Name: Oliver P. Stafford Date of Degree: August 7, 1965

Institution: Oklahoma State University

Location: Stillwater, Oklahoma

Title of Study: GENETIC CODING

Pages in Study: 45 Candidate for the Degree of Master of Science

Major Field: Natural Science

Scope and Method of Study: A survey of the literature was made to determine history of the search for the gene during the formative years as a science to recent findings in molecular genetics. The work of several investigators who are exploring the code between transfer RNA and messenger RNA to explain the route of synthesis of amino acids into protein is discussed in some detail. Recent findings on the behavior of genetics on the molecular level are then described. Possible explanations of classic genetic mechanisms are offered with indications that non-chromosomal genes, hormones and enzymatic feedback to the DNA from the environment must now be considered in any selective gene action.

Findings and Conclusions: The genetic code between messenger RNA and transfer RNA has essentially been solved with 19 of the 64 possible code words remaining to be confirmed.

ADVISER'S APPROVAL L. Herbert Bruncan

GENETIC CODING

By

OLIVER P. STAFFORD

Bachelor of Science in Education
Oklahoma State University
Stillwater, Oklahoma
1959

Submitted to the Faculty of the Graduate School of
the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE
August, 1965

GENETIC CODING

Report Approved:

L. Herbert Ramsey
Report Adviser

Dean of the Graduate School

INTRODUCTION

For thousands of years the curious have observed and speculated regarding genetic principles. A form of genetics was practised through selective breeding to domesticate plants and animals; though before the latter part of the seventeenth century and the invention of the microscope, little was written that could contribute to the field. In fact, not until 1900 and the rediscovery of Gregor Mendel's work of 1866, could genetics be carried on in a quantitative or qualitative fashion. Although Mendel and other early workers showed that something which had the nature of a "gene" existed, their work was superficial since a gene particle could not be found. It is the purpose of this report to investigate the nature and, if possible, to determine the chemical combination of the gene in light of recent discoveries in the field of molecular biology.

TABLE OF CONTENTS

Chapter		Page
I.	A BRIEF HISTORY OF CLASSIC GENETICS.	1
II.	MOLECULAR EXPLANATION OF THE "GENE".	15
III.	POSSIBLE CHEMICAL EXPLANATIONS FOR CLASSIC GENETIC MECHANISMS	29
BIBLIOGRAPHY		40

LIST OF TABLES

Table — Page

I. Latest MRNA-TRNA Code Words 39

LIST OF FIGURES

Figure — Page

1. Purines, Pyrimidines and Ribose Sugars for DNA - RNA Synthesis 35

2. Amino Acids Used in Protein Synthesis 36

3. Stereochemistry of DNA and Single Stranded RNA 37

4. Schematic Transport of Amino Acids Into Protein Synthesis. . 38

CHAPTER I

A BRIEF HISTORY OF CLASSIC GENETICS

Mendel showed mathematically that a combination of unit characters could be traced through several generations of the pea plant; and that the hybridization of these characters was predictable. Other than referring to egg and sperm cells he made no statement as to the localization of the activity of these combinations within the cell. The perfection of the compound light microscope and staining techniques with analine dyes gave the cytologist and geneticist the tools to observe in meiosis the coincidence of genetic control and chromosome division. W. S. Sutton, concurrently with T. Boveri (1903), published the hypothesis that chromosomes bore the hereditary traits. A. H. Sturtevant (1913) published a paper in which he laid the basis for chromosome mapping by the relative percent of crossover between the genes on the chromosomes. During meiosis, chromosomes are often twisted or bent and come to lie, as someone had said, like a tangled pile of thread. Internal tensions cause breakage and rejoining within the tetrad to bring about the phenomenon known as crossover. The theory of crossover mapping is based upon the fact that breakage and rejoining (or copy choice is more likely) has a greater probability of occurance between two widely separated genes than between two closely spaced ones. The results of numerous hybridization experiments involving many gene pairs and multiple crossover relations strongly indicate that the genes are located in a linear fashion, probably on the chromosome since this is one of the few truly linear structures in the cell.

Creighton and McClintock (1931) showed conclusively that cytological crossing over is accompanied by genetic crossing over in Zea mays. By the use of a chromosome with a characteristic knob they were able to show that pairing chromosomes, heteromorphic in two regions, exchange parts, which are cytologically observable, at the same time they exchange genes assigned to these regions.

T. S. Painter (1933) published one of the first maps of Drosophila melanogaster showing a correlation between crossover maps and the actual banding of the giant chromosomes of the larva of this insect. Many were ready to call these bands the genes but Painter was careful to state that the crossover map represented relative positions rather than absolute positions.

The brief foregoing history of genetics in its formative years as a science provided effective tools to describe all the forthcoming data, but there was still a lack of understanding as to what the gene actually was and how it controlled life processes. This does not in the least detract from the contributions of these early workers since many domesticated plants and animals were greatly improved by the existing methods of genetics. The ever growing question of what the gene was and how it functioned was still to be answered. It was not clear what to look for to answer for the gene.

Sewell Wright (1917) had suggested that the gene might control enzymes which would bring about the terminal effect of a unit character. Beadle and Tatum (1941) investigated the genetics of X-irradiated Neurospora crassa and found that several biochemical reactions involving growth factors and vitamins were traceable to heritable mutations. This gave strong proof to the theory that the gene worked through either one or a

series of enzymatic reactions, and that the gene - whatever it was - was not always the simple expression of a terminal visible manifestation.

The interest in gene behavior then began to shift toward the biochemical level. As early as 1928, Griffith showed that certain strains of Pneumococcus could be transformed from rough non-encapsulated strains to virulent encapsulated strains. Mice injected with the rough non-encapsulated strains revealed no ill effects. However, if the non-virulent injection was administered with a large amount of heat killed virulent cells, the animal succumed to infection typical of the virulent cells. Also, fully virulent cells were recoverable from the animal's blood. Clearly, something within the killed bacteria had passed into or was in some way controlling the genetics of the living strain. The identity of this substance remained unexplained until 1944 when Avery, Macleod, and McCarty set about to define the chemical nature of the agent causing the transformation. They found that the fractional extract of desoxyribonucleic acid (DNA) from the virulent cells in concentrations as dilute as 1 : 600,000,000 were capable of bringing about transformation. This discovery is of outstanding importance to geneticists because of the precise way in which it shows DNA to be of primary importance in changing the hereditary make up of an organism. Once this change has taken place the defined chemical compound continues to reproduce itself exactly. This work - even though it is considered a classic in genetics - still did not define the gene, but it greatly limited the field.

Not long after the work on transformation, further evidence of the role of DNA came from studies by Lederburg and Tatum (1947) who published information indicating the presence of a form of genetic exchange in a strain of Echerichia coli. This has been popularly called bacterial sex.

The importance of this unsuspected phenomenon should be recognized as an experimental method of recombination of genes which can thus be acted upon by natural selection. That a system somewhat similar to this should be suspected for bacteria should not be surprising for it is partly through the sexual recombination in higher organisms that evolution is theorized to have taken place. Genetic exchange for one celled plants and animals is quite conceivable in order for evolution to take place to the level of the multiple celled creature. Simple cell division had been accepted for many years as the only way in which bacteria could divide. Lederburg's contribution was to design an experiment to test this assumption.

In this study mutants formed by X-rays or ultra-violet light were found to have lost their ability to synthesize certain amino acids. Two strains, each having three mutations for different amino acids, were grown in mixed complete culture. When fully grown the cells were washed and plated on selective media. Reversions for one, two or all three growth factors were found. No mutagenic agents could be found in the media. The conclusion was that genetic material had been exchanged even though a hypothetical zygote could not be found cytologically. As will be shown in the characterization of the DNA molecule in this report the length of the molecule is variable as the chromosome length is variable. Estimates for the upper limits of size vary from 500,000 to 15,000,000 molecular weights. It seems inconsistent that molecules of this size and complexity can enter the cell when it is necessary for digestive processes to break down the proteins into their constituent amino acids before passing through the cell membrane. These amino acids are then reconstituted to the "formula" compatible to and usable by the cell. That molecules of the

size of DNA could get through the cell membrane seems almost impossible but this is the conclusion to which one is lead. This discovery is important not only from the standpoint of finding a rudimentary form of sex for evolution in a wholly unsuspected organism but also as a method of genetic transfer, in whole or in part, of the chromosome of a bacterium.

How DNA is permitted to enter the cell may be explained by the structure of the membrane. Holter (1961) states that the membrane is a rather flimsy but highly sensitive and selective structure of some 75 angstroms in thickness. Many lipoidal substances penetrate the membrane quite easily. This would indicate that there are lipids (fatty) materials in the membrane, probably in association with phosphates. Water readily passes in both directions as well as various ions in solution. Proteins or their momomeric units, amino acids, also pass freely through the membrane. To bring all of these observed activities together the membrane was proposed to have a double layer of lipids with their long carbon chain tails pointing toward each other and the phosphate groups directed toward the outside and inside. Overlaying this structure is a protein layer which covers both the outside and inside surface of the membrane. Polar ends of the molecules are directed outward to both surfaces. This would permit charged substances to enter the cell while the lipids would permit oily or fatty molecules to enter. The membrane was predicted to have disruptions of the phospho-lipids wherein the outside protein layer met and joined the inside layer of protein to form a pore. The pore was to have a lining of the protein material. The pores would probably be on the order of 10 A. to 20 A. in width. This would account for the entrance of a single strand of RNA or DNA which is on the order of 10 A. in diameter. Current electron micrographs have borne out these predictions as Holter

shows. The phage tails of certain bacteriophage are specific for certain sites on the cell wall for it can be saturated until no more phage will attach. These attachment sites of the cell wall may be in association in some way with the pores. In order to add specificity to substances entering the membrane, many cells have evolved enzymes in the membrane. The enzymes or the pores may or may not be the site for tail attachment but the pore seems very likely to be the entrance point for DNA and RNA for cells without cell wall structures.

David Robertson (1962) supports the structure predictions of Holtzer (1961) with beautiful electron micrographs. He alters this concept slightly, however, because differential staining shows that the outer surface is mainly phospho-lipid and sugar while the inner surface is mainly protein. How the long DNA molecule enters the cell without being broken down into constituent nucleic acids is still open to question.

Zinder and Lederburg (1952) found that other bacteria of the Salmonella group were susceptible to genetic exchange by quite another method, that of transduction. It had been known for quite some time that bacteria taken from the "wild", but still of a pure strain, often developed cleared areas, or "plaques", of killed bacteria when plated on agar. Furthermore these "plaques", when carefully scraped from the agar, suspended in a liquid solution and centrifuged or filtered to remove all traces of bacterial cells, would still give a supernatant fluid which would reinoculate the bacteria. The name bacteriophage, or phage for short, had been given to these organisms even before they had been observed in the microscope. Actually, they were beyond the resolution of the light microscope and escaped detection until the electron microscope came into wide use. They were then found to be a type of virus which were rather narrowly

specific for bacterial types. Mutants of these phage organisms, which must be cultured on bacteria, were used in the experiments to show that genetic exchanges did occur and that latent bacteriophages played some role in the interchange. Appropriate experimental controls were set up to show that cultures were not contaminated and also that transductions were too frequent to be accounted for on the basis of genetic exchange alone, or through mutational changes.

To understand transduction, it will be necessary to follow the action of a lytic phage through its various "life" stages. The mature phage approaches the membrane of the host cell and attaches to the specific receptor site on the cell wall of the bacterium. The DNA, or RNA in some cases, is then injected into the host cell where it intermingles with the cell parts. The viral material then takes over the mechanism of cell biochemistry and replicates itself upwards of a hundred to two hundred times. These new viral DNAs or RNAs then build a protein shell or coat about themselves. The cell wall then lysates, spilling out the new virus to continue their cycle on other cells. The whole cycle takes only about twenty minutes or roughly the average time for bacterial division to take place during the fast log growth phase. The bacterium hardly has a chance, for while it divides into two parts the virus might replicate a hundred fold (dependent upon the amount of cell material available to it). The chemical assembly of one hundred or so DNAs in 20 minutes will roughly show the rate of these biochemical reactions. The cycle of phage replication here described is typical of E. coli. (Edgar, 1965) In the case of Salmonella lysis does not occur at every cell division.

Zinder and Lederburg actually had discovered a method whereby slightly lytic phage entered the cells of Salmonella sp. and while replicating

themselves they also incorporated parts of the bacterial genetic material into their own genetic material (DNA). That this was not simply a transfer of genetic material was shown by its resistance to RNAse and DNAse, the respective enzymes for the breakdown of RNA and DNA. Apparently the protein coat of the virus protected the transduction agent. The transductive material, which they call filterable agent (FA), was also resistant to such disinfectants as chloroform, toluene, and alcohol as well as to enzymes such as pancreatin and trypsin. Although the total activity of the filterate encompassed the genotype of its parental culture, each transduction transmits only a single trait per bacterium. This contrasts with genetic exchange in E. coli, strain K-12, where there is unrestricted recombination of several markers that differentiate two parental lines. The recipient cell of transduction not only has its own DNA, but that of the virus as well as the small section of the donor cell; all living in a state of harmony until the next lysis. The genetic material (DNA) is apparently not as species specific as once thought.

The behavior of the gene to this time had been greatly limited for it had been shown that DNA was the genetic material and that it could pass from one cell to another in three different ways: (1) by transformation or the passage of the filtrate of a killed organism into a viable organism (2) by transduction, or the incorporation of genetic material of one cell into the genetic material of the lytic phage which would infect and pass on the characteristics of both in a healthy host cell, and (3) by sexual recombination in which a zygote is apparently formed to permit recombination and linkage types. It is essential to keep these concepts in mind for the discussion of the fine structure of the gene which is to follow. However, chronologically, Watson and Crick announced their proposal for

the structure of the DNA molecule in 1953 and since this information is also essential to the fine structure of the gene it will be considered next.

Watson and Crick (1953) proposed the structure for the nucleic acids to be that of two intertwined helices. Linus Pauling had made great contributions in the field of chemical bonding, bond angles and bond strength as well as the gross polymerized structures of proteins from their monomeric units. Pauling's work as well as the X-ray diffraction data published by M. H. F. Wilkins prompted their proposal of this structure. As is noted in their publication, Chargaff and Wyatt had shown that the ratio of amounts of adenine to thymadine and the ratio of guanine to cytosine, were always near unity. This suggested a pairing of each of these two groups. The structure of each of these nucleic acids as well as that of uridine, D-ribose, 2 de-oxy-ribose, and their method of phosphate linkage will be found on page 35 of this report. The novel feature of their proposed structure was the paired strands rather than the three stranded structure which Pauling had suggested. Also, the hydrogen bonding of the bases, which were to be directed toward the central axis and perpendicular to the axis of the helices, is a new inovation to the structure. This permits the double hydrogen bond between adenine and thymadine and a triple hydrogen bond between cytosine and guanine. This structure not only provides a code of two and three for "recognition" between the base pairs, but also provides the basis for replication of the strands since hydrogen bonds are rather weak when compared to the usual chemical bond. The two strands fit the observed copy and separation of the chromosomes during mitosis and also in meiosis. The "back bone" of the single stranded structure is the de-oxyribose sugar and phosphate chain. A purine or

pyrimidine is attached to the ribose sugar as is illustrated in the diagram. Although not discussed in this article RNA will be included in these diagrams because of its close relationship to DNA. The main difference between DNA and RNA is the substitution of uracil for thymadine in the apparently single strand of RNA. There are at least two types of RNA or three as recent evidence indicates (Scientific American, March 1963). There are ten base pairs per turn of the helices. Each turn is 34 angstroms in length which would give the axial distance of 3.4 A. between each base pair. The phosphate-sugar backbone is from the fifth carbon of one sugar through the phosphate to the third carbon of the next sugar. Although the purines, pyrimidines, phosphates, and sugars show an OH or an H at the attachment points, a molecule of water is lost in the reactions that join these units together.

Seymour Benzer (1955) designed experiments using the T4 bacteriophage for E. coli to test for the size of the gene. The technique of crossover genetics was an extension of the experiments of A. H. Sturtevant (1913). For the most practical purposes the fruit fly had lent itself well to the exhibition of Mendelian genetics as well as the mapping techniques of its four chromosomes. Its incubation period was rather easily handled and fairly large generations made for wide genetic recombinations. The giant salivary chromosomes gave visible evidence of the material under study. The small size of the individual also permits large populations in a small space. There were many drawbacks to Drosophila genetics, however. Benzer chose the T4 bacteriophage for E. coli mainly for the following reasons: (1) crossover information would be exceedingly rare in the range of a single gene. It would be necessary to have an organism which could provide vast populations in a very short time, (2) the small

size of the phage would present the least problem in storage and handling of large populations, (3) the single phage DNA would present pure "working" genetic material - this is not the case with diploid creatures for only half of their genetic material is believed to be working at one time, (4) haploid genetic material would circumvent the problem of phenotypic and genotypic ratios. The offspring of the phage is essentially a direct copy of the "parental" material except for the occasional mutant. There would be no problem of recessive traits. The design of his experiment was fairly simple but the work was tedious. Two strains of T4 phage were selected; a standard which could grow on both bacterial strains B and K as well as a mutant strain which could grow only on B cultures. The mutant phage could infect the K bacterial strain but normal lysis did not occur. Thus the two phages could be differentiated. Phage strains with mutations in the rII region were selected which gave characteristicly different plaques. This would facilitate a quick read out of the phages on the bacterial "fodder" as he refers to them.

To map the gene, independently arising mutants on plated strain B were removed. They were then crossed over in a liquid culture of B cells. The volume of B cells was controlled so that at least three phages were present for each bacterial cell. This gave a high probability of infection by both phage strains. Some mutants may crossover to revert to the standard type while most phages will simply replicate themselves. Since the number of crossovers is the all important factor, the recombinant phages were then plated on K cells. The recombinants or crossovers to the standard will grow on K cells whereas the mutant phages will not grow. The mapping technique requires the use of phages with deletions of known length. Deletions of successively shorter length are used until

crossovers occur. This is taken as evidence of the position of the mutation on the chromosome. By back crossing interchromosomal distances can be established between the known mutation sites and any new ones which might occur. This technique is so sensitive that mutants differing in only one base pair on the DNA chain can be resolved. Recombinants in the order of one in ten million can easily be detected by this method.

Benzer is apparently the first to question the concept of the gene and design experiments to test these questions. Some workers had previously defined the gene as the area of a chromosome necessary to produce one enzyme. That this definition is unsatisfactory can be seen in the corellary statement that an enzyme is that which is made by one gene. He states that if this is true it would be necessary to show that one enzyme caused each of the mutations in such experimental animals as the fruit fly Drosophila. He suggests that a genetic unit should be defined by a genetic experiment. (1962a)

Some mutant forms are known to be only substitutions of one amino acid in their resulting protein for another amino acid. V. M. Ingram (1958) discusses this for the genetic disease, sickle cell anemia. Three types of hemoglobin are known in which the substitution of one of three amino acids at a specific site can alter the behavior of hemoglobin drasticly enough to cause the severe sickle cell disease. There are some 300 amino acids in the hemoglobin molecule but the substitution at this single site can be fatal to its possessor. Work on the ACTH molecule which began some twenty years ago has recently been published by Choh Hao Li (1963). This work shows that beef, pig, and sheep ATCH are essentially the same except for minor substitutions in one small section of the 200 amino acid A 39 unit polypeptide extract of each of these macromolecules is identical

except for the 9 units between 24 and 34. Zuckerkandle (1965) has recently published similar findings on hemoglobin; even to the point that substitutions may prove to be an important tool in determining the evolution of several animals in question. These evidences as well as others indicate that the classic definition of the gene as the information contained for one enzyme would have to be changed.

Benzer found that the region rII which is necessary for the colon bacteriophage to multiply on K type bacteria consists of two cistrons and that each of these are composed of many mutons. Each of these were divided into recons which were divided into the smallest unit, the condon. Benzer is not too careful to define his terms which makes reading of his article confusing at times. It seems that the following definitions are consistent with his article. He calls the largest segment of the chromosome related to a single gene a cistron. The muton is the smallest segment capable of mutation. The recon is the smallest unit of recombination. The codon is the smallest unit which will code for one amino acid in an enzyme or other protein. In light of present information the muton and recon can be as small as the three nucleic acid codon. The use of these two terms have been all but dropped from present literature.

The frequency of spontaneous mutation might be suspected to be the same for all points along the chromosome segment. Benzer found however that some sites were six hundred times as sensitive to mutation as other sites. Other sites were so refectory that they did not respond to his tests. What is more surprising, perhaps, is the comparison of the frequency of induced mutations to those occurring spontaneously. Artificial mutagens, such as nitrous acid, 5-bromouracil, ultraviolet light and

others, showed broad specificity for various sites. Some sites show no effect at all while others show a greater tendancy to mutate spontaneously than under the effect of a mutagen. He suggests that adjoining genes may have some effect on mutation. Some three-hundred and fifty sites are known for their mutation while some one-hundred or more may be as yet undiscovered. In his earlier publication (Benzer, 1955) he states that each of the cistrons contain something on the order of four-thousand nucleotide pairs. From this information he was able to state that preliminary studies of this type indicate that the units of recombination are not larger than the order of one dozen nucleotide pairs and mutations may involve various lengths of chromosome. Although he does not take the experiments to their ultimate, that of determining the length of the codon, he suggests that his methods could be used to do so.

At this time the search for the "gene" shifts to the molecular level to find the codon. So far the gene had been found to be on the chromosome, that the chromosome was DNA, and that the structure of DNA which had been proposed was essentially correct. The system of genetic control pervaded all living things for DNA, or its counterpart RNA, had been found in all "living" things examined from virus to man and algae to flowering plants. Indeed, DNA might be used as one of the tests for living material. However, the identity of what the gene is and how it controls the biochemical behavior was still unknown. The chromosome could be seen in the electron microscope but the resolution was such that the fine structure could not be discerned. Experiments now were to go to the molecular heart of the information the DNA contained and how this was to be interpreted into what can be seen; the codon.

CHAPTER II

MOLECULAR EXPLANATION OF THE "GENE"

F. H. C. Crick (1962) continued work on his own suggestion made at the time the structure for DNA was proposed, that the order and number of nucleotide pairs may specify how the amino acids are incorporated into proteins. He gave a clear summary of work which had occurred to that time. It had become apparent that approximately three base pairs would be necessary for one codon. If only one of these base pairs was necessary for a codon, only four amino acids could be directed into a protein. This is insufficient for the twenty known to be necessary to life processes. Two base pairs ($4 \times 4 = 16$) might suffice if some were to do double duty, but this is doubtful. Three base pairs ($4 \times 4 \times 4 = 64$) provides ample combinations and thus seems more likely. However, a three base code raises many questions, such as, what are the functions of the other forty-four code words? Are they nonsense? Do they control some other genetic function, such as "telling" the gene to "turn on or turn off"? Does an amino acid have more than one code word, leading to degeneracy? Are these code words even utilized by a living organism? Immediately it can be seen that almost half the words such as AUG might be read as its reciprocal GUA. AAA, or polyA, certainly can be read in either direction as well as the other three poly bases. Is RNA specific for the reading of the codon in one direction? Although some combinations might not code for a specific amino acid they may control some other biochemical function. There is the possibility of a degenerate code

(Weisblum, 1962). This may explain the spontaneously mutable gene sites. The substitution of one amino acid for another on the resulting protein would likely change the character of the "gene", as long as the site on the DNA does not affect the active site on the resulting enzyme (Stein, 1961; Champe, 1962; Neurath, 1964; Nossal, 1964). Some sites certainly seem to be more mutable than environmental radiation or mutagens can account for. A degenerate code has not been conclusively determined in vivo at this time, but it seems likely. The structure of the chemicals, their spatial arrangements, bond angles and strength, activating energy levels, and numerous other factors must be kept in mind if the breaking of the code is not to degenerate into a game of anagrams. In short a chemical exchange sequence must be elucidated for the genetic code.

Crick states in his article that Gamow had suggested that the codons could overlap along the DNA sequence. Thus a sequence of ATCGAAG could read out in any of these triplets: ATC, TCG, CGA, GAA, AAG. This would imply that the code might "plug in" at any triplet point along the molecule. The overlap idea has fallen into disfavor because there is no indication of a restrictive sequence. The change of one code letter would change three code words, while evidence shows that the typical mutant has only one amino acid changed in the resulting protein.

Crick saw in Benzer's work a method to determine how the information in the gene-or cistron as Benzer calls it-could be read. Benzer (1962b) had noted that some deletion mutants, when located near another mutation on an adjacent cistron, would often show activity in the adjacent cistron. He had also noted that recombination would deactivate one of the adjacent mutants. To Crick this suggested that the deletion of one nucleic acid in mutant DNA would inactivate it at that site. The addition of a

nucleic acid, although rare, would also cause deactivation. One other eventuality, that of changing a nucleic acid while in the site, would also cause deactivation or a mutant. The most obvious solution to the whole problem then would be to test for "read out". To do this he used an addition mutant (+) with a very small map distance from a deletion mutant (-) in a crossover experiment. The resultant recombination reactivated the gene. When + and + were recombined, the gene did not work; neither did - and -. However, +++ or --- worked. Apparently these recombinants had somehow "shifted back into phase" the remaining segment of the chromosome between the two or three mutants. Loss of a small segment of chromosome did not seem critical for the gene to function, if it was not too long or did not have a vital activity. This is consistent with the findings of Choa How Li (1963) in his report of ACTH work. The shift back into phase would necessarily have to begin at some point on the chromosome. Apparently a small amount of the enzyme chain, which is the end result of DNA, could be dispensed with so long as the active site remained intact.

To test the effect of adjacent cistrons Crick used a mutant which had a deletion at the juncture of the A and B cistron. This was crossed with an acridine mutant (Lerman, 1963) which has an amino acid addition at the far beginning of the A cistron. The recombinant form showed the B cistron to be deactivated. When genes are joined without a stop point between them a change far away in one can in some way deactivate the other. This is the strongest evidence to this time for a space between the genes which behaves as a stopping and starting point as well as evidence for a sequential readout from one end. Since there was no stop between A and B, the gene would continue to read to the stop at the end of B. This

experiment by Crick may explain the classic concept of position effect, however, he does not discuss this.

To test for the size of the codon Crick used acridine mutants to crossover three additions into a resultant phage. Lerman suggests that acridine complexes vertically to the DNA helix axis (within the side space). Since the acridine molecule is one base pair long (ca. 3.4Å) this serves to shift pairing one step out of register or phase. This mechanism has the property of generating insertion or deletion mutants. In essence Crick successively built three addition mutants into one phage. These were each tested against the mutant with the small deletion between the A and B cistron. The first addition deactivated the cistron as expected. The second addition still deactivated the B cistron, but the third again showed activity of the B cistron. This does not conclusively prove three nucleotides per codon for six, nine, or twelve might also show the same evidence. Current investigations have confirmed that the original hypothesis of three nucleotides per codon was the correct assumption. In a recent note in the SCIENTIFIC AMERICAN (March, 1964) under the title "Hypothesis Confirmed" evidence is given for three being the correct number. By studying the rate at which polyribosomes settle out in the ultra-centrifuge, the molecular weight of the ribosome which carries the code for hemaglobin was computed to have about 450 nucleotides. Since the structure of hemaglobin is known to have some 150 amino acid subunits the ratio of nucleotides to amino acids is three to one. Alexander Rich (1963) also discusses this in his article on polyribosomes.

Crick concludes his article on coding by suggesting that this method might be of some value as a check against the synthetic codons which Ochoa and his group are using, as well as those that Nirenberg and Matthei are

making. He announces that other workers are attacking the problem of phage lysozyme. By "plugging in" an addition and/or a deletion the segment in between could be studied in the resulting enzyme. This would confirm codons in a living organism. The enzyme could then be systematically analyzed to determine its sequence. The sequence between the synthetic codons could then be inferred and the code broken.

Nirenburg (1963) gives perhaps the most comprehensive restatement of all information gained on Nucleic coding to that date of publication. He is continuing the attempt to find the code word or words through the use of synthetic RNA. Severo Ochoa (Speyer, 1962) of the New York University School of Medicine and his group of workers found that the enzyme polynucleotide phosphorulase would link nucleotide bases together in random fashion when in the presence of an energy source such as APT (adenine-tri-phosphate). Nirenberg uses this enzyme to polymerize separately all the possible three way combinations of the four bases. The resulting chains are then used in cell free systems (Bretscher and Manago, 1962) to see which of the radioactively tagged amino acids the transfer RNAs of a cell will direct into protein synthesis. For example, to test the coding of uracil, polynucleotide phosphorulase, ATP and uracil are placed in a container to polymerize. These long chains of polyU serve as the messenger RNA upon which the transfer RNA of the cell free system can deposit. The polyU is then added to twenty separate cell free systems which have been supplied with a full compliment of amino acids. One of the amino acids in each of the twenty containers has been made radioactive by the incorporation of carbon 14 into its structure. After incubation each of the twenty systems is percipitated with tri-chlor acetic acid. The percipitate is filtered onto paper disks to measure radioactivity.

A high retention of radioactivity is taken to indicate the incorporation of that specific amino acid into a polymer by the known bases present. This method was the touchstone to the code breaking technique. When first tried with polyU as the messenger RNA, the protein formed was poly-phenylalanine. The next code word discovered was that of polyC (poly-cytadylic acid) which was found to code for poly-proline. Other code words followed in rapid succession with confirmation from other laboratories working on this problem by other methods of attack. A summary of all code words to date of this report (June, 1965) is included in the section of tables on page 39 of this report.

Nirenberg also summarizes all information thus far reported on the activity of the amino acid from the time it is introduced into the cell by the activating enzyme to its incorporation into the complete and separate enzyme or protein. A diagram of this activity is to be seen on page 38 of this report. An explanation of this diagram follows.

Beginning at the cell membrane, which has been previously characterized in this report, there are some twenty, at least, activating enzymes (E_1) which serve to get the amino acid (A.A.) into the cell. This combination derives its energy from ATP. During this attachment a di-phosphate (pp) is released. The amino acid with the activating enzyme and Adenine monophosphate (AMP) now enter a reaction with transfer RNA (same as soluble RNA) in which the E_1 is released to activate other amino acids while the adenine monophosphate returns to be recombined with the diphosphate in the Kreb's cycle. The transRNA is now linked with the amino acid. There is good reason to believe that all the activity thus far discussed takes place in or at the cell membrane. TrRNA + A.A. now are ready to attach to the mesRNA. It will be noted that the amino acid has not

been identified. This will be considered later in this discussion.

While the foregoing activities are taking place in the membrane the nucleus prepares the RNA messages for transcription. RNA has recently been confirmed to be of three types; all made on the DNA template. This was discussed in "Source of the Third RNA" in SCIENTIFIC AMERICAN (March, 1963). Another source states that Ribosomal RNA shows evidence of being made on the nucleoli in higher forms of life (Liau, 1963). One of the surprising discoveries was that DNA has approximately forty sites for the manufacture of transfer RNA. Why the cell needs forty types of transfer RNA to deliver twenty types of amino acids has not yet been discovered. No reason was offered in this note. It seems, however, that the code anagrams such as TAC and CAT may offer a clue. These may "plug in" in a reversed fashion but still carry the same amino acids. This may also explain highly mutable sites in certain DNAs. The developments on this will be interesting to watch. Nirenberg makes no point of this in his table of code words, whether they read one way or the other. He simply gives the combination, not the order (1963). His later experiments (1965) use oligoribo-nucleotides of known order after the method of Ochoa (Speyer, 1962; Wabah, 1963) and Khorana (Byrd, 1965).

Whatever the case may be, there are three types of RNA. One type is the messenger (mRNA), so named because it carries the templated information of what is to go into the protein from the "gene" on the DNA to the ribosome in the cytoplasm where the protein is to be synthesized. A second type of RNA is the Ribosome RNA which makes up about half of the ribosome; the other half being the mesRNA. The third type, of course, is the "soluble" or transferRNA (tRNA). The structure of a tRNA for alanine from yeast extract has recently been published by Holley (1965a).

The discovery that brief treatment with takadiastase RNAse T1 and with Pancreatic RNAse by Penswick et al. (1965) would cleave the alanine RNA into two large fragments opened the way to systematic degradation of this molecule. Holley and his collegues have recently achieved this by proving the structure of this molecule. He gives the complete sequence as well as three possible conformations. To this time the structure was thought to be a twined helices with the codon somewhere in the middle, but he has found that only short stretches of one to no more than seven nucleotides are complementary. This gives many possible structures within the tRNA that may account for other enzymatic activity. The anti-codon sequence is not known; but he proposes some suspected sites. Strangely, these are not at the middle as some had assumed. The tRNA has a guanilic phosphate terminus and a cytosine-cytosine-adenine terminus. The amino acid is attached to transfer RNA by the activating enzyme making an amino acid bond. The three RNAs migrate from their DNA source to the cytoplasm as illustrated in the diagram.

Upon arrival at the active site of mRNA the tRNA anti-codon "recognizes" its complementary codon. Recent work by Alexander Rich (1963) seems to indicate that the ribosome or polyribosome performs the function of assembly for the protein. The secondary structure of the tRNAs as noted above may also aid in this enzymatic activity. Since the linear structure for alanine tRNA has been found, the structure for the forty or fifty others will surely follow. There may be some adjustment of this technique to give the structure of ribosomal RNA and protein. After the ribosome has apparently done its work and moved off the end of the mRNA, it then attaches to another mRNA. The assembled protein or enzyme then moves out into the cytoplasm to perform its functions. The ribosome now removes the

hydrogen bonding between the code bases of mRNA and tRNA. The tRNAs are then returned to pick up more amino acids. This is the continuous route of protein synthesis. Rich indicates that four, five or six ribosomes can perform their work at one time on mRNA. If the amino acid back at the membrane was phenylalanine the special activating enzyme attaches this amino acid to its own particular tRNA which has the anti-codon AAA to "recognize" polyU codons. Thus polyphenylalanine is the protein formed.

That each of the tRNAs are specific for one amino acid was exhibited in a unique experiment by Chapeville (1962). In the diagram notice that the structure of cystine is similar to alanine except for the sulphur atom on cystine. By attaching each of these to their respective tRNA then carefully removing the sulphur from the cystine with Rayney Nickel catalyst the cystine was converted to alanine without affecting the tRNA. The tRNA continued to code for UUG which is the code for cystine rather than CCG or UGG which is the code for alanine. (See chart, Nirenberg et al., page 39 of this report.) This would indicate that each of the tRNAs is specific for one amino acid.

Davies (1964) had cast some doubt on the codon dictionaries because streptomycin caused a binding to mRNA which would flood the cell with defective protein. Pestka (1965) has checked this possibility but found that their resistant strain ribosomes were not effected by the presence of Streptomycin. This summary of work is in agreement with findings of Bernfield (1965), Frankel-Conrat (1965), Jones (1962), Roberts (1962), Tsugita (1960) and Turpin (1965).

Nirenberg, in his latest article (1965), has proved the codons for approximately 45 tRNAs of E. coli in cell free systems. This is the corrected chart given on page 39 of this report.

The technique of hybridization of nucleic acids has been very fruitful in answering many questions about the behavior of nucleic acids. Spiegelman (1964) discusses the recent history of RNA and DNA experiments. The techniques they used involved the discovery by Marmur et al (1960-61) that gentle heat denaturing of DNA would cause the strands to separate to their single components. Reheating with slow cooling would cause the strands to reform into the double helix. Recombination occurs only between DNA from the same organism or closely related organisms. Spiegelman first used P-32 to label E. coli virus T2 m RNA and radioactive hydrogen (H-3) to label DNA from the same organism. By cesium chloride density centrifugation at 30,000 RPM for three days the bottom most fraction was found to contain RNA with RNA-DNA hybrids in the middle and DNA is one of the upper most fractions. Ultraviolet light at 260 m u was later used to identify DNA instead of the radioisotope label. Hybrids did occur and only between RNA and/or DNA of the same or closely related organisms. Hayashi, one of Spiegelman's collegues, used ribonuclease on RNA in hybridization studies. This enzyme can remove any uncoupled RNA from RNA-DNA hybrids.

In other experiments Spiegelman and his student Yankofsky (1963) tested for the origin of ribosomal RNA and tRNA. Since the cistron for either of these is only about a ten-thousandth the length of the DNA template radioactive labels were again used to tag the RNA. He discusses the difficulty of designing a controlled experiment to obtain this information at these difficult levels. However, their results clearly show that colon bacillus ribosomal RNA is produced on a DNA template. They also found that Bacillus megaterium requires approximately .18 percent of the DNA for one fraction of RNA and .14 percent for the other. This information was obtained by labeled RNA-DNA saturation hybrids.

Transfer RNA was tested for origin by the same methods. Although more difficult than the ribosome RNA tests, it was found to hybridize with about 0.25 percent of the DNA's total length. Further tests on tRNA showed that it is highly species specific, contrary to the findings of others. TRNA from E. coli did not hybridize with B. megaterium. This question is yet to be resolved; but he postulates that the structure of the amino acids other than the three codon acids may have some bearing on this. Holley (1965a) has done recent work which may present techniques to resolve this question. He has recently published the sequential readout of the nucleic acids in alanine tRNA from yeast extracts and three conformities which they may take in their secondary structure. These secondary structures, as noted elsewhere in this report, may explain certain enzymatic activity either associated with or due to tRNA. Proof of structure of all the other twenty to sixty tRNAs is sure to follow in rapid succession now that the technique for systematic degradation has been found.

At the end of this article Spiegelman presents evidence for which strand of DNA is readout for RNA replication. Of the three possibilities: (1) all of both strands, (2) one or the other at any given point, (3) only one, he shows by experimentation with a special case of virus that only one strand is used. Frankel-Conrat (1964) and Edgar (1965) both state that the cellular mechanism is disrupted and taken over by the virus. This would cast some doubt on Spiegelman's hypothesis. Zubay (1962) postulates a three stranded structure for the synthesis of mRNA. As can be noted from the structure proposed by Watson and Crick, the complementary base pairs are offset slightly from the center of the helices. This leaves a slot running around the entire helices from one end to the other. Zubay presents biochemical evidance for mRNA being synthesized in

this slot. This gets around the problem of uncoiling and which coil does the work. However, his proposed structure raises more questions than it answers. Regardless of its elegance of structure, it does not take into account the role of histones in genetic control. How would the spatial arrangement permit RNA polymerase to produce its effects? And most of all, this proposed structure ignores the vast amount of evidence confirming chromosome puffs and lampbrush configurations of DNA as the active sites of RNA replication. Guild and Robison (1963) state that RNA polymerase copies only one of the strands into a useful message from the two stranded structure. The question of which of the two strands of the DNA are read-out into mRNA is far from settled at this writing. Solution of this question may shed some light on what the classic dominant or recessive is.

Although the original technique of hybridization of nucleic acids was exceedingly fruitful in answering many questions about genetics at the molecular level, recent technical discoveries may supplant this method. The long periods of centrifugation are being replaced by the use of packed chromatographic or adsorption columns. Cellulose acetate, cellulose nitrate, agar, ion exchange resins and various other media are being used to couple the DNA. RNA is then passed through the column to hybridize. RNAse can strip any uncoupled RNA leaving only the hybrid. Gentle heat will remove the RNA. Single gene areas have been identified in a similar technique by Bautz (1962). A suggested technique by an Oklahoma State University group (Erhan et al, 1965) advocates using the proper polyribonucleotide attached to a chromatographic material. Cell extracts could then be filtered through the column to attach the tRNAs to the synthetic polyribonucleotide. These could then be uncoupled for

use in further experiments, perhaps analysis by the method of Holley. Although the literature has not so stated DNAse might be used to dissolve any DNA not coupled to RNA. To this writer's knowledge this has not been tried. If such were the case mRNA for some known protein, perhaps hemiglobin made in reticulytes, could be coupled back to the original DNA and also coupled to its respective tRNAs. A complete series of experiments could be set up which would elucidate each step from DNA to the resulting enzyme. Alexander Rich (1963) describes the handling of reticulytes and mRNA.

Michael Beer and Moudriankis (1962-1965) have been using an interesting technique in the quest for the genetic code. They proposed, in 1962, the use of a heavy marker atom or molecule which would selectively bond to one of the bases in DNA. The marked DNA would then be viewed under extremely high resolution in the electron microscope. Their first report stated that a diazonium salt of 8-amino-1,3,6-napthalenetrisulfonic acid stained with uranyl acetate would provide such a marker. The structure for this molecule is approximately 10 Angstroms across. This is well within the resolution of present electron microscopes as they have shown. A later report (1965) stated that the coupling of 2-amino-p-benzene-disulfonic acid was highly selective for guanine. The mild reaction caused no excessive degradation of the DNA chain. Markers of the uranyl stain were observed at 10-12 Angstroms. Other sites showed a proximity of the markers as close as 6-7 Angstroms, indicating two successive guanines. Some difficulty was encountered, however, for only eighty percent of the guanines were labeled. Successive photographs may be necessary to assure detection of all the guanine sties if the unmarked residues are not the same ones in each case. This technique

may be useful in confirming suspected sequences from living material rather than the synthetic materials so far used by most of the investigators. Other markers will need to be found for the other three bases. They show several clear electron micrographs with a fine dotted line running across the field of view. The "gene" can now be seen.

As can be seen, the concept of the "gene", if such exists, is now drastically altered from the Mendelian concept. The whole sequence of events from DNA, mRNA, tRNA, codon, enzyme to the resulting protein or by product can be interrupted or altered to cause a mutation. The whole scheme is not complete at this writing, but some order begins to evolve.

CHAPTER III

POSSIBLE CHEMICAL EXPLANATIONS FOR CLASSIC GENETIC MECHANISMS

How do the chromosomes work in the cell? Are all of the genes operative at once, manufacturing all of the possible RNAs which, in turn, will make all the possible enzymes the cell is capable of making? This is highly unlikely in light of reason and current evidence. Specialized cells tend to make the substance for which they are destined to function. For example, epithelial cells would tend to make proteins of the keratonus type for hair, scales or feathers. Connective tissue would produce lipo proteins, while nerve cells would more likely produce the basic proteins. It is known that each cell of the body has a full complement of chromosomes; why do they not all operate at once?

Recent evidence by Ru Chu and Bonner (1962) indicates that histones play a controlling role on the chromosomes by covering in some way those parts which are nonfunctional in the cell. They found that the pea plant embryo had eighty percent of its chromosomes covered with histone. They further showed that the remaining twenty percent was actually doing the cell's work. This would give some indication of the small amount of genetic material actually working in a mature differentiated cell, for these are immature, undifferentiated embryonic cells. Izawa, Allfrey and Mirsky (1963) have found much of the same true in amphibious oöcytes and cell free chromosomes. Actinomycin D, which inhibits RNA synthesis,

causes the disappearance of loops on chromosomes. Arginine rich histones also stop chromosome activity and disappearance of loops. Histone coverage may in some way explain the phenomenon of heterosis. Bloch (1962) suggests that histones are probably not made by the gene with which it is associated. Histones are composed of the basic amino acids, arginine and lysine, and are probably made on much simpler strips of the chromosome. This indicates a dynamic control of one gene over another. If histones are made on the chromosome, this may or may not be the explanation of the classic enhansor or suppressor gene. Another explanation of the suppressor gene might be the presence of an enzyme coded at another site which would interfere with the particular phene under surveillance. This interference could take the form of using the protomaterial of the phene to catalize some other product. The enhansor may work in the reverse wherein the enzyme coded elsewhere would aid in the catalytic action to the phene. Whatever the total effect, Hnilica (1964), Liau (1965), Beerman (1964), and Junis (1965) have all shown that histones show various control over nucleic synthesis. At this time it is not known how this fits into the overall genetic control.

The entire genetic mechanism of an organism may not be as simple as the plan outlined in the foregoing chapter may seem. Sanger (1965) has studied the effects of non-chromosomal genes. These have been known since 1908 when Carl Corens, a German botanist, first described them. Many organisms exhibit non-chromosomal systems, among them: insects, flowering plants, algae, yeasts, and fungi. These systems are identified by their failure to conform to Mendelian crosses. Sanger has found a type of DNA in the chloroplasts of Chlamydomonos species that does not correspond to nuclear DNA of that species. These DNA's can be distinguished

on the basis of base composition percentages of the two. Other workers have found that a special type of DNA is to be found in the mitochondria of animal cells. She presents evidence for the orderly replication and division of these nonchromosomal systems during cell division. It is her opinion that the evolution of these systems present a method of flexibility for organelle growth in response to changing environmental conditions. Any explanation of genetic control must also take into account this system as well as that found in the nucleus. Ray (1964) gives similar findings and characterizes the DNA from Euglena chloroplasts. The environment apparently controls genetic activity after it is conceived more than early investigators were led to believe.

That hormones show a control over genes is being investigated by Davidson (1965) and many others in this field. Differentiation of tissue and specialization of the cells of which tissue is composed depends upon the development of a specific pattern of protein synthesis in these cells. Feedback controls from the environment to the DNA master control of the cell is necessary for the adaptions which cells are called upon to make. Hormones, or the enzymatic proteins which affect reactions remote from their source, are proving to be that feedback mechanism within the limits of the system. Hormones are believed to attach to the outer membrane of the cell and there regulate the genetic mechanisms. That hormones control genetic action is evidenced by the presence of actinomycin D which is a specific inhibitor of mesRNA synthesis on the DNA template. "Genes" thus inhibited do not show some or all of the effects of a hormone injection.

Steroid hormones show a strong control over protein synthesis. Estrogen administered to ovarectomized females increases protein synthesis

in uterine cells upwards of 300 percent. Puromycin, a specific inhibitor of protein synthesis, negates the protein synthesis of estrogen when administered concurrently. Estrogen has been shown to not only increase the synthesis of RNA but also increases the activity of RNA-DNA polymerase. How one hormone, estrogen, can activate these as well as the various other steps, such as the some forty to sixty species of transfer RNA as well as ribosomal RNA, is still to be discovered.

Many other hormonal effects are cited in this article as evidence of "gene" activation. How these relatively large molecules can exert a control over the genetic mechanism from the cell membrane remains a mystery at this writing, but these controls must be considered in the dynamic balance of selective gene activation.

Changeux (1965) gives two possible pathways in which environmental control can occur. In one, the end product L-isoleucine signals the regulation of the first step, enzyme L-threonine deaminase, while the high concentration of the end product, L-isoleucine, suppresses each of the five other enzymatic steps.

Others propose the binding of a repressor to the operator gene. Some regulating action caused by the environment may make the enzyme change conformity, thus altering the active site. X-ray diffraction maps of oxygenated hemoglobin suggest this change of conformity and structure. Once an enzyme has made sufficient product it may turn itself off by changing its shape.

As previously noted, the presence of histones may in some way explain heterosis since most of the genetic material would be working if the histone uncovered the gene. This would not necessarily be mRNA for the same total effect might be observed if more ribosomes were made to

readout the available information. More fundamental, perhaps, is the determination of the cause of dominance and recessiveness. Cleavenger (1964) has worked with flower pigments to find that classic dominant and recessive traits, as well as their suppressors or enhansors, show a correlation to the chromatographed pigments. The determination of dominance or recessiveness will probably have to await the resolution of how the cistron is readout in mRNA synthesis-whether a single or double strand readout.

Hadorn (1962) has found that chromatograms of fruit fly eye pigments also show close correlation to the genetic make-up of the individual. He shows that recessive genes in heterogeny do have a distinct biological and biochemical effect even though they produce no visible effect. Totally recessive "genes" occur rarely, if at all, he states. His findings tend to show that both cistrons are readout on the "gene" to produce two enzymes.

The "gene", it now seems clear, is the dynamic balance of the enzymes of a system. That one enzyme may predominate others so that its product in the enzymatic chain of events will be the visable one seems to explain dominance. The recessive is probably the effect of what would be in the organism's enzymatic system if the "dominant" enzyme were not present, provided the system is viable in both cases. This may also be an explanation of hypostasis in a modified form.

The discussion of Crick's work (1962) included the statement that the "shifting into phase" of the three base codon may explain position effect. Adjacent cistrons without a "stop" between them may activate or deactivate each other by addition or deletion mutants.

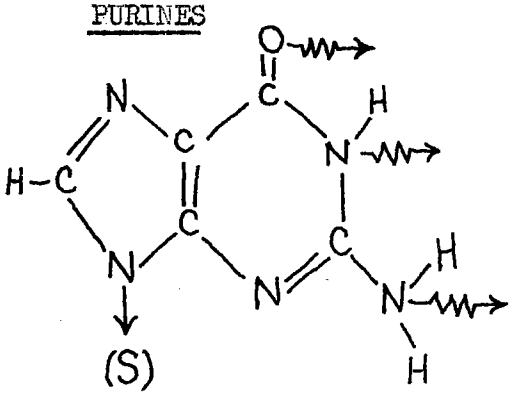
Pleiotropy would seem to be deceptively simple to explain since the

cistron required for one enzyme would be present on the chromosomes in all the cells. Whether this particular cistron was operative to complete a certain chain of enzymatic events would determine the extent of pleiotropy. Here the whole feedback system in differentiated cells would also determine genetic control.

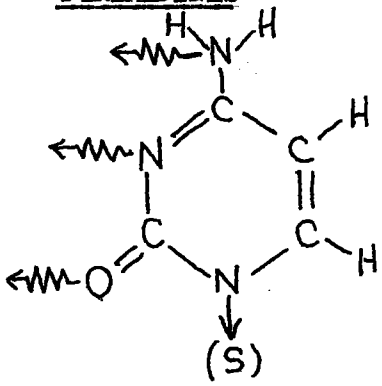
Edward Frankel (1964) has published a book written for the high school or an interested individual with moderate background. This book explains the effects of this current research without going into much detail. Asimov (1962) has also written a book about the current research in molecular biology and genetics. He assumes a slightly higher background for his readers than does Frankel, however. Both of these books are recommended for further reading on this subject.

It is now apparent that the chromosomes carry (1) the information to make enzymes and all the other proteins which go into making up the gross structure of the individual (2) information coding for the three types of RNA which will eventually do the work of making these proteins and (3) information to make histones to turn these various "genes" on or off. This probably accounts for the long blank expanses of distance seen on chromosome maps, for these chromosomes cannot now be detected unless a mutation occurs. Any explanation of genetic control must now take into consideration the presence of non-chromosomal genes, either detected or not, as well as hormones, histones, and enzymatic feedback to the DNA from the environment.

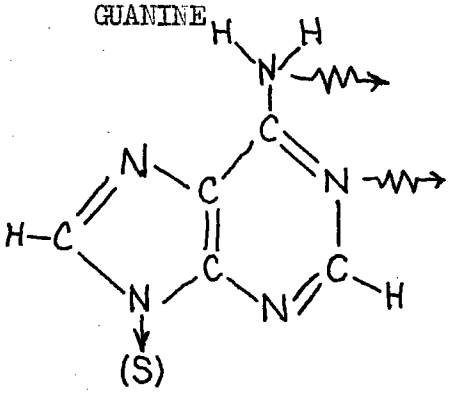
PURINES



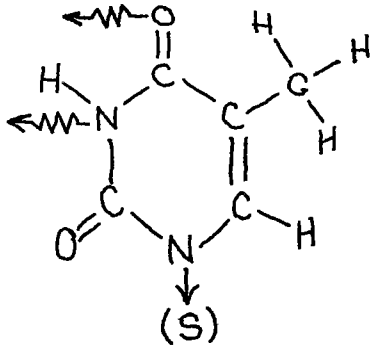
PYRIMIDINES



GUANINE

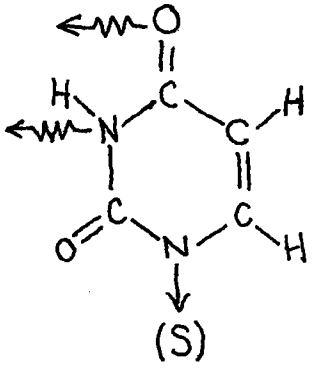


CYTOSINE



ADENINE

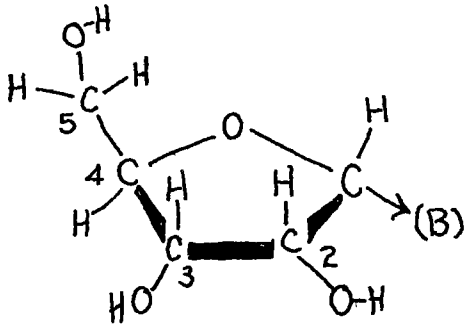
THYMINE



(S) Indicates attachment to sugar-phosphate chain

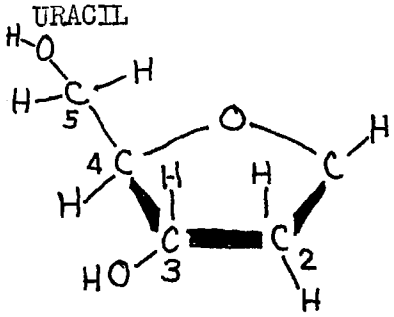
(B) Indicates attachment to base

Indicates hydrogen bond to base complement



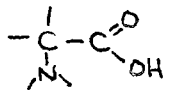
D-ribose

URACIL

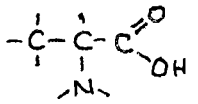


2-deoxy-D-ribose

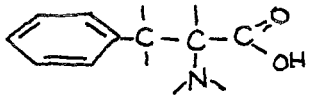
AMINO ACIDS



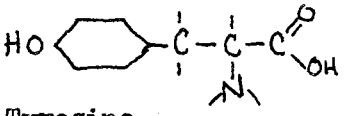
Glycine



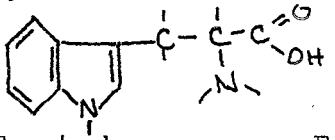
Alanine



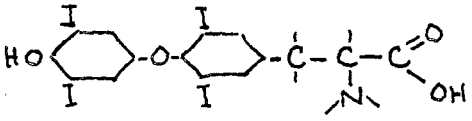
Phenylalanine E



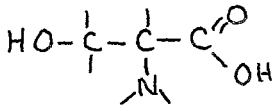
Tyrosine



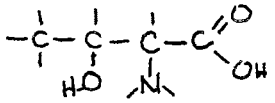
Tryptophane E



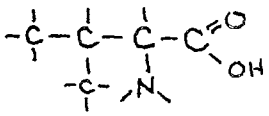
Thyroxin E



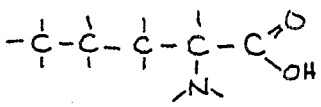
Serine



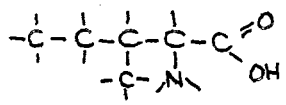
Threonine E



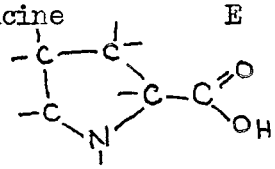
Valine E



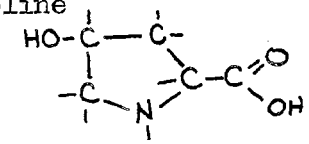
Leucine E



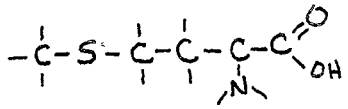
Isoleucine E



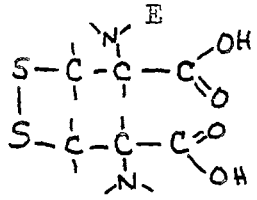
Proline



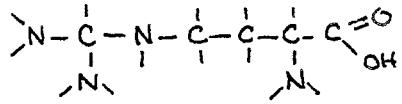
Hydroxy Proline



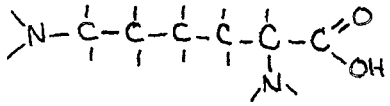
Methionine E



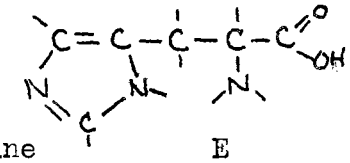
Cystine



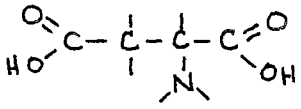
Arginine E



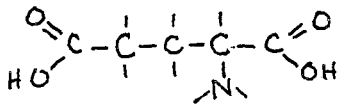
Lysine E



Histidine E



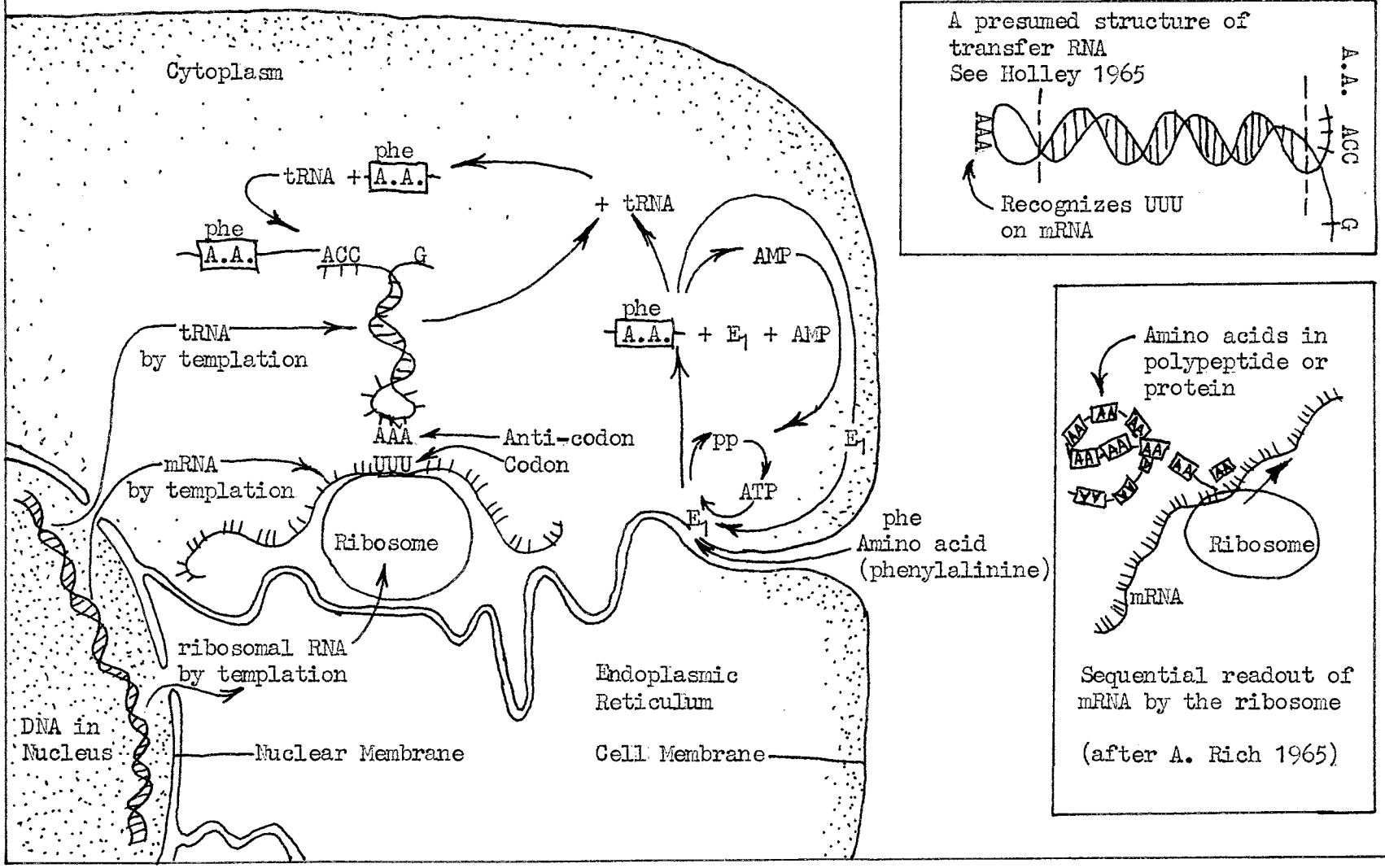
Aspartic acid



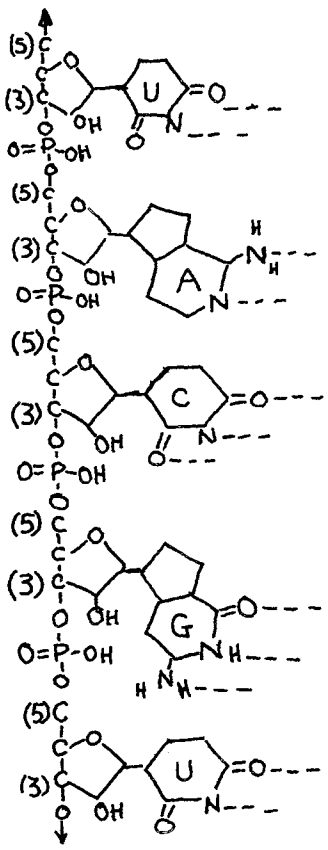
Glutamic Acid

E indicates essential to human consumption

SCHEMATIC TRANSPORT OF AMINO ACIDS INTO PROTEIN SYNTHESIS
BY POLYURIDYLIC ACID mRNA of NIRENBERG, 1962



STRUCTURE OF
RIBONUCLEIC ACID
(single strand)



Linear form shown
Only pertinent atoms shown
See page 35 for complete structures

Hydrogen bonds - - -

STRUCTURE OF
DEOXYRIBONUCLEIC ACID

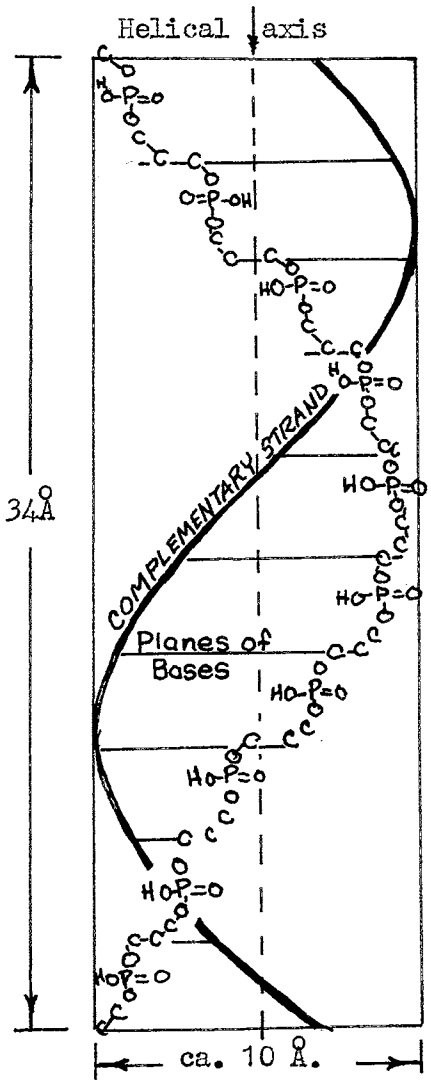
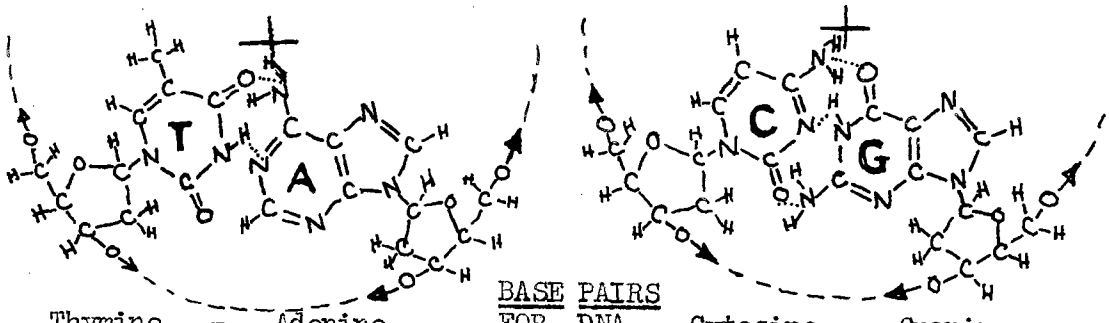


Diagram of phosphate linkages between bases for one strand of DNA



Thymine - Adenine

BASE PAIRS FOR DNA

Cytosine - Guanine

TABLE I

NUCLEOTIDE SEQUENCES OF RNA CODONS[1]

*UpUpU *UpUpC	Phe	*UpCpU *UpCpC	Ser	*UpGpU UpGpC	Cys	*UpApU *UpApC	Tyr
UpUpA *UpUpG	Leu	UpCpA *UpCpG	Ser	*UpGpA UpGpG	Nonsense# or Trypt	*UpApA *UpApG	Nonsense/
*CpUpU *CpUpC	Leu or Nonsense#	*CpCpU *CpCpC	Pro	CpGpU *CpGpC	Arg	*CpApU *CpApC	His
CpUpA *CpUpG	Leu	*CpCpA CpCpG	Pro	*CpGpA CpGpG	Arg	*CpApA *CpApG	Glu-NH_2
*ApUpU *ApUpC	Ileu	*ApCpU *ApCpC	Thr	*ApGpU *ApGpC	Ser	*ApApU *ApApC	Asp-NH_2
ApUpA *ApUpG	Met	*ApCpA *ApCpG	Thr	*ApGpA ApGpG	Arg. or Nonsense#	*ApApA *ApApG	Lys
*GpUpU GpUpC	Val	*GpCpU GpCpC	Ala	*GpGpU GpGpC	Gly	*GpApU *GpApC	Asp
GpUpA GpUpG	Val	GpCpA GpCpG	Ala	GpGpA GpGpG	Gly	*GpApA GpApG	Glu

*indicates trinucleotides which have shown experimental evidence of coding for the amino acid, others are predicted.

/UpApA and UpApG may correspond to terminator, or Ser codons, in other strains of E. coli.

#It is possible that these sequences are readable internal, but nonreadable terminal, codons.

[1]Nirenberg, M., Leder, P., Bernfield, M., Brimacrombe, R., Turpin, J., Rottman, F., and O'Neal, C. (1965).

CITED LITERATURE

Allfrey, Vincent G., and Alfred E. Mirsky 1961. How cells make molecules. Scientific American. 205(3): 78-82.

Allfrey, Vincent G., V. C. Littow, and A. E. Mirsky 1963. On the role of histones in regulating ribonucleic acid synthesis in the cell nucleus. Proc. Natl. Acad. Sci., U. S. 49:414.

Apgar, Jean, George A. Everett, and Robert Holley 1965. Isolation of large oligonucleotide fragments from the alanine RNA. Proc. Natl. Acad. Sci., U.S. 53(5):546-548.

Asimov, Isaac 1962. The genetic code. New York: Clarkson N. Potter, Inc.

Avery, O. T., C. M. MacLeod, and M. McCarty 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. J. Expl. Med. 79:137-158.

Bautz, E. K. F. 1963. Physical properties of messenger RNA of bacteriophage T4. Proc. Natl. Acad. Sci., U.S. 49:68.

Beadle, G. W. and E. L. Tatum 1941. Genetic control of biochemical reactions in neurospora. Proc. Natl. Acad. Sci., U. S. (27)409-416.

Beer, Michael and Evangelos N. Moudrianakis 1962. Determination of base sequence in nucleic acids with the electron microscope: visibility of a marker. Proc. Natl. Acad. Sci., U. S. 48(6):409-416.

Beerman, Wolfgang and Ulrich Clever 1964. Chromosome puffs. Scientific American. 210(4):50-58.

Benzer, Seymore 1962a. The fine structure of the gene. Scientific American. 206(1):70-84.

Benzer, Seymore 1955. Fine structure of a genetic region in bacteriophage. Proc. Natl. Acad. Sci., U. S. 41:344-354.

Benzer, Seymore 1962b. A change from nonsense to sense in the genetic code. Proc. Natl. Acad. Sci., U. S. 48:1114.

Bernfield, Merton R. and Marshall W. Nirenberg 1965. RNA codewords and protein synthesis. Science. 147(3657):479-484.

Bloch, David P. 1962. Derivation of histone specificity. Proc. Natl. Acad. Sci., U. S. 48:324-326.

Bretscher, M. S. and M. Grunberg-Manago 1962. Polyribonucleotide directed protein synthesis using an E. coli cell free system. Nature. 195(4838):283-284.

Byrd, C., E. Ohtsuka, M. W. Moon and H. G. Khorana 1965. Synthetic deoxyribo-oligonucleotides as templates for the DNA polymerase of E. coli: new DNA like polymers containing repeating nucleoide sequences. Proc. Natl. Acad. Sci., U. S. 53(1):79-84.

Champe, Sewel P. and Seymore Benzer 1962. Reversal of mutant phenotypes by 5-fluoro-uracil, an approach to nucleotide sequences in messenger RNA. Proc. Natl. Acad. Sci., U. S. 48:532-546.

Changeux, Jean Pierre 1965. Control of biochemical reactions. Scientific American. 214(4):36-45.

Chapaeville, Francois, Fritz Lipman, Gunter Von Ehrenstein, Bernard Weisblum, William J. Roy and Seymore Benzer 1962. On the role of soluble RNA in coding for amino acids. Proc. Natl. Acad. Sci., U. S. 48:1086-1092.

Chiu, Ru, C. Huang and James Bonner 1962. Histone, a suppressor of chromosomal RNA synthesis. Proc. Natl. Acad. Sci., U. S. 48:1215.

Clevenger, Sarah 1964. Flower pigments. Scientific American. 210(6): 84-92.

Creighton, Harriet S. and Barbara McClintock 1931. A correlation of cytological and genetical crossing over in Zea mays. Proc. Natl. Acad. Sci., U. S. 17:492-497.

Crick, F. H. 1957. Structure of nucleic acid. Scientific American. 197(3):47-53.

Crick, F. H. and L. Barnet et al 1961. Nature. 192:1227.

Crick, F. H. 1962. The genetic code. Scientific American. 207(4):66-74.

Davidson, Erick H. 1965. Hormones and genes. Scientific American. 212 (6):36-45.

Davies, Julian E. and Luigi Gorini 1964. Streptomycin and the genetic code. Scientific American. 211(1):44-46.

Edgar, R. S. and R. H. Epstein 1965. The genetics of a bacterial virus. Scientific American. 212(4):70-75.

Eikenberry, Eric F. and Alexander Rich 1965. The direction of reading messenger RNA during protein synthesis. Proc. Natl. Acad. Sci., U. S. 53(3):668-678.

Erhan, Semih, Lulla G. Northrup and Franklin P. Leach 1965. A method potentially useful for establishing base sequences in codewords. Proc. Natl. Acad. Sci., U. S. 53(3):646-652.

Frankel-Conrat, Heinz 1965. The genetic code of a virus. Scientific American. 211(4):47-54.

Frankel, Edward 1964. DNA - Ladder of Life. New York: McGraw Hill Co., Inc.

Geiduschek, Peter, John W. Moohr and Samuel B. Weiss 1962. The secondary structure of complimentary RNA. Proc. Natl. Acad. Sci., U. S. 48 (6):1078-1087.

Guild, W. R. and M. Robinson 1963. Evidence for message reading from a unique strand of Pneumococcal DNA. Proc. Natl. Sci. Acad., U. S. 50:106-111.

Hadorn, Ernst 1962. Fractionating the fruit fly. Scientific American. 206(4):100-110.

Hnilica, Lubomor S. and Daniel Billen 1964. The effect of DNA histone interactions on the biosynthesis of DNA in vitro. Biochem. Biophys. Acta. 91(2):271-280.

Holley, R. W. 1965a. Structure of a ribonucleic acid. Science. 147 (3664):1462-1465.

Holter, Heinz 1961. How things get into cells. Scientific American. 205(3):167-180.

Hurwitz, Jerard and J. J. Furth 1962. Messenger RNA. Scientific American. 206(2):41-49.

Ingram, Vernon M. 1958. How do genes act. Scientific American. 198 (1) January 1958.

Izawa, Mitsuo, V. G. Allfrey and A. E. Mirskey 1963. The relationship between RNA synthesis and loop structure in lamp brush chromosomes. Proc. Natl. Acad. Sci., U. S. 49:544.

Jones, Oliver W. and Marshall W. Nirenberg 1962. Quantitative survey of RNA codewords. Proc. Natl. Acad. Sci., U. S. 48(12):2115-2123.

Junis, George J. 1965. Interphase deoxyribonucleic acid condensation late deoxyribonucleic acid replicatin, and gene inactivation. Nature. 205(4968):311-312.

Kohrana, H. Gobind 1964. DNA made to specifications. Scientific American. 210(6):56.

Krakow, Joseph S. and Severo Ochoa 1963. Ribonucleic acid polymerase of Azotobacter vinelandii, I. Priming by polyribonucleotides. Proc. Natl. Acad. Sci., U. S. 49:88.

Lederburg, Joshua and Edward L. Tatum 1946. Gene recombination in Escherichia coli. Nature. 158:558.

Lerman, L. S. 1963. The structure of the DNA- acridine complex. Proc. Natl. Acad. Sci., U. S. 49:94.

Liau, Ming C., Lubomir S. Hnilica and Robert B. Hurlburt 1965. Regulation of RNA synthesis in isolated nucleoli by histones and nucleolar proteins. Proc. Natl. Acad. Sci., U. S. 53(3):626-633.

Li, Choh Hao 1963. The ATOH molecule. Scientific American. 209(1):46.

Marmur, Julius and Paul Doty 1961. Thermal renaturation of deoxyribonucleic acids. J. Molec. Biol. 3(5):585-594.

Marmur, Julius and Dorothy Lane 1960. Strand separation and specific recombination in deoxyribonucleic acids: biological studies. Proc. Natl. Acad. Sci., U. S. 46(4):453-461.

Mondriankis, Evangelos N. and Michael Beer 1965. Base sequence determination in nucleic acids with the electron microscope, III. Chemistry and microscopy of Guanine-labeled DNA. Proc. Natl. Acad. Sci., U. S. 53(3):564-571.

Newrath, Hans 1964. Protein digesting enzymes. Scientific American. 211(6):68-79.

Nirenberg, Marshall W. and J. Heinrich Matthaei 1961. The dependence of cell free protein synthesis in E. coli upon naturally occurring or synthetic polyribolucleotides. Proc. Natl. Acad. Sci., U. S. 47(10):1588-1602.

Nirenberg, Marshall W. 1963. The genetic code. Scientific American. 208(3):80-90.

Nirenberg, M., P. Leder, M. Bernfield, R. Brimacrombe, J. Turpin, F. Rottman and C. O'Neal 1965. RNA codewords and protein synthesis, VII. on the general nature of the RNA. Proc. Natl. Acad. Sci., U. S. 53(5):1161.

Nishimura, S. and David Novelli 1965. Disassociation of amino acid acceptor function of soluble RNA from its transfer function. Proc. Natl. Acad. Sci., U. S. 53(1):179-184.

Nossal, G. J. V. 1964. How cells make antibodies. Scientific American. 211(6):106-114.

Painter, T. S. 1925. Salivary chromosomes and the attack on the chromosome. J. of Heredity. 25:464-476.

Penswick, John Robert and Robert Holley 1965. Specific cleavage of the yeast Alanine RNA into two large fragments. Proc. Natl. Acad. Sci., U. S. 53(3):543-546.

Pestka, Sidney, Richard Marshall and Marshall Nirenberg 1965. RNA codewords and protein synthesis, V. Effect of streptomycin on the formation of ribosome-RNA complexes. Proc. Natl. Acad. Sci., U. S. 53(3):639-636.

Ray, Dan S. and P. C. Hanawalt 1964. Properties of the satellite DNA associated with the chloroplasts of Euglena gracilis. J. Mod. Biol. 9(3):812-824.

Rich, Alexander 1963. Polyribosomes. Scientific American. 209(6): 44-53.

Roberts, Richard B. 1962. Alternative codes and templates. Proc. Natl. Acad. Sci. 48(5):897-900.

Robertson, J. David 1962. The membrane of the living cell. Scientific American. 306(4):65-71.

Sager, Ruth 1965. Genes outside the chromosomes. Scientific American. 212(1):70-79.

Speyer, J. F., P. Lengeyel, C. Basilio and S. Ochoa 1962. Synthetic polynucleotides and the amino acid code, IV. Proc. Natl. Acad. Sci., U. S. 48(3):441-448.

Spiegelman, S. 1964. Hybrid nucleic acids. Scientific American. 210(5):48-56.

Staehelin, Theophel 1964. Hypothesis confirmed (science and the citizen). Scientific American. 210(3):54.

Stein, William H. and Stanford Moore 1961. The chemical structure of protein. Scientific American. 204(2):81-92.

Sturtevant, A. H. 1913. The linear arrangement of six sex linked factors on Drosphila, as shown by their mode of association. J. Expml. Zool. 14:43-59.

Tsugita, A. and H. Frankel-Conrat 1960. The amino acid composition and C terminal sequence of a chemically evoked mutant of TMV. Proc. Natl. Acad. Sci., U. S. 46(5):636-642.

Tsugita, A., D. T. Gish, J. Young, H. Frankel-Conrat, C. A. Knight, and W. M. Stanley 1960. The complete amino acid sequence of the protein of tobacco mosaic virus. Proc. Natl. Acad. Sci., U. S. 46(11):1463-1469.

Turpin, Joel S., F. M. Rottman, R. L. C. Brimacombe, P. Leder, M. R. Bernfield and M. W. Nirenberg 1965. RNA codewords and protein synthesis, VI. On the nucleotide sequences of degenerate codeword sets for Isoleucine, Tyrosine, Asparagine and Lysine. Proc. Natl. Acad. Sci., U. S. 53(4):807-811.

Watson, J. D. and F. H. Crick 1953. Molecular structure of nucleic acids. Nature. 171(4356):737-738.

Weisblum, Bernard, Seymore Benzer and Robert Holley 1963. A physical basis for degeneracy in the amino acid code. Proc. Natl. Acad. Sci., U. S. 48(8):1449-1453.

Weiss, Samuel B. and Tokumasa Nakamoto 1961. The enzymatic synthesis of RNA. Proc. Natl. Acad. Sci., U. S. 47(9):1400-1405.

Wright, Sewell 1917. Color inheritance in mammals. J. of Heredity. 8:224-235.

Yankfofsky, S. A. and S. Spiegelman 1963. Distinct cistrons for the two ribosomed RNA components. Proc. Natl. Acad. Sci., U. S. 49:538-544.

Zinder, Norton D. and Joshua Lederberg 1952. Genetic exchange in Salmonella. J. of Bact. 64:679-699.

Zubay, Geofry 1962. A theory on the mechanism of messenger RNA synthesis. Proc. Natl. Acad. Sci., U. S. 48(4):456-462.

Zuckerkandle, Emile 1965. The evolution of hemoglobin. Scientific American. 212(5):110-118.

VITA

Oliver Palow Stafford Jr.

Candidate for the Degree of

Master of Science

Report: GENETIC CODING

Major Field: Natural Science

Biographical:

Personal Data: Born in Drumright, Oklahoma, June 4, 1927, the son of Oliver P. and Martha Dora Stafford.

Education: Attended grade school in New Lima, Oklahoma; graduated from Wewoka High School in 1945; attended George Peabody College for Teachers as a special student in 1954; received the Bachelor of Science degree in Education with a major in Education and Natural Science in May, 1959; completed the requirements for the Master of Science degree July, 1965.

Professional Experience: Began teaching experience in General Science and Mathematics in Burbank Jr. H igh School, Berkeley, California in 1959; left Berkeley and began teaching General Science and Physical Science at Stillwater Jr. High School, Stillwater, Oklahoma from 1960 to 1962; received a National Science Foundation grant for study in the Oklahoma State University during the academic year of 1962-63; returned to Stillwater Jr. High School to teach Biology in 1963 and Physical Science in 1964; have been a member of National Science Teachers Association, National Education Association, and Oklahoma Education Association.