Molecular Disease, Evolution, and Genic Heterogeneity[1]

EMILE ZUCKERKANDL[2] AND LINUS PAULING

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California

I. Generalities

A. The Notion of Molecular Disease and Its Field of Application

Life is a relationship between molecules, not a property of any one molecule. So is therefore disease, which endangers life. While there are molecular diseases, there are no diseased molecules. At the level of the molecules we find only variations in structure and physicochemical properties. Likewise, at that level we rarely detect any criterion by virtue of which to place a given molecule "higher" or "lower" on the evolutionary scale. Human hemoglobin, although different to some extent from that of the horse (Braunitzer and Matsuda, 1961), appears in no way more highly organized. Molecular disease and evolution are realities belonging to superior levels

[1] Contribution No. 2774 from the Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California.

[2] On leave from the Centre National de la Recherche Scientifique, Paris.

of biological integration. There they are found to be closely linked, with no sharp borderline between them. The mechanism of molecular disease represents one element of the mechanism of evolution. Even subjectively the two phenomena of disease and evolution may at times lead to identical experiences. The appearance of the concept of good and evil, interpreted by man as his painful expulsion from Paradise, was probably a molecular disease that turned out to be evolution. Subjectively, to evolve must most often have amounted to suffering from a disease. And these diseases were of course molecular.

Relationships between molecules, which define states of health and disease, may be altered by environmental factors, or by factors of aging, or by inherited internal factors. The two last types of factors are partly the same, inasmuch as aging is itself determined by genetic factors. The two first types of factors are also partly the same, inasmuch as aging is due to the cumulative effect of external agents. The term molecular disease in its more restricted sense, the only useful one, relates to the third type of factors, to altered relationships between molecules traceable to altered genes.

To the extent to which we have grounds to believe today that inheritance is linked to nucleic acids and that the primary products of nucleic acids, beside perhaps other nucleic acids, are proteins and no other types of molecules, the notion of molecular disease relates exclusively to the inheritance of altered protein and nucleic acid molecules. Abnormal glycogens, for instance (Cori, 1954), are to be traced to abnormal enzymes, proteins responsible for their production.

The abnormal nucleic acid and the abnormal protein produced under its control represent two aspects of the same reality. A great many more protein molecules are present at a given time than nucleic acid molecules responsible for their production, and this is one of the reasons why it is much easier at the moment to study the phenomenon of molecular disease from the protein end.

An abnormal protein causing molecular disease has abnormal enzymatic or other physicochemical properties. Changes in such properties are necessarily linked to changes in structure. It seems unwarranted at the present time to draw a basic distinction between two types of structural changes as causes of molecular disease, change in folding of the polypeptide chains and changes in the sequence of amino acids, the building stones of the chains. It becomes increasingly probable that the changes in spatial configuration (conformation) are direct expressions of changes in sequence (Crick, 1958; Mizushima and Shimanouchi, 1961). Therefore a molecular disease can probably be defined at the molecular level in a way that is potentially complete by determining the alteration of the amino-acid sequence of a protein (or the nucleotide sequence in the corresponding nucleic acid). This statement applies of course to a given intracellular and extracellular environment. If there are changes in the environment, the spatial configuration of a protein may be altered without any change in amino-acid sequence and a pathological condition may ensue. Such a change in environment results either from external influences, which do not concern us here, or ultimately from the change in amino-acid sequence in other proteins, to which the molecular disease must then be traced.

B. Molecular Disease, Evolution, and Environment

The study of molecular diseases leads back to the study of mutations, most of which are known to be detrimental. All loss mutations in a broad sense of the word—involving either the total loss of a protein or the loss of protein function through a structural alteration of the protein—are molecular diseases. Loss mutations, on the other hand, are among the conditions of adaptation of the organism to changes in its environment and adaptation, the conditions of evolution. A loss of function, when compatible with survival thanks to the nature of the environment, may make cellular energy and genic raw materials available for the acquisition of new functions. More highly evolved organisms have lost powers of synthesis that more primitive organisms possess (Lwoff, 1943). It thus appears possible that there would be no evolution without molecular disease. A maintenance of molecular health, although in the interest of the individual, is opposed to evolution. However, only a small fraction of the molecular diseases that occur are used by and turned into evolution.

A bacterium that loses by mutation the ability to synthesize a given enzyme has a molecular disease. The first heterotrophic organisms suffered from molecular diseases, of which they cured themselves by feeding on their fellow creatures. At the limit, life itself is a molecular disease, which it overcomes temporarily by depending on its environment. Every vitamin we need today bears testimony to a molecular disease our ancestors contracted sometimes hundreds of millions of years ago. These molecular diseases are not experienced as such under normal circumstances, because our environment con-

stantly supplies palliative drugs. Conversely, if phenylalanine happened to be present only in low amounts in our usual diet, the mutation leading to phenylketonuria, characterized by the inability to convert toxic amounts of phenylalanine into tyrosine, would also not be experienced as a molecular disease, whereas it actually is one under the prevailing circumstances. We might say that evolution is based in part on the appearance of molecular diseases whereof the environment can cure the symptoms. Since our remote ancestors must have been autotrophic, we may consider ourselves as degenerate autotrophic organisms. Whereas, in order to achieve superiority, it is not sufficient to be degenerate, it is however necessary.

In many cases the notion of molecular disease is thus closely linked to the nature of the environment. It is not so in other cases, such as a structural change of the hemoglobin molecule leading to the loss of its ability to combine reversibly with oxygen. Both types of cases are similar in that chemicals available in the environment either cannot be used (oxygen or vitamin precursors from which vitamins are built) or cannot be disposed of (phenylalanine). The difference between the two types of molecular diseases resides in the fact that in one case, that of vitamin need and of phenylketonuria, the environment can make up for the lost biochemical reaction either by furnishing its product (the vitamin) or by ceasing to furnish its starting material (phenylalanine), and in the other case it is not that products are needed or that toxic substances must be excluded, but the process itself of making the product is essential to the organism. Thus when oxygen cannot be carried to the tissues efficiently it would be of no avail to furnish the tissues with oxidation products. The oxidation must be carried out by the organism itself, mainly because living matter requires chemical energy to be set free at the right time in the right place. Thus molecular diseases are defined in relation to the environment when the requirement involved is that of a substance or of less specific forms of energy such as heat, and they are not so defined when the requirement involved is that of a process fundamental to the existence of living matter, that is, of a high degree of specificity in the release of energy in relation to time and space. Life can get everything out of the environment except a degree of specificity approaching its own.

Considering molecular disease and the environment in relation to evolution, we are faced with a two-way relationship. Evolution has probably been influenced not only by heritable changes in the

organism that the environment could prevent from being deleterious, but also by changes in the environment that molecular changes in the organism could prevent from being deleterious. Molecular disease can be selected for as a defense against diseases caused by external agents. For example, the past incidence of malaria has been shown to be positively correlated with the abnormal hemoglobin, HbS, present in sickle-cell anemia; with thalassemia, another type of genetically controlled hemoglobin disease; with glucose-6-phosphate dehydrogenase deficiency; and with color blindness (Allison, 1957; Siniscalco *et al.*, 1961). Apparently the presence in the environment of the agent of a highly dangerous disease, *Plasmodium falciparum*, favors the conservation and spreading in the human species of molecular diseases that afford protection against the infectious agent by unknown mechanisms. The molecular diseases, at least in the heterozygous ("trait") condition, are less lethal than the infectious disease. Observations of this kind extend the interaction pattern between molecular disease and environment. A "molecular disease" may be maintained in the species because certain agents in the environment render it innocuous; or it may, on the contrary, be maintained because it renders relatively innocuous certain agents present in the environment. On account of this latter effect it seems possible that external disease-causing agents, notably infectious agents, have played a role in evolutionary sequences of noncompensated degenerative nature, such as those leading to parasitism.

The sickle-cell gene increases the life expectancy of the individual in the heterozygous state, while in the homozygous state it decreases it probably at least as radically as does malaria. When two carriers of sickle-cell trait marry, on the average half of their offspring will again be heterozygotes. The other half will have a decreased life expectancy, because of either sickle-cell disease (sickle-cell gene homozygotes) or malaria ("wild-type" homozygotes). In malaria-infested countries the sickle-cell gene will thus have a tendency to spread in the population. This would hardly be the case if the mutant gene were advantageous in the homozygous instead of the heterozygous condition. A newly appearing mutation that would be retained only in the homozygote would usually have no chance of establishing itself in the population. We may point out that the replacement in a population of a given gene by a mutant gene may often require two successive mutations, except when the population is very small (close inbreeding). After the first mutation the mutant is selected for in

the heterozygous state. This allows the mutant gene to establish itself, but at the same time the corresponding wild-type gene is preserved. To eliminate the wild-type gene a second mutation must now occur, such that the doubly mutated gene is most advantageous in the homozygous state. This double mutant would not cause what would appear as a molecular disease, while the single mutant that precedes it might. In this sense molecular disease may be a frequent intermediary step in those evolutionary sequences that lead to the total replacement of a gene by a mutant allele, and that require that the heterozygous condition be at first advantageous. During this phase the homozygous condition may indeed often be deleterious.

II. Hemoglobin, Its Multiplicity and Evolution

A. Introduction to the Hemoglobin Molecule

If one examines molecular disease in relation to evolution it is unavoidable at the present time to center the discussion on the hemoglobin molecule. So far this molecule is the only one that has been studied in many pertinent respects: amino-acid sequence, structure of the site directly involved in function, structural changes leading to molecular disease, normal structural multiplicity, different rates of synthesis of structurally distinct "editions" of the molecule, and their change in the course of time.

All vertebrates save the most primitive ones seem to have hemoglobins composed of four polypeptide chains, linked to each other by bonds much weaker than the peptide bonds that are instrumental in lining the amino acids up unidimensionally within the chains. Each of these chains is composed of slightly less than 150 amino acids and carries one heme group that contains the iron atom capable of binding oxygen reversibly. The string of amino acids winds about in space in a highly specific fashion that may be common to all vertebrate oxygen-carrying pigments, since even sperm whale myoglobin (muscle hemoglobin) shows a similar conformation in spite of the fact that its amino-acid sequence differs very considerably from the sequences so far found in blood hemoglobins (Watson and Kendrew, 1961). Nature has produced a great many such hemoglobin and myoglobin chains that differ in their amino-acid sequence and therefore in various physicochemical properties and yet apparently remain similar in their over-all conformation and in the fundamental characteristics of their relation with the heme group. Not only does the amino-acid sequence of these chains always vary in different animal species, except perhaps in some extremely closely related ones, but any one individual of any given species produces a number of different hemoglobin chains, in part successively and in part simultaneously.

So far the tetrahemic hemoglobin molecule of higher vertebrates has always been found to be normally made up of two kinds of chains that combine two by two. Thus human adult hemoglobin, HbA, contains two so-called α- and two so-called β-chains (Rhinesmith *et al.*, 1957, 1958). This type of hemoglobin is predominant only from the time of birth on, while during intrauterine life by far the greatest proportion of the hemoglobins produced is represented by fetal hemoglobin, HbF, composed of two α- and two γ-chains (Schroeder and Matsuda, 1958; Hunt, 1959, Schroeder *et al.*, 1959b, Shelton and Schroeder, 1960). There exists another human hemoglobin chain during postnatal life, but normally never in more than small amounts, the δ-chain. Two α-chains combine with two δ-chains to form the minor component known as HbA_2 (Kunkel and Wallenius, 1955; Kunkel *et al.*, 1957; Ingram and Stretton, 1961). This is as far as the list of structurally distinct *normal* human hemoglobin chains goes at present. Occasionally, in the diseases called the thalassemias, one or two of the chains are present in subnormal amounts. Sometimes a relative excess of the partner chain is produced, which then associates with its own kind to form tetramers. This leads to the formation of abnormal hemoglobin such as HbH, composed of four β-chains (Jones *et al.*, 1959), or Hb "Bart's," composed of four γ-chains (Lehmann, 1959). Thus the association of two different types of chains is not an absolute requirement for the formation of chain tetramers but probably a matter of preferential affinity. All the known normal human hemoglobins have one type of chain in common, the α-chain, but in other species this is not necessarily so, as has been shown for chicken (C. J. Muller, 1961).

B. Hemoglobin Heterogeneity

It is remarkable that hemoglobin-chain heterogeneity has been found in all species. So far only vertebrates have been examined, but of widely different classes, ranging from mammals to fish and Cyclostomes (reviewed by Gratzer and Allison, 1960; also Huisman *et al.*, 1960, and unpublished results from this laboratory). The

Cyclostomes belong to the most primitive group of vertebrates, whose living representatives, hagfish and lampreys, seem to have hemoglobins composed of single, unassociated polypeptide chains (Svedberg, 1933; Roche and Fontaine, 1940; Lenhert *et al.*, 1956), presumably of the same general type as those which in higher forms associate into tetramers. In this respect Cyclostome hemoglobins resemble the myoglobins. Even though lamprey hemoglobin chains do not normally associate into higher molecular units, these animals also possess several distinct types of chains (Andinolfi *et al.*, 1959). Apparently the multiplicity of hemoglobin chains is not an evolutionary consequence of their association, but their association is an evolutionary consequence of their multiplicity. This conclusion is confirmed by the observation that there exist also several distinct types of myoglobins in all individuals (Rossi-Fanelli and Antonini, 1956; Rumen, 1959; Rossi-Fanelli *et al.*, 1960; Edmundson and Hirs, 1961), while myoglobin polypeptide chains usually do not associate to yield molecular units of a higher order. [Evidence of the presence of dimers in solutions of some invertebrate myoglobins has recently been reported by Manwell (1958b, 1960)]. It appears justified at the present time to extrapolate from these and other observations to polypeptides and proteins in general and to state that proteins and polypeptides of all kinds may usually be expected to coexist within every individual in structurally distinct "editions." Several of these may be synthesized in the same cells, as fetal and adult hemoglobins often are (Kleihauer *et al.*, 1957; Itano, 1956); or they may be produced in different tissues. This generalization rests now on a number of investigations, among which are those of Markert and Møller (1959) and of Kaplan *et al.* (1960). The latter authors showed that in vertebrates as well as invertebrates lactic dehydrogenases extracted from different tissues of one animal are different from one another. By analogy with hemoglobin we may suppose that for most kinds of proteins there will be found in every organism, in addition to major components that succeed each other in time or coexist in different tissues, structurally distinct minor components. The importance of these minor components is probably negligible from the point of view of function, but not of evolution. Molecular diseases will of course relate to the quantitatively important "major" components only. From what precedes it will be recognized that the greater the role of minor components in evolution, the smaller that of molecular disease.

If this picture is correct and it probably is, the number of distinct proteins in the human organism that has been estimated by one of us (L.P.) to be of the order of 100,000 will have to be multiplied by a factor of presently unknown magnitude.

It must be pointed out that a number of minor hemoglobin components apparently do not differ in amino-acid sequence from one of the major ones (Jones, 1961). Some arise apparently through a secondary combination with some other molecule, such as glutathione (C. J. Muller, 1961). Others may be oxidation or denaturation products or chromatographic artifacts, or unusual combinations of hemoglobin chain dimers or monomers, or they may be hemoglobin polymers. Finally, it remains probable, though the contention still awaits experimental confirmation, that some minor components might have an altered amino-acid sequence without being produced under the control of an altered gene. Such components would express "errors" in the synthesis of normal hemoglobin chains (Pauling, 1957a).

If we consider structurally different hemoglobin chains found in one given species, man, we may divide them into two groups. The chains in one group differ from each other by more than one amino-acid substitution. Thus the number of changes in sequence, when the α-, β-, γ-, and δ-chains are compared with each other, varies from about 6 to a little less than 80 (computed from Braunitzer *et al.*, 1960a,b; G. Braunitzer, 1961, personal communication; Konigsberg *et al.*, 1961; Schroeder *et al.*, 1961; W. A. Schroeder and R. Shelton, personal communication; Ingram and Stretton, 1961). These chains which show marked differences have all been found to be controlled by distinct genetic loci (Itano, 1957; Smith and Thorbert, 1958; Cepellini, 1959a,b, Ingram and Stretton, 1961), and they are present in all normal human individuals. In the second group we find chains that differ from one of the others by only one amino-acid substitution. In all cases that have been examined, the gene that controls such a chain has been found to be an allele of the gene in control of the nearly identical chain. Each of these different alleles occurs only in small proportions of the population in different areas of the world. They are the abnormal hemoglobins. While some go unnoticed by their carriers, others lead to characterized molecular diseases in the homozygous condition (see for instance Itano and Pauling, 1957; Itano, 1957; Neel, 1959; Ingram, 1961a). Of course chains differing by more than one amino-acid substitution and controlled by allelic genes may be discovered, but they will presumably remain a small

minority, and, for reasons that will become clear presently, we may expect that chains differing by more than a very small number of changes in amino-acid sequence or, more accurately, chains that have been affected by more than a very small number of mutational events will generally be traceable to distinct genetic loci.

On the other hand, when we compare hemoglobin chains from different species, chains controlled by corresponding genetic loci may differ considerably in amino-acid sequence. This is of course only a presumption, since we have no means of matching genetic loci of different species. For instance, there seem to be two differences between the human and gorilla α-chains (Zuckerkandl and Schroeder, 1961), yet there is no reason to suppose that the genic loci controlling their production are not homologous.

We may venture the following generalization. While in different species markedly different hemoglobin chains may conceivably be and probably quite often are controlled by homologous loci (by genes that would be shown to be allelic if fertile crosses between the species were possible), within one species a greater difference between chains is associated with greater independence in their genetic control. In this respect it is suggestive that the α- and β-chain genes, among the most different within the species, have been shown to be on separate chromosomes, or at least not to be closely linked (Smith and Thorbert, 1958), while the β- and the δ-chain genes, which resemble each other most, appear to be linked (Cepellini, 1959b).

C. The Evolution of the Hemoglobin Chains

The foregoing observations can be understood at once if it is assumed that in the course of time the hemoglobin-chain genes duplicate, that the descendants of the duplicate genes "mutate away" from each other, and that the duplicates eventually become distributed through translocations over different parts of the genome. Different non-allelic genes are thus thought to have arisen from an original mother gene. Since it seems justified to consider effective (i.e., viable) translocations as phenomena that occur more rarely than effective amino-acid substitutions, one would expect that genes related to a common ancestor but not closely linked differ from each other by a number of mutational changes, in accordance with observation. Ideas of this kind have been evolved by Bridges (1935), Metz (1947), and notably Lewis (1951), and have been applied to hemoglobin evolution by Itano (1957), ably developed by Ingram (1961b), and

elaborated quite independently by ourselves in 1960. It seems likely that the *intraspecific* multiplicity of proteins of a given type is to be explained in these terms. Considerable *interspecific* differences between proteins of a certain type may, on the other hand, as stated before, be compatible with homology of genic loci and not require the intervention of gene duplication. As species gradually get to be more different from each other, so presumably do the genes at the homologous loci.

All we can check at present are homologies of chain structure as expressed by correspondences between amino-acid sequences in hemoglobin chains, and such homologies, whether inter- or intraspecific, suggest a common evolutionary origin. An alternate hypothesis would be convergence by selection for functionally adaptive hemoglobin-chain structures. While convergence may play a significant role, this role is most likely confined to a relatively small number of features of amino-acid sequence. The over-all similarity must be an expression of evolutionary history. This is indicated by the gradually increased amount of differences found when human hemoglobin is compared with hemoglobins from progressively more distant species (Zuckerkandl *et al.*, 1960; C. J. Muller, 1961). The difference between human and fish hemoglobins is such that no common features, except the presence of free lysine, could be detected by the comparison of peptide patterns obtained by spreading the products of a tryptic digestion two-dimensionally over filter paper by successive electrophoresis and chromatography. The absence of common features in these patterns in no way implies the absence of significant stretches of similar amino-acid sequences, but nevertheless expresses qualitatively a degree of difference. A comparable result has been obtained by comparing mammalian and fish insulin (Wilson and Dixon, 1961). Insulin, on the whole, seems less variable than hemoglobin, even taking into account its smaller molecular weight.

At the other extreme we may compare human and gorilla adult hemoglobins. From the amino-acid analysis of separate gorilla α- and β-chains it appears that there are only two differences in the α-chain and one in the β-chain. The amino-acid analysis of isolated tryptic peptides from gorilla hemoglobin, two thirds of which has been completed (unpublished), has so far furnished no evidence of further changes. It is therefore possible that the gorilla β-chain and the human normal and abnormal β-chains form one single population. (As mentioned before, the abnormal human chains differ from

the normal ones by only one amino-acid substitution.) Since gorillas get along well with their hemoglobin, as they prove by existing, it is not likely that the gorilla β-chain, if it were present in humans, would cause molecular disease. The required oxygenation properties of hemoglobin must be rather similar in the two species that are otherwise so much alike. Thus, if the gorilla β-chain occurred in a human family the physician's attention would probably not be attracted to it. Moreover, it would probably go unnoticed in general surveys, because the nature of its difference with the human β-chain —probably a substitution of a lysyl for an arginyl residue—seems to be of the kind that current scanning techniques do not detect. Conversely, it is also possible that the human β-chain occurs in some gorillas.

Some of the hemoglobin chains coexisting within one individual differ from each other as much as or more than corresponding chains may be expected to differ in the most distantly related vertebrates. While human β- and γ-chains are only moderately different, they are much more different than gorilla and human β-chains. Therefore with respect to hemoglobin an adult man resembles an adult gorilla much more than his own human embryo. Morphological observation also suggests this relationship, which is now confirmed at the biochemical level. The human α-chain differs much more from the β-, γ-, and δ-chains than the latter from each other. Nevertheless even in the case of the α-chain the remaining similarities with the others are striking. When the sequence of the first 30 amino acids of the β-chain became known from G. Braunitzer's laboratory, our own knowledge of the α-chain was limited to amino-acid sequences in peptides isolated from the chain by tryptic digestion by W. A. Schroeder, R. T. Jones, J. R. Shelton, and their collaborators. The succession of these peptides along the α-chain was unknown. By fitting them into the β-chain according to the principle of maximum homologies, a sequence of the first 31 residues of the α-chain was predicted that was later confirmed by Braunitzer *et al.* (1961). This showed for the first time that the homology principle could be put to work effectively, even between chains that differ considerably.

It is possible to evaluate very roughly and tentatively the time that has elapsed since any two of the hemoglobin chains present in a given species and controlled by non-allelic genes diverged from a common chain ancestor. The figures used in this evaluation are the number of differences between these chains, the number of differences between corresponding chains in different animal species, and the geological age at which the common ancestor of the different species in question may be considered to have lived. Braunitzer and Matsuda (1961) have recently found that there are a minimum of 15 differences in sequence between the horse and human α-chains. (Only one of the two main horse hemoglobin components was analyzed.) This number is not likely to be increased very much by subsequent work. If we estimate that the real number of differences in sequence is between 15 and 20, we may take 18 as a probable mean. From paleontological evidence it may be estimated that the common ancestor of man and horse lived in the Cretaceous or possibly in the Jurassic period, say between 100 and 160 millions of years ago (Piveteau, 1955; Dodson, 1960). For the sake of the calculation it is assumed that most effective mutations result in single amino-acid substitutions, as evidence from abnormal human hemoglobins indicates, and that the evolutionarily effective mutation rate, i.e., the rate of the mutations that have not been eliminated by natural selection, fluctuated during the time of evolution of hemoglobin around a mean without showing a predominant trend to increase or to decrease. Under these conditions the presence of 18 differences between the human and horse α-chains would indicate that each chain averages 9 evolutionarily effective mutations in 100 to 160 millions of years. This yields the figure of 11 to 18 million years per amino-acid substitution in a chain of about 150 amino acids, with a medium figure of 14.5 million years. Our results for the gorilla hemoglobin chains yield somewhat different figures. Because of considerable fluctuations that may be expected in cases where the number of evolutionarily effective mutations has been very small, it seems advisable to use the figure derived from the horse α-chain alone. As the amino-acid sequences of more animal hemoglobin chains become known and paleontological dating is improved, the calculation will have to be revised. Also the number of differences between the human chains is subject to moderate revision, especially the comparisons involving the γ- and δ-chains, based on the results of Schroeder *et al.* (1961), W. A. Schroeder and J. R. Shelton (personal communication, 1961), and Ingram and Stretton (1961).

As Table I shows, the evaluation of the time elapsed since the β- and δ-chains differentiated places their common ancestor at the time of origin of the Primates or somewhat earlier. This checks with the fact that so far δ-chains have been found only in Primates (Kunkel

et al., 1957) and furnishes evidence that, at least with respect to more recent evolution, the present evaluation is not unreasonable. Also the time of derivation of man and gorilla from their common ancestor as calculated on the basis of the figures derived from man and horse, 11 million years, falls on the lower limit of the range estimated on paleontological grounds, 11 to 35 million years.

TABLE I

THE APPROXIMATE TIME OF DERIVATION OF DIFFERENT HEMOGLOBIN CHAINS FROM THEIR COMMON ANCESTOR

Chains being compared	Number of differences[a]	Estimated time of derivation from common chain ancestor	Corresponding geological period
β and δ	~6	44×10^6 years	Eocene
β and γ	~36	260	Beginning of Carboniferous
α and β	78	565	Toward end of Pre-Cambrian
α and γ	~83	600	Toward end of Pre-Cambrian
Gorilla α and human α	2	14.5 } Mean 11	Pliocene
Gorilla β and human β	1	7.3 }	

[a] The presence or absence of one to several contiguous amino-acid residues in one of the chains is counted as one mutational change.

Of course, the uncertainty increases as we go further back in time. The common ancestry of the β- and γ-chains is placed at the beginning of the Carboniferous period, that is, about at the time of the first amphibians. Differences between fetal and adult hemoglobins have however been found also in contemporary fish (Manwell, 1957, 1958a). It is conceivable that these have arisen from a gene duplication independent of the one that led to the differentiation of β- and γ-chains.

The α- and β-chains are so different that the present evaluation places their common chain ancestor in the Pre-Cambrian, before the apparent onset of vertebrate evolution. The differences between the α- and γ-chains check reasonably well with those between the α- and β-chains. If the figures were taken at face value, it would seem that vertebrate hemoglobin with its differentiation into hemoglobin polypeptide chains derives from an invertebrate hemoglobin. Thus the ancestors of the present α- and β-chains would have to have been already present as differentiated chains in those primitive vertebrates in which the hemoglobin chains presumably did not associate into tetramers, as in the contemporary Cyclostomes, mentioned earlier. The finding of hemoglobin heterogeneity in the lamprey (Andinolfi *et al.*, 1959) is suggestive in this respect.

The figures in Table I also make it appear unlikely that corresponding chains, say α-chains and their homologs in animals, when the most distantly related vertebrates are compared, will be found to differ from each other more than human α- and β-chains differ. The figures also strengthen the presumption that hemoglobin has not been evolved independently more than once during vertebrate evolution and suggest, as stated, that even the most primitive among the ancestral vertebrates had already inherited their hemoglobin from other forms.

Polypeptide chains that are clearly not homologous, such as horse-heart cytochrome c (Margoliash and Tuppy, 1960) and mammalian hemoglobin chains, may still have a common molecular ancestor, in the sense in which all protein molecules of a given organism may conceivably have one, but such an ancestor would have existed so far back in Pre-Cambrian times that comparative studies on contemporary organisms have no significant chance of revealing a kinship. For all practical purposes it is therefore correct to say that horse-heart cytochrome c and horse hemoglobin chains have evolved independently.

Our best excuse for making the present evaluation is that it affords us the opportunity to point out why it is probably wrong. The sources of error involved are factors in gene evolution that deserve to be mentioned here.

We do not know whether the present "major" hemoglobin components have once been derived from "minor" components. The contribution of minor components as oxygen carriers is mostly negligible. Unless they have other unknown functions, natural selection will not be expected to act upon them. Thus all mutations will probably be preserved in a minor component until one of the three following possible events occurs: a mutation that makes it unrecognizable as a hemoglobin chain; a mutation that brings about a total inhibition of its synthesis, or a mutational change that transforms it into a major component. If the ancestors of the human hemoglobin chains

that are important quantitatively (the α-, β-, and γ-chains) have started out as minor components, they will during that remote period have retained many more mutations per unit time than we have assumed, and from this point of view the figures given in the table would be overestimates.

They tend to be underestimates for other reasons. In the comparison between the chains, possible back-mutations, of which we have no knowledge, had to be neglected, as well as successive different effective substitutions at the same amino-acid residue. The likelihood of these events increases with the increase in the number of amino acids affected by change in a given chain. Thus the number of effective mutational events that have actually occurred since the α- and β-chains have evolved from their common ancestor may be significantly greater than is presently apparent.

Furthermore, even if we assume the intrinsic mutation rate of the hemoglobin chain genes to have remained fairly constant throughout the geological periods, disregarding the probable effects of changes in temperature and in intensity of ionizing radiations, the effective mutation rate may have varied widely according to the "ecological" conditions of hemoglobin within the organism and of the organism within its environment. In particular, during evolutionary transition periods such as afforded by a change from aquatic to terrestrial habits the effective mutation rates may have been much higher than at other times. The size of the populations at every stage is also of paramount importance in determining the evolutionarily effective mutation rate. Some other factors, the impact of which is equally difficult to evaluate, should also be taken into account. Fortunately the over-all result of the interplay of all factors is expressed in the speed of evolution, which has been evaluated. The general finding is that in the course of time evolution has become accelerated (Rensch, 1954). The more recent terrestrial groups of animals, on the average, have evolved faster than the more ancient aquatic groups. We may expect that this generalization, based on morphological characteristics, has its counterpart in the speed of evolution of deoxyribonucleic acid (DNA) and of the proteins. On the average, a lesser number of evolutionarily effective mutations per unit time may thus have affected the hemoglobin molecules during the initial phases of vertebrate evolution than in later periods. Our guess is that the numbers given in Table I are more likely to be underestimates than overestimates.

In the preceding evaluations we have equated one mutational event to one amino-acid substitution in the polypeptide chain. As mentioned, present evidence tends to show that this is the most frequent type of evolutionarily effective mutation in hemoglobin genes. It is too early to generalize to structural genes at large. Work in progress in several laboratories has a direct bearing on this question, such as the work of Yanofsky's group on *Escherichia coli* tryptophan synthetase at Stanford University, that of Fraenkel-Conrat's group on tobacco mosaic virus (TMV) mutants at Berkeley, that of Levinthal's group on *E. coli* alkaline phosphatase at M.I.T., and that of Brenner's group on phage-head protein at the Cavendish laboratory. There is ample evidence that mutational events other than single amino-acid substitutions exist. Some such evidence is derived from the comparison of the hemoglobin α- and β-chains themselves. In several regions of both chains sequences of contiguous amino acids numbering from 1 to 5 (or 6, if we include in the comparison the C-terminal sequence of sperm whale myoglobin, to which no counterpart is as yet known in hemoglobin chains) are missing. Three types of mechanisms might account for such observations: terminal growth of the chains, in the case of terminal differences; deletions; or insertions. Insertions may be duplications of chain segments, associated or not with a reversion of the segments. Although the events at the level of the chromosome and of a single DNA molecule may be qualitatively quite different, one must not discount the possibility that events of the type first described by Sturtevant (1925)—the insertion into a chromosome of a duplicate of a chromosomal region—may have its counterpart within one structural gene.

There are two one-amino-acid "holes" in the α as compared to the β-chain and one two-amino-acid "hole" and one five-amino-acid "hole" in the β as compared to the α-chain. If each of these is tentatively considered to be attributable to a single mutational event, then of a total of 78 mutational events that have led to the present differentiation of the α- and β-chains four events, i.e., 5% of the total, represent deletions, insertions, or terminal chain growth. It is possible that the actual proportion of mutations that result in events other than the substitution of a single amino acid in the polypeptide chain is much higher, if such mutations are more often lethal than "substitution mutations," as seems likely indeed.

"Substitution mutations," such as have so far always been found in mutant alleles of hemoglobin chain genes, if occurring in a major

chain gene and advantageous either will be eliminated rather quickly or will eventually replace entirely the wild-type gene. Likewise, "substitution mutations" that cause molecular disease either will be eliminated before there is time for a second mutation to occur in the mutant or, if a selective advantage exists under special circumstances for the heterozygote, as in the case of sickle-cell hemoglobin, will be confined to a small enough number of individuals so that the appearance of a second mutation in the same gene will be improbable. Considering a given species, there are only two types of cases where the appearance, by repeated single substitutions, of more than one difference between originally identical genes would be favored: (1) when the heterozygote is universally favored over the wild-type homozygote; this would apply to sickle-cell hemoglobin, if humans were universally exposed to malaria; (2) when a gene has duplicated and the conformity of the duplicate gene to the original model is not selected for. Assuming that the duplicate gene contributes to protein production, this conformity will be selected for only if an increase in output of a given polypeptide chain is advantageous.

D. The Destiny of Duplicate Genes and the Function of Genic Multiplicity

If gene duplication is one of the means of increasing the output of a given protein, one may distinguish two phases in this respect. Up to an optimum number of duplications, the duplicate gene will be selectively retained with a structure identical to and a position rather near to that of the mother gene. Beyond this point, duplicate genes will be progressively more strongly selected against. During this latter phase they will be in part eliminated with their carriers, and in part subjected to progressive change. When they are preserved and changed, their destiny may be of three types. They might evolve new useful functional properties. In this case they will be retained as active genes, and to the extent to which polypeptide output depends on gene duplication their own duplicates will be kept unchanged by natural selection. Secondly, functionless or unfavorable duplicates will not maintain duplicates to their own likeness and may themselves be translocated to other chromosome parts and be reduced to minor-component genes by a position effect. Some such minor-component genes, more or less profoundly changed in the meantime, may be selected for later in evolution and be changed

into major-component genes. Thirdly, the activity of the duplicate may be reduced to zero.

This elimination of gene activity may again take three forms. The changed structural gene may be bodily eliminated through the loss of the part of the chromosome that carries it; or it may be modified to such an extent that its products, although significant in amount, are no longer recognizable in terms of the original protein; or it may be preserved in a modified state, but totally or subtotally deprived of the power of expression.

The existence of such "dormant genes," although difficult to verify, is a plausible inference from two types of observations: firstly, that within a given tissue, say the hematopoietic tissue, major and minor structurally distinct components of a given type of polypeptide chain are found. Since at a given time the relative quantities of hemoglobin chains vary between 100 and 1%, other genetically distinct minor components may be present in such small amounts that they are practically undetectable. Secondly, there are numerous examples of proteins that are produced exclusively in one type of tissue, and of which no trace is found with presently available analytical means in other tissues. The nonproduction of hemoglobin in muscle cells and of myoglobin in reticulocytes is one example. This example shows that some among even relatively closely related structural genes may, within a given tissue, be the ones strongly expressed, the others unexpressed. We must assume that all the structural genes have during embryological development been communicated to all cell lineages. It is therefore quite likely that there exist in every organism numerous structural genes that do not find in any of the existing tissues conditions favorable to their expression and thus remain permanently dormant.

Furthermore the relative structural similarity of minor hemoglobin components to one of the major components affords yet another argument in favor of the existence of dormant genes. Indeed, as we have seen, the human δ- and β-chains are quite similar. Likewise structurally distinct components of orangutan hemoglobin have been found to be quite similar, and the same holds for pig hemoglobin components (E. Zuckerkandl, R. T. Jones, Y. Nishiwaki and L. Pauling, 1959–1961, unpublished). If duplicate genes remained usually expressed, one would expect to find a series of minor-component chains differing from all the other chains as much as the human α- from the human β-chain. This is, however, not

the case. One is led to think that, in the long run, duplicate minor-component genes most often cease to be expressed. There is no apparent reason why one should assume that they have most times been bodily eliminated. Other possibilities are that they have been implied in a further translocation, that a mutation or transposition of a controlling element (repressor) has occurred, or that the structurally modified mutated gene possesses specificity characteristics that fail to comply with the specificity requirements for polypeptide production under the conditions prevailing in the cell.

Dormant genes of course are conceived as dormant only as far as their expression, and not as far as their mutability goes. Mutations in dormant genes and in minor-component genes will never be lethal, unless the latter have some distinct specific function, which would then lead us to consider them as "major components" of another protein type. Minor-component genes and, mainly, dormant genes may thus furnish an important and perhaps the principal part of the genic raw material for macro-evolutionary experiments of nature. A new translocation, or the transposition of a controlling element such as those described by McClintock (1956), or some other genetic modification may reactivate the dormant gene after a very long period of time during which mutations have changed it enough so that it now controls the production of a new kind of protein. In this fashion new enzymes, new functions can arise without the corresponding loss of old enzymes, old functions. We have recalled earlier the importance to evolution of the loss of functions through the mutations of active genes. But it is evident that evolution, while it makes the best of such losses and of molecular disease, could not be based on them alone. There are a great many more different functions to be carried out by a great many more different types of enzymes than we are allowed to suppose can have existed in early evolutionary times. Primordial living matter must have been limited to a few simple functions. Therefore the notion of evolution by gain is a necessary complement to the notion of evolution by loss.

Horowitz (1945) made a lasting contribution to our thinking about evolutionary gain at the enzymatic level. He described how new reaction chains might arise in certain circumstances through the chance combination of the necessary genes and furthermore proposed a general mechanism for the stepwise building up of complex enzyme-systems, presenting us with a plausible scheme of macro-evolution at the molecular level. Obviously, as enzyme systems

become more complex, more different enzyme molecules are needed. There are reasons to think that the same molecules cannot usually be expected to carry out several different enzymatic functions. Therefore, new genes are needed for the building up of new functions. This is where the concept of the mutational reactivation of dormant genes complements Horowitz's picture. Minor-component genes are not to be excluded from a similar role, but may usually not yet be different enough from their parental genes to be fit for carrying out novel functions. Most minor-component genes, as stated above, are liable to be eventually turned into dormant genes because natural selection will not prevent their transfer to synthetically inactive chromosome regions, nor their coming otherwise under the influence of a repressor gene, nor structural changes that might place the genes outside the range of specificity requirements of the available macromolecules that collaborate with the structural gene in protein synthesis. As dormant genes become reactivated after periods of cryptic existence corresponding perhaps to geological ages, they may produce potential enzymes that do not disrupt existing chains of reactions, but are able to add new processes to the old ones, perhaps in the ways described by Horowitz. One of the possible mechanisms of the reactivation of dormant genes is the reactivation of the chromosomal region where it is located through a change in intracellular environment. The hope of demonstrating the existence of dormant genes rests on this possibility. Such a change in intracellular environment could result from the adaptation of the organism to changes in the external milieu. In adaptational changes, during an initial phase, gain and loss mutations may be balanced, or loss mutations only may occur, so that the total complexity of the organism either remains constant or tends to decrease. In the process, however, the environment of the chromosomes may be altered in some tissue and, on account of this alteration, genes be activated in some regions of the chromosomes, inactivated in others. Inactivations of this kind will be mostly lethal and genomes with corresponding inactivation-resistant chromosome regions will be selected for. The newly activated genes on the other hand will now be available to respond to further adaptational needs and will furnish a series of gain mutations without corresponding losses. Parts of this concept are supported by observation. Changes in conformation of the genic DNA molecules are presumably related to changes in the activity of the genes, and Schmitt (1956) has shown that the state of chromosomal DNA

depends on the chemical environment of the chromosomes. Furthermore, genes in mice that display the same activity characteristics during the individual's life time have been found to be closely linked on the same chromosome (Paigen and Noell, 1961).

Through the reactivation of dormant genes by a modification of the intracellular environment an initial adaptational stress of great magnitude would appear instrumental in producing a rise in complexity of the organism. Thus would be solved the old paradox expressed in this question: why should organisms get to be more complex, since simpler organisms are evidently adapted as well to their environment? Once more Biology will show that it can do without any "élan vital" or "entelechy."

The present concept leads one to predict that the fastest evolution toward more highly organized forms should take place after the occurrence of major environmental stresses. Evolutionary history bears out this expectation. For a long time paleontologists have noted that the initial phases in the development of new types of forms has an "explosive" character (consult, for instance, Rensch, 1954). According to the present theory, we assume that at the onset of such an explosive evolutionary phase a change in intracellular environment has brought about the reactivation of a number of previously dormant genes. Rensch believes that the phenomenon of explosive evolutionary phases is adequately accounted for by increased natural selection accompanying the conquest of new biotopes. He thus considers as instrumental a change in *external* environment only. This however does not explain the trend toward increased complexity of the forms.

In the sense that has been laid out here a marked change in environmental conditions may lead to what a layman would call a shake-up of the genome, and this may be the part of reality behind the poorly documented and falsely interpreted observations of the Soviet anti-geneticist A. Lysenko.

To sum up, mutations of active genes controlling "major" protein components suffice to explain how an organism can adapt to changes in its environment. Mutations of minor-component genes and dormant genes, however, seem to be able to furnish the organism with the genic raw materials that eventually allow it not only to adapt to a new environment but also, in the process, to become more highly organized. Minor-component genes and dormant genes may thus prepare the major steps in evolution.

Beside a minor-component multiplicity, there is also a major-component multiplicity in protein production, especially the one characterized by a succession of major components in time. Why are hemoglobin β-chains substituted for γ-chains? The reason may not be that β-chains would not be fit to meet the respiratory prerequisities of intrauterine life and the γ-chains those of adult life, but, rather, that the structural genes corresponding to the two chains are located in two distinct chromosomal regions, one of which is activated in the particular intracellular environment determined by the organism at early developmental stages, and will not be active under other conditions, while the second chromosomal region, on the contrary, will be activated only at the end of embryonic development due to the presence or absence of some particular factors in the cell at that time.

To carry out a given function it is thus not sufficient for a cell to possess a favorable structural gene. A further prerequisite is that this gene either be located in a chromosome region that remains active in protein synthesis in spite of the changes of intracellular environment that occur during development, as is the case of the hemoglobin α-chain gene; or that there exist several duplicates of the gene, each of which is located in a chromosomal region that is active during certain phases of development, and that these duplicates are distributed in such a way over the genome that at any time of development at least one of them remains "on duty." These different genes will never be identical because, although they have supposedly arisen by duplication of an original gene, translocation is apparently a rarer event than amino-acid substitution, so that translocation genes will be expected to differ from each other by more than one amino-acid substitution.

Constant vital functions thus frequently need to have at their disposal several editions of a given type of genes in several regions of the genome that are successively activated and inactivated, or vice versa, with respect to protein-synthesizing ability. This view is advanced as an explanation of the generality of the most important types of "major-component" multiplicity in proteins. These types are on the one hand the successive embryonic and adult editions of a protein; and on the other hand the different editions found at any one time in different tissues of the same animal. The latter type of protein heterogeneity has been referred to earlier and may be interpreted in the same terms. In each tissue the particular intracellular

environment provides different conditions for the distribution of active and inactive regions throughout the genome, and thus different duplicates of a given gene that have undergone different translocations will be "on duty" in different tissues.

It thus appears that the unavoidable change in intracellular environment during embryonic development is a great challenge to embryonic development itself, because of the obligation that many critically important proteins be produced throughout and in spite of this change. One may therefore venture to say that there could be no embryonic development without gene duplication followed by gene translocation. In this theory the same events that furnish the genic raw materials for evolution also furnish the genic raw materials for ontogeny.

Other types of genes, of course, are active in protein synthesis during one or the other phase of ontogenetic development only. Certain functions, for instance, are developed only in the adult organism. When their expression is delayed or hastened, paedomorphosis or palingenesis ensues (see for instance Rensch, 1954). Chromosomal events of the type described may be postulated to cause these phenomena also.

After having insisted on the cause and on the function of genic multiplicity, in particular in relation to hemoglobin structural heterogeneity, we must here point out that there is a limit to hemoglobin heterogeneity, at least so far as phenotypic expression goes. A universal intraspecific structural multiplicity of all types of proteins is on the way to being established and this finding may be an important step forward in the recognition of biological reality. Yet part of this step is the rejection of older speculations about protein heterogeneity. These postulated a continuous spectrum of structural variants in the case of each protein. As a result, the apparent amino-acid composition of a protein would be only a mean composition, and proteins would not exist as strictly defined chemical species. Work of recent years has shown that, except for a possible low percentage of "errors" in synthesis, the chemical formula of proteins is as rigorously defined as that of simple molecules. We can no longer agree with J. S. Haldane and J. G. Priestley, who wrote in 1935 "It does, in fact, appear to be fairly certain that each individual has a specific kind of hemoglobin, just as he has a specific nose." Concordant amino-acid sequence analyses in three laboratories (Braunitzer in the Max Planck Institute in Munich; Schroeder in the California

Institute of Technology; Hill and Konigsberg in the Rockfeller Institute) of human hemoglobin chains obtained from different individuals, as well as the comparison of tryptic peptide patterns (Ingram, 1958) obtained with different human hemoglobins in many more instances, shows that this is not so. Of course a greater number of structurally distinct human hemoglobin chains are probably produced than we presently know about. This applies in particular to the abnormal human hemoglobins, about thirty of which have so far been described, and whose inventory one will probably never be able to consider as complete. This also applies to the normal human hemoglobin chains controlled by non-allelic genes, of which a few more may eventually be found.

Some workers go further and believe that a number of undetected hemoglobin alleles may be present in the population. They may escape notice if they behave identically in electrophoresis, but differ in uncharged amino acids. One of us (L.P.) and Itano (1957) have proposed this idea as a possible explanation for the inhibition of hemoglobin synthesis in the molecular diseases known under the name of the thalassemias. Changes in structural genes may indeed lead to an inhibition and even to a total loss of the ability to synthesize a protein. Ingram and Stretton (1959) have developed this idea as one of several possible explanations of the thalassemias. It has ceased to be likely that such cryptic mutants are of very general occurrence in normal populations. The great similarity between gorilla and human hemoglobin chains is a piece of evidence against such a view. With the exception of one difference relating to a serine residue in one of the chains, the uncharged amino acids may all be the same. This similarity suggests that nonpathogenic undetected structural variants of hemoglobin chains must be rather rare, unless a human is more often like a gorilla than like a human. It is true that this hypothesis is supported by more observations than most biological theories.

III. Three Types of Molecular Diseases

After having considered molecular disease in its relation to evolution and genic variability in its relation to evolution, we may comment on genic variability in relation to molecular disease. Three types of molecular disease may be distinguished. Mutations may (1) interfere with molecular function, (2) interfere with the adaptation of the molecule to the intracellular environment, and (3) interfere

with the rate of synthesis of a molecule that is functionally fit. Hemoglobin mutants offer examples of all three types.

A. Interference with Function

The only known alterations of amino-acid sequence that lead directly to an interference with the oxyphoric function of hemoglobin are those of the pigments collectively known under the name of HbM. They are all characterized by the formation of methemoglobin, in which the iron is oxidized to the tripositive state (Gerald, 1960). It was recently shown that several distinct mutations give rise to abnormal methemoglobin formation (Gerald, 1958, and personal communication, 1960). It is interesting that structural studies have revealed that the amino-acid substitutions leading to HbM formation all affect a certain region of the hemoglobin polypeptide chains. This region has been called the "basic center," because it comprises a relatively large amount of basic amino acids, and we may call it "basic center I" to distinguish it from a second basic center further along the chain. The two basic centers have in common the property of containing a histidine that seems to be in relation with the heme iron. While the main linkage of the heme iron to the globin is supposed to be to a histidine of basic center II, the histidine of basic center I is, according to present evidence (Watson and Kendrew, 1961), placed opposite the sixth coordination position of the iron, the one that binds oxygen in the oxygenated state and water in the deoxygenated state (Haurowitz, 1949). It may be that the second histidine is the one mentioned by Conant (1934) and Coryell and Pauling (1940). Table II shows the variability of some features of sequence and the constancy of others of the peptide region around basic center I in different animal species. Abnormal human mutants in which this region is affected are also listed. Corresponding results relative to basic center II have not yet been forthcoming.

The seven consecutive amino-acid residues shown in the table comprise residues numbers 56 to 62 in the α-chain and numbers 61 to 67 in the β-chain, counting from the amino end. Three of the seven residues are shown to be substituted in some normal respiratory pigments and the changes, as in the ovine α-chain and in one of the orangutan chains, may involve the substitution of an acid for a neutral amino acid without any fundamental interference with hemoglobin function. Of course such substitutions would be expected to affect one or several of the physical parameters of oxygenation.

TABLE II
"BASIC CENTER I"

List of Hemoglobin Chains Differing in This Region from the Human α- or β-Chain

Chain	Sequence	Reference
Human α, normal	–lys–gly–his–gly–lys–lys–val–	Schroeder *et al.* (1961)
Human β, normal	–lys–ala–his–gly–lys–lys–val–	Braunitzer *et al.* (1961)
Human α, Norfolk	–lys–*asp*–his–gly–lys–lys–val–	Baglioni (1961)
Human β, $M_{Milwaukee}$	–lys–ala–his–gly–lys–lys–*glu*–	P. S. Gerald *et al.* (personal communication, 1960)
Human α, M_{Boston}	–lys–gly–*tyr*–gly–lys–lys–val–	P. S. Gerald *et al.* (personal communication, 1960)
Human β, M_{Emory}	–lys–ala–*tyr*–gly–lys–lys–val–	P. S. Gerald *et al.* (personal communication, 1960)
Human β, Zürich	–lys–ala–*arg*–gly–lys–lys–val–	C. J. Muller and Kingma (1961)
Bovine α	–lys–gly–his–gly–*ala*–lys– (or arg)	C. J. Muller (1961)
Ovine α	–lys–gly–his–gly–*glu*–lys– (or arg)	C. J. Muller (1961)
Goat α	–lys–gly–his–gly–*glu*–lys– (or arg)	C. J. Muller (1961)
Horse α	–lys–*ala*–his–gly–lys–lys–	Inferred from Braunitzer and Matsuda (1961)
Orangutan α or β	–lys–*asp*–his–gly–lys–lys– (or *glu*)	C. Baglioni (personal communication, 1961)
Sperm whale myoglobin	–lys–*val*–his–gly–*ileu*–glu–val– (or glu)	Watson and Kendrew (1961); Edmundson and Hirs (1961)

The reduction of the basicity of this peptide region is still more marked in sperm whale myoglobin. While both types of chains of human, gorilla, chimpanzee, beef, and horse hemoglobins, one of the chains of orangutan hemoglobin, and the β-chains of sheep and goat possess four basic amino acids in the stretch of seven shown in the table, sperm whale myoglobin probably retains only the histidine

opposite the heme group as well as the initial lysine. In various types of fish the corresponding peptide must also be less basic than in man, as tryptic peptide patterns indicate (Zuckerkandl *et al.*, 1960). However, four amino-acid residues out of the seven have so far not been shown to vary in any of the normal respiratory pigments, and may be essential to the oxyphoric function.

Methemoglobin formation in man has so far been shown to be caused by substitutions at two different residues, and one of these substitutions has been observed in both α- and β-chains. In $HbM_{Milwaukee}$ there is a substitution at the valine in the fourth position C-terminally with respect to the histidine that is supposedly in relation with the heme iron. In the two other kinds of HbM that have been structurally analyzed, HbM_{Boston} and HbM_{Emory}, it is this histidine that is affected; in both cases it is replaced by tyrosine. A replacement of this histidine by another amino acid does however not necessarily lead to methemoglobinemia, since methemoglobinemia has not been reported as a feature normally observed in the family possessing $Hb_{Zürich}$ (Hitzig *et al.*, 1960), although oxydation may be facilitated. In $Hb_{Zürich}$ the histidine in question is replaced by arginine. Thus the basic character of histidine seems to be more essential in protecting the heme iron from oxidation than its particular configuration, and the affinity for heavy metals, which is much greater in histidine than in arginine (Albert, 1952), seems also not to be involved.

Since the "basic center I" region is probably part of an α-helix, as is the corresponding region in sperm whale myoglobin (helical region E, Watson and Kendrew, 1961), the fourth residue after the critical histidine should lie next to it on the helix. It is plausible that the acid residue found in this position in the case of $HbM_{Milwaukee}$ interferes with function. Thus all types of HbM that have so far been analyzed seem characterized by a change in state or in kind of the critical histidine residue. On the other hand, the aspartic acid that replaces a neutral amino acid in $Hb_{Norfolk}$ does not seem to interfere basically with the function of the neighboring histidine, and this is probably due to the fact that because of the different orientations of neighboring residues in a helix the side chains of neighboring residues can be further apart than the side chains of residues four removed along the helix.

A further interesting observation related to the basic center I is that the change in the orangutan peptide is apparently the same as

in the abnormal $Hb_{Norfolk}$ (C. Baglioni, 1961, and personal communication, 1961). In this respect human carriers of $Hb_{Norfolk}$ have orangutan hemoglobin. The disease is mild; but this hemoglobin has not yet been observed in the homozygous state. The case shows that what may appear as a "molecular disease" in one species may be the norm in another.

B. Interference with Normal Intermolecular Relations

Mutations that do not significantly affect the oxyphoric function of hemoglobin may nevertheless lead to severe molecular diseases if they alter the physicochemical properties of the molecules that are of importance in its relation to sister molecules and to other constituents of the red cell. Among the most important of these properties is solubility. Considering the great proportion of nonpolar groups usually found in proteins, the building of a soluble protein molecule appears to be a difficult accomplishment. The readiness with which solubility is lost upon changing spatial comformation (denaturation) demonstrates that the solubility of most proteins is very sensitive to the distribution of their side-chains in space. Mutations that result in the substitution of a polar by a nonpolar amino-acid residue, as well as other mutations that weaken the conformational stability, may thus be expected to interfere frequently with the functionally required solubility characteristics of a protein. One might surmise that many molecular diseases should involve such losses of solubility. If it is justified to consider aging as a multiple molecular disease arising through somatic mutations, then aging probably also expresses in part the loss of solubility of certain proteins.

Hemoglobin is expected to be particularly sensitive in this respect to mutational change because it is in solution in the red cell at a concentration of about 30%, a concentration that not many molecules are able to achieve. Even a slight change in the properties of molecular interaction may under such circumstances lead to a drastic effect. Under certain conditions the abnormal human hemoglobins H [composed of four β-chains (Jones *et al.*, 1959a)] and Zürich (Hitzig *et al.*, 1960) tend to precipitate in the red cell. Among the cases in point, the most well known, and the one that typified molecular diseases in general (Pauling *et al.*, 1949), is that of HbS, the hemoglobin of sickle-cell anemia chemically characterized by the substitution of a valyl for a glutamyl residue in the sixth position from the N-terminus of the β-chain (Ingram, 1959). This substitution does not lead to a

decrease in solubility of the oxyhemoglobin (Perutz and Mitchison, 1950), but upon deoxygenation the molecules, when in sufficiently concentrated solution, interact and align along fibers that seem to form a network reminiscent of gelification (Bessis *et al.*, 1958), and at the level of cellular dimensions the alignment of the fibers is expressed by the formation of tactoids (Harris, 1950), which deform the cell membrane and lead to an early destruction of the red cells and to interference with blood flow in capillaries.

The oxygen-dissociation curve, its position, shape, and dependence on pH, and thus the fitness of the hemoglobin with respect to a given set of circumstances may be affected indirectly by amino-acid substitutions that alter the relation between the respiratory pigment and its cellular environment. The oxygen affinity within the erythrocytes of HbS is lower than that of HbA (Becklake *et al.*, 1955), whereas very similar affinities have been reported for these two pigments in cell-free solution (Allen and Wyman, 1954). Recent evidence indicates however that also in cell-free solution the oxygen affinity of HbS is lower than that of HbA (Riggs and Wells, 1961). As to HbF, it has a greater oxygen affinity than HbA within the red cell, yet in cell-free dialysed solution the affinities of HbA and HbF appear to be nearly identical (Allen *et al.*, 1953), except at pH values below 7 (Manwell, 1960). Red cells containing fetal or adult hemoglobin, respectively, differ in certain parameters such as surface, volume, and thickness (Riegel *et al.*, 1959). This difference may possibly express the influence of structural characteristics of the hemoglobin molecule on structural characteristics of the red cell. Admittedly other factors, such as the relative rate of synthesis of different hemoglobins, will also intervene here. This leads us to consider briefly the third type of molecular diseases, associated with the hampering of protein synthesis.

C. Interference with Synthesis

Hemoglobin synthesis may be interfered with by various mechanisms. One of the most obvious is the absence of red-cell production. This apparently occurs as a normal feature in a family of antarctic bony fish, the *Chaenichthydae* (Ruud, 1954). When red cells are present a decrease in the output per cell of individual hemoglobin chains is known as a heritable character in many human families. Most of the known abnormal human hemoglobins, while not unfit as oxygen carriers, are present in the red cells in subnormal amounts.

HbJ, which is present in higher quantities per red cell than HbA, is so far the only exception to this statement (Thorup *et al.*, 1956). The decreased ratio of abnormal to normal hemoglobins may be assumed to express a decreased relative rate of synthesis.

Itano (1953, 1957) was the first to point out that the structure of a gene may have a direct influence on its synthetic activity. The decrease in rate of synthesis of most abnormal hemoglobins may be visualized in at least two different ways. It may be due to an interaction between the structural gene and its substrate on the chromosome; or it may be due to an interaction between the gene and extrachromosomal macromolecules, which might be called co-determinant factors. The collaboration of co-determinant factors with the genes or the primary gene products is assumed to be necessary for protein synthesis. An alteration of the gene structure would upset the balance of attractive and repulsive forces (of electrostatic, van der Waals and steric nature) between these molecules. A decreased rate of synthesis would then result from a decreased degree of fitting between the two entities. We have referred earlier to the probable change of state in the genes during various phases of the development of an organism. Co-determinant factors also may change in tertiary structure as a result of altered intracellular conditions, or they may combine with smaller molecules, perhaps hormones, that change their specificity and activity. The rate of protein synthesis as a function of the intracellular environment may thus be determined by changes of state of the genes on the chromosomes as well as by changes of state of other macromolecules in solution.

Evidence is accumulating to the effect that large parts of protein molecules are not connected with any of their known specific functions. Various hypotheses can be used as partial tentative explanations of this fact such as the stabilization of the tertiary structure in larger proteins, the random survival of structures that were useful during earlier evolutionary stages, and the resistance to diffusion of large molecules (Pauling, 1957b). Consideration should also be given to the possibility that many parts of the protein molecule that appear functionally neutral may have a function in connection with the rate of synthesis of the protein. Natural selection may well act not only at the level of the finished protein product, but also at the level of its production.

The decrease in the rate of synthesis of hemoglobin chains is often more considerable in thalassemias than in the case of abnormal

hemoglobins. While a single genetic event probably causes both the amino-acid substitution and the change in rate of synthesis in abnormal hemoglobins (Itano, 1959), there is evidence that more than one type of genetic event may be involved in thalassemias. The mechanism of these diseases is being investigated in many quarters and has been recently discussed by Ingram and Stretton (1959) as well as other authors. None of the theories so far proposed is completely satisfactory. Whatever the theory, it should be kept in mind that the inhibition of protein synthesis is a nonspecific effect that may have different causes and that there may be several types of thalassemia, not only, as is commonly recognized, with respect to the particular hemoglobin chain that is most severely affected, but also with respect to the underlying mechanism.

IV. Fighting Molecular Disease

In the more or less distant future an enzyme deficiency such as the one that causes phenylketonuria may be met by endowing the organism permanently with a certain quantity of a stable artificial enzyme (Pauling, 1956). Of course, such a solution could hardly be considered in the case of a respiratory pigment, where the quantities required with respect to the "substrate" are not catalytic, but stoichiometric.

Another conceivable means of fighting molecular disease that cannot at present be theoretically excluded is the activation in a certain tissue through drug action of a minor-component gene or a dormant gene, representing a functionally fit duplicate of the functionally unfit "major-component" mutant. Whether or not such a treatment is possible depends on the nature of the control of gene activity. If this activity is in part controlled by position effects in relation to factors present in the intracellular environment, medical research might take such a course, although of course the problems of cell permeability and of the toxicity to the activity of other important genes of the agents capable of modifying specifically the intracellular environment might create major obstacles. The fact that this question can be raised shows however that research on the mechanisms of control of gene activity is not only of fundamental importance to biology, but of great interest to medicine as well.

None of the means of meeting the challenge of molecular disease will, however, be as satisfying as the elimination of the disease-causing mutant genes from the human populations. This is theoretically

the best, and it is at present the only concrete efficient measure that can be proposed. H. J. Muller (1959, 1961) who has given considerable attention to eugenics, has recently proposed the creation of germ-cell banks, from which prospective parents could draw the choicest human genomes. Such discussions and proposals are of paramount importance, even though one might not share Muller's optimism, which leads him to believe that in a process of free choice of genomes on the part of the populations the greatest human values would all get their fair share. It may be anticipated that governments would advocate and propagandize the choice of the socially minded, the active, the efficient, at the expense of the contemplative person; and the "good fellow" who represents the majority of humans would tend to procreate by choice a "super good fellow," a super corporation man, more able in conforming than in intellectual accomplishments. Without the contemplative, endowed with refined perceptive abilities of the qualitative and of the significance of forms, the human race would deteriorate. They are indeed humanity's greatest asset, both in the realm of doing and in the realm of being. It is true that the present development of the world toward a huge hive where nothing can stand in the path of technology and mass production promises to individuals so endowed little in the future except frustration and unhappiness, and one might contend that to relieve their suffering it might be a good deed to eliminate them genetically.

On the other hand, no objection can be legitimately raised, it seems to us, against the ambition to eliminate from human heredity those genes that lead to clearly pathological manifestations and great human suffering. The means of achieving this goal need to be discussed. We know now that in the United States about 10% of the Negro population carry one HbS or HbC gene. Therefore about one 400th of the children born to Negro parents have the deadly disease sickle-cell anemia. A simple test permitting the detection of the heterozygous carriers of a sickle-cell-hemoglobin gene exists, and as a first protective step there should be a law requiring all persons within a population in which this gene is present to any significant extent to submit themselves to this test.

If people carrying the mutant gene were to refrain from marrying one another, but married normal individuals, the incidence of the gene would remain constant in the population, and the problem of eliminating the gene would not be solved. To eliminate the mutant

gene the following rules may be proposed. If two heterozygotes marry they should have no children of their own. If a heterozygote marries a normal person they should have a number of progeny smaller than the average. In this way the mutant gene would be eliminated in the course of time in a way not involving the suffering caused by the birth of the defective children. Similar measures should be taken in the case of phenylketonuria and other molecular diseases.

In a marriage of heterozygotes, who do not suffer from the diseases caused by recessive genes, the chance that each child is homozygous for the defective gene is 25%. This percentage is much too high to let private enterprise in love combined with ignorance take care of the matter. And although interference of law in questions that are to a great extent of a very personal nature is to be avoided whenever possible, it would be clearly unethical to oppose such an interference in the case of molecular disease, at the very least in those cases, such as sickle-cell anemia, where presently available palliative measures are inadequate.

We may accordingly have hope that the increase in knowledge about molecular disease will in the course of time lead to a significant decrease in the amount of human suffering in the world.

Acknowledgment

We thank Professors E. B. Lewis and N. H. Horowitz for stimulating discussions about some of the topics referred to in this article.

References

Albert, A. (1952). *Biochem. J.* **50,** 690.
Allen, D. W., and Wyman, J., Jr. (1954). *Rev. hématol.* **9,** 155.
Allen, D. W., Wyman, J., Jr., and Smith, C. A. (1953). *J. Biol. Chem.* **203,** 81.
Allison, A. C., (1957). *Exptl. Parasitol.* **6,** 418.
Andinolfi, M., Chieffi, G., and Siniscalco, M. (1959). *Nature* **184,** 1325.
Baglioni, C. (1961). *Federation Proc.* **20,** 254.
Becklake, M. R., Griffith, S. B., McGregor, M., Goldman, H. I., and Schreve, J. P. (1955). *J. Clin. Invest.* **34,** 751.
Bessis, M., Nomarski, G., Thiery, J. P., Breton-Gorini, J. (1958). *Rev. hématol.* **13,** 249.
Braunitzer, G., and Matsuda, G. (1961). *Z. physiol. Chem.* **324,** 91.
Braunitzer, G., Rudloff, V., Hilse, K., Liebold, B., and Müller, R. (1960a). *Z. physiol. Chem.* **320,** 283.
Braunitzer, G., Hilschmann, N., Hilse, K., Liebold, B., and Müller, R. (1960b). *Z. physiol. Chem.* **322,** 96.

Braunitzer, G., Hilschmann, N., Rudloff, V., Hilse, K., Liebold, B., and Müller, R. (1961). *Nature* **190,** 480.
Bridges, C. B. (1935). *J. Heredity* **26,** 60.
Cepellini, R. (1959a). *Acta Genet. Med. et Gemellol.* **8,** Suppl. II, p. 47.
Cepellini, R. (1959b). *In* "Biochemistry of Human Genetics" (Ciba Foundation Symposium), p. 133. Churchill, London.
Conant, J. B. (1934). *Harvey Lectures Ser.* **28,** 159.
Cori, G. T. (1954). *Harvey Lectures Ser.* **48,** 145.
Coryell, C. D., and Pauling, L. (1940). *J. Biol. Chem.* **132,** 769.
Crick, F. H. C. (1958). *Symposia Soc. Exptl. Biol.* **12,** 138.
Dodson, E. O. (1960). "Evolution: Process and Product," 352 pp. Reinhold, New York.
Edmundson, A. B., and Hirs, C. H. W. (1961). *Nature* **190,** 663.
Gerald, P. S. (1958). *Blood* **13,** 936.
Gerald, P. S. (1960). *In* "The Metabolic Basis of Inherited Disease" (J. B. Stanbury, J. B. Wyngaarden, and D. S. Fredrickson, eds.), pp. 1068–1085. McGraw-Hill, New York.
Gratzer, W. B., and Allison, A. C. (1960). *Biol. Revs.* **35,** 459.
Haldane, J. S., and Priestley, J. G. (1935). "Respiration," New ed., 493 pp. Clarendon Press, Oxford.
Harris, J. W. (1950). *Proc. Soc. Exptl. Biol. Med.* **75,** 197.
Haurowitz, F. (1949). *In* "Haemoglobin" (P. J. W. Roughton and J. C. Kendrew, eds.), pp. 53–56. Interscience, New York.
Hitzig, W. H., Frick, P. G., Betke, K., and Huisman, T. H. J. (1960). *Helv. Paediat. Acta.* **15,** 499.
Horowitz, N. H. (1945). *Proc. Natl. Acad. Sci. U.S.* **31,** 153.
Huisman, T. H. J., van de Brande, J., and Meyering, C. A. (1960). *Clin. Chim. Acta* **5,** 375.
Hunt, J. A. (1959). *Nature* **183,** 1373.
Ingram, V. M. (1958). *Biochim. et Biophys. Acta* **28,** 539.
Ingram, V. M. (1959). *Biochim. et Biophys. Acta* **36,** 402.
Ingram, V. M. (1961a). "Hemoglobin and its Abnormalities," 153 pp. Charles C Thomas, Springfield, Illinois.
Ingram, V. M. (1961b). *Nature* **189,** 704.
Ingram, V. M., and Stretton, A. O. W. (1959). *Nature* **184,** 1903.
Ingram, V. M., and Stretton, A. O. W. (1961). *Nature* **190,** 1079.
Itano, H. A. (1953). *Am. J. Human Genet.* **5,** 34.
Itano, H. A. (1956). *Ann. Rev. Biochem.* **25,** 331.
Itano, H. A. (1957). *Advances in Protein Chem.* **12,** 215.
Itano, H. A. (1959). *In* "Abnormal Haemoglobins" (J. H. P. Jonxis and J. F. Delafresnaye, eds.), pp. 1–17. Charles C Thomas, Springfield, Illinois.
Itano, H. A., and Pauling, L. (1957). *Svensk Kem. Tidskr.* **69,** 509.
Jones, R. T. (1961). Chromatographic and Chemical Studies of Some Abnormal Human Hemoglobins and Some Minor Hemoglobin Components. Ph.D. Thesis, California Institute of Technology.
Jones, R. T., Schroeder, W. A., Balog, J. E., and Vinograd, J. R. (1959a). *J. Am. Chem. Soc.* **81,** 3161.

Jones, R. T., Schroeder, W. A., and Vinograd, J. R. (1959b). *J. Amer. Chem. Soc.* **81,** 4749.
Kaplan, N. O., Ciotti, M. M., Hamolsky, M., and Bieber, R. E. (1960). *Science* **131,** 392.
Kleihauer, E., Braun, H., and Betke, K. (1957). *Klin. Wochschr.* **35,** 637.
Konigsberg, W., Guidotti, G., and Hill, R. J. (1961). *J. Biol. Chem.* **236,** PC55.
Kunkel, H. G., and Wallenius, G. (1955). *Science* **122,** 288.
Kunkel, H. G., Ceppellini, R., Muller-Eberhard, U., and Wolf, J. (1957). *J. Clin. Invest.* **36,** 1615.
Lehmann, H. (1959). *Nature* **184,** 872.
Lenhert, P. G., Love, W. E., and Carlson, F. D. (1956). *Biol. Bull.* **111,** 293.
Lewis, E. B. (1951). *Cold Spring Harbor Symposia Quant. Biol.* **16,** 159.
Lwoff, A. (1943). "L'évolution physiologique. Etude des pertes de fonctions chez les microorganismes," 308 pp. Hermann, Paris.
McClintock, B. (1956). *Cold Spring Harbor Symposia Quant. Biol.* **21,** 197.
Manwell, C. (1957). *Science* **126,** 1175.
Manwell, C. (1958a). *Physiol. Zool.* **31,** 93.
Manwell, C. (1958b). *J. Cellular Comp. Physiol.* **52,** 341.
Manwell, C. (1960). *Comp. Biochem. Physiol.* **1,** 267.
Margoliash, E., and Tuppy, H. (1960). Presented at the 138th Annual Meeting of the American Chemical Society, New York, September, 1960.
Markert, C. L., and Møller, F. (1959). *Proc. Natl. Acad. Sci. U.S.* **45,** 753.
Metz, C. W. (1947). *Am. Naturalist* **81,** 81.
Mizushima, S. I., and Shimanouchi, T. (1961). *Advances in Enzymol.* **23,** 1.
Muller, C. J. (1961). A Comparative Study on the Structure of Mammalian and Avian Hemoglobins. Ph.D. Thesis, Groningen, Netherlands.
Muller, C. J., and Kingma, S. (1961). *Biochim. et Biophys. Acta* **50,** 595.
Muller, H. J. (1956). *Acta genet. statist. med.* **6,** 157.
Muller, H. J. (1959). *Prespectives in Biol. and Med.* **3,** 1.
Muller, H. J. (1961). *Science* **134,** 643.
Neel, J. V. (1959). *In* "Abnormal Haemoglobins" (J. H. P. Jonxis and J. F. Delafresnaye, eds.), p. 158. Charles C Thomas, Springfield, Illinois.
Paigen, K., and Noell, W. K. (1961). *Nature* **190,** 148.
Pauling, L. (1956). *Am. J. Psychiat.* **113,** 492.
Pauling, L. (1957a). *In* "Arbeiten aus dem Gebiet der Naturstoffchemie. Festschrift Arthur Stoll," p. 597. Birkhäuser, Basel.
Pauling, L. (1957b). *Am. Inst. Biol. Sci. Publ. No.* **2,** 186.
Pauling, L., Itano, H. A., Singer, S., and Wells, I. C. (1949). *Science* **110,** 543.
Perutz, M. F., and Mitchison, J. M. (1950). *Nature* **166,** 677.
Piveteau, J. (1955). *In* "Traité de Zoologie" (P. P. Grassé, ed.), vol. 17, p. 1. Masson, Paris.
Rensch, B. (1954). "Neuere Probleme der Abstammungslehre. Die transspezifische Evolution," 2nd ed., 346 pp. Ferdinand Enke, Stuttgart.
Rhinesmith, H. S., Schroeder, W. A., and Pauling, L. (1957). *J. Am. Chem. Soc.* **79,** 4682.
Rhinesmith, H. S., Schroeder, W. A., and Martin, N. (1958). J. *Am. Chem. Soc.* **80,** 3358.
Riegel, K., Bartels, H., and Schneider, J. (1959). *Z. Kinderheilk.* **83,** 209.

Riggs, A., and Wells, M. (1961). *Biochim. Biophys. Acta* **50,** 243.
Roche, J., and Fontaine, M. (1940). *Ann. inst. oceanog.* **20,** 77.
Rossi-Fanelli, A., and Antonini, E. (1956). *Arch. Biochem. Biophys.* **65,** 587.
Rossi-Fanelli, A., Antonini, E., and Giuffrè, R. (1960). *Nature* **186,** 896.
Rumen, N. M. (1959). *Acta Chem. Scand.* **13,** 1542.
Ruud, J. T. (1954). *Nature* **173,** 848.
Schmitt, F. O. (1956). *Proc. Natl. Acad. Sci. U.S.* **42,** 806.
Schroeder, W. A., and Matsuda, G. (1958). *J. Am. Chem. Soc.* **80,** 1521.
Schroeder, W. A., Jones, R. T., Shelton, J. R., Shelton, J. B., Cormick, J., and McCalla, K. (1961). *Proc. Natl. Acad. Sci. U.S.* **47,** 811.
Shelton, J. R., and Schroeder, W. A. (1960). *J. Am. Chem. Soc.* **82,** 3342.
Siniscalco, M., Bernini, L., Latte, B., and Motulsky, A. G. (1961). *Nature* **190,** 1179.
Smith, E. W., and Thorbert, J. V. (1958). *Bull. Johns Hopkins Hosp.* **101,** 38.
Sturtevant, A. H. (1925). *Genetics* **10,** 117.
Svedberg, T. (1933). *J. Biol. Chem.* **103,** 311.
Thorup, O. A., Itano, H. A., Wheby, M., and Leavell, B. S. (1956). *Science* **123,** 889.
Watson, H. C., and Kendrew, J. C. (1961). *Nature* **190,** 670.
Wilson, S., and Dixon, G. H. (1961). *Nature* **191,** 876.
Zuckerkandl, E., and Schroeder, W. A. (1961). *Nature* **192,** 984.
Zuckerkandl, E., Jones, R. T., and Pauling, L. (1960). *Proc. Natl. Acad. Sci. U.S.* **46,** 1349.