THE ORIGIN OF LIFE
The Emergence of Organized
Self-Replicating Molecular Systems

by

Bulent Atalay
Professor of Physics
Mary Washington College
Fredericksburg, Virginia   USA

"Life is an interaction between nucleic acids
and proteins, that is biochemically speaking..."

Cyril Ponnamperuma (1981)

PROLOGUE

Humankind has long been preoccupied with the origin of life.
Every religion known bespeaks of a time before life existed. Each
has its own set of myths. Soil, in the form of dust, mud or clay
appears to be a virtually ubiquitous building ingredient. In the
Bible it was the 'dust of the ground' that formed man. One novel
alternative was offered by the Mayan epic tome in which the Gods
had initially created mankind from clay, but with rain their
creations disintegrated. The Gods tried again, forming their
creations out of straw, only to see them perish in a spontaneous
fire. Success was finally achieved when maize was used as the
major ingredient for modeling living creatures.

Aristotle in metaphysics had explained that fireflies had
issued from morning dew, a theory which found acceptance later by
medieval scholastics, and later still by a number of Seventeenth
Century scientists, including Isaac Newton, René Descartes,
William Harvey and Johannes Baptista von Helmont. In fact,
von Helmont had added his own recipe: A mixture of grain and dirty
laundry, incubated for twenty-one days, would result in spontaneous
generation of life. He reported soberly, "Mice.... neither weanlings,
nor sucklings, nor premature.... jump out fully formed. The twenty-
one day period, however, was critical. Yet a thousand years
earlier, the Greek philosopher Anaximander had shown remarkable

prescience in suggesting that human beings had emerged from the sea, as had the Indian sages who had recorded similar beliefs in the Rig-Vega and the Atar-Vega. All these, however, were born of mysticism no less than the claims of Aristotle or von Helmont.

Serious scientific questions concerning the origin of life date to the time of Darwin and Wallace in the last century and are epitomized in the notion that each living species is traceable to its immediate predecessor, that life comes from life. In this sense, the necessary regression should take all living species back to a common ancestor. But then one is confronted by the ultimate question of the "first living" (organism) coming from the nonliving. This question certainly gave substantial pause to a scientific community already reeling from the Darwinian picture, perhaps because of the perception of the gap which exists between the living and nonliving. And in the late Nineteenth Century, Pasteur demonstrated the impossibility of the growth of micro-organisms in a sterile medium, and thereby disproved the theory of spontaneous generation. He announced his discovery to the French Academy of Sciences in 1861, saying "Never will the doctrine of spontaneous generation recover from this mortal blow." But by "spontaneous generation" Pasteur referred to creation of complete biological systems--whether fireflies from morning dew or mice from dirty undergarments.

I. A. Chemical Evolution

It was not until the early 1920's with the work of the Soviet biochemist A. I. Oparin and the British biologist J. B. S. Haldane that seminal ideas again began to emerge. The two researchers independently theorized that components of living entities could more readily be formed in the earth's primordial conditions. In their models, before there were systems able to evolve indefinitely under natural selection, a period of chemical evolution took place, leading gradually to an accumulation of a stock of the types of organic molecules characteristic of all living organisms (A). Oparin's work pointed to the formation of organic molecules (such as oils and waxes which constitute petroleum deposits) in an atmosphere abounding in hydrogen. Haldane's work was complimentary to Oparin's in saying that organic molecules could more readily be produced and amassed in an atmosphere devoid of free oxygen, which would react with and degrade the organic molecules. Both models elicited external perturbations, such as ultraviolet light, to provide the energy for the requisite endothermic reactions.

Here, two ironies present themselves immediately: First, according to this picture the atmosphere which is suitable and

---

(A) It was Darwin himself who had conceived of chemical evolution. He had written to a friend Hooker, "We can conceive of a warm little pond with all sorts of ammonia and phosphoric salts- with light, heat, electricity, etc. present.... a protein compound already formed, ready to undergo more complex changes." C. Ponnamperuma (81)

necessary for life now is precisely wrong as the conditions to spawn life in the first place, and conversely life, as we know it, would perish if it were subjected to the atmosphere which made it possible. And just as ironic is that solar ultraviolet light, which destroys many molecules essential to life, would have been the primary energy source to drive the chemical reactions producing organic molecules in the first place.

The early atmosphere of the earth included the elements hydrogen, carbon, oxygen, and nitrogen in the chemical combinations methane (CH_4), ammonia (NH_3), water vapor and free hydrogen (a great deal of which would have escaped the earth's gravitational field). This mixture of gases is rather evocative of the present atmospheres of Jupiter, Saturn, and Saturn's moon Titan. This early atmosphere was drastically modified through the action of various energy sources. These include ultraviolet light and x-radiation from the sun, cosmic radiation, terrestrial lightning, geothermal sources such as hot springs and volcanoes, terrestrial radioactivity in rocks, shock waves generated by the collision of meteorites, and even the shock waves generated by thunder or by ocean surf. Free oxygen did not appear until plant life capable of photosynthesis developed.

Experiments attesting to the viability of such a model have been carried out repeatedly since 1953 when S. Miller and H. Urey first performed their experiments, passing electrical discharges through a mixture of gases simulating the atmosphere of the primitive earth.

Their experiments, replete with vaporizing and condensing stages, representing the water cycle in nature, saw the synthesis of water-soluble organic molecules, with at least fifteen percent of the carbon atoms originally present in methane having been transformed. In all, four of the amino acids of the twenty which constitute proteins had been produced, including glycine (NH_2CH_2COOH) in reactions such as,

$$NH_3 + 2CH_4 + 2H_2O + \text{(energy)} \longrightarrow NH_2CH_2COOH + 5H_2.$$

One objection to the Miller-Urey hypothesis has been that it presupposes a rich admixture of hydrogen in the primordial atmosphere, an assumption which may not be valid in light of this element's propensity to escape the gravity of the earth. This difficulty, however, may not be intractable. P. Abelson has shown that complex organic molecules can be formed in a weak reducing atmosphere, one low in hydrogen content. Carbon dioxide, carbon monoxide and nitrogen in the presence of only small quantities of hydrogen, when irradiated with ultraviolet light, can react to produce hydrogen cyanide (HCN) and water. The former (HCN), combining with itself in an alkaline bath (such as that presumably available in the early oceans), can also produce amino acids. Here the reaction is given by

$$3HCN + 2H_2O + \text{energy} \longrightarrow NH_2CH_2COOH + CN_2H_2,$$

where the products are glycine and cyanamide molecules, respectively.

Although nucleic acids were not produced directly in the Miller-Urey experiments, it was shown by J. Oro in the early 1960's that cyanide (HCN) molecules could join together to form adenine, one of the bases of DNA and RNA. Meanwhile, another of the organic molecules synthesized by Miller-Urey, formaldehyde (CH_2O) had been known for the past century to possess an affinity to join together to form ribose, another constitutent of RNA.

A serious objection to the Miller-Urey scenario, however, has existed all along. Their's and subsequent experimenters' work have all succeeded in creating one of the necessary components of living cells, the various amino acids, or the building blocks of proteins. The other necessary component, the nucleotides which make up the DNA and RNA molecules, and in turn genes, however, had hitherto eluded synthesis in the laboratory. That is, until recently, when C. Ponnamperuma announced their detection along with the amino acids. Using improved detection techniques, designed for analysis of organic molecules in meteorites, Ponnamperuma found all four bases which comprise DNA--adenine, cystosine, guanine and thymine, and the fifth, uracil, which substitutes for thymine in RNA. Evidently, nucleic acids were being formed along with the amino acids all along; the experiments themselves were not new, but the detection techniques, more sensitive.

As for the new techniques' application to the study of meteorites, there was revealed the existence of both amino acids and nucleic acids, molecules clearly formed in the lifeless

conditions of interstellar space.

I. B. Polymerization

The preceding discussions illuminate the possible mechanisms for the formation of organic molecules, as the building blocks for the next stage in the creation of life, namely their assembly into macromolecules--into proteins and the doublehelix of the DNA.

This process, 'polymerization,' represents a far greater puzzle, and invites more speculation and more model building than had the variety of mechanisms which created the organic molecules in the first place. On one hand, it has been incontrovertibly shown that primordial molecules can be transformed into organic (X), and on the other, that DNA placed in a suitable bath will replicate (XX). But the actual assembling of the first DNA from constituent molecules has not been achieved in a laboratory.

The prevailing model for this essential intermediate step is still somewhat vague and involves the accumulation of the various important organic molecules at the edges of shallow ponds by a process of evaporation and/or freezing. There the organization

---

X 'Organic' in the sense of those present in organism.

XX 'Suitable bath' for the invitro synthesis of RNA would be a mixture of three ingredients: 1) A crude wheat germ (or RNA-) polymerace enzyme preparation, 2) The buffer Tris-HCl and salts, e.g. $MnCl_2$, $MgCl_2$, NH_4So_4 and 3) Equal proportions of the substrates Guanine Triphosphate (GTP), Adenine Triphosphate (ATP), Uracil Triphosphate (UTP), and Cystosine Triphosphate (CTP).

into polymers is thought to have occurred, perhaps around particles of clay, which abounds on the edges and the bottom of bodies of water. Certain specimens of clay offer the maximum surface area for a given grain size and allow an enrichment in the concentrations of the organic molecules at their surface, and also by their chemistry catalyze a variety of reactions. The actual organization into macromolecules may come from the lattice structure of the clay. The mineral montmorillonite, for example, which abounds on this planet, has been shown to align some organic substances such as guanine and adenosine in ways to promote polymerization.

Work of the past several decades, expecially by S. Fox, has shown that if high concentrations of amino acids are heated, the molecules will link together in protein-like polymers. Immersion in water will see the formation of minute spheres, 'microspherules' evocative of bacterial cells. One might conjecture that these serve as a prototype for membrane-bound cells, capable of enclosing nucleic acids. But the next step toward life would be to incorporate the self-replicating system within these proto-cells, a system capable of directing the formation, maintenance and renewal of the entire cell.

At this juncture it is possible to offer an alternative scenario to a primordial soup as the spawning ground for the first organisms. Instead of warm shallow ponds sitting below a primordial atmosphere, conditions which have long since disappeared, the correct site (or an additional site) for the origin of life may have been superheated seawater at great ocean depths, conditions

which still exist on earth. H. Yanagawa recently reported that
protein-like spheres had been produced in the laboratory from amino
acids, in the environment of 240° C and 130 atmospheres. Such
conditions exist naturally around hydrothermal vents, deep in
oceans where plumes of hot water and dissolved chemicals spew
forth. The manifest implication is that life is still being
created in the sea.

II. Crystals of Clay Genetic Takeover

The traditional view of the origin of life and the one which has been generally favored by biochemists ever since Oparin and Haldane propounded their ideas, and Miller and Urey gave these ideas weight, has been the two phase theory consisting of molecular or chemical evolution of organic molecules, followed by their polymerization.

Now there are two incontrovertible facts:

1. The most central molecules of life are essentially the same in all organisms on earth today. This is the principle of 'Unity of Biochemistry.'

2. Most of the twenty-eight molecules which comprise life can be synthesized under conditions simulating the primitive earth. These happy facts notwithstanding, there is the problem of explaining their polymerization into extremely sophisticated structures, possessing an entirely symbiotic relationship, and then resolving the question whether it was the proteins or the nucleic acids that came first.

A. G. Cairns-Smith (1984, 1985) has offered a model which reverses the order of prebiotic evolution, in putting Phase Two, the polymerization, ahead of Phase One, the formation of organic molecules. Thus the scaffolding precedes the accumulation of the building blocks. The scaffolding is provided by inorganic clay minerals. Cairns-Smith calls the Principle of the Unity

of Biochemistry "red herrings," or misleading clues. The originality of his approach and the fundamental nature of the questions he poses most certainly warrant a discussion of his model.

The complexity of the structures of the proteins and the nucleic acids, the existence of a unique code for converting RNA messages into protein sequences (even though innumerable others can be envisioned), suggests that all life now on earth descended from a common ancestor, which was quite high on the evolutionary tree, and the central biochemical system was fixed by that time. Meanwhile the curious interdependent kind of complexity, with the ability to have remained fixed for so long, suggests that the 'high tech' engineering itself was itself the product of evolution, and that the Unity of Biochemistry refers to a much later stage of evolution and not to its beginning.

As for our ability to synthesize the various organic molecules essential to life, Cairns-Smith points out that these molecules are minor constituents of tars, and that it is difficult to imagine that normal geological processes, which would ordinarily make the organic mixture more of a jumble, would instead separate and purify them.

The model calls for three requirements for the first organisms:

1. They must be able to evolve,

2. They must be 'low tech,' and

3. They must consist of inorganic geochemicals.

These lead him to conclude that the first organisms were naked genes

of some unknown material. These then would have evolved to control their immediate environments by specifying the production of increasingly elaborate phenotypes, which would have helped in their survival and propagation.

Certain clay minerals represent the suitable combination of structural characteristics, growth patterns and cleavage properties. Defects in crystal structure could supply multiple, stable alternative configurations for the storage of information. Clay minerals are crystallized from dilute solutions of silicic acid and hydrated metal ions formed by the weathering of rocks. The two mechanisms which power the formulation of clay minerals are the geologic cycle, in which radioactivity inside the earth bakes sediments and thrusts them above the surface, and the water cycle, in which the action of the Sun causes the evaporation of water from the sea, followed by the condensation in the atmosphere.

The introduction of organic molecules, e.g. amino acids, di- and tricarboxylic acids, can make metal ions such as aluminum more soluble in water, and thereby provides a catalytic service for clay synthesis. This could answer the question, 'Why?' for a genetic takeover in time of inorganic genes by organic ones.

The actual takeover machinery in Cairns-Smith's view is offered by photosynthesis, using carbon dioxide from the atmosphere to make molecules such as formic acid. The formation of more complicated molecules, the nucleotides, before the existence of enzymes, is a question open to speculation. Also open to speculation are 1) Why crystal genes are no longer with

us? and 2) What experiments can one perform to test the model? About the first question, Cairns-Smith offers several possible answers, including that there are perhaps no crystal genes, or they have evolved into unrecognizable forms, or suitable conditions for their formation no longer exist. About the second question, the author reports that NASA biochemists have found that zinc and copper can mediate the binding of nucleotides to clays, and that ions in clay can exert selective catalytic effects on amino acids. Also it is reported that clays subjected to cycles of wetting and drying can cause linking in the amino acid glycine.

Some of the experimental results concerning clay, in fact, can be used by the proponents of the chemical evolution model to support their thesis that crystal structure of clay can direct the polymerization of the already-formed organic molecules. One test that would be most revealing would be provided by the discovery of crystal genes. This could reveal the plausibility of the mineral versions of replicating systems.

III. A. Unresolved Questions and a Mathematical Model

A number of questions emerge and wax significant. With a
view toward framing these questions with some degree of precision,
we shall present next a mathematical model formulated essentially
for this purpose.

Sophisticated mathematics for population genetics was developed
early, around the time of Oparin and Haldane's work, by R. A. Fisher
(1930). It was refined further by M. Kinumura (1970), and notably
applied to aging in cells by Kirkwood (1980). A mathematical model
specifically for the origin of life was formulated by F. J. Dyson
(1982). It is this last treatment that we shall call upon here.
It should be reiterated at the outset that this is a 'model,'
rather than a 'theory' on the origin of life. The latter would
call for a detailed examination of reaction rates and dissociation
rates and much more knowledge of the primitive earth than is
presently available.

Before sketching out the mathematics underlying Dyson's model
we identify some of the questions which the model may aspire to
answer. The questions in most cases are significant for the 'Genetic
Take-Over Model' of part II.

1. As we have seen, organisms in the modern sense have two
    essential components: nucleic acids which are essential
    to genetic continuity, and proteins which have catalytic
    and structural functions. Their relationships, totally

symbiotic and interdependent, leads to an obvious question:
'Were the first living creatures composed of proteins or
of nucleic acids, or of a combination of the two?'

2. At what state did random genetic drift give way to natural
selection?

3. Is there a violation of the Central Dogma of molecular
biology, which states that genetic information can pass
from nucleic acids to nucleic *acid* or from nucleic to protein,
but cannot pass from protein to nucleic acid or to
protein? (The information is contained in the nucleic
acid, and not the proteins.)

4. What is the origin of nucleic acids?

5. How did these proto organisms faithfully reproduce before
the modern genetic apparatus evolved, and how did the
present system evolve?

III. B. A Mathematical Model

The assumptions underlying the Dyson Model are the following:

1. Molecular or chemical evolution takes place in small 'islands'
or colonies. These could be colloidal droplets or solid
particles, such as those of clay, with molecules adsorbed
on the surface. The islands exchange molecules with the
environment which acts as a source of chemical free energy
to catalyze reactions.

2. Chemical evolution toward more organized polymerization in

the island is by randon genetic[X] drift and not by natural selection. (The latter process would be ascribed to a later development when the islands begin to grow and compete for nutrients.)

3. Each island contains a fixed number N of 'monomers' or molecular units. Some of the monomers may be free and others, combined in random ways into polymers.

4. The population of polymers changes by discrete mutation, with monomers being added, substracted or exchanged, one monomer at a time.

5. In normal living systems reactions are stimulated by enzymes. The efficiency of an enzyme in catalyzing a reaction is a function of its site. Thus a monomer will be regarded as 'active' or 'inactive' according to whether or not it is correctly placed as part of a structure catalyzing the synthesis of other catalytic structures. Multidimensional random walk of polymer mutation will be mapped onto one dimensional random walk by counting the numbers of 'active' and 'inactive' monomers.

With k designating the number of active monomers, the parameter $x \equiv k/N$ will represent the fraction of active monomers in the popula-tion. _Thus $0 \le x \le 1$_ A simulated potential energy U (x) is formulated with an unstable maximum point at β, and a pair of minima at α and γ (see Appendix). Thus U (α) and U (β) represent two different levels of

---

[X] 'Genetic' in this sense is not meant to connote the existence of genes as much as the ability to transfer information.

organization. Starting from a random population of monomers, each
island will approach the 'disordered' state or one of limited
molecular organization at x=α rapidly, and remain there for a long
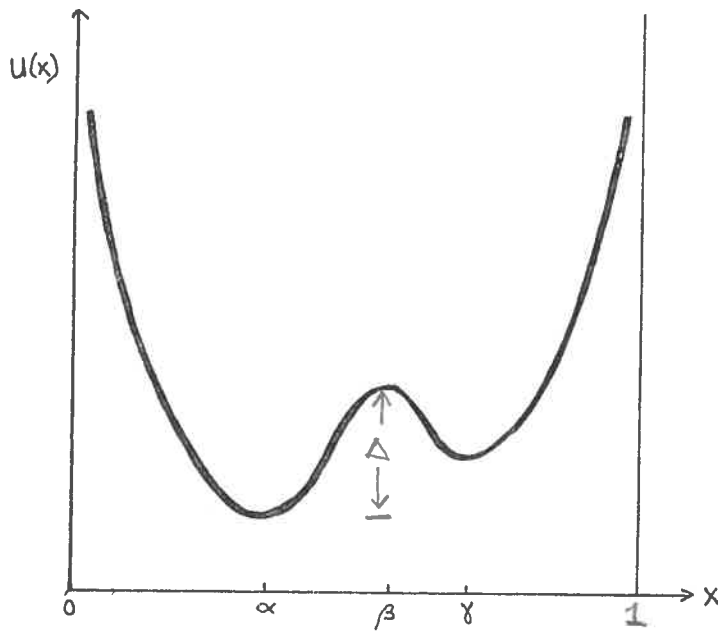time, executing statistical fluctuations about the disordered state.

Figure 1.

With very low probability there will take place large enough
statistical fluctutations to carry the population over the saddle
point to the 'ordered' state or one of higher organization *at x=γ*. Then
in the ordered state there will again be statistical fluctuations,
which after some time, will take it back over the saddle point to
the disordered state. Without the activation of an enzyme the
population cannot jump over the saddle point.

Regarding in some sense the ordered state at x=γ to represent
the 'living' state of the population, transitions from x=α to x=γ
is the crucial step in the origin of life; the transitions in the

reverse direction, that of 'death'.

The population distribution will be centered around the two minima at $x=\alpha$ and γ, and will have an approximate Gaussian form. In terms of this distribution one can proceed to find the ratio of the populations in the ordered to the disordered states. However, it is not the size of the ordered population that is crucial to the question of the origin of life as much as the rates of transition in the two directions. For this, we need the transition rate for the disordered to the ordered state, i.e. the probability per mutation, (polymerization or increase of order), that a population crosses the barrier from α to γ: (see equation (A-4) in the Appendix). In terms of the barrier height $\Delta=U(\beta) - U(\alpha)$, a parameter related to the activation energy, the meantime for a population to make the transition from disorder to order is computed (see A-6).

In the absence of any real clues for the number of island populations that may have been present and their duration time, one can only supply 'reasonable' guesses. With Dyson's suggestion of perhaps 10^{10} islands existing for 10^5 mutation times, N monomers would have made the disorder to order transition with an upper limit given by N_c, where

$$N < N_c = 30/\Delta \qquad (1)$$

Since each catalyst lowers the activation energy for correct placement, an increase in the ~~discrimination factor b~~ *catalytic effectiveness* manifests itself as a decrease in Δ. This, in turn, sees a corresponding

increase in N_c.

U(x) is expressed in terms of an autocatalytic function of ϕ(x),
which represents the probability that the mutated unit be active in an
island with k monomers present. A mathematical description of ϕ(x)
is relegated to the Appendix. Let it suffice for now that ϕ(x) is
a monotonically increasing function defined for the closed inter-
val $0 \leq x \leq 1$ crossing ϕ(x)=x at the three points α, β, γ, with
$0 < \alpha < \beta < \gamma < 1$. (See Figure A-1 in the Appendix.)

The detailed chemical mechanisms by which population transitions
occur are incorporated in the function ϕ(x) which relates the
catalytic activity of the newly placed monomer to the catalytic
activity of the extant population. In the absence of any detailed
knowledge of prebiotic chemistry, a primitive ϕ(x) is chosen based
on two parameters a and b. The former specifies, as (a+1), the
number of species of monomers; the latter, b, is the measure of their
effectiveness in catalyzing reactions, a discrimination factor.

The properties of ϕ(x) impose certain constraints on the para-
meters a and b: for the discrimination factor a lower limit exists
(see Appendix):

$$b > e^4 = 54.6 \tag{2}$$

Thus a value of 10^2 for b would be reasonable for this prebiotic molecular
evolution, compared with typical discrimination factors between
10^3 and 10^4 in present day enzymes, which are suitable for the fine-
tuned present day organisms. In order to guarantee the existence
of a pair of distinct stable ordered and disordered states,
separated by saddle point, hence to make the model viable, a lower

limit on a emerges:

$$a > e^2 = 7.4 \tag{3}$$

Thus, there must be nine or more species of monomers. Pure nucleic acid systems (constructed from four bases), without amino acids cannot be accommodated by the model. However, ten to twenty amino acids, or a mixed system consisting of ten or more amino acids and nucleotides can be accommodated. For a $\leq$ 7 distinct ordered and disordered states cannot coexist. Either one finds a disordered equilibrium state (if the catalysts are weak), but no ordered state, suggesting that cells are incapable of living; or an ordered equilibrium state (if the catalysts are strong), but no disordered state, suggesting here that the cells are incapable of dying. The problems of interest span the range from 'marginally alive' (α, β = γ), to the symmetric (α, β=0.5, γ), to the 'marginally immortal' (α=β, γ). Defining a pair of parameters A = ln a and B = ln β, these problems all lie in a wedge shaped area in A-B space. See Figure A-2 in the Appendix.

In the 'marginally alive' case there is no minimum in U(x) at γ but rather a point of inflection at β=γ (Figure 2). In the 'marginally immortal' case, there is no minimum at x=α, but a point of inflection at α=β (Figure 3). An infinitesimally stronger catalyst in the marginally alive case would produce the disorder-order transitions. Conversely, an infinitesimally weaker catalyst in the marginally immortal case would produce the order disorder transitions.
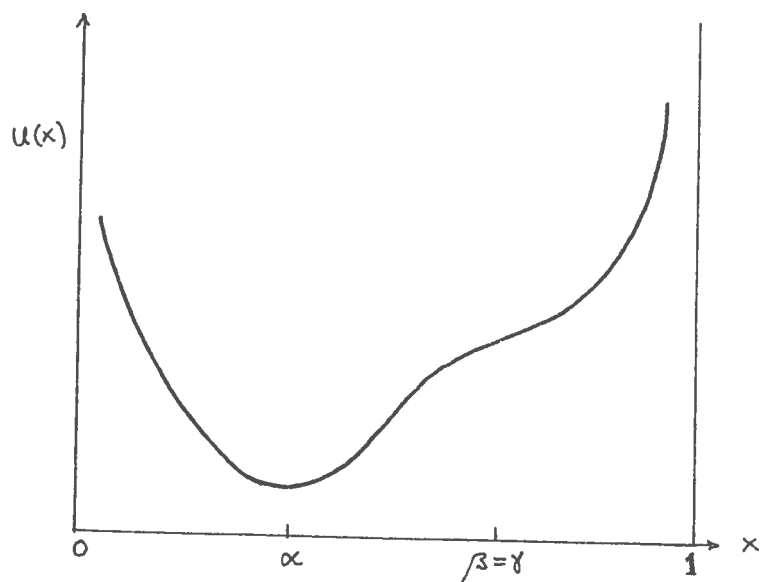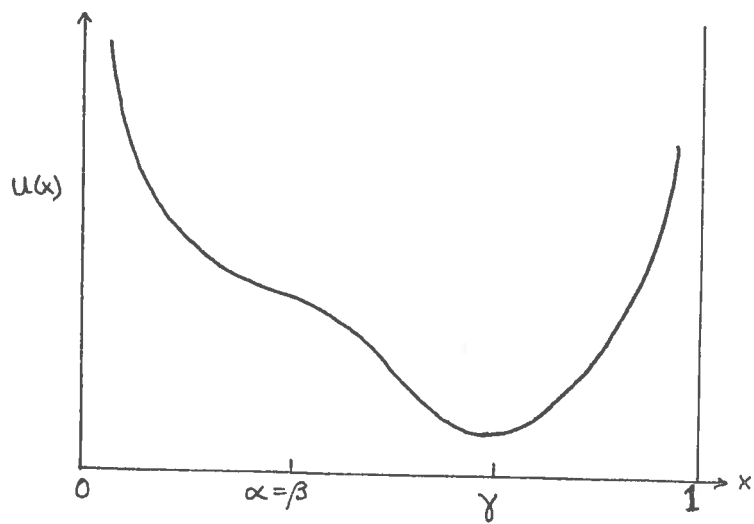
Figure 2.

Figure 3.

I. For a=8 (i.e. nine species of monomers)

    a) The marginally alive model with a discrimation factor
b=62.9 yields the critical values of α=0.32, β=γ=0.59
(This is to be interpreted as 32% of the catalysts in
the disordered state are active, and 59% of the
catalysts in the ordered states are active). The
barrier height Δ=0.002695 and a maximum population of
N_c = 11,131 can make the disorder→order transition.

    b) The symmetric model with a discrimination value b=a^2=64
yields α=0.33, β=0.50, γ=0.67. The barrier height is
reduced to Δ=0.001129 and the maximum population making
the transition increases to N_c = 26,566.

    c) Finally in the case of the marginally immortal, an
enzyme discrimination factor of b=65.7 yields α=β= 0.39,
γ=0.70. The barrier height Δ=0 and the maximum number
making the transition becomes infinite.

II. For the case of a = 19 corresponding to a+1 = 20 species
of monomers, and akin to the present organisms with 20
amino acids.

    a) The marginally alive model with an enzyme discrimination
factor of b=219.3 yields α=0.07 and β=γ=0.75. The barrier
height Δ is given by 0.1906 and the maximum number making
the transition by 157.

    b) The symmetric model with a discrimination factor b=a^2=361
yields α=0.08, β=0.50 and γ=0.92 barrier height is lowered

to 0.1051 and the maximum population is raised to 285.

    c) The marginally immortal case with the discrimination factor b=3195 yields $\alpha=\beta=0.16$ and $\gamma=0.99$. The barrier height Δ vanishes, and the maximum population goes to infinity.

III.   An intermediate case with a=10, (11 species), and b=100. The symmetric model $\alpha=0.20$, $\beta=0.50$ and $\gamma=0.80$ presents a barrier height $\Delta=0.0145$ and a maximum population making the transition as 2068. The significance of the α and β is that in the disordered state 20% of the monomers are active and in the ordered state it is 80% that are active.

The implication of the mathematical model is that nucleic acids with four species of nucleotides alone could not have conspired to evolve into organisms. Amino acids, however, numbering nine or more, and polymerized into protoenzymes with a polypeptide structure, could have reached the ordered state. Thus the Dyson Model seems to suggest that enzymes preceded nucleic acids and genes.

The model does not allow for growth of a population of monomers within an island. Once an island population has reached the ordered state, however, one could visualize its passing into a new phase where chemicals from the environment could be absorbed and a growth in the island's population realized. Competition for nutrients between such islands could then follow, with those that suffer shortage of nutrients dying, and others flourishing. Thus, natural selection would come at this later state of evolution.

The Central Dogma of molecular biology is violated in that the

enzymes are allowed to pass on the genetic information to other enzymes. The model suggests that at this prebiotic stage of evolution the Central Dogmas were not yet operational.

If this model is correct and amino acids did precede nucleic acids, organisms in the beginning may have used nucleotides such as adenine triphosphate (ATP) as energy carriers. A description of the recipe for assembling such nucleotides by way of an enzyme Q_β replicase appears in a paper by Biebricher et al (1981).

It is possible to speculate about a number of different pathways along which the genetic code could have evolved. One possible channel would involve the non-specific binding of RNA to free amino acids, leading to transfer RNA; specific binding to catalytic sites leading to ribosomal RNA. In this particular pathway described by Dyson, catalytic sites evolve from special-purpose to general-purpose by using transfer RNA instead of amino acids for recognition. Then the recognition unit splits off from the ribosomal RNA and becomes messenger RNA; the former becoming unique, as the genetic code takes over the function of recognition.

As for the question of when the latest common ancestor of modern organisms emerged, it appears likely that the evolution of the genetic code was an extremely lengthy and tortuous process, the emergence of the ancestral prokaryotic cells taking much longer perhaps than the evolution of the prokaryotic to the eukaryotic, and subsequently multicellular organisms. Evolution certainly would have been accelerated after the establishment of

the genetic code. At this juncture it should be pointed out that in the last few years the universality of the genetic code itself has come into question. In the early 1980's exceptions in mitochondria had been found. More recently exceptions were reported in the genetic code in bacteria as well. The question which arises immediately is, 'Did all organisms evolve from a single ancestor or did two separate evolutionary lines occur, commencing from separate origins?' The existence of inordinate similarities in the genetic codes of all organisms suggest that the most plausible explanation is the former, and in fact that the branching occurred very late in the evolutionary process.

Concomitant with the question of how life began on earth is the one of whether similar conditions have occurred elsewhere and hence of the existence of extraterrestrial life. We live in a metagalaxy, the universe, of approximately 10^{23} stars. Although we have not seen a single planet outside our solar system, our faith in the universality of physical laws, coupled with an understanding of the formation of our solar system, albeit incomplete, suggests that planets are commonplace, and conditions suitable for life are also commonplace.

The earth represents the only successful experiment that we know of for the origin of life. Or can we even be certain of this? Indeed, in the early part of this century the Swedish chemist Swante Arrhenius proposed a theory of 'panspermia,' or life having originated elsewhere, and of the spores having dispersed about as seeds, giving rise to life on earth. Arrhenius argued that these spores could have been propelled through interplanetary

or interstellar space by electrostatic forces. The theory simply offers an alternative site for the primordial soup and does nothing to remove the formidable difficulties in resolving the plethora of questions. In fact, it introduces the additional problem of explaining how these spores survived for any length of time the frigid cold (3° K) of interstellar space, and the unrelenting bombardment by cosmic and solar-radiation. A whimsical scenario offered by Thomas Gold however, alleviates the latter difficulties by suggesting that it was a spaceship of aliens who stopped off for a picnic on earth and left their crumbs behind. The theory has life evolving from these crumbs, and has been aptly named "Gold's Garbage Theory."

EPILOGUE

The foregoing scenarios point toward how, within the framework of physico-chemical processes, life could have originated, but demonstrating the possibility of a process is not the same as show-ing its probability, and certainly not its inevitability. Was the evolution of primitive life a necessary (therefore inevitable) consequence of the conditions from which it arose, or was there an ephemeral episode whose chances of occurrence in the first place were so unlikely that it would occur again, given the substantial time?

In Dyson's mathematical model sizeable populations of monomers could make the transition from the disordered to the ordered state subject to reasonable conditions. (For example, with eleven species

of monomers in the symmetric model "populations of around 2000 could make the transition without requiring a miracle.") In Ponnamperuma's experiments nucleic acids were synthesized along with amino acids after electrical discharges were applied to gases simulating a primordial atmosphere. In Yanagawa's experiments protein microspherules were formed in superheated water, simulating conditions near hydrothermal vents in the ocean floor. The confluence of all of these results appears to reduce the origin of life to a level no less unlikely than an incident which took place in Cleveland in 1895: There were only two cars in the entire city and they collided.

NEW APPENDIX

APPENDIX

In terms of the parameter $x=k/N$, representing the fraction of active monomers in the population, a function $\phi(x)$ is constructed. $\phi(x)$ describes the autocatalytic capability (for polymerization) of the assemblage of active monomers. The function is chosen to depend on ~~the~~ *two* parameters a and b, where a+1 specifies the diversity of the population of monomers, *of which a are inactive* and b, ~~this~~ called *the discrimination factor, measures the* precision of the polymerizing catalysts. The form of $\phi(x)$ is based on a simple model in which each site in the island is assumed to be occupied by either an active or an inactive monomer. In the absence of autocatalysis, the probability that a site will be correctly occupied is

$$\phi(0) = (1+a)^{-1} \tag{A-1}$$

If each catalyst can discriminated active from inactive monomers by a factor b, then the probality that a newly placed monomer will be active is given by

$$\phi(1) = (1 + a/b)^{-1} \tag{A-2}$$

For intermediate values of x, $0 \leqslant x \leqslant 1$,

$$\phi(x) = (1 + ab^{-x})^{-1} \tag{A-3}$$

Here each perfect catalyst is assumed to lower the activation energy (required for correct placement) of a monomer by

$$d = k \, T \ln b \tag{A-4}$$

The function $\phi(x)$ is a monotonically encreasing function over the closed interval $0 \leqslant x \leqslant 1$. *It is possible to show that* ~~the equation $\phi(x)=x$ has three~~

the simulated potential energy U(x) is given in terms of $\phi(x)$ *by*

$$U(x) = \int_0^x \left\{ \ln\left[x/(1-y) \right] - \ln\left[\phi(y)/(1-\phi(y)) \right] \right\} dy \tag{A-5}$$

U(x) for the simple choice $\phi(x)$ given by equation (A-3) can be written as

$$u(x) = x \ln x - (1-x) \ln(1-x) + x \ln a - \tfrac{1}{2} x^2 [\ln b]^2 \qquad (A-6)$$

From Eq. (A-5) it is very simple to verify that the maxima and minima of $u(x)$ are given by the equation

$$\phi(x) = x \qquad\qquad (A-6)$$

This equation has three solutions, corresponding to ~~solutions~~, $x = \alpha, \beta, \gamma$, with $0 < \alpha < \beta < \gamma < 1$. The condition that $\phi(x)$ crosses $\phi(x) = x$ at three points imposes a constraint on the derivative of $\phi(x)$ at $\alpha, \beta,$ and γ :

$$\phi'(\alpha) < 1, \quad \phi'(\beta) > 1, \quad \phi'(\gamma) < 1 \qquad\qquad (A-7)$$

Thus $\phi(x)$ resembles the letter "S" stretched out at angle of 45° as seen in Figure A-1.

The constraints on the number of monomers (a+1) and on the discrimination factor b arise from the restrictions imposed on $\phi(x)$. A necessary (but not sufficient condition) for $\phi(x)$ to cross $\phi(x)=x$ three times is for $\phi'(x)$ at the point of inflection
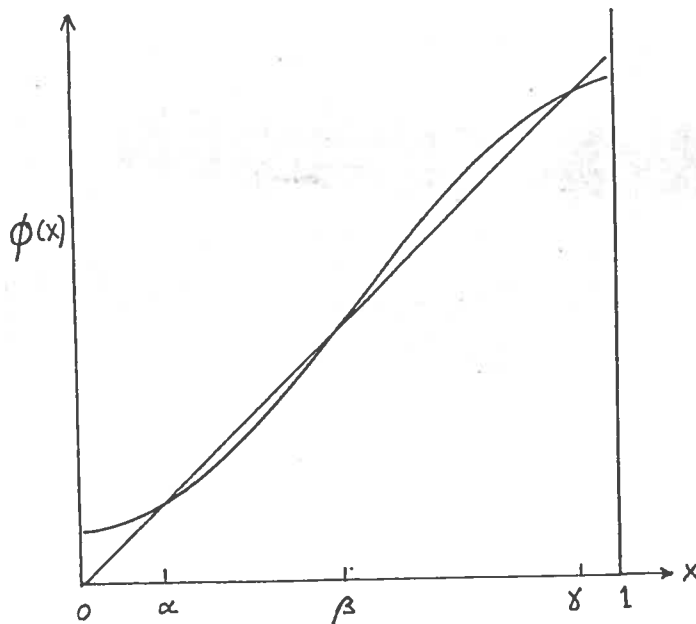

Figure A-1. The autocatalytic function
(x) as a function of the fraction of
active monomers.

to be greater than unity. We can compute this readily:

$$\phi(x) = (1+e^{A-Bx})^{-1}$$
$$\phi'(x) = Be^A \left[e^{-Bx}(1+e^A e^{-Bx})^{-2} \right]$$
$$\phi''(x) = Be^A \left[-Be^{-Bx}(1+e^A e^{-Bx})^{-2} + 2Be^{-Bx}(1+e^A e^{-Bx})^{-3} \right] = 0$$
$$\frac{Be^{-Bx}}{[1+e^A e^{-Bx}]^2} = \frac{2Be^{-Bx}}{[1+e^A e^{-Bx}]^3}$$
$$1 + e^A e^{-Bx} = 2$$
$$e^{Bx} = e^A$$

which yields the point of inflection as x=A/B = ln a/ ln b.

Substituting this value into the first derivative, we obtain

$$\phi'(x) = Be^A \left\{ \frac{e^{-B(A/B)}}{(1 + e^A e^{-B(A/B)})^2} \right\} > 1$$

$$B e^A e^{-A} / (1 + e^A e^{-A})^2 > 1$$

$$B/2^2 > 1$$

Finally, we have $B > 4$, or $b > e^4 = 54.6$ (A-8)

Meanwhile, another necessary condition for $\phi(x)$ to have three solutions is that at the unstable saddle point $x = \beta$, $\phi'(x) > 1$. A similar derivation to the foregoing yields a lower limit on A,
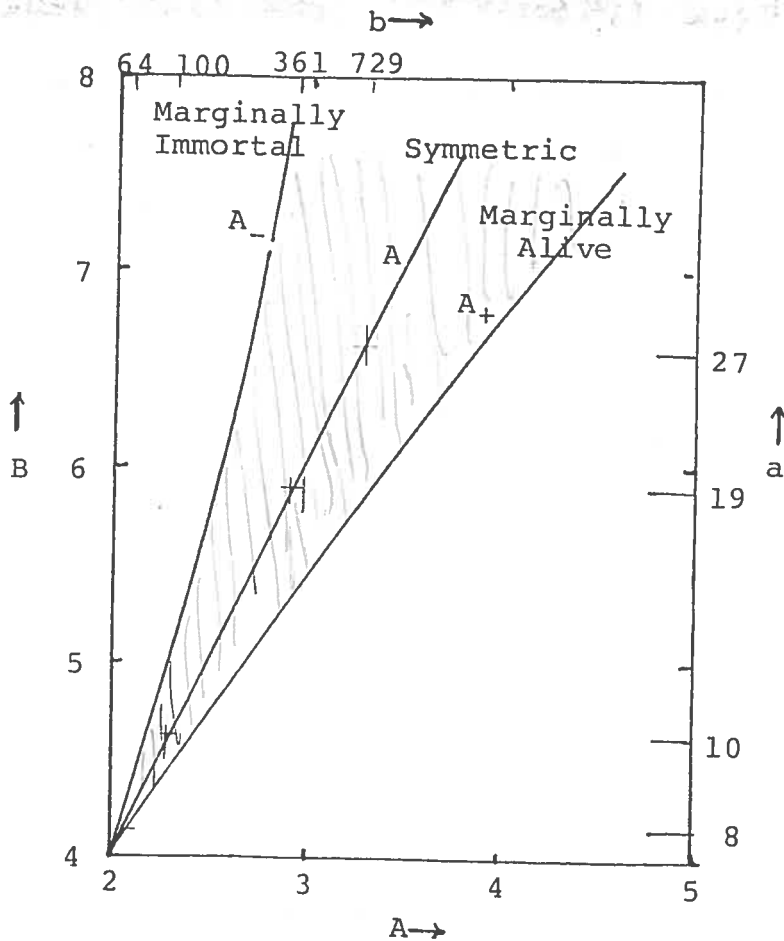
$$A > 2, \quad a > e^2 = 7.4 \tag{A-9}$$

Figure A-2. The wedge-shaped region in A-B space corresponds to the range of values for which the potential energy curve U(x) has viable minima at α and γ, and a maximum at $x = \beta$. Thus order-disorder transitions become relevant only for this region.

It can be shown that
the transition rate from disorder to order is given by

$$\varepsilon = (2\pi N)^{-1} \left\{ [1 - \phi'(\alpha)][\phi'(\beta) - 1] (\alpha - \alpha^2)/(\beta - \beta^2) \right\}^{1/2} \exp\left\{ N[u(\gamma) - u(\alpha)] \right\} \quad (A-10)$$

and the transition rate for the reverse process by

$$\eta = (2\pi N)^{-1} \left\{ [1 - \phi'(\gamma)][\phi'(\beta) - 1] (\gamma - \gamma^2)/(\beta - \beta^2) \right\}^{1/2} \exp\left\{ N[u(\gamma) - u(\beta)] \right\} \quad (A-11)$$

The ratio of the two rates (A-10) and (A-11), which is also

the inverse of the population ratio

$$\frac{\eta}{\varepsilon} = \sqrt{\frac{1 - \phi'(\gamma)}{1 - \phi'(\alpha)} \frac{[\phi'(\beta) - 1]}{[\phi'(\beta) - 1]} \left(\frac{\gamma - \gamma^2}{\beta - \beta^2}\right)\left(\frac{\beta - \beta^2}{\alpha - \alpha^2}\right)} \exp\left[N\left(u(\gamma) - u(\beta) - u(\alpha) + u(\beta) \right) \right] \quad (A-12)$$

The mean time for a population to make the transition from

disorder to order is given by

$$t = \tau e^{\Delta N} \qquad (A-13)$$

where τ is the average time between mutations at a given site.

(For details of the derivation: Dyson (82)).

ACKNOWLEDGMENTS

LIST OF REFERENCES

C. K. Biebricher, M. Eigen, R. Luce (1981), Kinetic Analysis of Template Instructed and de novo RNA Synthesis by Q_β Replicase, Journal of Molecular Biology, 148, 391-410.

A. G. Cairns-Smith (1984), Genetic Takeover and the Mineral Origin of Life, Cambridge University Press.

(1985) The First Organism, Scientific American (June 1985).

F. J. Dyson (1982), 'Origin of Life,' Journal of Molecular Evolution, 18, 344-355.

Lecture given at Oxford University, Department of Theoretical Physics, May 1982.

R. A. Fisher (1930), The General Theory of Natural Selection, Dover Reprint Series (1958), Dover Publications, Inc., N. Y.

D. Goldsmith, T. Owens (1980), The Search for Life in the Universe, Benjamin-Cummins Publishing Co., Inc.

M. Kimura (1970), Mathematical Topics in Population Genetics, ed. K. Kojima, Springer-Verlag, N. Y., 178-209.

T. B. L. Kirkwood (1980), Error Propagation in Intracellular Transfer, Journal of Theoretical Biology, 82, 363-382.

S. Miller, L. Orgell (1974), The Origin of Life on the Earth, Prentice Hall, Englewood Cliffs, N. J.

A. I. Oparin (1953), Origin of Life, Dover Publication, Inca, N. Y., Dover Reprint Series.

C. Ponnamperuma (1981), "Search for Life in the Universe," Lecture given at Mary Washington College, Fredericksburg (May 1981).

A. Weiss (1981), Replication and Evolution in Inorganic Systems, Angew, Chem. Int. Engl., 20, 850-860.