EVOLUTION OF PROTEINS

by

June Vernon

Submitted as an Honors Paper
in the
Department of Biology

The University of North Carolina
at Greensboro
1969

Approved by:

Bruce M Eberhart

Director

Examining Committee:

Laura G. Anderston

Wm K Bates

James F. Wilson

TABLE OF CONTENTS

INTRODUCTION

As the topic Evolution of Proteins suggests, this paper is concerned with processes which occur in evolution at the molecular level to actually change the protein structure. The changes in protein structure become greater with increasing time so that a comparison of homologous proteins from organisms widely separated in time should reflect their evolutionary history and relationships. In addition to studies where proteins of different species are compared, comparisons of homologous proteins isolated from different organisms of the same species should reflect what evolutionary trends are at work at present in given species. This was the theory behind the present study in which the aryl-B-glucosidase from eleven strains of Neurospora crassa, which were geographically isolated, were compared electrophoretically to detect variations in structure. This study has many parallels. Two examples of particular interest here are Horowitz's discovery that the tyrosinase of two different strains of Neurospora crassa were distinguishable when subjected to electrophoresis (Horowitz, Fling, 1961). Mahadevan and Eberhart have also found differences in the thermolability of aryl-B-glucosidase from geographically separated strains (Mahadevan, Eberhart, 1964). These examples can be directly compared with the present study.

The evolutionary process at the molecular level is primarily due to mutations which alter the base sequence of DNA, resulting in the replacement of the residue originally coded for. The manifestation of evolutionary change at the structural level of the protein molecule is in the amino acid sequence, or the primary structure. A mutagenic alteration of DNA which would delete or add amino acids along the protein chain would produce such large effects that it is usually selected against. The mutagenic effect on protein that would represent the most probable evolutionary mechanism then is amino acid substitution. By a comparison of the different amino acids occurring at identical positions on the

homologous proteins, the amount of variation is determined.

The degree of variation is limited by the function of the protein in the organism. Change in amino acid sequence can occur only if the original function is maintained, otherwise the protein becomes useless unless a new function evolves with the change in residues. This, however, is very improbable. In addition to the limits on change imposed by the necessity of maintaining function, is the need to preserve structural features of the molecule. Since the residues do play an important role in maintaining the structure of the protein, especially at the tertiary level, the amount of variation is further restricted. Despite these limitations, considerable variations do exist among homologous proteins. Thus a determination of the amino acid sites which vary and those that tend to remain invariant will provide important clues to the role of the amino acids in the overall functioning protein. Examples of such studies being performed with cytochrome c and hemoglobin will be discussed later in this paper.

Whereas invariance prohibits polymorphism, variation provides a clue to the evolutionary relationships of the species compared and the time of their separation into distinct species. In all cases, the use of variation between homologous proteins as a criterion for the amount of time since divergence of the species has been valid, although the method cannot predict evolutionary distance between species precisely. The question of how one can be sure that each variable residue is the result of a single mutational step has been raised; but, as will be pointed out later in more detail, evolution seems to tend toward favorable substitutions by a single step. Zuckerkandl and Pauling (Bryson and Vogel, p. 148, 1965) believe that the comparisons of species at the polypeptide level is much more significant than at the organismal level. Their reasoning is based on the belief that the type, rather than the number of substitutions, is important. For instance, the substitution of a given residue may greatly alter the function; this would in turn affect the functioning of the organism, whereas, most substitutions are going to maintain the protein's original function. Thus, variations in amino acid sequence that are proportional to evolutionary time are not necessarily related

to changes due to adaptive requirements. The larger number of alterations would be conservative substitutions that worked as a function of evolutionary time. A second feature of this system is that the rate of effective substitution reflects the flexibility of the molecule concerning a particular function and the sequence required to carry out this function. With more flexibility the evolutionary rate will be high, whereas more rigid structural requirements will limit the number of successful substitutions and, thus, lower the rate at which species divergence occurs. The result is a more or less constant average rate since the latter events are rare. Reinforcing the predictability which the system allows is the possibility that the evolutionary process does not occur in a totally random fashion.

Based on findings thus far in this field of study, it has been postulated that natural selection has worked on the genetic mechanism of evolution at the level of the coding triplets to give a more non-random pattern of evolution. Recalling that function and structure are the conditions requiring constancy of amino acid sequence, it becomes apparent that the residues are the features of importance in their interactions within the protein. Since the amino acids can be classed into groups of similar chemical properties, a residue similar in its reactive potential to one it replaces in a given structure could perform the function effectively if that chemical property were the important feature of the residue initially present. This then would maintain the old function while introducing some new property in the molecule. This being the case, it would seem reasonable that natural selection had acted upon the base triplets so that mutations producing a single substitution tended to produce an amino acid of similar chemical properties.

Thus the genetic code could have undergone a period of evolution in which the base sequences that were most favored by natural selection were maintained. This would agree with the fact that the most probable substitutions occur in a single mutational step and further support the validity of these methods of comparing amino acid differences as an index to evolutionary separation.

MATERIALS AND METHODS

The aryl-B-glucosidase of eleven strains of Neurospora crassa was extracted and subjected to electrophoresis. The enzyme solutions studied were obtained from the following strains of Neurospora crassa (identified by FGSC numbers): 420 (Puerto Rico), 430 (North Africa), 433 (Phillipine Islands), 434 (Liberia), 435 (Fiji), 847 (Lein), 851 (Costa Rica), 852 (Costa Rica), 961 (Liberia), 967 (Liberia), and the laboratory wild type strain 74A-23-1A (USA). The strains were grown on "Glycerol complete media" in two-liter flasks for seven days. After the seven days, the conidia were harvested by pouring 100 ml. of sterile water into the flasks and shaking. This mixture was then filtered through spun glass to remove mycelia. The resulting conidial suspension was allowed to sit for thirty minutes during which time the conidia secreted aryl-B-glucosidase into the surrounding water solution. The conidia were then filtered out of the solution through a Buchner funnel on Watmann #1 filter paper. The remaining solutions from each strain were put into dialysis tubing and covered with carbowax. After nine hours in the refrigerator, the enzyme had been concentrated to a volume of fifty milliliters. These concentrated enzyme solutions were then used in all of the experiments. The enzyme preparations were kept frozen when not being used.

Preliminary tests of PNPG and the Lowry Protein assay were run to determine specific activity of the enzyme in each case. These tests indicated that there was enough activity to show up in the electrophoresis experiments and gave some idea as to the comparative activities of each. Specific activities of the enzymes were determined by using data from the PNPG and Lowry assays in the following formula:

$$\frac{A/C \times B}{D \times E} \times 1000$$

where, for each enzyme A = O.D. reading from the PNPG assay

B = dilution factor in the PNPG assay

C = reaction time in minutes (ten minutes)

D = mg/ml of enzyme solution actually used in the Lowry assay

E = ml. of enzyme used in the PNPG assay.

Electrophoresis was chosen as the means of detecting variations among homologous proteins. The basic principle behind electrophoresis is that charged particles will migrate in an electric field so that a characteristic migration pattern is associated with a given protein due to the charged character of the amino acids. Thus, an amino acid substitution might change the charged nature of the molecule and, hence, affect its migration in an electric field.

The electrophoresis was the primary experiment which should indicate variations in protein structure. All electrophoreses were carried out in a Gelman Electrophoresis cell on Gelman Sepraphore III strips (1 in. X 6.75 in.). The buffer used was 0.1 M $KHPO_4$ at pH 6.0. One milliliter samples of each enzyme were applied with a micropipette to the strip, which was presoaked in the 0.1 M $KHPO_4$ buffer (pH 6.0). The electrophoresis was carried out in the refrigerator at a temperature of $12^{o}C$. for one hour at 250 volts. In the first run of samples, each enzyme was kept separate; but in later experiments two enzymes that appeared to give different migration rates were combined as a single sample. After electrophoresis, the strips were placed on umbiliferone (.15 mg/50 ml) soaked strips of Watman #1 filter paper cut to the size of the Gelman strips to detect the migration patterns. The umbiliferone provided a substrate for aryl-B-glucosidase, which fluoresced when broken down. Thus, an ultraviolet lamp would indicate a light area where there was enzyme.

RESULTS

A comparison of the calculated specific activities which are given in Figure 1, indicated that there were differences in the activities of the enzymes. These differences must be due to structural differences in the proteins that have affected reactivity. This clearly reflects variability in the enzyme from strain to strain. Hopefully, electrophoresis would clarify the cause of variation by showing different rates of migration in an electric field. The conclusion then would be that some change in charge caused different migration rates, and that, in turn, was caused by residue substitution through mutation.

If differences among the enzymes existed, the fact that they migrate at different rates would cause a separation of bands where two enzymes with different rates of migration were mixed. From the first electrophoresis experiment, I selected the enzymes that showed the greatest differences when tested individually. These were 74-A23-1A and 430; 961 and 967; and 434 and 433. The enzymes from these strains were mixed together as listed and again subjected to electrophoresis. Figures 2, 3 and 4 show the results of this experiment. There was no separation of the bands where two were mixed in any case. The only differences were variations in the thickness of the bands of a mixed as compared to a single enzyme. This indicates that (a) the variations were so small that they were not detectable or (b) this method is insufficient to detect differences due to variations in amino acid sequence for this enzyme. A possible explanation for the inadequacy of this method could be the bulky nature of aryl-B-glucosidase (molecular weight 90,000). The bulky molecules would migrate slowly so that differences due to surface charge would not separate to be detected as distinct bands. Since distribution of charge is probably as important as the actual charge-bearing residue at a given site, a given residue could change without altering the electrophoretic pattern (Zuckerkandl, Pauling, 1965). This method would not be suitable for detecting variation in such a case. The conclusion that must be drawn is that electrophoresis is suitable for detecting large differences due to variable surface charge features but cannot serve as a general method where the variation is slight or actually hidden.

The degree of variation would be limited by the enzyme's role in the cell at any rate. This factor will be discussed in more detail later; but at this point I should like to point out that any difference which could be detected by electrophoresis would be unlikely due to the mechanism whereby aryl-B-glucosidase is excreted from the interior of the cell to the mural space, where it is normally found. In studying this mechanism, Eberhart has pointed out that the excretion into the mural space must be based either on distinctive charged groups or on the tertiary or quaternary configuration of the enzyme (Eberhart, 1968). Such a mechanism would limit the degree of variation of amino acids since the excretion depends on certain residue characteristics which must be maintained to carry out the function.

Mahadevan and Eberhart conducted a study similar to the present one in which they compared aryl-B-glucosidase of Neurospora crassa strains collected from different parts of the world. They used the thermal inactivation point as the criterion for comparison. In their findings, they reported that some of the strains showed different rates of inactivation and concluded that the differences reflected allelic variation at the level of the structural gene (Mahadevan, Eberhart, 1964). The differences in this physical feature indicates that there are variations in the aryl-B-glucosidase molecule, just as do the specific activities obtained for this study. However, the differences must be relatively small due to structural and functional limitations, and the types of variation are limited to those which could not be detected by electrophoresis.

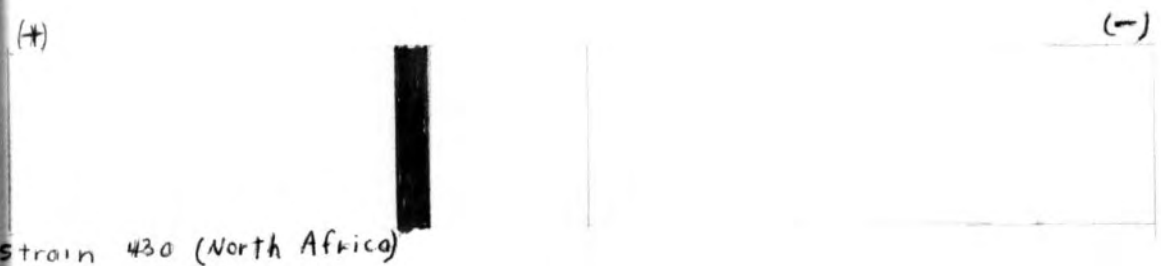
Figure 1

SPECIFIC ACTIVITIES

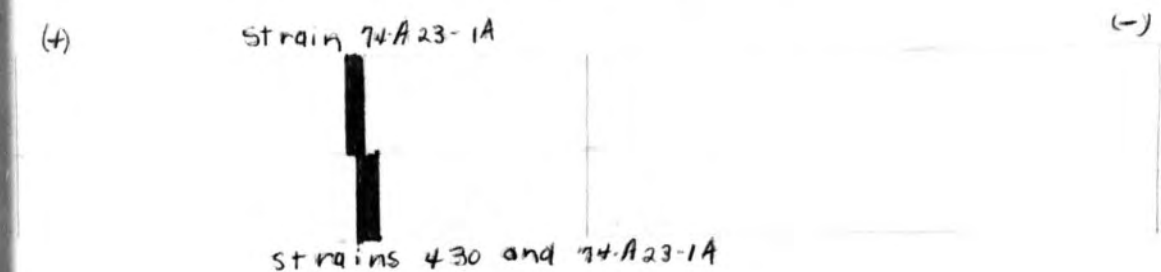
STRAIN	SPECIFIC ACTIVITY
852	2930
967	1000
WT-74A23-1A	2571
435	1667
847	5858
430	12,770
961	4000
851	16,800
434	1765
429	4193
433	4500

Figure 2

Electrophoretic Patterns of Separate Enzymes



Electrophoretic patterns of mixtures of two Enzymes



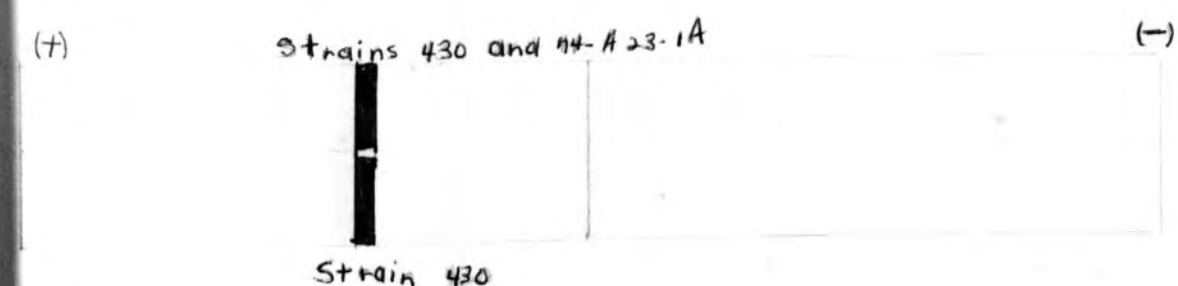
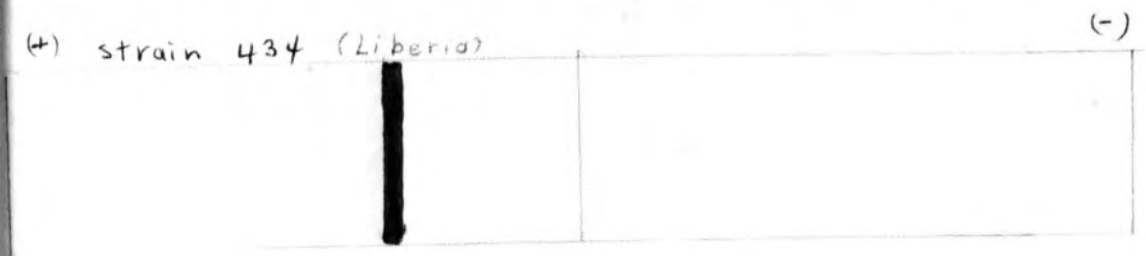
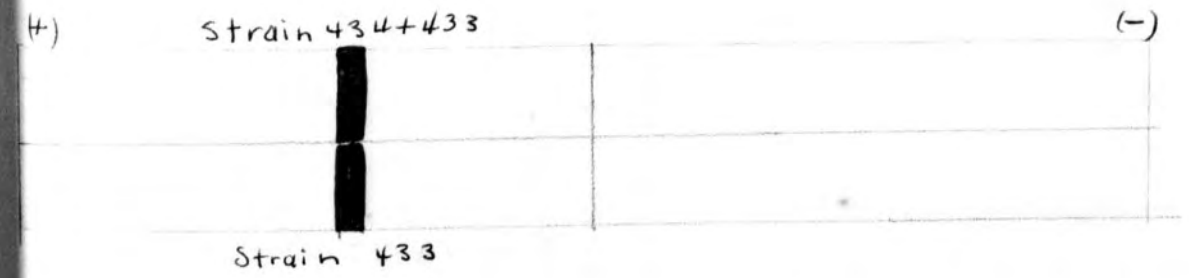


Figure 3

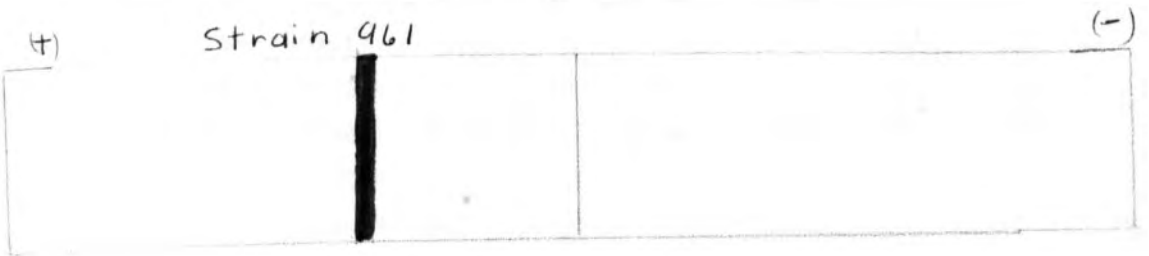
Electrophoretic Patterns of Separate Enzymes

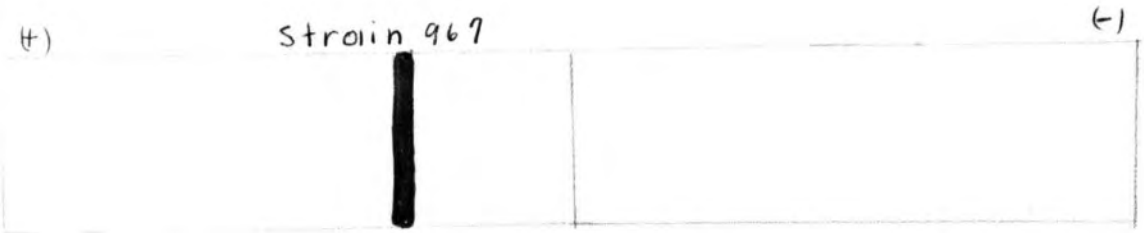


Electrophoretic patterns of mixtures of two Enzymes

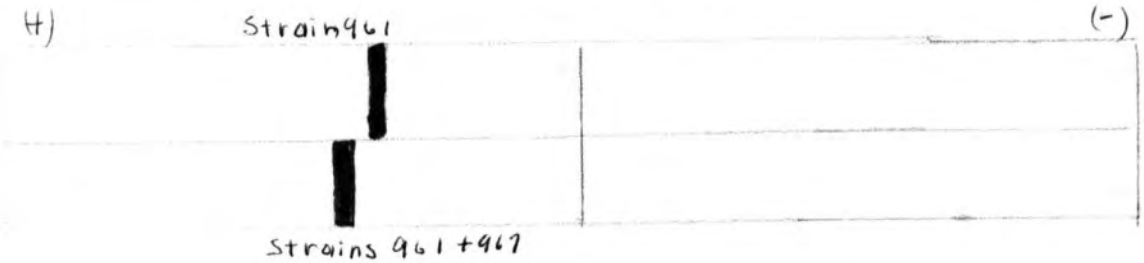


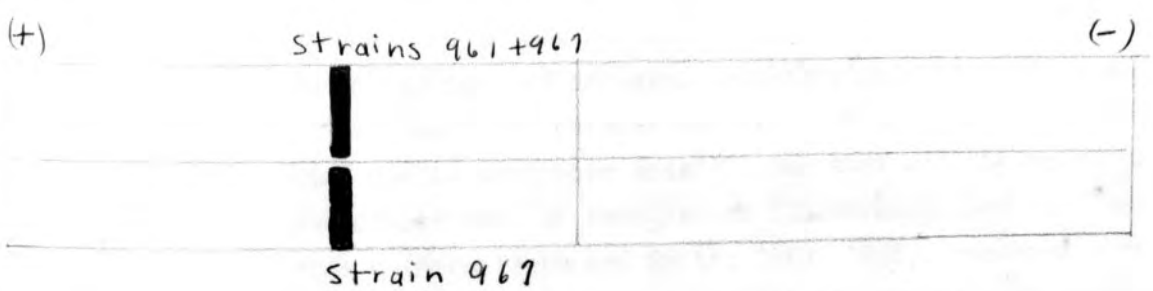
Electrophoretic Patterns of Separate Enzymes





Electrophoretic patterns of mixtures of two Enzymes





DISCUSSION

Although unsatisfactory in this case, electrophoresis has been used as an effective index of protein function. Horowitz had previously used this method to detect the intraspecial polymorphism of the enzyme tyrosinase in Neurospora crassa (Horowitz, Fling, 1961). Horowitz postulated that the variation in the tyrosinase could exist because, as a nonessential enzyme, it is not subject to the rigorous natural selection. Although not essential to the life of the cell, the role of aryl-B-glucosidase is such that it offers selective advantage when functioning. Thus, natural selection would tend to maintain the general function but, at the same time, would not so harshly select against any change. The functions referred to are the inactivation of toxins and transfer activity in synthesis of glucosides (Eberhart, 1968).

This then implies the importance of function in determining the rate at which a protein molecules will actually undergo evolution. If an enzyme performs some vital role upon which the organism depends for life, an alteration which would change the protein's performance in any way should be subject to natural selection. Since a functioning enzyme has been selected for its ability to perform its role and is thus a stable structure, any mutational alteration will usually be detrimental so that selection works against the deviant form. However, for an enzyme which does not perform an essential function, or for which the cell has an alternative mechanism of giving the same results, variations that altered function slightly could occur without killing the organism. The selective forces in such a case would be much slower so that more variation would be permitted.

The remainder of the paper will be concerned with the degree of polymorphism that can occur in homologous proteins and the actual mechanism of evolution at the molecular level. The studies conducted related to this involve comparisons of proteins from strains within the same species as well as comparison between species. These latter studies have been carried out in much more detail. Two that will be referred to often are the comparisons of hemoglobins (Zuckerkandl and Pauling, 1965) and of cytochrome c (Margoliash and Smith, 1963, 1965), isolated from widely separated species. The studies involve the analysis of amino acid

composition so that differences can be noted directly. By such studies, more general principles pertaining to the mode of evolution have been formulated. The primary factors of concern in arriving at principles of molecular evolution have proven to be the invariance of protein structure versus variance. There are limitations to variation which are based on the protein's function. The rigidity of such limitations to the structure will be important since evolutionary divergence can occur no faster than do the molecular components of the organisms.

The actual process of evolution is due to the mutation whereby alterations of DNA give a slightly different form of the protein. Mutation at the level of the gene produces a change which most commonly results in an amino acid substitution at the primary level of the protein. This is the force producing variation in the protein which is consequently responsibly for the actual evolution of the protein. If the alteration in structure produces a protein less suited to its function than the original, the mutant form will be diluted into the available gene pool or eliminated by natural selection.

Hopefully, the comparison of homologous proteins in different organisms will be a more accurate guide to the study of evolution than is the comparative study of more gross features of organisms. The basis for such studies at the molecular level is the assumption that the variable features represent a record of the interplay of genetic change and natural selection on the structural gene of the protein during the evolutionary history of the species studied. Generally, the more closely related species show less variation in the structure of a given protein while those species that diverged early have more variable features in the same protein. This indicates that elapsed time is the one most important criterion in determining differences.

With time as the most important factor in producing protein variations, the next problem is the different rates of evolution for many of the proteins. For example, cytochrome c isolated from species as diverse as yeast to man show some variability but are remarkably similar, with 50%of the amino acids invariable (Margoliash, 1963). On the other hand, tyrosinase in the same species of Neurospora was found to vary in

two different strains, indicating a greater rate of change and degree of variability for this particular protein. The explanation for these differences in degree of variability lies in a consideration of the function of the protein. Where proteins such as cytochrome c perform an indispensable function in all the species in which it is found, natural selection works to maintain this function, selecting against variations in the portions which participate in the protein function. Anfinsen has pointed out that individuals completely lacking the protein serum albumin were only slightly abnormal (Anfinsen, p. 219, 1959). Horowitz' tyrosinase exemplifies another nonessential protein that varied within the species Neurospora crassa. The protein can vary more in these cases because the utmost efficiency in performance of function is not necessary to the organism, and thus, alterations that would affect function would not be lethal.

There is a definite relationship between the protein's role in the overall metabolism and the amount of variability which can occur. The protein's function must be maintained if it is vital to the organism. This limits the degree of variability since those amino acids that are necessary for the protein's action must be retained. The forces of evolution and the livelihood of the organism seem to be contradictory in this sense. On the one hand, evolution would alter a protein while, on the other, the protein must perform the same function. There is, however, the possibility that the two forces would not oppose each other if evolution tended to replace amino acids by residues with similar chemical and bonding properties. Such residues would be more likely to interact in the protein molecule like its predecessor did since they were similar to begin with. Margoliash believes that mutations are not completely random. He gives examples of "conservative" mutation in cytochrome c which act as a built-in mechanism of genetic variability. As indirect evidence for such a mechanism, he pointed out that all the RNA base triplets that code for hydrophobic acids contain uridylic acid (Margoliash, 1963). Further argument for this is given by Freese, who emphasized that each amino acid can usually be represented by three other amino acids where base pair transitions occur, and this is limited

in some cases where transitions occur in only one direction (e.g., from GC--AT) (Bryson and Vogel, p. 349, 1965).

In addition to functional stability, the structure of the molecule must be maintained for the protein to remain effective. Due to the structural characteristics of the protein, there is more room for amino acid substitution at the primary level than at the tertiary level. Zuckerkandl and Pauling have outlined a systematic guide to some qualifications to variability imposed by the protein's physical structure. They have listed areas where natural selection would oppose change as follows:

1. inside the molecule if the substitution produced an increase in bulk (where there would be no space) or polarity (which would destabilize the structure);

2. outside the molecules if there were a notable change in polarity (which would change the molecule's solubility);

3. at bends if substitutions destabilize them;

4. at helices, where substitution would destabilize them;

5. at sites of high conformational specificity; these would include active sites, sites where prosthetic groups, and cofactors combine, at allosteric sites, and at sites where binding to other polypeptides occurs. (Bryson and Vogel, p. 112, 1965).

Such factors were considered in the discussion of aryl-B-glucosidase when the requirements for a certain charge distribution and configuration were referred to as necessary for moving the enzyme into the mural space.

The restrictions on structure are numerous, yet much variation does occur. In Zuckerkandl's comparison of hemoglobin's α chain and the whole myoglobin chain, he found a great deal of similarity in tertiary structure and function, but there were 107 differences in amino acids; this represented variation at 76% of the total number of sites. The explanation put forth for the mechanism by which variability could be so great was related to the interation between two residues rather than the necessity of a single residue at some site (Bryson and Vogel, p. 122, 1965).

In comparing the globin molecules Zuckerkandl and Pauling were able

to make several interesting generalizations about the types of substitutions that existed. The types of sites they listed in considering the nature of the substitutions were (1) charged,(2) acidic, (3) basic, (4) hydrogen bond forming (serine, threonine, asparagine, and glutamine), (5) serine and threonine alone, (6) alanine and glycine, (7) apolar with bulky side chains (valine, leucine, isoleucine, phenyl-alanine), plus methionine, and (8) uncharged. They found that the residues of each type appear together at a given site on homologous proteins. The most significant result was the finding of the apolar bulky residue alone at twenty-six sites. This led to the conclusion that apolar bonding was the most specifically determined process of the globin molecule since so many sites specialized in binding to apolar bonding sites. Charged sites occurred at over 60% of the molecular sites with the great majority of them located on the outer surface. Combining this information with their finding that polar sites were more variable than apolar sites, Zuckerkandl and Pauling concluded that the outside of the molecule was more variable than the inside. This sounds reasonable when one recalls the restrictions on tertiary configuration previously listed. According to this analysis of variation, the most significant aspect of polarity is the distribution of the charge over the surface rather than the presence of a charge at a particular site (Bryson and Vogel, p. 126, 1965). Thus, much variability can occur at this level as long as the original relationships are maintained.

Applying this to aryl-B-glucosidase, the fact that no differences due to surface charge were detected and that surface charge is believed to be important in the mechanism whereby aryl-B-glucosidase is transported across the cell membrane would indicate that the variations within the molecule are small, although some do exist.

This then puts more emphasis on the chemical nature of the amino acids as opposed to the older concept of amino acids working like a puzzle piece to bring about a chemical reaction. Thus, amino acids similar in chemical properties could be interchanged without a change in the original function. These would constitute conservative substitutions.

In observing the mechanism whereby a chemically similar residue

replaced some amino acid, Zuckerkandl and Pauling noted that the most frequent substitutions were those that resulted in some change in the residue properties while preserving other properties. This then could be important in understanding how evolution actually occurs at the molecular level. They postulate further that the basis for extensive variation without radical change in tertiary structure and protein function is the fact that each residue has several important functional properties; the set of amino acids coded is thus selected so that one or more of these functions can change while the others remain constant.

This mechanism of a not completely random mode of genetic mutation would then allow variability while maintaining the basic interactions of the molecule. Such a method of variation in the process of evolution is further indicated in the large percentage of conservative substitutions observed by Zuckerkandl and Pauling for globin and by Margoliash for cytochrome c. Types of substitutions based on the frequency with which a given substitution appears are listed by Zuckerkandl and Pauling as (1) very conservative (these occur at 40% or more of the sites of at least one of the members of any couple); (2) fairly conservative (these appear at more than 25% of the sites of at least one member in each couple); and finally, (3) the radical substitution (these occur at any percentage less than 25 of the sites) (Bryson and Vogel, p. 129, 1965). Since the set of amino acid residues coded for by the genetic code possesses overlapping properties, there can be a large degree of variability in the amino acid sequence without changing the characteristic tertiary structure and function of the molecule. Zuckerkandl and Pauling cite as an example the appearance of a new trait associated with charge, while the former functions -- perhaps apolar interactions --are maintained. They continue to propose that this simultaneous appearance of conservation and nonconservation with a single substitution is one of the basic conditions for protein evolution. The chemical similarities between given amino acids allow for some variation in the actual residues present while the function remains unchanged. "The possibility of extensive variations of the primary structure within the limits of a given function probably provides the richness in combinatory resources that is necessary for

making mutations with radical structural effects sometimes successful in relation to novel function" (Bryson and Vogel, p. 134, 1965).

The limits imposed upon protein structure are of primary importance since they occur at areas which must not change throughout the evolutionary process.

Those areas on a protein which do not vary must define the actual properties of the functioning enzyme. Thus, recognition of residues which remain constant from species to species provides a significant clue as to the actual mechanism by which the protein functions.

The invariance of protein structure then refers to the stability of a given molecular function in different species. Mutations at the level of the structural gene are responsible for the differences in molecular structure of an enzyme common to different species which are believed to be of the same origin. An enzyme such as cytochrome c can be present in a wide range of species, perform the same function in all, and be very similar in structure in the different species. Yet there will be some differences in structures of proteins for different species and even within one species. The similarity suggests common evolutionary origin while the dissimilarity is indicative of the means by which species have diverged. To maintain the same function throughout the many species in which it is present, an enzyme has to possess the same structure at points which determine its activity. At active sites, which combine with the enzyme in the reaction, the amino acid sequence must be invariable; that is, at identical positions there will be corresponding identical amino acids for an enzyme common to different strains and species. Zuckerkandl and Pauling reported eleven invariant sites, representing 8% of the total number of sites in a comparison of sperm whole myoglobin, human myoglobin, human hemoglobin B and α chains, horse α and B chains, cattle B and α chains, cattle fetal chains, gorilla α and B chains, carp α chain and a lamprey hemoglobin chain. Cytochrome c isolated from species as diverse as yeast to man show 50% invariance in structure (Margoliash, 1963). This high degree of stability emphasizes other factors in considering why natural selection would keep a given amino acid at a given site over millions of years. The structural requirements of the

enzyme limit variation in this case; cytochrome c, is located in the mitochrondria where it binds with the structural protein, respiratory enzyme, cofactors, and ATP. Margoliash and Smith, for example, have examined the area where the cytochrome protein is bound to a prosthetic group by covalent thioester links between sulfhydral groups of cysteine residues. In addition, this area contains at least one of the side chains that binds with the heme iron, and there is a site of electron transfer in the oxidative process. In this functionally significant area, nine out of twenty-three amino acids have been found to vary. The knowledge of which sites do vary discloses the minimal functionally necessary amino acids.

Added to this type of restriction are the restrictions to variation imposed when the polypeptides occur in combination with other subunits. In the case of cytochrome c, which reacts with a macromolecule, that in turn reacts with other molecules, a large degree of invariability would be reasonable. The association with the structural protein also limits variation so that the surface molecules that reversibly interact with soluble proteins should be highly invariant. The logical conclusion is that the active sites of molecules interacting with the cytochrome would also be constant in structure. There will be different rates of evolution for polypeptides forming subunits of different versions of the same protein in the quaternary structure and polypeptides forming only one form of the protein. The structural requirements of the former are much more rigid since they have to fit with several different polypeptides, whereas, the polypeptides of the latter adapt to only one form. Since like chains will also interact over certain surfaces, Zuckerkandl and Pauling generalize this to say that polypeptides evolve more rapidly when they exist as monomers than when in combination with other polypeptides in the quaternary structure (Bryson and Vogel, 1965). Here then is another factor which would limit the variability of aryl-B-glucosidase, which does possess quaternary structure.

A more obvious structural requirement is the necessity of maintaining the actual structure of the protein molecules. An example of such a

requirement has been cited by Zuckerkandl and Pauling where glycine at a given site is invariant (Bryson and Vogel, 1965). It has been shown that this glycyl residue interacts with another glycyl at another position which forms a crossing point between the two helices. This suggests that the invariance is necessary to maintain the structure since no other residue permits a contact short enough to connect the two helical segments. Alanine has been found to replace glycine at both sites with the maintenance of the bond. This substitution is still limited in that it had to be by amino acids with the required bonding properties. This represents a conservative substitution in that residues with similar characteristics could replace the original amino acid.

A third type of invariance is a constancy which maintains the overall physiochemical character of segments of the protein molecule such as areas of hydrophobic or basic character (Anfinsen, 1959). Such limitations would be over a regional functional unit of the protein molecule.

CONCLUSION

In cases where some substitutions can occur, but those that are successful are limited to amino acids of a certain chemical or bonding nature, the variation is referred to as a conservative substitution. Most mutations leading to a residue alteration that are preserved by natural selection introduce a substitution of the conservative type. Zuckerkandl and Pauling state that the most active stage of evolutionary change is in the early history of the enzyme before its stability is established (Bryson and Vogel, 1965). As the enzyme becomes established in function, natural selection works to maintain the function so that only the conservative substitution appears.

The one factor that determines the amount of variation for proteins of different organisms is time. However, the degree of change that will be effective in the evolutionary process is determined by the protein's function in the cell and its interrelations with other structures. Aryl-B-glucosidase of different strains of Neurospora crassa was seen to vary in some ways which give different specific activities, but the limits imposed by function and the mechanism of putting it where it belongs probably prevented detectable variations from occurring.

Within the framework of maintaining the functional and structural integrity of enzyme, conservation substitution can occur where the new amino acid meets the needs of that residue position. An effective substitution by a closely related amino acid could perform the same functions as the initial residue did, and at the same time introduce some new features to the protein. With evolution working in a way that tends to produce residues of similar properties, the process of organic evolution would move toward successful variations in a less random way.

BIBLIOGRAPHY

Ainsworth, G. C. and Alfred S. Sussman. 1968. The Fungi-An Advanced Treatise. New York and London, Academic Press.

Anfinsen, Christian B. 1959. The Molecular Basis of Evolution. New York, Science Editions, John Wiley & Sons, Inc.

Bryson, Vernon and Henry J. Vogel. 1965. Evolving Genes and Proteins: A Symposium. New York and London, Academic Press.

Crow, James. 1961. "Population Genetics," American Journal of Human Genetics, 13, 137-155.

Dobzhansky, Theodosius. 1963. Evolution, Genetics and Man. New York, Science Editions, John Wiley & Sons, Inc.

Dobzhansky, Theodosius. 1961. Genetics and the Origin of Species. New York, Columbia University Press.

Eberhart, Bruce M. 1968. The Biochemical Genetics of the B-Glucosidases of Neurospora. Research Proposal to the National Science Foundation.

Ford, E. B. 1965. Genetic Polymorphism. All Souls Studies V. Massachusetts, The M.I.T. Press.

Horowitz, N. H. and Marguerite Fling. 1953. "Genetic Determination of Tyrosinase Thermostability in Neurospora," Genetics, 38, 360-374.

Horowitz, N. H., Marguerite Fling, Helen MacLeod, and Noboru Sueka. 1961. "A Genetic Study of Two New Structural Forms of Tyrosinase in Neurospora," Genetics, 46, 1016-1024.

Horowitz, N.H. 1945. "On the Evolution of Biochemical Syntheses," Proceedings of the National Academy of Science 31, 153-157.

Horowitz, N. H. 1957. "The Genetic Modification of Enzyme Structure," Supplement Volume of Cytological Proceedings of the International Genetics Symposa 623-626.

Hubby, J. L. and R. C. Lewonten. 1965. "A Molecular Approach to the Study of Genetic Heterozygosity in Natural Populations. I. The Number of Alleles at Different Loci in Drosophila pseudoobscura," Genetics, 54, 595-609.

Jukes, Thomas H. 1966. Molecules and Evolution. Columbia Univeristy Press, New York and London.

Lewonten, R. C. and J. L. Hubby. 1966. "A Molecular Approach to the Study of Genic Heterozygosity in Natural Populations. II. Amount of Variation and Degree of Heterozygosity in Natural Populations of Drosophila pseudoobscura," Genetics.

Mahadevan, P. R. and B. Eberhart. 1964. "Aryl-B-Glucosidase of some Neurospora Strains." Biochim. Biophys. Acta 90, 214-215.

Margoliash, E. 1963. "Primary Structure and Evolution of Cytochrome c", National Academy of Science Proceedings U. S., 50, 672-679.

Mayr, E. 1963. Animal Species and Evolution. Cambridge, Mass., Belknap Press of Harvard University Press.

Neurath, Hans; Kenneth A. Walsh; William P. Winter. 1967. "Evolution of Structure and Function of Proteoses," Science, 158, 1638-1644.

Sibley, Charles G. 1962. "The Comparative Morphology of Protein Molecules as Data for Classification," Systematic Zoology, II.

Stebbins, G. Ledyard. 1966. Processes of Organic Evolution. Englewood Cliff, New Jersey, Prentice-Hall, Inc.

Wallace, Bruce. 1966. Chromosomes, Giant Molecules and Evolution. New York, W. W. Norton and Co., Inc.

Wright, Sewall. 1968. Genetics and the Biometric Foundations, Volume I. Chicago and London, University of Chicago Press.

Zamenhof, Stephen and Herbert H. Eichhorn. 1967. "Study of Microbial Evolution Through Loss of Biosynthetic Functions: Establishment of Defective Mutants," Nature, 216, 456-458.