Self-Assembly of the Protocell from a Self-Ordered Polymer

Paper presented to
International Convention of Biochemists,
Bangalore, India
7 September 1967

GPO PRICE $____________
CFSTI PRICE(S) $____________

Hard copy (HC) __
Microfiche (MF) __

Since 1960, this research has been aided by the National Aeronautics and Space Administration, currently Grant no. NsG-689. Contribution no. 096 of the Institute of Molecular Evolution.
Sidney W. Fox

Institute of Molecular Evolution
University of Miami
Coral Gables, Florida, U.S.A.
The problem of the origin of life, or in truly perceptive nineteenth century terms, the problem of spontaneous generation, has often been regarded as one of overwhelming complexity. Upon analysis, with the aid of hindsight, this problem loses some of its imponderability. The aspect of evolution which first received major attention was that of the progression, in principle, from primitive cell to contemporary cell and to contemporary multicellular organisms. This stage is the one that has been illuminated mechanistically by Darwin's theory of selection. We can now regard this stage as far more intricate and involved than the emergence of primitive life from the primordial reactant gases. By such an analysis, the primordial cell is emphasized, the highly ramified later stages are removed from purview, and the limits of the meaningful problem are identified.

The preorganismic stage can also be analyzed. For intellectual convenience, it may be divided into two or three parts. The first of these parts is that of the spontaneous organic synthesis involved in the production of the small organic molecules which are necessary for contemporary and, presumably for, primitive organisms. The second step is the spontaneous synthesis of the polymers and of cells. This latter constitutes in turn, however, two stages. These two steps were collectively most forbidding in quality, and are particularly significant to life and therefore, to its origin. Our most modern knowledge requires that we recognize that a primitive cell can not be a synthesized entity in the true meaning of "synthesized". The precursor macromolecule can be conceived of as synthetic. When the appropriate macromolecule has been formed, the final and crucial stage, leading to a primitive organism, would then be one of self-assembly. The term "self assembly" and the concept have recently been receiving increasing recognition e.g. in the biochemistry of contemporary systems.

One way in which students of the total problem have dealt with the seemingly great complexity has been to postulate a long chemical evolution extending over, say, 25 million years. I will explain here why our experiments lead to the interpretation that the essential steps from primordial gases, amino acids, primitive protein, a primitive organized structure having simultaneously many lifelike properties including the ability to participate in its reproduction could have occurred many times in a very short period, say 25 hours. My immediate problem is to present the salient experimental material in 25 minutes. This problem exists because of the careful devotion to it by many associates during 14 years of continuous study in our laboratory. Accordingly, I shall rely heavily on summaries and upon examples from our laboratory and others.
Our approach to this problem was based on clues from contemporary cells. The results of experiments have been evaluated in part by how well they lead to an increasing appearance of the properties that are associated with contemporary cells. The experiments have however been based on very simply derived initial systems and simple processes. These employ conditions that have proved to be plausible not only for geologically ancient times; the conditions identified are widespread now and through recorded history.

Models of the prebiotic synthesis of small organic compounds such as monosaccharides, amino acids, purine and pyrimidine bases, ATP, porphyrins, etc. have been described from many laboratories including those of Ponnamperuma, Orgel, and our own. Since the essence of life is generally recognized as being that of the biopolymers, protein and nucleic acid, this paper will focus on questions involving the primordial formation of protein and nucleic acid and on the attributes of the polymers formed in the laboratory. It will deal also with complexes of the two.

Turning first to the question of proteins, we find that a number of studies of the synthesis of peptide bonds, mostly in aqueous solution, have been carried out in a number of laboratories. Akabori employed the progressive substitution of polyglycine as a model of the first protein, and Matthews has reported a similar process. The model of our laboratory which relies on heat and hypohydrous conditions is the only one that has yielded polymers of molecular weight in the thousands, a content of all eighteen amino acids common to protein, several protoenzymic activities, and it is the only model which has been demonstrated to yield organized structures with a lengthy roster of the properties of the contemporary cell. This synthesis has the simplicity appropriate to geologically spontaneous occurrences, and it yields both polymer and organized units in abundance.

One other synthesis, which has most of the attributes enumerated, uses as intermediates the reactive Leuchs anhydrides of the amino acids. This synthesis was also first performed in our laboratory, by Dr. T. Hayakawa. Of these two syntheses, only the thermal process has the simplicity appropriate to geologically spontaneous occurrences.

The thermal syntheses, first attempted in 1953, were indicated as thermodynamically possible following studies of Borsook, Huffman, Ellis, and Fox. The results of calculations from the tabulated physical constants have shown that one could expect in an open aqueous solution only small yields of small peptides unless the reaction were somehow coupled to an endergonic one.
The fact that organisms are nearly always aqueous entities has led some to assume that hot, dry conditions would not have been appropriate to early life and they have somehow projected such thinking to precursor molecules. Our chemical experience however tells us that macromolecules can easily survive conditions lethal for ordinary cells, and our biological experience reminds us that bacterial spores are relatively resistant to heat and dryness.

The reaction involving formation of peptide bond with its attendant Gibbs free energy change is:

\[
\begin{align*}
H_2NCHRCOOH + H_2NCHR'COOH &= H_2NCHRCONHCHR'COOH + H_2O \\
\Delta G^0 &= 2000 \text{ to } 4000 \text{ cal.}
\end{align*}
\]

As the number of peptide bonds per molecule increases, the equilibrium constant becomes geometrically more unfavorable. Dixon and Webb\(^{13}\) have calculated that the volume of 1 M amino acid solution in equilibrium with one molecule of protein of molecular weight 12,000 would be \(10^{50}\) times that of the Earth! Stated otherwise, uncoupled synthesis from amino acids in water should be expected to give small yields of small peptides only.

In order to shift this equilibrium to favor synthesis, one can postulate removal of either product. Theoretically, one contribution of a membrane in contemporary protein-synthesizing systems may be the overcoming of an energy barrier by separation of synthesized macromolecules from the aqueous solution. This process could not apply, however, until a membrane composed of macromolecules had first formed. Our attention therefore shifts to removal of the other product, water. This route to peptide bond synthesis can be visualized as a geochemical possibility. It has also been demonstrated experimentally\(^{7}\).

One mode of removal of water, as thus suggested by the thermodynamic analysis, would be that of heating the amino acids above the boiling point of water. When this possibility was initially contemplated, the probability of gross decomposition had to be considered. Such a consequence of heating \(\alpha\)-amino acids above the boiling point of water has been recorded in the literature a number of times and was also common knowledge. We were led to attempt the thermal condensation by employing an inference from comparative studies of organismic protein, the fact that the amino acids which most dominate the composition of proteins are glutamic acid and aspartic acid\(^{7}\). These contents were taken hypothetically to be an evolutionary reflection of a circumstance required for the primordial formation of prebiotic protein.

Another consideration that had to be dealt with was the somewhat vague feeling that, without nucleic acids present, the necessary systematic sequences of amino acid residues would not
result. This problem was conceptually eliminated, in principle, by studies of enzymic acylpeptideanilide synthesis which demonstrated that interactions of amino acid residues would alone select the sequence formed. (This principle and the inference that prior nucleic acids may have been unnecessary has since been corroborated by Steinman in another system of reacting amino acids.) Since the difficulties were thus conceptually surmountable in 1953, heating was employed (Fig. 1). The discussion now deals with experimental observations.

A typical thermal condensation used at first a mixture of 1 part of aspartic acid, 1 of glutamic acid, and 1 of an equimolar mixture of the 16 other amino acids common to protein. This mixture was heated at 170° for 6 hours. The resulting light amber glassy product, not depicted, is entirely soluble in water by salting-in, and can be purified by salting-out. Such products yield amino acids 100% by acid hydrolysis, they contain some proportion of each of the amino acids common to protein (or fewer as desired), and molecular weights of many thousands. They have many other properties of protein and are called proteinoids. The proteinoid described in this example is, because of the proportions reacted, a 1:1:1 type. More recently, Dr. Waehneldt has shown in our laboratory that aspartic acid and glutamic acid may be merely equimolar with the 16 other amino acids. Proteinoids are produced even so. Yields in the usual syntheses are typically in the range of 10–40%, higher yields being obtained by the addition of phosphates. Many other laboratories have repeated this synthesis and have confirmed it and its simplicity, which is in turn crucial to the geological validity. The spontaneous occurrence of a carbobenzyoxy synthesis or the formation and condensation of N-carboxy amino acid anhydrides cannot, of course, be defensibly imputed to the geological environment.

The numerous properties which the proteinoids have in common with proteins are described in detail in the literature, and these have been reviewed a number of times. Table I lists the many simultaneous structural, chemical, and biological properties all of which are described in literature cited bibliographically, except for recently demonstrated hormonal activity and a few other aspects. Time will be devoted here to reviewing only three salient properties on which many new data are available: limited heterogeneity and the related question of systematic sequences, 2nd, catalytic activity, and 3rd the property of forming regular and highly structured particles.

The first indication that a thermal condensate of eighteen amino acids yields no more than two electrophoretic individuals was from a study of Dr. Carl Vestling. Subsequently Dr. Harada in our laboratory showed that two fractionations of a 2:2:3-proteinoid from hot water resulted in virtually no change
Table I

Properties of Thermally Prepared Proteinoids

Limited heterogeneity
Quantitative composition identical to that of protein
Quantitative compositions resembling those of proteins
Quantitative recoverability of amino acids upon hydrolysis
Range of molecular weights like those of smaller protein molecules
Positive color tests as for protein
Solubilities resembling those of protein classes (albumins, globulins, etc.)
Some optical activity
Tendency to be salted-in
Tendency to be salted-out
Precipitability by protein group reagents (phosphotungstic acid, etc.)
Hypochromicity
Infrared absorption maxima as found in protein
Some susceptibility to proteases
Nonrandom distribution of amino acid residues
Many catalytic activities
Inactivatability of catalytic power
Nutritive quality
Melanophore stimulating activity
Morphogenicity
in amino acid composition. Also, N-terminal or C-terminal analyses of many polymers demonstrated marked disparities from the total compositions, indicating that systematic non-random sequentialization occurred. This could be due only to selective interactions of amino acids during thermal condensation. Since that finding, many new data obtained by Dr. Nakashima and a review of all of the data have been published in Biochimica et Biophysica Acta.

The elution pattern from fractionation on DEAE-cellulose is shown in Fig. 2.

While no random assortment of polyanhydro-\(\alpha\)-amino acids has been prepared such that a comparison would be possible, we would expect theoretically an elevated nearly horizontal line for an elution pattern of a disordered polymer. What is repeatedly found, instead, is a pattern of six major peaks, of which some are already symmetrical.

For comparison is presented an elution pattern of turtle serum protein also fractionated on DEAE-cellulose. Eight major peaks, with less spread in each peak, are observed. The individual peaks of the eluate sediment in the Spinco Model E (Fig. 3) to indicate for the fractions the degree of ultracentrifugal homogeneity which is observed.

Fig. 4 shows the complete and partial hydrolyzates of three fractions from DEAE-cellulose. The bottom three are partial hydrolyzates in each group. The patterns are highly similar, as can be seen.

A total picture of nonrandom sequences in the linear distribution, discrete macromolecular fractions, and relatively uniform composition and sequences throughout the entire polymer is supported also by high voltage electrophoresis, by separation on columns of Sephadex, by polyacrylamide fractionations, and by gel electrophoresis.

Reports of catalytic activity in proteinoids are summarized in Table II. These findings are from six laboratories and they include catalysis of the hydrolysis of esters and of ATP and decarboxylations of a number of natural substrates. Krampitz has recently recorded an example of transamination. Michaelis-Menten kinetics have been reported in several of the studies. The activities are mostly weak. As Calvin pointed out for the iron-containing enzymes and as I pointed out in 1953, weak primitive enzymic activity would be selected and enriched by Darwinian processes in organisms. Of relevance, also, is the fact that some gross specificities have been identified and that individual proteinoid preparations each have an array of catalytic activities. A metabolic sequence has been recorded, for example, for oxaloacetic acid + pyruvic acid + acetic acid + CO\(_2\).
Table II

Catalytic Activities in Proteinoids

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Salient Finding</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-Nitrophenyl Acetate</td>
<td>Activity&lt;sub&gt;p&lt;sub&gt;td&lt;/sub&gt;=Activity&lt;sub&gt;hsd&lt;/sub&gt;</td>
<td>(26)</td>
</tr>
<tr>
<td>p-Nitrophenyl Acetate</td>
<td>Activities</td>
<td>(27)</td>
</tr>
<tr>
<td>p-Nitrophenyl Acetate</td>
<td>Inhibition by organic phosphates</td>
<td>(28)</td>
</tr>
<tr>
<td>p-Nitrophenyl Acetate</td>
<td>Detailed treatment</td>
<td>(29)</td>
</tr>
<tr>
<td>p-Nitrophenyl Acetate</td>
<td>&quot;Active site&quot;, and inactivation</td>
<td>(30)</td>
</tr>
<tr>
<td>Glucose + glucuronic acid + CO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>First natural substrate reported</td>
<td>(31)</td>
</tr>
<tr>
<td>ATP + ADP, et al.</td>
<td>Biochemical energy source</td>
<td>(32,33)</td>
</tr>
<tr>
<td>p-Nitrophenyl phosphate</td>
<td>Second phosphate hydrolysis</td>
<td>(34)</td>
</tr>
<tr>
<td>Pyruvic acid + acetic acid + CO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>Decarboxylation</td>
<td>(35,33)</td>
</tr>
<tr>
<td>Oxaloacetic acid + pyruvic acid + CO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>Catalyzed by ptlds of type not active on pyruvic acid</td>
<td>(36)</td>
</tr>
<tr>
<td>α-Ketoglutaric acid + urea + glutamic acid</td>
<td>Proteinoid and Cu each needed</td>
<td>(37)</td>
</tr>
</tbody>
</table>
We can explain the evolutionary potential of polyanhydro-amino acid catalysts by the fact that in the same macromolecule are found not only a variety of chemically functional groups, but also the products of interaction of the fields of force of two or more of these groups. This picture of chemical polyfunctionality is also of course applicable to proteins and emphasizes what A. E. Needham has referred to as "the uniqueness of biological materials". Proteinoid is in one sense perhaps even more unique than any one protein in that it is less specialized, and recalls thereby for serious consideration the concept of "urprotein".

The property of forming structurally organized units on contact with water is crucial to a comprehensive theory of the origin of the first cell. This tendency is intrinsic to many thermal polymers of amino acids, as we reported in 1960. The need for a macromolecular precursor of the first cell has been emphasized in theoretical discussions by Oparin by Wald by Lederberg, and by others. The degree to which thermal proteinoid meets this need by providing organized microscopic units having numerous associated properties such as are found in contemporary protein and in contemporary cells could not have been predicted. Only some of the more salient properties will be presented here. Others are documented with references to supporting literature. The properties include, among others, a cellular type of ultrastructure, double layers, abilities to metabolize, to grow in size, to proliferate, to undergo selection, to bind polynucleotides, and to retain some macromolecules selectively. The structures of Fig. 5 are usually produced merely by heating the proteinoid with water or aqueous solution. The clear liquid, on cooling, deposits huge numbers of microscopic units, 0.5-80 μ in diameter, of quite uniform size in any one preparation. These are usually found as spherules, as in this photomicrograph, but they occur also as filaments, budded microspheres, as twinned units, and in other shapes. They are exceedingly numerous; one gram of heated amino acids can produce many billion spherules. As pointed out by us and by others, they have physical stability comparable to that of contemporary cells; they can, for instance, be sectioned for electron microscopy. In their uniform size and in other respects they are readily distinguished from oily droplets or from the usual coacervate droplets. While they can be produced as entirely separate units, they also tend to associate. By adjustment of the basic amino acid content, they can be produced to stain either Gram-negative or Gram-positive. They also have been shown to have some of the catalytic activities which have been carefully identified in the polymer of which they are composed.

The ultrastructure of the microsphere is shown in the electron micrograph of Fig. 6. On the left is a section of Bacillus cereus which has been fixed by osmium tetroxide, sectioned, and electron micrographed, by Murray. On the right
is seen a section of a proteinoid microsphere fixed with osmium tetroxide and sectioned. While some bacteria reveal a more organized pattern than Bacillus cereus, this micrograph reveals only a boundary and granular cytoplasm. The granular appearance is found also in the proteinoid microsphere, and the latter has a more definite boundary. In one place the boundary appears to be a double layer. In the experiment preceding the electron micrography of Fig. 7, the polymer in the interior was first caused to diffuse out through the boundary by raising the pH in the suspension of proteinoid microspheres by one to two units. Double layers are clearly evident. We can also see part of the results of other phenomena. The diffusion depicted is one manifestation of selective behavior in the boundary or membrane. This diffusion is further illustrated in Fig. 8, in which the effect was followed by photographing in ultraviolet light through the quartz optics of Dr. Philip Montgomery's microscope at the University of Texas.

The proteinoid, like protein, absorbs at ultraviolet wavelengths. These pictures and related data indicate that polymer is not condensing on the membrane but is passing selectively through a membrane composed of very similar polymer, as has been shown by analysis.

Models of primitive polynucleotides have also been examined. With Dr. Waehneldt we have reported how nucleoside mono- and tri-phosphates could have been prebiotically synthesized in quantity, following an earlier report on the production of ATP by Ponnamperuma. Thermal polymers of mononucleotides have been shown, by Schwartz, in our laboratory, to be attacked by ribonuclease and by venom phosphodiesterase. Recently, we reported with Dr. Waehneldt and others that such polynucleotides, as well as calf thymus DNA and yeast RNA, bind with appropriate proteinoids to yield models of the various nucleoprotein particulates found in the contemporary cell, such as ribosomes, chromosomes, chromatin, etc. (Fig. 9).

In Fig. 10 is provided an example which shows fibers produced from lysine-rich proteinoid and calf thymus DNA. Very small microspheres result when RNA or thermal polyribonucleotides are used instead of DNA. The ratio of polynucleotide to basic proteinoid in such complexes tends to be quite constant. The fibrous and spherical morphologies are reminiscent of contemporary analogs as in chromosomes and ribosomes respectively. Those proteinoids that bind to form such particles have a ratio of basic amino acid to dicarboxylic amino acid above a minimum of 1.0.

With polyphosphoric acid and temperatures of 60-100°, experiments in our laboratory have shown that either mononucleotides or amino acids might be polymerized. Accordingly, these processes might ordinarily have occurred simultaneously. The suggestion of Calvin that proteins and nucleic acids might
have arisen simultaneously in a primordial event is thus consistent with the experimental demonstrations. Although model studies of prebiotic polynucleotides have been pursued, as indicated, a basic question that persists is that of how many properties models of primitive protein systems might display without polynucleotides. This question especially deserves to be asked in view of the fact that proteinoids contain their own information and have sharply limited variation without any control by nucleic acids. Also, the self-assembling properties of proteinoid yield ultrastructure, double layers, fission, etc. without nucleic acid control. Can then, for example, proteinoid microspheres multiply without polynucleotides present? In Figs. 11-14 we observe how in a very simple manner proteinoid microspheres do, in fact, participate in the reproduction of their own likeness.

In the first photomicrograph are seen a number of proteinoid microspheres which have been allowed to stand in their mother liquor for two weeks. On these are found "buds" which in appearance resemble buds on yeast. We first observed such buds in 1959. These "buds" can grow in size either while attached to the parent microsphere, or after separation. Removal can be accomplished at various stages of growth in size by shock - electrically, thermally, or mechanically. In the first two modes, we believe that some interfacial material is dissolved. "Buds" are seen in the second picture. In order to demonstrate the next event rigorously, the separated buds were stained with Crystal Violet. When the stained separated buds are allowed to stand in a solution of proteinoid saturated at 37°, and this is allowed to cool to 25°, the buds, in an appropriately sized vessel, "grow" by accretion to the size shown within one hour. The resultant units tend to be very uniform in size; the opposing forces are evidently precisely balanced under any given set of conditions.

In this manner we can visualize how proteinoid microspheres could simply first have developed the ability to participate cyclically in the reproduction of their own likeness. In manifesting such a process, a primitive organized structure would be functioning as a nearly complete heterotroph in that it obtained its large molecules by feeding on the environment instead of synthesizing them itself. Many theorists on the subject of abiogenesis have reasoned that the first organisms must have been heterotrophs. One may find the arguments in the writings of Oparin, Haldane, Pirie, Pringle, Horowitz, Van Nieuwenhuis and others. To repeat, the hypothetical need for nucleic acid-mediated constraints on primitive protein are seen not to apply because of the internal constraints on the primary structure of proteinoid.

Experiments in our laboratory have indicated how multiplication could occur also through the model of a primitive kind of binary fission and growth by accretion, as well as through budding.
With cyclical proliferation due to budding or fission, as depicted, Darwinian selection can, conceptually, occur. Increasingly incisive experiments on Darwinian selection from non-biological precursors have in fact been performed with acid proteinoid and neutral proteinoid.

The condensed description of this paper (in conjunction with other reports) explains how a primitive organism capable of a kind of self-multiplication and possessing other salient properties could have emerged from primitive gases through the amino acids and subsequently through protein-like polymer. As perhaps need not be re-emphasized, this model of a kind of primitive unit is clearly not a contemporary cell, at least not of the usual contemporary organism.

From this model, however, we visualize that in evolving to a contemporary organism a primitive self-replicating heterotroph would especially have had to develop an internal synthesis of protein and of polynucleotide.

The model processes which have been described are extremely simple. They consist of a) heating above the boiling point of water, and b) the intrusion of water. This simple sequence requires a) geologically anhydriding temperatures, e.g., those above 100° and b) sporadic rain or other common geological events of water such as drought or recession of the seas. These conditions a) and b) have been widespread geologically and, in fact, are quite widespread on the Earth today. The reactions are rugged, their occurrence is not easily disturbed by added substances since they are not solutes in aqueous solution, and the products have long-term stability. All details in the processes are found to be sequentially compatible.

To summarize, I use a 1966 quotation from Lederberg. "The point of faith is: make the polypeptide sequences at the right time in the right amounts and the organization will take care of itself. This is not far from suggesting that a cell will crystallize itself out of the soup when the right components are present." The results described here (and others) show that when amino acids are simply and suitably heated, polypeptide sequences to at least some degree make themselves, and they do this in "the right amounts." The organization of the polypeptides does indeed take care of itself when water is added to thermal proteinoid, which appears to be the right component to crystallize out as a cell. Individual properties of proteins, polynucleotides, and of cells can be mimicked by other substances and units. Thermal proteinoid and its organized particles are the only chemically synthetic products, however, which have been shown to possess in simultaneous association and in life-like inclusiveness properties of the contemporary cell and its structural polymer, with indications of the potential for further evolution to contemporary cells.
Figure 1. On left, tube containing mixture of amino acids heated to above the boiling point of water. On right, granular polymer prepared by heating a mixture of amino acids containing sufficient proportions of aspartic acid and glutamic acid.
Figure 2. Distribution of 1:1:1-proteinoid on elution from a DEAE-cellulose column by tris-HCl buffer.
Figure 3. Distribution of turtle serum proteins on elution from a DEAE-cellulose column by sodium phosphate buffer.
Figure 4. Top 3 profiles of hydrolyzates of 3 fractions of proteinoidamide show great similarity in amino acid content. Bottom 3 profiles show similarity in peptides obtained on partial hydrolysis.
Figure 5. Proteinoid microspheres.
Figure 6. Electron micrographs of *Bacillus cereus* and of proteinoid microsphere (right). Each has been fixed with osmium tetroxide and sectioned.
Figure 7. Proteinoid microsphere subjected to raised pH. Double layers are evident.
Figure 8. Photograph of pH effect followed by photograph in ultraviolet light through quartz optics.
Figure 9. Microspheres composed of RNA and lysine-rich proteinoid.
Figure 10. Photograph of fibrous complex of calf thymus DNA and of lysine-rich proteinoid.
Figure 11. "Budded" proteinoid microsphere.

A. microsphere
B. "bud"
Figure 12. Separated "buds".
Figure 13. Accretion of proteinoid particle around Crystal Violet-stained "buds"

A. proteinoid particle
B. stained center
Figure 14. "Daughter bud" on microsphere.

A. stained center
B. microsphere by accretion
C. "daughter bud"


14. Fox, S. W., Amer. Scientist, 44 (1956), 347.


23. Fox, S. W., Nature, 00 (1968), 000.


25. Fox, S. W., Amer. Naturalist, 87 (1953), 253.


