UCRL-19733

CHEMICAL EVOLUTION: CONDENSATION REACTIONS OF NUCLEOTIDES AND AMINO ACIDS IN AQUEOUS SOLUTION

Loring Kent Bjornson
(Ph. D. Thesis)

March 1970

AEC Contract No. W-7405-eng-48

LAWRENCE RADIATION LABORATORY
UNIVERSITY of CALIFORNIA BERKELEY

UCRL-19733

PAGES ___i___to___ii___

WERE INTENTIONALLY

LEFT BLANK

TABLE OF CONTENTS

jeg

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

CHEMICAL EVOLUTION: CONDENSATION REACTIONS OF NUCLEOTIDES AND AMINO ACIDS IN AQUEOUS SOLUTIONS

Loring Kent Bjornson

Laboratory of Chemical Biodynamics
Lawrence Radiation Laboratory
University of California
Berkeley

March 1970

ABSTRACT

The condensation reactions in aqueous solution of several nucleotides and amino acids have been studied since it is believed that this type of reaction was important in the formation of polypeptides and polynucleotides on the prebiotic Earth.

A number of condensing agents were examined to determine which were the most effective in reactions with nucleotides. The condensing agents included four water-soluble carbodiimides, cyanamide and two of its derivatives, and cyanoacetylene. In a model reaction of AMP-5' and ethanol it was found that only the carbodiimides gave satisfactory yields of the ethyl phosphate ester under mild conditions. AMP-5' reacts with the carbodiimides to form diadenosine pyrophosphate, $A^{5'}ppA$. It was found that 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EPC) produced $A^{5'}ppA$ in larger yields than the other carbodiimides and a small amount of pApA was observed. Numerous attempts to demonstrate a catalytic effect for the basic polypeptides on the self-condensation of AMP-5' using the charged condensing agent, EPC, were unsuccessful. This catalytic effect was finally found using a new uncharged carbodiimide, 1,3-bis-(2-methoxyethyl) carbodiimide (MEC), in an acetone-water solvent. In this mixed solvent (acetone 37% by volume), AMP-5' reacts with MEC to give $A^{5'}ppA$ in a 15% yield in the presence of poly-L-Lys-ε-$\overset{+}{N}(CH_3)_3$, whereas in its absence the yield is 8.6%.

The reaction of glycine with MEC gave two major products, the N-glycylurea and diglycine. When the reaction was studied as a function of pH, it was found that the N-acylurea formation increased as the pH increased, whereas the maximum amount of diglycine was formed at pH 2. The reaction of phenylalanine with EPC was studied in the presence of various polynucleotides, basic polypeptides and complexes of both polymers. No catalytic effect was observed in these cases. However, when lysine, in an acetone-water solvent, was reacted with MEC in the presence of polyadenylic acid, the yield of dilysine (2.5%) was found to be three times greater than in its absence.

The catalytic effect that has been demonstrated between the monomer and polymer systems of adenylic acid and lysine is postulated to have been important in the prebiotic formation of these types of polymers.

I. HISTORY OF CHEMICAL EVOLUTION STUDIES

The term "chemical evolution" is used to refer to the chemical and physical events which led to the formation of living systems. Other terms also used to describe these events are protobiochemistry, primordial biogenesis, abiogenic molecular evolution, experimental organic cosmochemistry, or, simply, origin of life studies. People have speculated on the origin of life throughout history but the modern chemical evolution area of study really started around 1950. Oparin[1] and Haldane[2] are independently credited with having proposed that the origin of life on earth was preceded by an immensely long period of "abiogenic molecular evolution." The Oparin-Haldane theory was advanced in the mid 1920's but the crucial experiment in which amino acids were made from a simple gas mixture of methane-ammonia-hydrogen-water was not done until 1953 by S. Miller.[3] Although a similar type of experiment had been done earlier without a nitrogen-containing compound[4] it was the amino acid synthesis that really marked the beginning of a whole new experimental scientific study. The experiments have been designed primarily to test various hypotheses about how chemical evolution could have taken place and what events could have been important in this process.

One of the appealing features of the chemical evolution theory is that it presents a continuity of evolution starting with very simple molecules and proceeding through today's living systems. It is also a logical backwards extrapolation of Darwinian evolution to the point where the first living system appeared and further to the point where only the constituent chemicals of living systems were present. It has been noted[5] that Charles Darwin had also conceived of this possibility in 1871. However, only recently has scientific knowledge and technology advanced to a stage where this problem could be investigated experimentally.

A. General Considerations

In discussing chemical evolution a number of questions arise about the primordial environment, the time span, etc. The exact answers to many of these questions are not known but a general picture has been drawn based on evidence from a variety of scientific disciplines, such as astronomy, geology, chemistry and physics.

The question of the length of time available for chemical evolution has been given an approximate answer of one billion years. This is an estimate of the time between the formation of the earth and the emergence of living systems.

After the arrival of life, chemical evolution would rapidly decrease since the chemicals necessary for it would be used by living organisms as in today's environment. It is currently accepted that the earth was formed by condensation of a stellar dust cloud 4.5-4.8 x 10^9 years ago.[6] Evidence has recently been found for the existence of simple unicellular organisms at least 3.1 x 10^9 years ago[7,8] and multicellular organisms 2-2.5 x 10^9 years ago.[9]

The question of the environment of the primitive earth is very important since most experiments considered relevant to chemical evolution are carried out under conditions which are presumed to be like those of the primitive earth.

There is a general consensus that the early atmosphere contained little or no molecular oxygen. There is some question as to the exact composition of this atmosphere (see Chapter 3 of Ref. 10) with W. Rubey[11] favoring an atmosphere dominated by N_2 and CO_2 and H. Urey[12] favoring CH_4 and NH_3 as the dominant species. H. Holland[13] has proposed a model containing elements of both the above theories wherein he has three different stages of development for the atmosphere. The initial stage is highly reduced and contains CH_4, NH_3 and H_2O, the second stage has N_2, CO_2 and H_2O, and the final stage is oxidized (molecular oxygen is present). The third stage is like our present atmosphere and is thought to take place after living systems are formed, especially photosynthetic systems which evolve oxygen. Probably the most compelling evidence for the early reduced atmosphere is the fact that when energy in the form of ionizing radiation or heat is put into a mixture of the reduced gases (CH_4, NH_3 and H_2O) many biologically important compounds are formed, but when the conditions become oxidizing the organic synthesis stops.[14]

The sources of energy available on the primitive earth have been categorized by Miller and Urey.[15] They conclude that by far the largest source of energy was the sun. This is not surprising since life on our planet today is dependent on the sun for its energy ultimately. Other sources of energy in their estimated order of abundance are electric discharges, radioactivity, volcanoes and cosmic rays. The CH_4-NH_3-H_2O mixtures show little UV absorption above 2000 Å and since there is considerable technical difficulty in using lower wavelength light most experimental work has simulated the other types of energy sources.

Life is generally assumed to have arisen in the sea. Certainly water is necessary for all forms of life both aerobic and anaerobic, and it constitutes the major part of all organisms. Thus, in chemical evolution studies most experimental approaches emphasize the importance of water either directly or indirectly. For example, in the irradiation of the reduced gases water serves as a means for concentrating the products formed and protecting them from degradation. Also, the polymerization of the biomonomers in dilute aqueous solution with simple condensing agents has been studied by several investigators and shows considerable promise as a prominent event in chemical evolution. The pH, temperature and salt concentration of the primitive ocean are not known.

It is usually assumed that these variables are not greatly different than in the present oceans except for probably a lower salt concentration in the primitive oceans. It should be emphasized that there probably existed a great variety of microenvironments on the primordial earth since the earth presently contains locales which differ greatly in the above-mentioned variables. It is interesting to note that organisms exist today in surprisingly different environments. Different organisms have been reported[16] to exist at temperatures as low as -18°C and as high as 104°C, at pH values as low as 0 and as high as 13, and salt concentrations of essentially zero and as high as in saturated brines. These facts taken together with the fact that many biomonomers can be produced under a wide variety of conditions seem to indicate that the environmental requirements for the origin of life may not be particularly rigid. This is to say that the origin of life may not have come about as the result of an improbable event where everything had to be exactly so, but rather may have come about as the result of many probable events where the exact conditions were not so crucial.

The current picture of the primitive earth then is one where the atmosphere is reduced, the temperatures are moderate and there are water and land masses existing somewhat like there are today. Chemical evolution may be divided into three different stages of development (see fig. 1).

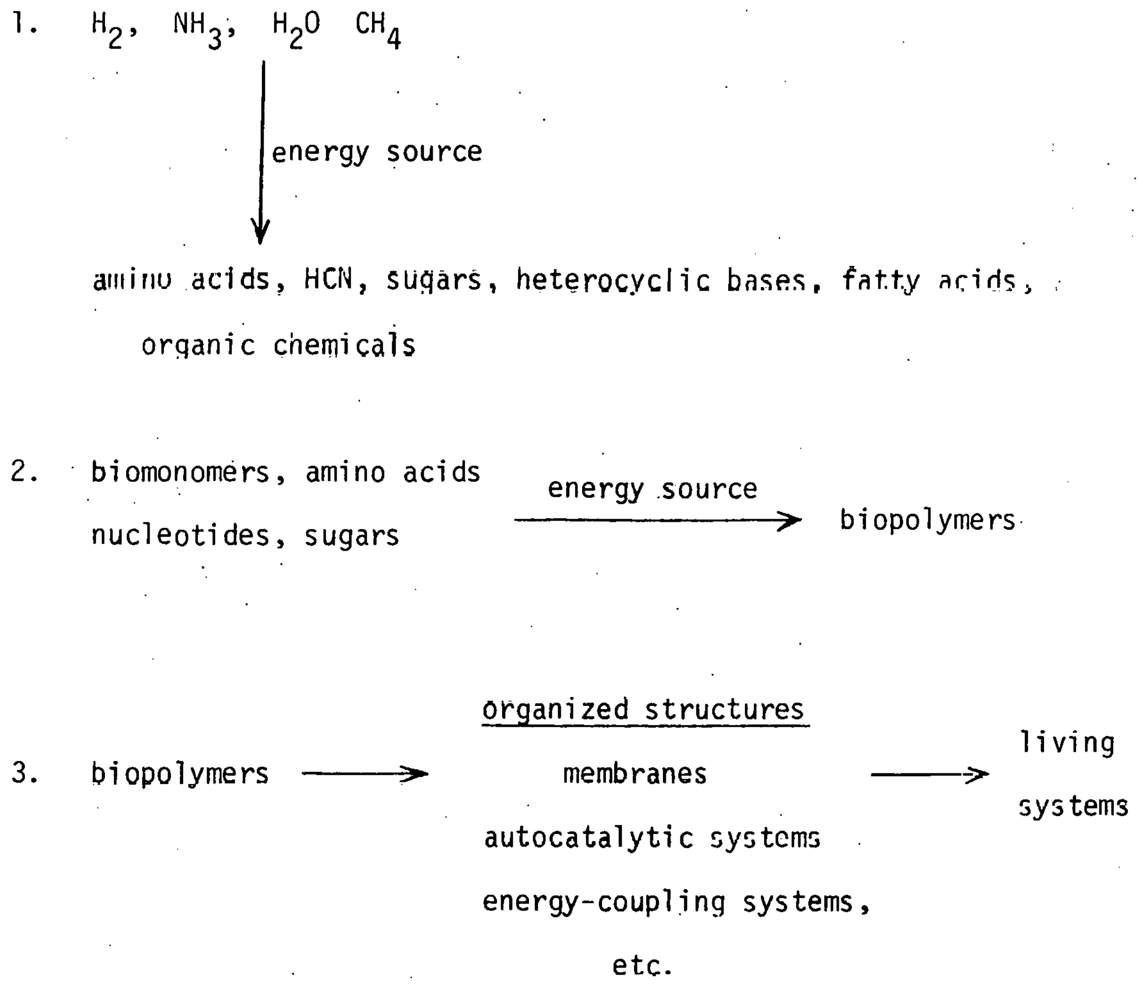


Figure 1. Major steps in chemical evolution.

This is an oversimplified scheme of events but it serves as a model to illustrate the various experimental approaches to the problem. Step one has been examined experimentally in the most detail with the interesting result that given the initial conditions it would have been difficult if not impossible to prevent the formation of the biomonomers. Step two has also been examined in some detail, as will be discussed later, whereas step three which is the most complex has received the least attention. The following sections contain a brief review and discussion of the most important experimental work on the three stages of chemical evolution.

B. Abiogenic Synthesis of Biomonomers

The biomonomers which will be presented for discussion are the amino acids, heterocyclic bases of the nucleic acids, nucleosides, nucleotides sugars, fats and porphyrins. These compounds have been synthesized by two types of reactions. The first of these types will be called a primary reaction and is usually the simulation of reactions that took place in the primitive atmosphere. This involves an energy source and various combinations of water, methane, ammonia, hydrogen, carbon monoxide, carbon dioxide and nitrogen in the gas phase. The other type of reaction will be called a secondary reaction and this involves reacting simple organic compounds that have been formed from the primary reactions. An example of this would be the formation of adenine from hydrogen cyanide and ammonium hydroxide. Of the simple organic compounds formed in the primary reactions such as hydrogen cyanide, formaldehyde, formic acid, acetic acid, etc., hydrogen cyanide is usually formed in the largest yield and is thought to have been very important in chemical evolution. Not only have a number of biomonomers been formed from hydrogen cyanide, but it was proposed earlier (Ref. 5, p. 165) that it could function as a condensing agent in polymerizations. Although this has not been demonstrated with hydrogen cyanide, several derivatives of it have been shown to promote condensation reactions.

1. Amino Acids

Of all the biomonomers the amino acids have been produced the easiest and under the most widely varying conditions. The amino acids which are referred to here are the α-amino acids which are common to all living systems now (p. 10, Ref. 17). The twenty or so amino acids which are the building blocks of protein are all of the L-configuration. Although the D-amino acids do not occur naturally in proteins they are widely distributed in living tissue, particularly as constituents in bacterial cell walls (p. 15, Ref. 17). The amino acids synthesized under simulated primitive earth conditions are D,L-mixtures and the origin of optical activity is yet another unsolved problem.

Since the earliest synthesis of amino acids under simulated primitive earth conditions (electric spark and CH_4, NH_3, H_2O and H_2) by Miller,[3] numerous investigators have repeated and extended this type of work (for a comprehensive review see Ref. 18). Nearly all of the biologically important amino acids have been produced from various simulation experiments.

Using the most highly reduced gases CH_4, NH_3 and H_2O (with or without H_2 since this is generated from the other reactants) and an electric discharge, a number of amino acids have been formed. These include glycine,[19-22] alanine,[19-22] aspartic acid,[19-21] glutamic acid,[19-21] leucine,[21] isoleucine,[21] lysine[20,21] and serine.[21] Amino acids have also been produced from this type of experiment when the gases are in a higher oxidation state, i.e., CO, CO_2, N_2, H_2 and H_2O.[23,24]

Using UV radiation as an energy source, glycine, alanine and α-aminobutyric acid have been synthesized[25] from the reduced gases, but only when ethane was substituted for methane. Ionizing radiation in the form of electrons has also been used to produce the amino acids.[26]

When the reduced gases are passed through a heated column (around 1000°C) of quartz sand or silica gel, a large number of amino acids are formed including such complex ones as tyrosine and phenylalanine, which had not been produced previously.[27]

Most of the amino acids have also been formed from secondary type simulation reactions using simple organic compounds as starting materials. Thus, the action of sunlight on paraformaldehyde, H_2O, NO_3^- and Fe^{+3}, and the action of UV radiation on aqueous solutions of H_2NOH, HCHO and CO_2,[29] HCHO, NH_4Cl and NH_4NO_3,[20] or HCN[30] (followed by hydrolysis), results in the synthesis of amino acids. Success has also been reported using ionizing radiation (γ-rays[31] or electron beam[32]) and by the moderate heating of aqueous solutions of NH_4CN[33,34] or HCHO and H_2NOH.[35]

These experiments are illustrative of the remarkably wide variety of conditions under which the amino acids are formed. Thus, in spite of our lack of knowledge concerning the exact conditions that existed on the primitive earth, it seems almost certain that there would have been many amino acids present.

2. Heterocyclic Bases of Nucleic Acids

The basic building blocks of the nucleic acids are the nucleotides. These in turn are composed of a heterocyclic base (either purine or pyrimidine), a sugar moiety and a phosphate group.[17] There are two main types of nucleic acids termed deoxyribonucleic acids (DNA) and ribonucleic acids (RNA). The main differences between the two are that in DNA the sugar moiety is a 2-deoxyribose and thymine is found instead of uracil. In RNA the sugar moiety is D-ribose and uracil is present along with three other bases, adenine, cytosine and guanine which are common to both types of nucleic acids. There exists one type of RNA called transfer RNA where a number of other bases are found but these are usually derivatives of the above-mentioned bases.

Adenine is the only base that has been synthesized from the reduced gases[36] and it appears to be the most readily formed. Adenine, which is a pentamer of HCN, has been produced by heating an aqueous solution of NH_4CN.[37] Adenine and guanine have both been synthesized by the action of UV light on HCN[38] or diaminomaleonitrile.[29] Cytosine has been formed by heating an aqueous solution of cyanoacetylene and ammonium cyanide.[39] Uracil has also been synthesized from simple starting materials,[40,41] but thymine has not.

3. Nucleosides and Sugars

The nucleosides are formed from the heterocyclic bases and ribose or deoxyribose. The formation of sugars (both hexoses and pentoses) has long been known to result from the base catalyzed reation of formaldehyde.[42] Formaldehyde is, of course, one of the products found after the irradiation of the reduced gases[15] and both ribose[38,43] and deoxyribose[38,44] have been formed from it using various catalysts.

However, linking the base to the sugar is apparently more difficult. Adenosine[45] and deoxyadenosine[46] have been synthesized by the action of UV light on a dilute aqueous solution of adenine, phosphoric acid, and ribose or deoxyribose. There has been no report of similar reactions for the other bases as yet.

Another method of synthesizing nucleosides has recently been discovered[53] which involves the formation of the base on the sugar. In this way cytosine has been formed by heating an aqueous solution of ribose, cyanoacetylene, and either cyanamide or cyanogen.[53]

4. Nucleotides

As mentioned previously the nucleotides are the constituent units of the nucleic acids and can be obtained by mild hydrolysis of RNA.

The nucleotides have been synthesized by phosphorylating the nucleosides. One method of accomplishing this involved heating the five nucleosides (adenosine, guanosine, cytidine, uridine, and thymidine) with sodium dihydrogen phosphate.[47] This reaction yields many of the various nucleotides (2'; 3'; 2'-3'-cyclic, and 5'-monophosphates) as well as some dimers. Similar experiments have been done using other inorganic phosphates.[48] Another method of phosphorylating the nucleosides is to react them with polyphosphoric acid.[49] This reaction works well at temperatures in the range of 0-22°C.

5. Fats

The fatty acids or aliphatic acids have been formed by exposure of methane and water to an electric discharge.[50] From this reaction the monocarboxylic acids from C_2 to C_{12} have been identified. It has been known for some time that hydrocarbons with a wide range of molecular weights can be synthesized by the passage of ionizing radiation through methane.[51] Also, ionizing radiation can directly put CO_2 onto an alkane making the fatty acid.[52]

6. Porphyrins

Chlorophyll, hemoglobin, the cytochromes, and other important biological compounds contain the basic porphyrin structure. Thus, the appearance of porphyrins must have been an important step in chemical evolution since they are needed for efficient oxidation and reduction systems.

It has been reported recently that porphyrins are produced from an electric discharge into a CH_4-NH_3-H_2O mixture.[54] Also, since the porphyrins contain four molecules of pyrrole it should be mentioned that pyrrolic compounds have been produced by the UV irradiation of aqueous solutions of δ-aminolevulinic acid.[55]

The preceding reports of the various simulated primitive earth experiments on the abiogenic synthesis of the biomonomers makes it clear that a number of these compounds can be produced easily and under a variety of conditions. Thus, while not every simple molecule that exists in living cells has been produced in simulation reactions, enough of the most important ones have to make it probable that they all could have been produced on the primordial earth.

Making this assumption the next aspect of chemical evolution to be discussed is the polymerization of the biomonomers.

C. Abiogenic Synthesis of the Bipolymers

The polymerization of the biomonomers to produce the biopolymers must certainly have been a critical event in chemical evolution. The reason for this is that in present living systems the macromolecules are essential for life. Two polymeric systems in particular are thought to be of central importance. These are the polynucleotides and polypeptides. In the polynucleotide system the DNA contains the genetic information necessary for the reproduction of living systems. The DNA is transcribed into RNA (messenger RNA) and the information in the RNA is then translated into the correct amino acid sequences in polypeptides.[56] The polypeptides (proteins and enzymes) are the other class of important macromolecules that are necessary for life processes. For example, the enzymes in cells catalyze all the various biochemical reactions that take place including the above-mentioned translation-transcription sequence.

Thus, most of the experimental work on the synthesis of the biopolymers under simulated primitive earth conditions has centered on the polypeptides and polynucleotides. However, some polysaccharides have been produced. For example, glucose has been polymerized by heating in the presence of HCl[57] or phosphoric acid[58] and by reaction with polyphosphate esters.[59] It should be noted that the polyphosphate esters are very rapidly hydrolyzed by water and their importance as a prebiotic condensing agent is questionable.

It is an interesting fact that all of the biopolymers mentioned can be formed by removal of one molecule of water from the monomeric units (see fig. 2). This type of reaction is called a dehydration condensation reaction. This similarity in the formation of the biopolymers seems to make it likely that they could have been formed under similar conditions, e.g., by the same condensing agent.

1. Polypeptides

Just as the amino acids are the most readily formed of the biomonomers, the polypeptides are the most readily formed of the biopolymers. This fact has been used as evidence that the polypeptides preceded the polynucleotides on the primitive earth.

One of the most successful methods of polymerizing the amino acids has been developed by S. W. Fox and coworkers. They have produced high molecular weight polymers by simply heating to 150-200°C a mixture of all the amino acids under anhydrous conditions.[61] This thermal synthesis has also been carried out at temperatures below 100°C by using polyphosphoric acid.[62] Polypeptides which are basic, acidic, or neutral have been thermally synthesized with the molecular weights ranging from about 3,000 to 9,000.[63] These polymers have been called "proteinoids" because of their similarity to proteins. For example, they give infrared spectra with bands typical of peptide bonds, positive biuret reactions, and the amino acids after hydrolysis. They are also subject to proteolytic enzyme attack and show some weak catalytic activity in the hydrolysis of p-nitrophenyl esters.

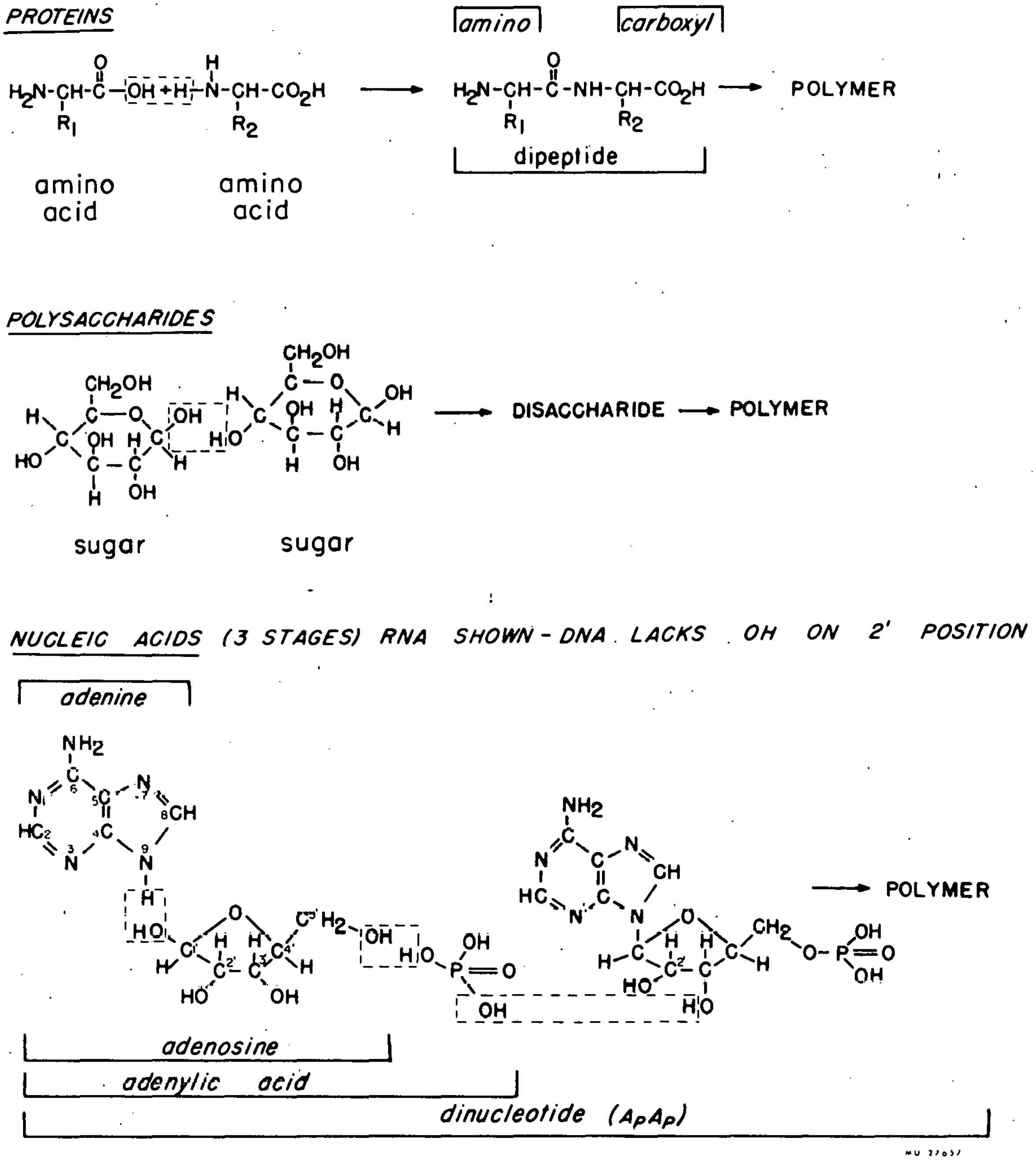


Figure 2. Formation of the biopolymers from biomonomers.[60] The elimination of water is indicated by the dotted rectangles.

The most recent report from this group is the synthesis of polypeptides from the adenylates of the amino acids.[64]

Polypeptides have also been found after passage of an electric discharge through the reduced gases.[21,65] Since polypeptides have been formed from the hydrolysis of HCN polymers[67,68] and HCN is produced very readily from the reduced gases, it has been suggested that this route is pertinent to chemical evolution.[66] Another similar method for forming the polypeptides involves the polymerization of aminoacetonitrile,[69] which itself could be formed from HCN. This would give rise to polyglycine. The other amino acids could be produced in the polymer by reaction of various aldehydes with the active hydrogens in polyglycine.[69]

Some amino acids have been condensed in aqueous solution by reaction with simple organic condensing agents. This type of approach has been successful using cyanamide,[70-72] and several of its derivatives, dicyandiamide[73] and dicyanamide.[74,75] Halmann[72] has studied the cyanamide-induced dimerization of glycine in dilute aqueous solution at pH 6.5. He found that in the presence of orthophosphate diglycine is produced in larger amounts, but concludes that its yield under these conditions (about 1 per cent) is so low that there must have been a more selective condensing agent on the primitive earth. Cyanamide ($H_2N{-}C{\equiv}N$) has been formed in simulation type experiments[76] and is a tautomer of the parent member of the carbodiimide series ($H{-}N{=}C{=}N{-}H$). The carbodiimides have been used by organic chemists for some time to make peptide bonds[77] and the condensation of glycine with dicyanamide has been postulated[74] to proceed through a carbodiimide type mechanism. The cyanamide derivatives work the best at lower pH's. For example, glycine reacted with dicyanamide at pH 1 gives more than 80 per cent of reaction products of which diglycine accounts for almost 30 per cent.[74] This method has been used to produce small peptides (mainly dimers), but, in theory, after long periods of time large polymers could be built up.

It has recently been reported[78] that diglycine can be produced from the reaction of glycine and linear or cyclic polyphosphates in aqueous solution at pH 7-8. The best yields are found using trimetaphosphate, but both tripolyphosphate and pyrophosphate give detectable amounts (about 0.5 per cent) of the dimer. It has been postulated that in this reaction an amino acid reacts with the polyphosphate to give an aminoacyl polyphosphate which then reacts with the amine group of another amino acid to form a peptide bond. This method of synthesis is relevant to chemical evolution since it is carried out at neutral pH with condensing agents that could well have been found on the primitive earth.[70,79,80] Also, there is some similarity between this synthesis and the biochemical synthesis of proteins in cells which proceeds through aminoacyl adenylates (aminoacyl phosphate) formed from the enzymic reaction of adenosine triphosphate (ATP) with the amino acid.[17,81] The pyrophosphate bond in ATP is a prime source of energy in contemporary organisms, but it has been suggested that the earliest organisms may have used polyphosphates or pyrophosphate.[82] Also, polyphosphates are found in many microorganisms.[83]

Peptides have also been reported as products from heating glycine in aqueous NH_4OH[84] or by the action of γ-irradiation on N-acetylglycine.[85] Large polypeptides have been produced from amino acids in the presence of polyphosphate esters.[88]

2. Polynucleotides

As mentioned previously, the nucleotides are more complex than the amino acids. Also, the formation of the polynucleotides (nucleic acids) from the monomers is more difficult to accomplish than the polypeptides both in chemical evolution type syntheses and in the regular organic syntheses. The following reactions are believed to be relevant to chemical evolution.

Cytidine 2'(3') monophosphate (cytidylic acid) has been polymerized by the action of polyphosphoric acid in the absence of water.[86] The average chain length of the oligomer was 5.6. This reaction gave a large proportion of 5'-phosphodiester linkages and some 3'-phosphodiester linkages as demonstrated by the appropriate enzymatic assays. The phosphodiester linkage in RNA is 3'→5'. It seems quite probable that this same method could be used to polymerize the other nucleotides also.

Uridine oligomers have been prepared by simply heating uridine and uridine 2',(3')-monophosphate under anhydrous conditions.[87]

Two other methods have been used to condense nucleotides. Schramm et al.,[59] have produced high molecular weight polynucleotides by reacting the nucleotides with the polyphosphate esters. Oligonucleotides have been synthesized in aqueous solution by reacting the nucleotides with water-soluble carbodiimides.[117-119] From a chemical evolution standpoint both of these reactions have some drawbacks. As mentioned, the polyphosphate esters are extremely reactive with water and the carbodiimides that have been used for these condensation reactions are probably too complex to have existed in any appreciable amount on the primitive earth. Nevertheless, both types of reactions are relevant to chemical evolution since they demonstrate in one case that large polynucleotides can be produced easily from simple starting materials and in the other case that nucleotides can be condensed in aqueous solution. These reactions will be discussed in more detail in the next section since they both have been used to demonstrate the possibility of nonenzymatic catalysis.

D. Interactions and Organization of Biopolymers

The evidence presented in the preceding sections gives us reason to assume that the biomonomers and the biopolymers could have been formed on the prebiotic earth. The next step would be for the biopolymers to interact and organize themselves in such a way as to give rise to the first living systems. This is a very complex step and involves many smaller sub-steps, such as the formation of membrane structures, auto-catalytic systems, energy-coupling reactions, etc. The magnitude of this problem is huge indeed, and it is made more difficult by the fact that there are many details of the biochemistry of cells that are obscure. Perhaps if all the details of the protein synthesizing apparatus or of the photosynthetic apparatus, etc., were known it would be much easier to conceive of how living systems originated. However, as in attacking any very large problem the best approach is to divide it into many smaller ones which have some possibility of being solved. In this way several studies pertinent to chemical evolution have been carried out on non-enzymatic nucleic acid replication and membrane structure. This is an area where chemical evolution studies could conceivably contribute to our understanding of biological processes.

1. Formation of Membrane Structures

Probably one of the first types of organization of the biopolymers was the formation of some kind of membrane structure. This would be necessary for the concentration of the chemicals in the prebiotic ocean since they would be present in dilute amounts. All contemporary organisms have adopted a semi-permeable membrane as a means of protection from their environment and as a means for concentrating certain chemicals needed for life processes. This membrane is usually a double layer of lipid covered on each external side by protein. It has been suggested[89] that such a membrane could arise from the agitation of a protein solution covered by a lipid scum, i.e., that droplets could form by film collapse. Vesicular structures have been formed in the laboratory from solutions of phospholipids and proteins such as cytochrome c.[90] The mechanical disturbance of the lipoprotein layer gives rise to closed vesicles of various sizes, shapes, and dimensions. The boundary layer of these structures is about 60-80 angstroms thick which is approximately the thickness of most naturally occurring lipid layers. These vesicles show some of the properties of natural membranes in that they are osmotically active and selectively permeable.[91,92] This seems to be a plausible method for producing prebiotic membrane structures in the primitive ocean where the wind and waves could have acted as the mechanical disturbances.

It might be noted that inorganic models for semipermeable membranes have also been discovered.[93] Mixing $CuSO_4$ and $K_3Fe(CN)_6$ produces a membrane of $CuFe(CN)_5$ which has been found to be selectively permeable.

Another approach to the problem of how the first membrane structures came into being is the thermal polymerization of the amino acids. The polyamino acids formed from this type of synthesis have been termed proteinoids and they exhibit a number of properties that make them seem quite plausible as a precellular kind of structure.[94] When the proteinoid is dissolved in hot water or hot salt solution and the resulting mixture allowed to cool microspherical particles are formed. These microspheres have a definite shape and are approximately two microns in diameter. They swell and shrink in hypotonic and hypertonic solutions respectively, but not as much as biocells do. The microspheres are quite stable and retain their shape over a period of many weeks or after being centrifuged.[94] These spherules under certain conditions are capable of associating in algae-like chains or of multiplying by a process resembling budding,[95] (they are apparently split in two by a fission process). When zinc is incorporated into the microspheres, they exhibit ATP hydrolyzing activity and finally they also show what appears to be a double layer or "membrane" type structure. Although the microspheres have not yet been shown to function in many ways like contemporary cells their resemblance to cells and their ease of formation is truly remarkable.

A. I. Oparin has been the leading proponent of yet another type of prebiotic membrane structure called a coacervate. The coacervates are particles or droplets formed in solution by associations of various polymers such as polypeptides and polynucleotides, or serum albumin, gum arabic and RNA.[96] One of the main drawbacks of this type of structure is that it is quite unstable and most will not survive even mild centrifugation. Nevertheless, some very interesting results have been reported concerning the properties of the coacervates. These droplets can concentrate both large organic polymers and small monomers either on their surfaces or inside of them. Coacervates have been shown to markedly enhance several enzymatic reactions such as the polymerization of glucose-1-phosphate to starch or of adenosine-5'-diphosphate to polyadenylic acid.[96] Thus, because of their ease of formation and their ability to concentrate organic compounds the coacervates could well have been important at an early stage in chemical evolution.

2. Self-Replicating Systems and the Genetic Code

The central problem of the genetic code[115] is to identify the factors giving rise to the observed codon assignments. One of the key issues is whether or not amino acid-oligonucleotide specific interactions play or have played a role in determining these assignments, and if so, to what extent. In the beginning there were no pertinent facts bearing on this matter, although there was certainly no lack of theoretical speculation. For example, Pauling and Delbruck[97] and Gamow[98] assumed amino acid-nucleic acid steric interactions to be the sole factors determining codon assignments. Crick, on the other hand vigorously challenged this dogma of amino acid-nucleic acid complimentarity, arguing that since neither evidence for such interactions nor a reasonable model for them existed, the interactions themselves did not exist under any circumstances. From this basic tenet, Crick[99] then derived the now famous adapter theory, whereby the code is determined by the cell's set of amino acid activating enzymes and s-RNA's, with the specificity of attachment of amino acids to the s-RNA-residing in the activating enzyme. If there is no possibility of an oligonucleotide's recognizing an amino acid then the assignment of a codon to a particular amino acid would be a chance event, purely a matter of "historical accident," which would make the reason for any particular assignment totally uninteresting.[100]

There have been a number of attempts to detect nucleic acid-amino acid interactions (e.g., by equilibrium dialysis, and ORD measurements), all of which have yielded negative results (except for the obvious charge-charge interactions involving lysine or arginine). Only a few of these results have been published (e.g., Zubay and Doty,[101]), since they have been negative.

Weinstein[102] presented results which are in accord with accumulating evidence that the genetic code, at least in its general features, is the same in bacterial, plant, and animal cells. Since the genes controlling the synthesis of the amino acid-activating enzyme and t-RNA have undergone considerable mutational alterations during the course of evolution, he questions whether natural selection is adequate to prevent species divergence of the genetic code according to a strict interpretation of the adapter theory. Weinstein[102] then proposed, "at the risk of committing heresy," that there exists some as yet unknown stereochemical fit between each amino acid and a group of nucleotides on its corresponding t-RNA. If the amino acid-t-RNA recognition site were the same as the t-RNA-codon recognition site, changes in the genetic code of an organism would be sharply restricted, since mutations which alter the ability of an t-RNA to deliver an amino acid to a particular codon would simultaneously alter the ability of that t-RNA to accept a given amino acid.

Based on the genetic code as published by Nirenberg, et al.,[103] the following correlations between codons and amino acids were deduced:[104] the amino acids with hydrocarbon residues have uridine (U) or cytosine (C) as the second base, those with branched methyl groups U as the second base; the basic and acidic amino acids have adenine (A) or guanine (G) as the second base; the aromatic amino acids and amino acids derived from a common organic acid are grouped together; a simple structural change in an amino acid is accompanied by a change of one base only, such that 16 out of 20 amino acids are interconnected; if the third base of a triplet influences the coding, it does so only by being a purine or a pyrimidine. It seems unlikely that such a logical arrangement could have arisen by chance, and Pelc[104] suggested that a structural relationship between coding triplets and amino acids exists. Similar suggestions have been made by Nirenberg and his collaborators,[103] and Woese.[105] Pelc and Welton[106,107] have tested this hypothesis by constructing models using Courtaulds space-filling components in the usual way. Amino acid molecules were assembled separately, and then fitted into the nucleotide triplets by means of hydrogen bonds, stacking bonds and covalent bonds. Amino acids with hydrocarbon residues possess no sites for the formation of hydrogen bonds on the hydrophobic residue. Thus, it was assumed that a covalent bond between residue and base could be formed by removing a hydrogen from a methyl group on the residue and from a base. This assumption is made somewhat more acceptable by the fact that methylated bases in t-RNA and the required enzymes (t-RNA methylases) have been found.[108,109] However, the acidic or basic amino acids could be fitted by hydrogen bonds alone to their codons in a satisfactory manner, e.g., asparagine can form five hydrogen bonds with its codons, AAU and AAC. Possible bonds to the ribose sugar or linking phosphate were not used. These workers thus found good, though not perfect agreement with the genetic code for most of the amino acids and their codons.

Woese, et al.,[110] have studied the interactions of amino acids and organic bases (pyridine and derivatives) by paper chromatography. They have concluded that organic base-amino acid interactions do exist and comprise two aspects, a polar interaction of ring N's on the bases with the polar portions of the amino acids, plus a nonpolar interaction, between the more organic portions of the bases and amino acids. The amino acids were characterized on the basis of their polar requirements and a correlation was observed between these requirements and the polar characteristics of the respective amino acid codon (assuming U < C < A < G as π electron donors, U << G < A < C as ring nitrogen donors, and U < C < G as oxygen electron donors). These authors,[110] after reviewing other known facts about amino acid-codon relationships, state that "all in all, the conclusion that the genetic code derives basically from some sort of codon-amino acid-pairing interactions, is essentially unavoidable."

Another approach to the origin of the genetic code has been to study the interactions of the mononucleotides with basic polypeptides. In this way, the interactions of the four 5'-mononucleotides with poly-L-lysine have been examined with respect to the amount of turbidity produced.[111] The turbidity increase is probably produced by a two step process which involves first the ionic binding of the nucleotides to the poly-L-lysine chain followed by interchain nucleotide-nucleotide interactions of the bound monomers. The effectiveness with which the nucleotides produce turbidity, GMP > AMP > CMP > UMP, is in the same order as the extent of the self association of the mononucleotides. Using Corey-Pauling three-dimensional atomic models these workers found that when the polyamino-acid is in an α-helix it is possible to associate a complementary strand of mononucleotides with the polyamino-acid. This model has three nucleotides for each amino acid and has been used to theoretically predict that GGG would be a codon for glycine.[111] Thus, it has been postulated that the mononucleotide-polyamino acid interaction must have been the defining one during the prebiotic formation of complementary proteins and polynucleotides.[111] However, because the interactions are very complex and only one amino acid has had its codon predicted much more work needs to be done before the theory can be accepted.

Woese[112] has also found the same type of selectivity in the interactions of the nucleotides with polyarginine, i.e., G > A > C > U. He postulates that the first primitive system was composed of polybasic amino acids and polynucleotides which may have been somewhat like the DNA:histone complexes. In this model translation would be a "direct templating" and it would be a reciprocal matter. However, there are a large number of conceptual hurdles in getting this type of system to evolve into a system that translates in the sense of the modern cell.

Calvin[124] has proposed yet another model for coupling the amino acid and nucleic acid systems. He suggests that there would be some sequence determination in the primordial polypeptides by "growing end control." This idea is supported by the fact that amino acids do react with each other with some selectivity,[125] e.g., glycine will react much more readily with another amino acid than phenylalanine. Also, it has been pointed out that there is a pentapeptide synthesis in bacteria that does not go through a templating mechanism at all.[126] This particular antibiotic synthesis proceeds by growing chain determination with catalyst specificity, i.e., the enzyme will only attach D-glutamic acid to the L-alanine. This sort of synthesis could be a carryover of the manner in which the first proteins might have been generated. The growing end control mechanism would not provide a very accurate means of reproducing polypeptides so that Calvin has proposed that this accuracy could be improved by having the amino acid react first with a nucleotide or oligonucleotide and then be transfered to the polypeptide chain. This would be roughly analogous to the coupling of an amino acid to a t-RNA and the prediction is made that the base of the nucleotide will influence the reaction of it with an amino acid. There is some evidence for this selectivity already and it is predicted that this selectivity will become larger when two or more bases are participating. Thus, in this model both the polypeptides and polynucleotides can replicate themselves with the polypeptides depending on oligonucleotides to enhance their reproducibility. However, whether the selectivity of the reactions of the amino acids with themselves and with the nucleotides is large enough to give the beginning of a coupling system for the two linear polymers remains to be determined.

The origin of the genetic code has been theoretically discussed by Crick[113] and Orgel.[114] Orgel considers two possibilities; life based on proteins in the absence of nucleic acids; and life based on nucleic acids in the absence of proteins. He finds neither system satisfactory since the polypeptides are catalytic but cannot replicate themselves and the polynucleotides could replicate themselves but are not catalytic. However, Orgel believes that nucleic acids are more likely to be capable of behaving as crude catalysts (though there is no hard experimental evidence for this idea) than proteins are of being self-replicative. He bases his hypothesis on the fact that short lengths of nucleic acids are able to fold themselves up to form specific structures. It is this same ability which makes proteins so superbly fitted for the role of catalysts today.

In this system then random polymerization of nucleotides formed short chunks of nucleic acids. These were able to replicate themselves even in the absence of catalysts. Once "replicative doodling" had gone on for some time, some of the "doodles" may have had catalytic properties. From then on, replication would have been much easier and a primitive version of present-day living systems could have developed which would have been capable of improving itself through natural selection.

Crick[113] agrees with this idea and goes on to discuss how such an all-nucleic acid system began to synthesize proteins. Specifically, he considers the development of the informational pathway by which instructions flow from nucleic acids to the proteins through the medium of the genetic code, and the way in which this development determined the nature of the code.

The present system for translating the code involves it first being copied onto messenger RNA, and then carried to the ribosomes (themselves constructed of RNA and protein) where the transfer RNA adapter molecules do the actual job of translation. Transfer RNA's act as the link between each codon and the amino acid that codon specifies. Thus, they have two "working regions," one of which recognizes the codon on the messenger molecule, while the other recognizes the relevant amino acid. The transfer RNA molecules bring each amino acid into the position specified by the messenger so that they can link together in the correct order.

Nucleic acids play a large role in this process, especially RNA, and much of it is employed in roles which make little or no use of its information-carrying capacity. In the present system, several important steps, such as the charging of a transfer RNA with an amino acid need proteins as catalysts. But Crick believes that in the primitive version, RNA might have acted as its own catalyst. Indeed, he postulates that the whole machinery of protein synthesis might once have been carried out by RNA. Also, it is proposed that the code was always a triplet, since if it was once something else (like a doublet), a change in codon size would make nonsense of all previous messages and would almost certainly have been lethal. The reason for a triplet being chosen rather than anything else might simply be a structural one, dictated by the size of the transfer RNA molecules. Next, it seems likely that primitive protein synthesis involved only a few amino acids, and that even these were coded for rather imprecisely (possibly by only two different bases instead of the present four). Probably the whole of the code then available was used, even for these few amino acids, so that each amino acid was coded for by a large number of codons.

At this stage, it is obvious that the "proteins" made under the instructions of this crude code were themselves crude and can have had only a poorly defined structure. The final stage in the evolution of the code consisted of an increase in the precision of recognition and the introduction of new amino acids. Each new amino acid would have needed a new transfer RNA, which was probably derived from an existing transfer RNA and might thus continue to recognize its "old" codon. For this reason it would have been an advantage if the "new" amino acid was like the "old" amino acid, since this would minimize the disruption caused by its insertion into a protein instead of the "old" amino acid. Thus, transfer RNA's with similar codons would tend to carry similar amino acids which is what is observed today.

As this "bootstrap" process continued, more amino acids being made use of and the coding process becoming more sophisticated, the proteins made would have become increasingly complex and increasingly susceptible to further coding changes. At a very early time, when the proteins were crude, the introduction of new amino acids could be tolerated. But the point would come where any further additions would have disrupted too many proteins. At this stage the code was frozen. According to this theory many of the present features of the system like there being 20 amino acids are simply the result of an accident. Crick calls this theory of the code's evolution the "frozen accident" hypothesis. Its' stress all the time is on minimizing the harmful effects of increasing sophistication. In this way he is able to account for similar amino acids having similar codons and for the code's universality. Once codons were assigned in the primordial "organism" from which all life has subsequently evolved, changes would have been too damaging to be accommodated.

But Crick admits that the frozen accident theory is an extreme view. On the other extreme one finds the "stereochemical" theory, which has it that related amino acids possess related codons because they must, i.e., there is a stereochemical relationship between the two. Both Crick and Orgel find this theory somewhat unattractive although they are careful to guard against finding themselves too far out on a limb by stressing that "it is essential to pursue the stereochemical theory." Crick also admits that the frozen accident theory "is too accommodating; in a loose sort of way it can explain anything." However, he apparently feels that this theory is the correct one anyway.

It should be noted that there exists some experimental evidence for the idea that polynucleotides possess the ability to replicate themselves with some degree of specificity without the help of enzymes. Some of the earliest work was reported by Schramm and coworkers[59] who found that in the presence of polyadenylic acid the polymerization of uridylic acid by polyphosphate esters is speeded up tenfold. Other work[116] in aqueous solutions demonstrated that polyadenylic acid catalyzed the carbodiimide coupling of hexathymidylic acid to make a dodecathymidylic acid (an oligomer containing twelve units). The most recent and significant work has been done by Orgel and collaborators. They have demonstrated that polyuridylic acid forms a triple helix with adenylic acid (or adenosine) and that this structure apparently aligns the nucleotide (AMP-5') such that when a water-soluble condensing agent is added a much higher yield of oligoadenylate is realized.[117,118] They have shown that while polyuridylic acid enhances the yield of adenylate products it has little effect for the other nucleosides. Also, polycytidylic acid has no effect on adenylate condensation, but does facilitate both the self-condensation of guanosine-5'-phosphate and the condensation of guanosine-5'-phosphate with guanosine.[119] Thus, at least in these cases studied the replication does proceed with specificity in accord with the Watson-Crick base pairing rules. However, it is apparently not easy to do these experiments in reverse, i.e., to incorporate the pyrimidine nucleosides into oligomers on templates containing purine nucleosides. This is probably due to the fact that organized helices are not formed by pyrimidine nucleoside derivatives with polyadenylic acid or polyguanylic acid. It may be possible to use alternating oligonucleotides to demonstrate complete non-enzymatic replication. Thus, the evidence for the theory that non-enzymatic replication of polynucleotides was an early step in chemical evolution is encouraging even though it is not conclusive as yet.

II. REACTIONS OF NUCLEOTIDES WITH VARIOUS CONDENSING AGENTS

Since the nucleic acids occupy such a central role in biology, their formation on the primitive Earth would be crucial to chemical evolution. Several methods now exist for polymerizing the nucleotides with the most successful ones employing anhydrous or semi-anhydrous conditions. While it seems quite possible that some polymerization of nucleotides could have occurred under anhydrous conditions on the prebiotic Earth, and then have been washed into an ocean or lake by rain water, the problem of further polymerization or replication in aqueous solution still exists. Furthermore, it is conceivable that most of the polymerizations took place in aqueous solution in the first place. In any case it is obvious that it would be important to study the dehydration condensation reactions of the nucleotides in aqueous solution. The reaction of two nucleotides to produce a dinucleotide is not favored thermodynamically since the standard free energy is about +6 Kcal per mole.[56] In contemporary organisms the energy needed for this reaction to proceed is supplied by the energy rich pyrophosphate bonds in adenosine-5'-triphosphate (ATP). However, these reactions with ATP are catalyzed by enzymes and coupled with other biochemical reactions so that at present there has been no report of ATP being used as an energy source in chemical evolution studies. It seems likely that another source of chemical energy was used early in chemical evolution such as polyphosphates, cyanamides or other energy-rich compounds that could be synthesized under primitive-Earth conditions. Thus, the common sources of energy available on the primitive Earth like heat, light and ionizing radiation would be used to produce compounds which by virtue of their chemical bond energies can promote polymerization reactions, etc. The largest problem in carrying out these reactions, e.g., condensing nucleotides, is that one is actually doing a dehydration reaction in the presence of water, which is not favored thermodynamically. There are only a few compounds that have been reported to be able to accomplish this sort of dehydration condensation reaction in aqueous solution. Carbamyl phosphate[120] and several cyanamide derivatives[74,121] are able to accomplish this to some extent.

When the present work was started, one of the few reports of condensation reactions of nucleotides in aqueous solution was that by Naylor and Gilham,[116] where they used special water-soluble carbodiimides to effect these reactions. Thus, one of the first projects that was attempted was to compare the reactions of the nucleotides with the water-soluble carbodiimides and the cyanamide derivatives. However, the direct dimerization of the nucleotides (e.g., deoxyadenosine-5'-monophosphate) had already been unsuccessfully tried using the cyanamide derivatives.[121] This was not too surprising in view of the known unreactivity of secondary alcohols. The linkage of the naturally occurring dinucleotide is 5' to 3', so that in this reaction the 3' (secondary) alcohol of one molecule would react with the 5'-phosphate of another. The polymerization of nucleotides with carbodiimides in nonaqueous solutions has been studied in detail by Khorana and co-workers.[122,123] This reaction goes very well in the organic solvents yielding fairly large oligomers, whereas in aqueous solutions the products are much smaller.

Thus, originally it was thought that the cyanamide derivatives would be able to condense the nucleotides if the nucleotides were aligned on a template so that the reactive ends of each molecule would be very close to each other. One of the reactions that was envisioned was to condense an oligonucleotide with another oligonucleotide where both were templated on a polynucleotide. This is essentially the experiment that Naylor and Gilham[116] carried out with water-soluble carbodiimides, where they demonstrated that polyadenylic acid had to complex with hexathymidylate before the thymidine oligomer could self-condense. Since the cyanamide derivatives are chemically such simpler compounds than the fairly complex water-soluble carbodiimides, they are more relevant to chemical evolution syntheses, i.e., they would be much more likely to have been present in significant amounts on the primitive Earth. However, before trying the reactions with the oligonucleotides, which are difficult to obtain, it was thought to test the reactivity of the cyanamide derivatives in several other standard nucleotide reactions. Two of these reactions are the formation of the 2':3'-cyclic phosphate from the 2' or 3'-mononucleotide, and the reaction of AMP-5' with a primary alcohol. Both of these reactions have been carried out with carbodiimides in both aqueous and non-aqueous solvent systems.

The mechanism of the carbodiimide mediated esterification of the monophosphates is illustrated in Fig. 3 for the reaction of uridine-2',3'-monophosphate (UMP-2',3'). The first step is the addition of OH across the carbon double-bond nitrogen, where the H probably adds to the

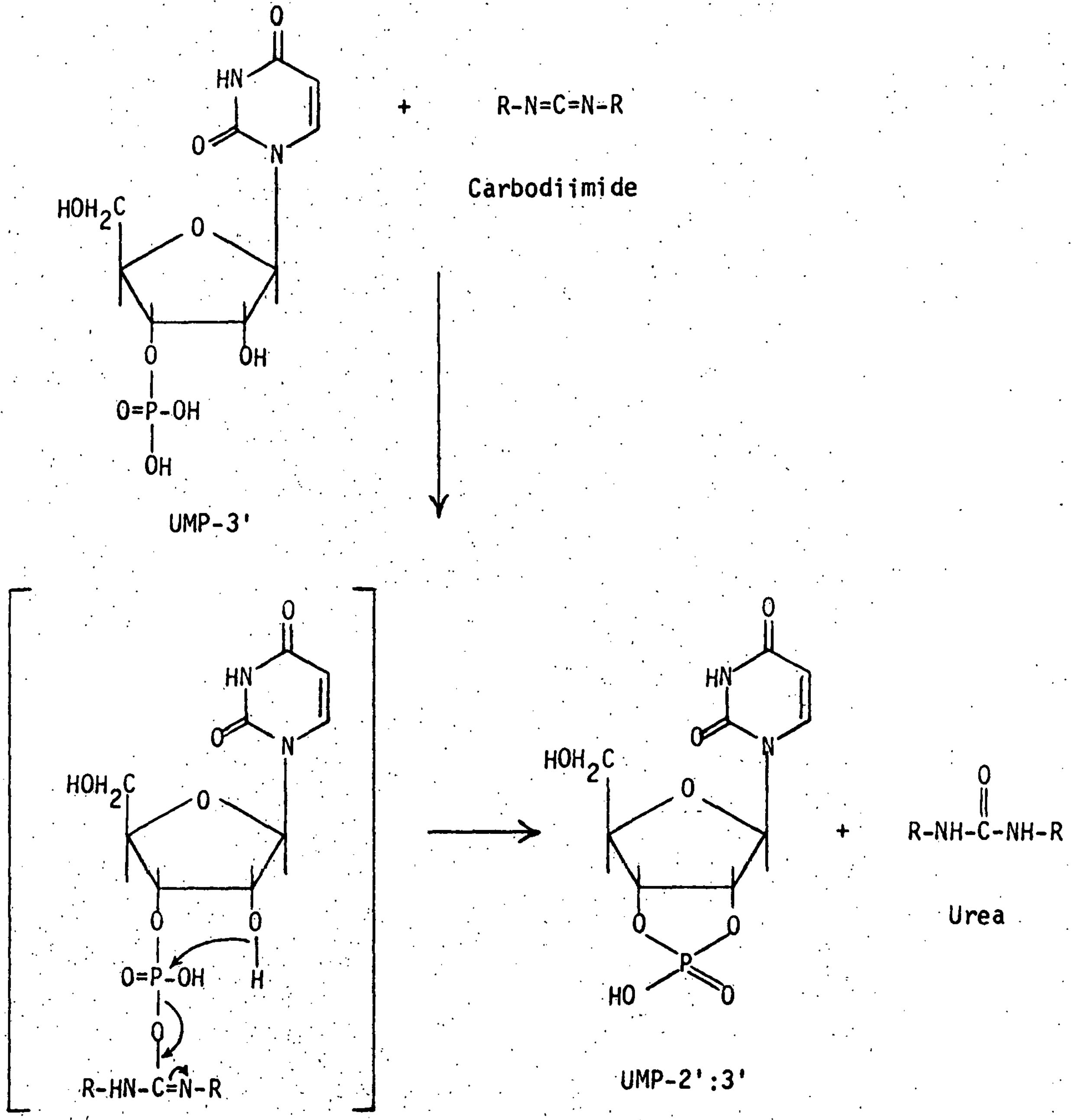


Figure 3. Mechanism of the reaction of UMP-2':3' with a carbodiimide.

nitrogen first. The attack of the phosphoryl oxygen on the carbon of the carbodiimide results in the formation of an o-phosphorylisourea intermediate.[128] This intermediate is highly unstable and has never been isolated in the reactions of phosphates with carbodiimides. The final step in the reaction is the nucleophilic attack of the alcohol on the phosphorous atom which gives the five-membered cyclic phosphate.

In the case of the five membered cyclic phosphate the reaction can proceed further with dicyclohexylcarbodiimide in aqueous pyridine to give the N-phosphoryl urea.[129] This product is not formed when tri-n-butylamine is present or when other carbodiimides are used. Also, the N-phosphorylureas are not observed with the six- or seven-membered cyclic phosphates.

A. Reactions of UMP-2',3'

The reaction of UMP-2',3' with carbodiimides has been studied in detail in organic or partially organic solvents, as has been mentioned.[128,120,130] More recently, this same reaction has been examined in completely aqueous solution using water-soluble carbodiimides,[116] and it was found that the yields of the 2':3'-cyclic phosphate were very high. This is a clear indication that the hydroxyl next to the phosphate is in very close proximity since it competes quite effectively for the activated intermediate against a large excess of water. Sheehan and co-workers[121,132] had first synthesized the acid-soluble and water-soluble carbodiimides for use in peptide synthesis since the by-product ureas and acylureas can be easily removed by washing with dilute acid or water. This is an advantage over the more commonly used dicyclohexyl- and diisopropylcarbodiimides, since the corresponding ureas and acylureas[133] of these compounds frequently have solubility properties similar to the peptides, rendering separation of the products difficult.

The structures, names and abbreviations of the water-soluble carbodiimides and other compounds used as condensing agents are given in Fig. 4. Chronologically, with the carbodiimides, the cyclohexylmorpholinylcarbodiimide (CMC) was first used in this work since it was easily available. This was replaced by the ethyldimethylaminopropylcarbodiimide (EPC) since EPC gave higher yields of reaction products than CMC. The EPC was subsequently displaced in favor of the uncharged water-soluble carbodiimide, 1,3-_bis_-methoxyethylcarbodiimide (MEC) for the study of the interactions of nucleotides and polypeptides, which will be discussed later.

Thus, since the reaction of UMP-2',3' was one of the first reactions studied, the carbodiimide used was CMC. The UMP-2':3' was formed in a 54% yield from the reaction of ^{14}C-UMP-2',3' (0.01 _M_) and CMC (0.05 _M_) at pH 6. The products were separated by paper chromatography and the autoradiogram is depicted in Fig. 5. The UMP-2':3' was identified by co-chromatography with commercially available UMP-3':3' and by its base catalyzed hydrolysis back to UMP-2',3'.

However, under these same conditions, _i.e._, five-fold excess of condensing agent, pH 6, room temperature, and reaction time of about one day, sodium dicyanamide (DCA) produced no observable reaction. Although DCA has been reported to be the most reactive of the cyanamide derivatives,[74] this result is not too surprising since these derivatives react best at low pH's. Unfortunately, this particular reaction could not be easily run at lower pH's since the five-membered cyclic phosphate is

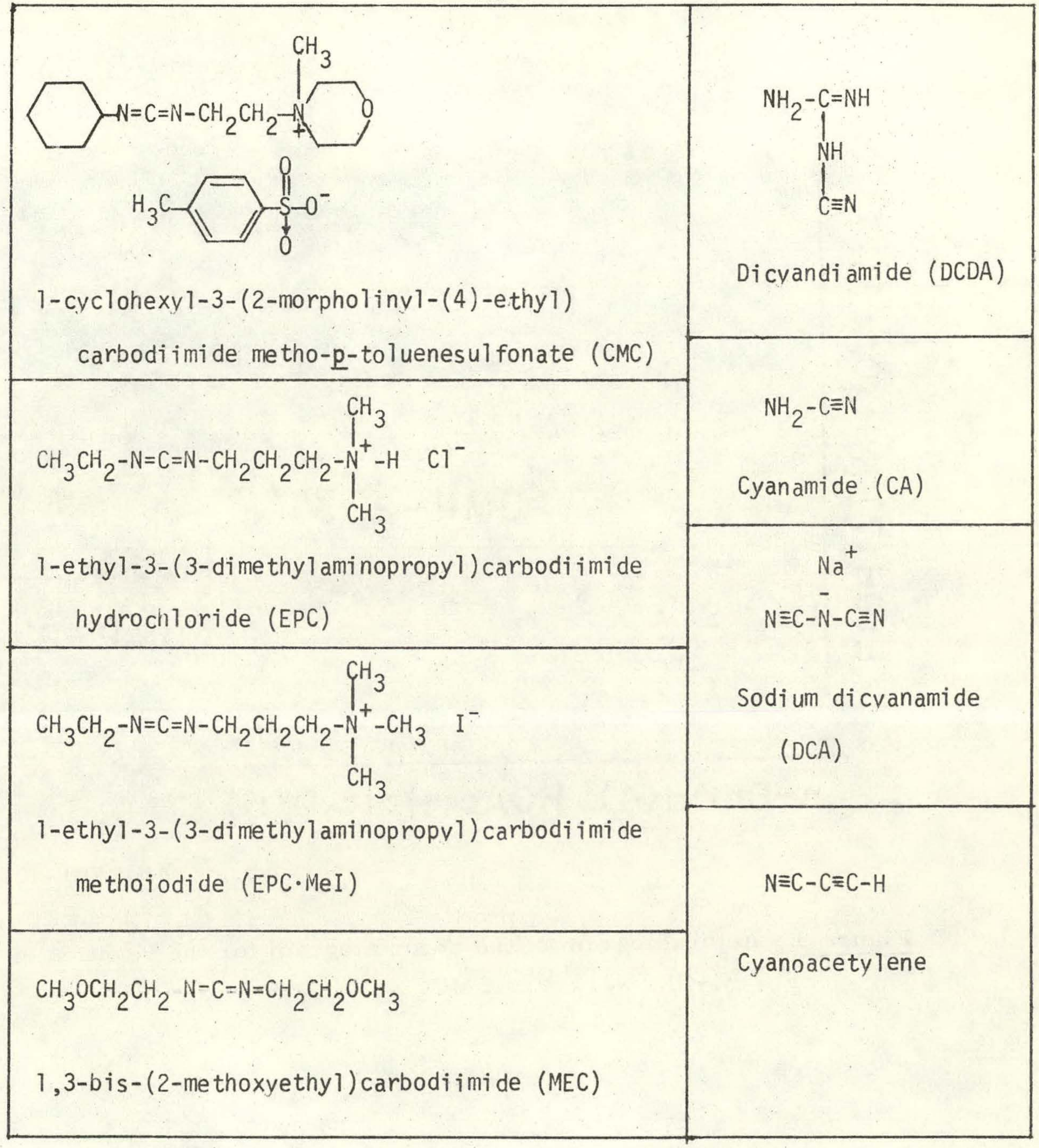
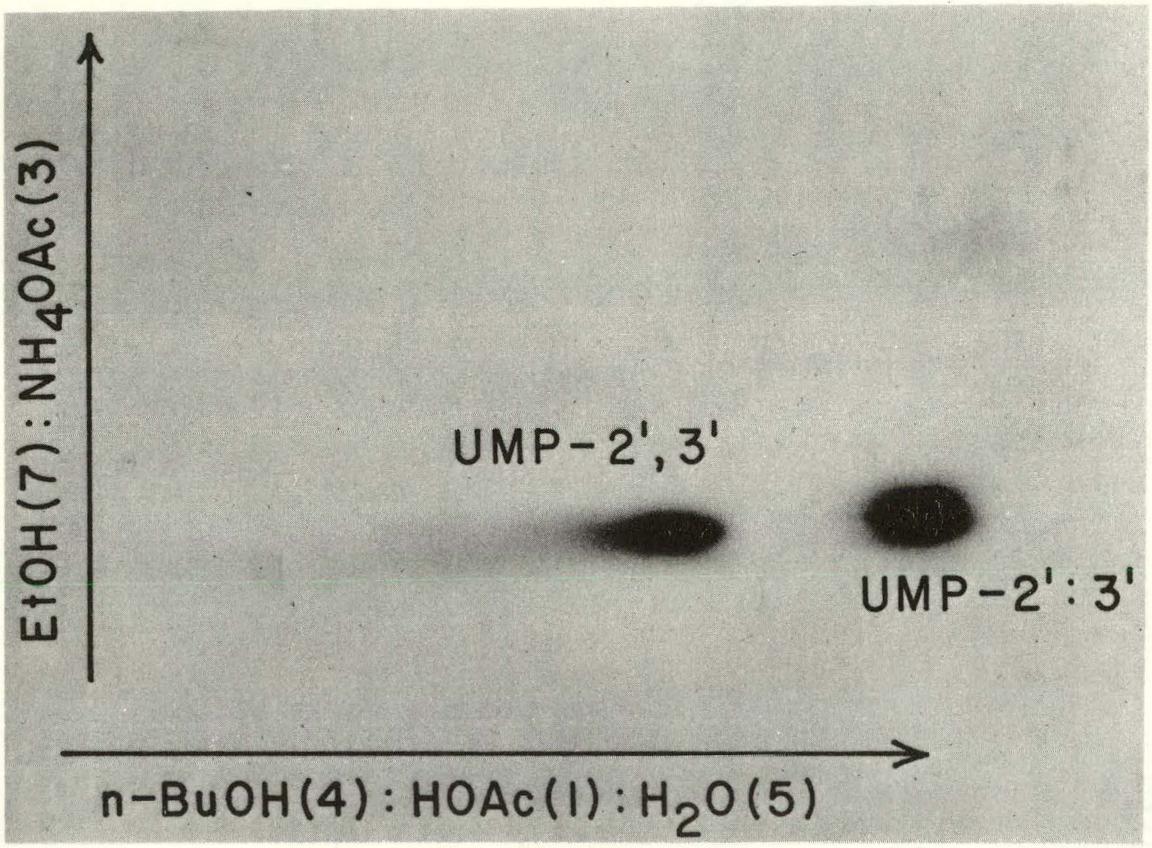


Figure 4. Chemical compounds used to effect dehydration condensation reactions in aqueous solution.



XBB 701-407

Figure 5. Autoradiogram of the chromatogram for the reaction of ^{14}C-UMP-2,3' and CMC.

very labile, and is readily hydrolyzed back to the 2',3'-phosphate mixture in dilute acid media. The order of stability of the cyclic phosphates is as follows:[130] the seven-membered cyclic phosphate is about as stable as the acyclic phosphodiester, which is slightly more stable than the six-membered cyclic phosphate, which in turn is much more stable (10^6 to 10^7 times in alkali) than the five-membered cyclic phosphate. As an example of the stability of a typical six-membered cyclic phosphate, adenosine-3':5'-cyclic phosphate has a half-life of 30 minutes in 1 M HCl at 100°C.

Thus, the reaction of the 2',3'-nucleotides was not studied in any further detail since another model nucleotide reaction existed which could be studied at various pH's. This is the esterification of AMP-5', which will be discussed next.

It should be noted here that sodium dicyanamide was also examined for condensing ability with AMP-2',3', with AMP-5', and with AMP-5' and adenosine with no reaction observed.

However, more recently Lohrmann and Orgel[135] have studied phosphorylations in aqueous solution using a number of plausible prebiotic condensing agents which include cyanogen, cyanoformamide, cyanate, cyanamide, thioformate, ethylisocyanide, and a water-soluble carbodiimide for comparison. The reactions studied were the formation of UMP-5' from uridine and phosphate and the cyclization of UMP-2',3'. They found that cyanamide would give the UMP-5' product in about a 2% yield after heating the reaction mixture for 30 days at 65°C. More important, they also showed that cyanamide would cyclize UMP-2',3' in about a 40% yield after running the reaction for ten days at 37°C. From this study these authors[135] concluded that the most plausible "prebiotic" condensing agents now available are cyanogen and cyanoformamide, but that cyanamide and dicyandiamide should also be considered carefully.

B. Reaction of Adenosine-5'-Monophosphate with Ethanol

The reaction of ethanol with AMP-5' in aqueous solution was thought to be a good model system for the polymerization of the nucleotides. The similarity between the two systems is that they both involve the formation of a phosphodiester linkage. One of the differences is that in the polynucleotide system the reaction is with a 3'-sugar alcohol (secondary alcohol) and monophosphate, whereas in the other reaction the primary alcohol, ethanol, is reacting with the monophosphate. However, the reaction of AMP-5' and ethanol was studied first using a water-soluble carbodiimide and the three cyanamide derivatives since this reaction gives very high yields of the ethylphosphodiester when a large excess of ethanol is present. The idea was that if the cyanamide derivatives could promote this reaction then they could be used to promote nucleotide or oligonucleotide self-condensation on a template. On the other hand, if they could not promote this reaction fairly readily then they would not be expected to promote internucleotide condensations even if the nucleotides were templated.

The reaction of AMP-5', ethanol and carbodiimide is illustrated in Fig. 6. For this reaction the ethanol is approximately 7 $\underline{M}$ and the carbodiimide concentration is five times that of AMP-5'. Since the adenosine-5'-ethylphosphate would be expected to be as stable as other phosphodiesters, this reaction was studied at various pH values of 1, 3, 6 and 8.5. The results of this study using the cyclohexylethylmorpholinylcarbodiimide are presented in Table I. The reaction gives the best

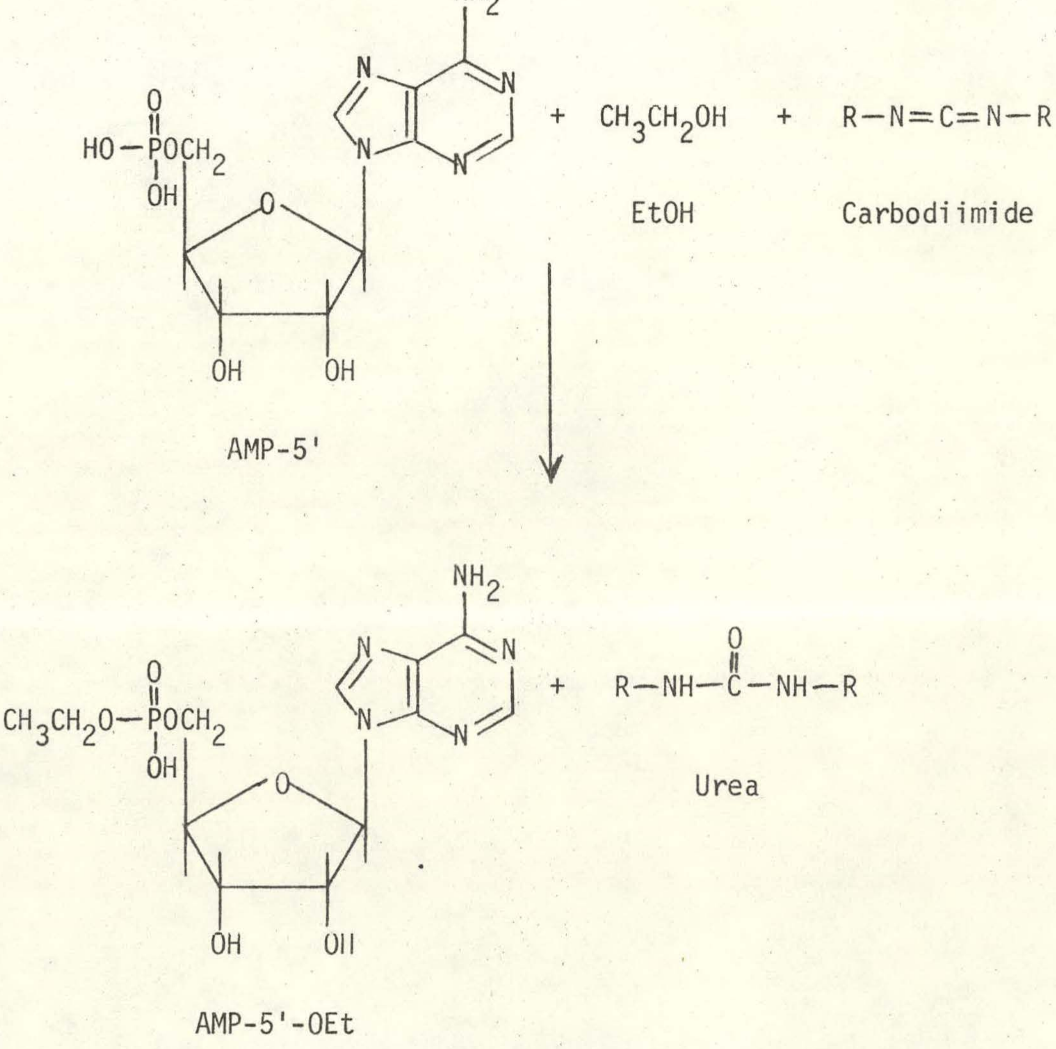


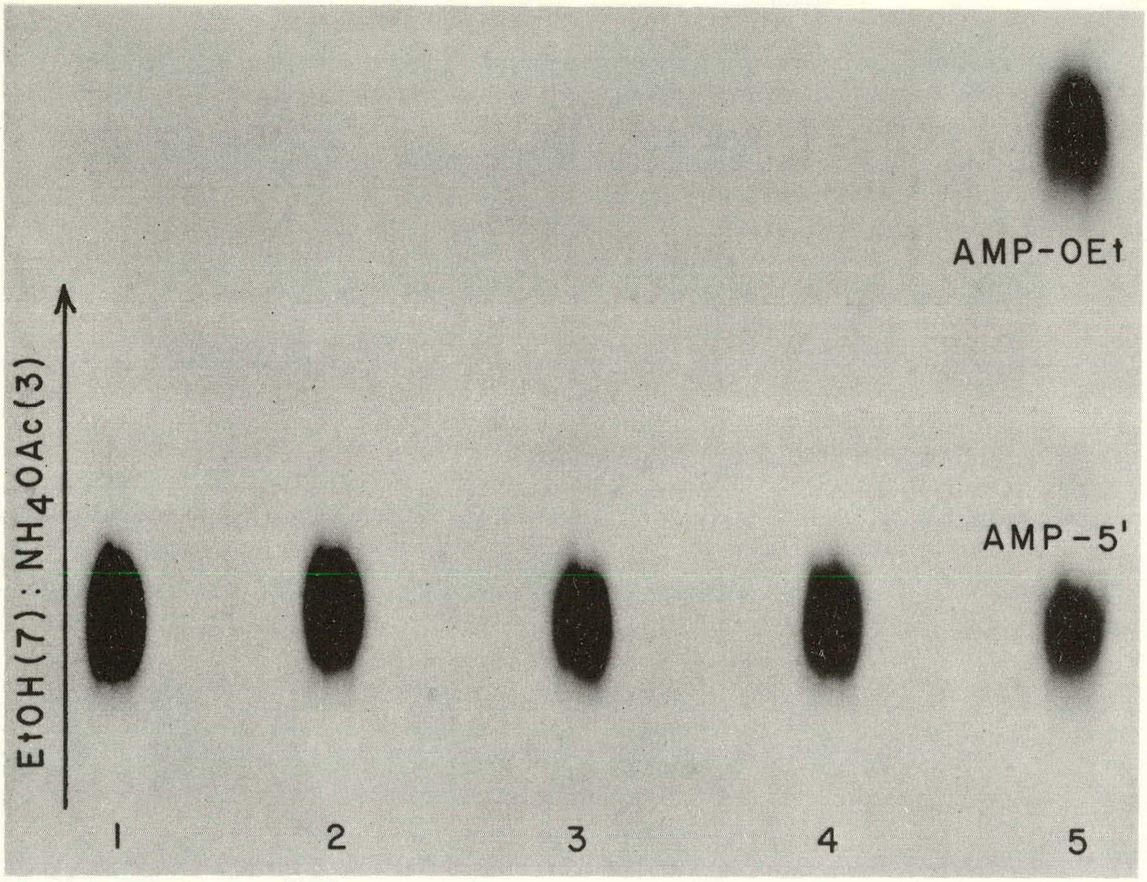
Figure 6. Reaction of AMP-5', ethanol, and carbodiimide.

Table I. Reaction of AMP-5', aqueous ethanol (7 M) and cyclohexylethylmorpholinylcarbodiimide (CMC).

pH	5'-AMP moles/liter	% Yield of adenosine-5'-ethylphosphate	Reaction time in hours	Other products
1	0.05	2	24	---
3	0.05	27	24	---
6	0.05	82	24	---
6	0.014	53	72	2%
8.5	0.014	10	72	12%

yields of product at pH 6, which is in agreement with literature reports[116] for the optimum pH value for carbodiimide mediated phosphate condensations. At lower and higher pH's the carbodiimide is probably being hydrolyzed more rapidly so that the nucleotide cannot compete as well for the carbodiimide. The carbodiimides in general are readily hydrolyzed in either acidic or basic media.[77] This same reaction was examined using cyanamide, sodium dicyanamide, and dicyandiamide as potential condensing agents at various pH's. Fig. 7 shows an autoradiogram with the comparison of the condensing agents, the cyanamide derivatives and carbodiimide, in this reaction at pH 6 and a reaction time of 72 hours.

Cyanamide and its derivatives gave no reaction at pH 3 or above. But at pH 1 both cyanamide and dicyandiamide formed a product in 35% and 39% yield, respectively. In both cases the new product had the same R_f. However, it was apparent that this new product was not the



XBB 701-405

Figure 7. Autoradiogram of the chromatogram of the reaction of an aqueous solution (pH 6) of ^{14}C-AMP-5$'$ (0.014 M), ethanol (7 M), and various condensing agents (0.07 M).
1 - no agent added; 2 - cyanamide; 3 - Na Dicyanamide;
4 - dicyandiamide; 5 - CMC.

adenosine-5'-ethylphosphate, since their R_f values were vastly different. The ethyl phosphate has a higher R_f (0.70) than AMP-5' (0.30) in an ethanol:1 M NH_4OAc (7:3) solvent system, whereas the new product has a lower R_f (0.20) than that of AMP-5'. It was concluded that this new product did not contain ethanol since it also resulted from an aqueous mixture of AMP-5' and DCDA without ethanol present. The new compound with low R_f was shown to contain both AMP-5' and DCDA by running the same experiment again, but this time having the ^{14}C-label in the DCDA. The ^{14}C spot on the X-ray film was in the same position relative to AMP-5' as it had been when ^{14}C-AMP-5' had been used. Shadowgrams made from the paper chromatograms showed that this new product absorbed UV light, and also indicated the position of unreacted AMP-5' when it was unlabeled. This new product which contained both AMP-5' and DCDA probably results from the reaction of DCDA with either the heterocyclic base or the ribose hydroxyls of AMP-5' and not the phosphate group. If the phosphate group were involved in the reaction the R_f of the compound would be much higher.

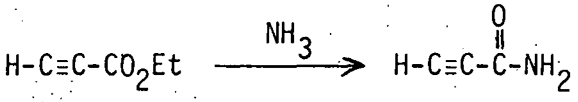
In the reactions using ^{14}C-labeled DCDA there also appeared to be a compound resulting from the reaction of DCDA and ethanol. This product was not present in the reaction without ethanol, and it did not absorb UV light. It gave a ^{14}C spot on X-ray film only with ^{14}C-labeled DCDA, and had a higher R_f (0.90) than DCDA (0.80). Sodium dicyanamide did not appear to promote any reaction with the ^{14}C-AMP-5', even at pH 1.

One other condensing agent was studied in this reaction somewhat later, and this was cyanoacetylene. This compound had been reported[137] to be a major product resulting from electrical discharges through a number of gases and gas mixtures, including hydrogen cyanide, cyanogen and acetylene, hydrogen cyanide and acetylene, and methane and nitrogen. In these cases, cyanoacetylene was the most abundant volatile nitrogen-containing compound other than HCN. It was then demonstrated[137] that cyanoacetylene under appropriate reaction conditions (reaction with NH_4CN) can yield asparagine, aspartic acid, and cytosine.

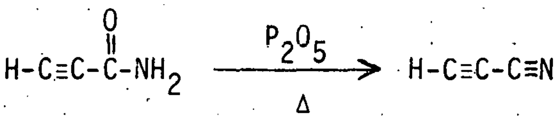
Since cyanoacetylene possesses a doubly-unsaturated (acetylenic) center, it should be able to promote dehydration condensation reactions. This expectation is reinforced by its relationship to ethoxyacetylene, which is known to promote the formation of peptides from amino acids[138-140] and of anhydrides from carboxylic acids.[141] Cyanoacetylene has been shown to react with inorganic phosphate to give cyanovinyl phosphate.[142] The cyanovinyl phosphate has then been used as a phosphorylating agent in the conversion of uridine to uridine monophosphate and phosphate to pyrophosphate.

Cyanoacetylene is not available commercially and was prepared by a modified procedure of Moureu and Bongrand,[143] illustrated in Equations 1-3.

$$\mathrm{H{-}C{\equiv}C{-}CO_2H} + \mathrm{EtOH} \xrightarrow{\mathrm{H_2SO_4}} \mathrm{H{-}C{\equiv}C{-}CO_2Et} \tag{1}$$



$$\mathrm{H{-}C{\equiv}C{-}CO_2Et} \xrightarrow{\mathrm{NH_3}} \mathrm{H{-}C{\equiv}C{-}\overset{\overset{O}{\|}}{C}{-}NH_2} \tag{2}$$



$$\mathrm{H{-}C{\equiv}C{-}\overset{\overset{O}{\|}}{C}{-}NH_2} \xrightarrow[\Delta]{\mathrm{P_2O_5}} \mathrm{H{-}C{\equiv}C{-}C{\equiv}N} \tag{3}$$

The reaction of AMP-5' and ethanol with cyanoacetylene was studied at various pH's also. At pH 8.5 several products were formed but their yield was so small (about 1%) that this study was not pursued further.

From the results of these experiments a major decision was made. This decision was to study the reactions of the nucleotides with the carbodiimides, since the cyanamide derivatives did not appear to be able to give satisfactory reactions. However, in view of Lohrmann and Orgel's[135] work showing that cyanamide can produce UMP-2':3' from UMP-2',3' it is now thought probable that cyanamide could also produce the adenosine-5'-ethylphosphate in the reactions just studied, when this reaction is carried out under much longer reaction times and with heating. Even if this were the case these derivatives would not be satisfactory to use to promote templated condensation reactions since this type of reaction usually requires low temperatures, c.g., the polynucleotides interact the strongest at low temperatures, about 0°C, and at higher temperatures this interaction ceases to exist as the double helix is "melted out". The interaction of smaller oligonucleotides with polynucleotides is even more sensitive to temperature increases, where "melting out" temperatures are in the 5-30°C range.[116]

The adenosine-5'-ethylphosphate which had been synthesized and identified by other workers[116] was identified in these experiments by its reaction with snake venom phosphodiesterase to give AMP-5' and by changing the ^{14}C-label from AMP-5' to the alcohol. When this reaction was studied further at lower concentrations of ethanol and higher concentrations of carbodiimide it was discovered that additional reaction products were formed. This came about when the two water-soluble carbodiimides, cyclohexylethylmorpholinyl (CMC)- and ethyldimethylaminopropylcarbodiimide (EPC), were compared with respect to their ability to promote the esterification of AMP-5'. One of the additional products formed was the symmetrical diadenosine pyrophosphate ($A^{5'}ppA$), and it was demonstrated that at pH's higher than 7 this was the major product in spite of the large excess of ethanol present. This is probably due to the increased amount of phosphate dianion present which would be a much stronger nucleophile. The pK_A's of AMP-5' are approximately 1 and 6.5.[56] The smaller water-soluble carbodiimide, EPC, gave much better yields of both the adenosine-5'-ethylphosphate (20%) and the $A^{5'}ppA$ (40%) than CMC, which gave the products in 10 and 13% yields, respectively. Thus, in the next section the reactions of AMP-5' and EPC under various conditions will be discussed.

It should be mentioned again that the water-soluble carbodiimides in general are not very plausible primitive Earth type condensing agents because of their complexity. However, these compounds were used because they react much faster and give much better yields of products than the usual prebiotic condensing agents. The carbodiimides probably react by a mechanism similar to that of the cyanamide derivatives and could be used as a model for condensing agents in general--i.e., what the carbodiimides can do very well in a short time the other condensing agents can do possibly less well under more severe reaction conditions and longer times. Finally, it is not improbable that some simpler types of carbodiimides, such as 1,3-bis-dimethylcarbodiimide, could have been formed under primitive Earth conditions, although this has not been reported.

C. The Reaction of Adenosine-5'-Monophosphate with Water-Soluble Carbodiimides

Since the reaction of AMP-5' with the carbodiimides in aqueous solution gave diadenosine pyrophosphate (Fig. 8) as a major product even with some ethanol present, it was thought that it would be worthwhile to study the reactions of just the nucleotide and carbodiimide. One of the reasons for doing this was that we had hoped to eventually show a templating interaction between the mononucleotide and its complementary polynucleotide, in this case between AMP-5' and polyuridylic acid. It was predicted that this templating mechanism would serve to orient the mononucleotides on the polymer such that when the condensing agent was added the polymerization of the monomer would be facilitated. This is essentially what Naylor and Gilham[116] had already done using thymidine oligonucleotides and polyadenylic acid. However, they could not detect any interaction between the oligomer and the polymer when the oligomer was four units long or less. Miles and coworkers[136] have reported that they could find no evidence for interaction of AMP-5' and polyuridylic acid under conditions where complexes having helical structures were formed with adenosine and adenine. Thus, this problem of the templated replication of polynucleotides was not simple at the start. But before the replication problem could be attacked, the reaction of AMP-5' with the water-soluble carbodiimides had to be studied in detail.

This reaction was carried out using the four water-soluble carbodiimides, CMC, EPC, EPC·MeI, and MEC (see Fig. 4). CMC and EPC are commercially available. EPC·MeI was synthesized from EPC as illustrated in Equation 4.

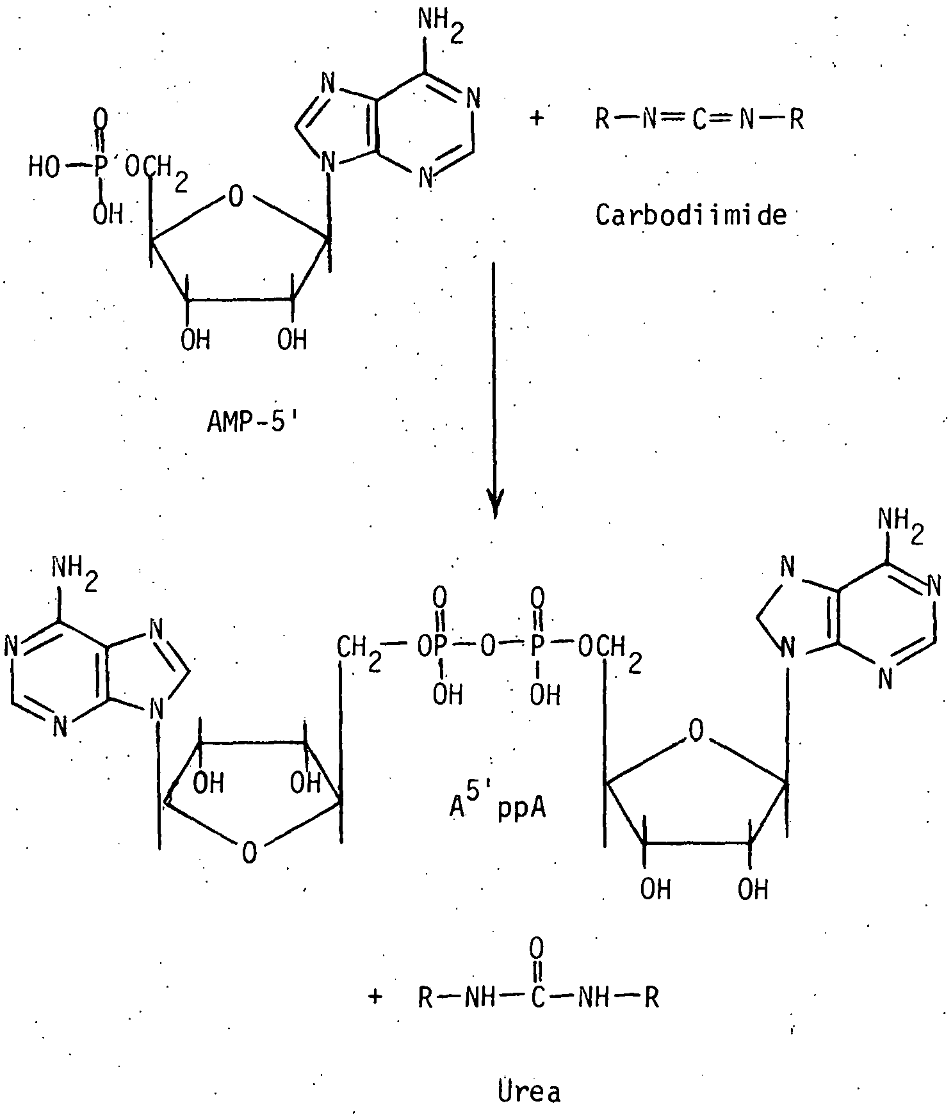
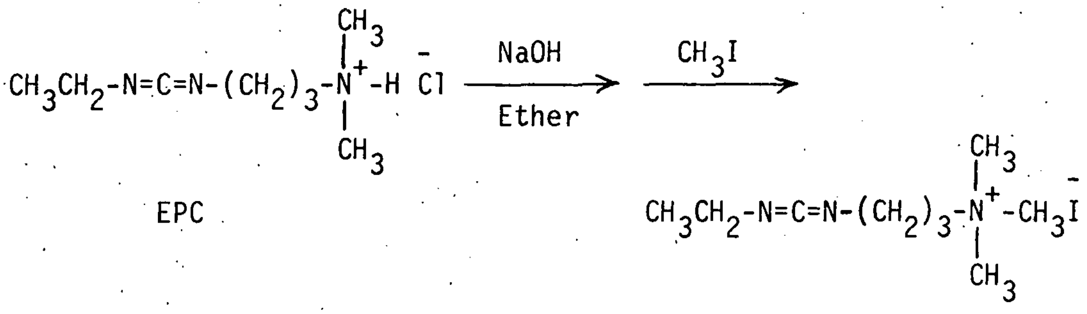


Figure 8. Reaction of AMP-5' with carbodiimide.

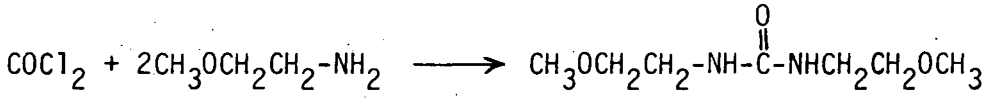


$$CH_3CH_2\text{-}N{=}C{=}N\text{-}(CH_2)_3\text{-}\overset{\overset{\displaystyle CH_3}{|}}{\underset{\underset{\displaystyle CH_3}{|}}{N^+}}\text{-}H\ Cl^- \xrightarrow[\text{Ether}]{\text{NaOH}} \xrightarrow{CH_3I} CH_3CH_2\text{-}N{=}C{=}N\text{-}(CH_2)_3\text{-}\overset{\overset{\displaystyle CH_3}{|}}{\underset{\underset{\displaystyle CH_3}{|}}{N^+}}\text{-}CH_3I^- \qquad (4)$$

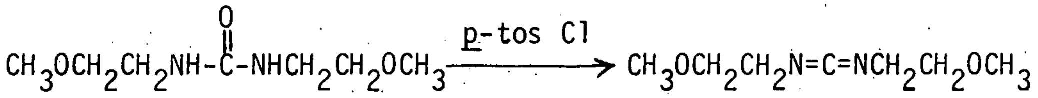
EPC

EPC·MeI

The symmetrical bis-methoxyethylcarbodiimide was synthesized according to Equations 5 and 6.



$$COCl_2 + 2CH_3OCH_2CH_2\text{-}NH_2 \longrightarrow CH_3OCH_2CH_2\text{-}NH\text{-}\overset{\overset{\displaystyle O}{\|}}{C}\text{-}NHCH_2CH_2OCH_3 \qquad (5)$$



$$CH_3OCH_2CH_2NH\text{-}\overset{\overset{\displaystyle O}{\|}}{C}\text{-}NHCH_2CH_2OCH_3 \xrightarrow{\underline{p}\text{-tos Cl}} CH_3OCH_2CH_2N{=}C{=}NCH_2CH_2OCH_3 \qquad (6)$$

The carbodiimides CMC, EPC·MeI, and EPC are related in that they are all very water-soluble by virtue of the positively-charged quaternary ammonium group. Thus, they were all expected to give quite similar reactions with AMP-5'. In fact, the EPC·MeI was synthesized ^{14}C-labeled by Equation 4 using $^{14}C\text{-}CH_3I$, in the expectation that it would react like EPC. This was done because the EPC·MeI-^{14}C could be synthesized very easily, whereas to synthesize the EPC-^{14}C would have been a major undertaking. However, it turns out that EPC is somewhat different than EPC-MeI and CMC in that it reacts with AMP-5' to give $A^{5'}ppA$ in such larger yields. Under the same conditions, AMP-5' (0.025 $\underline{M}$) at pH 7.5, reacts with EPC to give $A^{5'}ppA$ in more than 40% yield, whereas CMC and EPC·MeI give only about a 5% yield with a ten-fold excess of the condensing agent. CMC and EPC·MeI both formed more side products which are thought to be adducts containing both AMP-5' and carbodiimide. The uniqueness of EPC as a condensing agent is probably due, at least in part, to the fact that it can exist in several tautomeric forms (Fig. 9). For example, Sheehan and collaborators[132] have postulated that EPC exists in the carbodiimide form in chloroform solutions, but in the solid state as the non-carbodiimide tautomers.

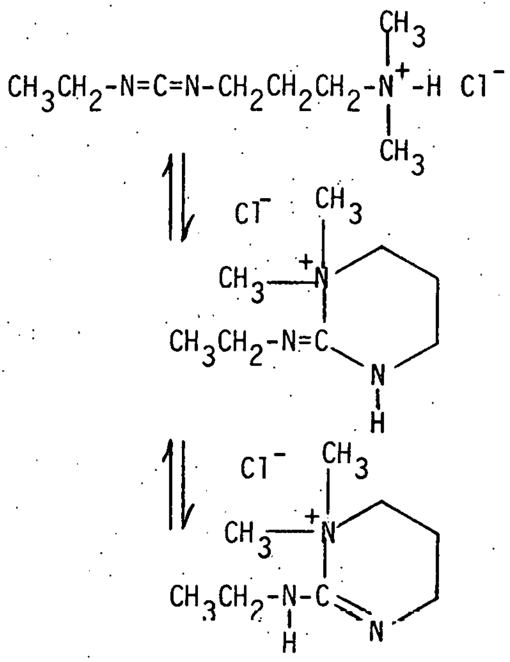


Figure 9. Tautomers of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EPC).

The evidence for this is the infrared spectra,[132] where in chloroform solution one finds the characteristic carbodiimide absorption band at 2130 cm^{-1}, and in nujol or KBr suspensions this band is absent. In the crystalline state there are two other absorption bands at 3250 cm^{-1} and

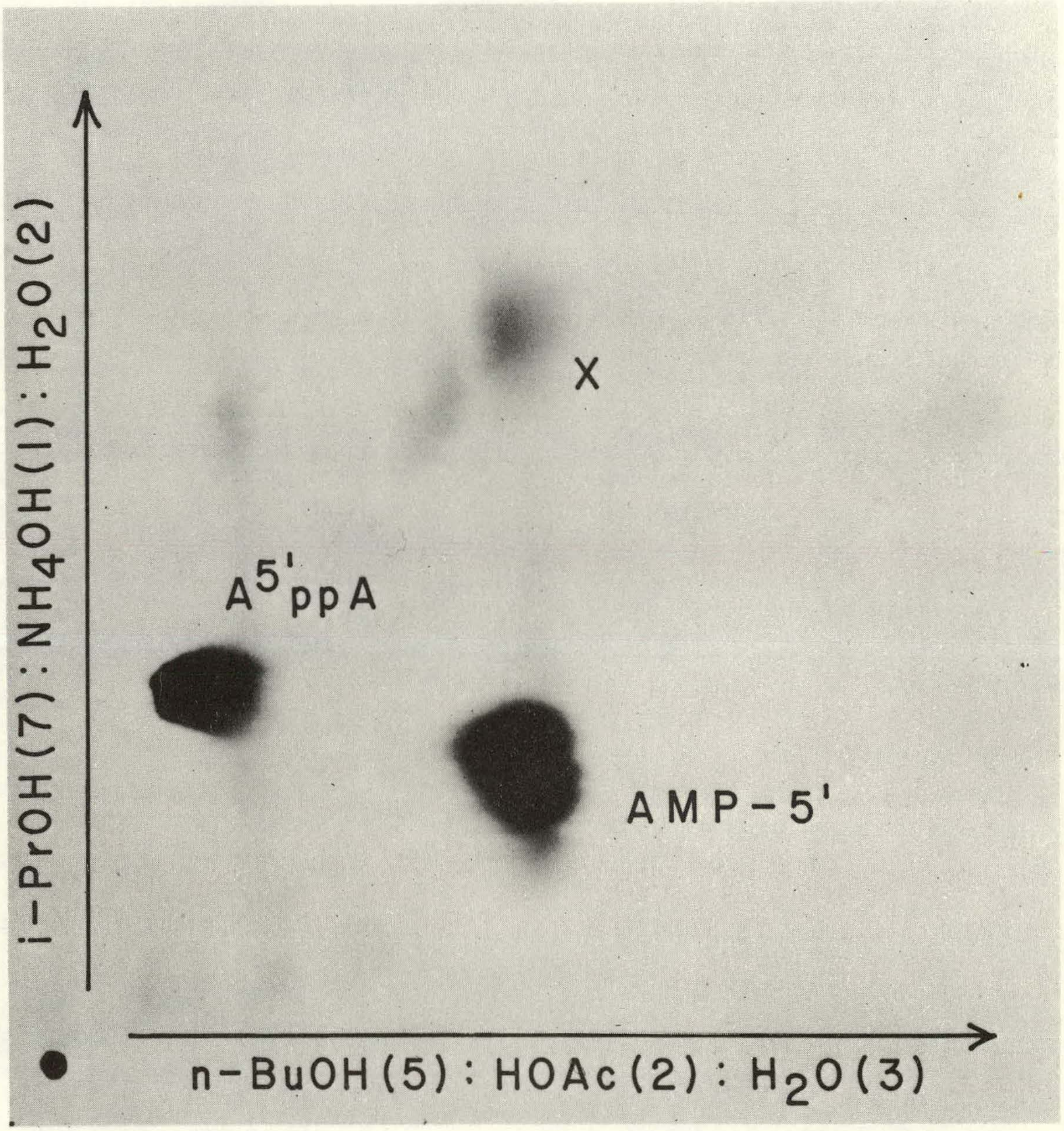
1700 cm^{-1} which are characteristic of the -NH- and -C=N-, respectively. These observations have been verified in this laboratory.

In later experiments the reactions of EPC and MEC with AMP-5' were compared, and it was again found that EPC gave much more $A^{5'}$ppA (20%) than MEC (5%). The autoradiograms of the chromatograms from the reaction of AMP-5' and EPC and MEC are shown in Fig.10 and Fig. 11, respectively.

The reaction of AMP-5' and EPC was studied as a function of pH, and it was found that the best yields of diadenosine pyrophosphate were obtained around pH 5 to 6. This is about the same pH for optimum yields as in the esterification of AMP-5'. The kinetics of this reaction were also studied. For this study the reaction was run in an ice bath at 0°C and aliquots were withdrawn and spotted on chromatography paper at various time intervals. The reaction turns out to be very fast and is about 90% complete in 5 minutes or less. The product distribution changes very little after that, even after several days time.

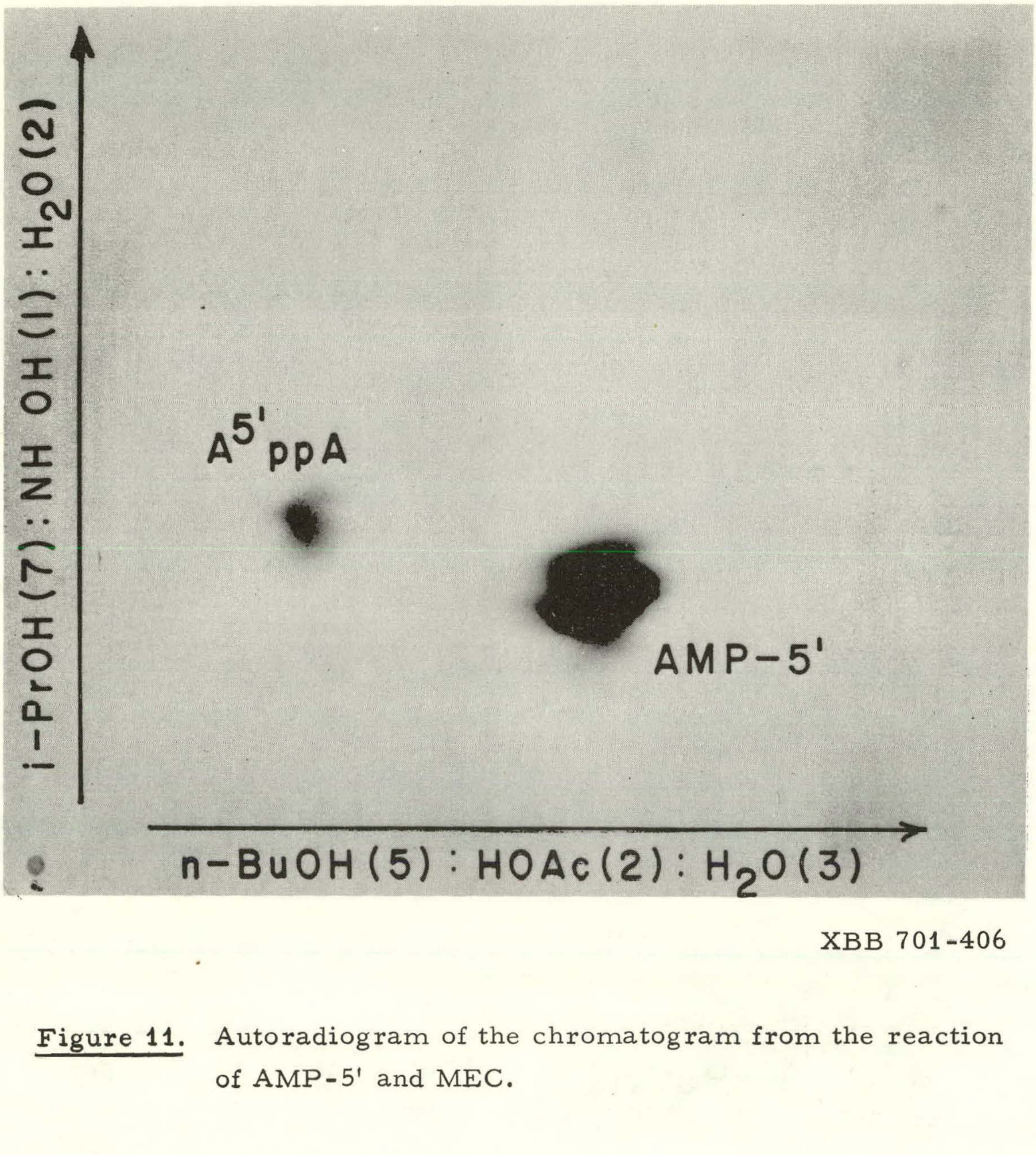
The reaction of AMP-5' and EPC was carried out at neutral pH with varying amounts of EPC added. The results of this study are presented in Table II. After the molar ratio of EPC to AMP-5' reaches 10 there is very little increase in the amount of $A^{5'}$ppA produced by increasing this ratio further. However, the amounts of side products tends to increase. These products are thought to be adducts between the carbodiimide and AMP-5' by virtue of their high R_f's.

The diadenosine pyrophosphate formed in this reaction was identified by co-chromatography with an authentic sample of $A^{5'}$ppA, and by basic and acidic hydrolysis back to AMP-5'. The $A^{5'}$ppA used as a



XBB 701-403

Figure 10. Autoradiogram of the chromatogram from the reaction of AMP-5$'$ and EPC.



XBB 701-406

Figure 11. Autoradiogram of the chromatogram from the reaction of AMP-5' and MEC.

Table II. Percent yield of products from the reaction of ^{14}C-AMP-5' (0.025 M) and EPC.

Molar ratio EPC/AMP	AMP-5'	$A^{5'}ppA$	X	Other products
0.25	98.7	1.3	---	---
0.50	96.8	2.9	0.3	---
1.00	92.5	6.6	0.9	---
10.0	41.8	40.5	10.1	7.6
20.0*	31.1	43.2	10.4	15.3

*EPC was added in two equal portions 24 hours apart.

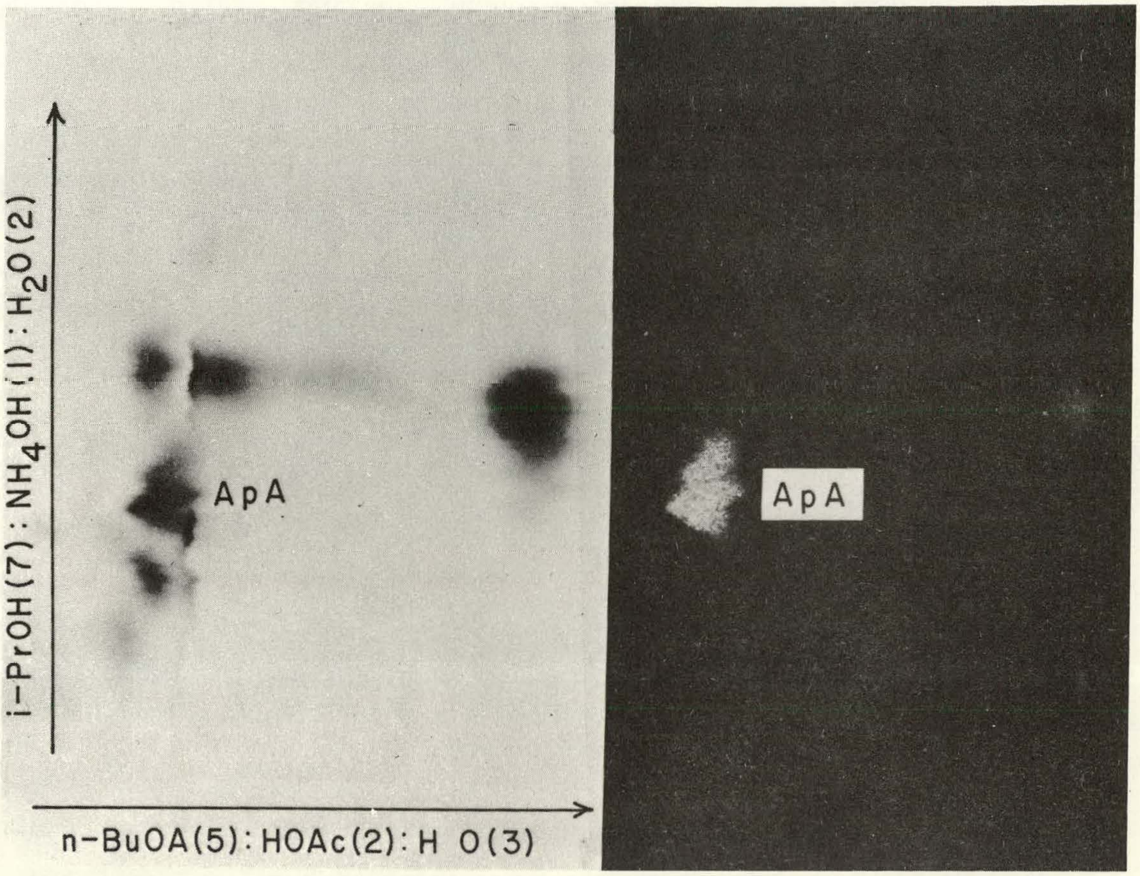
standard was synthesized according to a literature procedure[144] from the reaction of AMP-5', dicyclohexylcarbodiimide, and n-Bu$_3$N in anhydrous pyridine.

In the synthesis of oligonucleotides from 5'-nucleotides and carbodiimides in anhydrous pyridine, it was demonstrated that the symmetrical dinucleoside pyrophosphate was formed first and the linear oligonucleotide later.[144,145] In the reaction of thymidine-5'-monophosphate these oligonucleotides have been isolated and characterized up to the undecamer [(pT)$_{11}$].[145] In these reactions the dinucleoside pyrophosphate is the only product formed when tri-n-butylamine is present. The mechanism of the dicyclohexylcarbodiimide-mediated polymerization of nucleotides is still unknown,[77] although several attempts have been made to account for it.[146-148] Khorana and collaborators[148] have suggested that the reaction probably proceeds through meta- or polyphosphate intermediates.

With this information in mind, it was thought that the diadenosine pyrophosphate product which was formed from the reaction of AMP-5' in aqueous solution might be reacted further to give higher oligomers. However, all attempts to do this were unsuccessful. For example, $A^{5'}ppA$ did not appear to react with EPC nor to react with ^{32}P-TMP-5' in the presence of EPC.

Compound X (Fig. 10) which is formed from the reaction of AMP-5' and EPC in the second largest yield was initially thought to be a trimer of AMP-5. This compound is quite unstable and breaks down (probably by hydrolysis) to give X (53%), AMP-5' (43%) and $A^{5'}ppA$ (4%) under the mild conditions of elution and re-chromatography. Compound X could be some sort of meta- or polyphosphate intermediate but this seems a bit doubtful since $A^{5'}ppA$ would not react further. It seems more likely that this product is an unstable adduct between the EPC and AMP-5'.

The reaction of ^{14}C-AMP-5' with EPC was carried out using a large amount of radioactivity in an attempt to detect the natural dimer, pApA. In this experiment the products were reacted with alkaline phosphatase, which cleaves monophosphates to alcohols and inorganic phosphate. The resulting products were then co-chromatographed with standard ApA which is commercially available. In this way any pApA that resulted from the initial reaction would be enzymatically hydrolyzed to ApA which could be identified by its coincidence with the standard ApA. Figure 12 shows this coincidence of the ^{14}C-ApA which appears on the X-ray film and the standard ApA, the position and shape of which appear on the shadowgram. Both the autoradiogram and shadowgram were made from the same paper that the ^{14}C-products and standard ApA were co-chromatographed on (see the



XBB 701-404

Figure 12. Autoradiogram and shadowgram showing the coincidence of ^{14}C-ApA and standard ApA.

Experimental Section for further details of this technique). The pApA was formed in about 0.4% yield and was identified in this way. The ^{14}C-ApA was further characterized by reacting it with snake venom phosphodiesterase to give AMP-5' and adenosine. The internucleotide linkage is unspecified as to whether it is 2'-5' or 3'-5'. The overall reactions for the identification of pApA are given in Equation 7.

$$\text{pApA} \xrightarrow[\text{phosphatase}]{\text{alkaline}} \text{ApA} \xrightarrow[\text{phosphodiesterase}]{\text{snake venom}} \text{pA} + \text{A} \qquad (7)$$

It was at about this stage in this work when the first of the papers from Orgel's laboratory[117-119] appeared on the non-enzymatic replication of polynucleotides. This work was discussed in the latter part of the first section. What these workers have done is to show that polyuridylic acid forms a stable three-stranded helix with AMP-5' and that this complex greatly facilitates the condensation of AMP with the commercially available EPC. This same catalysis has been shown to exist with polycytidylic acid and guanosine-5'-monophosphate, but has yet to be demonstrated with the pyrimidine nucleotides and purine polynucleotides--the problem being that stable complexes between them are difficult to form. However, this will probably be worked out, so that the idea of non-enzymatic replication of polynucleotides on the primitive Earth will have substantial experimental support.

Since the work on the polynucleotide replication had already been started in Orgel's laboratory, the research effort changed directions at this point. The new direction was towards finding experimental evidence for interactions between the nucleic acid and polypeptide systems. These experiments are discussed in the following section.

D. Reactions of Adenosine-5'-Monophosphate in the Presence of Polypeptides

For life to have evolved from non-living matter a great many events must have occurred. It should be possible to reproduce many of these events in the laboratory under conditions believed to have existed on the prebiotic Earth. This has been accomplished for the synthesis of many biomonomers and biopolymers. However, in the more advanced stages of chemical evolution most of the necessary events have not been reproduced in the laboratory with the notable exceptions of the formation of membrane-like structures and possible non-enzymatic polynucleotide replication. Probably one of the most important questions in chemical evolution at this time is how the genetic code originated. That is, how did the present relationship of the polynucleotides and polypeptides evolve, wherein the polynucleotides direct polypeptide synthesis *via* a complex series of enzymatic steps. The genetic code appears to be universal, so that the question of why a particular codon codes for one amino acid and no other is a major part of the above question.

There are two extreme views on the genetic codes evolution, one being that it was an accident that any particular codon codes for a specific amino acid and the other that each codon and its amino acid are related stereochemically (see Chapter I for a detailed discussion). The "frozen accident" theory has it that the codon for a particular amino acid was assigned by chance and that this particular assignment was then frozen since any later change would be lethal to the system in which it arose. According to the stereochemical view the assignment of a particular codon to an amino acid was the result of the interaction of this amino acid with the codon--i.e., this amino acid interacted stereochemically more strongly than the other amino acids with this particular codon. With this theory if life evolved somewhere else the genetic code would be the same, whereas with the frozen accident theory the code would probably be quite different.

This author feels that the stereochemical theory is more attractive than the "frozen accident" theory, although both are too extreme. It seems likely that the genetic code is the result of some sort of interaction between the nucleic acid and amino acid systems since they must have interacted in some way initially to have become as intimately coupled as they are now. It also seems probable that the original interactions between these systems were not anywhere near selective enough to give an all-or-nothing relationship between an amino acid and its codon. Thus, what is envisioned is that there exists weak interactions between the nucleic acid and amino acid systems that are slightly selective and that with the help of enzymes this slight selectivity has been amplified in the course of evolution to the complete specificity that is now observed in the genetic code.

However, the fact remains that there is really no hard experimental evidence that can substantiate one theory or invalidate the other. Most of the experiments having to do with the interactions of polypeptides and polynucleotides have been done with the basic amino acids and this interaction is primarily due to the opposite charges on the two systems.

In these systems there appears to be another interaction besides the primary ionic one, since Leng and Felsenfeld[149] have reported that poly-L-lysine exhibits an almost perfect selectivity for interaction with DNA rich in adenine and thymine as judged by preferential precipitation under reversible conditions (1 M NaCl). Thus, polylysine interacts in order of decreasing preference with the homopolymer poly dA:dT, _P. mirabilis_ DNA (62% AT), _S. marcescens_ DNA (42% AT), _M. lysodeikticus_ DNA (28% AT), and poly dG:dC. Poly-L-arginine's specificity is reversed and not as pronounced as that of polylysine--_e.g._, polyarginine interacts about equally well with the DNA having 40-60% GC, but exhibits a slight preference for the DNA of 72% GC, and a marked preference for the homopolymer, poly dG:dC. These workers have concluded that under their conditions these interactions are cooperative, reversible, and selective; however, the influence of counter ions in solution is important since they reported that the selectivity of polylysine is reversed when 1 M tetramethylammonium chloride was used instead of 1 M NaCl.

It has also been found[150] that there is some selectivity in the stabilization of native calf thymus DNA against melting with various cationic polypeptides. In this study the polypeptides listed in decreasing order of stabilization are poly-L-ornithine, poly-L-lysine, poly-L-arginine, and poly-L-homoarginine. This difference in degree of stabilization does not appear to be explainable simply on the basis of electrostatic neutralization. There is a correlation between the side-chain length and complex stability--_i.e._, the shorter the side chain, the greater the stabilization, as manifested by a progressive increase in melting temperature. This differential stabilization may be a consequence of steric limitations, of differing hydrophobicity of the various side-chains, or different over-all charge density. Analysis of the dispersion of hyperchromicity revealed that polyornithine and polylysine preferentially stabilize A-T regions, whereas polyarginine and polyhomoarginine appear less discriminating.[150]

Other experiments on the interaction of the mononucleotides with polylysine[111] and polyarginine[112] reveal that the mononucleotides interact in the order of decreasing strength, GMP > AMP > CMP > UMP. This interaction primarily involves the electrostatic neutralization and the base stacking interactions of the ionically bound nucleotides.

Thus, there existed evidence for interactions of the nucleic acids and the basic polypeptides so that this appeared to be a good system to work with. It was thought that the basic polypeptides would bind the mononucleotides in such a way that when a condensing agent was added the amount of polymerization would be enhanced. The experimental approach here was patterned after that used in the experiments on non-enzymatic polynucleotide replication, where it was shown that the condensation of a mononucleotide was greatly enhanced when it was bound to a polynucleotide. These experiments were carried out with the nucleotides in the presence of the polypeptides and with several amino acids in the presence of polynucleotides to determine whether either system would be catalytic for the other.

In the first experiments on the condensation reactions of AMP-5' in the presence of polypeptides, the charged water-soluble carbodiimide, ethyldimethylaminopropyl carbodiimide hydrochloride (EPC), was used. The results of all of these experiments using EPC were negative--*i.e.*, no enhancement of condensation products was observed when polypeptide was added. The reaction of AMP-5' and EPC was studied in the presence and absence of poly-L-lysine·HBr. The polylysine appeared to decrease the yield of $A^{5'}$ppA. Although the free amino groups on polylysine did not appear to react appreciably with the AMP-5' in this system it was thought that it would be better to block this group so that it couldn't react with either the AMP or carbodiimide.

Thus, in the second series of experiments a copolymer, poly-L-Phe-L-Lys (1:1), was used, where the ε-amino group of lysine had been reacted with CH_3I making the tetra-alkylammonium derivative. After reaction with methyl iodide the copolymer was dialyzed. The results of the reaction of AMP-5' and EPC in the presence of copoly-L-Phe-L-Lys-ε-$\overset{+}{N}(CH_3)_3$ are presented in Table III. The effect of the added copolymer is not large , but the yield of $A^{5'}$ppA does decrease with increased

Table III. Reaction of AMP-5' (0.025 M) and EPC (0.125 M) in the presence of copoly-L-Phe-L-Lys-ε-$\overset{+}{N}(CH_3)_3$. Reaction time was 7 hours at pH 7 and 0-5°C.

	AMP/Lys	% Yield		
		Compound X	$A^{5'}$ppA	Unreacted AMP-5'
1	---	4.3	30.2	65.5
2	5/1	4.0	27.3	68.7
3	1/1	4.5	22.0	73.5
4	1/2	4.5	18.3	77.2

amounts of the polypeptide. The copoly-Phe-Lys was used with the expectation that AMP would be bound ionically to the lysine part and the phenylalanine would help to stabilize this complex by interacting with the adenine portion of the nucleotide. Several of these copolymers are commercially available, but it turns out that when the Phe-to-Lys ratio is 2 or more these polymers are insoluble in water. Many experiments of this kind were carried out using EPC as the condensing agent with negative results. One explanation for this could be that the addition of a charged condensing agent was actually disrupting the very interaction that was expected--i.e., that the ionic interaction between AMP and lysine was disrupted by EPC. Thus, it was to test this possibility that the new uncharged carbodiimide, 1,3-bis-(2-methoxyethyl) carbodiimide (MEC), was synthesized and used in place of EPC. This explanation was not completely correct, since it was found that under analogous conditions using MEC the yield of $A^{5'}ppA$ in the presence of poly-L-Lys-ε-$\overset{+}{N}(CH_3)_3$ was about the same as in its absence.

However, the effect that was expected was finally found using somewhat different conditions. It was known that decreasing the dielectric constant of an aqueous solution of polypeptide and mononucleotide by adding a less polar organic solvent resulted in precipitation of both species. Thus, what was done was to add acetone to the aqueous solution of polylysine and AMP-5' just short of the point of precipitation (about 37% by volume of acetone). Table IV presents the results of these experiments. It is apparent that when acetone is present, the polylysine (M.W. - 100,000) apparently catalyzes this reaction so that the yield of $A^{5'}ppA$ is almost double that when it is absent.

Table IV. Reaction of AMP-5' (0.025 M) and MEC (0.125 M) in aqueous solution at pH 7.

	*Acetone	**Poly-Lys-ϵ-$\overset{+}{N}(CH_3)_3$	% Yield		
			$A^{5'}ppA$	AMP-5'	Origin
1	-	-	6.9	92.1	1.0
2	-	+	6.5	92.0	1.5
3***	+	+	15.1	81.9	3.1
4***	+	-	8.6	90.0	1.4

*When acetone was added it was about 37% by volume.

**Polylysine (M.W. = 100,000) was added so that the lysine-to-AMP ratio was one.

***AMP (0.0125 M) and MEC (0.0625 M).

It is believed that the reason this effect is observed in the acetone-water solution is that in this medium the polar compounds, AMP and polylysine, are less soluble and more attracted to each other. Thus, probably more AMP is bound to polylysine and it is bound more tightly than in the purely aqueous solution. This theory is supported by experiments with lysine and polyadenylic acid where they precipitate in a one-to-one complex (lysine-to-AMP) from a solution of dimethylformamide and water (3:1). Other workers[111,112] have shown that the mononucleotides and basic polypeptides precipitate, usually in a one-to-one ratio, when the concentration is high enough. In these experiments the exact amount of AMP-5' actually bound to the polymer is not known, but it probably is not in a one-to-one ratio since the complex would precipitate if it were. However, the increased yield of $A^{5'}$ppA is believed to result from more AMP being complexed by polylysine in the acetone-water solution.

It is still too early to assess the importance of this result for chemical evolution. It could be pertinent in that this sort of catalysis would facilitate polynucleotide formation. However, in these experiments the effect has only been observed for the formation of the unnatural dimer, $A^{5'}$ppA, although the earliest polynucleotides probably were a mixture of the various types of linkages. The other problem in evaluating this result is whether the acetone-water solution is similar to any sort of microenvironment on the primitive Earth. Since it is thought that the main effect of the added acetone is to increase the amount of complexing of AMP to polylysine, this may be analogous to what would happen inside of a membrane or on some catalytic surface.

Thus, after many unsuccessful experiments it was found that polylysine and AMP-5' do interact under special conditions so that the self-condensation of AMP-5' is enhanced by the presence of polylysine. This effect has been demonstrated under conditions not believed to have existed on the primitive Earth--*i.e.*, the condensing agent, MEC, is too complex and the acetone-water solution is an unlikely solvent. However, the interaction does exist and this type of catalysis could well have been important in an early stage of polynucleotide formation under conditions more relevant to chemical evolution. This has yet to be demonstrated under primitive Earth conditions, but the model exists now and this should be possible in the future.

E. Summary

A number of condensing agents that can effect condensation reactions in aqueous solution were examined to determine which were the most effective in reactions with nucleotides. The condensing agents included several water-soluble carbodiimides, cyanamide and two of its derivatives, and cyanoacetylene. In a model reaction of AMP-5' with ethanol it was found that only the carbodiimides gave satisfactory yields of the ethyl phosphate ester. In studying the reactions of AMP-5' with four different carbodiimides it was found that $A^{5'}ppA$ was formed in the highest yield using EPC (see Fig. 4). The optimum pH for these reactions is about 6, and the optimum ratio of carbodiimide to AMP-5' is a five- to tenfold excess. Numerous attempts to demonstrate a catalytic effect for the basic polypeptides on the self-condensation of AMP-5' using the charged condensing agent, EPC, were unsuccessful. This catalysis was finally found using an uncharged water-soluble carbodiimide, MEC, in a solvent system of acetone-water. In this mixed solvent (acetone 37% by volume), AMP-5' reacts with MEC to give $A^{5'}ppA$ in about a 15% yield in the presence of poly-L-Lys-ε-$\overset{+}{N}(CH_3)_3$, whereas in the absence of the polypeptide the yield is 8.6%.

F. Experimental Section

1. Standard Procedures for ^{14}C-Labeled Compounds.

The ^{14}C, ^{3}H, or ^{32}P isotopically labeled compounds used in these experiments were usually supplied commercially in a dilute solution of either water or a water-alcohol mixture. These compounds generally had a very high specific activity so that the radioactive (hot) samples had to be diluted with non-radioactive (cold) samples. The specific activity is expressed as either the number of millicuries per millimole or microcuries per micromole where one microcurie (10^{-6} curies) is defined as that amount of radioactivity giving 2.220×10^{6} disintegrations per minute.[127] The usual procedure was to pipette the required amount of radioactive sample into a pear-shaped flask and evaporate this sample to dryness on a rotary evaporator. A standard solution of cold compound was then prepared at the desired molarity, pH, etc. The cold compound was then directly pipetted into the pear-shaped flask. Usually there would be only a small correction in the final molarity of the resulting radioactive solution since the initial hot compounds had such a high specific activity.

After the reaction was over one of the methods used to identify the products was to add a fairly large amount of a compound which was thought to be produced in the reaction. Enough of the standard compound was added to permit easy detection by an established method, such as spraying or ultraviolet photography, and an aliquot of the resulting mixture was spotted on a sheet of Whatman No. 1 chromatography paper. Two-dimensional chromatography was then carried out.

After the chromatography process was completed and the sheet was dried, radioactive ink spots were placed as positional references at selected points on the paper and the chromatogram was folded to the size of a standard X-ray film.

In those experiments where the added standard was UV-absorbing, positive contact photographic paper was placed under the chromatogram and the paper illuminated for about 15 seconds in the dark room. A Hanovia 125 watt type 16200 mercury ultraviolet lamp was used at a distance of about two feet from the paper which was then developed by standard techniques. UV-Absorbing carriers and the ink spots appeared white on a black background.

Carbon-14 has a half-life of 5770 years and emits a beta-particle of 0.156 MeV, making it readily detectable with standard X-ray film.[127] A sheet of blue-sensitive X-ray film was placed in juxtaposition with the chromatography paper and together both were fitted into a light-tight envelope. After sufficient exposure time, the film was withdrawn in the dark-room and developed by standard techniques. The dried film was then placed on top of the contact print made earlier. Alignment of the two was achieved by adjusting their positions so that the radioactive reference spots appearing as black marks on the X-ray film were directly over the corresponding white marks on the contact print. In this manner it was possible to ascertain if one of the radioactive products appearing on the film corresponded both in position and shape to the standard outline on the UV contact print (shadowgram) of the same chromatogram.

Those carriers which responded to chemical sprays, such as ninhydrin for amino acids, were located and compared to the radioactive products in a corresponding manner.

2. Synthesis of Cyanoacetylene

Cyanoacetylene was prepared by a modified procedure of Moureu and Bongrand.[143] The first step was to add 25 gms of propiolic acid (0.36 moles) to 50 ml of methanol and 3 ml of sulfuric acid in a 250 ml Erlenmeyer flask. This reaction mixture was allowed to stand four days at room temperature and was then refluxed for half an hour. Seven hundred fifty ml of a saturated ammonium sulfate solution was added and the resulting mixture extracted with four 150 ml portions of ether. The ether was distilled from the solution leaving the crude methylpropiolate as a pale yellow liquid.

Twenty ml of the methylpropiolate was added to 200 ml of liquid ammonia in a 500 ml round-bottomed flask cooled in an acetone-dry ice bath. The reaction mixture was stirred with a magnetic stirrer for four hours and then left overnight in the dry ice bath. All the volatile material was removed by first warming the flask up to room temperature and then drawing a vacuum on it. This gave the amide of propiolic acid as a greyish-white solid.

Fourteen gms of the impure amide was added to 100 gms of P_2O_5 and 100 gms of sand in a 500 ml round-bottomed flask. The flask was then connected to a vacuum distillation apparatus and heated to 120°C in an oil bath. A vacuum was applied to the system and cyanoacetylene was distilled under a CO_2 atmosphere. The product was collected in a

50 ml round-bottomed flask cooled in an acetone-dry ice bath. The overall yield was about 20% based on propiolic acid.

Cyanoacetylene is a colorless liquid which forms white crystals when frozen. It has a melting point of 5°C and boils at 42°C. The compound decomposes slowly in light and air but can be kept for long periods of time when frozen. The UV spectrum in 95% ethanol gives three maxima at 229, 219 and 211 millimicrons. Its infrared spectrum showed absorptions at 2070 cm^{-1} for the $-C{\equiv}C-$ stretching vibration, 2270 cm^{-1} for $-C{\equiv}N$, and 3320 cm^{-1} for the stretching vibration of the acetylenic hydrogen. For the elemental analysis:

Calculated: C = 70.61%; N = 27.42%; H = 1.96%

Found: C = 70.43%; N = 27.18%; H = 2.18%

3. Synthesis of 5'→5'-Diadenosine Pyrophosphate

The synthesis of the di(tri-n-butylammonium) salt of $A^{5'}$ppA was carried out according to a literature procedure[128] that was used for the synthesis of diuridine pyrophosphate. One gram of AMP-5' (2.88 mmoles) was added to 115 ml of dry pyridine with 1.4 ml of tri-n-butylamine. To this solution was added 2.72 grams of dicyclohexylcarbodiimide (13.2 mmoles) and the reaction mixture was allowed to stand at room temperature for three days. The solution was then poured into a 250 ml round-bottomed flask and evaporated under reduced pressure. One hundred ml of water was added to the residue and the mixture filtered to remove any carbodiimide and urea present. The filtrate was again evaporated under reduced pressure and dried over P_2O_5. The resulting product was chromatographically homogeneous and was identified by its hydrolysis back to AMP-5' in 0.1 M NaOH. In basic solvent systems (solvent A) $A^{5'}ppA$ had the same R_f as AMP-5', but in an acidic solvent system (solvent F) it had a lower R_f (0.15) than AMP (R_f 0.35). Elemental analysis was as follows:

Calculated: C = 50.51%; H = 7.72%; N = 16.11%; P = 5.93%

Found: C = 49.56%; H = 7.70%; N = 15.70%; P = 5.90%

4. Synthesis of 1-Ethyl-3-(3-Dimethylaminopropyl) Carbodiimide Methiodide (EPC·MeI)

EPC·MeI was prepared from EPC (Cyclo Chemical Co.). Fifteen gms of EPC (0.078 moles) were added to 400 ml of diethylether in a 1 liter Erlenmeyer flask equipped with a magnetic stirrer. To this suspension was added 30 gms of KOH which had been ground up under ether. The reaction mixture was stirred for one hour while being cooled in an ice bath. The solids were removed by filtration and the ether solution dried over Na_2SO_4. The Na_2SO_4 was removed by another filtration step and 10 ml of methyl iodide (0.17 moles) was added to the ether solution of carbodiimide. A copious white precipitate formed immediately, and after letting the reaction proceed overnight the EPC·MeI was collected by filtration. The product was recrystallized from chloroform-ethyl acetate. Its infrared spectrum showed the 2130 cm^{-1} band characteristic of the -N=C=N- chromophore. Melting point was 106-107°C; literature reports[132] 106.5 - 107.5°C. Elemental analysis was as follows:

Calculated: C = 36.42%; H = 6.74%; N = 14.15%; I = 42.40%

Found: C = 36.36%; H = 6.87%; N = 14.42%; I = 42.66%

5. Synthesis of Poly-L-Lysine-ε-$\overset{+}{N}(CH_3)_3$

Poly-L-lysine is available from Pilot Chemicals or Miles Laboratories. In this synthesis poly-L-Lys·HBr from Pilot Chemicals was used with a molecular weight of about 100,000. This corresponds to a degree of polymerization of about 500. Four hundred eighteen mg of polylysine (2×10^{-3} moles of lysine) was dissolved in 20 ml H_2O in a 100 ml Erlenmeyer flask equipped with a magnetic stirrer. To this solution was added 1 ml of methyl iodide. The methyl iodide was ^{14}C-labeled, but at a very low specific activity (3.12×10^3 dpm/μmole), so that the concentration of the final product could be easily determined. The resulting reaction mixture was stirred and 1 M NaOH was added by pipette. The reaction could be followed roughly by the amount of NaOH added, and appeared to be complete in about five hours. The solvent was evaporated under reduced pressure. Twenty-five ml of water was added to the solid and the solution was dialyzed against 0.2 M NaCl (2 liters) overnight to remove all the iodide ions. After this dialysis the solution was dialyzed against distilled water for 2 days. The final solution was concentrated to about 20 ml and the concentration of the poly-L-Lys-ε-$\overset{+}{N}(CH_3)_3$ determined to be 0.8 μmoles Lys/10λ by measuring the radioactivity of an aliquot of this solution.

6. Reactions of 8-^{14}C-AMP-5' with Carbodiimides

8-^{14}C-AMP-5' was obtained from Schwarz BioResearch Co. The sample came in a 50% ethanol-water solution and had a specific activity of 42.5 mC/mmole. The procedure described for this reaction was used in most of the other reactions of this type. One ml (10 μC) of the AMP-5'-^{14}C solution was added to a 25 ml pear-shaped flask and evaporated to dryness under reduced pressure. Ten ml of a 0.025 M solution of unlabeled AMP-5' was prepared and the pH adjusted to 7. Five hundred λ of the cold solution was added to the dried AMP-5'-^{14}C in the pear-shaped flask. 2.4 mg of EPC (1.25 x 10^{-5} moles) was then added directly to 100 λ of the resulting ^{14}C-AMP-5' (0.025 M) solution and the reaction allowed to stand overnight at room temperature. Ten λ of the reaction mixture was then spotted on Whatman #3 chromatography paper. Two dimensional paper chromatography was then carried out for 36 hours with solvent A in the long dimension and solvent F in the short dimension. The paper was then exposed to X-ray film as described in the first part of this section.

7. Synthesis of 1,3-bis-(2-methoxyethyl) carbodiimide

1,3-*bis*-(2-methoxyethyl) urea was synthesized from the reaction of 2-methoxyethylamine and phosgene. The 2-methoxyethylamine (b.p. 91-92°) was distilled from NaOH pellets. Thirty-two gms (0.427 moles) of the amine and 100 ml of benzene were added to a 500 ml two-necked round-bottomed flask, which was equipped with a magnetic stirrer. This mixture was placed in an ice bath and 150 ml of a 12.5% phosgene in benzene solution (0.214 moles of phosgene) added dropwise, with a dropping funnel, to the cold amine solution. The reaction was very exothermic.

After the last addition of phosgene, the solution was refluxed for 4 hours. After cooling, 20 gms of NaOH (powdered) was added to the solution. The resulting mixture was filtered and the solvent evaporated under reduced pressure. The residue contained the 1,3-*bis*-(methoxyethyl) urea which was purified by recrystallization from an ether-petroleum ether solution. The final product (10 gms, m.p. 68-69°) was determined to be pure by thin-layer chromatography on silica gel with the solvent system acetone (80):benzene (15):water (5).

1,3-bis-(2-methoxyethyl) carbodiimide was prepared from the 1,3-bis-(2-methoxyethyl) urea by a procedure used by Sheehan and coworkers[132] to synthesize some other carbodiimides. The urea (35 gms, 0.2 moles) in 600 ml of CH_2Cl_2 and 120 ml of triethylamine was added to a 2-liter round-bottomed flask. This solution was cooled to below 5°C and 76.3 gms (0.3 mole) of p-toluene sulfonyl chloride in 450 ml of CH_2Cl_2 added. After this addition, the reaction mixture was allowed to warm to room temperature and then refluxed 3-4 hours. This solution was extracted three times with 300 ml portions of 40% aqueous K_2CO_3 and these portions then washed with CH_2Cl_2. The CH_2Cl_2 solutions were combined and the solvent evaporated under reduced pressure. The residual oil remaining was extracted with petroleum ether and the ether solution then washed with a small amount of water to remove a trace of unreacted urea still present. The petroleum ether was evaporated under reduced pressure and the pure 1,3-bis-(methoxyethyl) carbodiimide obtained by vacuum distillation, b.p. 70-75°C at 0.4 mm. The carbodiimide was shown to be pure by thin-layer chromatography in the acetone (80):benzene (15):water (5) solvent system. R_f for the carbodiimide is 0.85 and R_f for the urea is 0.63 on silica gel. Further evidence for the identity of the compound was provided by the infrared spectrum, which had a strong absorption at 2130 cm^{-1}, characteristic of carbodiimides, and by elemental analysis. Calculated for 1,3-bis-(methoxyethyl) carbodiimide:

Calculated: C = 53.10%; N = 17.72%; H = 8.86%

Found: C = 53.36%; N = 17.51%; H = 8.79%

The compound was also subjected to gas-liquid chromatography on a S.E. 30 column and NMR and mass spectra were obtained. All of the data were consistent with characterization of the compound as the desired carbodiimide.

8. Solvent Systems and Buffers

A. i-PrOH (70): NH_4OH (10):H_2O (20)

B. 95% EtOH (70):1 M NH_4OAc (30)

C. i-butyric acid (66):2 M NH_4OH (34)

D. t-BuOH (40):methylethylketone (30):formic acid (15):H_2O (15)

E. 85% saturated aqueous solution of NH_4HCO_3, 16 gms/100 ml; equilibrate 16 hours and run 3 hours

F. n-BuOH (50):HOAc (20):H_2O (30)

H. n-PrOH (55):NH_4OH (10):H_2O (35)

J. n-BuOH (40):HOAc (10):H_2O (50) , upper phase only

K. 95% EtOH (7):H_2O (3)

M. Acetone (80):benzene (15):H_2O (5)

N. 95% EtOH (7):NH_4OH (3)

P. Benzene (95):acetone (5)

R. Acetone (80):benzene (15):NH_4OH (5)

Electrophoresis buffers

pH 9.2: phosphate borate buffer. 120 gms of $Na_2B_4O_7 \cdot 10\ H_2O$ and 15.24 gms of $Na_2HPO_4 \cdot 7\ H_2O$ dissolved in 12 liters of water.

pH 3: ammonium formate buffer. Add 30 ml of formic acid to 4 liters of 0.05 M ammonium formate (12.61 gms) solution.

III. REACTIONS OF AMINO ACIDS WITH VARIOUS CONDENSING AGENTS

A. Introduction

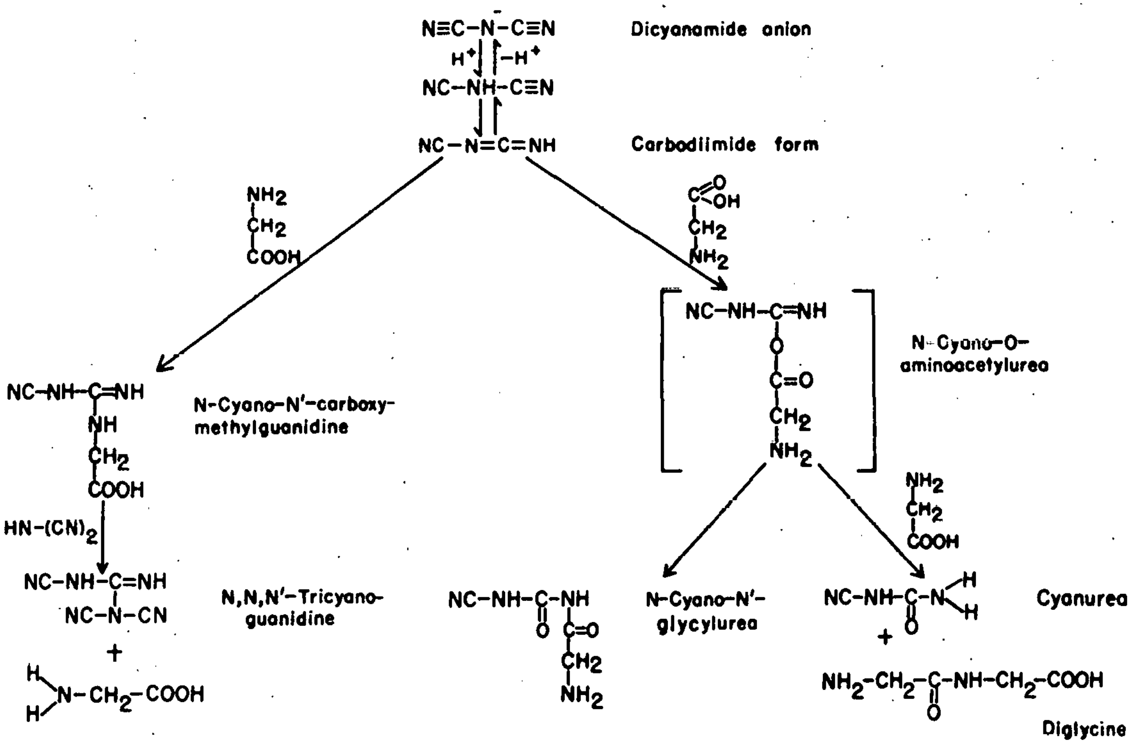
One of the goals of this work was to determine whether there exists interactions between the amino acid and nucleic acid systems that could provide a basis for how these two systems were coupled together in the evolutionary process to give the final result observed now--*i.e.*, the universal genetic code of contemporary living systems. The method used in this search for interactions between the systems was to condense the monomer of one system in the presence of the polymer of the other--*e.g.*, lysine in the presence of polyadenylic acid. The expectation was that the polymer would exert a catalytic effect on the condensation of the monomer.

However, before this type of experiment could be carried out, a condensing agent had to be chosen and the products of its reaction with the monomer had to be identified. The condensing agent should be one which would react in aqueous solution for two reasons. First, because the experiments should simulate conditions that are thought to have existed on the primitive Earth--*i.e.*, if interactions between the amino acids and nucleic acids were important in chemical evolution they would have taken place in aqueous solution. Also, the polynucleotides are quite soluble in water and quite insoluble in most aprotic organic solvents.

Since this work is related to the work with the nucleotides discussed in the previous section, the logical choice was to continue with the condensing agents that had already been used (Fig. 4). Thus, the condensation reactions of amino acids with carbodiimides and the cyanamide derivatives in aqueous solution were studied.

Cyanamide and its derivatives, dicyandiamide and sodium dicyanamide, were first used in simulated primitive Earth experiments as condensing agents by Steinman *et al.*[70,74,121,134] It was found that these compounds can effect dehydration condensation reactions in aqueous solution and that this type of reaction gives the best yields at lower pH's--*e.g.*, at pH 1 to pH 3. A number of biologically important linkages have been formed using these compounds--*e.g.*, phosphate esters, esters, pyrophosphates and peptide bonds. The reaction that gave the highest yields of products and was studied in the most detail was that of the amino acids to give peptides. The proposed mechanism for the reaction of glycine and sodium dicyanamide (DCA) is illustrated in Fig. 13. This reaction is postulated to proceed through a carbodiimide intermediate, and this intermediate then gives the expected carbodiimide reactions with glycine. The condensing agent, NaDCA, dimerizes in the acid solution and this side reaction is catalyzed by the presence of glycine.

Dicyanamide is a fairly strong acid with a pK of -0.8.[121] The crystalline hydrate of DCA can be isolated, but it is somewhat unstable.[151] When DCA is hydrolyzed with one mole of water cyanurea is formed, and if it is hydrolyzed further a second mole of water can be taken up to give biuret.[152,153] Free DCA is difficult to isolate



MUB 13012

Figure 13. Proposed mechanism of operation of dicyanamide (from Ref. 74).

since it tends to polymerize, but it is reasonably stable in solution, especially as its sodium salt. It is interesting that whereas DCA is a moderately strong acid, its parent compound, cyanamide, is a weak base with a dissociation constant[154] of the order of 10^{-11}.

Cyanamide upon hydrolysis gives urea, but slightly acid aqueous solutions of it are stable for several months.[155] Upon standing or heating both solid cyanamide and solutions of the compound tend to dimerize to give dicyandiamide (DCDA). This dimerization is accelerated by both acid and base with the optimum rate at pH 9.6.[156] Both cyanamide[157] and DCDA[154] when heated form the cyclic trimer of cyanamide, melamine.

Aqueous solutions of DCDA are neutral and decompose very slowly with time.[154] The hydrolysis of this compound yields guanylurea which is a strong base. It has been suggested that because of the electron-withdrawing capacity of the nitrile group of dicyandiamide, the hydrogen of the secondary amine would be quite labile in aqueous solution. This hydrogen would then be available to satisfy the basicity of the amidinium group, so that the most probable structure of DCDA would be that one which contains two amino groups. In addition, since this compound displays a high melting-point, salt-like character, and high temperature coefficients of solubility in water and alcohol, it may exist to some degree as a zwitterion.[160] The X-ray diffraction studies[164] of the solid DCDA have indicated that this compound does not exist in one simple chemical structure but is a composite of several structures illustrated in Fig. 14. It should be noted that one of the contributing forms of DCDA is actually a carbodiimide. This fact would support the proposed mechanism of reaction of these compounds.

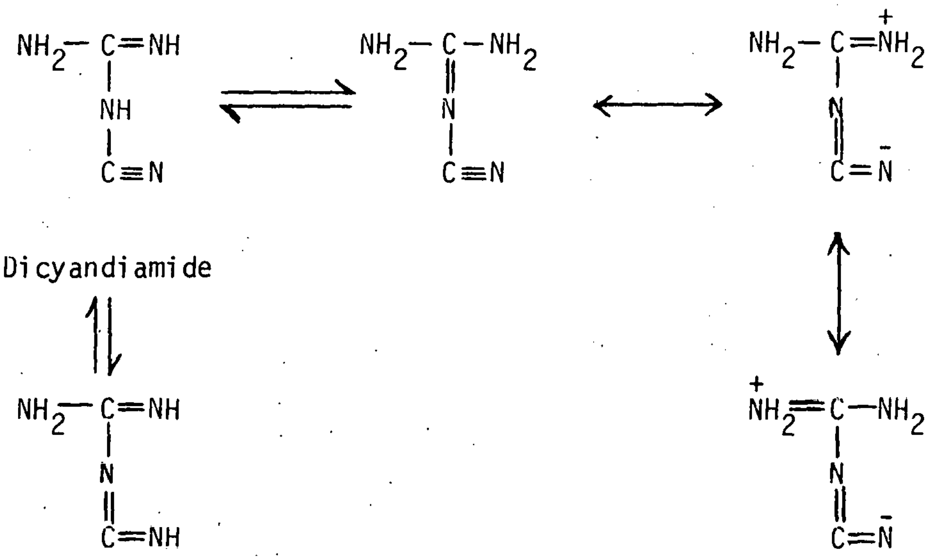


Figure 14. Structures contributing to the overall formulation of dicyandiamide.

Chemicals with active hydrogen are known to add across the nitrile group of dicyandiamide.[160] For example, hydrazine reacts with cyanoguanidine (DCDA) to give aminobiguanide.[161] DCDA also reacts with esters of acetoacetic acid, disubstituted malonic acid, or cyanoacetic acid to form various pyrimidines and barbituric acids.[158,162,163] 1-Amidino-O-alkylureas have been prepared from the reaction of dicyandiamide and alcohols with copper acetate as a catalyst.[165]

Carbodiimides, as the other class of condensing agents studied, have for some time been used routinely as reagents for peptide synthesis.[77,166] This type of reaction has generally been carried out with high yields in organic solvents with two suitably protected amino acids, one containing a free carboxyl and the other a free amino group. The major side reaction in this synthesis is the formation of the N-acylurea, the amount of which is increased by solvents such as dioxane and tetrahydrofuran,[133,167-169] or by heating.[170] With aromatic carbodiimides and monocarboxylic acids the N-acylurea formation greatly predominates.[133,170-173]

Since the reaction conditions are very mild and the acylureas are well-defined crystalline compounds, Zetzsche[174] has suggested carbodiimides as reagents for characterizing carboxylic acids and recommended di-(p-dimethylaminophenyl) carbodiimide as the compound most suitable for this purpose.[175,176] The acid may be regenerated from the N-acylurea by mild alkaline hydrolysis.

Aliphatic carbodiimides and carboxylic acids normally yield disubstituted ureas together with acid anhydrides,[177,178] but afford exclusively N-acylureas if organic bases (e.g., pyridine or triethylamine) are present.[172,179,180] Thus, gradual addition of carboxylic acids to dicyclohexylcarbodiimide in hot pyridine affords the acylurea as the main product.[170,172]

Another possible side reaction which apparently doesn't take place in peptide synthesis is the addition of the amino group to the carbodiimide. Ammonia and amines have been reported to react additively with carbodiimides to form the expected di-[181-184] and trisubstituted guanidines.

Thus, in the experiments to be discussed in the next section where the unprotected amino acid is reacted with carbodiimides, several side reactions are possible. This type of reaction is somewhat unique in that the usual procedure for peptide synthesis is to react protected amino acids to minimize side reactions and control the desired reaction, whereas, when the free amino acid is used these controls are absent. Needless to say, the reactions with the free amino acid are more inefficient, and this inefficiency is compounded when the reaction is carried out in aqueous solution. The mechanism and possible products from this reaction are illustrated in Fig. 15. The reaction is believed to proceed through an unstable intermediate,[130,144] the O-acylisourea, which may either rearrange, by way of a cyclic electronic displacement, to the stable N-acylurea, or react with another amino acid to give the peptide or anhydride. Support for this mechanism has been provided by the isolation of a cyclic O-acylisourea and by kinetic studies.[191]

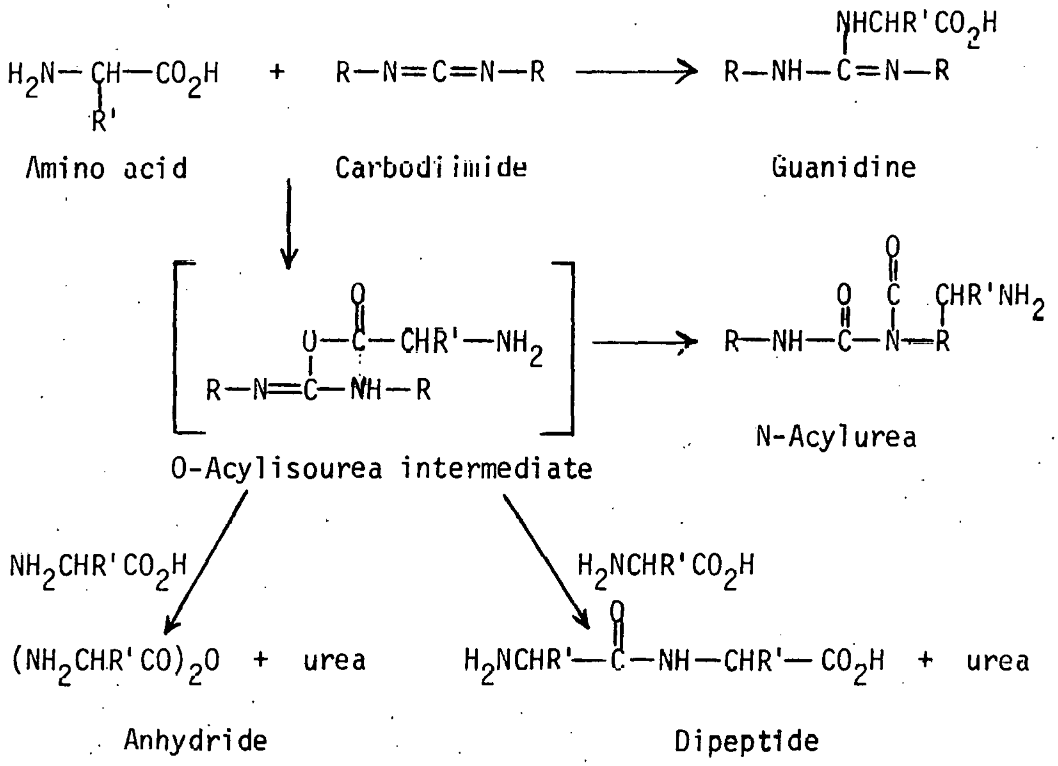


Figure 15. Mechanism and possible products from the reaction of an amino acid and carbodiimide.

B. Condensation Reactions of Glycine

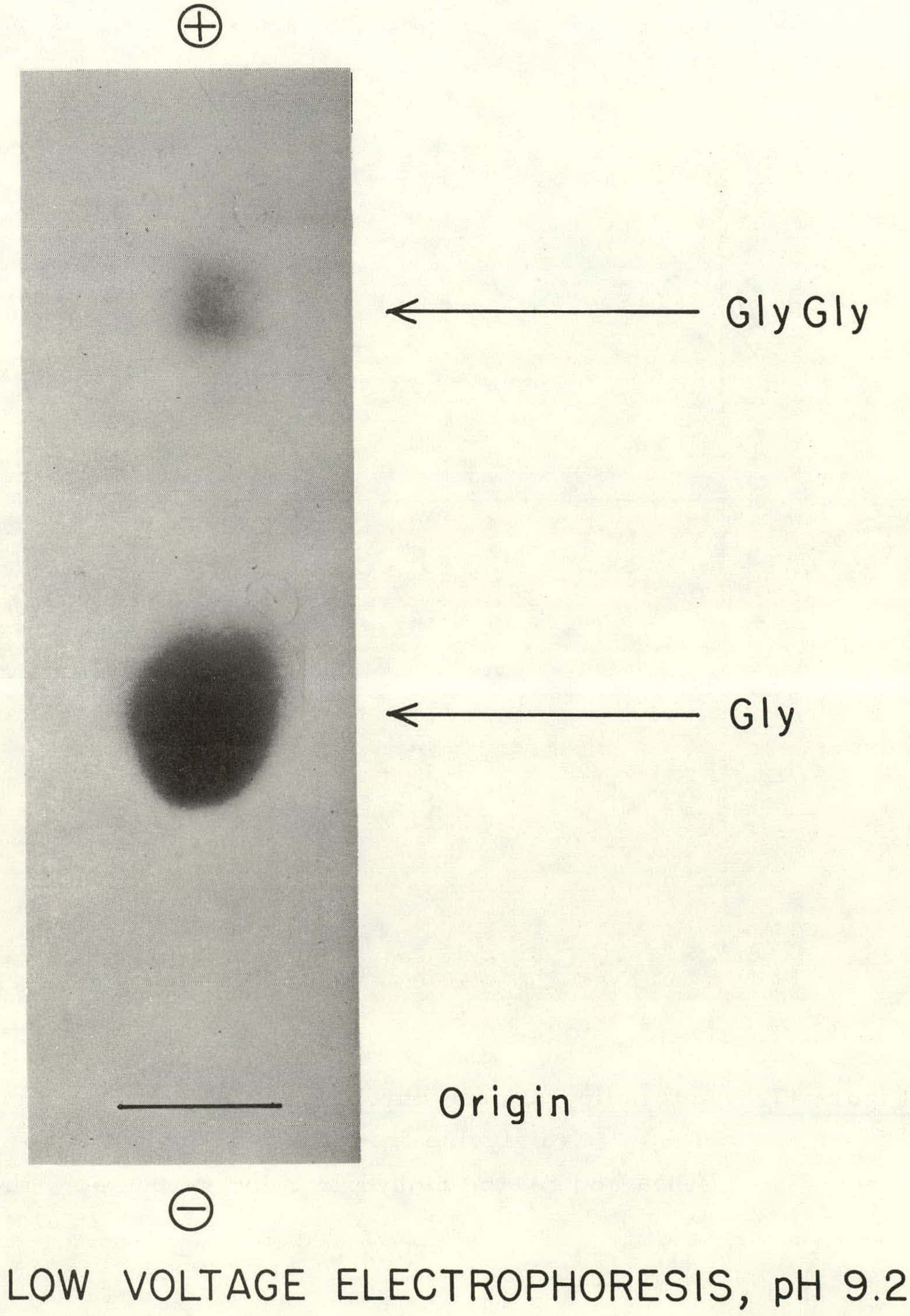
Glycine (Gly) is the simplest and most abundant of the naturally occurring amino acids. Since it has no side chain and is not sterically hindered in this respect it usually gives higher yields than the other amino acids in condensation reactions. It is for this reason that glycine has been studied as a model for the condensation reactions of amino acids. Glycine has a pK_A of 2.35 for the dissociation of its carboxyl group and a pK_A of 9.78 for its amino group.[192] Its dimer, glycylglycine, has pK_A values of 3.14 and 8.25,[192] so that it is apparent that there is a marked change in pK's in going from the monomer to the dimer. As a result of this difference in pK's, the two compounds are well separated by paper electrophoresis, whereas, this separation is difficult to accomplish by paper or thin layer chromatography.

1. Reaction with Cyanamide Derivatives

Cyanamide, dicyandiamide, and sodium dicyanamide are all able to effect dehydration condensation reactions in aqueous solution.[121] NaDCA has been reported to be the most effective condensing agent with DCDA being less reactive and cyanamide being the least reactive.[74,121] The reaction of glycine with cyanamide has been studied in detail by Halmann.[72] At pH 6.5, with long reaction times and heating to 60°C he reports yields of diglycine of about 1%. However, it is known that the cyanamide derivatives work best at lower pH's.[70,74,121]

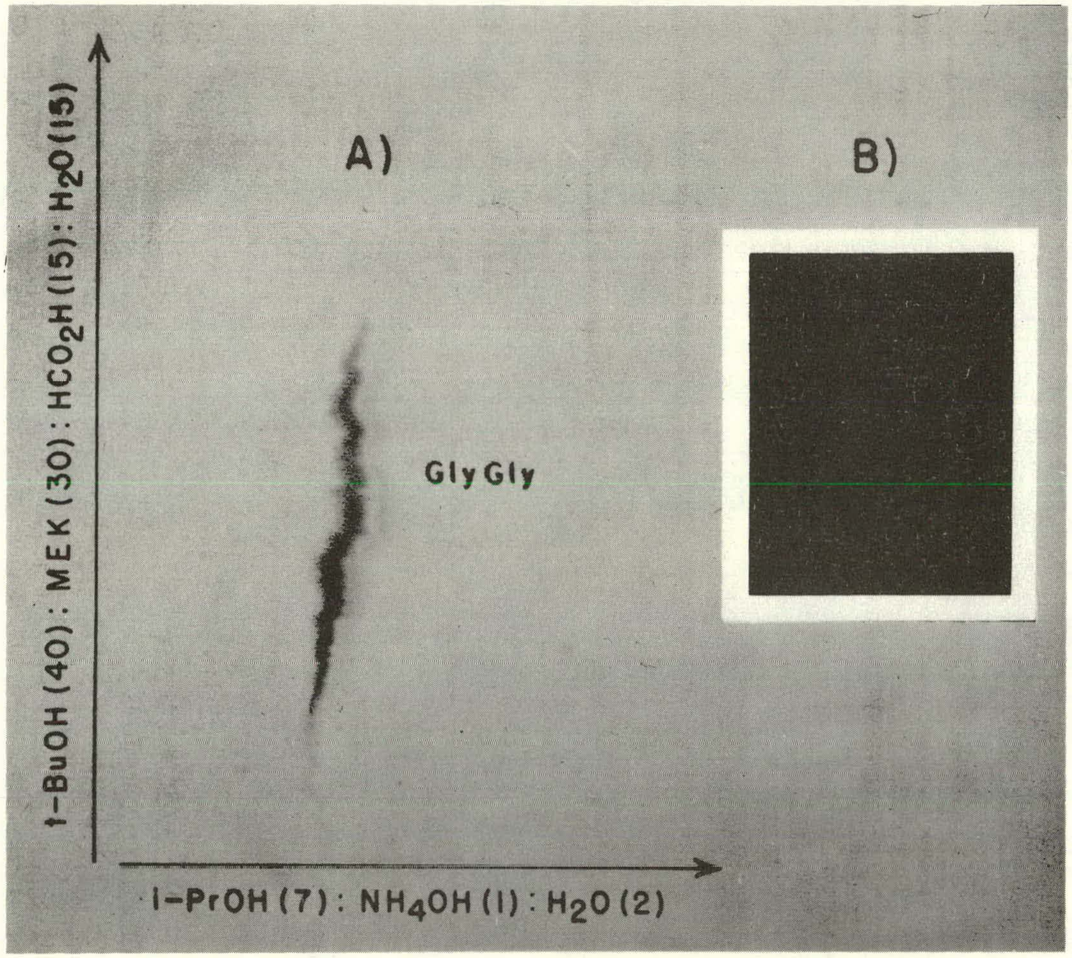
In the reaction of ^{14}C-glycine (0.1 M) with dicyandiamide (0.5 M) at pH 1, glycylglycine was formed in a 7.6% yield after a reaction time of one day. The products were separated by low voltage electrophoresis in a phosphate-borate buffer (pH 9.2). The autoradiogram of the electropherogram is shown in Figure 16. The resultant ^{14}C-diglycine was eluted from the paper and co-chromatographed with standard glycylglycine in an acidic and a basic solvent system. Figure 17 shows that the ^{14}C-compound on the autoradiogram and the ninhydrin color on the paper due to the standard glycylglycine are coincident. This confirms the identification of the GlyGly synthesized in the reaction.

In another experiment glycine (0.1 M) was reacted with sodium dicyanamide (0.5 M) at pH 1. This reaction was carried out under the identical conditions of the previous one and the yield of product with the same electrophoretic mobility as diglycine was about 20%. However, when this product was co-chromatographed with standard diglycine it was found that there were several products actually present. Thus, the overall yield of pure diglycine using DCA was about 6.5%, which is slightly less than that obtained with DCDA. The autoradiogram of this reaction is shown in Figure 18. The compound with a higher mobility than diglycine has been identified as N-cyano-N'-carboxymethylguanidine,[74] and it was formed in a 32% yield. The compound with the lowest mobility (8% yield) has been shown to be N-cyano-N'-glycylurea.[74] In this reaction no triglycine was detected as ascertained by two-dimensional co-chromatography of the products with standard triglycine.



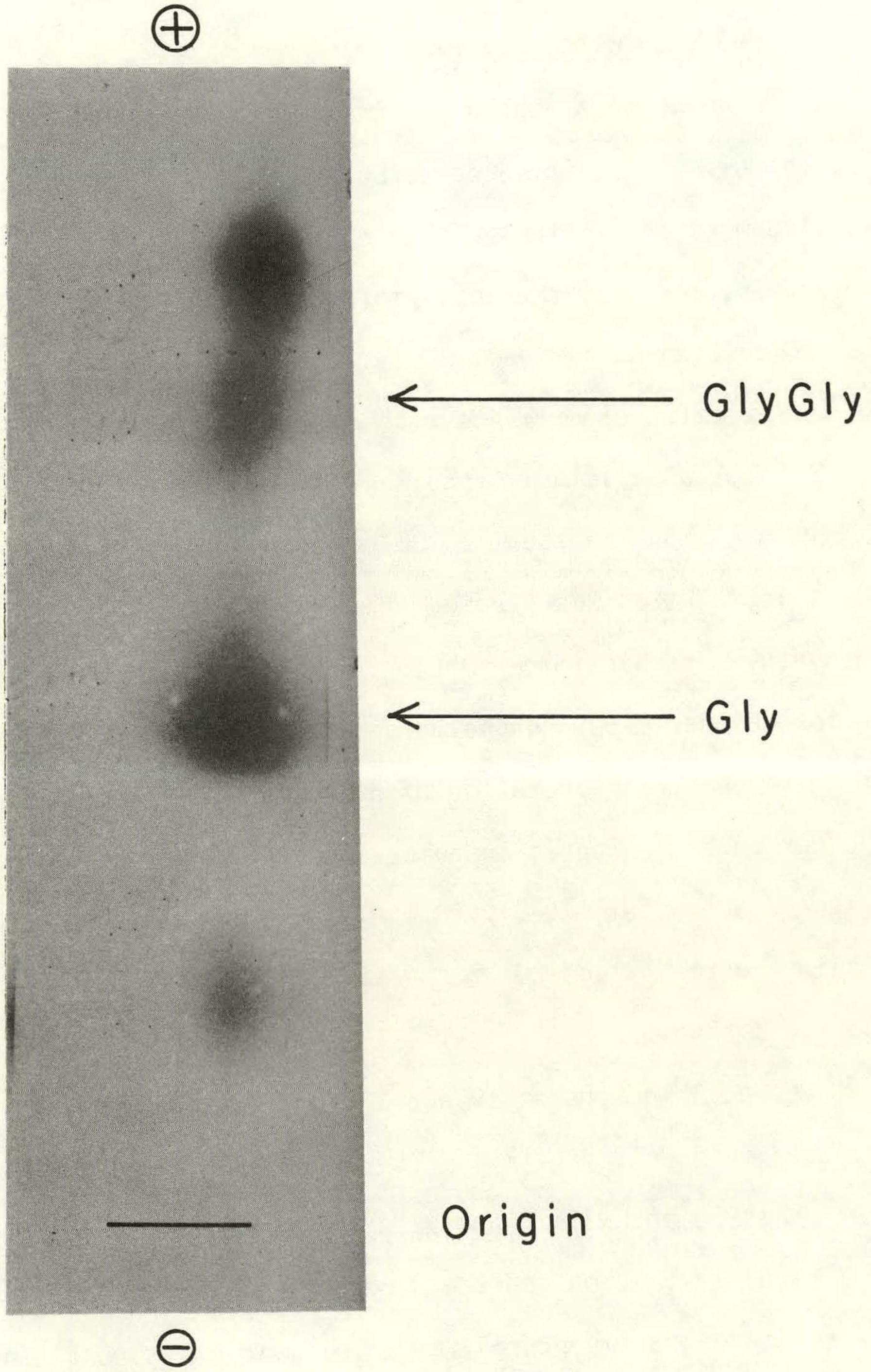
XBB 702-674

Figure 16. Autoradiogram of the electrophoretic separation of products from the reaction of ^{14}C-glycine (0.1 M) and DCDA (0.5 M) at pH 1.



XBB 702-681

<u>Figure 17.</u> Autoradiogram and chromatogram showing the coincidence of A) ^{14}C-diglycine from the reaction of ^{14}C-glycine and DCDA and B) the ninhydrin color spot of standard diglycine.



LOW VOLTAGE ELECTROPHORESIS, pH 9.2

XBB 702-679

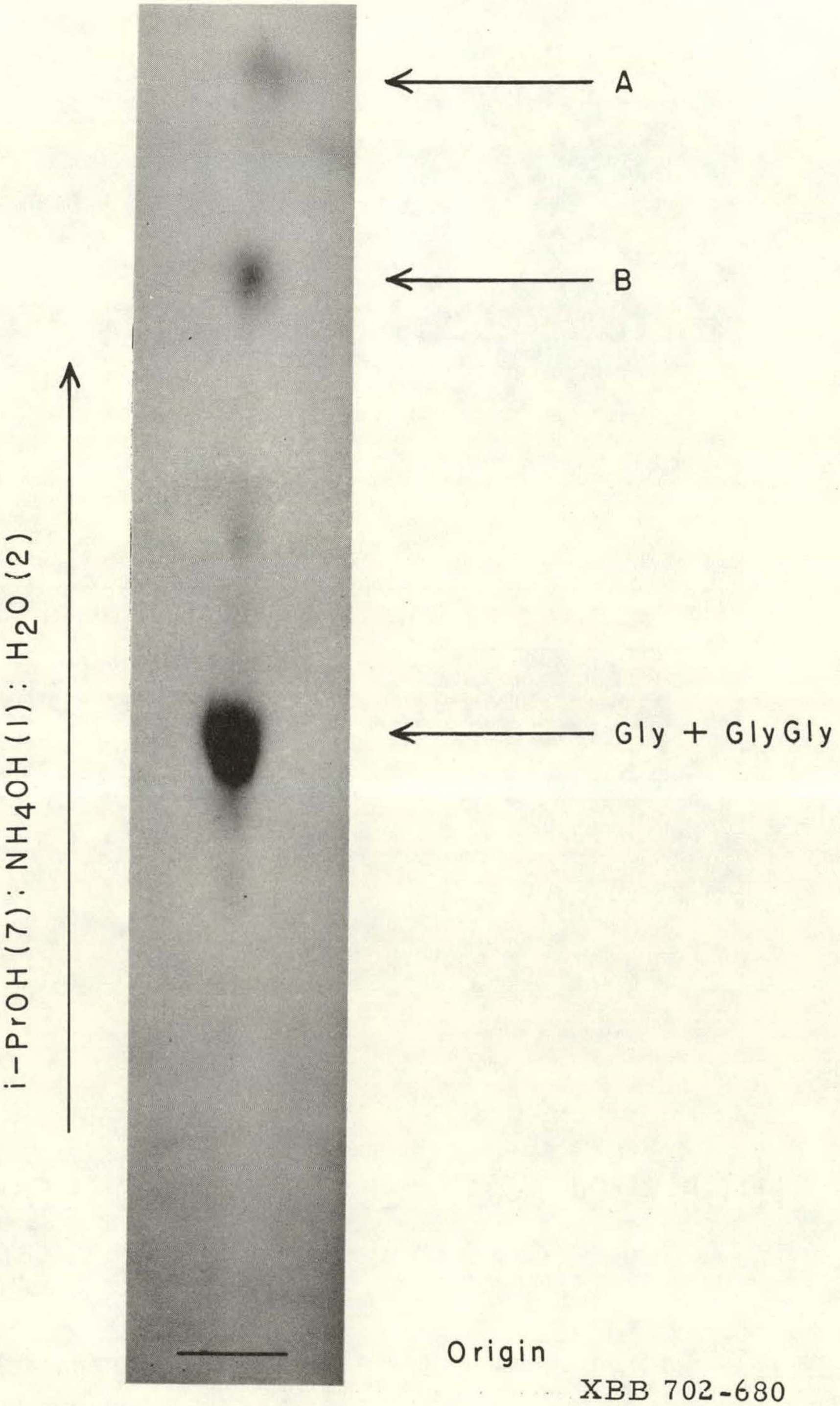
Figure 18. Autoradiogram of the electrophoretic separation of products from the reaction of ^{14}C-glycine (0.1 M) and NaDCA (0.5 M) at pH 1.

2. Reaction with Carbodiimides

It is known that peptide synthesis can take place in aqueous solutions using disubstituted carbodiimides.[166] One report states that oligomers of glycine up to eight units are formed in this type of reaction, and that the best yields are obtained at pH 5 with water-soluble carbodiimides.[193]

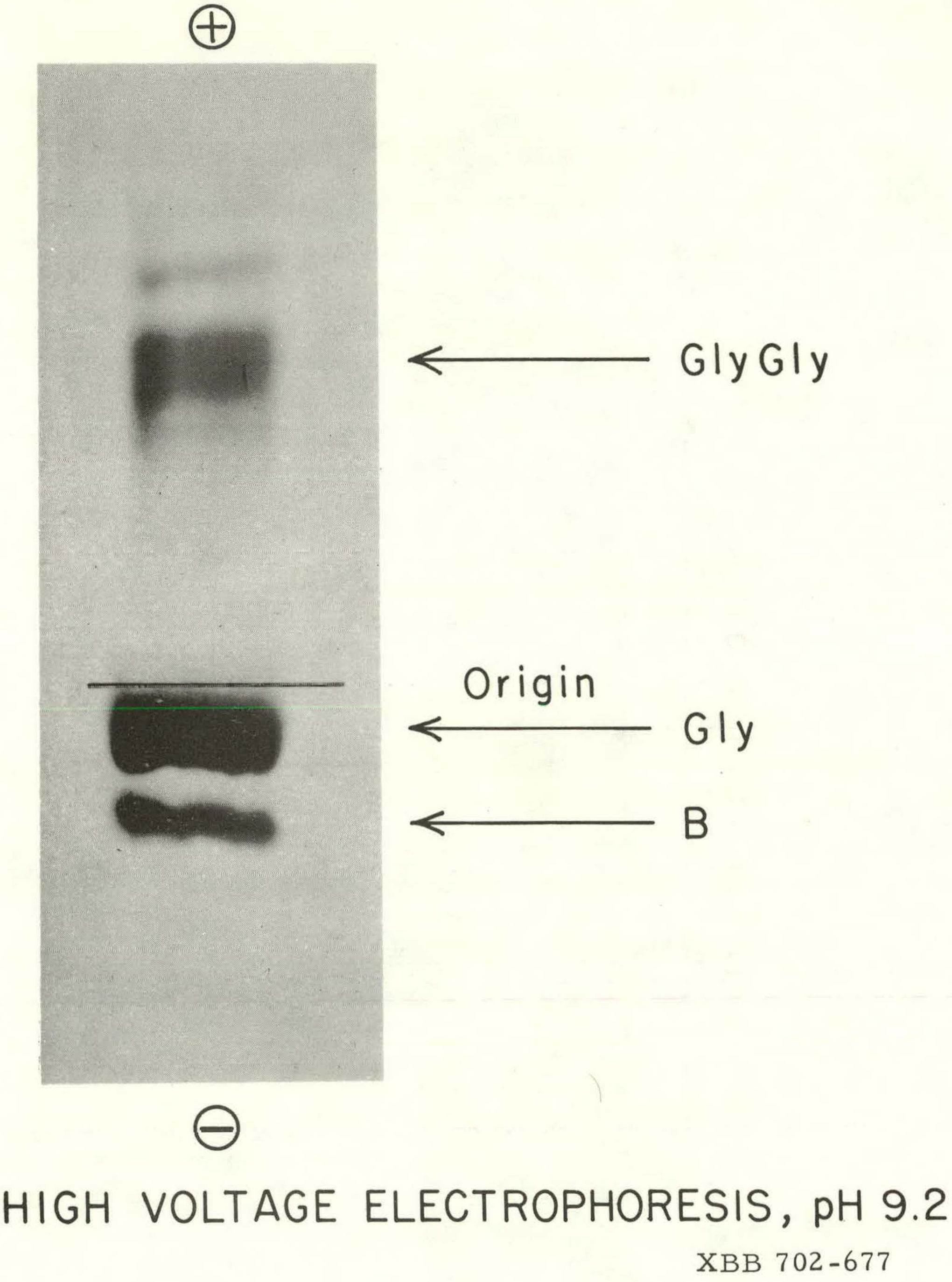
In the first experiments with glycine and water-soluble carbodiimides, cyclohexylethylmorpholinyl carbodiimide (CMC) was used. CMC gave diglycine in about a 10% yield at pH 3. In later experiments this reaction was studied in more detail using 1,3-bis(2-methoxyethyl) carbodiimide (MEC). It was found that for this reaction neither paper chromatography nor paper electrophoresis alone gave complete separation of products. The autoradiograms of the paper chromatographic and electrophoretic separation of products from the reaction of ^{14}C-Gly (0.1 M) and MEC (0.075 M) at pH 3 are shown in Figures 19 and 20, respectively. The reaction times were one day, as usual.

Compound B in Figures 19 and 20 has been identified as the N-glycylurea (Figure 21) of MEC. The evidence for this identification is that this compound is hydrolyzed readily under basic conditions (0.1 M NaOH) to glycine and the dimethoxyethylurea. N-acyl urea's are known to be easily hydrolyzed under mild basic conditions to the starting acids and ureas.[77] Also, the fact that this compound has a fairly high R_f (0.75) in a basic solvent system and a low mobility in electrophpresis in a basic buffer is a good indication that there

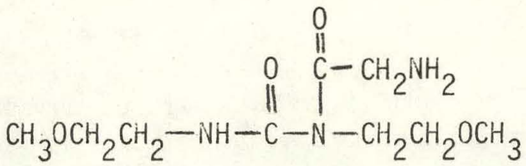


XBB 702-680

Figure 19. Autoradiogram of the paper chromatogram from the reaction of ^{14}C-glycine (0.1 M) and MEC (0.075 M) at pH 3.



<u>Figure 20</u>. Autoradiogram of the electrophoretic separation of products from the reaction of ^{14}C-glycine (0.1 M) and MEC (0.075 M) at pH 3.



$$\mathrm{CH_3OCH_2CH_2{-}NH{-}\overset{\overset{O}{\|}}{C}{-}\overset{\overset{\overset{\overset{O}{\|}}{C}{-}CH_2NH_2}{|}}{N}{-}CH_2CH_2OCH_3}$$

Figure 21. N-glycyl-N,N'-bis(2-methoxyethyl) urea.

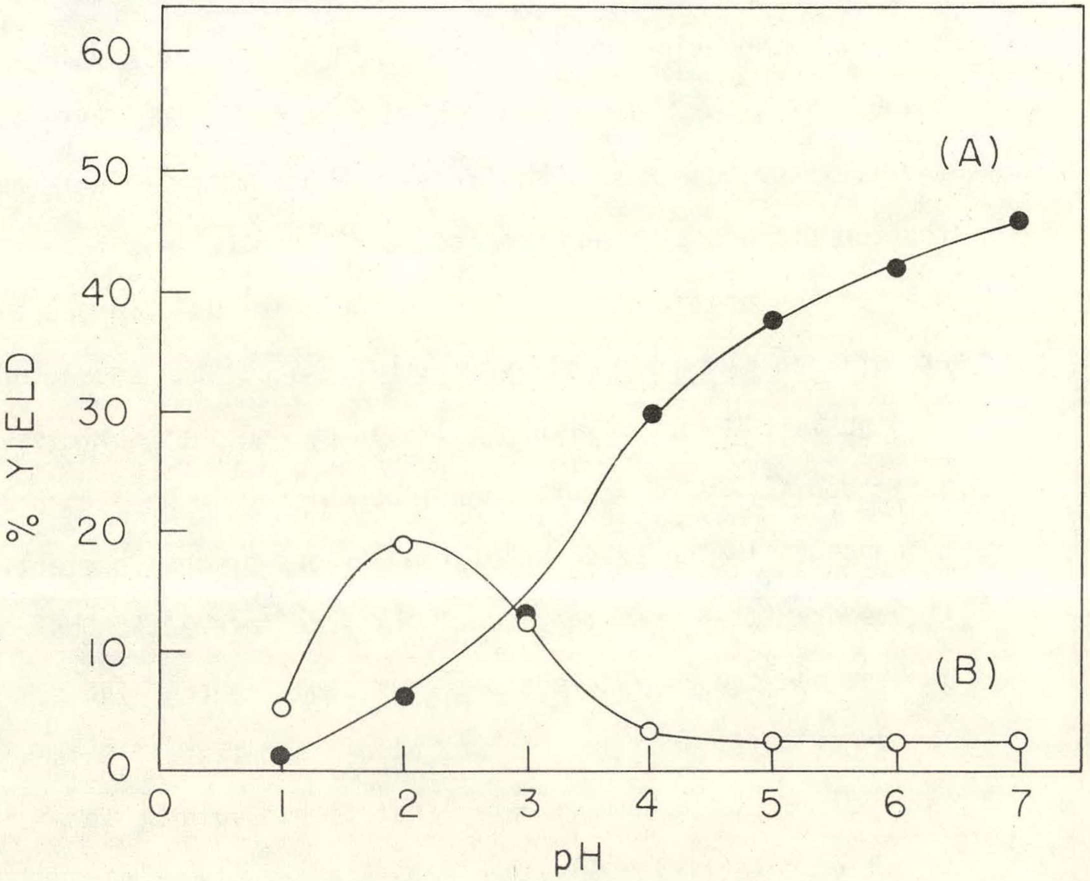
---

is not a free carboxylic acid group present. Finally, ^{14}C-labeled MEC was synthesized according to Equations 5 and 6 from ^{14}C-phosgene, and reacted with ^{3}H-glycine. The specific activity of ^{3}H-glycine was 51 times greater than that of the ^{14}C-MEC, to make the simultaneous counting of both isotopes on the liquid scintillation counter more accurate. Thus, the theoretical value for a product containing one mole each of ^{3}H-glycine and ^{14}C-MEC was $^{3}H/^{14}C = 51$, and the experimental $^{3}H/^{14}C$ ratio for compound B was about 49. This is conclusive evidence that there is one mole each of glycine and carbodiimide in compound B. Also, this product gives a positive ninhydrin test which is characteristic of primary amines and indicates that the amino group is free. All of the evidence taken together clearly indicates that compound B is the N-glycylurea of MEC.

The ^{14}C-diglycine was identified by the usual co-chromatography with standard diglycine and the resulting coincidence of the ninhydrin and ^{14}C-spots. Also, there was no radioactivity in the diglycine area when ^{14}C-MEC was used.

Compound A has not been identified. It is formed in about a 5% yield. It is thought that this product could be an adduct resulting from the addition of two molecules of carbodiimide to one of glycine, where the carbodiimides would add to the carboxyl and amino groups, forming the N-acylurea and trisubstituted guanidine structures, respectively. However, due to the low yields of this product and to the fact that it has chromatographic properties similar to those of the 1,3-bis-methoxyethylurea, the double label experiment (^{3}H-Gly and ^{14}C-MEC) did not provide a definitive answer.

The reaction of glycine and MEC was studied as a function of pH, and the results of this study are illustrated in the graph in Figure 22. It can be seen that the yield of diglycine reaches a maximum at pH 2, whereas the yield of side product B, the N-acylurea, keeps increasing as the pH is raised. This would indicate that the acyl migration from the oxygen to nitrogen is facilitated by the increase in pH, as previous studies have predicted.[170,172]



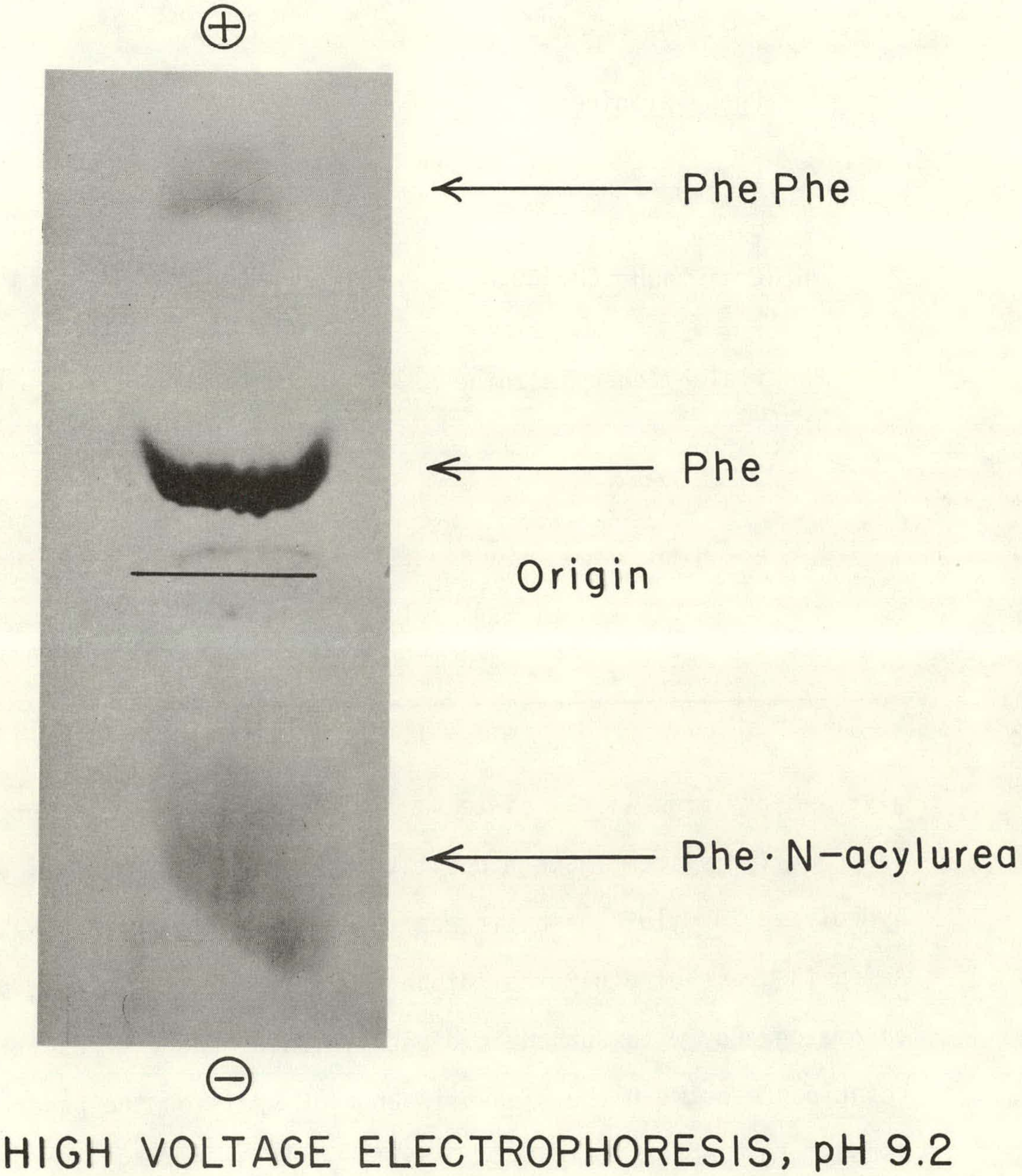
XBL 702-5053

Figure 22. The effect of varying the pH on the yield of (A) N-glycylurea, and (B) diglycine, in the reaction of glycine (0.1 M) and MEC (0.075 M).

C. Reactions of Phenylalanine with Carbodiimides

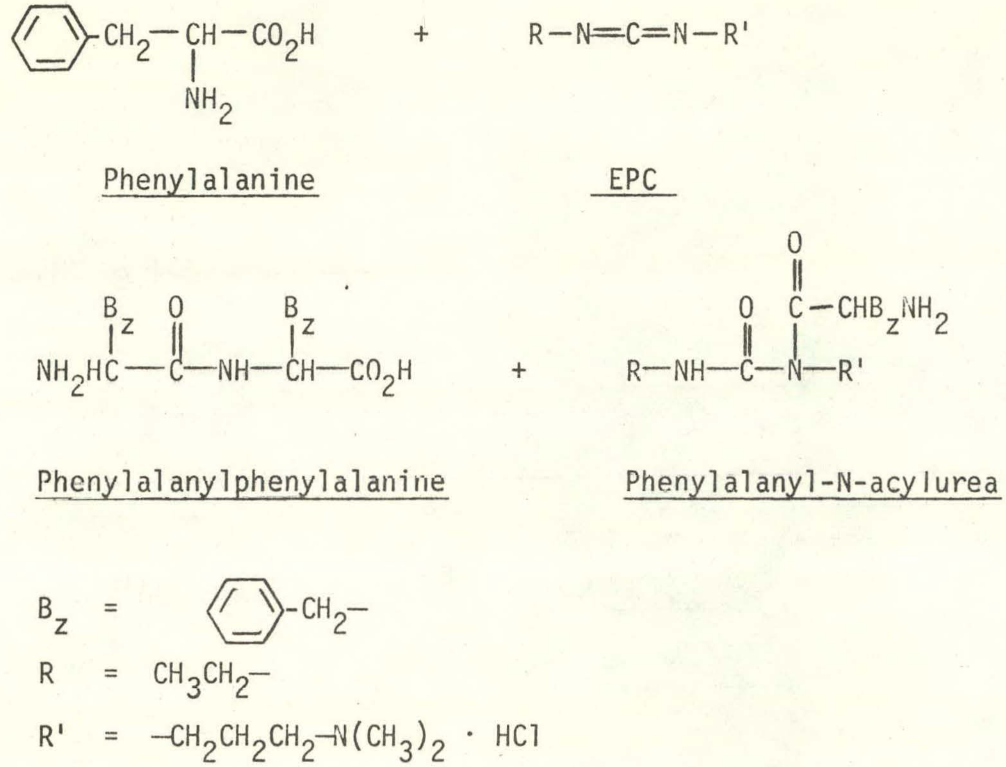
Phenylalanine is an aromatic amino acid and is not as common as glycine in naturally occurring proteins. It is one of the more complex amino acids and because of its benzene moiety is less soluble in water. Its reactions with carbodiimides were studied for two reasons. One was to determine whether this aromatic amino acid reacted differently than glycine, and the other was to determine whether the aromatic part of the amino acid could interact with the aromatic bases of the polynucleotides. Since the codon for phenylalanine is UUU (three uridines linked by phosphate bridges) and the anticodon is AAA it was thought that this would be a good amino acid to condense in the presence of the various homopolynucleotides. Thus, it was predicted that there would be some interaction between phenylalanine and the polynucleotides that could enhance the self-condensation of this compound. It was also hoped that if this sort of catalysis could be demonstrated that there would be some detectable amount of selectivity involved--_e.g._, that polyuridylic acid would be more catalytic than polycytidylic or polyguanylic acid. However, in this case neither the predicted catalytic effect nor any sort of selectivity was observed.

The autoradiogram of the electrophoretic separation of products from the reaction of ^{14}C-phenylalanine (0.005 M) and ethyl-dimethylaminopropyl carbodiimide (EPC, 0.05 M) is reproduced in Figure 23. The yield of diphenylalanine is 10% and the yield of the N-phenylalanylurea is about 30%. The reaction is illustrated in Equation 8. The N-acylurea was characterized in a manner analogous to the



XBB 702-676

Figure 23. Autoradiogram of the electrophoretic separation of products from the reaction of phenylalanine (0.005 M) and EPC (0.05).



(8)

---

glycyl-N-acylurea--i.e., since it migrated towards the cathode in a basic buffer system, gave a positive ninhydrin test and was easily hydrolyzed in dilute base, it was demonstrated to be the N-acylurea. As in the case of diglycine, diphenylalanine was identified by co-chromatography with authentic diphenylalanine and the resulting coincidence between the ninhydrin and ^{14}C-spots on the paper and X-ray film, respectively.

The reaction of phenylalanine with EPC in the presence of various polynucleotides and polypeptides was studied in some detail. As mentioned, this study gave completely negative results. These reactions were usually carried out at 0° - 5°C to maximize any interactions between the systems. In the case where this reaction was carried out in the presence of the polynucleotides, the yield of diphenylalanine decreased from about 10% to 8% and this decrease was the same whether polyadenylic or polyuridylic acid was used. It is thought that this decrease probably results from an increased rate of hydrolysis of the carbodiimide, since it is known that phosphates do catalyze the hydrolysis of carbodiimides.[135]

Phenylalanine was reacted with EPC in the presence of the copolypeptide, poly-L-Phe-L-Lys · HBr . This copolymer was used since it had hydrophobic parts, the phenyl group of phenylalanine and the hydrocarbon portion of the lysine sidechain, and a hydrophylic part, the ε-amino group of lysine which made it water-soluble. Thus, it was thought that since phenylalanine also has hydrophobic and hydrophylic parts that it would interact with the copolymer so that when EPC was added the amount of dimer would be increased. This turned out not to be true and the amount of dimer again decreased as the amount of copolymer increased from a 1/1 to 10/1 molar ratio (moles of Phe in the polymer to free Phe).

In another series of experiments this same reaction was carried out in the presence of both polynucleotides and polypeptides. This was an attempt to simulate a possible crude ribosome-type structure where the polynucleotide would be complexed with the basic polypeptide and the reactants would be concentrated on the surface of this structure. It has been reported[96] that coacervates made from polyadenylic acid and polylysine have this sort of catalytic activity as judged by increased rates for several enzymatic reactions that have been studied in their presence. However, in this case the effect of adding various combinations of polyadenylic acid, polyuridylic acid, or both complexed with either polylysine or copoly-Phe-Lys, was to lower the yield of dimer.

In view of all these negative results, which were discouraging to say the least, the study of the reactions of phenylalanine under these types of conditions was abandoned. Instead, it was thought more worthwhile to study another system where the interactions are known to exist. This system is the lysine:polyadenylic acid one, and the interaction is predominently, if not totally, one of electrostatic attraction.

D. Reactions of Lysine with Carbodiimides

After the unsuccessful attempts to demonstrate a catalytic effect by the polynucleotides on the condensation of phenylalanine, it was thought that this could more easily be demonstrated in the case of lysine. The reason for this is that lysine is a basic amino acid and in aqueous solutions around neutrality would exist with a net positive charge. Since the polynucleotides have one negative charge per base, the phosphodiester bridge (pK_A approximately 1), there should be an electrostatic attraction between these polymers and lysine. However, the reaction of lysine, which has two amino groups and one carboxyl group, with a carbodiimide was expected to be more complex than that for glycine or phenylalanine, and for this reason it was not studied earlier.

The fact that basic polypeptides, e.g., polylysine and polyarginine,[149,150] and several small cationic molecules, like spermine and spermidine,[197] interact with polynucleotides, provided some evidence that lysine also could interact. The expectation was that the lysine molecules would be attracted to the polynucleotide and in effect be concentrated on them so that when the condensing agent was added there would be an increase in the amount of dilysine formed over that when the polymer was absent.

In this search for interactions which could have led to the coupling of the nucleic acid and amino acid systems, the only obvious one is the interaction between the basic polypeptides and the polynucleotides. This interaction is quite important in biological systems, e.g., in the complexing of histones and DNA.[198] It also could play some role in the structure of ribosomes since many of the polypeptides contained in this structure are basic.[199] The fact that polylysine exhibits a marked selectivity by interacting more strongly with polynucleotides having a high adenine-thymine content should be mentioned again.[149] Since one of lysine's codons is AAA this fact seemed suggestive of a possible relationship between this amino acid and its codon.

However, before this relationship could be investigated, the reaction of lysine with carbodiimides in aqueous solution had to be studied in detail. The carbodiimide chosen for this study was the 1,3-bis-(2-methoxyethyl) carbodiimide (MEC) since it is uncharged and moderately water-soluble. From previous experience it was realized that a charged carbodiimide would probably disrupt any electrostatic interaction, so that an uncharged condensing agent was a necessity. As expected, the reactions of lysine were more complicated than those of glycine and phenylalanine.

The pK_A values for lysine are 2.16, 9.18, and 10.79 for the carboxyl, the α-amino and ε-amino groups, respectively. The pK values for L-Lys-L-Lys are 3.01, 7.53, 10.05, and 11.01, so that in this case the separation of these two compounds was best accomplished by high voltage electrophoresis at pH 3. At this pH the dimer has three positive charges and about one-half of a negative charge, and the monomer has two positive and approximately one negative charge. Thus, the dimer migrates somewhat faster towards the cathode than the monomer, although this difference is not as pronounced as with glycine or phenylalanine and their dimers. Lysine and dilysine could not be separated by paper chromatography in the usual solvent systems or by electrophoresis in the basic buffer system.

The reaction of lysine with MEC is illustrated in Figure 24. It can be seen that there is another side reaction possible in this system that was not possible for glycine or phenylalanine. This is the cyclization of lysine through attack of the ε-amino group on the carbonyl carbon to form the seven-membered lactam, 3-aminohomopiperidone-2. This type of cyclization has been observed for the methyl ester of N-α-tosyllysine in ammonia and alcohol, and for the analogs, ornithine and α,γ-diaminobutyric acid.[194-196] It was found that the five-membered lactam was formed in the best yield (88%) and the seven-membered lactam in the least yield (49%), which indicates that the larger ring is less easily formed. However, the cyclization of the lysine N-acylurea in 0.1 N NaOH at 70°C is very nearly

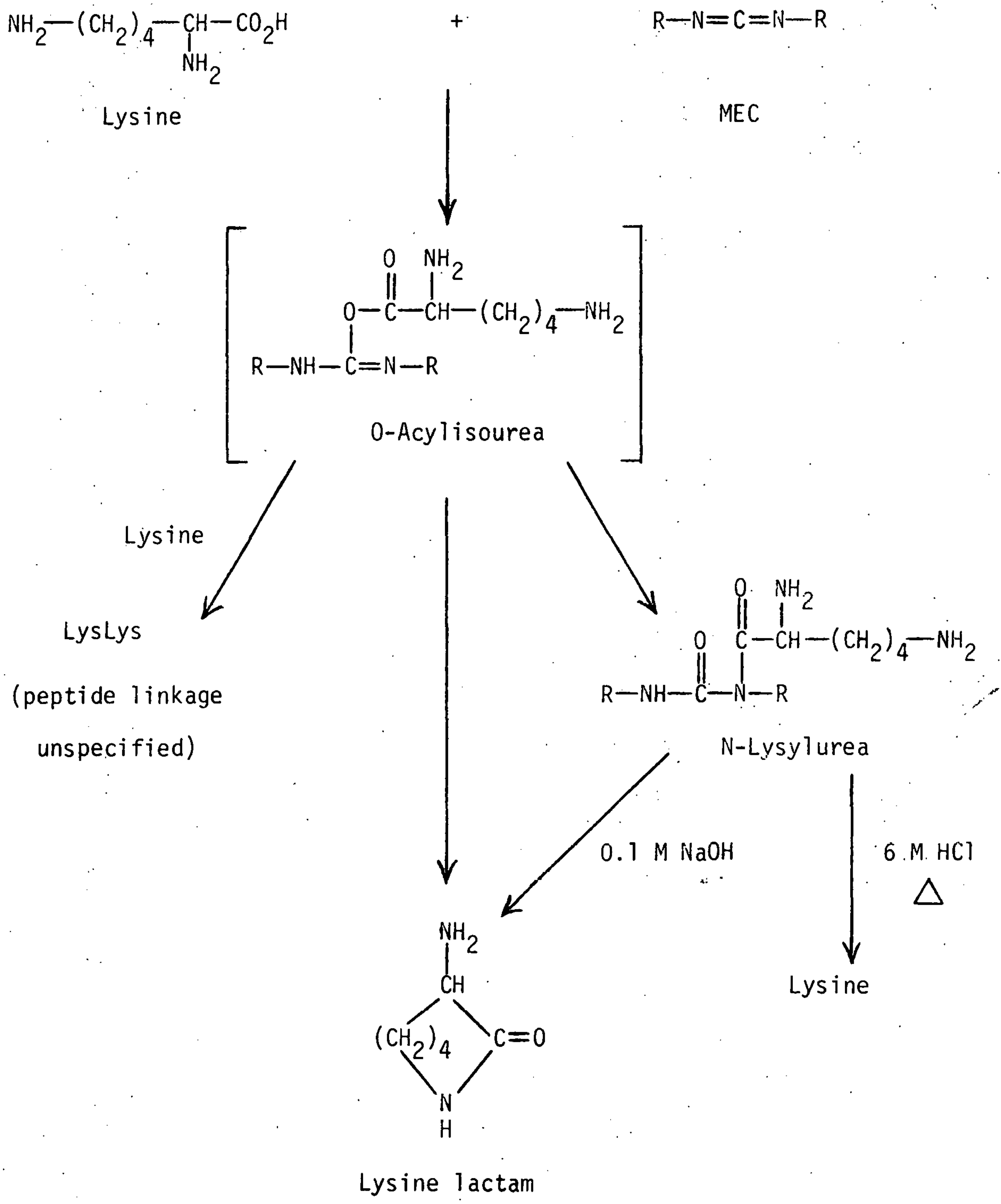


Figure 24. Reaction of lysine and 1,3-*bis*-(2-methoxyethyl) carbodiimide.

R = $CH_3OCH_2CH_2^-$.

quantitative with only 1 or 2 percent of lysine being produced. The lysine N-acylurea is stable to heating at 70°C in 1 N HCl for one hour, but is hydrolyzed to lysine in 6 N HCl when heated to 100°C for twelve hours.

The lactam is quite stable, and is unaffected by heating in 0.1 N NaOH or 1 N HCl for short periods of time. Heating at 100°C in 6 N HCl for twelve hours only partially hydrolyzes this compound to lysine. The hydrolysis was best carried out at 70°C in 4 N NaOH,[194] and was about 70% complete after twelve hours.

There were two major difficulties encountered in the identification of the products from the reaction of lysine and MEC. The first problem was in the electrophoretic separation of these products. Figure 25 shows the results of the electrophoresis at pH 3 (ammonium formate buffer) of the reaction of ^{14}C-Lys (0.025 M) and MEC (0.125 M). There appear two products, A and C, which are seemingly quite well separated from unreacted lysine. When this separation was first observed it was thought that product A was dilysine since they had exactly the same electrophoretic mobility. However, this turns out not to be the case as was demonstrated when the ^{14}C-label was changed from lysine to MEC. Figure 26 shows the electropherogram of the reaction of lysine and ^{14}C-MEC, and it can be seen that both products A and C contain the carbodiimide. Thus, A and C must be adducts of lysine and carbodiimide.

At this point a new method of separation was tested and found to be quite satisfactory. This was thin-layer chromatography (TLC) with silica gel.[200,201] It was discovered that both A and C contain

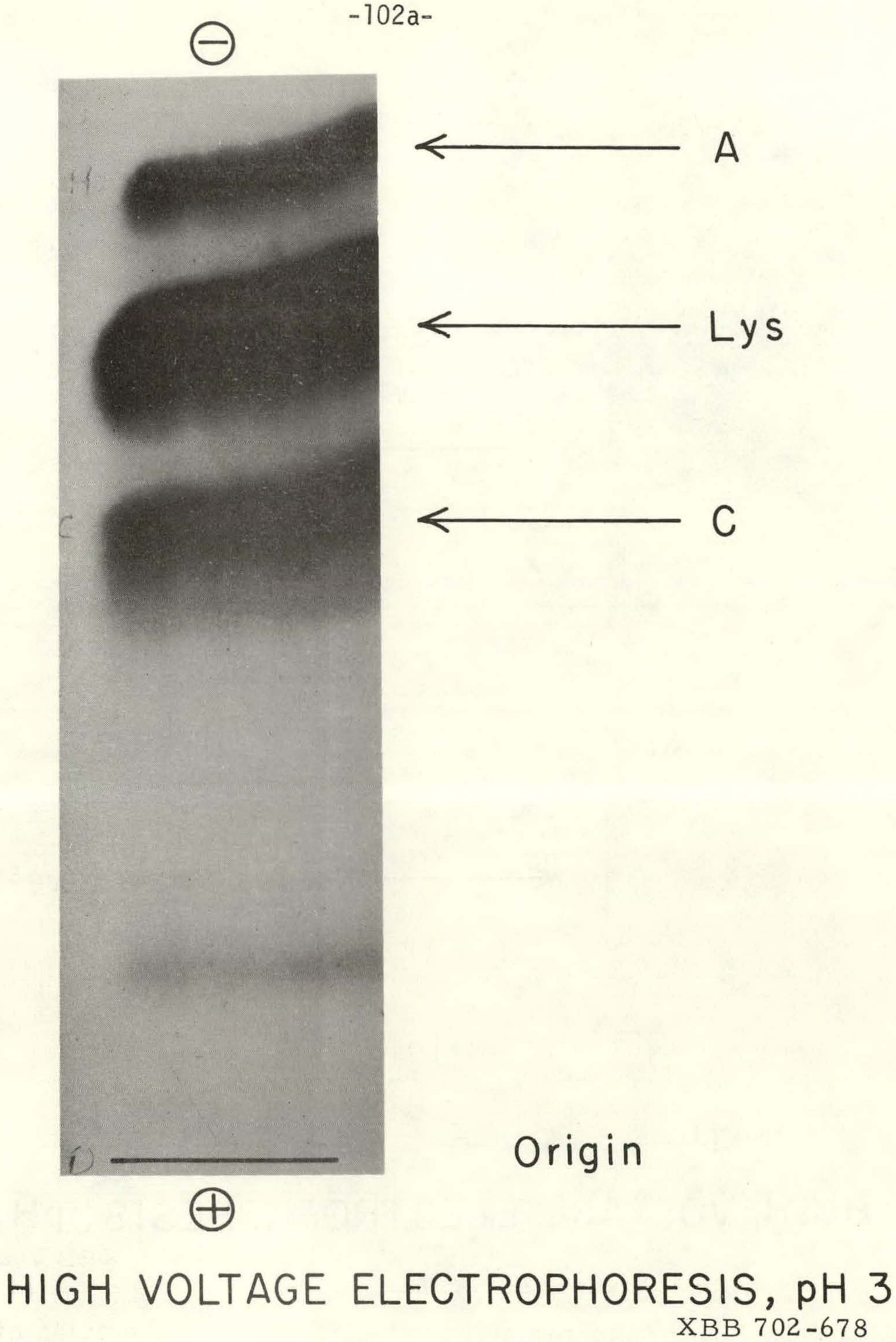
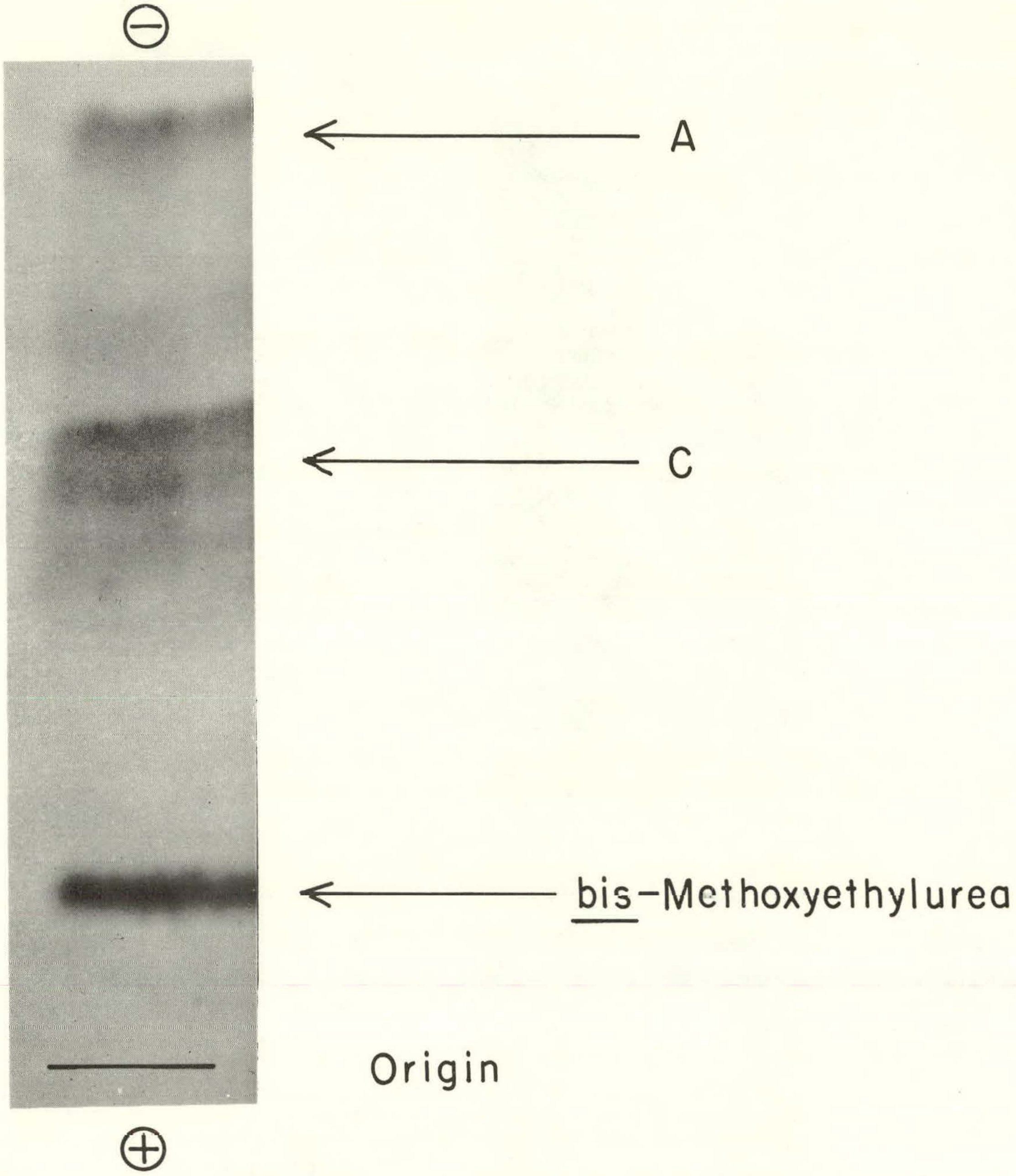


Figure 25. Autoradiogram of the electrophoretic separation of products from the reaction of ^{14}C-lysine (0.025 M) and MEC (0.125 M)



HIGH VOLTAGE ELECTROPHORESIS, pH 3

XBB 702-675

Figure 26. Autoradiogram of the electrophoretic separation of products from the reaction of lysine (0.025 M) and ^{14}C-MEC (0.125 M) at pH 7.

two major peoducts with the same R_f's in the solvent system 95% EtOH (7): NH_4OH (3). At first it was not thought possible that these products were the same, but after finding that they had the same R_f's on TLC in several solvent systems (R, N, Q), and in both cases the hydrolysis products were the same, it was concluded that they were identical. The products were the N-acylurea and the lactam of lysine which were formed in about a 5 to 1 ratio, respectively. It is not understood at this time why the electrophoresis gave such a result. There are several possible reasons, such as, overloading the paper with compound, interaction of the compounds during electrophoresis, or even tautomeric structures existing of the same compound. However, since it was reasonably certain that this anomaly did not arise from simply spotting too much sample on the paper, the electrophoretic separation procedure was abandoned in favor of thin-layer chromatography.

It should be mentioned that before it was realized that the products in A and C were the same, considerable effort was spent in trying to isolate these products and react them with 2,4-dinitrofluorobenzene (DNFB). DNFB is a common reagent used for determining free amino groups and in protein end-group analysis.[201-203] The procedure that was tried was: 1) to separate the products by preparative electrophoresis, 2) to elute products from the papers, 3) to separate the products in A and C on TLC, 4) to elute these products from the TLC plates and react with DNFB, 5) to separate again on TLC, and 6) to elute DNP derivatives and determine the moles of DNP per mole of lysine from the UV absorbance at 360 nm and the specific activity of ^{14}C-lysine. This project was unsuccessful for two main reasons. One was that the compounds proved very difficult to elute from silica gel, and even after heating the silica gel suspension in 0.04 N HCl for several hours only about 20% of the products could be recovered. The other problem was that the silica gel could not be removed completely from the aqueous solution even after methanol was added and centrifugation carried out for 30 minutes. Thus, when the products were reacted with DNFB, with some silica gel present, only a small amount of reaction occurred.

Although the DNFB method was unsuccessful in the above case, it was used to check the specific activity of both ^{3}H-Lys and ^{14}C-Lys solutions. The reaction is illustrated in Figure 27.

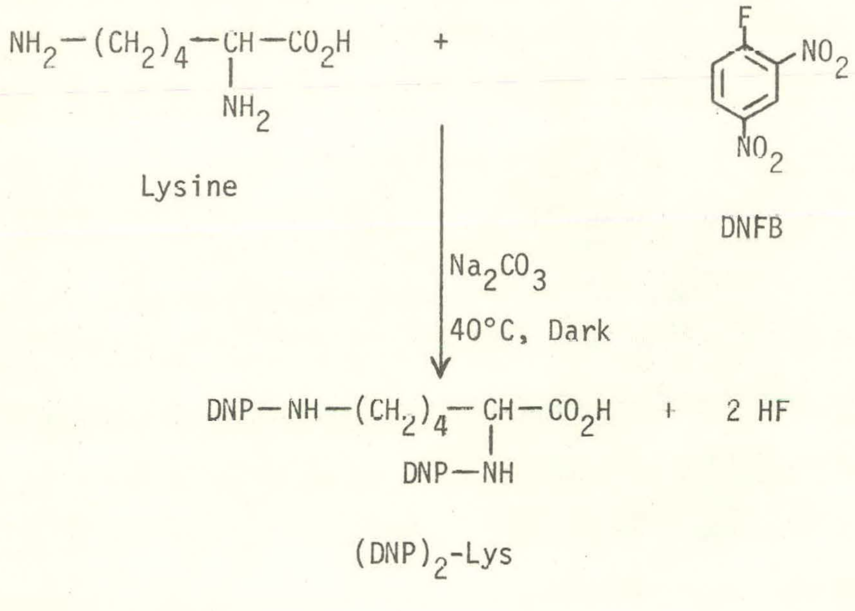


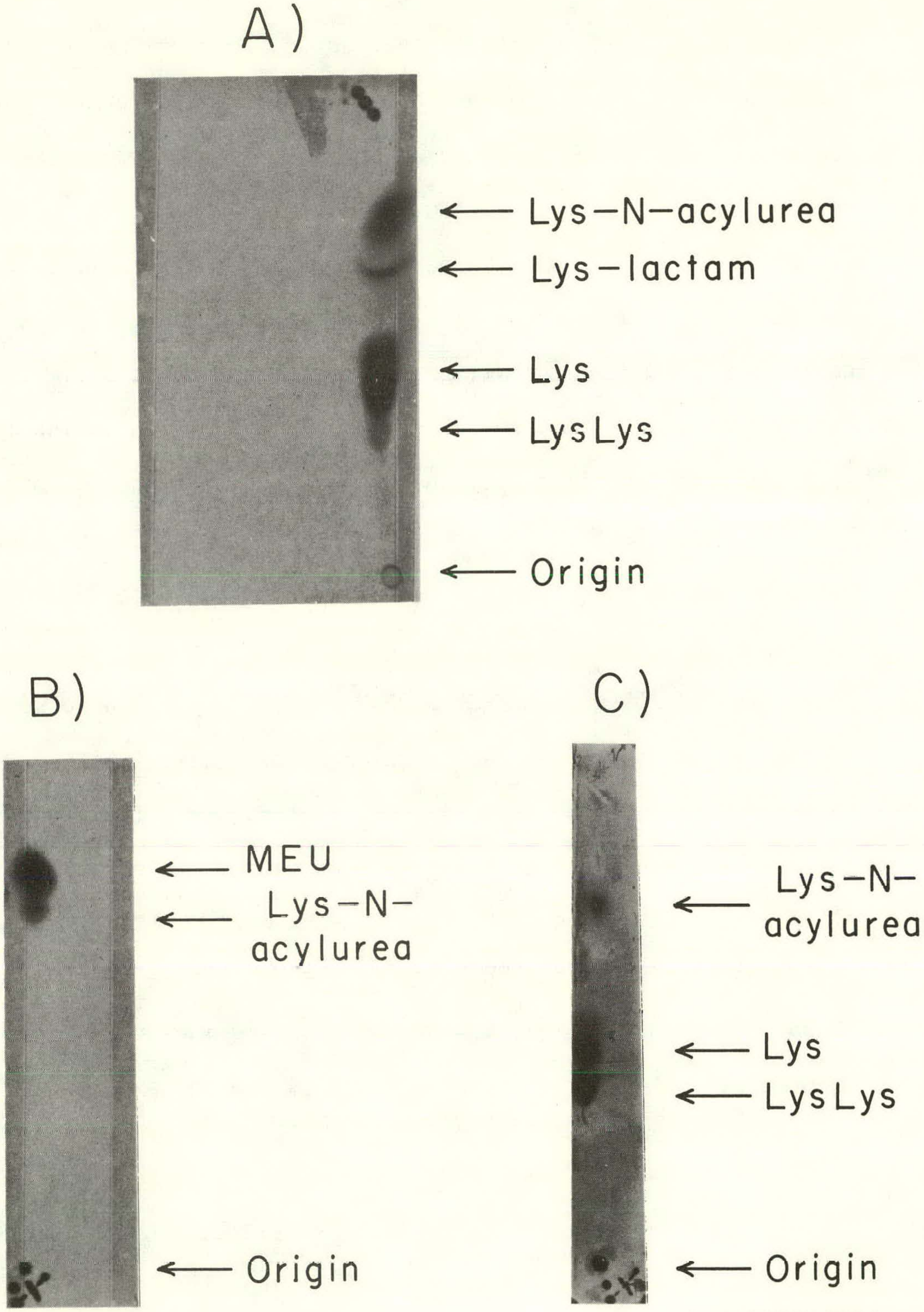
Figure 27. Reaction of lysine with 2,4-dinitrofluorobenzene. DNP stands for 2,4-dinitrophenyl.

The procedure was to separate the reaction products by thin-layer chromatography, elute the desired $(DNP)_2$-Lys and determine the number of moles of DNP and the amount of radioactivity. The molarity of the DNP solution was determined from the optical density at 360 nm ($\varepsilon_{360} = 1.64 \times 10^4$ per DNP) and the radioactivity was determined by liquid scintillation counting. The di-DNP-lysine could be eluted from the silica gel with water quite easily, indicating that it was the free amino groups that caused lysine to bind so tightly to the silica gel in the previous elution attempts.

The products from the reaction of lysine and MEC were separated on TLC as shown in Figure 28. The yield of the N-acylurea was about 30%, the lactam 8% and dilysine 2%, in the reaction of ^{14}C-lysine (0.025 M) and MEC (0.05 M) at pH 2. It can be seen in this figure that when the reaction is carried out with ^{14}C-MEC only the N-acylurea has both ^{14}C and lysine, where the ninhydrin positive response indicates lysine's presence. An experiment was also carried out where ^{3}H-Lys and ^{14}C-MEC were reacted. It was found that the product thought to be the N-acylurea had the correct ^{3}H to ^{14}C ratio for a one-to-one adduct of lysine and MEC. The theoretical value for this ratio was 15.9 and the experimental was 14.6.

1. Reaction of lysine in the presence of polynucleotides

Once the products from the reaction of lysine and 1,3-bis-(2-methoxyethyl) carbodiimide had been identified, the next step was to determine what effect polyadenylic acid would have on their



XBB 702-999

<u>Figure 28.</u> Autoradiograms from the TLC (silica gel, 95% EtOH: NH_4OH, 7:3) of the reaction of A) ^{14}C-Lys + MEC, and B) Lys + ^{14}C-MEC.

yield. It was thought that since the two systems are oppositely charged it would be possible for lysine to be attracted to the polyadenylic acid. If this attraction were strong enough, the polymer could act as a catalyst by concentrating the amino acid on its surface. Thus, when a condensing agent was added the amount of self-condensation would be increased by the presence of the polymer. The model for this interaction is illustrated in Figure 29.

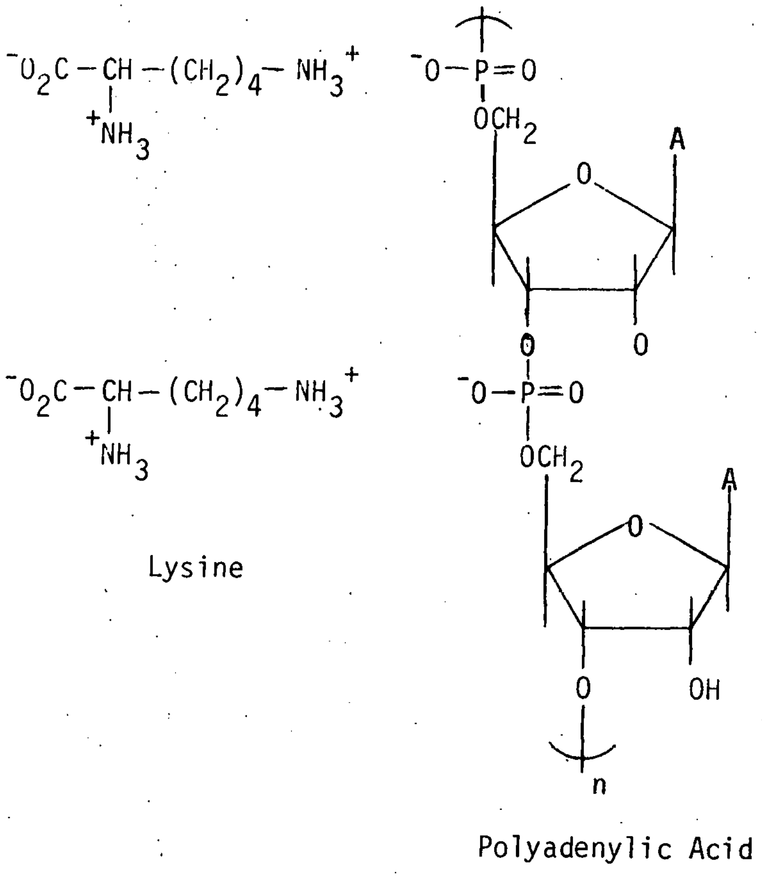


Figure 29. Model for the interaction of lysine and polyadenylic acid.

In the experiments designed to test this hypothesis, lysine was reacted with MEC in the presence and absence of polyadenylic acid. All other conditions for the reactions were the same--_i.e._, the temperature, concentrations, and pH. When this reaction was first studied in completely aqueous solution, using electrophoresis as the means for separating the products, it was found that there was an increase in the amount of product A (Figure 25) when polyadenylic acid was present. In view of the difficulties with electrophoresis that were encountered, this result is not readily interpretable. However, it has been discovered, subsequently, that polyadenylic acid does enhance the amount of dilysine that is formed when the reaction is carried out in an acetone-water (4:1) solvent. Table V summarizes the results of these experiments. The most pronounced effect of the polynucleotide is to

---

Table V. Percent yields of products from the reaction of ^{14}C-L-lysine (0.002 M) and MEC (0.004 M) in the presence and absence of polyadenylic acid. The solvent is 80% acetone-20% water, and the pH is about 7.

	LysLys	Lys	Lys Lactam	Lys-N-Acylurea
Polyadenylic Acid	2.5	62.2	5.3	30.0
----	0.8	58.0	4.2	37.0

3

triple the amount of dilysine that is formed. The amount of lactam is slightly increased and that of the N-acylurea slightly decreased in the presence of polyadenylic acid. In these experiments polyadenylic acid and lysine were precipitated in a 1 to 1 complex from a solution of 75% dimethylformamide-25% water. This precipitate was then purified and used directly in the reactions discussed.

There is now some experimental evidence that a catalytic system could exist between adenylic acid and lysine, where the polymer of either monomer would catalyze the polymerization of the other monomer. However, it must be remembered that this catalytic activity for both polyadenylic acid and poly-L-Lys-ε-$\overset{+}{N}(CH_3)_3$ was demonstrated under somewhat special conditions--*i.e.*, the solvent for the reaction was acetone and water. Acetone was added to the aqueous solution to reduce the dielectric constant of the solution, and cause the monomer to be complexed more strongly to the polymer.

Two main problems exist in evaluating these results in terms of significance for chemical evolution. These are that neither the acetone-water solvent nor the specific carbodiimide used could have existed to any appreciable extent on the prebiotic Earth. However, the carbodiimide was used as a model for more plausible primitive Earth condensing agents, and the acetone-water solvent used as a model for what could happen if the monomers and polymers were concentrated in some type of membrane structure--*e.g.*, a coacervate, or on a clay or mineral surface. If these models are correct, then the catalytic effect that has been demonstrated would have been important for the prebiological formations of polynucleotides and polypeptides.

C. R. Woese has put forth the hypothesis[112] that evolution of the cell began with two very general kinds of polymers, polynucleotides that were purine-rich and polyamino acids composed largely if not solely of basic amino acids. The one type of polymer is viewed as catalyzing, in particular ways, the synthesis of the other, and vice versa. According to this model, translation began as a "direct templating" and in addition, "translation" was initially a reciprocal matter, not unidirectional as it is now. The evidence presented here would tend to support such a model. However, the conceptual difficulties in going from such a model to the actual translation process which takes place in living systems are extremely large.

E. Summary

The condensation reactions of glycine in aqueous solution with water-soluble carbodiimides and the cyanamide derivatives, dicyandiamide and sodium dicyanamide, were studied. In the reaction of glycine with 1,3-bis-(2-methoxyethyl) carbodiimide (MEC), the two major products were identified as the N-glycylurea and diglycine. When this reaction was studied as a function of pH (from pH 1 to 7) it was found that the N-acylurea formation was enhanced as the pH increased, whereas, the maximum amount of diglycine was formed at pH 2. With the cyanamide derivatives, diglycine formation (5-7%) was observed at pH 1.

The reaction of phenylalanine with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EPC) was studied in the presence of polyadenylic and polyuridylic acids, poly-L-Phe-L-Lys-ε-$\overset{+}{N}(CH_3)_3$, and various complexes of the polypeptides and polynucleotides. The products from this reaction were the dimer and N-acylurea of phenylalanine, and no catalytic effect was observed for the polymers.

The reaction of lysine with MEC in aqueous solution was investigated and the products identified as the N-lysylurea, the ε-lactam of lysine and dilysine, where the interpeptide linkage is unspecified. When the reaction was carried out in an 80% acetone-20% water solvent in the presence of polyadenylic acid it was found that the yield of dilysine (2.5%) was three times greater than it was in the absence of the polymer. It is believed that polyadenylic acid exerts a catalytic effect on this reaction by complexing lysine via an electrostatic interaction. This catalytic effect could have been important in the prebiotic formation of basic polypeptides.

F. Experimental Section

1. Synthesis of ^{14}C-1,3-bis-(2-methoxyethyl) carbodiimide

^{14}C-Phosgene was obtained from New England Nuclear with a specific activity of 1 millicurie per 33 mg. Three millicuries of phosgene were transferred by vacuum line into a 25 ml pear-shaped flask containing 10 ml of β-methoxyethylamine. The reaction of phosgene with the amine was very fast, and after an hour at room temperature the reaction mixture was transferred to a 250 ml flask. One gm of NaOH in 20 ml of ether was added, and the resulting mixture stirred magnetically for 30 minutes. The NaOH was removed by vacuum filtration and the solution evaporated to dryness with a rotary evaporator. A small amount of white solid remained after the evaporation, which was the ^{14}C-1,3-*bis*-(2-methoxyethyl) urea (MEU). This ^{14}C-labeled urea was mixed with 3 grams of unlabeled *bis*-methoxyethylurea and recrystallized from diethylether and petroleum ether. The specific activity of the ^{14}C-MEU was determined by weighing out 10 mg of the product, dissolving this in 100 ml of ether and counting 100 μl of the ether solution. The specific activity was found to be 1.724×10^5 μC/mole, and would not change when the urea was converted into the carbodiimide.

3.17 gm of ^{14}C-MEU (0.018 mole) was dissolved in 100 ml of CH_2Cl_2 and 11.1 ml of Et_3N. The resulting solution was cooled to 0°C in an ice bath and boiling chips and 6.86 gm of *p*-toluenesulfonyl chloride (0.036 moles) in 75 ml of CH_2Cl_2 were added. After this addition the reaction mixture was allowed to come to room temperature and then refluxed for 3-4 hours. The solution was then extracted three times with 40 ml portions of a 40% aqueous K_2CO_3 solution. The aqueous portions were combined and extracted with CH_2Cl_2. All of the CH_2Cl_2 solutions were combined and evaporated. The residual oil left was extracted with three 50 ml portions of petroleum ether, and these solutions combined and evaporated. The remaining ^{14}C-MEC was purified by vacuum distillation, and about 360 μl (20% yield) were obtained. The product was determined to be pure by thin-layer chromatography in solvent M.

2. Preparation of lysine: polyadenylic acid complex

L-Lysine·HCl was obtained from CalBiochem Co., ^{14}C-lysine from Schwarz, and polyadenylic acid from Miles Chemical Co. Fifty μC of ^{14}C-Lys were added to 0.0244 gm of L-Lys·HCl (0.15 mmoles) in 4 ml of a dimethylformamide (DMF)-water (3:1) solution. The ^{14}C-Lys solution was added to 20 ml of another 75% DMF-H_2O solution of polyadenylic acid (36.7 mg, 10^{-4} moles of AMP). A white precipitate was formed immediately upon mixing, and the resulting suspension stirred gently with a magnetic stirrer for 1 hour. The precipitate, a complex between lysine and polyadenylic acid, was collected by centrifugation. It was purified by re-suspension and subsequent centrifugation in 75% DMF-H_2O and then acetone. The poly A:Lys complex was dried over P_2O_5 under vacuum and the ratio of AMP to Lys was determined to be 1 from the UV absorption of AMP and the specific activity of ^{14}C-Lys. The extinction coefficient for the AMP units in poly A at 259 nm was 9.8×10^3, and the specific activity of Lys was 7.44×10^{11} dpm/mole.

3. Reaction of ^{3}H-Lys and 2,4-dinitrofluorobenzene

^{3}H-Lys was obtained from Schwarz BioResearch, and 2,4-dinitrofluorobenzene (DNFB) from Sigma Chemical Co. This reaction was carried out to determine the specific activity of ^{3}H-Lys, since tritium-labeled impurities are much more difficult to detect than ^{14}C-labeled ones. The procedure that was used was one developed by Levy and Chung.[203]

Forty μl of ^{3}H-Lys (1 μmole) were placed in a 1 ml test tube covered with aluminum foil. 2.5 μl of a 10% solution of DNFB in acetone and 5 μl of a 2 M Na_2CO_3 solution added. The aluminum foil was then folded over the test tube so that no light would pass through and the reaction mixture placed in a water bath heated to 40°C. The reaction was allowed to proceed for one hour with agitation. The solution was then extracted with three 0.5 ml portions of ether to remove any unreacted DNFB. The remaining aqueous mixture was separated by thin-layer chromatography (silica gel, solvent R). The N,N'-DiDNP-Lys (R_f 0.35) was identified by co-chromatography with a commercial sample. The specific activity of the initial lysine solution was then determined by eluting the DNP_2-Lys from the silica gel with 3 ml of water and determining the molarity of the dinitrophenyl groups from the optical density at 360 nm ($\varepsilon_{360} = 1.64 \times 10^4$ per DNP). Since the lysine concentration was one-half that of the DNP groups, the specific activity was calculated from the number of moles of lysine in an aliquot of this solution and the radioactivity as determined by liquid scintillation counting. The specific activity of this ^{3}H-Lys solution was found to be 6.26 x 10^6 dpm/μmole. This value was very close to that obtained by counting the ^{3}H-Lys (0.025 M) solution directly, indicating that the ^{3}H-Lys was radiochemically very pure.

4. Liquid scintillation counting of tritium and carbon-14 simultaneously

The liquid scintillation counting[127] was done on a Packard tri-carb spectrometer which has three channels. The first channel is used for counting tritium, the second for carbon-14, and the third for phosphorus-32 or carbon-14. In counting ^{3}H and ^{14}C at the same time, channels 1 and 3 were used, where channel 3 had special settings. These settings were such that less than 0.05% of the tritium present was counted in channel 3 (E = 150, F = 1000, and % gain = 8.5). Thus, the procedure was to calculate the amount of ^{14}C from channel 3 and then subtract this amount of ^{14}C-activity from the net counts per minute in channel 1 to determine the amount of tritium present from the efficiency curve. To make this type of counting accurate, the tritium to carbon-14 ratio should be at least 10 to 1 since in channel 1 both isotopes are counted and the efficiency of ^{14}C is higher than that for ^{3}H.

5. Reactions of ^{14}C-amino acids

The reactions of the ^{14}C-amino acids were carried out in the same manner as the reactions of ^{14}C-nucleotides (see the experimental section of Chapter II). The labeled amino acids were obtained from Schwarz BioResearch, Inc.

The standard procedure was to make up 10 ml of the unlabeled amino acid at the desired concentration and pH, and add a certain amount of this solution to the ^{14}C-amino acid. The ^{14}C-amino acid solution was pipetted into a 25 ml pear-shaped flask and evaporated on a rotary evaporator. The usual amount of radioactivity was 50 μC, and this was then dissolved in 500 μl of the unlabeled amino acid solution.

The reactions were carried out with 100 μl amounts of the resulting radioactive solutions in 2 ml centrifuge tubes, and the condensing agent was added directly to this solution. After the usual reaction time of one day, the products were separated by thin-layer chromatography, paper chromatography, or paper electrophoresis, and the percent yields of the various components were calculated by counting the amount of ^{14}C activity in the various spots after exposure of the paper to X-ray film. This carbon-14 counting was usually done on a Geiger-Müeller double-hand counter.[127]

ACKNOWLEDGMENTS

The author is grateful to Dr. Melvin Calvin, who is one of the pioneers in chemical evolution studies, for his many valuable suggestions and comments in the progress of this work.

Many interesting and stimulating discussions were had with Dr. Richard Lemmon and Dr. Ercole Cavalieri. Their advice and friendship were crucial to the completion of this work, and are gratefully acknowledged here.

The assistance of Evie Litton is deeply appreciated. Her cheerful cooperation in preparing the figures and typing the manuscript has made the burden of thesis writing much lighter.

Financial support is acknowledged from the Lawrence Radiation Laboratory. This work was sponsored by the United States Atomic Energy Commission.

REFERENCES

1. A. I. Oparin, "Proiskhozhdenie zhizni," Izd. Moskovskii Rabochii, Moscow, 1924; "The Origin of Life," The Macmillan Company, New York, 1938.
2. J. B. S. Haldane, Rationalists Ann., 148 (1929).
3. S. L. Miller, Science 117, 528 (1953).
4. W. M. Garrison, D. C. Morrison, J. G. Hamilton, A. A. Benson and M. Calvin, Science 114, 416 (1951).
5. M. Calvin, "Chemical Evolution," Oxford University Press, New York and Oxford, 1969.
6. G. R. Tilton and R. H. Steiger, Science 150, 1805 (1965).
7. E. S. Barghoorn and J. W. Schopf, Science 152, 758 (1966).
8. J. W. Schopf and E. S. Barghoorn, Science 156, 508 (1967).
9. H. J. Hofman, Science 156, 500 (1967).
10. D. H. Kenyon and G. Steinman, "Biochemical Predestination," McGraw-Hill Book Company, New York and London, 1969.
11. W. W. Rubey, Geol. Soc. Am. Spec. Papers No. 62, p. 531, 1955.
12. H. Urey, "The Planets," Yale University Press, New Haven, Conn., 1952.
13. H. D. Holland, in "Petrologic Studies," (A. E. J. Engel, H. L. James and B. F. Leonard, eds.), Geological Society of America, p. 447, 1962.
14. I. S. Shklovskii and C. Sagan, "Intelligent Life in the Universe," Holden-Day, Inc., San Francisco, p. 231, 1966.
15. S. L. Miller and H. C. Urey, Science 130, 245 (1959).
16. J. R. Vallentyne, in "Current Aspects of Exobiology," (G. Mamikunian and M. H. Briggs, eds.), Pergamon Press, 1965.

17. H. R. Mahler and E. H. Cordes, "Biological Chemistry," Harper and Row, New York and London, 1966.

18. R. M. Lemmon, Chem. Rev., in press.

19. S. L. Miller, J. Am. Chem. Soc. 77, 2351 (1955).

20. T. E. Pavlovskaya and A. G. Pasynskii, in "The Origin of Life on Earth," (Proc. 1st Int'l. Symp., Moscow, 1957), Pergamon Press, New York, p. 151, 1959.

21. K. A. Grossenbacher and C. A. Knight, in "The Origins of Prebiological Systems," (S. Fox, ed.), Academic Press, New York, p. 173, 1965.

22. K. Heyns, W. Walter and E. Meyer, Naturwissenschaften 44, 385 (1957).

23. P. H. Abelson, Science 124, 935 (1957).

24. P. H. Abelson, Ann. N. Y. Acad. Sci. 69, 274 (1957).

25. W. E. Groth and H. von Weyssenhoff, Planet. Space Sci. 2, 79 (1960).

26. C. Palm and M. Calvin, J. Am. Chem. Soc. 84, 2115 (1962).

27. K. Harada and S. W. Fox, in "The Origins of Prebiological Systems," (S. Fox, ed.), Academic Press, New York, p. 187, 1965.

28. K. Bahadur, Nature 173, 1141 (1954).

29. C. Reid, in "The Origin of Life on Earth," (Proc. 1st Int'l. Symp., Moscow, 1957), Pergamon Press, New York, p. 619, 1959.

30. P. H. Abelson, Proc. Nat. Acad. Sci. U.S. 55, 1365 (1966).

31. K. Dose and C. Ponnamperuma, Rad. Rev. 31, 650 (1967).

32. T. Hasselstrom, M. C. Henry and B. Murr, Science 125, 350 (1957).

33. C. U. Lowe, M. W. Rees and R. Markham, Nature 199, 219 (1963).

34. J. Oró and A. P. Kimball, Arch. Biochem. Biophys. 96, 293 (1962).

35. J. Oró, A. Kimball, R. Fritz and F. Master, Arch. Biochem. Biophys. 85, 115 (1959).

36. C. Ponnamperuma, R. M. Lemmon, R. Mariner and M. Calvin, Proc. Nat. Acad. Sci. U.S. 49, 737 (1963).

37. J. Oró, Biochem. Biophys. Res. Comm. 2, 407 (1960).

38. C. Ponnamperuma, in "The Origins of Prebiological Systems," (S. Fox, ed.), Academic Press, New York, p. 221, 1965.

39. R. A. Sanchez, J. P. Ferris and L. E. Orgel, Science 154, 784 (1966).

40. J. Oró, in "Origins of Prebiological Systems," (S. Fox, ed.), Academic Press, New York, p. 137, 1965.

41. S. W. Fox and K. Harada, Science 133, 1923 (1961).

42. A. Butlerov, Ann. 120, 296 (1961).

43. N. W. Gabel and C. Ponnamperuma, Nature 216, 453 (1967).

44. J. Oró and A. C. Cox, Fed. Proc. 21, 80 (1962).

45. C. Ponnamperuma, C. Sagan and R. Mariner, Nature 199, 222 (1963).

46. C. Ponnamperuma and P. Kirk, Nature 203, 400 (1964).

47. C. Ponnamperuma and R. Mack, Science 148, 1221 (1965).

48. A. Beck, R. Lohrman, and L. E. Orgel, Science 157, 952 (1967).

49. T. V. Waehneldt and S. W. Fox, Biochim. Biophys. Acta 134, 1 (1967).

50. W. V. Allen and C. Ponnamperuma, Currents in Modern Biology 1, 24 (1967).

51. S. C. Lind and D. C. Bardwell, J. Am. Chem. Soc. 48, 2335 (1926).

52. B. C. McKusick, W. E. Mochel and F. W. Stacey, J. Am. Chem. Soc. 82, 723 (1960).

53. R. A. Sanchez and L. E. Orgel, 5th Western Regional Meeting of the Am. Chem. Soc., Anaheim, Calif., 1969, as reported in Chem. Eng. News, Oct. 6, 1969.

54. G. W. Hodgson and C. Ponnamperuma, Proc. Nat. Acad. Sci. U.S. 59, 22 (1968).

55. A. Szutka, Nature 212, 401 (1966).

56. J. D. Watson, "Molecular Biology of the Gene," W. A. Benjamin, Inc., New York, N. Y., 1965.

57. A. Thompson, A. Kimika, M. L. Wolfrom and M. Inatome, J. Am. Chem. Soc. 76, 1309 (1954).

58. P. T. Mora and J. W. Wood, J. Am. Chem. Soc. 80, 685 (1958).

59. G. Schramm, H. Grötsch and W. Pollmann, Angew. Chem. int. Edn. 1, 1 (1962).

60. M. Calvin, Proc. Roy. Soc. 288A, 441 (1965).

61. S. W. Fox and K. Harada, J. Am. Chem. Soc. 82, 3745 (1960).

62. K. Harada and S. W. Fox, in "The Origins of Prebiological Systems," Academic Press, New York, N. Y., p. 289, 1965.

63. S. W. Fox and T. V. Waehneldt, Biochim. Biophys. Acta 160, 246 (1968).

64. G. Krampitz and S. W. Fox, Proc. Nat. Acad. Sci. U.S. 62, 399 (1969).

65. C. N. Matthews and R. E. Moser, Proc. Nat. Acad. Sci. U.S. 56, 1087 (1966).

66. C. N. Matthews and R. E. Moser, Nature 215, 1230 (1967).

67. T. Wadsten and S. Anderson, Acta Chem. Scand. 13, 1069 (1959).

68. C. W. Lowe, M. W. Rees and R. Markham, Nature 199, 219 (1963).

69. S. Akabori, in "The Origin of Life on Earth," Pergamon Press, London, p. 189, 1959.

70. G. Steinman, R. M. Lemmon and M. Calvin, Proc. Nat. Acad. Sci. U.S. 52, 27 (1964).

71. C. Ponnamperuma and E. Peterson, Science 147, 1572 (1965).

72. M. Halmann, Arch. Biochem. Biophys. 128, 808 (1968).

73. G. Steinman, R. M. Lemmon and M. Calvin, Science 147, 1574 (1965).

74. G. Steinman, D. H. Kenyon and M. Calvin, Biochim. Biophys. Acta 124, 339 (1966).

75. G. Steinman, Arch. Biochem. Biophys. 119, 76 (1967).

76. A. Schimpl, R. M. Lemmon and M. Calvin, Science 147, 149 (1964).

77. F. Kurzer and K. Douraghi-Zadeh, Chem. Rev. 67, 107 (1967).

78. F. Rabinowitz, J. Flores, R. Krebsbach and G. Rogers, Nature 224, 795 (1969).

79. J. Rabinowitz, S. Chang and C. Ponnamperuma, Nature 218, 442 (1968).

80. S. L. Miller and M. Paris, Nature 204, 1248 (1964).

81. W. D. MacElroy, M. DeLuca and J. Travis, Science 157, 150 (1967).

82. F. Lipmann, in "The Origins of Prebiological Systems," Academic Press, New York, N. Y., p. 259, 1965.

83. F. M. Harold, Bact. Rev. 30, 772 (1966).

84. J. Oró and C. L. Guidry, Arch. Biochem. Biophys. 93, 166 (1961).

85. K. Dose and C. Ponnamperuma, Rad. Res. 31, 650 (1967).

86. A. W. Schwartz and S. W. Fox, Biochim. Biophys. Acta 134, 9 (1967).

87. J. Morávek, Tetrahedron Let. 18, 1707 (1967).

88. G. Schramm and H. Wissmann, Chem. Ber. 91, 1073 (1958).

89. R. J. Goldacre, "Surface Phenomena in Chemistry and Biology," Pergamon Press, London, p. 278, 1958.

90. D. Papahadjopoulos and N. Miller, Biochim. Biophys. Acta 135, 624 (1967).

91. A. D. Bangham, J. DeGier and G. D. Greville, Chem. Phys. Lipids 1, 225 (1967).

92. D. Papahadjopoulos and J. C. Watkins, Biochim. Biophys. Acta 135, 639 (1967).

93. M. Traube, Arch. Physiol. 87, 129 (1967).

94. S. W. Fox, in "The Origins of Prebiological Systems," Academic Press, New York, N. Y., p. 361, 1965.

95. R. S. Young, in "The Origins of Prebiological Systems," Academic Press, New York, N. Y., p. 347, 1965.

96. A. I. Oparin, in "The Origins of Prebiological Systems," Academic Press, New York, N. Y., p. 331, 1965.

97. L. Pauling and M. Delbrück, Science 92, 77 (1940).

98. G. Gamow, Nature 173, 318 (1954).

99. F. H. C. Crick, J. Griffith and L. Orgel, Proc. Nat. Acad. Sci. U.S. 43, 416 (1957).

100. F. H. C. Crick, in "Progress in Nucleic Acids," (J. Davidson and W. Cohn, eds.), Academic Press, New York, N. Y., Vol. 1, p. 163, 1963.

101. G. Zubay and P. Doty, Biochim. Biophys. Acta 29, 47 (1958).

102. I. B. Weinstein, Cold Spring Harb. Symp. Quant. Biol. 28, 579 (1963).

103. M. Nirenberg, P. Leder, M. Bernfield, R. Brimacombe, J. Trupin, F. Rottman and C. O'Neal, Proc. Nat. Acad. Sci. U.S. 53, 1161 (1965).

104. S. R. Pelc, Nature 207, 597 (1965).

105. C. R. Woese, Proc. Nat. Acad. Sci. U.S. 54, 71 (1965).

106. S. R. Pelc and M. G. E. Welton, Nature 209, 868 (1966).

107. M. G. E. Welton and S. R. Pelc, Nature 209, 870 (1966).

108. P. R. Srinivasan and E. Borek, Science 145, 548 (1964).

109. B. S. Baliga, P. R. Srinivasan and E. Borek, Nature 208, 555 (1965).

110. C. R. Woese, D. H. Dugre, W. C. Saxinger and S. A. Dugre, Proc. Nat. Acad. Sci. U.S. 55, 966 (1966).

111. J. C. Lacey, Jr. and K. M. Pruitt, Nature 223, 799 (1969).

112. C. R. Woese, Proc. Nat. Acad. Sci. U.S. 59, 110 (1968).

113. F. H. C. Crick, J. Mol. Biol. 38, 367 (1968).

114. L. E. Orgel, J. Mol. Biol. 38, 381 (1968).

115. C. R. Woese, "The Genetic Code," Harper and Row, New York, N. Y., 1967.

116. R. Naylor and P. T. Gilham, Biochemistry 8, 2722 (1966).

117. J. E. Sulston, R. Lohrmann, L. E. Orgel and H. T. Miles, Proc. Nat. Acad. Sci. U.S. 60, 409 (1968).

118. J. E. Sulston, R. Lohrmann, L. E. Orgel and H. T. Miles, Proc. Nat. Acad. Sci. U.S. 59, 726 (1968).

119. J. Sulston, R. Lohrmann, L. E. Orgel, H. Schneider-Bernloehr, B. J. Weimann and H. T. Miles, J. Mol. Biol. 40, 227 (1969).

120. S. L. Miller and M. Parris, Nature 204, 1248 (1964).

121. G. D. Steinman, Ph.D. Thesis, University of California, Berkeley, UCRL-16566, January, 1966.

122. H. G. Khorana, W. E. Razzell, P. T. Gilham, G. M. Tener and E. H. Pol, J. Am. Chem. Soc. 79, 1002 (1957).

123. P. T. Gilham and H. G. Khorana, J. Am. Chem. Soc. 80, 6212 (1958).

124. M. Calvin, "Chemical Evolution," Oxford University Press, New York and Oxford, pp. 162-183, 1969.

125. G. Steinman and M. N. Cole, Proc. Nat. Acad. Sci. U.S. 58, 735 (1967).

126. Ito, Eiji, and J. L. Strominger, J. Biol. Chem. 325, Prelim. Comm. (1960).

127. C. H. Wang and D. L. Willis, "Radiotracer Methodology in Biological Science," Prentice Hall, 1965.

128. M. Smith, J. G. Moffatt and H. G. Khorana, J. Am. Chem. Soc. 80, 6204 (1958).

129. G. M. Tener and H. G. Khorana, J. Am. Chem. Soc. 77, 5349 (1955).

130. H. G. Khorana, "Some Recent Developments in the Chemistry of Phosphate Esters of Biological Interest," John Wiley & Sons, Inc., 1961.

131. J. C. Sheehan and J. J. Hlavka, J. Org. Chem. 21, 439 (1956).

132. J. C. Sheehan, P. A. Cruickshank and G. L. Boshart, J. Org. Chem. 26, 2525 (1961).

133. J. C. Sheehan, M. Goodman and G. P. Hess, J. Am. Chem. Soc. 78, 1367 (1956).

134. G. Steinman, D. H. Kenyon and M. Calvin, Nature 206, 707 (1965).

135. R. Lohrmann and L. E. Orgel, Science 161, 64 (1968).

136. F. B. Howard, J. Frazier, M. Singer and H. T. Miles, J. Mol. Biol. 16, 415 (1966).

137. R. A. Sanchez, J. P. Ferris and L. E. Orgel, Science 154, 784 (1966).

138. J. F. Arens, Rec. trav. Chim. 74, 769 (1955).

139. H. J. Pannemann, A. F. Marx and J. F. Arens, Rec. trav. Chim. 78, 487 (1959).

140. J. C. Sheehan and J. J. Hlavka, J. Org. Chem. 23, 635 (1958).

141. H. Wasserman and P. S. Wharton, J. Am. Chem. Soc. 82, 1411 (1960).

142. J. P. Ferris, Science 141, 53 (1968).

143. C. Moureu and J. C. Bongrand, Ann. Chim. 14, 5 (1920).

144. M. Smith, J. G. Moffatt and H. G. Khorana, J. Am. Chem. Soc. 80, 6204 (1958).

145. G. M. Tener, H. G. Khorana, R. Markham and E. H. Pol, J. Am. Chem. Soc. 80, 6223 (1958).

146. A. R. Todd, Proc. Soc., 187 (1961).

147. R. V. Tomlinson and G. M. Tener, J. Org. Chem. 29, 493 (1964).

148. G. Weimann and H. G. Khorana, J. Am. Chem. Soc. 84, 4329 (1962).

149. M. Leng and G. Felsenfeld, Proc. Nat. Acad. Sci. U.S. 56, 1325 (1966).

150. D. E. Olins, A. L. Olins and P. H. von Hippel, J. Mol. Biol. 24, 157 (1967).

151. W. L. Burdick, J. Am. Chem. Soc. 47, 1485 (1925).

152. W. Madelung and E. Kern, Ann. 427, 1 (1922).

153. E. Baumann, Ber. 8, 708 (1875).

154. K. Grube and G. Kruger, Z. Physikal. Chem. 86, 65 (1913).

155. L. A. Pink and H. C. Hetherington, Ind. Eng. Chem. 27, 834 (1935).

156. J. Buchanan and B. Barsky, J. Am. Chem. Soc. 32, 198 (1930).

157. D. Drechsel, J. pr. Chem. 13, 331 (1876).

158. H. Hale and H. Vibrans, J. Am. Chem. Soc. 40, 1046 (1918).

159. W. D. Kumler, J. Org. Chem. 20, 700 (1955).

160. American Cyanamid Company, "Cyanamid's Nitrogen Chemicals Digest," Vol. III, A.C.C., New York, 1949.

161. S. Stolle and K. Krauch, J. Prakt. Chem. 88, 306 (1913).

162. N. Pohl, J. Prakt, Chem. 77, 533 (1908).

163. K. Taub and T. Kropp, Chem. Abst. 28, 2016 (1934).

164. E. W. Hughes, J. Am. Chem. Soc. 62, 1258 (1940).

165. G. D. Diana, E. S. Zalay and R. A. Culter, J. Org. Chem. 30, 298 (1965).

166. J. C. Sheehan and G. P. Hess, J. Am. Chem. Soc. 77, 1067 (1955).

167. H. G. Khorana, Chem. Ind. (London), 1087 (1955).

168. G. I. Lavrenova and N. Y. Poddubnaya, Zh. Obshch. Khim. 31, 2474 (1961).

169. J. C. Sheehan, M. Goodman and G. P. Hess, Federation Proc. 14, 226 (1955).

170. E. Schmidt, F. Hitzler and E. Lahde, Chem. Ber. 71, 1933 (1938).

171. R. Schwyzer and P. Sieber, Helv. Chim. Acta 41, 2190 (1958).

172. F. Zetzsche and A. Fredrick. Chem. Ber. 72, 1735 (1939).

173. F. Zetzsche and G. Roettger, Chem. Ber. 72, 2095 (1939).

174. F. Zetzsche and G. Baum, Chem. Ber. 75, 100 (1942).

175. F. Zetzsche, H. E. Meyer, H. Overbeck and H. Lindler, Chem. Ber. 71, 1516 (1938).

176. F. Zetzsche, H. E. Meyer, H. Overbeck and W. Nerger, Chem. Ber. 71, 1512 (1938).

177. D. Aldridge, J. R. Hanson and T. P. C. Mulholland, J. Chem. Soc. 3539 (1965).

178. F. Zetzsche and A. Fredrick, Chem. Ber. 73, 1114 (1940).

179. W. S. Johnson, V. J. Bauer, J. L. Margrave, M. A. Frisch, L. H. Dreger and W. N. Hubbard, J. Am. Chem. Soc. 83, 606 (1961).

180. I. Muramatsu and A. Hagitani, Nippon Kagaku Zasshi, 80, 1497 (1959); Chem. Abstr. 55, 6394 (1961).

181. I. G. Farbenindustrie Akt.-Ges., German Patent 550, 571 (1927).

182. C. P. Joshua, J. Indian Chem. Soc. 37, 621 (1960).

183. J. MacHoldt-Erdness, Chem. Ber. 91, 1922 (1958).

184. W. Weith, Chem. Ber. 17, 1303 (1874).

185. A. Huhn, Chem. Ber. 19, 2404 (1886).

186. H. G. Khorana, Chem. Rev. 53, 145 (1953).

187. F. Kurzer and P. M. Sanderson, J. Chem. Soc., 3240 (1960).

188. W. Marckwald, Ann. 286, 343 (1895).

189. E. Schmidt and W. Carl, Ann. 639, 24 (1961).

190. G. Doleschall and K. Lempert, Tetrahedron Letters, 1195 (1963).

191. D. G. Knorre and T. N. Shubina, Kinetika i Kataliz 1, 519 (1960); Chem. Abst. 55, 18610 (1961).

192. Biochemists Handbook, C. Long, ed. (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1961).

193. J. H. Bradbury and D. C. Shaw, Australian J. Chem. 12, 300 (1959).

194. E. Schröder and K. Lübke, "The Peptide," Vol. I (Academic Press, New York and London, 1965).

195. B. C. Barras and D. T. Elmore, J. Chem. Soc., 4830 (1957).

196. E. F. Curraghand and D. T. Elmore, J. Chem. Soc., 2948 (1962).

197. H. Tabor, Biochemistry 1, 496 (1962).

198. H. Busch, "Histones and Other Nuclear Proteins," (Academic Press, New York, 1965).

199. W. Möller and J. Widdowson, J. Mol. Biol. 24, 367 (1967).

200. K. Randerath, "Thin-Layer Chromatography," (Academic Press, New York and London, 1966).

201. G. Pataki, "Techniques of Thin-Layer Chromatography," (Ann Arbor Science Publishers, Inc., 1968).

202. F. Sanger, Biochem. J. 39, 507 (1945).

203. A. L. Levy and D. Chung, J. Am. Chem. Soc. 77, 2899 (1955).

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