ON THE ORIGIN OF LIFE EVENT

F. G. MOSQUEIRA

Department of Chemistry, Centro de Estudios Nucleares, Universidad Nacional Autónoma de México, Ap. Postal 70-543, 04510-México, D.F., Mexico

(Received 14 November, 1986)

Abstract. On the assumption of a uniform sample space probability hypothesis it is estimated a maximum number of polypeptides (or other kind of polymers) that could be synthesized in the prebiotic Earth. Besides, on the basis of five premises that are postulated as indispensable requirements for the origin of a living system, under the constraints of a protein-nucleic acid chemistry, it is concluded categorically that the origin of life event could not be the result of unbiased polymerization phenomena. On the contrary, biased and specific patterns of polymerization had to be an essential component in this fundamental event.

Finally, several theories on the origin of life and complementary concepts like hypercyclic organization and self-organization phenomena in dissipative structures are discussed in the light of the conclusions arrived at in this work.

1. Introduction

A fundamental matter in chemical evolution theory that has remained unanswered has been the origin of information in a prebiotic environment. Ordering of monomers into specific sequences in a polymer is a distinctive feature of living systems. Thus, how could it come forth in the absence of a coding system? How can one conceive a fairly faithful reproduction of a macromolecule (say a polypeptide) on such environments? Is it feasible to obtain a minimum 'coherent system' from a prebiotic environment that could in turn sustain by itself?

In this article I introduce some simple calculations to provide some clues into these questions. To this end it is calculated a probability p associated with the event of obtaining *a set* of n polymers *simultaneously*. Such set n of polymers (refered to as $\{n\}$ for short) is aiming to constitute a fundamental chemical machinery that might be qualified as a most primitive 'living system'. That is, $\{n\}$ is aiming to represent the minimal and most primitive living entity constituted chiefly by proteins and nucleic acids polymers, which should be able to perform some vital tasks. To this end $\{n\}$ is characterized with simple molecular parameters (as degree of polymerization for example) and their values are suggested by the physicochemical properties of the polymers and the functions to be fulfilled by $\{n\}$. In this way, known physicochemical characteristics of the participating macromolecules are a valuable criteria to search into the nature of this fundamental event. In doing this, I try also to give a sharper image of the general and blurred belief that given the appropriate conditions (i.e.: (i) a set of suitable monomers; (ii) a variety of energy sources; and (iii) a long enough period of time) then life appeared somewhere on the primitive Earth (of course, under the assumption that life arose on this planet).

Origins of Life and Evolution of the Biosphere **18** (1988) 143–156.

Another important aspect that is discussed in this paper is the extent of bias in the physicochemical laws governing the polymerization of monomers. To illustrate the extent of bias in the polymerization processes giving rise to a protein-nucleic acid living system we calculate p under a uniform sample space hypothesis, as it is known in probability theory. In the context of polymerization processes of a set of l different monomers, this hypothesis would assume that the probability of reaction of monomer M_i into a growing chain headed by monomer M_j, $(i, j = 1, 2, \ldots, l)$, is the same for every possible combination.

This calculation is valid whether the polymerization process giving rise to 'the first' living system constituted by protein and nucleic acid was performed (a) *ab initio* or (b) by means of an undetermined primitive precursor, as it will be discussed later on. In any case, the purpose of these calculations is to have a quantitative estimate of the extent of bias occurring in real polymerization processes, with respect to an hypothetical situation represented by the uniform sample space hypothesis that has been used to estimate p. Undoubtedly, these considerations emphasize the importance of mechanisms producing nonrandom polymers, such as the self-ordering principle of α-amino acids (Fox and Dose, 1977; Fox, 1978).

In the next section, we shall evaluate the order of magnitude of the total number of polypeptides that could be synthesized on Earth under several physical conditions. This estimate is necessary to calculate p. It is important to notice that such estimate can be also applied to other polymers, like nucleic acids, as it will be shown in Section 3. In the light of these results, I comment on several theories on the origin of life that have been proposed to sketch roughly a series of stages leading to the origin of life. I also point out a fine distinction in the meaning of the term 'self-organization' as it is used in the thermodynamic theory of nonequilibrium systems (Nicolis and Prigogine, 1977) and compare it with the self-organization process associated with the origin of life event in the context of a protein-nucleic acid chemistry.

2. An Estimate on the Maximum Number of Polypeptides Synthesized on the Earth

To calculate an approximate order of magnitude of such number we need figures on the following parameters: total volume of reaction; volume of the sites producing the polypeptides; rate of polypeptide synthesis on such sites; and total time alloted to this process to occur.

We will adopt a pair wise calculation. On the one hand, the parameter values are intentionally exaggerated. It will be called the over estimated version (OEV). This is done purposefully to illustrate an important aspect of this problem to be discussed later on. At the same time an estimation is made on another figure which intends to follow more closely realistic conditions and will be referred to as the realistic version (RV). Obviously, as we vary the parameter values the final figure will correspondingly be altered. However, it does not affect the main conclusions to be shown.

Total volume of reaction. On OEV let us consider a volume the size of an ocean

3.2 km depth (which is the average depth of contemporary oceans) covering the *entire* Earth. The calculated volume is 1.630055×10^{18} m^3. On RV we use the volume of contemporary oceans, about 1.2×10^{18} m^3. Notice that in using this value we are in fact overestimating such parameter as we will assign synthesis activity under conditions (as extreme pressure) which are not compatible with the kind of system we have in mind.

Volume of the sites producing the polypeptides. On OEV we take regular cubes 10^{-3} nm^3 of volume. This corresponds to a cube in which its edges are about the length of a covalent bond. Of course, in such reduced space there is not even room to fit in a water molecule. In the RV calculation the edge of the regular cube is increased to 4.0 nm, which is about the length of one subunit of the hemoglobin molecule.

Rate of polypeptide synthesis on such sites. On the OEV we will assign a rate of 10^6 synthesized polypeptides per second in *every* regular cube. Notice that this is well overestimated, as under optimal conditions the time required for the synthesis of an *E. coli* polypeptide chain containing 300 to 400 amino acids is 10 to 20 s (Watson, 1970). On the RV we assign a rate of 1 polypeptide synthesized per hour in every cube, that is 2.78×10^{-4} polypeptides per second per site.

Total time alloted to this process to occur. On OEV, we use the age of the Earth, about 4.6×10^9 yr. On RV we consider the lapse in between the origin of the Earth and the age of the oldest fossils of prokaryotes, estimated in about 3.4×10^9 yr. That is a lapse of about 1.2×10^9 yr.

In fact, the period available for polymerization processes is still shorter as strictly speaking it should be taken into account time for Earth's cooling and ocean formation. However my purpose is to estimate the order of magnitude of time's availability using crude data. This will save going into details and unnecessary complex calculations as it might be the time's calculation for cooling of an object such as the Earth. Anyhow, the use of less favorable data (as a longer period available for polymerization processes) will not alter in the least the point I want to underline in this work.

With the above data it is calculated the total number of polypeptides synthesized on Earth under OEV and RV. The order of magnitude of such numbers are, respectively, 10^{71} and 10^{56} polypeptides and will be refered to as N_{OEV} and N_{RV}.

3. On the Extent of the Sample Space Associated with $\{n\}$

Up to this moment we have estimated the equivalent to the number of times we can participate in a game of chance. The probability to obtain a desired outcome it all depends on that, the number of times we dispose to play such game. Consider a chance game in which there exists 30 possible outcomes, then at every time we play we have a probability of 1/30 to win. But if we play such game several times, our chance to win increases correspondingly. Suppose we are allowed to bet eleven times to a number in this game, then we would have a probability of 11/30 to win. If we

may bet a hundred times then we have all chances to win such game, even though the probability of each individual throw is 1/30.

So, as a second part of our calculations we should establish how big is our sample space, as it is called in probability theory, in order to compare it with N_{RV} (and N_{OEV}) and finally estimate the degree of feasibility or probability of the event we are interested in. To do that we would need to specify the properties and characteristics of $\{n\}$. We shall then pause for a while to think about a series of conditions or requirements that should fulfill the participating polymers in $\{n\}$ as components of a self-sustaining system or 'minimum living system'. As it has been said previously, $\{n\}$ stands for a minimal chemical machinery made of proteins and nucleic acids that is able to perform some vital tasks inherent to a living entity. The reason to select this chemical composition should be fairly obvious. Life has adopted proteins and nucleic acids as their chief building blocks. Thus, its appearance in chemical evolution may be explained as a convergence problem. A minimal chemical organization made of proteins and nucleic acids had to come forth either as a result of (i) *ab initio* polymerization processes in solution or surfaces, or (ii) as the result of some action of precursors of different chemical composition. Whatever the route, we have to explain the appearance and minimal chemical organization of a *protein and nucleic acid system*. (In Section 5 I will discuss precursors in more detail and show that this situation do not alter the main conclusions of this paper).

It would be immaterial to try to assign to $\{n\}$ specific tasks as replication, translation, etc. Nevertheless, one can propose some general and indispensable activities. A living entity had to have at least the ability to reproduce itself, although crudely. This clearly suggests $\{n\}$ was able at least to: (a) synthesize proteins and/or nucleic acids and (b) perform other vital activities. Anyhow, the important point to be highlighted is that no matter what were the working abilities of this early entity, they had to rely on molecules, or more precisely macromolecules. Indeed, some pysical chemical properties of proteins and nucleic acids are used to set minimal limits in molecular parameters and describe $\{n\}$ more accurately.

Now, let us enumerate and comment on the following premises postulated as essential characteristics of $\{n\}$ and its components:

(1) Plurality. This should appear to be the most obvious premise even if it is not always completely realized. The functioning of a coherent system or a living system should depend on the cooperation of *several* macromolecules. We can not equate a living system with a *single* macromolecule. This is why we are concerned with a set of macromolecules that has been called $\{n\}$.

(2) Simultaneity. A coherent system will need not only a set of macromolecules but it will also be indispensable that every macromolecule will collaborate or participate in a precise manner in a kinetically connected network of chemical reactions. If one macromolecule is missing in $\{n\}$, the coherent action will surely be lost. Thus, we need the whole set $\{n\}$ *simulatensouly* at some physical locality. Such premise imposes a very restrictive mathematical condition. That is, if event I has the probability p_I to occur and another independent event II has a probability p_{II}, then the proba-

bility to have simultaneously both events I and II is $p_I p_{II}$. Of course, if p_I, $p_{II} < 1$, then $p_I p_{II} < p_I$, p_{II}. In a three member polymer system such constraint may be illustrated as follows. Let us call such polymer as α, β, and γ. The associated probabilities to obtain such polymer are, correspondingly, p_α, p_β, and p_γ. Thus, the probability to have α, β, and γ simultaneously at the same physical locality would be $p_\alpha p_\beta p_\gamma$. Again, as p_α, p_β, $p_\gamma < 1$ then $p_\alpha p_\beta p_\gamma \ll p_\alpha$, p_β, or p_γ.

A most obvious prebiotic condition to assure the set of macromolecules in $\{n\}$ will act into physical proximity is to suppose $\{n\}$ is located into a compartment like a primitive membrane. If we do not make such assumption, polymers will simply diffuse away and the simultaneity condition that is required would be lost.

The participation of a primitive membrane will not be considered in the following probability calculation. It is assumed its formation is an entropy driven process and therefore it does not represent a limiting factor for the origin of a coherent molecular system.

(3) Number of Participating Macromolecules in $\{n\}$. It has been previously suggested that an early entity had to have at least the capacity to: (1) Synthesize protein and/or (2) nucleic acids. Besides such *essential* tasks, the early entity had also to cope with activities derived either from an (3) autotrophic or (4) heterotrophic start. In turn, these quite plausible suppositions imply of course molecular support. If we assume an autotrophic start then we should assign a number of macromolecules to secure a most primitive 'metabolism'. On the other hand, an heterotrophic start faces the problem of the transport of required solutes across a primitive membrane. That is, a plain membrane acts rather as a barrier with respect to charged ions and certain indispensable solutes. Even molecules containing OH groups, as it is the case in simple carbohydrates as glucose, would be hindered to pass across a membrane (Lehninger, 1982). Such situation would suggest to assign a small number of 'carriers' into a primitive membrane. So, whatever our hypothesis is about the nature of our more distant common ancestor, we are constrained to include a *set* of macromolecules that collaborate to a given task.

On this basis we will assume that $\{n\}$ is constituted at least by a number of macromolecules falling into the following range: $8 < n < 14$. Again, this is not by any means an overestimate. On the contrary, one could hardly conceive how such a minimum chemical equipment of macromolecules could fulfill the most vital requirements as (1) and/or (2), and (3) or (4). Besides, we should bear in mind that in contemporary biochemistry the number of participating macromolecules just required in protein synthesis is greater than 300 (Lehninger, 1982).

(4) Degree of Polymerization x. If we deal with proteins and nucleic acids like RNA macromolecules, a relatively rigid tertiary structure seems to be an essential feature of a functional macromolecule. So, it would be required to reach a certain degree of polymerization in order to obtain a sufficient number of weak forces, like hydrogen bonds, to give the polymer the necessary rigidity to perform its chemical tasks. Otherwise, oligomers with a low degree of polymerization would predominate in a random configuration and it would be impossible to assign any function to such oligomers.

Thus, we postulate $x \approx 40$ (monomers) as an indispensable characteristic of the macromolecules that constitute $\{n\}$.

Notice that a degree of polymerization of about 40 is indeed a very low estimate as in contemporary biochemistry it is common to have about 300–400 amino acids per polypeptide and sometimes it goes up to about 3000 monomers.

(5) Alphabet a. It has been frequently postulated that a primitive genetic code would have started with a very coarse and rudimentary capability to distinguish amino acids. For that reason it has been assumed that a two letter code would be the more likely beginning, with the ability to distinguish only among hydrophylic and hydrophobic amino acids. On this basis we will adopt a two letter alphabet ($a = 2$) and avoid the more complicated start with $2 < a \leq 20$.

Observe that strictly speaking, in a prebiotic situation we should include the two optical forms L and D for every letter. This of course would bring the alphabet into a four letters one. However, we will adopt the more adverse hypothesis and consider $a = 2$ to perform the calculations. With this assumption we gain generality as it may also refer to the RNA type macromolecules as the constituent polymers of $\{n\}$. Of course the two letter hypothesis is taking into account (in large excess) the rare possibility that two or more different polypeptides may perform the same chemical action.

4. On the Order of Magnitude of p

Finally, we are ready to specify the main characteristics of $\{n\}$ and so calculate the extent of its associated sample space. Let us first summarize the parameter values we will use for $\{n\}$: $n = 10$; $x = 40$; and $a = 2$.

Then it is calculated the number of possible polypeptides that could be synthesized with $a = 2$ and $x = 40$. Notice that the first position of the polypeptide may be occupied by any of two kinds of monomers, the second position again may be occupied by any of two possibilities, and so on. Thus, the total number of polypeptides that might be constructed under such constraints is 2^{40} ($\approx 10^{12}$). In general, the total number of polypeptides is a^x.

The probability to obtain any predetermined polypeptide in this group is 10^{-12} and the probability to obtain 10 predetermined such objects simultaneously (see premise 2) is 10^{-120}. This result means that in order to have a non vanishing probability to get the predetermined set $\{n\}$ with the aforesaid characteristics ($n = 10$; $x = 40$; and $a = 2$) we would have to have about 10^{120} 'throws'. Only if we play with such high 'intensity' (about 10^{120} polypeptides or throws) we would have a finite chance to get the predetermined set $\{n\}$ with the aforesaid characteristics. Before going on we define $E = (a^x)^n$ as a parameter which measures the extent of the sample space.

Now we ask, how intensive has been the polymerization game on Earth? We have previously estimated and compiled it in N_{RV} (and N_{OEV}). It rises to $N_{RV} = 10^{56}$. Thus the ratio of the total number of throws N_{RV} to the extent of the sample space E would give us the value of the probability p we have been looking for. It is the proba-

bility to obtain a set of objects $\{n\}$ with its predetermined characteristics (i.e., $n = 10$; $x = 40$; and $a = 2$) and evaluated as $p_{RV} = N_{RV}/E$. In this case $p_{RV} = 10^{-64}$. But this number is so small that it is just another way of writing zero. It would be completely unreasonable to expect to 'win' with such low probability value. Even with N_{OEV}, the corresponding probability value would be zero: $p_{OEV} = 10^{-49}$.

It might be argued that such polymerization game is being played in the entire Universe and that the planet Earth has been lucky enough to win in the Life Game. Let us inquire into this possibility by using again over estimated parameter values. At most, the mass of the visible Universe is 10^{60} g. On the other hand the order of magnitude of the Earth's mass is 10^{27} g. So, there is enough mass to have 10^{33} Earths. Notice that it is a very crude over estimate, of course, as the majority of the mass of the Universe would be either dispersed or in the plasma state. Besides note that even if matter were condensed in a planet, it has still to fulfill a series of conditions before it can be a candidate to support life.

Let us then suppose that the polymerization game has been played in 10^{33} objects as the Earth. The total number of throws would rise to $N_{*RV} = N_{RV} \times 10^{33} = 10^{89}$, thus $p_{*RV} = N_{*RV}/E = 10^{-31}$ ($\equiv 0$). Even the overestimated version is insufficient: $N_{*OEV} = N_{OEV} \times 10^{33} = 10^{104}$, and $p_{*OEV} = N_{*OEV}/E = 10^{-16}$, still a very small number

As we have outlined previously, a change in the parameter values would correspondingly alter the result. However, in any case we arrive to extremely small values for p_{RV} (and p_{OEV}, p_{*RV}, and p_{*OEV}) that in fact may be equated with zero. Indeed, the purpose of introducing a parallel over estimated version (OEV) has been to discard the possibility of supposing that if data on the realistic version (RV) were chosen otherwise (perhaps with higher values in the parameters and conditions of polypeptide synthesis on Earth), then figures would be sufficient to make the appearance of $\{n\}$ as specified a feasible phenomenon. The point to make it clear is that though we exaggerate in the parameter values we choose, under a uniform sample space hypothesis we cannot even explain the appearance of $\{n\}$, constituted by 10 small polymers, 40 monomers of length each of them, with a coarse two letter alphabet, and supposing as many Earths as possible from the entire mass of the visible Universe.

We should outline, however, that these results are valid under a *uniform* sample space hypothesis. That is, it is assumed that different monomers react exactly with the same afinity among them to constitute a polymer. However, the above calculations show very clearly that it has not been the case. On the contrary, polymerization processes that gave rise to the first living entity had to be associated with extremely *biased* mechanisms. This aspect will be amply discussed in subsequent sections.

5. The Nature of the Problem and Theories on the Origin of Life

Up to this point we have deepened into several important aspects in the origin of life event. On the one hand, a first coherent or living system had to rely on some minimum molecular equipment $\{n\}$ consituted by several macromolecular entities. Con-

ceptually, we cannot elude premises 1 to 5 as it cannot be conceived a simpler start for the kind of life we are trying to explain: one based on a protein-nucleic acid chemistry. A 'start' cannot be conceived without the collaboration, as rudimentary and primitive as it might be, of several such polymers that interact simultaneously into physical proximity to accomplish vital chemical activities. If we have less than such a minimum molecular equiment $\{n\}$, then we cannot expect life phenomena. That is, it would be useless to have, say 1/3, or 3/4, or even 9/10 of $\{n\}$. In order to have a coherent or living system it is required $\{n\}$ wholly. Integrity of vital functions is an essential feature of living matter.

This discussion then poses a general difficulty to theories on the origin of life. Indeed, such theories usually assert that the origin of a living system was performed in a stepwise manner. However, how many conditions in turn make up one 'step'? We have seen that for an early coherent system made of protein and nucleic acid it is required to fulfill several condition *simultaneously*, in order to have a functional and complete entity $\{n\}$. In the hypothetical case of a polymerization process modeled by a uniform sample space assumption (as it has been used in this paper) it is certainly an unsurmountable step to get $\{n\}$ simultaneously, as we have verified in the numerical example above. Thus, we are conveyed to conclude that polymerization process associated with the emergence of life had to rely on biased or preferential polymerization process. That is, real polymerization processes had to conform rather to a *non-uniform* sample space assumption in order to make feasible the appearance of a minimum nucleation set of coherently linked macromolecules $\{n\}$.

Let us make then a simple appraisal of the extent of bias of real polymerization processes with respect to the hypothetical uniform sample space used in this paper to calculate p_{RV}. In order to make the appearance of $\{n\}$ a likely event, a number such as 10^{120} (in the example I worked out above) has to be reduced to an order of magnitude similar to N_{RV}. This would mean a reduction by a factor as small as 10^{-64} to give this polymerization game a possibility of being 'won'. That is, the non-uniform sample space associated with the polymerization process giving rise to a living system should have been stringent enough as to 'pick up' a fraction as small as 10^{-64} (with respect to 10^{120}). We can thus conclude that polymerization processes involved in the origin of a living system made of protein and nucleic acids had to be astonishingly biased.

Now, let us examine the effect of primitive precursors on these conclusions. Theories on the origin of life have proposed different stages before the emergence of an early living system made of protein and nucleic acids. Classical chemical evolution theory postulates that more elaborate and complex systems arose from a prebiotic aqueous medium with the aid of a variety of monomers and energy sources (Oparin, 1957; Calvin, 1969; Miller and Orgel, 1974). Alternative theories, like the protenoid (Fox and Dose, 1977) and the mineralic (Cairns-Smith, 1982) origin of life, put much more emphasis on the appearance of primitive precursors as an intermediate stage before an early version of a protein-nucleic acid biochemistry could have emerged. These theories propose then a more gradual evolution from more primitive and

rudimentary ancestors made of a different chemical nature. Nevertheless, it should be stressed that *whatever* the nature of the precursor (that is, either it is a chemical, a protenoid, or a mineralic precursor) *it is compelled to give rise* to a minimum and *complete* set of macromolecules $\{n\}$ made of protein and nucleic acids. So, premises 1 to 5 should also remain valid, and the conclusions conerning the extremely biased character in the polymerization process giving rise to $\{n\}$ will also apply whether we consider primitive precursors or not.

These considerations can be generalized to *any* theory on the origin of life. It can be asserted that any theory on this subject should be able to provide suitable mechanisms to produce heavily biased polymers of amino acids and nucleotides. That is, a non-uniform sample space constraint should remain valid either one is considering polymerization processes *ab initio* or via primitive precursors.

These are, I think, some of the features and problems to elucidate before we can give a plausible explanation to the origin of life event.

6. Some Corollaries

From this analysis we can derive a series of important observations and corollaries. First and more important, the origin of life event had to be associated with a wholly *biased* mechanism of polymerization as it has been outlined above. Otherwise, we are conveyed to deal with superastronomical numbers and consequently constrained to assign a zero probability to the origin of life event. Second, in relation to the ultimate source of information in a prebiotic system it can be said that, if it resides somewhere in this physical world, it should be at the *monomer level*, specifically that of α-amino acids. In the absence of a coding system, information has to reside at the monomer level, it cannot be at the polymer level as it is in contemporary biochemistry. Under prebiotic conditions, a synthesized polypeptide should be the result of the expression of information already contained in the residue group of α-amino acids. The main aspects to analyze and study would be then to make an appraisal of the degree of information within α-amino acids as a factor to induce polypeptide reproduction with a reasonable degree of fidelity.

It is noteworthy that the polymerization of another kind of monomers fundamental for life, as nucleotides, in itself cannot give rise to any ordering principle as the polymerization of nucleotides involves always the same chemical groups. In this case the chemical variety of bases do not have any (chemical) chance to induce a self-ordering principle with respect to nucleotides. In contrast, the proximity of the residue groups of α-amino acids during their polymerization opens the possibility of a self-ordering principle (Fox and Dose, 1977; Fox, 1978).

Another corollary related to the previous one is that a nucleic acid molecule in a prebiotic scene cannot have the ability to transfer information to macromolecules of a different chemical nature, for example proteins, as it occurs in present day biochemistry. This observation had been previously suggested from the study of a theoretical model on a primitive translation mechanism (Mosqueira, 1979) in which

nucleic acids acts only as a support molecule, not as an informational molecule, because in a prebiotic medium it is missing the necessary molecular recognition mechanism to give 'sense' to a nucleic acid molecule. It is nothing more than a prejudice from our present knowledge of molecular biology to associate information to a nucleic acid molecule. This is true in contemporary biochemistry but not in a prebiotic milieu. In the former the nucleic acid molecule is supported by a chemical equipment of more than 300 macromolecules to give sense to a nucleic acid molecule (Lehninger, 1982). In a prebiotic system, obviously, it is lacking such macromolecular support. Thus a prebiotic nucleic acid molecule can only perform as much as its physical chemistry properties can afford and certainly information transfer capabilities to proteins cannot be assigned to them in a prebiotic system (Mosqueira, 1979). At most, under prebiotic conditions nucleic acids may have the capacity to transfer information only within molecules of the same chemical nature through a process of prebiotic replication. Alternatively, it may be assumed that the molecular equipment able to give sense to a prebiotic nucleic acid is made up of a very small number of macromolecules, for example $n = 10$, and make a probabilistic appraisal of its consequences under a uniform sample space hypothesis as it is done in the present article.

Further, it is noteworthy to mention that strictly speaking a sequence of nucleotides in the absence of a coding system bring us again to face superastronomical numbers. Consider the following simple calculation pointed out by Ayala (1976) in which it is refered to the present day biochemistry with 21 different units of information (20 amino acids plus a termination signal) and a 600 nucleotides long polynucleotide, as a typical length for a structual gene. Under such conditions, the total number of possible messages for a chain that length is $21^{200} = 10^{264}$, a number exceedingly larger than the total number of atoms in the visible Universe (about 10^{80}). Such enormously complicating aspect brings us to consider another very important requirement in the origin of life event. Suppose we have a set $\{n\}$ of macromolecules that are able to express information from a sequence of nucleotides, how is it that the expression of a sequence (which is random under prebiotic conditions) gives as a result primitive *life phenomena* in spite of the huge number of possible and different messages? Is it then that *any* combination of nucleic acids and proteins produce life phenomena? This could hardly be the answer as we know that there exists a very strong tendence during evolution to conserve certain proteins, as for example histones, cytochromes, globines, etcetera. It should be admitted that we really do not have any basis, neither experimental nor theoretical, to give even a faint clue to answer this important question.

Anyhow, even if it could be clearly established the possibility of production of small molecules of nucleic acids under prebiotic conditions (either in solution or with the aid of mineral organisms or a protenoid system) we should avoid to assign to it any informational content as it would require *in addition* the assistance of a select number of polymers to express such 'information'.

On this perspective, hypercyclic organization (Eigen, 1971; Eigen and Schuster, 1977; Eigen and Schuster 1978a; Eigen and Schuster 1978b) should be considered as

a *postbiotic* event rather than a *prebiotic* one, as the expression of information from a polynucleotide I_i to give a polypeptide E_i presuppose already the existence of a primitive genetic code: a set of molecular tools $\{n\}$ able to perform such information expression task.

7. Comments on Self-Organization Phenomena

Finally, a few words about the self-organization phenomena observed in nonequilibrium system (Nicolis and Prigogine, 1977) seem in order now. This theory has been an important theoretical tool in several areas in science. Its dissipative structure and order through fluctuations concepts have been successfully used in chemistry and biology (Nicolis and Priogogine, 1977). However, a fine distinction in the use of the word 'self-organization' must be pointed out.

As I have stressed in this article, self-organization phenomena related to the origin of life event had to be associated with totally *biased* mechanism of polymerization. Thus, physicochemical processes inducing the appearance of biased and specific patterns in the sequencing of monomers into a polymer are of special relevance to the origin of life event. This condition will be called requirement **I** for the origin of life with a protein-nucleic acid chemistry. Besides, the geometry and specific fitting and binding of polymers and monomers, although primitive and rudimentary as it might be, should be another indispensable condition for the origin of life to occur as *molecular recognition* phenomena is an essential feature of living matter. Let us call requirement **II** to this condition. Taking into consideration that it has not been always realized, it is quite important to emphasize that under prebiotic conditions (i.e., in the absence of a coding system) requirement **II** *must be* a consequence of requirement **I**. That is, in the absence of a coding system any recognition phenomena (for example structural or catalitic abilities) in 'the first' protein and nucleic acid living entity $\{n\}$ should be the result of extremely biased physicochemical features of the polymerization process producing the polymers embodied in $\{n\}$.

So, the conjunction of requirement **I** and **II** *might* then give rise to a coherent or even a living system. These requirements represent two aspects of self-organization of matter necessary to the emergence of life under prebiotic conditions.

Now we ask what is the relation between self-organization associated to the origin of life event and that observed in dissipative structures, which arises because of particular nonlinearities in the kinetic equations governing a nonequilibrium system (Nicolis and Prigogine, 1977). The emergence of spatio-temporal order in the later is indeed spectacular and striking. However, its influence on the origin of life event appears to be limited due to several reasons. To estimate the degree of influence of self-organization from dissipative structures to the origin of life event we should pose the question otherwise. To what extent requirement **I** is affected by dissipative structures phenomena? In connection with this aspect, it has been recently suggested a possible mechanism to induce specific patterns in the sequencing of monomers into polymers by means of an oscillatory dissipative structure (Prigogine, 1984). This ef-

fect may arise as a consequence of an oscillatory mechanism of the type of a Belousov-Zhabotinski reaction. The reasoning is as follows: suppose there was a far-from equilibrium chemical system with three reactants X, Y, and Z that oscillate. As in the case of the Belousov-Zhabotinski reaction, let us assume that the concentrations of these variables reach their maxima in a well defined order: X reaches its maximum first followed by Y and Z successively. The order X, Y, and Z is thus determined and fixed by nonequilibrium kinetics. Besides, it is assumed that such system is coupled to a polymerizing catalyst. In consequence, this chemical system would produce polymers like X-Y-Z-X-Y-Z ... (A) in much greater quantities than other polymers like X-Z-Y-X-Z-Y ... (A) in much greater quantities than other polymers like X-Y-Z-X-Y-Z ... (B). After some time, if the oscillations die out due to some reason, this chemical system would have left behind an abundance of polymer (A).

There are several remarks to be pointed out and discussed concerning this possible mechanism. Of course, first it should be experimentally established that monomers like α-amino acids or nucleotides may be involved in an autocatalytic kinetics of the ture of the polymer. If the polymer (A) is a nucleic acid then, as it has been stressed in the present article, we would need in addition a minimum chemical equipment, like $\{n\}$, in order to express the 'information' from such polymer. Besides, even if such nucleic acid has specific patterns of polymerization as polymer (A), we cannot assign to it any informational content (except its informational ability to reproduce itself) as there is a lack of molecular specificity mechanisms in a prebiotic system (Mosqueira, 1979). Thus, the appearance of specific patterns in *nucleic acid* polymers as a result of an oscillatory dissipative structure *is of no aid* to make feasible the origin of life event.

If the polymer (A) is a polypeptide, then an oscillatory behaviour should be overlaped to another physicochemical effect: the self-ordering principle of amino acids (Fox and Dose, 1977; Fox, 1978). This principle accounts for the experimental production of biased polymers. Such effect should be directly related to the physical chemistry aspects of the very process of polymerization of α-amino acids as it involves the molecular interaction among the different residue groups of the α-amino acids. Thus, either the self-ordering principle alone or a joint action of the self-ordering principle and an oscillatory kinetics (provided it can be experimentally established the ability of α-amino acids to participate in such oscillatory kinetics) may produce biased polymerization patterns. In consequence those phenomena may be relevant to the origin of life event. Thus, in this case the self-organization behaviour derived from oscillatory dissipative structure *may be a complementary* factor to induce biased polymerization. To discuss if biased polypeptides will satisfy requirement **II**, that is will produce molecular recognition capabilities, is still premature as there exist neither theoretical nor experimental basis to speculate on the catalytic properties of such biased polypeptides produced under prebiotic conditions.

To sum up: the thermodynamic theory of nonequilibrium systems *is* an important *complement* as a theoretical tool in the study and comprehension of the origin of life. It is the right framework to deal with a nonequilibrium system as a prebiotic one cer-

tainly was. However, the spatio-temporal organization typical of a dissipative structure *may be* of *limited* influence to the origin of life event, as it has been discussed above. True, both situations, dissipative structures and the origin of life event imply self-organization phenomena and strictly require open system conditions to permit the essential flow of energy and matter into the system. However, the origin of life event imply a *much higher degree of self-organization* in comparison to what it can be afforded by a dissipative structure. The origin of life event should involve many more conditions or self-organization aspects of matter to be fulfilled before we could conceive the emergence and functioning of a living system. The typical self-organization level of a dissipative structure may be considered as a *complement*, but cannot account for the higher degree of organization involved in the origin of life event.

Several requirements in the origin of life event appear to be indispensable: a set of macromolecules $\{n\}$ that (i) is chemically and coherently articulated (premises 1 to 5), (ii) is able to work on the basis of a primitive molecular recognition mechanism (requirement **II**), which in turn emerged from a heavily biased polymerization mechanism (requirement **I**), and (iii) is able to self-replicate repeatedly with reasonable fidelity. These requirements certainly represent *other aspects* of self-organization of matter which exceeds dissipative structures phenomena.

8. Conclusions and Perspectives

The probabilistic hypotheses used in the aforegoing calculations have assumed an *equal* probability of reaction among all different amino acids. Such hypothesis would be consistent with the production of unbiased polymers. However, experimental work, mainly from the thermal condensation of α-amino acids (Fox and Dose, 1977) has shown that such monomers react differently among them. In fact, this observation supports the postulation of the self-ordering principle of α-amino acids (Fox and Dose, 1977; Fox, 1978). Thus, the polymerization game has been played with charged dice, not with fair dice.

Now, the self-ordering principle of course would induce a reduction in the number of possible polypeptide synthesized under prebiotic conditions. However, would this reduction be so large as to make the appearance of $\{n\}$ a feasible event? Or more generally, would the primitive precursors (be it a protenoid, mineralic, or else) generate a polymerization process stringent enough to make likely the appearance of an early protein-nucleic acid living system $\{n\}$? And if it is so, what would this mean? In order to make the appearance of $\{n\}$ a feasible event, a number such as 10^{120} (in the present example) has to be reduced drastically by a factor of 10^{-64} as it has been pointed out in Section 4. That is, the self-ordering principle or primitive precursors should be powerful enough as to 'choose' a fraction as small as 10^{-64} (with respect to 10^{120} in this example performed under the uniform sample space assumption) to give this polymerization game a possibility of being won. So, to evaluate the reach of the preferential polymerization phenomena (which is intimately associated with

the origin of life event) appears to be of paramount importance. From these preliminary calculations it would be rather a big surprise that the self-ordering principle or an undetermined primitive precursor could bring about such drastic reduction. Eventually, it should be tested the joint effect of an oscillatory source of α-amino acids (in case such monomers could participate in an oscillatory kinetics of the type of a Belousov-Zhabotinski reaction) and the self-ordering principle to quantify their influence in the production of biased polymers. Anyhow, this matter should be further analyzed and clarified with the aid of probabilistic methods to give a rigorous answer to this fundamental question.

References

Ayala, F. J.: 1976, in F. J. Ayala (ed.), *Molecular Evolution*, Sinauer Associates, Inc., Sunderland, Massachusetts.

Cairns-Smith, A. G.: 1982, *Genetic Takeover and the Mineral Origins of Life*, Cambridge University Press, London.

Calvin, M.: 1969, *Chemical Evolution*, Oxford University Press, London.

Eigen, M.: 1971, *Naturwiss.* **58**, 465–523.

Eigen, M. and Schuster, P.: 1977, *Naturwiss.* **64**, 541–565.

Eigen, M. and Schuster, P.: 1978a, *Naturwiss.* **65**, 7–41.

Eigen, M. and Schuster, P.: 1978b, *Naturwiss.* **65**, 341–369.

Fox, S. W. and Dose, K.: 1977, *Molecular Evolution and the Origin of Life*, Marcel Dekker, New York.

Fox, S. W.: 1978, in W. H. Heidcamp (ed.), *The Nature of Life*, University Park Press, Baltimore.

Lehninger, A. L.: 1982, *Principles of Biochemistry*, Worth Publishers, New York.

Miller, S. L. and Orgel, L. E.: 1974, *The Origins of Life on the Earth*, Prentice Hall, Englewood Cliffs, New Jersey.

Mosqueira, F. G.: 1979, *BioSystems* **11**, 233–242.

Nicolis, G. and Prigogine, I.: 1977, *Self-Organization in Nonequilibrium Systems*, John Willey & Sons, New York.

Oparin, A. I.: 1957, *The Origin of Life on Earth*, Academic Press, New York.

Prigogine, I.: 1984, *Adv. Chem. Phys.* **LV**, 43–62.

Watson, J. D.: 1970, *Molecular Biology of the Gene*, W. A. Benjamin, New York.