amino-n-Butyric acid (9), α -Aminoisobutyric (0), Nva, γ -Aminobutyric acid (7), β -Aminoisobutyric acid (6), β -Amino- η -butyric acid (5), β -Alanine (4), N-Methylalanine (3), N-Ethylglycine (2), and Sar, were classified as prebiotic amino acids. The 21 amino acids adopted the inverse abundance distribution shown in Table 9. We used two inverse polarity distributions, one that induced a bias (Table 8-A) and one without bias (Table 8-B). With these restrictions 3000 peptides were generated.

The Fox & Harada approach

The initial conditions of the Fox & Harada experiments used for a hypothetical peptide building scenario

Table 3. Past trend of B matrix profile.

$B^{0.9997}$ matrix profile									
$\downarrow(i,j)\rightarrow$	Asp	Glu	Gly	Pro	Lys	His	Ala	Leu	Val
Asp	0.3711	0.2264	0.1646	-0.0144	1.0552	-0.0056	0.3490	0.1025	-0.0113
Glu	0.1601	0.7860	0.4254	-0.0164	0.0042	-0.0063	0.2266	-0.0085	-0.0028
Gly	0.6139	0.2522	5.5743	0.7668	0.2325	0.1743	0.9253	0.4800	0.7390
Pro	-0.0091	-0.0037	0.4615	-0.0154	-0.0052	-0.0058	0.2283	0.0021	-0.0019
Lys	0.4024	0.0027	0.7893	-0.0295	0.2496	-0.0113	0.5983	-0.0151	-0.0129
His	-0.0161	-0.0065	0.4006	-0.0273	-0.0093	0.3192	0.8114	0.1859	0.4989
Ala	0.1631	-0.0011	0.6788	1.5051	0.2643	0.3023	1.7959	0.2468	0.2249
Leu	0.0504	-0.0104	2.5150	0.7779	-0.0285	0.1579	0.0827	0.2574	0.0955
Val	0.1036	-0.0191	1.6517	0.0425	-0.0169	1.2999	1.0597	0.3399	0.8994
$B^{0.9998}$ matrix profile									
$\downarrow(i,j)\rightarrow$	Asp	Glu	Gly	Pro	Lys	His	Ala	Leu	Val
Asp	0.3740	0.2276	0.1897	-0.0096	1.0568	-0.0037	0.3560	0.1050	-0.0075
Glu	0.1634	0.7873	0.4536	-0.0109	0.0061	-0.0041	0.2344	-0.0056	0.0014
Gly	0.6526	0.2681	5.9063	0.8312	0.2550	0.1995	1.0169	0.5133	0.7893
Pro	-0.0060	-0.0024	0.4876	-0.0103	-0.0034	-0.0038	0.2355	0.0047	0.0020
Lys	0.4083	0.0051	0.8395	-0.0196	0.2530	-0.0075	0.6122	-0.0100	-0.0052
His	-0.0107	-0.0043	0.4470	-0.0182	-0.0061	0.3227	0.8242	0.1906	0.5059
Ala	0.1720	0.0026	0.7558	1.5201	0.2695	0.3082	1.8173	0.2545	0.2366
Leu	0.0669	0.0036	2.6567	0.8053	-0.0190	0.1686	0.1218	0.2716	0.1170
Val	0.1191	-0.0127	1.7845	0.0683	-0.0079	1.3099	1.0965	0.3532	0.9196
$B^{0.9999}$ matrix profile									
$\downarrow(i,j)\rightarrow$	Asp	Glu	Gly	Pro	Lys	His	Ala	Leu	Val
Asp	0.3770	0.2288	0.2149	-0.0047	1.0584	-0.0018	0.3630	0.1075	-0.0038
Glu	0.1667	0.7886	0.4818	-0.0054	0.0081	-0.0020	0.2422	-0.0028	0.0057
Gly	0.6913	0.28406	6.2382	0.8956	0.2775	0.2248	1.1085	0.5467	0.83969
Pro	-0.0029	-0.0015	0.5138	-0.0051	-0.0017	-0.0019	0.2428	0.0073	0.0060
Lys	0.4142	0.0076	0.8898	-0.0097	0.2565	-0.0037	0.6261	-0.0050	0.0023
His	-0.0053	-0.0021	0.4935	-0.0091	-0.0030	0.3264	0.8371	0.1953	0.5130
Ala	0.1810	0.0063	0.8330	1.5350	0.2748	0.3141	1.8386	0.2623	0.2483
Leu	0.0835	0.0032	2.7984	0.8327	-0.0094	0.1793	0.1609	0.2858	0.1385
Val	0.1345	-0.0063	1.9173	0.0942	0.0011	1.3200	1.1333	0.3666	0.9398

The $B^{0.9997}$, $B^{0.9998}$, $B^{0.9999}$ matrices represent the past trend of A^1 matrix, where the superscript 0.9999 represents the oldest trend in the past (in mM). These matrices were calculated by polynomial extrapolation (Section *Composition of the past*), in the corresponding value. (na): Linkages not investigated yet. (nf): Linkages analyzed but not found. (tr): Linkages found with traces but not measurable.

Table 4. B^6 matrix profile.

$\downarrow(i,j) \rightarrow$	Asp	Glu	Gly	Pro	Lys	His	Ala	Leu	Val
Asp	3468.3539	1427.9499	29754.8119	5788.8016	2016.6014	2263.4513	8232.5218	2989.9498	4512.9504
Glu	3896.8746	1605.6718	33436.3624	6493.0943	2263.7709	2540.5702	9233.9082	3358.2850	5070.1504
Gly	45837.3684	18890.3222	393346.7766	76325.0849	26616.2255	29876.5149	108542.2963	39499.9936	59642.7125
Pro	3614.7631	1489.5075	31019.1610	6020.6646	2099.2198	2356.5065	8562.0379	3115.1917	4703.6017
Lys	6943.2927	2860.1526	59577.4586	11573.1421	4033.8101	4528.3688	16458.4902	5984.5106	9034.9666
His	6419.0169	2641.6774	55083.6518	10719.1386	3731.1447	4194.6284	15245.6483	5536.4891	8357.8872
Ala	10644.5081	4381.1817	91336.8505	17771.8123	6187.6481	6952.6196	25275.4589	9179.4633	13856.6804
Leu	19561.4276	8063.2945	167866.5916	32557.2473	11356.4518	12745.3528	46299.0242	16854.8332	25451.0315
Val	8336.9633	7550.8095	157358.1012	30583.9222	10653.7057	11970.3675	43497.0265	15810.2508	23869.7528

Trend to the future of peptide linkage composition (in mM) from $B^{0.9997}$ matrix (Section Composition of the past). (na): Linkages not investigated yet. (nf): Linkages analyzed but not found. (tr): Linkages found with traces but not measurable.

Table 5. Proximity between the A^6 and B^6 matrices.

$\downarrow(i,j) \rightarrow$	Asp	Glu	Gly	Pro	Lys	His	Ala	Leu	Val
Asp	0.5680	0.1787	0.3347	0.7303	0.4871	0.2247	0.4416	1.8257	0.7548
Glu	0.8418	0.5681	0.7560	0.9009	0.4544	0.5512	0.7952	0.0369	0.3570
Gly	0.7197	0.2357	0.5681	0.8245	0.0334	0.2038	0.6366	0.8353	0.1394
Pro	0.3080	0.8827	0.0626	0.5681	1.3813	0.9655	0.1031	3.5127	1.8089
Lys	0.8745	0.6581	0.8070	0.9217	0.5680	0.6441	0.8376	0.1804	0.4907
His	0.8477	0.5847	0.7658	0.9051	0.4758	0.5678	0.8030	0.0059	0.3812
Ala	0.6658	0.0897	0.4868	0.7919	0.1489	0.0525	0.5679	1.1779	0.3561
Leu	0.9340	0.8201	0.8984	0.9587	0.7724	0.8124	0.9144	0.5681	0.7317
Val	0.8937	0.7101	0.8363	0.9336	0.6338	0.6984	0.8624	0.3048	0.5681

Difference in mM between $|A^6_{(i,j)} - B^6_{(i,j)}| / |B^6_{(i,j)}|$, where the A^6 matrix is the future matrix of the A^1 matrix, i.e. $A^6 = A^1 A^1 \dots A^1$ (Section 2.1), and B^6 matrix is the future matrix of $B^{0.9997}$ matrix i.e. $B^6 = B^1 B^1 \dots B^1$ (Section Construction of the B^6 matrix). (na): Linkages not investigated yet. (nf): Linkages analyzed but not found. (tr): Linkages found with traces but not measurable.

Table 6. Proximity between A^1 and $B^{0.9997}$ matrices.

$\downarrow(i,j) \rightarrow$	Asp	Glu	Gly	Pro	Lys	His	Ala	Leu	Val
Asp	0.0240	0.2491	3.4350	1.0069	0.6020	1.0179	0.4556	0.0244	14.2743
Glu	0.4366	0.0051	0.2948	1.0061	1.3810	1.0159	0.9559	2.1765	1.0000
Gly	0.6091	1.0222	0.1786	0.2958	3.0430	2.0981	0.0165	5.1250	1.7740
Pro	1.0110	1.0270	1.0802	1.0000	1.0192	1.0172	5.7893	408.5238	-64.1579
Lys	1.6342	2.7037	0.6199	1.0034	0.0417	1.0088	0.5320	1.0066	1.7752
His	1.0062	1.0154	0.3759	1.0037	1.0108	0.0338	0.6056	0.0221	1.6659
Ala	1.2685	228.2727	0.7678	0.8339	1.4215	1.8118	0.0357	0.1896	4.2023
Leu	1.1825	1.0000	0.7694	0.9871	1.0035	0.2666	2.2648	0.1655	2.9791
Val	1.0000	1.5236	0.4612	0.7647	1.5917	0.6000	0.7546	0.5293	0.0674

Difference in mM between $|A^1_{(i,j)} - B^{0.9997}_{(i,j)}| / |B^{0.9997}_{(i,j)}|$, where A^1 matrix is Rode's matrix (Section Discrete dynamic system), and $B^{0.9997}$ matrix is calculated by polynomial extrapolation (Section Discrete dynamic system). (na): Linkages not investigated yet. (nf): Linkages analyzed but not found. (tr): Linkages found with traces but not measurable.

were simulated by us before (Polanco *et al.*, 2013a). It considered 18 amino acids. The proportions were 10 g Asp and 10g Glu as well as 5 g of the remaining 16 amino acids given in Table 10. We took these proportions and two polarity distributions for the amino acids, one of which induced a bias (Table 8-A), and one did not (Table 8-B). 3000 peptides were generated. In these simulations, Gly was considered in the neutral polar group in order to compare it to the Rode's experiment.

Polarity matrix

The polarity matrix is an array of 16 elements, 4 rows and 4 columns that correspond to the polar groups P+, P-, N, and NP, called for simplicity the M matrix. The M matrix was an essential part of the mathematical-computational polarity index method (Polanco *et al.*, 2012; Polanco *et al.*, 2013; Polanco *et al.*, 2014) and it was used to inform in an exhaustive way the polar profile of the analyzed peptides. In or-

Table 7. Rode matrix of pre-established values by abundance.

↓(i,j)→	Asp	Glu	Gly	Pro	Lys	His	Ala	Leu	Val
Asp	27	44	61	250	9	250	29	98	250
Glu	62	13	24	250	250	250	44	250	250
Gly	16	40	4	13	43	57	11	21	14
Pro	250	250	22	250	250	250	44	250	250
Lys	25	250	13	250	40	250	17	250	250
His	250	250	25	250	250	31	12	54	20
Ala	61	250	15	7	38	33	6	41	44
Leu	198	250	4	13	250	63	121	39	105
Val	97	250	6	235	250	8	9	29	11

Inverse relative abundances in $B^{0.9997}$ matrix (Section Discrete dynamic system). (na): Linkages not investigated yet. (nf): Linkages analyzed but not found. (tr): Linkages found with traces but not measurable.

Table 8. Polarity composition by lateral chain.

A bias				B bias					
	P+	P-	N	NP		P+	P-	N	NP
P+	99	21	85	95	P+	100	100	100	100
P-	21	99	85	95	P-	100	100	100	100
N	60	60	85	95	N	100	100	100	100
NP	60	60	85	95	NP	100	100	100	100

Inverse relative polarities by lateral chain: [P-] polar, [N] neutral, [P+] basic hydrophilic and [NP] non-polar amino acids. A bias: with polar bias, B bias: without polar bias.

der to build this matrix from the set of 3000 peptides taking into account the experiments of Rode, Miller and Fox & Harada with the described hypothetical peptide building extrapolations, we took the 3000 sequences in terms of their amino acids and translated them into the equivalent of their polar groups with the following convention: His, Arg and Lys were translated to the first group; Asp and Glu to the second group; Gly, Ser, Thr, Cys and Tyr to the third group and α -amino-n-butyric acid (9), α -aminoisobutyric (0), Nva, γ -aminobutyric acid (7), β -aminoisobutyric acid (6), β -amino- η -butyric acid (5), β -alanine (4), N-methylalanine (3), N-ethylglycine (2), and Sar, Ala, Leu,

Pro, Val, Trp, Met, Phe and Ile were translated to the fourth group.

In this way, the file of amino acid sequences was re-written in terms of an alphabet of 4 numbers {1, 2, 3, and 4}. After this step the number of polar interactions was counted, reading each sequence from left to right by pairs every time. To illustrate this procedure in the sequence **EEGPKHKDEV** the polar equivalent is **2234111224**. At this stage, the initial polarity matrix is equal to zero, i.e. $M_{(i,j)} = 0$. When we start reading the sequence, from left to right, we find the position (2,2), therefore we add 1 in M matrix, i.e. $M_{(2,2)} = 1$, after counting this first interaction we move one place to the right, to find the interaction (2,3), and we add 1 to this position, i.e. $M_{(2,3)} = 1$, and so forth until we find the interaction (4,1) and add 1 incident i.e. $M_{(4,1)} = 1$. Note that in the following two runs the interaction (1,1) is repeated, therefore interaction (1,1) is 2, i.e. $M_{(1,1)} = 2$, and so on successively until the end, then we continue with the next sequence.

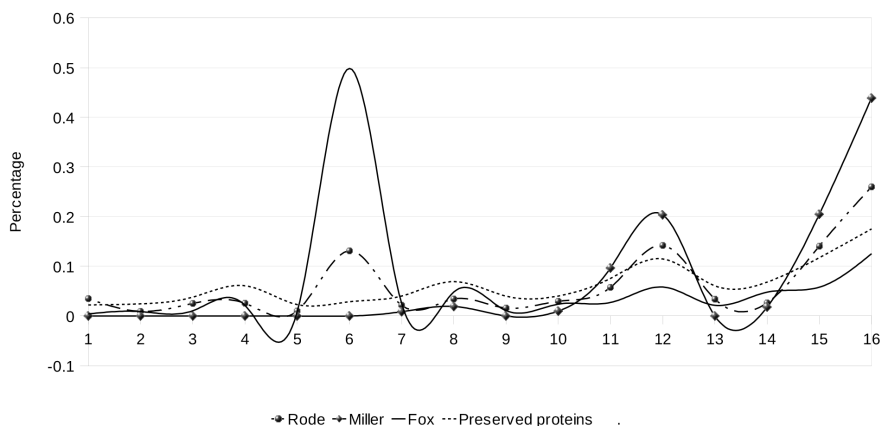


Figure 1. Linear polar interaction between simulated peptides formed in the Rode, Miller, and Fox & Harada approach. The 16 columns on the x-axis correspond to 16 polar interactions from the polarity matrix without polar bias (Table 11).

Polar profile of prebiotic peptides

The M polarity matrix collected all the peptide combinatorial interactions built with the prebiotic computational model. In

Table 9. Miller matrix of pre-established values by abundance

Gly	Ala	Val	Leu	Ile	Pro	Asp	Glu	Ser	Thr	Lys	0	Sar	2	3	4	5	6	7	Nva	9
Gly	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Ala	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Val	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41
Leu	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71
Ile	166	166	166	166	166	166	166	166	166	166	166	166	166	166	166	166	166	166	166	166
Pro	526	526	526	526	526	526	526	526	526	526	526	526	526	526	526	526	526	526	526	526
Asp	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23
Glu	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102	102
Ser	158	158	158	158	158	158	158	158	158	158	158	158	158	158	158	158	158	158	158	158
Thr	987	987	987	987	987	987	987	987	987	987	987	987	987	987	987	987	987	987	987	987
Lys	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160
0	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24
1	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
2	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26
3	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52
4	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42
5	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633
6	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633	2633
7	329	329	329	329	329	329	329	329	329	329	329	329	329	329	329	329	329	329	329	329
8	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
9	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3

Matrix of pre-established values by abundance of Miller's experiment. For the prebiotic amino acids 0-9, we maintained the initially adopted notation (Polanco *et al.*, 2013, Table 1) (Section The Miller approach).

order to interpret the M matrix, it was normalized to 1 and ordered in a *column-vector* of 16 positions (Table 11). In this way the *column-vector* contained the polar relative distribution of the sequences generated by the model. From this *column-vector*, a graph was drawn with smooth curves for the four scenarios described (Figs. 1, 2).

Preserved genes

The same number of *E. coli*, *M. jannaschii* and *S. cerevisiae* used by Delaye and coworkers (Delaye *et al.*, 2005) was used here, extracted from the KEGG data base (Kanehisa *et al.*, 2000) for a previous publication (Polanco *et al.*, 2013).

Past-future profile

The terms "remote past" or "distant future" should be understood as approximations. The past and future profiles result from matrix multiplications and the construction of analytical functions. It is not possible to quantify a time-scale and for that reason the kinetics of dipeptide formation in our simulated scenarios cannot be defined. However, it is possible to affirm that these approximations by analytic functions have enabled us to build a past-future scenario with a time period large enough to be compared with the set of preserved genes (Section Preserved genes).

The exponents or superscripts used in the estimation of the remote past (0.9997, 0.9998, and 0.9999) are not arbitrary. Integer values would have produced extremely high values in the final concentrations. Therefore the selection of the exponents was related to the analytic functions. In the case of the superscripts used for the distant future (1, 2, ..., 6), they were integer numbers, as the multiplication of the resulting matrices did not induce extreme concentration values.

RESULTS

The analysis of similarities between the A^6 matrix, which represents the future state of the dipeptides composition from Rode's experiment and the B^6 matrix, ob-

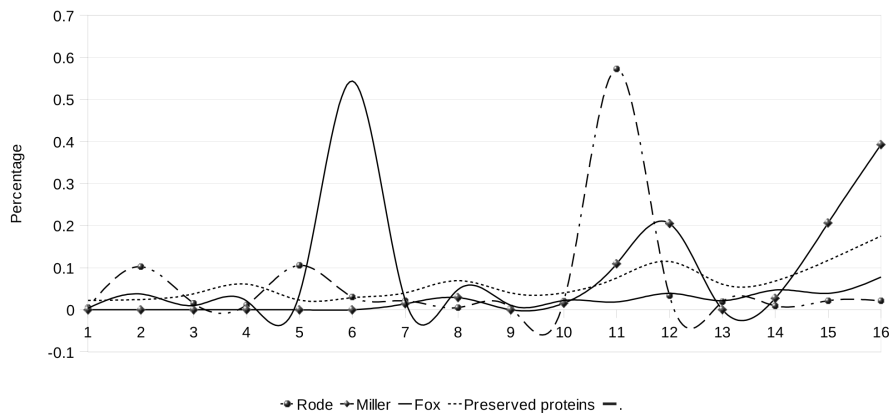


Figure 2. Same as Fig. 1 taking into account the polar bias during the peptide formation

tained by polynomial extrapolation, shows a small difference between the 81 elements (Table 5). The same small difference is observed between the A' matrix, representing the initial dipeptide composition from the Rode's experiment and the $B_{0.9997}$ matrix built with the discrete dynamic system (Table 6).

Interestingly, the bias by polarity did not alter the polar profile of the peptides significantly (Table 9). In all cases the percentage difference (+/-) between the two distributions with and without bias was not significant.

The curves of all three computational scenarios, either with or without the polarity bias (Figs. 1–2), almost preserved the same maximum and minimum points, despite the fact that the amino acid numbers and participation percentages were different.

The Fox & Harada distribution (Fig. 2, column 6) reveals a maximum of Glu and Asp as well as the Rode

distribution (Fig. 2, column 11). Something similar occurs with Gly.

The preserved protein distribution (Section Preserved genes) shows an almost total coincidence when the three scenarios without polar bias were compared (Fig. 1). It does not happen the same way for the scenarios with polar bias (columns 2, and 6; Fig. 2).

DISCUSSION

According to our simulations of short peptide formation, the polarity matrices of the discrete dynamics system based on the Miller, Fox & Harada, and Rode approach, were nearly coincident and converged into the same profile regardless of the bias induced by the polarity, the last profile is also consistent with the set of preserved genes (Polanco *et al.*, 2013). From the mathematical point of view, we consider the starting 9 amino acids used in the Rode experiments as a *basis* (Poole, 2011), i.e. the minimum number of elements in a set capable of generating that set. We do not know if 9 amino acids are in fact the minimum possible to induce the same profile as in the hypothetical peptide formation based on the Miller and Fox & Harada approach. Nevertheless, they represent 40% of those generated in the Miller experiment and 50% of those in the starting conditions of the Fox & Harada experiment. In this regard, the Rode experiment in itself is important, since it can open the discussion about the

Table 10. Fox & Harada matrix of pre-established values by abundance.

	Arg	Cys	Ala	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val	Glu	Asp	
Arg	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Cys	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Ala	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Gly	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
His	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Ile	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Leu	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Lys	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Met	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Phe	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Pro	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Ser	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Thr	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Trp	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Tyr	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Val	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	85	85	
Glu	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	3	3
Asp	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	3	3

Matrix of pre-established values by abundances used in Fox & Harada's hypothetical model (Polanco *et al.*, 2013a) (Section The Fox & Harada approach).

Table 11. Polar profile comparative

#	Polar interaction			Miller approach			Fox & Harada approach			Rode approach		
				With bias	Without bias	+/-	With bias	Without bias	+/-	With bias	Without bias	+/-
1	P+	-	P+	0.0000	0.0000	0.00	0.0038	0.0039	0.00	0.0054	0.0352	0.03
2	P+	-	P-	0.0000	0.0000	0.00	0.0375	0.0086	0.03	0.1024	0.0096	0.09
3	P+	-	N	0.0000	0.0000	0.00	0.0103	0.0104	0.00	0.0154	0.0248	0.01
4	P+	-	NP	0.0001	0.0001	0.00	0.0210	0.0213	0.00	0.0115	0.0257	0.01
5	P-	-	P+	0.0000	0.0000	0.00	0.0367	0.0083	0.03	0.1056	0.0100	0.10
6	P-	-	P-	0.0000	0.0000	0.00	0.5437	0.5018	0.04	0.0305	0.1309	0.10
7	P-	-	N	0.0143	0.0085	0.01	0.0229	0.0253	0.00	0.0214	0.0215	0.00
8	P-	-	NP	0.0289	0.0191	0.01	0.0481	0.0491	0.00	0.0052	0.0342	0.03
9	N	-	P+	0.0000	0.0000	0.00	0.0107	0.0107	0.00	0.0044	0.0160	0.01
10	N	-	P-	0.0158	0.0099	0.01	0.0221	0.0241	0.00	0.0207	0.0301	0.01
11	N	-	N	0.1092	0.0971	0.01	0.0187	0.0264	0.01	0.5723	0.0575	0.51
12	N	-	NP	0.2049	0.2034	0.00	0.0391	0.0578	0.02	0.0335	0.1416	0.11
13	NP	-	P+	0.0001	0.0001	0.00	0.0217	0.0215	0.00	0.0191	0.0342	0.02
14	NP	-	P-	0.0274	0.0176	0.01	0.0469	0.0483	0.00	0.0094	0.0260	0.02
15	NP	-	N	0.2064	0.2050	0.00	0.0390	0.0576	0.02	0.0217	0.1409	0.12
16	NP	-	NP	0.3927	0.4391	0.05	0.0777	0.1248	0.00	0.0214	0.0352	0.24

Comparison of the computed relative sequence distributions. (+/-): Percentage difference in the computational model for both biases: $|model\ with\ bias - model\ without\ bias|$, where $| |$ represents the absolute value.

minimum number of amino acids capable to generate a prebiotic profile of the proteins.

Our results indicate that the relative abundance of the amino acids is the most influential aspect for the sequential characteristics of the “first peptides” as it is shown by the coincidental distribution of the three scenarios that do not seem to be greatly affected by a polarity bias. This last observation could lead to the modeling of a prebiotic scenario with greater granularity, since it would be possible to prioritize the involved biases and use a *hierarchical hidden Markov model* (Fine *et al.*, 1998) where, particularly the abundance, would be a *non-visible* component and the amino acid profile would be the *visible* element to be determined. Computer simulations in this direction are under progress because the mathematical profile of this type of models allows considering several biases, without increasing the computational complexity.

SOFTWARE RESOURCES

We calculated the discrete dynamic system with the Bluebit.NET Matrix Library platform. NML <http://www.bluebit.gr/matrix-calculator/> accessed July 9, 2013; and the matrices: $B^{0.9997}$, $B^{0.9998}$, $B^{0.9999}$ with: GNU Octave <http://www.gnu.org/software/octave/> accessed July 16, 2013. The formation of short prebiotic peptides from mathematical-computational program (Polanco *et al.*, 2013) was written in FORTRAN 77 and executed on a Fedora 14 Unix-type platform (GNU). We run the program from 1 up to 50 generations in an HP Workstation Z210 — CMT — 4 x Intel Xeon E3-1270/3.4 GHz (Quad-Core) — RAM 8 GB — SSD 1 x 160 GB — DVD SuperMulti — Quadro 2000 — Gigabit LAN, Linux Fedora 14, 64-bits. Cache Memory 8 MB. Cache Per Processor 8 MB. RAM 8 GB.

CONCLUSIONS

Using the discrete dynamic system based on the percentage composition of peptide linkages from Rode's experiment on salt-induced peptide formation, we observed that, instead of the polarity bias, the abundance bias on the amino acids plays a major role in the sequential characteristics of the dipeptides. Our simulations based on the Miller, and Fox & Harada experiments converge with the simulation based on the Rode experiment, into a unique profile, being the latter, coincident with the experimental preserved genes.

Conflict of Interests

We declare that we do not have any financial and personal interest with other people or organizations that could inappropriately influence (bias) our work.

Author Contributions

Theoretical conception and design: CP, and JLS. Computational performance: CP. Data analysis: CP, TB, JACG, and JLS. Results discussion: CP, JLS, TB, and JACG.

Acknowledgments

The authors thank Prof. Bernd Rode for stimulating discussions, gratefully acknowledge financial support by the Mexican-French bilateral research grant CONACYT (188689) — ANR (12-IS07-0006), the technical support of Computer Science Department at Institute for Nuclear Sciences at the National Autonomous University of Mexico, and we also to thank Concepción Celis Juárez whose suggestions and proof-reading have greatly improved the original manuscript.

REFERENCES

- Abelson PH (1956) Amino acids formed in primitive atmospheres. *Science* **124**: 935.
- Bahadur K, Ranganayaki S, Santamaria L (1958) Photosynthesis of amino acids from paraformaldehyde involving the fixation of nitrogen in the presence of colloidal molybdenum oxide as catalyst. *Nature* **182**: 1668.
- Boschke FL (1972) *Die Herkunft des Lebens*, pp 187. Verlag GmbH, Dusseldorf un Wien.
- Bujdák J, Faybiková K, Eder A, Yongyai Y, Rode BM (1995) Peptide chain elongation: a possible role of montmorillonite in prebiotic synthesis of protein precursors. *Orig Life Evol Biosph* **25**: 431–441.
- Cloud P (1973) Paleo ecological significance of Precambrian banded iron-formations. *Econ Geol* **68**: 1135–1143.
- Derossi D, Chassaing G, Prochiantz A (1998) Trojan peptides: the penetratin system for intracellular delivery. *Trends Cell Biol* **8**: 84–87.
- Delaye L, Becerra A, Lazcano A (2005) The last common ancestor: what's in a name? *Orig Life Evol Biosph* **35**: 537–554.
- Fine S, Singer Y, Tishby N (1998) The Hierarchical Hidden Markov Model: Analysis and Applications. *Machine Learning* **32**: 41–62.
- Fox SW, Harada K (1958) Thermal copolymerization of amino acids to a product resembling protein. *Science* **128**: 1214.
- Fox SW, Harada K (1960) The thermal copolymerization of amino acids common to protein. *J Am Chem Soc* **82**: 3745–3751.
- Fraser DG, Fitz D, Jakschitz T, Rode BM (2011) Selective adsorption and chiral amplification of amino acids in vermiculite clay-implications for the origin of biochirality. *Phys Chem Chem Phys* **13**: 831–838.
- Garrison WM, Morrison DC, Hamilton JG, Benson AA, Calvin M (1951) Reduction of carbon dioxide in aqueous solutions by ionizing radiation. *Science* **114**: 416–418.
- Gautam A, Singh H, Tyagi A, Chaudhary K, Kumar R, Kapoor P, Raghava GP (2012) CPPsite: a curated database of cell penetrating peptides. *Database (Oxford)* **2012**: bas015.
- Groth W, Weyssenhoff H (1957) Photochemische Bildung von Aminosäuren aus Mischungen einfacher Gase. *Naturwissenschaften* **44**: 510–511.
- Hanic F, Morvova M, (1998) Eleventh symposium on elementary processes and chemical reactions in low temperature plasma. *Low Tatras, Slovakia*.
- Hulett HR, Bar-Nun A, Bar-Nun N, Bauer SH, Sagan C (1970) Amino acid synthesis in simulated primitive environments. *Science* **170**: 1000–1002.
- Isaacson DL, Madsen WR (1976) *Markov chains, theory and applications*. pp 14–16, Wiley Series in Probability and Mathematical Statistics. Inc. USA.
- Jakschitz TA, Rode BM. (2012) Chemical evolution from simple inorganic compounds to chiral peptides. *Chem Soc Rev* **41**: 5484–5489.
- Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* **28**: 27–30.
- Kobayashi K, Tsuchiya M, Oshima T, Yanagawa H (1990) Abiotic synthesis of amino acids and imidazole by proton irradiation of simulated primitive earth atmospheres. *Orig Life Evol Biosph* **20**: 99–109.
- Lawless JG, Boynton CD (1973) Thermal synthesis of amino acids from a simulated primitive atmosphere. *Nature* **243**: 405–407.
- Miller SL (1953) A Production of amino acids under possible primitive earth conditions. *Science* **117**: 528–529.
- Ochiai EI (1978) The evolution of the environment and its influence on the evolution of life. *Orig Life* **9**: 81–91.
- Oro J (1980) Prebiological synthesis of the organic molecules and the origin of life. *The Origins of Life and Evolution*, pp 47–63. Alan R. Liss Inc., USA.
- Palm C, Calvin M (1962) Primordial organic chemistry. I. Compounds resulting from electron irradiation of C¹⁴H₄. *J Am Chem Soc* **84**: 2115–2121.
- Pavolvskaya TE, Pasynskii AL (1959) *The origin of life on the earth* (Proceedings of the first international symposium held at Moscow 19–24 August 1957). New York: Pergamon Press Inc. USA.
- Plankensteiner K, Reiner H, Rode BM (2006) Amino acids on the rampant primordial Earth: electric discharges and the hot salty ocean. *Mol Divers* **10**: 3–7.
- Plankensteiner K, Reiner H, Rode BM (2005) Stereoselective differentiation in the Salt-induced Peptide Formation reaction and its relevance for the origin of life. *Peptides* **26**: 535–541.
- Plankensteiner K, Reiner H, Rode BM. (2004) From earth's primitive atmosphere to chiral peptides—the origin of precursors for life. *Chem Biodivers* **1**: 1308–1315.
- Polanco C, Samaniego JL (2009) Detection of selective cationic amphipatic antibacterial peptides by Hidden Markov models. *Acta Biochim Pol* **56**: 167–176.
- Polanco C, Samaniego JL, Buhse T, Mosqueira FG, Negron-Mendoza A, Ramos-Bernal S, Castanon-Gonzalez JA (2012) Characterization of selective antibacterial peptides by polarity index. *Int J Pept* **58**: 5027.
- Polanco C, Buhse T, Samaniego JL, Castañón González JA (2013a) A toy model of prebiotic peptide evolution: the possible role of relative amino acid abundances. *Acta Biochim Pol* **60**: 175–182.
- Polanco C, Buhse T, Samaniego JL, Castañón González JA (2013b) Detection of selective antibacterial peptides by the Polarity. *Acta Biochim Pol* **60**: 183–189.
- Polanco C, Buhse T, Samaniego JL, Castañón González JA, Estrada MA (2014) Computational model of abiogenic amino acid condensation to obtain a polar amino acid profile. *Acta Biochim Pol* **61**: 253–258.
- Polanco C, Castañón-González JA, Samaniego JL, Buhse T (2014) *Letter to Editor. A Global Challenge Sci. Transl. Med* DOI: 10.1126/scitranslmed.3009315.
- Poole D (2011) *Linear Algebra, A Modern Introduction*. pp 461–490. Brooks & Cole, Cengage Learning Inc, UK.
- Reiner H, Plankensteiner K, Fitz D, Rode BM (2006) The possible influence of L-histidine on the origin of the first peptides on the primordial Earth. *Chem Biodivers* **3**: 611–621.
- Rode BM (1999) Peptides and the origin of life. *Peptides* **20**: 773–786.
- Rode BM, Fitz D, Jakschitz T. (2007) The first steps of chemical evolution towards the origin of life. *Chem Biodivers* **12**: 2674–2702.
- Sagan C, Khare BN (1971) Long wavelength UV photoproduction of amino acids on the primitive earth. *Science* **173**: 417–420.
- Saetia S, Liedl KR, Eder AH, Rode BM (1993) Evaporation cycle experiments — a simulation of salt-induced peptide synthesis under possible prebiotic conditions. *Origins of Life Evol Biosph* **23**: 167–176.
- Schwendinger MG, Rode BM (1989) Possible role of copper and sodium chloride in prebiotic evolution of peptides. *Analyt Sci* **5**: 411–414.
- Schwendinger MG, Rode BM (1992) Investigations on the mechanism of the salt-induced peptide formation. *Orig Life Evol Biosph* **22**: 349–359.
- Thakur N, Qureshi A, Kumar M (2012) AVPPred: collection and prediction of highly effective antiviral peptides. *Nucleic Acids Res* **W199–W204**.
- Thom R (1975) *Stabilité structurelle et morphogénèse: essai d'une théorie générale des modèles*. pp 348–362. Addison-Wesley, Inc, USA.
- Vogel G (1998) A sulfurous start for protein synthesis? *Science* **281**: 627–629.
- Wang G, Li X, Wang Z (2009) APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucleic Acids Res* **37**: D933–D937.
- Watson JD, Crick FHC (1953) A Structure for deoxyribose nucleic acid. *Nature* **171**: 737–738.
- Wolman Y, Haverland WJ, Miller SL. (1972) Nonprotein amino acids from spark discharges and their comparison with the murchison meteorite amino acids. *Proc Natl Acad Sci USA* **69**: 809–811.
- UniProt Consortium Update on activities at the Universal Protein Resource (2013) *Nucleic Acids Res* **41**: D43–D47.
- Urey HC (1952) *The Planets: Their Origin and Development*. New Haven, CT: Yale University Press.
- Yanagawa H, Kobayashi Y, Egami F (1980) Genesis of amino acids in the primeval sea. Formation of amino acids from sugars and ammonia in a modified sea medium. *J Biochem* **87**: 359–362.
- Yoshino D, Hayatsu K, Anders E (1971) Origin of organic matter in early solar system — III. Amino acids: Catalytic synthesis. *Geochimica et Cosmochimica. Acta-Geochim Cosmochim Acta* **35**: 927–938.