

Talk presented at the Mitsubishi Kasei
Institute of Life Sciences, Tokyo,
Japan, May 18, 1974.

LBL-3319
c.2 repl.

CHEMICAL EVOLUTION

Melvin Calvin

May 18, 1974

RECEIVED
LAWRENCE
RADIATION LABORATORY

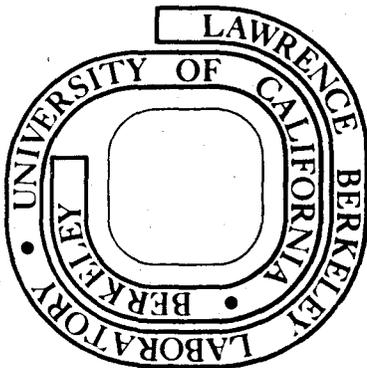
MAY 7 1975

LIBRARY AND
DOCUMENTS SECTION

Prepared for the U. S. Atomic Energy Commission
under Contract W-7405-ENG-48

TWO-WEEK LOAN COPY

This is a Library Circulating Copy
which may be borrowed for two weeks.
For a personal retention copy, call
Tech. Info. Division, Ext. 5545



LBL-3319

c.2 repl.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

CHEMICAL EVOLUTION^{*}

Melvin Calvin

Laboratory of Chemical Biodynamics, University of
California, Berkeley, California 94720^{**}

INTRODUCTION

The question of the origin of life on the earth and the nature of chemical evolutionary processes which could have given rise to it is one which has engaged the minds of men since they first contemplated the nature of their place on the earth and in the universe.¹ However, the most acceptable view in scientific terms today is one which stems primarily from the concepts first carefully and clearly enunciated by Charles Darwin in his early writings. The basis for his discussion was primarily the morphological and functional forms of today's living things, in their great variety, as well as what was known about the morphology of organisms that had disappeared in geological time as they were exhibited in the paleontological record in the rocks. The only record that was available then, and perhaps almost the only one which is still available, is the shape and structure of historical organisms as they have been preserved in the earth's sedimentary rocks. On that basis, Darwin was able to formulate

* Transcription of talk presented for Mitsubishi Kasei Institute of Life Sciences, Asahi-Kodo Hall, Tokyo, Japan, May 18, 1974.

** The preparation of this paper was sponsored by the U.S. Atomic Energy Commission.

his general hypothesis of biological evolution, which is, perhaps, best expressed by the title of a paper which he never used. The title which he and Wallace proposed was "On the Tendency of Varieties to Depart Indefinitely from Original Types". It seems to me that these words express best the fundamental idea of Darwinian Evolution, namely, that two species which exist today as independent species, if followed back in time, were originally two varieties of the same species. So, if one goes forward in time, individual variations would gradually separate in time to become new species. Looking forward in time, they separate; looking backward in time, they come back together. If one follows the reverse-time process sufficiently far back, one must arrive at a time, and a condition, in which there was only one original species -- one type of organism -- from varieties which became today's separate species. If one goes still further in such a backwards extrapolation, reaching that point in which the individual living thing was only one variety of many different kinds of molecular aggregations (collections of molecules), one can see that the transition from molecules to a living thing is a continuous one. Thus, one reaches back into the history of the earth to a period when the earth had no living things, only molecules. And, one can go even further back to that time when there were no molecules, only atoms, bringing us to the period of the evolution of the elements themselves.

As Darwin realized, and expressed in his unpublished paper, there was a continuity in evolution, ultimately arriving at a single successful starting point. That starting point is what we tend to mean by the "origin" of living matter. Darwin wrote a very interesting letter about this idea, in 1882 in response to a query, stating:

"You expressed quite correctly my views where you said that I had intentionally left the question of the Origin of Life uncanvassed as being altogether ultra vires in the present state of our knowledge, and that I dealt only with the matter of succession. I have met with no evidence that seems in the least trustworthy, in favour of so-called Spontaneous Generation. I believe that I have somewhere said (but cannot find the passage) that the principle of continuity renders it probable that the principle of life will hereafter be shown to be a part, or consequence, of some general law".

The statement to which Darwin refers, and which he had forgotten, was written earlier, in 1871:

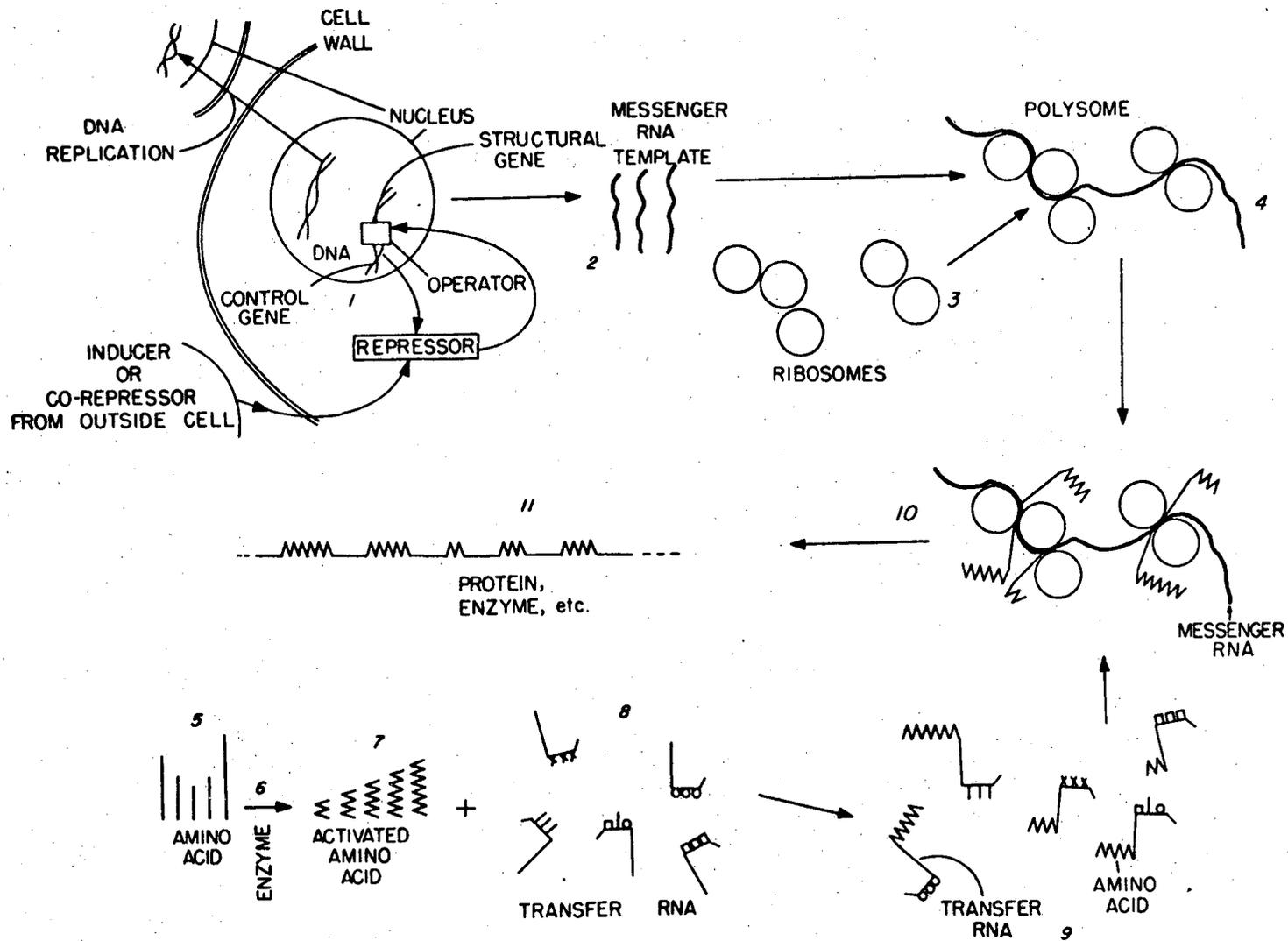
"It is often said that all the conditions for the first production of a living organism are present, which could ever have been present. But if (and oh what a big if) (this is Darwin's parenthetical remark) we could conceive in some warm little pond with all sorts of ammonia and phosphoric salts -- light, heat, electricity, etc. -- present, that a proteine compound was chemically formed, ready to undergo still more complex changes, at the present day such matter would be instantly devoured, or absorbed, which would not have been the case before living creatures were formed".

EVOLUTION OF CHEMICALS

What I propose to do is to discuss a certain period in that time sequence, that period in the time sequence, in which the molecules themselves were being formed and transformed and built up to reach, eventually, a size and complexity which could contain and sustain the living process leading to life as we now know it.² What we are going to do is to have a look at the essence of the molecular nature of living things as we understand them today, to see, first, what it is we must arrive at by chemical means. It is not yet possible for us (and I am not sure it ever will be) to find a record in the rocks of the molecular events which may have taken place prior to the appearance of what we call, today, a living

thing. Therefore, we must try to reconstruct those possible processes from what we know about today's chemistry and see how far we can carry them out, experimentally, in the laboratory.

Before we begin this "travelogue in time" amongst the molecules of the primitive earth, prior to the appearance of living things, let us make a quick examination of the molecular principles of today's living things. Figure 1 is a diagrammatic representation of the essential principles of today's living organisms and their construction.³ A living cell is represented in the upper left-hand corner, with the double line representing the cell wall. Inside the cell membrane is a cytoplasm and inside of that is the cell nucleus which contains the DNA, the genetic material of the cell. What is represented here, in diagrammatic and molecular terms, is what a cell is capable of doing. It must be able to copy the DNA of the cell nucleus to make messenger RNA templated on pieces of the DNA; the messenger RNA combines with ribosomes inside the cytoplasm to form a poly-ribosome (polysome), synthetic machines which eventually can hook together the amino acids in some order designated by the template of the messenger RNA to give a protein. Each one of the lines in 5, 7 and 9 represents a different amino acid which connected in some particular sequence, constitute the structural materials of the cell. Reaction 2,3,4 and 10 represent the information-transferring process. Group 5 represents the simple amino acids formed in the cytoplasm, group 7 are the activated amino acids, and group 9 are the activated amino acids on t-RNA. The loaded t-RNAs come together with the template of the m-RNA to make a specific sequence of amino acids. This process is a combination of information transfer and energy transfer, giving rise to the final product from which the cell is constructed.



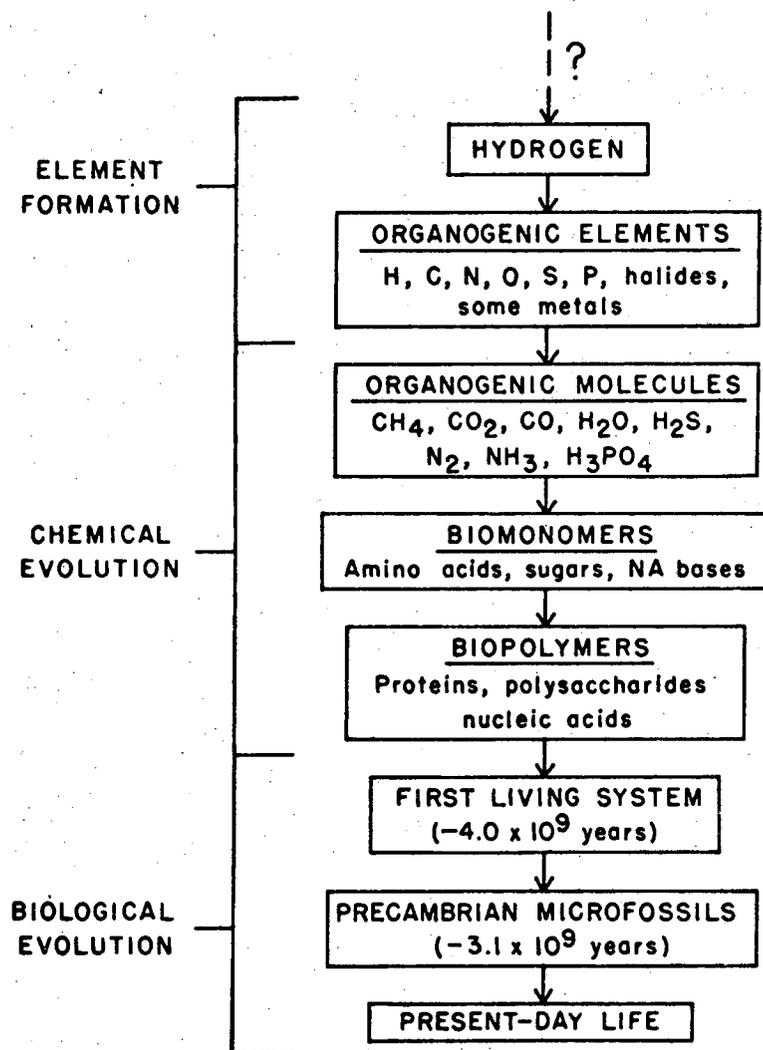
Calvin. Figure 1
 Mechanism of protein biosynthesis

MUB-938-A

Figure 1 contains all of the elements which we must eventually describe in molecular terms -- the appearance of small molecules (amino acids), appearance of polymers of those amino acids and nucleic acids to give large molecules which ultimately have specific structure, giving rise to specific shapes and sizes. Thus, we have the evolution of molecules, the evolution of polymers, the evolution of catalysts and finally the evolution of information such that these processes can be repeated over and over again, thus leading to biological evolution ultimately.

There are four elements of molecular evolution which we must try to understand: The first is the evolution of molecules themselves, the simple of molecules of which living things are made; the second is the evolution of large molecules (macromolecules) and structures; the third is the evolution of catalysis -- the ability to make specific reactions occur with a high degree of efficiency; and, the fourth is the evolution of information and information transfer processes which allow the two systems to coalesce.

Let us go back and look at the time sequence in which the evolutionary results are achieved, which is shown in Fig. 2. In the beginning, most of the elements of the universe were in the form of hydrogen, which eventually must undergo fusion reactions, giving rise to the higher elements in the periodic table, particularly those important to living things: carbon, nitrogen, oxygen, sulfur, phosphorus, halides and certain metals, particularly iron which is important for catalytic functions in living organisms. Then, the primitive (prebiotic, primeval) molecules formed from the organogenic elements with which the earth was coated to begin with: methane, ammonia, carbon monoxide, water, carbon dioxide, hydrogen sulfide and, of course, hydrogen. These first three



Calvin. Figure 2

XBL714-5137

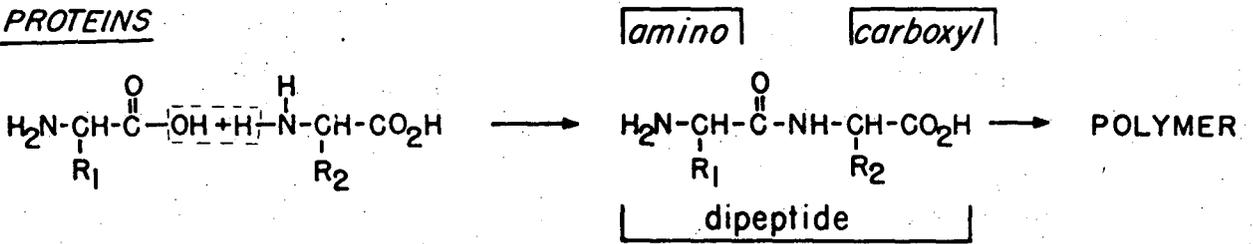
Evolution scheme from elements to present day

stages present no chemical problem, since the first two are nuclear and the third is simply the result of the presence of carbon, hydrogen, nitrogen and oxygen at a low enough temperature to produce the small, primitive molecules. However, the next stage of chemical evolution, from the organogenic molecules to the biomonomers, has been the one area of major progress in the last twenty years.^{2c,4} The conversion of organogenic molecules into amino acids, sugars, nucleic acid bases and other carboxylic acids (acetic acid, citric acid) has been achieved in the laboratory under the influence of a wide variety of energy sources, ranging all the way from the ultraviolet light of the sun, radioactive energy (in the form of ionizing radiation), to mechanical energy (in the form of meteoritic shock waves).⁵ All of these energy sources give rise to the transformation of the organogenic molecules to biomonomers. I will not review the details of this molecular transformation; this has been done repeatedly.

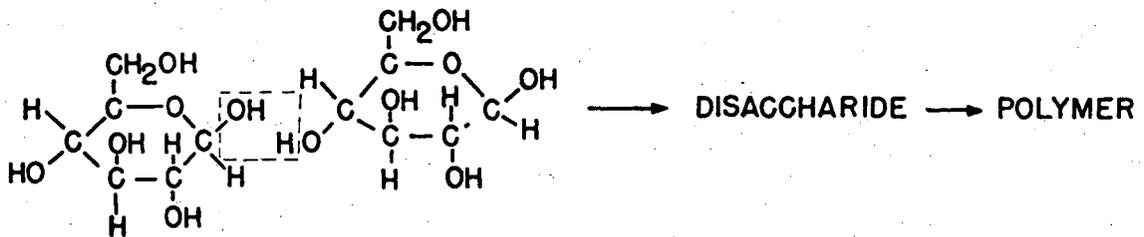
The next stage, the stage of the transition from biomonomers to biopolymers, is more difficult to achieve in terms of chemical evolution. Most of the rest of the discussion will be concentrated in this area -- formation of biopolymers from biomonomers³ and the formation of structures and function in the biopolymer region, eventually giving rise to the first living organism about four billion years ago.

With that as an outline, let us examine the nature of the problem involving the transition from biomonomers to biopolymers. That problem is best illustrated in the formation of polypeptides and nucleic acids, shown in Figures 3 and 4. In each of those cases, in order to make the biopolymer from the biomonomer, it is necessary to remove a water molecule

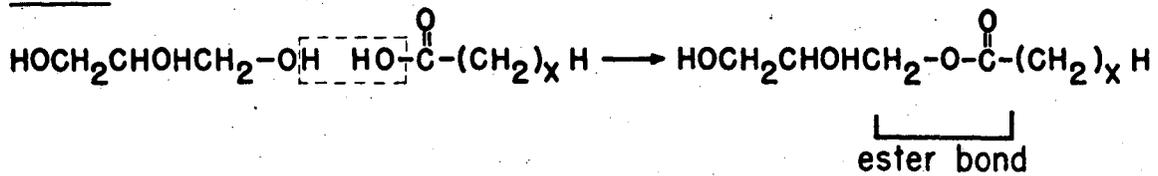
PROTEINS



POLYSACCHARIDES



LIPIDS



MUB-5261

Calvin. Figure 3

Dehydration condensation of polypeptides, carbohydrates and fats

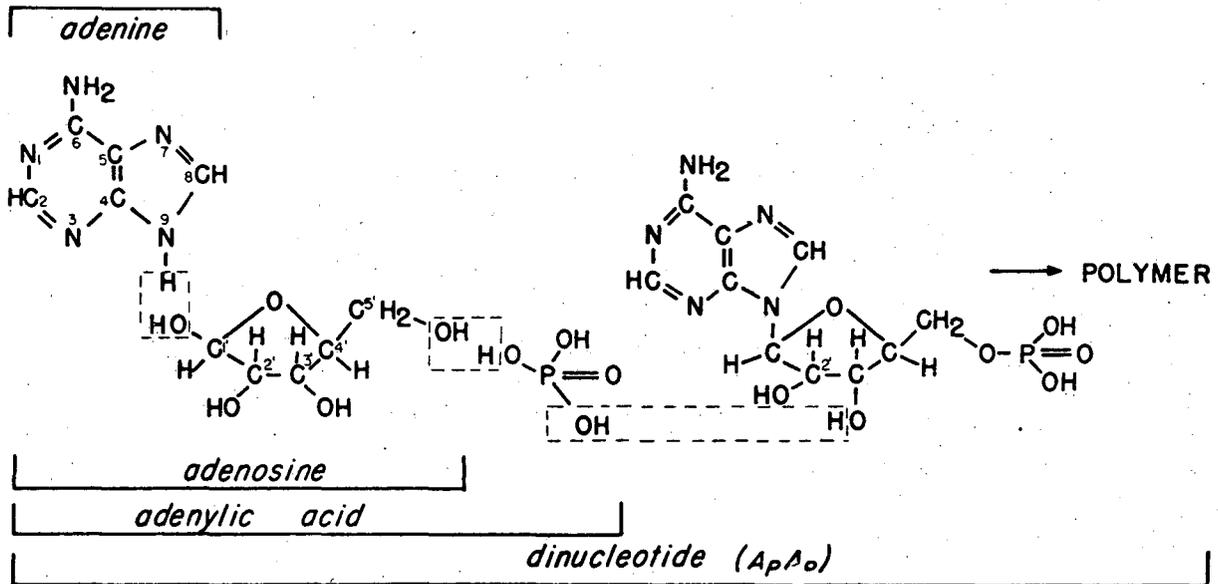
between the two monomers. The removal of the water molecule in every case is the essential chemical reaction which must be achieved.^{6,7}

To achieve that water removal in the presence of water itself is sometimes very difficult. However, we know that it can be done if the conditions are correct, because that is what occurs today in every living organism. Proteins, polysaccharides, lipids are all made by such a water-removal process, in the presence of water. Every organism is mostly water, and these processes are taking place in the presence of water itself.

It was necessary to devise a special kind of chemical reaction which would allow the condensation polymerization to take place. That same kind of reaction is required as well for nucleic acid formation, shown in Figure 4. In this case, there are three different places in which the removal of water is required to produce the polynucleotide: between the No. 9 hydrogen of the adenine and the hydroxyl of the ribose to form adenosine; the removal of water between the No. 5' carbon atom of adenosine with the phosphoric acid; and, finally, the removal of water between the second phosphoric acid hydroxyl group and the No. 3' carbon atom of another adenosine to form the polymer. All three reactions must take place for polymer formation.

In order to demonstrate that this dehydration condensation reaction could take place, without the intervention of biological agents, the following experiments were performed: We used a variety of chemical reagents which store the energy of ionizing radiation, or ultraviolet radiation, agents which are formed very readily from methane and ammonia. These are reagents in which the carbon-nitrogen multiple bond is contained (either a double bond as in the tautomer of cyanamide in which one of the hydrogens has moved or a triple bond such as in cyanide ion, HCN).

NUCLEIC ACIDS (3 STAGES) RNA SHOWN - DNA LACKS OH ON 2' POSITION

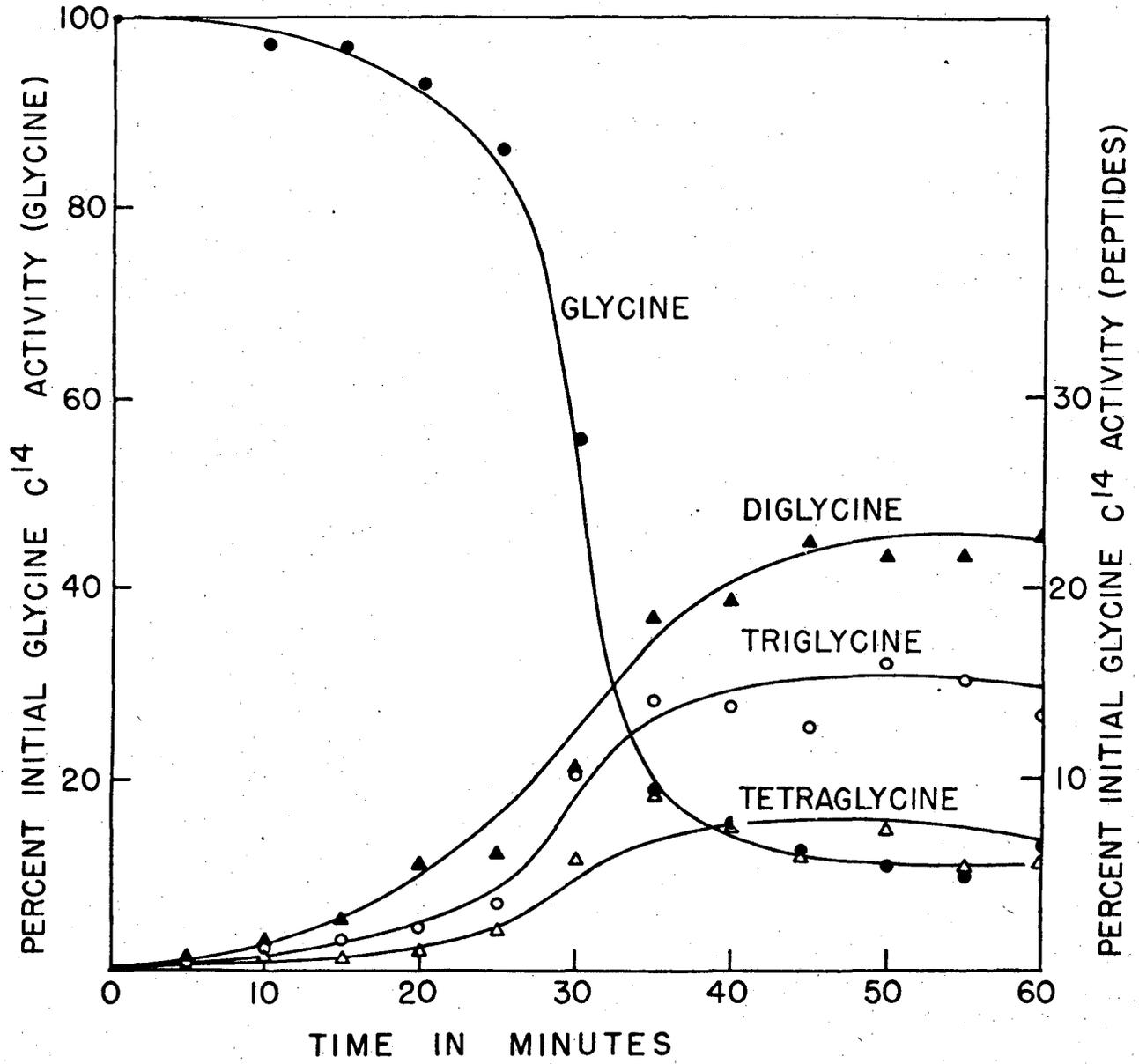


MUB-5262

Calvin. Figure 4

Dehydration condensation of
polynucleotides

These multiple carbon-nitrogen bonds are high-energy storage sources for absorption of water and they do not react very rapidly with water themselves; they react with the sources of water, preferably. We did that type of experiment with glycine, and we were able to make glycine polymers by hooking together two or more of the amino acids to form a peptide. The polymer so formed is one in which the glycine loses a water molecule between the carboxyl group of one molecule and the amino group of another to form a dipeptide. The formation of some polyglycines with cyanamide is shown in Figure 5.⁸ This occurs, of course, in water, and the products are diglycine, triglycine and tetraglycine. We have thus demonstrated that it is possible to hook together two amino acids to form a peptide link, even in water. In fact, that process takes place with some degree of specificity. When one takes a mixture of several amino acids together in the same solution and in the same reaction, it is possible to see a certain selectivity of amino acids for each other, as shown in Table 1.⁹ Already, we are beginning to see some intrinsic selectivity between certain amino acids. The data in Table 1 (from Steinman⁹) show the experimental value for the coupling of glycine-glycine taken as the standard and the relative rates of coupling for each of the other peptides are compared to diglycine. For example, glycy-alanine is only 0.8 as efficient as glycy-glycine and alanyl-glycine is only 0.8; if we go down a little further, valyl-glycine is only 0.5, and so on down the scale, until we reach phenylalanyl-glycine which is 0.1 as rapid and efficient as the coupling of the two glycines. So, there is already, just in the peptide formation itself, some selectivity in the coupling of amino acids with each other. The column on the right of table 1, labeled "calculated", the the value obtained from known proteins. You can see that the two numbers



MUB-8960

Calvin. Figure 5

Homogeneous polypeptide formation by dicyandiamide

Table 1

Comparison of Experimentally Determined Dipeptide Yields and Frequencies Calculated from Known Protein Sequences⁹

Dipeptide*	Values (Relative to Gly-Gly)	
	Experimental	Calculated
Gly-Gly	1.0	1.0
Gly-Ala	0.8	0.7
Ala-Gly	0.8	0.6
Ala-Ala	0.7	0.6
Gly-Val	0.5	0.2
Val-Gly	0.5	0.3
Gly-Leu	0.5	0.3
Leu-Gly	0.5	0.2
Gly-Ile	0.3	0.1
Ile-Gly	0.3	0.1
Gly-Phe	0.1	0.1
Phe-Gly	0.1	0.1

* The dipeptides are listed in terms of increasing volume of the side chains of the constituent residues. Gly, glycine; Alan, alanine; Val, valine; Leu, leucine; Ile, isoleucine; Phe, phenylalanine. Example: Gly-Ala = glycylalanine.

run very close together, suggesting that the original polypeptide which was formed was similar to the ones we have today. By suitable modification, that is, elongation, insertion and other such transformations, we get the variety of proteins which we now have --about a billion different ones.

Another method of hooking amino acids together in water solution was described by Katchalsky several years ago. He made use of the fact that the way in which polypeptides are formed today is via the formation of an amino acyl adenylate. He used these amino acid adenylates with Montmorillonite, a naturally-occurring clay, functioning as a catalyst.¹⁰

Katchalsky was able to demonstrate not only the formation of polypeptides but ones containing 20-40 amino acid units. The structure of the molecules which Katchalsky used, and how they behave, is described diagrammatically in Figure 6.¹¹ Here, the molecule with which we begin is best represented by the second line of the figure, in which you see an amino acyl on the phosphate group of adenylic acid. The reaction was catalyzed by clay and is represented by the small "a" under the words "alternative reactions". Reaction "b" does not take place. Notice that when reaction "a" goes, the result is a polypeptide adenylate. The polypeptide adenylate was the principal product when amino acyl adenylate in water is brought together with the Montmorillonite clay. The possibility of obtaining the amino acyl polyadenylate simultaneously, shown in the lower right-hand part of Figure 6, was also considered, but as yet it has not been demonstrated experimentally. The Montmorillonite clay apparently does not catalyze reaction "b".

Here, also, it was possible to examine the reaction for any specificity which might occur with respect to the combination of one amino acid with another. Paecht-Horowitz, one of Katchalsky's students, per-

formed this same type of reaction with mixed amino acid adenylates to determine if there was any selectivity of one amino acid over another. Table 2 shows the results of that study.¹² The relative yields of the bonds formed are shown in the right-hand column. You can see that when the reaction is done with a mixed amino acid adenylate, for example glycine and alanine, the homopolymer is preferred over the heteropolymer. This fact is true for almost all of the cases which were studied by Paecht-Horowitz. If we examine the heteropolymers themselves, we can see that in some cases there is no distinction as to which way the heteropolymer will go, and in other cases there is a large distinction. The importance of this observation is the fact that it demonstrates that even in a simple polymerization, in which there is no nucleotide template to guide the sequence of amino acids, there is already a selectivity in the order in which the polyamino acids will be formed. This, I believe, indicates that there is some possibility that the earliest protein catalysts (the primitive catalysts) which were formed around the trace metals such as iron were formed prior to the appearance of the information-storing and transferring process represented in Figure 1.¹³

EVOLUTION OF STRUCTURE

Once the polymers have evolved in the prebiotic earth situation, we know that they will have a secondary and tertiary structure which is intrinsic to the primary sequence of amino acids or nucleotides. That secondary structure is a helical one, determined primarily, in the case of polypeptides, by the hydrogen bonding between the amide hydrogen and the

Table 2

Relative Yields of Bonds in the Copolymerization Reactions
of Adenylates of Pairs of Amino Acids

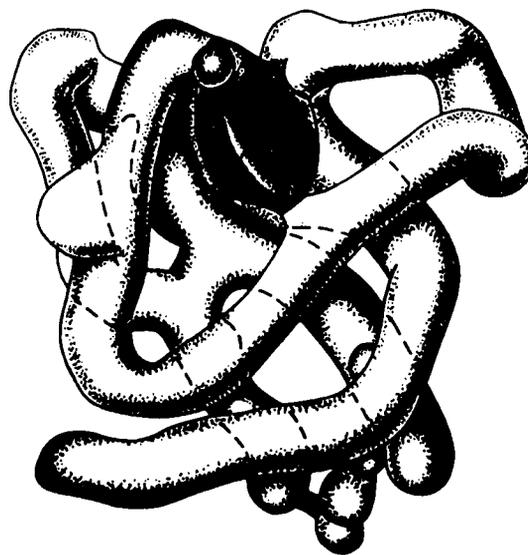
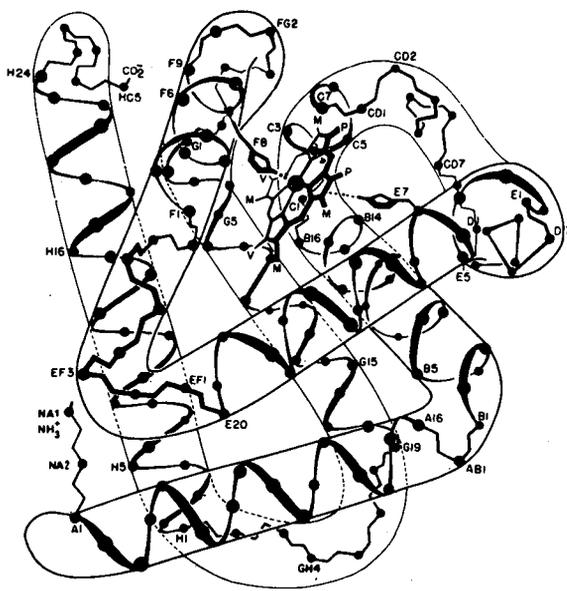
Interacting substances	Bonds	Relative yields of bonds (%)	Interacting substances	Bonds	Relative yields of bonds (%)
Alanine-Adenylate Glycine-Adenylate	Al-Al	40	Aspartyl-Adenylate Glycine-Adenylate	Asp-Asp	55
	Gly-Gly	32		Gly-Gly	21
	Al-Gly	15		Asp-Gly	9
	Gly-Al	13		Gly-Asp	15
Alanine-Adenylate Valine-Adenylate	Al-Al	23	Aspartyl-Adenylate Serine-Adenylate	Asp-Asp	59
	Val-Val	52		Ser-Ser	22
	Al-Val	12		Asp-Ser	10
	Val-Al	13		Ser-Asp	9
Alanine-Adenylate Aspartyl-Adenylate	Al-Al	47	Aspartyl-Adenylate Histidyl-Adenylate	Asp-Asp	36
	Asp-Asp	49		Hist-Hist	44
	Al-Asp	2		Asp-Hist	8
	Asp-Al	2		Hist-Asp	12
Alanine-Adenylate Serine-Adenylate	Al-Al	37			
	Ser-Ser	37			
	Al-Ser	12			
	Ser-Al	14			

M. Paecht-Horowitz, 1974

amide carbonyl, three peptides removed. That is the very familiar alpha-helical structure. In addition, there is a tertiary structure into which the coils are folded in a very specific manner, and this factor is very important for the catalytic and structural function of proteins.¹⁴

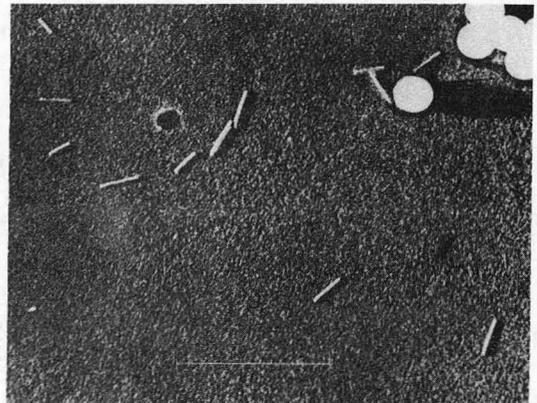
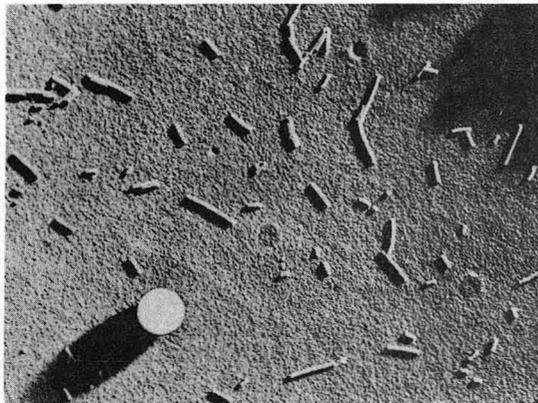
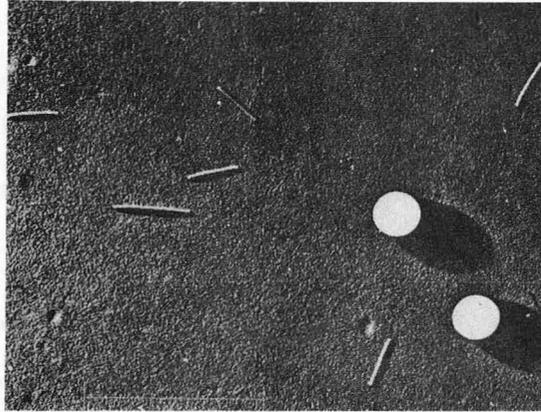
I want to show that the highly convoluted structure of a protein is contained in the primary sequence of the protein itself. Figure 7 shows the structure of myoglobin in two representations of the same molecule, with one of them showing clearly the sequential amino acids. It is easy to see the secondary structure by tracing the dots in the left-hand side of the figure. The tertiary structure, in which the helical parts are folded and bent in an individual and specific way, is best seen in the right-hand side. The complete structure, including the iron heme and represented by the black disc in the upper center of the molecule is contained in the sequence itself.

Higher orders of structure may also arise, and the one which I have used frequently to illustrate the degree to which structure is contained in the sequence is shown in Figure 8. The upper picture is the native tobacco mosaic virus (TMV), just as it is isolated; it has a specific diameter and length and is made up of a single strand of nucleic acid around which are packed one kind of protein molecules, in a group. We can take that virus particle apart to produce a solution containing the protein, on the one hand, and the nucleic acid in another solution, on the other. In doing this, we separate the genetic material (a single strand of RNA) of the virus from the protein molecules. When we then allow the proteins to reaggregate in the solution, they form little rod-like



XBL 727-4704

Calvin. Figure 7
Tertiary structure of myoglobin

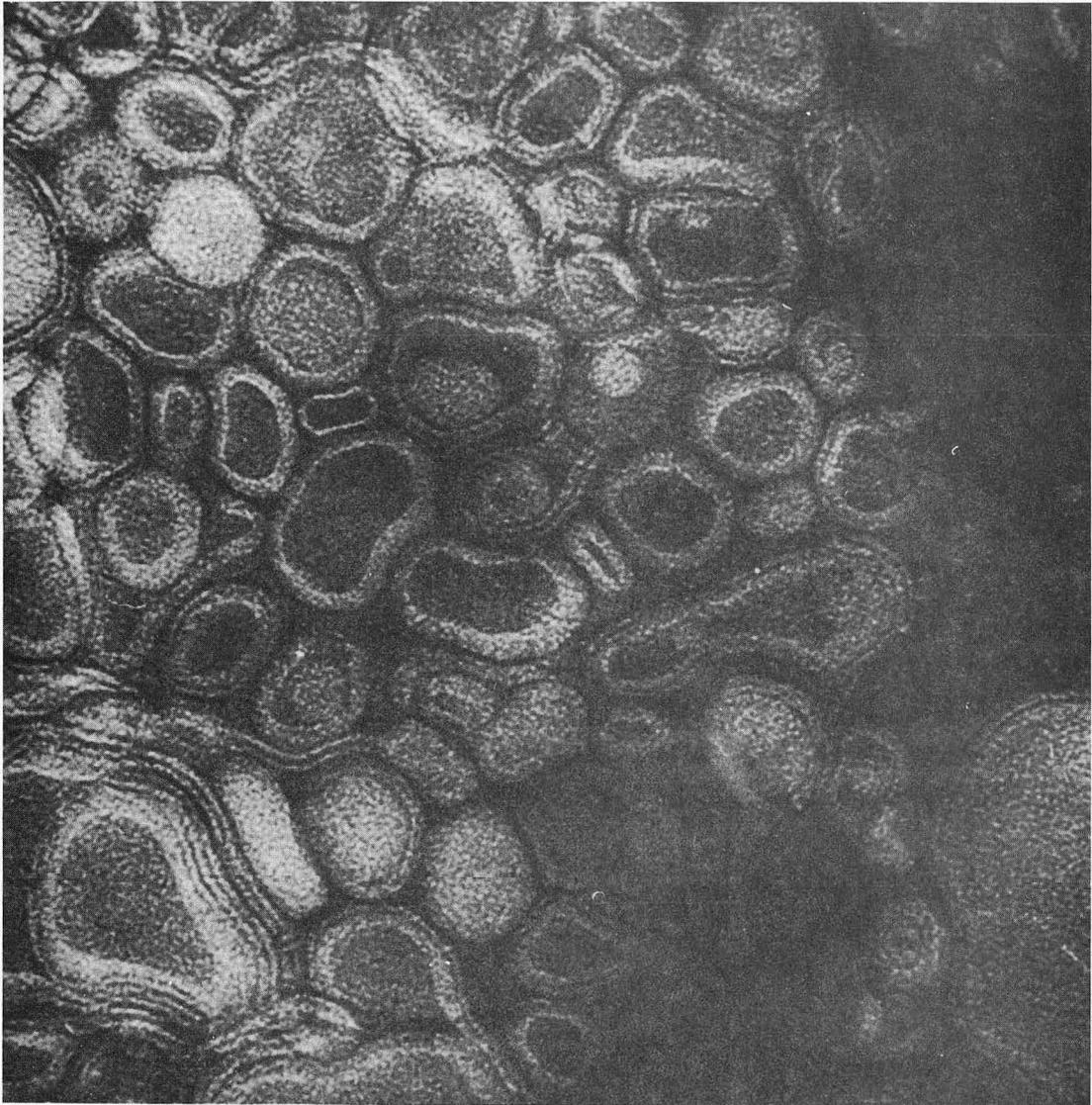


XBB 745-3081

Calvin. Figure 8
Tobacco Mosaic Virus:
Native, protein (reconstituted)
and nucleic acid (reconstituted)

particles of indeterminate length (lower left-hand portion of Figure 8). This, then, is a reaggregation of the TMV protein in the absence of the RNA genetic material of the TMV. If, however, one adds to the protein solution the RNA of the TMV, the intact virus particle is seen again, as shown in the right-hand view of Figure 8. The structure of the molecules (the RNA and the protein) contains all the necessary information to construct spontaneously the structure of tobacco mosaic virus.

I was seeking a reaggregated, or reassembled, cell membrane, but as yet a complete reassembled cell membrane was not available to me. However, a partial synthetic reaggregation of lipid and protein to give a membrane-like structure can be achieved in the laboratory. The characteristics of the membrane structure are contained in the structure of the molecules of which it is composed. Figure 9 shows the construction of a membrane by a mixture of cytochrome and phospholipid molecules.¹⁵ By subjecting this mixture to sonication, vesicles are created which are very reminiscent of cell membranes.¹⁶ Near the center is a single membrane surrounding a protein solution; the others are multiple membranes surrounding protein solutions. Since that figure was made, there have been many cases described in which a reassembly of phospholipids and proteins is achieved to give structural similarity to a cell, but some of the membrane functions are also recovered -- permeability, for example, ion pumping mechanisms, etc. The mechanisms by which selective permeability in the cell membrane is achieved are currently under very close investigation. The structure of the membrane is commonly believed to be a combination of protein and lipid molecules in which the protein molecules are embedded in the bilipid



XBB 729-4501

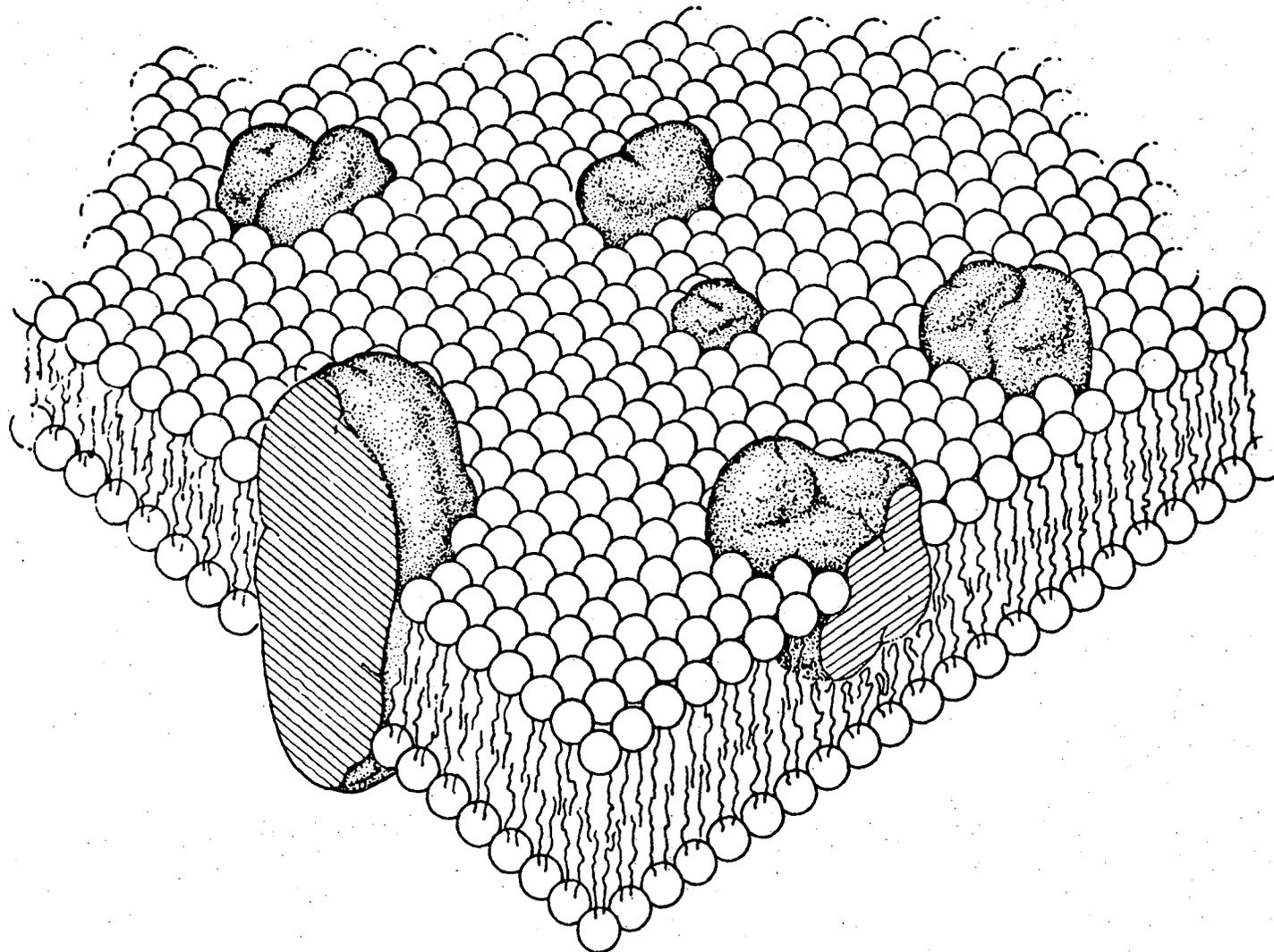
Calvin. Figure 9
Liposomes of phospholipid
with cytochrome (Horne &
Watkins)

membrane, a diagram of which is shown in Figure 10.¹⁷ In this representation of a membrane, the phospholipid molecules are represented by the circles with the fatty tails and the protein molecules are shown embedded in the lipid bilayer. This represents the current concept of the membrane of a living cell. Considerable progress has been made toward reconstructing such active biomembranes -- active in structure but also active in function as well.

Thus, even the highly selective membrane functions of living cells are contained in the structure of the phospholipids and proteins of which they are made. One can reassemble the membranes and get spontaneously formed membranes which have at least some of the structure and function of natural living membranes. We have, thus, passed through our second stage of evolution: We have evolved the biomonomers, the biopolymers and the structure.

EVOLUTION OF CATALYSIS

I would now like to turn to the problem of the evolution of the catalytic function itself. Here, we have a rather interesting development which can be traced in the laboratory as well as in the nature of existing catalysts in the animal and plant world. The basic notion in the evolution of catalysis is the role which reflexive catalysis, or autocatalysis, must play in all of biological evolution.¹⁸ One living organism catalyzes the organization of organic matter into another living organisms, similar to itself. In that sense, all living organisms are simply systems of reflexive catalysts for manufacturing themselves. Autocatalysis as the



The lipid-globular protein mosaic model with a lipid matrix (the fluid mosaic model); schematic three-dimensional and cross-sectional views.

XBL 728-1625

Calvin. Figure 10

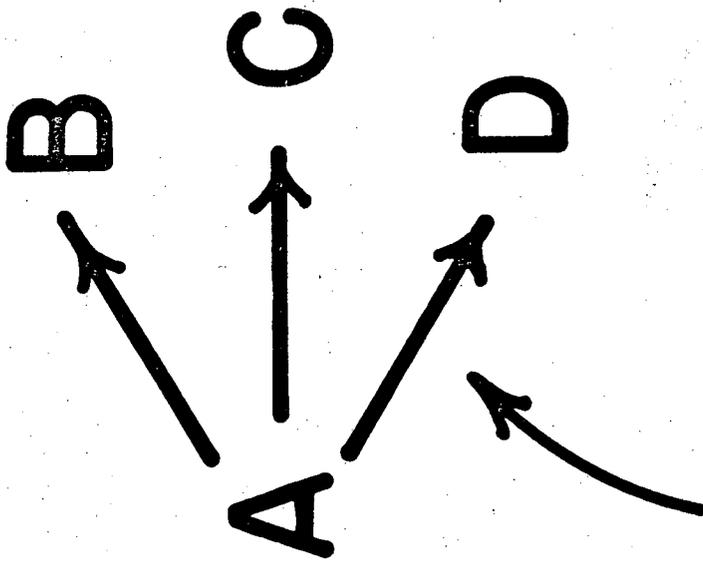
Present concept of membrane structure showing protein embedded in bilipid membrane

chemist normally knows it is simply a single stage of such a complex process which goes on in living organisms.

Figure 11 is a representation of the process of autocatalysis. Substance A is present in a medium and can be transformed either into B, C or D. If D is an autocatalyst it will very quickly transform all of A into D only. This is an example of Darwinian selection for adaptive properties in a particular environment.

The evolution of a catalyst for the reaction of hydrogen peroxide to water and molecular oxygen is a well known example of autocatalysis and is depicted in Figure 12. This illustrates how the catalytic function can be improved with time and selection. It shows what happens to the ability of iron ion to catalyze the decomposition of hydrogen peroxide to water and oxygen. The bare aqueous ferric ion has a catalytic ability represented by 10^{-5} . If, however, that iron ion is incorporated into a porphyrin molecule, such as a heme, the catalytic capability of the iron has been increased by a thousandfold, to 10^{-3} . If, further, one builds that heme into a protein such as catalase, which has a rather specific structure holding the two imidazole groups on either side of the iron, the catalytic function increases once more, to 10^5 . Thus, you can see the primitive catalytic function of the iron which is present in sea water as ferric ion can be evolved by, first, putting it into heme and then into a protein and then the protein itself developing the highly catalytic function of catalase.

The other half of the photodecomposition of water which takes place in photosynthetic organisms is the ability of some organisms to use or evolve hydrogen in various ways. We are just now beginning to learn how



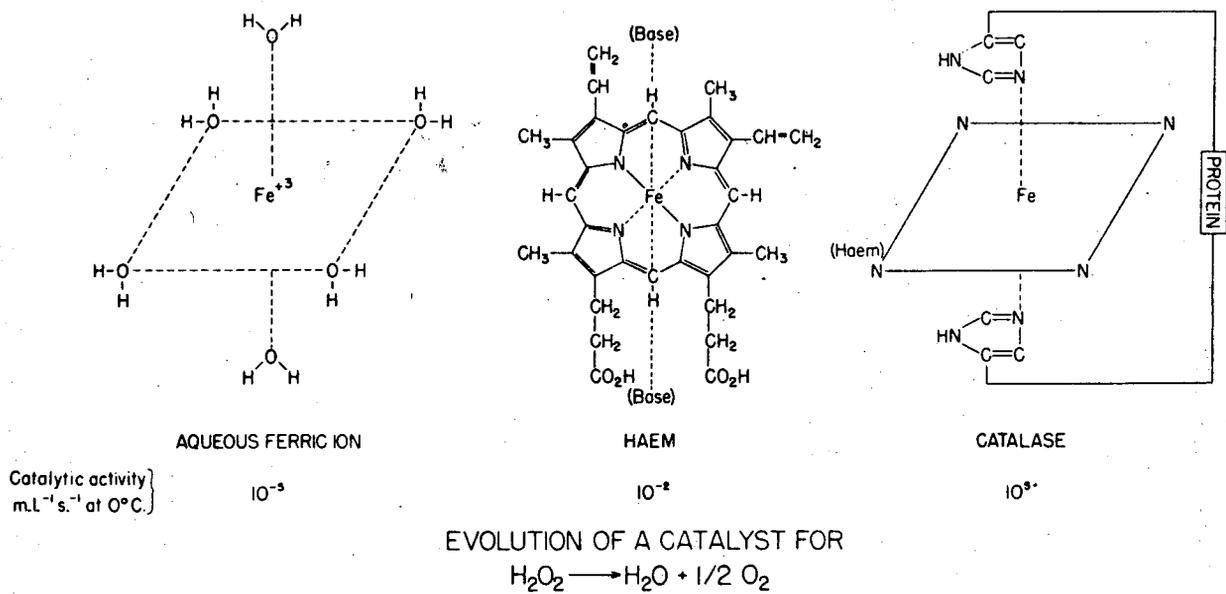
This transformation catalyzed by D

AUTOCATALYSIS

Calvin. Figure 11

Autocatalysis

MU-16090

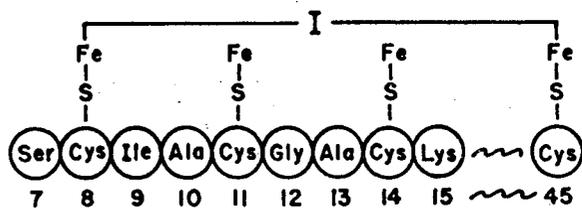
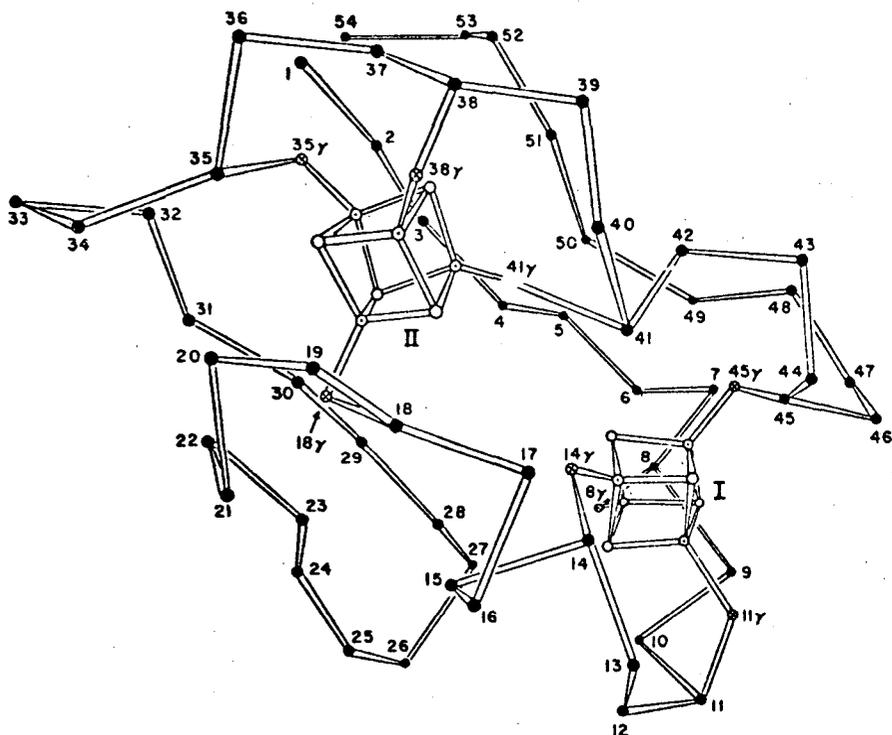


MU-4450

Calvin. Figure 12

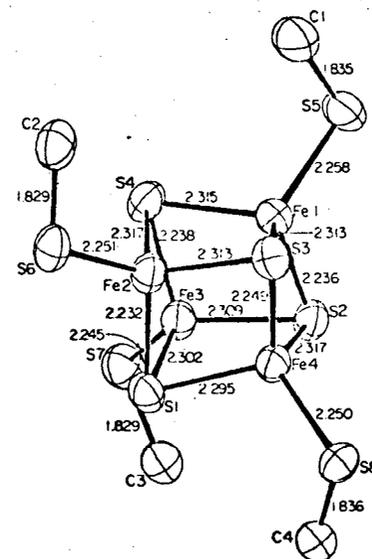
Evolution of a catalyst for reaction of hydrogen peroxide to water and molecular oxygen

we can organize iron into a catalyst which might be capable of reacting with hydrogen, or generating hydrogen.^{19,20,21} There exists in a number of primitive photosynthetic organisms, as well as in some bacteria, a catalyst known as hydrogenase, an enzyme which will catalyze the reaction of hydrogen with a variety of other materials. In some cases, the hydrogenases can be used to evolve hydrogen from the organism, when it is receiving its energy from the light, as in the case of photosynthetic organisms.²² The structure of this type of compound, in which the hydrogen is activated (and in which the nitrogen may be reduced)^{23,24} is very different from the structure of the compound and its iron shown in Figure 12. The one known structure which has been published is that of bacterial ferredoxin, shown in Figure 13 (on the left-hand side).²⁵ Bacterial ferredoxin is made up of a complex protein, containing two distorted cubes of iron and sulfur atoms which have exactly the same structure and they are in two different places in the protein. A synthetic material, very similar to that distorted cube in the bacterial ferredoxin, has been made, using iron and sulfur and benzyl mercaptan. The x-ray structure of that synthetic material is shown on the right-hand side of Figure 13.²⁶ There are four iron atoms, four sulfur atoms and four mercaptides. The important idea to remember here is that these ferredoxins are not involved in oxygen evolution but are involved in hydrogen transport -- the other side of the oxygen scale. Also, these distorted cubes (and things related to them) have been found in single units, in double units (bacterial ferredoxin) in a variety of bacteria and algae. The hydrogenase structure is simpler, containing only two iron atoms instead of four, the structure of which is shown in Figure 14. This structure is merely suggested, and not completely elucidated.²² Figure 14 shows green plant ferredoxin and

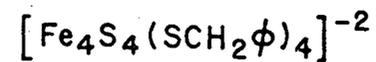


Bacterial Fd

(from Adman et al., 1973)



Distorted cube of



(from Herskovitz et al., 1972)

Calvin. Figure 13

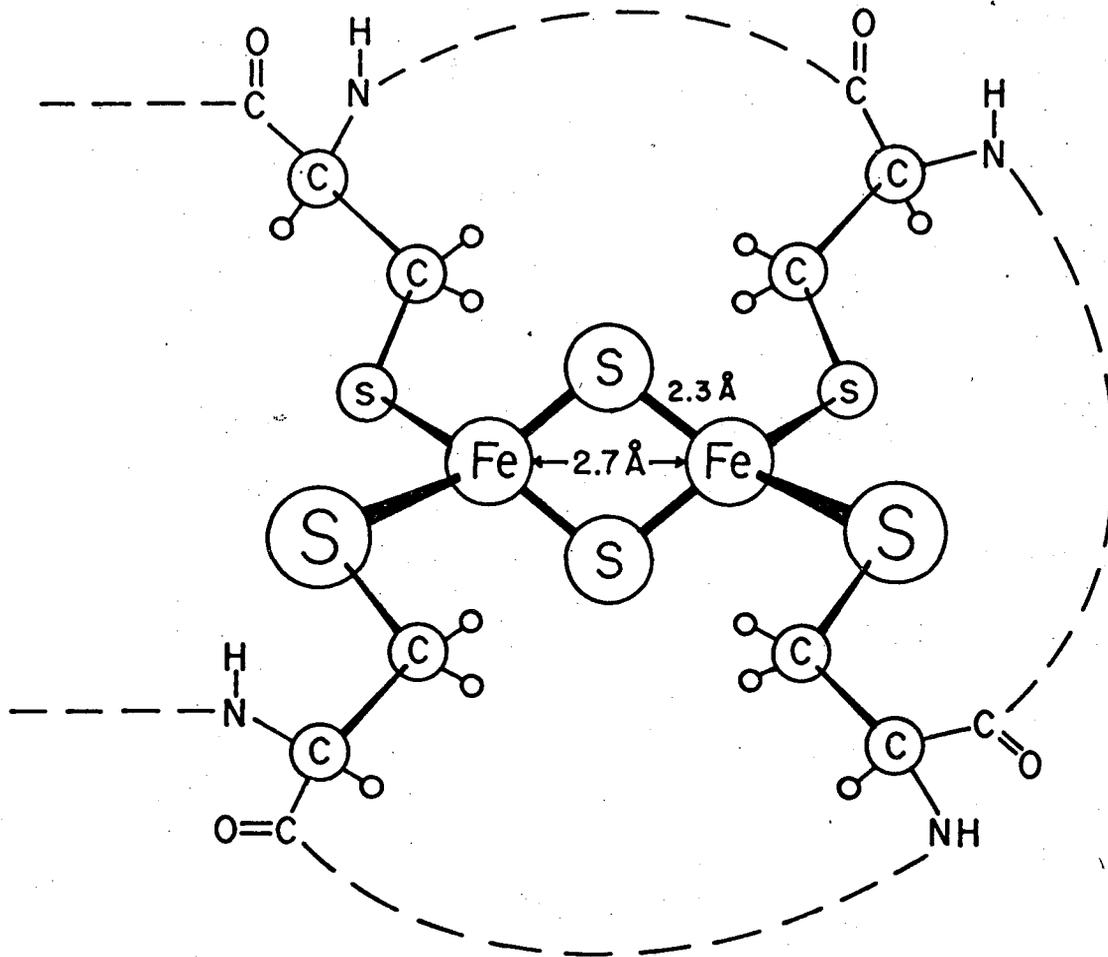
Structure of bacterial
ferredoxin

XBL741-5004

hydrogenase, with two iron atoms, two sulfur atoms and four mercaptides. The structure, which is very similar in properties to that of the known bacterial ferredoxin cube structure, has not yet been demonstrated unequivocally for either green plant ferredoxin or for hydrogenase. However, I feel confident that there is enough circumstantial evidence to say that this will be very close to the structure of green plant ferredoxin and of hydrogenase when those materials are analyzed by x-ray crystallography.

The question is: How could such a structure evolve? It occurred to me to examine the structure of the commonest iron-sulfur mineral, namely, iron pyrite, to see if that iron pyrite could give rise to a dimeric iron-sulfide structure by interaction with hydrogen sulfide of the primitive atmosphere. A possible route by which this process could have occurred is shown in Figure 15. Starting with iron pyrite structure and reacting with hydrogen sulfide (which was present in the primitive atmosphere of the earth) it is possible to form a dimeric structure which, upon further electron and hydrogen transport followed by ligand exchange²⁷, would give rise to a dimeric structure which would be exactly analogous to the structure of the hydrogenase and ferredoxin. It remains for experiments to determine whether such a system as this can indeed operate, and this is where the problem of the evolution of the hydrogen-generating system now stands.

There are many other, more sophisticated examples which could be used in our discussion of autocatalysis. However, the two discussed here show that the nature of the catalytic function is sharply effected by the nature of the surround of the iron atom: In one case, porphyrin and

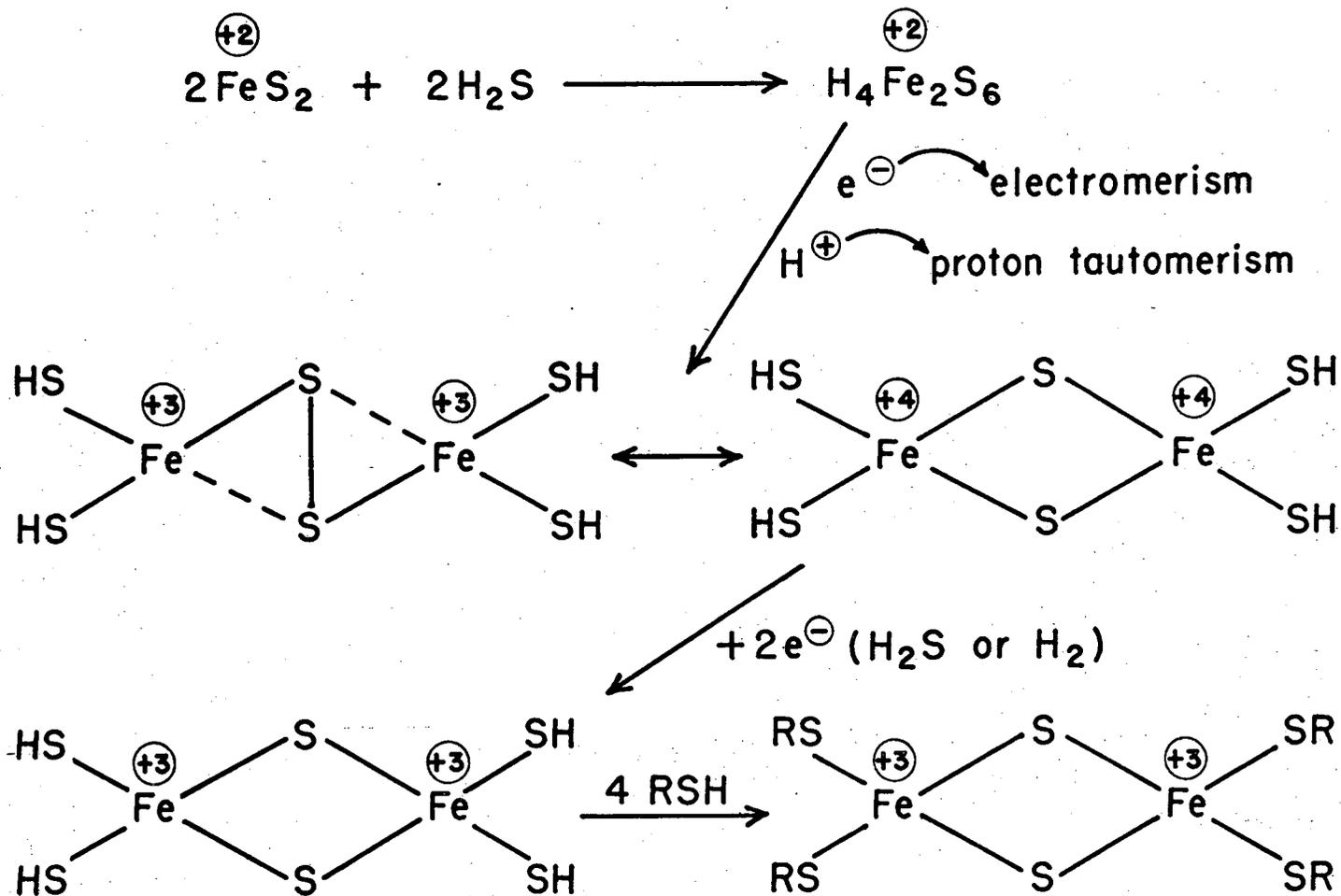


Suggested $\text{Fe}_2\text{S}_2(\text{CyS}^-)_4$ center of green plant Fd
and for hydrogenase

XBL741-5015

Calvin. Figure 14

Proposed structure of plant
ferredoxin and hydrogenase



MODEL FOR ORIGIN OF HYDROGENASE (Fd)

Calvin. Figure 15

Model for origin of hydrogenase

XBL 745-5139

protein is used to produce an iron atom which is involved in oxygen reactions and in the other case sulfur and sulfide are used to produce a set of iron atoms involved in hydrogen reactions.

EVOLUTION OF INFORMATION

This is the last of the evolutionary problems with which we have to deal -- the coupling of the information carrying polynucleotide and the catalytic function carrying polypeptide. How could that have happened? We now know what the code for such coupling is, shown in Figure 16. You will recall that in Figure 1 was depicted how the living cell produces a particular polypeptide and depends on a code in which three bases in a sequence code for a particular amino acid in order to produce a particular amino acid sequence from the polynucleotide which codes for it. For example, three adenine bases in a single sequence code for lysine, as shown in the upper left-hand corner of Figure 16; or, three uridine bases in a sequence code for phenylalanine. The question is: How could that code have evolved? It is clear that the code had some redundancy, some repetitiousness in it, and, in many cases, it looks as though the code started out as a two-letter rather than a three-letter code. For example, two A's in sequence is enough to code for lysine, whether the third is an A or G. However, two A's in sequence is also the code for asparagine, and one would have to distinguish between them by adding a third letter. Thus, the code probably started as a single base code, then a doublet and now a triplet.

How did the coupling evolve? I think this code arose not by accident but because of the peculiar chemistries of the various bases and the

		FIRST LETTER							
		A		C		G		U	
SECOND LETTER	A	AAA Lys	CAA Gln	GAA Glu	UAA Terminate	AAG Lys	CAG Gln	GAG Glu	UAG Terminate
	AAC Asn	CAC His	GAC Asp	UAC Tyr	AAU Asn	CAU His	GAU Asp	UAU Tyr	
	ACA Thr	CCA Pro	GCA Ala	UCA Ser	ACG Thr	CCG Pro	GCG Ala	UCG Ser	
	ACC Thr	CCC Pro	GCC Ala	UCC Ser	ACU Thr	CCU Pro	GCU Ala	UCU Ser	
G	AGA Arg	CGA Arg	GGA Gly	UGA Term.	AGG Arg	CGG Arg	GGG Gly	UGG Trp	
AGC Ser	CGC Arg	GGC Gly	UGC Cys	AGU Ser	CGU Arg	GGU Gly	UGU Cys		
U	AUA Ile	CUA Leu	GUA Val	UUA Leu	AUG Met	CUG Leu	GUG Val	UUG Leu	
AUC Ile	CUC Leu	GUC Val	UUC Phe	AUU Ile	CUU Leu	GUU Val	UUU Phe		

XBL729-4743

Calvin. Figure 16

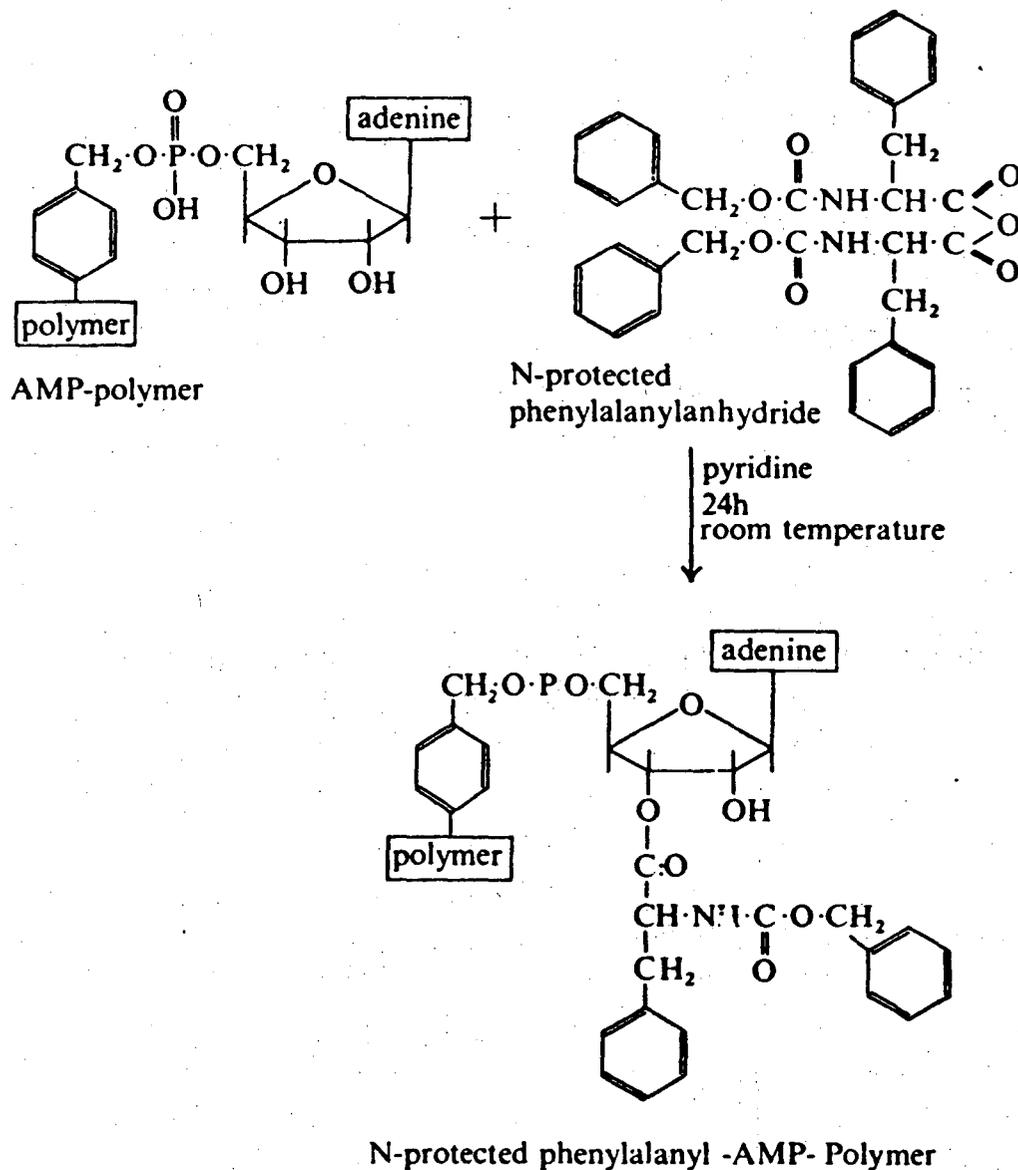
The code

peculiar chemistries of the various amino acids. I do not believe that the origin of the code is a "frozen accident" of some sort.^{28,29,30} I believe it is a consequence of the chemistries of the bases and of the amino acids.

There are two sets of experiments which I can describe which give some idea of how the coupling of information-bearing molecules, such as the polynucleotides, with the catalytic-bearing molecules, such as the polypeptides, might have occurred. In the laboratory, we tried to see if we could find specific relations between some amino acids and some particular bases. The first experiment of this sort was as follows: A particular base (either adenine or cytosine, A or C) was attached to a synthetic polymer and then the relative efficiency, or rate, by which different amino acids reacted with that base attached to a polymer, were measured. This was, if you like, a "model" of a very primitive transfer-RNA.³¹ This was done for two bases and two amino acids. The chemistry of the experiment is given in Figure 17, which shows the reactions used for the coupling of the polymer-AMP complex with the anhydride form of an N-protected amino acid. The percent of bound nucleotide reacted is shown in Table 3.³² We hooked two bases, adenine and cytosine to

Table 3

Percent of Bound Nucleotide Reacted		
Amino Acid/Base	Adenine	Cytosine
Phenylalanine	6.7	2.9
Glycine	10.0	6.5



The coupling of the polymer-AMP complex with the anhydride form of an N-protected amino acid.

XBL 729-4750

Calvin. Figure 17.

Coupling of polymer-AMP complex with the anhydride form of an N-protected amino acid

the polystyrene and then measured the efficiency with which phenylalanine is coupled to adenine and cytosine; we also measured the efficiency with which glycine is coupled to the bound adenine and cytosine. Here, again, there is a difference. We thus have the beginnings of the evidence that even with one amino acid and one base there is a kind of selectivity intrinsic in the structures. We now must explore this idea with larger groups, i.e., two (and more) bases and different amino acids.

A different experiment of this type was devised by one of my former students, Professor C. A. Ponnamperna. Instead of hanging the base on the polymer and measuring the coupling of the free amino acid to the base on the polymer, he did it the other way around. He put the amino acid on the polymer and then he examined the ability of that polymer to non-covalently hold various nucleotides. The results of this experiment are shown in Table 4.³³ Here, the results are with di- and tri-nucleotides, and several different amino acids. It is very interesting to notice the difference with which glycine-bearing polymer adsorbs the various bases and has a greater selectivity for AU than AGU, and this difference also appears with the tryptophan but in the opposite direction. This experiment shows even further the beginnings of selectivity of interaction which is chemical and not an accident.

I am convinced by my chemical background that the code has a chemical base and is not a sheer accident.³⁴ We now have two kinds of evidence that this can indeed be the case, and even some theoretical background.^{35,36} The high degree of selectivity which exists today has yet to be demonstrated, but the rudiments of such a selectivity have now been shown in at least these two preliminary types of experiments.

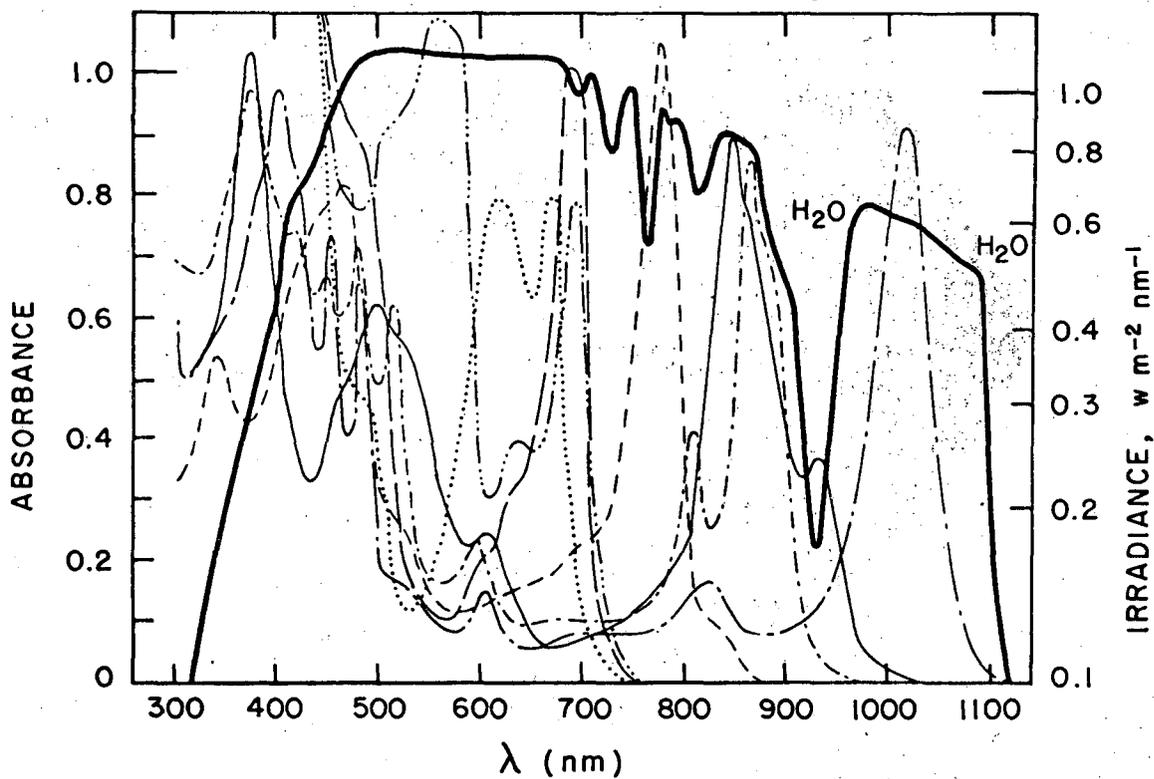
Table 4
Selectivity Coefficients for the Binding of Oligonucleotides to
Immobilized Amino Acids³³

	UpGp	GPUp	ApUp	ApApUp	GpApUp	ApGpUp
gly	10.32	14.05	23.6	63.9	16.6	10.9
	13.44	14.39	27.5	60.0	19.4	13.0
trp	95.1	42.1	187.5	2045	60.4	173
	101.3	56.2	177.6	1817	65.4	197

As a late comment on biological evolution in response to the physical environment, I would like to show how photosynthetic organisms have evolved over the years (4 billion years or so) to fill up all the useful light space available to them. The dark line in Figure 18 is the wavelength of light available to the photosynthetic organisms, and the other lines represent different organisms which have evolved over the years of the existence of photosynthetic organisms. You can see that the biological evolution has filled all of the physical light niches available to it. This is an example which is useful and instructive.

Finally, I would like to show you a picture (Figure 19) which a Dutch artist created about thirty years ago. Here is another way of describing the nature of the evolutionary process in very satisfying human terms. This picture, *Verbum* (as the word in the center says) was drawn by Maurits Escher, a Dutch artist, in 1942. It shows in a graphic way an artist's conception of the unitary character of all life on the earth. Life began with the word "Verbum" in the center, and as the day and the night wore on, as time went forward, expanding outwards, the variety of organisms gradually increased in a slow and steady evolution.

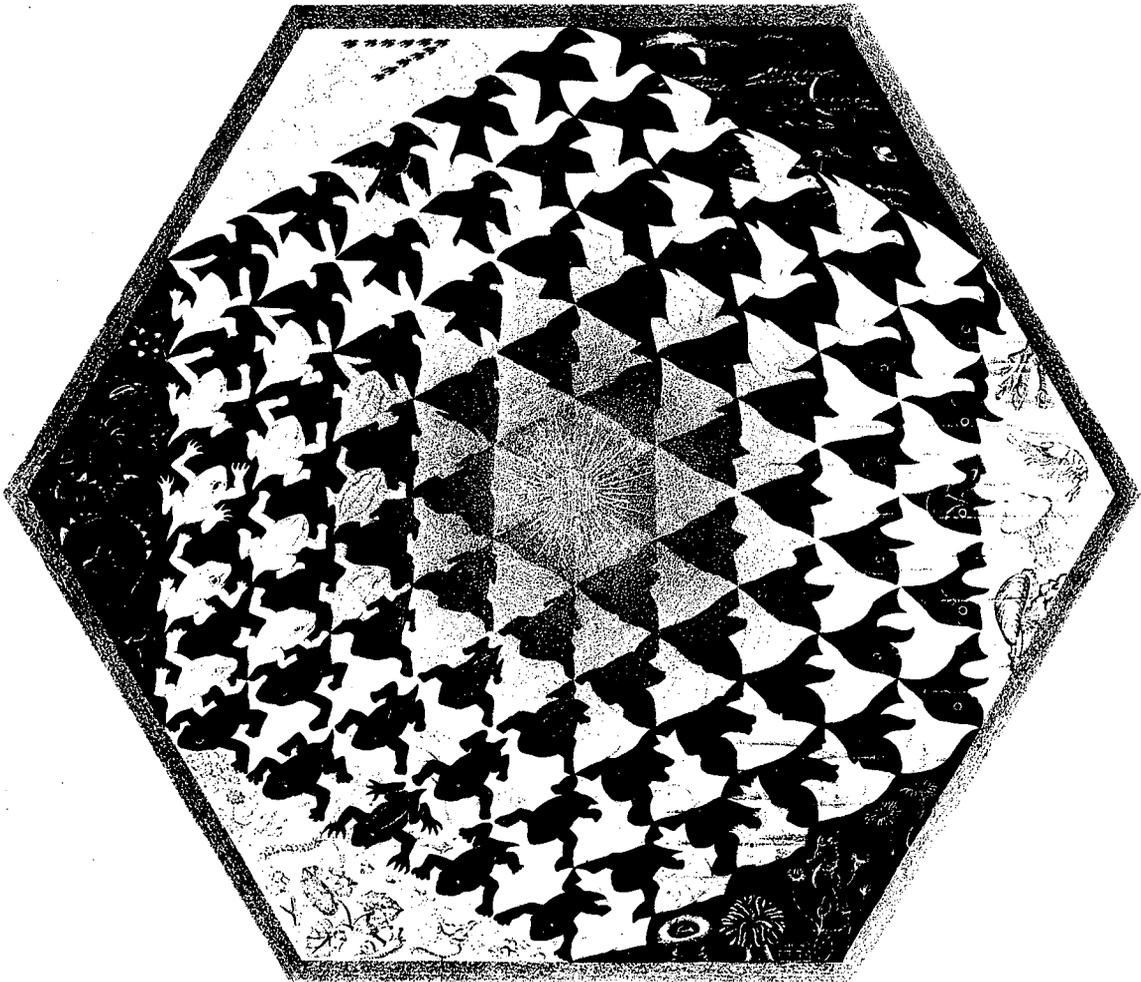
- *Porphyridium* (red alga)
- *Rps. spheroides* (red bac.)
- *Chlorella* (green alga)
- *Synechococcus* (blue-green alga)
- *Rps. viridis* (red bac.)
- *Chromatium vinosum* (red bac.)
- AVAILABLE SOLAR ENERGY
- *Chloropseudomonas ethylica* (green bac.)



XBL723-4565

Calvin. Figure 18

Evolution of photosynthetic organisms showing use of entire solar energy spectrum



MU-23353

Calvin. Figure 19
VERBUM (Escher)

REFERENCES

1. There have been many reviews and books written on the subject of Chemical Evolution. The ones listed below contain overall summaries of the work:
 - a. Melvin Calvin. "Chemical Evolution: Molecular Evolution Towards the Origins of Living Systems on Earth and Elsewhere". The Clarendon Press, Oxford, England (1969).
 - b. Sidney W. Fox and Klaus Dose (eds). "Molecular Evolution and the Origin of Life", W. H. Freeman & Co., San Francisco, California (1973).
 - c. L. E. Orgel. "The Origins of Life: Molecules and Natural Selection". John Wiley & Sons, New York (1973).
2.
 - a. Reference 1a, Chapters 5 and 6.
 - b. R. M. Lemmon, Chem. Rev. 70, 95 (1970).
 - c. R. M. Lemmon, Environ, Biol. & Med. 2, 1 (1973).
3. M. Calvin, Angew. Chem. Internat. Ed. 13, 121 (1974).
4. M. Calvin, La Recherche, #41, Jan. 1974, pp. 44-57.
5. A. R. Hochstim,, Proc. Nat. Acad. Sci. 50, 200 (1963).
6. D. H. Kenyon and Gary Steinman. "Biochemical Predestination", McGraw-Hill Publishing Co., New York (1969).
7. M. Calvin, Reference 1a, pp. 160-170.
8. D. H. Kenyon, G. Steinman and M. Calvin, Biochim. Biophys. Acta, 124, 339 (1966).
9. G. Steinman and M. N. Cole, Proc. Nat. Acad. Sci. 58, 735 (1967).
10. M. Paecht-Horowitz, J. D. Breger and A. Katchalsky, Nature, 228, 636 (1970).
11. M. Paecht-Horowitz, Angew, Chem. Internat. Ed. 11, 798 (1972).
12. M. Paecht-Horowitz, Origins of Life, 5, 173 (1974).
13. Roger Acher, Angew. Chem. Internat. Ed. 13. 186 (1974).
14. M. Calvin, Reference 1a, pp. 187-193.
15. D. Papahadjopoulos and N. Miller, Biochim. Biophys. Acta, 135, 624 (1967); D. Papahadjopoulos and W. Watkins, Biochim. Biophys. Acta, 135, 639 (1967).
16. A. D. Bangham, J. DeGier and G. D. Grevitch, Chem. Phys. Lipids, 1, 225 (1967).

17. S. J. Singer and G. L. Nicholson, *Science*, 175, 721 (1972).
18. M. Calvin, Reference 1a, pp. 145-152.
19. R. H. Wickramasinghe, *Space Life Sciences*, 4, 341 (1973).
20. D. O. Hall, R. Cammack and K. K. Rao, *Space Life Sciences*, 4, 455 (1973).
21. W. H. Orme-Johnson, *Ann. Rev. Biochem.* 42, 159 (1973).
22. For a discussion of this aspect of solar energy conversion, see M. Calvin, *Science*, 184, 375 (1974).
23. G. N. Schrauzer, G. W. Kiefer, K. Tano and P. A. Doemeny, *J. Amer. Chem. Soc.* 96, 641 (1974).
24. R. R. Eady and J. R. Postgate, *Nature*, 249, 805 (1974).
25. E. T. Adman and L. C. Silkin and L. H. Jensen, *J. Biol. Chem.* 248, 4987 (1973).
26. H. Herskovitz, B. A. Averall, R. H. Holm, J. A. Ibers, W. D. Phillips and F. J. Weiher, *Proc. Nat. Acad. Sci.* 69, 2437 (1972).
27. L. Que, Jr., M. A. Bobrik, J. A. Ibers and R. H. Holme, *J. Amer. Chem. Soc.* 96, 4168 (1974).
28. M. Eigen, *Naturwissenschaften*, 58, 465 (1971).
29. F.H.C. Crick, *J. Mol. Biol.* 38, 267 (1968)
30. L. E. Orgel, *M. Mol. Biol.* 38, 381 (1968).
31. M. A. Harpold and M. Calvin, *Nature*, 219, 486 (1968).
32. M. A. Harpold and M. Calvin, *Biochim. Biophys. Acta*, 308, 117 (1973).
33. C. Xainger and C. Ponnampereuma, *Origins of Life*, 5, 189 (1974).
34. M. Calvin, *Proc. Roy. Soc. Edinburgh*, 70, 273 (1969).
35. G. Melcher, *J. Molec. Evol.* 3, 121 (1974).
36. C. W. Carter and J. Kraut, *Proc. Nat. Acad. Sci.* 71, 283 (1974).

LEGAL NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Atomic Energy Commission, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

TECHNICAL INFORMATION DIVISION
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720