

Origins of the First Cell.

A New Model for the Spontaneous Formation of the First Living Cell Based on a Novel Approach

Semih Erhan

2101 Chestnut Street, Philadelphia

(Z. Naturforsch. **32 c**, 1003–1010 [1977]; received June 20/August 9, 1977)

Evolution, Origins of Life, Cell, Model

Whether proteins or nucleic acids were responsible for the emergence of life has been debated for a long time. Taking the observation that families of proteins display a remarkable invariance of their amino acid sequence around critical regions, such as active/binding sites, even though these proteins may represent considerable evolutionary diversity, as the naturally provided evidence of evolutionary selection of a working system, the idea is developed that:

1. Proteins had to have been first informational macromolecules that were necessary and sufficient to lead to the emergence of life;
2. it is impossible for a nucleic acid molecules to have formed, by chance, whose base sequence could yield a biologically active protein.

A model is proposed to account for the emergence of the first successful cell according to this approach.

Introduction

It has become almost customary, in certain scientific quarters, to consider life as *the* evidence for an act of God. Some scientists even go as far as stating that the emergence of life can only be conceived as a result of a lucky accident. A typical example of this attitude is seen in a book entitled “Science of Life”, by the prominent experimentalist Weiss who writes “In stressing the indispensable existence of those ultimate prerequisites of primordial life, I anticipate the inescapable conclusion that life, as we know it, cannot possibly have developed gradually in consecutive steps, but must have come about in a major cataclysm”¹. In other words Weiss is saying that, since there is so much complex interdependence in living cells now, this interdependence must have been created at the very beginning, because such a complex interdependence could never have evolved from simpler states. Theoretical considerations, however, make such a “lucky accident” inconceivable because the probability of its happening is low enough to be taken as impossible – 10^{-510} – ².

Against this background it is quite reasonable to ask the question: “Is the phenomenon called life indeed so incomprehensible that it could only appear as a result of an act of God or is it possible to analyze it scientifically and to draw some reasonable conclusions?”

The ever increasing wealth of knowledge in biochemistry and molecular biology makes one rather optimistic that a positive answer can be given to this question. This optimism stems particularly from the series of observations on the complicated phenomenon called morphogenesis whereby complex macromolecules tend to organize themselves, spontaneously, so that ruptured membranes are repaired, functional particles such as viruses and organelles such as ribosomes are formed. This optimism is further reinforced by the demonstration that the folding of proteins depends solely on their amino acid sequences and again occurs spontaneously³.

If furthermore, one is able to find:

- 1) the existence of simpler forms of life today;
- 2) phenomena which could explain how living organisms might have evolved more complicated proteins, and
- 3) ways to extrapolate towards mechanisms through which first macromolecules and still simpler forms of “life” might have been formed, and perform simulation experiments to test the various conclusions drawn and hypotheses formulated and obtain supporting evidence to certain steps in these hypotheses, then this optimism will be justified still further.

Mycoplasma fulfill the first requirement, as has been described by Morowitz, which have about 600



Dieses Werk wurde im Jahr 2013 vom Verlag Zeitschrift für Naturforschung in Zusammenarbeit mit der Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V. digitalisiert und unter folgender Lizenz veröffentlicht: Creative Commons Namensnennung-Keine Bearbeitung 3.0 Deutschland Lizenz.

Zum 01.01.2015 ist eine Anpassung der Lizenzbedingungen (Entfall der Creative Commons Lizenzbedingung „Keine Bearbeitung“) beabsichtigt, um eine Nachnutzung auch im Rahmen zukünftiger wissenschaftlicher Nutzungsformen zu ermöglichen.

This work has been digitalized and published in 2013 by Verlag Zeitschrift für Naturforschung in cooperation with the Max Planck Society for the Advancement of Science under a Creative Commons Attribution-NoDerivs 3.0 Germany License.

On 01.01.2015 it is planned to change the License Conditions (the removal of the Creative Commons License condition “no derivative works”). This is to allow reuse in the area of future scientific usage.

proteins that perform all of the functions needed for their existence⁴. There is no reason to believe that still simpler forms cannot be discovered yet or that they might not have existed earlier.

Gene duplication and crossing over are two phenomena, which have been described in detail, to account for many genetic experimental observations. Gene duplication can also explain how, during evolution, larger and more complex molecules could have formed rather simply⁵⁻⁷. There is even a suggestion that gene duplication might have been involved in the appearance of allosterism⁸.

Our own studies on amino acid homologies found among ancestrally unrelated proteins have led us to postulate that primordial protein-like molecules were formed, through hierarchical assembly, to larger and larger polypeptides, from a finite number of primordial peptides. We believe that the condensation between two peptides or more correctly which two peptides did condense with each other, was influenced primarily by the interaction among the terminal and penultimate amino acids and to a lesser degree by the interactions among internal amino acids of the two peptides and hence was to a certain extent nonrandom. The hierarchical assembly was also suggested by Simon based on theoretical considerations⁹. We also had stated earlier that it should be possible to discern what these primordial peptides were, from the homologies we observed and named "subsequences"¹⁰⁻¹². More recently we also found similar subsequences within t-RNA molecules, about 1/8th the length of the entire molecule, which we believe were the primordial oligonucleotides from which the t-RNA molecules finally evolved¹³. This last observation makes it likely that hierarchical assembly may have been the universal method for the formation of all macromolecules under primordial circumstances. Together with the results of the simulation experiments, obtained by many groups¹⁴⁻¹⁶ these observations suggest how the earliest cells might have been formed.

In this communication I shall attempt to describe how this might have occurred and will propose a model for the emergence of the first "successful cell".

Development of the Background for the Model

A living cell can be likened to a large manufacturing plant, an assembly line, highly integrated

and compartmentalized, which extracts energy from certain extraneous molecules, known collectively as "nutrients", and utilizes this energy for the production of cellular components needed for cell division, following what may be called "growth" of the cell. Energy is also used for the production, breakdown and modification of special molecules which control various activities essential for higher organisms.

The only difference or shortcoming of a machine, in this comparison, is its inability to reproduce itself, even though self reproducing machines — automata — are regularly treated in computer science. Indeed the reason why an automaton cannot physically reproduce itself, at the present, is due to an inability to pack all of the disparate manufacturing steps into a reasonably small volume and because of the enormous financial outlay required to realize this goal.

Regardless of this difficulty, a comparison between a cell and a self reproducing machine can help one to understand the structure and function of cellular components. Thus like a working assembly line which needs functional machines, cells must have, in order to live or survive, components which perform certain operations — movement of chromatids, contraction of muscles, etc. — as well as processes or reactions — oxidative phosphorylation, photosynthesis, etc. —. We know, today that these components are represented by biologically active proteins and they constitute the "hardware" of cells, borrowing a term from the computer science. Even though there are some proteins known as the "structural proteins", the fact remains that all proteins do contribute to the cellular structure and that the so called structural proteins, or certain portions of their structure, may not have some catalytic or other biological activity, at least some time during the life cycle of a cell, cannot be dismissed categorically^{17, 18}.

Unless the laws of chemistry have changed during the last four billion years, which is not very likely^{19, 20}, we are forced to conclude that proteins, more precisely biologically active protein-like molecules — proteinoids —, had to have been present in the primordial oceans for the first "living cell" to have evolved or to have been created, by an act of God. Which of these alternatives had been responsible for the emergence of the first cell is clearly irrelevant, in comparison with the *undeniable fact*

that protein had to have been there! There is indeed supporting evidence for this view which is found by going back, from the present, to two to three billion years ago and it presents itself as a striking similarity among proteins that perform the same task and that are known as a protein family. Chemically this relatedness is seen most clearly when proteins are aligned as seen in Atlas of Protein Sequence and Structure²¹. When amino acid sequences of such similar proteins, such as serine proteases, cytochromes c, etc. are compared, one finds a remarkable invariance of amino acids at certain positions along the chain. These positions, known as the critical regions, correspond usually to active/binding sites, etc. of biologically active proteins, and had to be preserved because the activity of these proteins depend on those amino acids being what they are. This observation tells us that when a molecule was arrived at, during evolution, which could successfully execute an operation, no matter how slowly and nonspecifically, then, that molecule was not altered. In evolutionary terms this means that "successful molecules were selected against others less capable of performing that same reaction". This may have occurred, for example, by the more active hydrolytic enzymes degrading the lesser active ones, giving rise to the phenomenon of the "survival of the fittest" to take place even at the sub- or pre-cellular level, as was suggested by Oparin²². Thus extrapolating toward the past, it is quite reasonable to assume that the primordial "protocells" had some of the same enzymatic activities that we have in more highly evolved cells/organisms, today. Naturally these proteinoids would have lower catalytic activity with very low selectivity; both the increase in selectivity and efficiency evolving gradually. The amino acid changes which are observed along the amino acid sequence of the members of a protein family represent, most probably, the consequences of evolution whereby new species were created^{23, 24}.

In summary, *it can unequivocally be stated that among the informational macromolecules proteins play the most significant role in a cell today, as they did under primordial conditions, for the first cell to emerge!*

A comparison of viruses, isolated nuclei and DNA with enucleated cells demonstrates this point: A virus is nothing but a nucleoprotein molecule which is utterly useless without a cell to act in. Erythro-

cytes and enucleated protozoa can function perfectly well for long periods, without their nuclei. On the other hand a nucleus removed from its cytoplasm as well as DNA isolated from a nucleus are bound to deteriorate very rapidly.

Most texts dealing with chemical evolution give three possibilities under which life could have evolved:

- 1) Life based on protein-like molecules having been the first and essential macromolecules,
- 2) Life based on nucleic acids having been the first macromolecules, and
- 3) life based on coevolution of proteins with nucleic acids.

So far, discussion presented above clearly favors the first alternative, by demonstrating that proteins had to be the first informational macromolecules present in the primordial oceans if a successful cell were to emerge at all.

This latter statement is supported by the results of many simulation experiments: that amino acids, when heated under hypohydrous conditions, give rise to proteinoids²⁵ with nonrandom sequence due to selective amino acid side chain interactions²⁶; that amino acids can yield smaller peptides also in aqueous solution²⁷; that proteinoid solutions, when left to stand, produce cell like structures, about 2 μ m in diameter, known as microspheres, which multiply by budding, cell division, which show bilaminar membrane like structure under electron microscope²⁸; that by repeated freezing and thawing microspheres could be converted to a state resembling coacervate droplets²⁹ and most importantly, that proteinoids were demonstrated to possess many enzymatic activities³⁰.

Because of the regularities found in their amino acid sequence — alternating hydrophilic and hydrophobic amino acids — some proteinoids might have been the forerunners of present day β sheets³¹. I propose that these proteinoids were the ones which formed the microspheres.

Where do nucleic acids, particularly DNA, fit in this picture? The primary function of cellular DNA is the preservation of organismic identity and storage of information needed for the production of the multitude of proteins and other macromolecules needed for the day to day activities of cells. Because of the need for biologically active protein-like molecules to have been present before the first cell emerged, as described above, DNA molecule ought

ta have evolved after the formation of these biologically active proteinoids. This statement is simply due to the fact that unless biologically active molecules were components of a "living cell" which had to reproduce itself, *they need not have been reproduced exactly* and hence there was no need to store that information!

One can also ask here, why is the DNA of a cell such a gigantic molecule, while all the proteins it codes for can fit into a much smaller volume? The answer lies in the fact that the greatest concern for a cell is to have an error-free replication mechanism. This was achieved by decreasing the number of elements used in the formation of DNA, during evolution, and as we know now, an optimum number for this purpose was reached in two complementary base pairs²³. Even this, however, cannot afford an absolutely error-free replication³² and a comparison of melting points or Cot data of normal DNA with poly(dA·dT) or poly A/dA·poly dT support this idea³³. Naturally a decrease in the number of building blocks occurred at the expense of information density, leading to one dimensional expression as opposed to the three dimensional expression of information by proteins²³. This resulted in enormously large molecules.

A corollary of this discussion is that DNA molecule could not have evolved before the formation of biologically active proteinoids confined to a "cell".

This does not mean, though, that DNA-like but much smaller and single stranded oligo(deoxy)ribonucleotides might not have been formed in the primordial oceans, even though simulation experiments make this event very unlikely, because so far no satisfactory way was found to account for the formation of nucleosides³⁴. What it really means is that any nucleic acid molecule, which was destined to become the genetic material — memory tape or software — for a living cell, had to form within the "protocells" in such a way that it possessed the information contained within the amino acid sequence of biologically active proteinoids. As will be demonstrated below, this could only have happened, if the nucleic acid molecules were synthesized on a proteinoid template. Even though this may sound like a heretical statement, any other nucleic acid molecule whose base sequence did not correspond to the amino acid sequence of a single biologically active proteinoid molecule would be an

utterly useless polymer. Since no oligonucleotide has been shown to possess any enzymatic activity, the internucleotide bond formation had to be under the influence of water scavenger molecules such as dicyanamide or HCN, but how could the replicated strands separate, without help from a preformed and highly specialized protein is a question begging for an answer. Furthermore, before one can even begin to think of internucleotide bond formation, one has to worry about the formation of nucleosides and phosphorylation of the nucleosides at the correct position — either 5' or 3' — consistently. The first oligonucleotides, however, were most probably formed from nucleosidemonophosphates and the emergence of higher phosphate precursors corresponds to a later stage in the evolution, when oxidation of ingested metabolites furnished the energy for cells and when this energy could be trapped internally (see the model, below).

Probability of a Nucleic Acid to Have the Correct Base Sequence to Yield a Biologically Active Protein

Is it possible for a single nucleic acid molecule, small enough to code for one biologically active protein, to form by chance, in such a way that its base sequence would correspond to the amino acid sequence of that particular protein?

The likelihood of this event can be calculated, very simply, for a 100 amino acid long protein, similar to cytochrome c.

The probability of a unique, biologically active, 100 amino acid long protein, to form from 20 amino acids, is:

$$p_P = 20^{-100} = 7.9 \times 10^{-131}.$$

The probability for a nucleic acid molecule, having a unique base sequence and 300 nucleotides long, to form from 4 different nucleotides, is:

$$p_{NA} = 4^{-300} = 2.4 \times 10^{-181}.$$

The probability for such a unique nucleic acid molecule to have a base sequence which could yield a unique and biologically active protein molecule, as the one mentioned above, is, because these two events are totally independent of each other, the product of the two probabilities calculated above:

$$\begin{aligned} p_{\text{overall}} &= p_P \times p_{NA} = 7.9 \times 10^{-131} \times 2.4 \times 10^{-181} \\ &= 1.9 \times 10^{-311}. \end{aligned}$$

It is obvious that it is highly improbable to have such a small nucleic acid molecule whose sequence

can yield one small, biologically active protein, by chance. It is still more highly improbable, more correctly impossible, to have a larger nucleic acid molecule which could have the correct base sequence to code for more than one biologically active protein, certainly needed for a living cell!

Thus, this discussion lays to rest the claims that nucleic acids were the first informational macromolecules involved in the emergence of "life" by eliminating erroneous arguments and underlining the critical issue here: The question is not whether or not nucleic acid molecules could have been formed first or whether they could replicate themselves without help from a preformed proteinoid. The critical issue is *the impossibility for any nucleic acid formed, by chance, to possess the correct base sequence corresponding to a single biologically active protein, unless it was formed on such a protein template.*

Some time ago a third alternative was proposed, as a reconciliatory idea, in this context: that proteins and nucleic acids could have evolved simultaneously^{35, 36}. Fox, too, has included this idea in his model for the emergence of protocells^{37, 38}. I am now going to demonstrate that this third alternative is nothing but a special case of the first alternative discussed above:

Coevolution could occur, if and only if, amino acids and nucleotides — one to one or one to three, respectively — could selectively interact with each other. I personally believe that not only is this interaction possible but that it actually took place during the formation of nucleic acids on proteinoid templates. More correctly, such an interaction must have occurred between preformed peptides and three nucleotides, as suggested by the model building studies of Lacey and Pruitt³⁹. Woese had also produced indirect experimental evidence for such interactions^{40, 41}. This again suggests that internucleotide bond formation, using di- and triphosphates of nucleosides, was a much later development during evolution. At the beginning mononucleotides could have been condensed, after they aligned themselves "correctly" along a peptide template, which was in α helical conformation. I wonder whether nucleoside formation itself might not have been facilitated by an alignment of bases and sugars against a peptide.

Since amino acids were shown to react preferentially, to yield nonrandom sequences²⁶, a reaction which was responsible for the formation of

biologically active proteinoids³⁰, without any help from nucleic acids, — while no such selective interaction is known to occur among the four nucleoside bases — the coevolution idea still requires some preferential interaction in order to yield biologically active proteinoids. And this interaction, indeed the only one we know, occurs among amino acids and is driven by the side chain interactions. Thus coevolution, whether or not it really occurred, is only a special case of the first alternative which states that proteinoids were responsible for the emergence of life.

The Model

The amino acids, formed under the impact of various forms of energy on a mixture of simple gases, were condensed either in solution by water scavengers such as HCN, dicyandiamide, etc., or under hypohydrous conditions on lava hot plates, to small polypeptides. Because of the preferential amino acid side chain interactions, there were only a finite number of these peptides produced¹¹. These peptides condensed with each other, in various combinations, this reaction also being influenced by the preferential side chain interactions of terminal and penultimate amino acids of the peptides, to yield a large number of still nonrandomly combined polypeptides. These larger polypeptides also interacted, under similar conditions, to give still larger polypeptides. Some of these larger polypeptides exhibited weak hydrolytic and other enzymatic activities. The hydrolytic peptides began to degrade other peptides that did not have as strong an activity themselves. Even though this degradation might have occurred very inefficiently, considering the long time available to them, it must have resulted in extensive destruction of many proteinoids and a selection for the active ones. The degradation products — smaller peptides — might have been used in the production of other proteinoids, in the same hierarchical way. Those peptides that possessed alternating hydrophilic and hydrophobic amino acids started to aggregate and form the microspheres. Sometimes enzymatically active peptides were trapped either on the outside of the microsphere surfaces or within the microspheres or both. The microspheres having proteolytic activity on the surface, could destroy others and utilize their contents for their own growth. Degradation

products of those microspheres thus became the "nutrients" of the attacking cells. Those enzymes which were trapped inside the microspheres helped the growth of those cells in a way similar to coacervate droplets⁴². It was already pointed out, before, that microspheres could be converted to a coacervate-like state by certain treatments that could have taken place under primordial conditions²⁹. Thus up to this stage one does not even have to have "life" for the growth and competition of the successful microspheres! Hence there was also no need to have a definite identity, either.

Over eons some microspheres did acquire peptides whose activities complemented each other so that a series of reactions could be executed. Because amino acid sequence alone controls the folding of a protein into its characteristic three dimensional structure and also because proteins which perform the same reaction have homology among themselves, it is quite clear that they would fold into similar three dimensional structures⁴³⁻⁴⁵. Since the active site of an enzyme which is responsible for its catalytic activity consists of a few amino acids, which themselves are part of the amino acid sequence of the protein, situated in a critical three dimensional geometry, proteins belonging to the same family should have the amino acids forming their active sites within peptides which are homologous among themselves^{21, 46}. Starting with a finite number of peptides which react preferentially among themselves, many proteinoids would form in which the order of some of the primordial peptides — subsequences — would be the same. In other words, in many proteinoids, a few subsequences would follow each other in a particular order, but not necessarily at the same position along the amino acid sequence of all proteinoids. Of these proteinoids, those that have similar overall folding, despite differences in their size, will show homology along certain portions of their sequences. If the homologous regions represent active site fragments of those proteins, they would perform the same reaction with differing facility. This is why, for instance, there is, in *E. coli*, a protein which cross-reacts with antiserum prepared against wild type β -galactosidase. The evolution of a new enzyme with β -galactosidase activity — Ebg, evolved beta galactosidase — in the laboratory, from a progenitor protein molecule, in a *E. coli* mutant with a deletion of β -galactosidase gene^{47, 48}, too, occurs because

the progenitor molecule is capable of binding lactose⁴⁹ and most probably has a similar three dimensional structure⁵⁰. Thus there are three known, and perhaps a few still to be discovered, proteins in *E. coli*, that can bind to lactose and have similar three dimensional folding. Of these, the one whose three dimensional structure was closest to wild type galactosidase was converted to Ebg.

Many of the small molecule were being degraded, within the microspheres, by active proteinoids, all along. In some cells, the energy released during this breakdown could be trapped by the formation of high energy bonds, because nucleoside monophosphates and phosphate ions or acetate and phosphate ions happened to be within close proximity of the degradation site. These high energy compounds could have been protected by binding to basic proteinoids. It may even have been possible to have anhydride or thioester type high energy bonds to have formed on amino acid side chains belonging to some proteinoids. This trapped energy could occasionally be released to drive some other reaction, inside the microspheres, again by some small molecules streaming close by the high energy bonds. The proteinoid protecting the high energy compounds could also conceivably be moving freely inside the microspheres, prior to the emergence of protective barriers — compartmentalization — and bringing the high energy bonds to a potential reaction site, where two molecules with potential to react were present. A basic proteinoid could also act as a matrix to help concentrate various acidic molecules which could later react. Small structural fluctuations occurring within the proteinoid molecule could bring the reactants and high energy compounds together. Again microspheres showing the ability to utilize the energy released during breakdown of other molecules have a selective advantage over others that cannot perform the same reaction. As the exogenous nutrient molecules began to be depleted by these competing sub-living cells, those microspheres which happened to have certain proteinoids on their "membranes" that showed some affinity toward one of those nutrients, began to evolve another selective advantage, over others which had to rely on diffusion alone. This modest beginning of transport phenomenon also meant that undesirable or harmful molecules had greater difficulty in entering such cells. I would suggest that the first

successful living cell did not appear until most of the exogenous nutrients were depleted.

Throughout this period the nucleic acid components — bases, sugars, phosphate ions — were continuously being taken up and released through simple diffusion. In some microspheres nucleosides were formed from the bases and sugars aligned along certain peptides and released. Some of the released nucleosides diffused outside the microspheres while others were phosphorylated within them and occasionally bound again to the peptides and were condensed to give short copies of the amino acid sequences of the peptides to which they were bound. The accuracy of translation was not very high and depended upon the degree of selectivity of the amino acid-nucleotide interactions. Incidentally, a comparison of amino acid side chain structural characteristics with code words yields some interesting rules (Erhan and Rasco, unpublished observations).

Some of the oligonucleotides having the information of amino acid sequence of proteinoids might have condensed in various combinations after being released from the proteinoids, giving rise to various recombinants, some of which possessing greater activity or specificity than the original proteinoids. Some of the peptides might have condensed to larger peptides possessing two enzymatic activity. Nucleic acids formed from them would be "multicistronic". Earliest nucleic acids were most probably single stranded ribonucleic acids, protected by association with basic proteinoids. The evolution of redundancy of nucleic acid sequences and double strandedness to increase stability and to protect against errors occurred after the first cells came about, probably in a similar way described about MDV-1 phage RNA⁵¹, the last step being the substitution of deoxyribose as an added precautionary measure. The presence of nucleic acids within the microspheres so necessitated a greater compartmentalization to protect them.

The formation of biologically active proteins in these early cells — early translation — could occur through the reversal of the direct interaction through which the early nucleic acids were produced. During later stages certain oligonucleotides which were formed within the microspheres, and which may even have had anticodon regions at one end, were utilized as mediators — adaptors — to carry amino acids to their sites of synthesis, because this would increase the accuracy of alignment¹³. Binding of the amino acids to these oligonucleotides still probably occurred through noncovalent interactions. The final size of the t-RNA molecules was predicated by the emergence of ribosomes. Translation phenomenon utilizing messenger molecules appeared only after DNA was fully evolved as the depository of information and had to be protected from increased chances of error production during the protein synthesis, which occurred more frequently than replication.

I think that the first really successful cells were the ones which were able to utilize the energy obtained from the nutrient molecules to drive other reactions on a regular basis, to obtain certain nutrients from the outside and to produce the needed building blocks internally.

Thus I favor the view that there were many partially successful early cells performing the necessary activities and competing with each other with varying degrees of efficiency. There were many alternatives evolving, some through the mechanisms suggested above and some through others, including possibly mutations occurring on the cellular nucleic acids. All of these cells were the products of incomplete endeavors. The one cell which could perform all of these vital operations with the least degree of error and greatest efficiency had such a great selective advantage that it easily outgrew all of the lesser successful cells with its offsprings and became our primordial ancestor.

¹ P. A. Weiss, *The Science of Life*, p. 73, Futura Publishing Co. New York, New York 1973.

² F. B. Salisbury, *Nature* **229**, 342 [1969].

³ C. B. Anfinsen, E. Haber, M. Sela, and F. H. White, Jr., *Proc. Nat. Acad. Sci. U.S.A.* **47**, 1309 [1961].

⁴ H. J. Morowitz, *Hosp. Prac.* p. 95, July [1975].

⁵ S. Ohno, *Evolution by Gene Duplication*, p. 160, Springer Verlag, New York 1970.

⁶ L. Hood, K. Eichmann, H. Lackland, R. M. Krause, and J. J. Ohms, *Nature* **228**, 1240 [1970].

⁷ J. Barnabas, M. Goodman, and G. W. Moore, *Comp. Biochem. Physiol.* **39B**, 455 [1971].

⁸ P. C. Engel, *Nature* **241**, 118 [1973].

⁹ H. A. Simon, in *Hierarchy Theory* (H. H. Pattee, ed.), p. 3, G. Braziller, New York 1973.

¹⁰ L. D. Greller and S. Erhan, *Int. J. Pept. Prot. Res.* **6**, 165 [1974].

¹¹ S. Erhan and L. D. Greller, *Int. J. Pept. Prot. Res.* **6**, 174 [1974].

- ¹² S. Erhan and L. D. Greller, *Int. J. Pept. Prot. Res.* **9**, 5 [1977].
- ¹³ S. Erhan, L. D. Greller, and B. Rasco, *Z. Naturforsch.*, in the press [1977].
- ¹⁴ S. W. Fox, *Nature* **205**, 328 [1965].
- ¹⁵ G. Hennon, R. Plaquet, M. Dautrevaux, and G. Biserte, *Biochimie* **53**, 215 [1971].
- ¹⁶ J. Kovacs, H. N. Kovacs, I. Konyves, J. Czarzar, J. Vajda, and H. Mix, *J. Org. Chem.* **26**, 1084 [1961].
- ¹⁷ S. Erhan and L. D. Greller, Abstract submitted to Int. Hemat. Soc. 4. Meeting, Istanbul, Turkey 1977.
- ¹⁸ R. L. Katzman, A. H. Kang, and E. H. Beachey, *Science* **181**, 670 [1973].
- ¹⁹ M. G. Rutten, *Geological Aspects of the Origins of Life on Earth*, Elsevier Publishing Co., Amsterdam, Holland 1962.
- ²⁰ K. Rankana, *Progress in Isotope Geology*, Interscience Publishers New York, New York 1962.
- ²¹ M. O. Dayhoff, *Atlas of Protein Sequence and Structure*, **Vol. 5**, 1972.
- ²² A. I. Oparin, *The Beginning and the Origin of Life on Earth*, Medizine, Moscow, USSR. 1966.
- ²³ S. Erhan, *J. Theor. Biol.* **63**, 111 [1976].
- ²⁴ T. R. Marzolf, L. D. Greller, and S. Erhan, *Int. J. Biomed. Comp.*, Accepted for publication 1977.
- ²⁵ S. W. Fox and K. Harada, *J. Am. Chem. Soc.* **82**, 3745 [1960].
- ²⁶ S. W. Fox and T. Nakashima, *Biochim. Biophys.* **140**, 155 [1967].
- ²⁷ G. Steinman, *Arch. Biochem. Biophys.* **119**, 76 [1967].
- ²⁸ S. W. Fox, P. J. McCauley, and A. Wood, *Comp. Biochem. Physiol.* **20**, 773 [1967].
- ²⁹ A. E. Smith and F. T. Bellware, *Science* **152**, 362 [1966].
- ³⁰ D. L. Rohlfsing and S. W. Fox, *Adv. Catalysis* **20**, 369 [1969].
- ³¹ A. Brack and L. E. Orgel, *Nature* **256**, 383 [1975].
- ³² T. A. Trautner, M. Swartz, and A. Kornberg, *Proc. Nat. Acad. Sci. U.S.* **48**, 449 [1962].
- ³³ R. J. Britten and D. E. Kohn, *Science* **161**, 529 [1968].
- ³⁴ S. L. Miller and L. E. Orgel, *The Origins of Life on Earth*, p. 112, Prentice-Hall, Englewood Cliffs, N. J. 1974.
- ³⁵ *ibid.* p. 161.
- ³⁶ J. T. Wong, *Proc. Nat. Acad. Sci. U.S.* **72**, 1909 [1975].
- ³⁷ S. W. Fox, *Mol. Cell. Biochem.* **3**, 129 [1974].
- ³⁸ S. W. Fox, J. R. Jungck, and T. Nakashima, *Orig. Life* **5**, 227 [1974].
- ³⁹ J. C. Lacey and K. M. Pruitt, *Nature* **223**, 799 [1969].
- ⁴⁰ C. Woese, *Proc. Nat. Acad. Sci. U.S.* **54**, 546 [1965].
- ⁴¹ C. Woese, Personal Communication (1975).
- ⁴² A. I. Oparin, *Adv. Enzymol.* **27**, 347 [1965].
- ⁴³ R. M. Stroud, L. M. Kay, and R. E. Dickerson, *Cold Spring Harbor Symposium*, **Vol. 36**, p. 125 [1971].
- ⁴⁴ J. J. Birkhoff and D. M. Blow, *J. Mol. Biol.* **68**, 187 [1972].
- ⁴⁵ A. D. McLachlan, *J. Mol. Biol.* **61**, 409 [1971].
- ⁴⁶ S. Erhan and L. D. Greller, *Int. J. Biomed. Comp.*, in the Press (1977).
- ⁴⁷ J. H. Campbell, J. A. Lengyel, and J. Langridge, *Proc. Nat. Acad. Sci. U.S.* **70**, 1841 [1973].
- ⁴⁸ D. L. Hartl and G. B. Hall, *Nature* **248**, 152 [1974].
- ⁴⁹ J. Arraj, Ph. D. Thesis, University of California [1973].
- ⁵⁰ S. Erhan, Submitted for publication 1977.
- ⁵¹ D. R. Mills, F. R. Kramer, and S. Spiegelman, *Science* **180**, 916 [1973].