20 December 1963, Volume 142, Number 3599 SCIENCE

# Correlation of Structure and Function in Enzyme Action

Theoretical and experimental tools are leading to correlations between enzyme structure and function.

D. E. Koshland, Jr.

Enzymes were first identified in 1833 (*1*), and since that time the central role of enzymes in biological processes has been established. Today over 700 different enzymes are known, and well over 100 have been crystallized. Most of the enzymes responsible for digestion, muscular contraction, metabolism of carbohydrates, and synthesis of fats, to name a few processes, are known. Even in the less well understood areas, such as vision and nerve conduction, substantial beginnings have been made. Moreover, there are great similarities in enzyme systems of species ranging from bacteria to man, indicating common patterns in the way these amazing molecules catalyze the dynamics of living systems.

The correlation of the function of an enzyme with its structure has ramifications in many areas, but I might mention two major ones. The first is in the field of medical science. As greater control over infectious diseases is achieved, hereditary diseases caused by "the inborn errors of metabolism" move nearer the center of the stage. In these diseases defective enzymes are to blame. Thus, one type of mental illness has been traced to the deficiency of one enzyme, and the deficiency of the enzyme may in turn result from a tiny change in its structure—for example, the loss of one of its several hundred amino acid residues. A change in an amino acid residue may be harmless, beneficial, or deleterious, depending on the type of change and the position of the amino acid in the enzyme which is altered. Thus, the one gene, one enzyme concept of Beadle and Tatum (*2*) makes it clear that knowledge of the structural basis of enzyme action will help us understand and cure hereditary diseases.

The second area in which this correlation will be of interest is that of chemical theory. No man-made catalyst has yet achieved the specificity or catalytic power of these fragile molecules formed by living systems. To understand them will require new developments in chemical theory, whose application will be of untold benefit.

In this article the structure of enzymes and the concept of an "active site," which has been so helpful in interpreting enzyme action are described first. After this some typical enzyme modification studies are outlined, to indicate the kind of information which has been obtained by these methods. Finally, an attempt is made to synthesize a consistent picture of the way in which enzymes act, based on the modification studies and appropriate theoretical models. Since it is impossible to cover the field comprehensively, the choice of enzymes for illustration is necessarily somewhat arbitrary. The entire area of oxidative enzymes and prosthetic groups is omitted, for example. However, the purpose of this article is to illuminate the state of present knowledge, and the omitted examples are, as far as we know now, consistent with the theories presented here.

## Structure of Enzymes

Enzymes are globular proteins whose molecular weights range from approximately 10,000 up into the millions. The average molecular weight of an individual amino acid residue is 100; the smallest enzymes are composed of approximately 100 amino acid residues, and the largest ones, of many more.

Twenty different amino acid residues are the building blocks of the protein. Since each position of the peptide chain can be occupied by any one of 20 different residues, the first question that arose was whether enzymes were indeed unique molecules or were, rather, agglomerations of molecules which were similar but not identical. Sanger showed that the protein insulin from a given species had a unique amino acid sequence (*3*). The same type of information has since been obtained for many proteins. The amino acid sequence in an enzyme is, therefore, precise—each position is occupied by one of the 20 amino acid residues, and the same sequence is repeated again and again as the enzyme is synthesized over and over by the cell's machinery.

The three-dimensional folding of the long linear arrays of amino acids also appears to be precisely defined. The very fact that an x-ray analysis has been resolved to the 1.5-angstrom level by Kendrew is sufficient to indicate that the three-dimensional folding is essentially identical for each molecule of myoglobin (*4*). Similar x-ray studies on enzymes, although less far advanced, indicate that enzymes too

The author is affiliated with the Biology Department of Brookhaven National Laboratory, Upton, N.Y., and with the Rockefeller Institute, New York. This article is based on material presented at the O. N. Smith lecture at Oklahoma State University, Stillwater.

have unique three-dimensional structures (*5*).

The x-ray diffraction data reveal that a globular protein looks like a piece of badly tangled rope. Not only is the sequence of the 200 or more amino acids in an average enzyme chain precisely defined, but—even more incredible—this chain is then folded in a complex way, so that it possesses a reproducible three-dimensional structure. A clue to the way this may come about has been obtained by Anfinsen and his co-workers (*6*), who believe that the interactions of the amino acid side chains of the protein determine the three-dimensional structure. In other words, the coded gene merely has to determine the linear sequence of amino acids, and the interactions of the side chains in this sequence will establish the proper three-dimensional geometry.

In a sense, this new knowledge of protein structure reinforces some of the theories derived from studies of enzyme action. From attempts to explain specificity and catalytic power, enzymologists concluded that a precise geometry of the enzyme surface must be a feature of the enzymatic process. The finding that enzymes do possess precisely machined structures therefore supports the theoretical conclusions.

## The Active Site

The early studies of kinetics of enzyme action by Michaelis and Menten showed that most enzyme reactions followed simple kinetics explainable by the reaction

$$E + S \rightarrow ES \rightarrow E + P \qquad (1)$$

In the first step, enzyme ($E$) and substrate ($S$) combine to give an enzyme substrate complex ($ES$), which then decomposes to give the products ($P$). Free enzyme is regenerated and is then available to react with another molecule of substrate. This process is repeated many times; in the case of many enzymes, thousands of molecules of substrates are converted into product per second per molecule of enzyme.

The substrate does not bind randomly to the enzyme but binds at specific sites—usually one or at most a few, per molecule of enzyme. The sites of binding and catalysis of the substrates have been called "active sites." They are highly specific, being able to distinguish between closely similar compounds. For example, urease hydrolyzes only urea and fumarase adds $H_2O$ to the double bond of fumaric acid only. Even compounds as similar to urea and fumaric acid as biuret and maleic acid are untouched by these two enzymes.

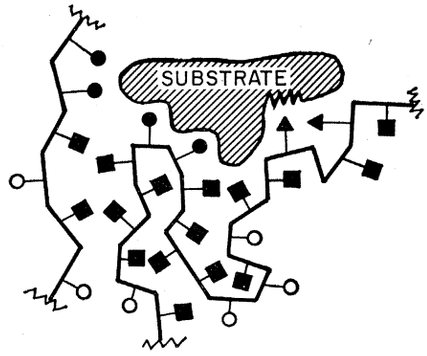


Fig. 1. A schematic representation of an active site. (Solid circles) "Contact" amino acid residues whose fit with substrate determines specificity; (triangles) catalytic residues acting on substrate bond, indicated by a jagged line; (open circles) nonessential residues on the surface; (squares) residues whose interaction with each other maintains the three-dimensional structure of the protein.

Most substrates are far smaller than the enzymes which act on them. In a few cases in which the enzymes are smaller—for example, ribonuclease—the enzymes act by attaching to a specific small portion of the large molecule, breaking a bond there, and then act again by attacking a second, similar portion of the very large molecule. Thus, the actual substrate for the enzyme action is usually a molecule of relatively low molecular weight or a small portion of a molecule of high molecular weight. Considerations of size alone would indicate therefore, that only a fraction of the protein could be in contact with the substrate. For example, appropriate molecular models show that a typical substrate of chymotrypsin, such as acetyl tyrosine ethyl ester, could be in contact with, at most, approximately 20 amino acid residues if these residues were free to adopt any position in space. In the actual protein, however, the residues are attached in a backbone structure which precludes complete freedom of movement, and therefore this number is probably too large. Since there are about 250 residues in this molecule, only a small fraction can be directly in contact with the substrate.

From such considerations a rough picture of an enzyme can be obtained. A schematic representation of an enzyme-substrate complex is shown in Fig. 1. From Fig. 1 it is clear that some amino acids are in direct contact with the substrate while others can play a role only indirectly. Some may play no role at all, at least as far as the enzyme action is concerned.

Identification of the role of each amino acid residue is the goal of these studies, and so far a wide variety of methods have been applied. Central among them are protein modification methods (*7*). These methods are based on chemical modification of amino acid side chains followed by examination of the properties of the modified proteins. Modification of side chains essential to activity should eliminate enzyme action, modifications which lead to unfolding of the protein should be detectable by measurements of protein structure, and so forth. Correlation of these results with results of studies of nonenzymatic models and of protein structure has led to significant advances in our understanding of enzyme action. In the following sections, illustrative examples from some of the most extensively studied enzymes are outlined.

## Ribonuclease

The first enzyme in which the complete amino acid sequence was determined was ribonuclease, a pancreatic enzyme which hydrolyzes nucleic acid. This classic work, carried out largely through the efforts of Hirs, Stein, and Moore (*8*), has placed this enzyme in the forefront of research on the relation of function to structure. The complete amino acid sequence, with cysteine residues in proper relation to each other for formation of S-S bridges, is shown in Fig. 2. Although such a picture cannot describe the three-dimensional structure of the protein, it establishes important limitations and provides a framework for the structure-function correlation.

Certain features of the three-dimensional structure have also emerged. The x-ray crystallographic study with resolution to the 4-angstrom level indicates that there is little, if any, alpha helix in the protein (*9*). Overall, the molecule is fairly compact, the size of the unit cell being approximately 30 by 30 by 48 angstroms (*10*). From chemical reactivities it is possible to say that the methionines in positions 13, 29, 30, and 79 are "buried" in the interior of the protein (*11*), and that the lysines in positions 1, 7, 31, 37, 41, and 104 are "exposed" on the surface (*12*).

Treatment of ribonuclease at *p*H 5.5

with iodoacetic acid alkylates the imidazole side chain of the histidine residues at positions 12 or 119 (*11, 13, 14*). Alkylation of either residue leads to loss in activity, and moreover, as Crestfield, Stein, and Moore have shown, alkylation of one residue prevents alkylation of the other residue (*14*). From this finding they have deduced that the two residues are very close to each other, perhaps no more than 5 angstroms apart. Hirs *et al.* (*15*) have shown that the amino group of lysine at position 41 can be arylated by fluorodinitrobenzene. In this case the activity of the protein is completely lost. These lysine and histidine residues can be kept from reacting with modifying reagents by means of the substrate, cytidylic acid. Thus, it appears that all three of these residues are at "the active site." The amino group of lysine would be protonated, and thus positively charged, at *p*H 7, and it might be involved in attracting the negatively charged substrate, ribonucleic acid, to the active site (*15*). Moreover, the *p*H-rate profile indicates the presence of two groups of *p*K approximately 7; one of these must be present as an acid, the other, as a base (*16*). These would appear to be the two histidine residues.

Deamidation of the glutamine residue at position 11 (*17*) and oxidation or deletion of the methionine residue at position 13 (*18*) lead to great loss in activity. Removal of the aspartic acid at position 121 and of the alanine at position 122 have also been shown to cause loss of activity (*19*). Whether these amino acids play a direct role in the enzyme action or merely help to maintain the three-dimensional structure of the molecule is not yet known.

On the other hand, modification of many regions of the protein can be carried out without serious effect on the activity of the enzyme. For example, Richards has shown that breaking of the covalent bond between residues 20 and 21 does not diminish the activity of the enzyme (*20*). Hofmann *et al.* (*21*) have synthesized a peptide containing amino acids 1 through 13, have added it to the peptide containing amino acids 21 through 124, and have found that these recombined fragments have an activity 70 percent that of the native protein. Since the activity is only slightly diminished by the removal of residues from positions 14 to 20, it is apparent that these residues do not play any vital role in the enzyme action. Treatment of the protein with carboxypeptidase removes the C-terminal valine without loss in activity; thus it appears that this residue is not required for enzyme action either.

The most significant feature of the evidence available so far is the implication that three positions widely separated in the sequence of the protein must be in close proximity in three-dimensional structure if the enzyme is to function. The amino acids involved are the two histidines and the lysine at positions 12, 119, and 41, respectively. The glutamine at position 11, the methionine at position 13, and the aspartic acid at position 121 appear to be involved in some way, although their roles are obscure at present. When a model of ribonuclease is made, it is possible to arrange the three essential residues in juxtaposition without introduring any serious strains in the protein model.

## Chymotrypsin

Chymotrypsin, a digestive enzyme of molecular weight 25,000, isolated from the pancreatic juices, is also in the forefront of studies on active sites. Although the complete sequence of amino acid residues in this molecule is not known, major portions of it have been determined (*22*). Its structure is shown schematically in Fig. 3; it consists of three peptide chains held together by disulfide bridges.

When the nerve gas diisopropyl fluorophosphate (DFP) is allowed to react with chymotrypsin, it phosphorylates the hydroxyl group of a single serine in the C chain rapidly and stoichiometrically, with deactivation of the enzyme (*23*). When the enzyme is treated with *p*-nitrophenyl acetate (*24*) at *p*H 5, the same serine is acylated

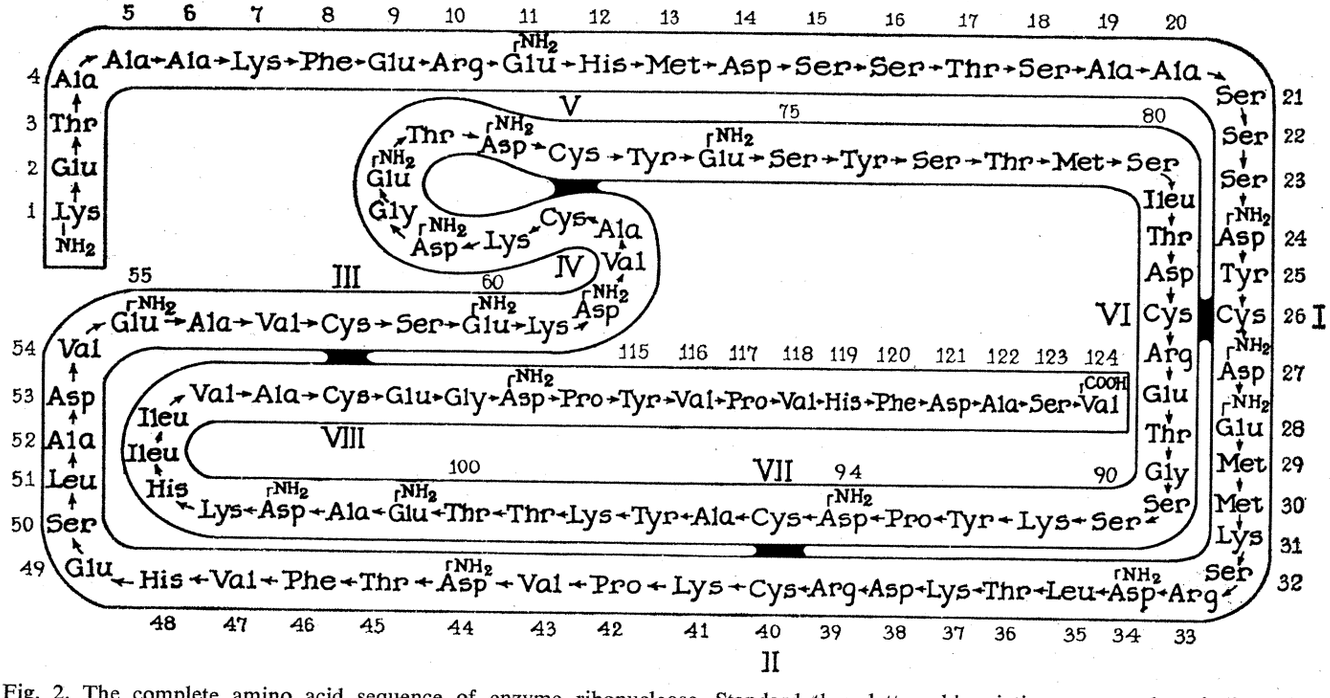


Fig. 2. The complete amino acid sequence of enzyme ribonuclease. Standard three-letter abbreviations are used to indicate individual amino acid residues. [Reproduced from the *Journal of Biological Chemistry*, with permission]

stoichiometrically. The acylated enzyme hydrolyzes to the original enzyme at *p*H 7. Classical substrates, such as N-acetyl-L-tyrosine ethyl ester, keep this serine from reacting with DFP or nitrophenyl acetate, presumably by covering up the active site.

One of the histidine residues in the B chain has also been implicated in the action of chymotrypsin. Weil and Seibles first indicated that histidine and tryptophan were photooxidized at rates that corresponded to loss of enzyme activity (*25*). Kinetic methods were developed in our laboratory by means of which it was established that histidine was the residue responsible for loss in activity and that modification of histidine produced a protein which had no detectable activity when assayed by the most sensitive techniques available (*26, 27*). Schoellman and Shaw prepared a compound which was similar to a typical substrate but which possessed an alkylating group ($COCH_2Br$) instead of the ester group. This compound alkylated histidine specifically, showing that this residue was also at or near "the active site" and, perhaps even more significantly, near the substrate carboxyl group where the bond-breaking and bond-forming occurs (*28*).

Other residues have been modified also, but the evidence shows that none of these are directly involved in enzyme action. Oxidation of the methionine 3 residues from the active serine produces a chymotrypsin which has one-third the binding capacity of the native protein but is otherwise completely active (*27*). Modification of the methionine 15 residues from the active serine produces a partially active protein. All of the lysine residues and a number of tryptophan and tyrosine residues can also be modified without loss in enzyme activity.

## Serine Enzymes

A number of enzymes in addition to chymotrypsin are inactivated by reaction with the nerve gas DFP. In each of these cases, acid hydrolysis of the protein has yielded a single molecule of serine phosphate. The amino acid sequences around this serine have been determined and are listed in Table 1. A recurring sequence of a dicarboxylic amino acid, serine, and a small neutral amino acid is seen to occur in most of the esterases and in alkaline phosphatase, suggesting that the presence of

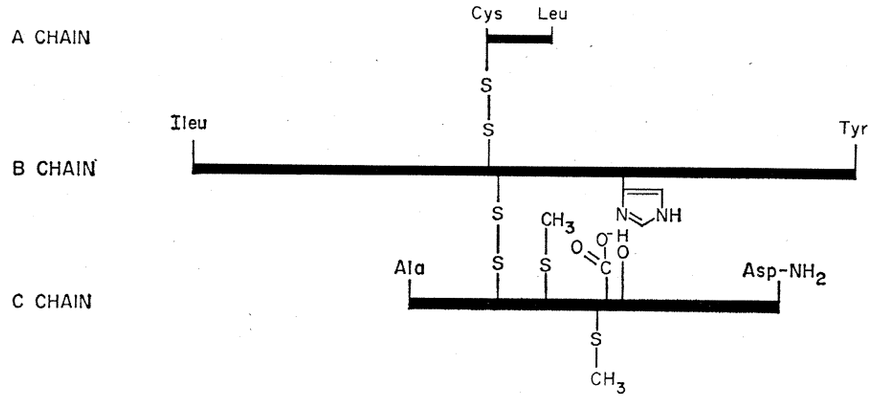


Fig. 3. Schematic representation of the enzyme chymotrypsin. Lengths of the black lines are proportional to lengths of the peptide chains. The carboxyl group of aspartic acid is adjacent to serine hydroxyl in the peptide sequence. Histidine in the interior of the *B* chain is apparently near the reactive serine of the *C* chain.

these residues at so many active sites may possibly be significant. However, different sequences are present in phosphoglucomutase and subtilisin, which have properties similar to those of the other enzymes. At the moment, therefore, the evidence allows only the conclusion that the presence of serine in so many enzymes probably indicates a common feature in the action of these diverse species.

## Role of the Serine Residue

The unusual reactivity of serine and its labeling in so many enzymes has focused attention on the role of this residue. That serine is at or near the active site in these enzymes is little disputed, but its function in the catalytic process is still controversial, despite the fact that 14 years have passed since it was demonstrated that DFP reacts with serine, and 7 years, since publication of Gutfreund and Sturtevant's proposed acylation-deacylation mechanism (*29*). The arguments have been summarized elsewhere (*27, 30, 31*), and I will not give them here in detail.

Briefly stated, the arguments for serine's having a central role in the catalytic process are (i) that its modification by agents such as DFP leads to loss in activity, (ii) that modification can be prevented or retarded by the presence of substrates, and (iii) that the kinetics of acylation and deacylation can be explained on the basis of an acyl-enzyme intermediate. Those who argue against its having a central role accept the idea that serine is near the active center of the enzyme but assert (i) that it is an alternate acceptor, competing with the true acceptor—for example, water or glucose—and hence plays no compulsory role in the catalytic process; (ii) that its modification by reagents prevents enzyme action by sterically preventing access to the true catalytic residue—for example, histidine; and (iii) that the kinetics of reaction in many cases are difficult or impossible to reconcile with an acyl-enzyme intermediate. The controversy emphasizes two important facets of this work: (i) the difficulty in establishing a precise role for a particular amino acid; (ii) the importance of resolving the role of serine, since an understanding of the unusual reactivity of this alcoholic hydroxyl group may lead to a general understanding of enzyme action.

William White, David Strumeyer, and I recently completed experiments in which we attempted to resolve this dilemma. Since previous modification studies in which the alcoholic group of serine ($CH\text{-}CH_2OH$) was converted to a bulkier group ($CH\text{-}CH_2OR$) were controversial because of the steric access argument, we decided to convert this residue to a smaller group ($C{=}CH_2$), by converting the active serine in chymotrypsin to a dehydroalanine residue.

Table 1. Enzymes with serine at the active site.

| Enzyme | Amino acid sequence adjacent to reactive serine | References |
|---|---|---|
| Chymotrypsin | Asp Ser Gly | (*49*) |
| Trypsin | Asp Ser Gly | (*50*) |
| Thrombin | Asp Ser Gly | (*51*) |
| Elastase | Asp Ser Gly | (*52*) |
| Liver aliesterase | Glu Ser Ala | (*53*) |
| Pseudocholinesterase | Glu Ser Ala | (*54*) |
| Alkaline phosphatase | Glu Ser Ala | (*55*) |
| Subtilisin | Thr Ser Met | (*56*) |
| Phosphoglucomutase | Ala Ser His | (*57*) |

This was accomplished by treating chymotrypsin with $C^{14}$-labeled tosyl chloride, forming mono-tosyl-chymotrypsin with the tosyl group on the active serine, and then removing the tosyl group by a base catalyzed elimination reaction. Control experiments in which tosyl chloride was omitted showed no loss of activity, but complete loss of activity was observed in the procedure in which the active serine was converted to a dehydroalanine residue. Since the alcoholic group ($CH_2OH$) of the serine has been replaced by a smaller group ($=CH_2$) in these experiments, the alternate-acceptor and steric-access arguments against a central role for serine seem to have been eliminated. It seems, therefore, that the alcoholic group of serine plays a central role as a key catalytic group in chymotrypsin and by inference is also a key residue in the other "serine enzymes."

## Carboxypeptidase

The modifications described above have emphasized experiments seeking to explain the catalytic power of the enzyme. The specificity of enzymes also plays an important part in their physiological function. An illustrative example of modification studies which throw light on specificity properties is provided by carboxypeptidase A, an enzyme of molecular weight 34,000 containing one atom of zinc, which hydrolyzes both esters and peptides. Vallee found that acetylation of the enzyme with acetyl imidazole destroyed the peptidase activity, while the esterase activity was increased seven-fold (*32*). Chemical analyses have shown that two tyrosine residues are acetylated and that the deacetylation of these tyrosines restores activity to the protein. A similar change in specificity has been observed when the atom of some other metal is substituted for the zinc atom. For example, the substitution of cobalt for zinc enhances the esterase activity and decreases the peptidase activity. The zinc atom is apparently not involved in the binding of substrate to the enzyme, since loss of this zinc atom does not affect the affinity of substrate for the enzyme surface, even though the remaining zinc-free protein is incapable of catalyzing hydrolyses. $\beta$-Phenylpropionate, a competitive inhibitor of carboxypeptidase A, protects the tyrosine residues against acetylation and also prevents deacetylation of the tyrosine residues by hydroxylamine.

## Modification Studies on Other Enzymes

Many other enzymes have been studied with amino acid reagents. Although none of these have been examined as extensively as ribonuclease and chymotrypsin have been, the results are similar to those described. For example, the sulfhydryl and amino groups of cysteine and lysine, respectively, appear to be essential in some enzymes.

Alkylation or mercuration of the sulfhydryl group in papain blocks enzyme action (*33*). An acetyl-enzyme can be isolated in the case of triosephosphate dehydrogenase, and the acylated group has been shown to be cysteine (*34*). A Schiff base appears to be formed between the lysine residue of aldolase and substrate during the action of that enzyme (*35*). Similar intermediates have also been identified in analogous hydrolyzing and condensing enzymes, suggesting roles for lysine and cysteine paralleling those of serine in the esterases.

The finding that many of the residues in a protein can be modified without loss in enzyme activity has also been demonstrated in many systems. Perhaps the most dramatic example is papain; Smith and his co-workers removed two-thirds of the amino acids of this molecule without change in the catalytic velocity or specificity of the enzyme (*36*). Other examples are beta amylase, in which all of the amino groups can be modified without essential loss of the activity of the enzyme, and phosphoglucomutase, in which photooxidation was shown to damage ten of the eleven surface methionine residues and six of the seven surface histidine residues without appreciable loss of enzyme activity.

Two general conclusions emerge from these many studies. (i) Other reactive groups may play roles similar to the roles of serine, histidine, and lysine residues in chymotrypsin and ribonuclease. (ii) Common patterns of behavior exist—that is, some residues are apparently essential to enzyme action, some residues can be modified with only partial change in catalytic velocity and specificity, and some residues play no part in enzyme action. Hence, it appears probable that the generalizations developed from the more completely investigated cases will apply to all enzymes.

Now that we have considered instances in which modifications of enzymes have led to an identification of amino acids involved in enzyme action, it is perhaps worth while to summarize the present status of our correlation of these results with the function of the enzyme in catalyzing the reaction of substrates to products. It should be emphasized that a field as vast and complex as enzymology cannot possibly be covered exhaustively in an article of this length. Many areas—the elegant kinetic investigations of Theorell and Chance, the entire field of isozymes, the role of metals and so forth—are omitted, not because such studies are less important but only because the illustrative rather than the comprehensive approach seems most appropriate here.

## Basic Nature of the Catalytic Process

In the early studies of enzyme action, enzymes were found to possess properties of specificity and catalytic power so extraordinary that it was supposed that some basically nonchemical mechanism must be responsible. Because the properties of enzymes were so different from the properties of ordinary chemical catalysts it seemed probable that some totally new theory would be needed to explain enzyme action. For example, it was proposed that enzymes had the properties of semiconductors and thus were able to cause chemical changes at velocities which were qualitatively different from the methods of action of ordinary chemical catalysts such as acids and bases. Such a conclusion was further supported by the observation that some enzymes transformed substrates into products which were completely unexpected from analogy with usual chemical transformations.

As more experimental information accumulates, the assumption that the enzyme is behaving like a physical tool such as a semiconductor seems increasingly unwarranted. In the first place, in all the enzyme experiments in which unusual products have been formed, formation of these products has been shown to be the result of the sequential action of several enzymatic species in an impure enzyme preparation. When the enzymes were separated it was found that the individual reactions had familiar counterparts in the chemical laboratory. In many cases the enzymatic reactions can be precisely duplicated in the test tube. In other cases the enzymes catalyze reactions with higher

yield and fewer side products than are obtained from the nonenzymatic process, but in each case a simple chemical analog could be found. Excellent reviews of model systems have been written by Westheimer and by Bender and Breslow (*37*).

Secondly, the modification studies have demonstrated that the total structure of the protein is not necessary for enzyme action. As discussed earlier, large sections of some enzymes can be removed without loss of enzyme activity, and these sections are frequently adjacent to amino acids identified as part of the active site. Any electronic-network theory would seem to require that the region immediately adjacent to the reactive residues be maintained intact. Hence, it appears increasingly clear that enzyme action must result from chemical catalysis in which mechanisms yet to be delineated by physical organic chemistry play a fundamental role.

## Comparison of Enzymic and Nonenzymic Velocities

The conclusion that enzymes are organic catalysts introduces an anomaly of its own. The velocities at which enzymatic reactions occur are extraordinarily more rapid than their nonenzymatic analogs. For example, a mixture of amino acids can be added to a solution of starch and no detectable hydrolysis will occur. However, beta amylase (containing the same amino acids) can hydrolyze 4000 glycosydic bonds per second per molecule of enzyme.

A possible explanation for the unusual reactivity of the amino acid side chains is shown schematically in Fig 4, which illustrates a reaction between two substrates, *A* and *B*, catalyzed by three amino acid side chains, *R*, *S*, and *T*. It is apparent from the simple chemical analogs that the orientation of *A* and *B* relative to each other and of *R*, *S*, and *T* relative to *A* and *B* will be crucial. Yet, the probability of a five-body oriented collision in aqueous solutions is extremely remote. If, however, the three catalysts are held in a fixed and appropriate position because of the three-dimensional structure of the enzyme, and if the substrates *A* and *B* are absorbed in a precise geometrical relationship to each other and to the catalytic groups, the probabilities of reaction are enormously increased.

Although this idea is qualitatively satisfying, it is extremely important to be able to calculate the degree of acceleration which might be expected from the enzyme's ability to juxtapose and orient substrates and catalysts. Until recently, no method of calculation was available, but a formula for such a calculation has now been derived, and its application has led to two interesting results (*38*). The first of these is that a complex of the sort shown in Fig. 4 can indeed provide enormous velocities of acceleration. In a case where there are three catalysts and two substrates, acceleration by a factor of $10^{23}$ might be expected. This acceleration is more than adequate to explain some enzymatic velocities. This result is consistent with the findings of Bender, Bruice, and others that intramolecular rates can be made enormously higher than the intermolecular rates of analogous reactions (*39*). In the case of the complex of Fig. 3, the simple intramolecular compounds are presumably providing the kind of proximity and orientation effects which occur in the hypothetical model of Fig. 4.

On the other hand, the formula reveals that in many cases, even after the proximity and orientation advantages of such an active site have been calculated, the enzymatic velocity is still too high to be accounted for. In the case of beta amylase, for example, the enzymatic reaction velocity is $10^7$ times the rate calculated with values for the best nonenzymatic catalyst known and with orientation and juxtaposition assumed to be optimum (*38*). Similar ratios have been found for other reactions.

It is clear that orientation and juxtaposition can account for very large accelerations over nonenzymic velocities and that they probably make important contributions to enzymatic catalysis. They are not, however, of themselves sufficient to explain the enormous velocities of enzyme-catalyzed reactions. The solution to this puzzle has not been found and is one of the major goals of current research in enzymology.

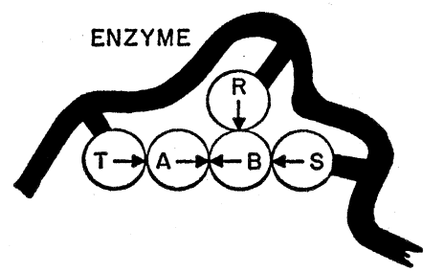


Fig. 4. A schematic representation of an active site, showing reaction between substrates *A* and *B* catalyzed by catalytic side chains *R*, *S*, and *T*. Orientation of the side chains is indicated by arrows. Orientation and juxtaposition are possible because of the structure of the protein and because of the attraction of the substrate to the active site.

## Covalent and Noncovalent Intermediates

Since the time of Michaelis and Menten it has been thought that the enzyme forms a complex with substrate or substrates during the catalytic action. However, the specific role of the enzyme is not delineated in this kinetic picture. One question that may be asked is whether the enzyme forms a covalent bond to substrates or whether it acts by polarizing electrons in the substrate to accelerate reaction.

One of the most important developments in this regard was the finding of Doudoroff, Barker, and Hassid (*40*) that sucrose phosphorylase catalyzed an exchange of phosphate between glucose-1-phosphate and inorganic phosphate. They explained this exchange reaction by postulating that the enzyme forms an enzyme-glucose intermediate which, in the course of a normal enzymatic reaction, reacts with fructose to give sucrose. In the absence of the fructose, the enzyme-glucose intermediate can revert to the glucose-1-phosphate by reaction with inorganic phosphate. In the presence of phosphorus-32 this reversion can be measured. Although the isolation of a glucose enzyme from sucrose phosphorylase has not yet been achieved, this concept that an amino acid residue of the enzyme could form a covalent bond with the substrate in the course of enzyme action was a landmark in the understanding of enzyme action, and it has been supported by findings with other enzymes—for example, chymotrypsin and phosphoglucomutase.

The mode of action of enzymes in this mechanistic category, which includes sucrose phosphorylase and chymotrypsin, was called the double-displacement mechanism to indicate that two successive attacks occurred, the first by the enzyme on one substrate to produce a covalent intermediate, the second by the second substrate on the intermediate to produce a final product (*41*). By correlating stereochemical, specificity, and exchange studies it was possible to deduce that other mecha-

nisms, in which no covalent intermediate is formed between enzyme and substrate, also exist (*41*). In these cases the second substrate makes a direct attack on the first substrate. In recent years many other enzymic reactions have been found to fit into these categories.

The existence of a single-displacement mechanism shows that enzymes can catalyze chemical transformations by affecting the environment around the bonds in a substrate, probably by polarizing the electron distribution. In the double-displacement reactions the enzyme must also polarize the electrons, since the alcoholic group of serine, for example, does not appear to be of itself an especially reactive residue, but in this case the enzyme also serves as a stoichiometric reactant during the enzyme action.

### Enzyme Flexibility and Specificity Mechanisms

A necessary feature of enzymes is their specificity. The catalytic control of the cellular environment depends on the enzyme's ability to guide compounds through an intricate series of steps to final products without excessive losses due to side reaction and without the accumulation of intermediates in toxic amounts. The high specificity of enzymes was originally explained by Fischer (*42*) on the basis of the template hypothesis. According to this theory a reactive substrate is adsorbed to the surface of an enzyme in such a way that the bond to be broken is in juxtaposition with the catalytic group on the enzyme surface. The fit of substrate and enzyme, like a key in a lock, is required for this juxtaposition.

This classical specificity mechanism has recently been shown to require an important modification. The concept that there is a fit between a negative and a positive, as in a jigsaw puzzle or a template, was questioned on the basis of kinetic evidence (*43*). Analysis of the data led to the theory (*43*) that the specificity properties of a number of enzymes demand a flexible active site in which the substrate induces a conformation change leading to an appropriate alignment of catalytic groups. An example of such a flexible-site specificity mechanism is illustrated in Fig. 5. This hypothesis predicts that changes in protein structure will accompany the binding of the substrate to the enzyme surface, and these predictions have been given strong support in studies of these conformation changes in our own and in other laboratories (*38, 43, 44*).

The requirement for a flexible active site is in some ways unpleasant, since it makes certain experimental problems more formidable. In other ways the concept makes enzymology easier, since it has already made it possible to explain such diverse phenomena as the ability of many enzymes to exclude water from the active site, the phenomenon of noncompetitive inhibition, and the phenomenon of sequential absorption of substrates, all of which are difficult or impossible to explain on the basis of the template hypothesis (*43*). This concept of a flexible active site has also been applied, by Pardee (*45*) and Monod (*46*), to the problem of feedback inhibition in bacterial systems, and the accumulating evidence suggests that a flexible site must indeed be a crucial factor in these control mechanisms, allowing an organism to stop producing a substrate once an adequate supply is achieved. It seems quite possible that hormone action may also proceed by means of a flexible site (*43, 44*). Thus, the concept of a flexible site which allows the enzyme to achieve a high level of specificity in an individual step apparently may also be particularly useful in elucidating the mechanism of control of enzymatic processes involving many enzymes.

Another important feature of the specificity process is the correlation of

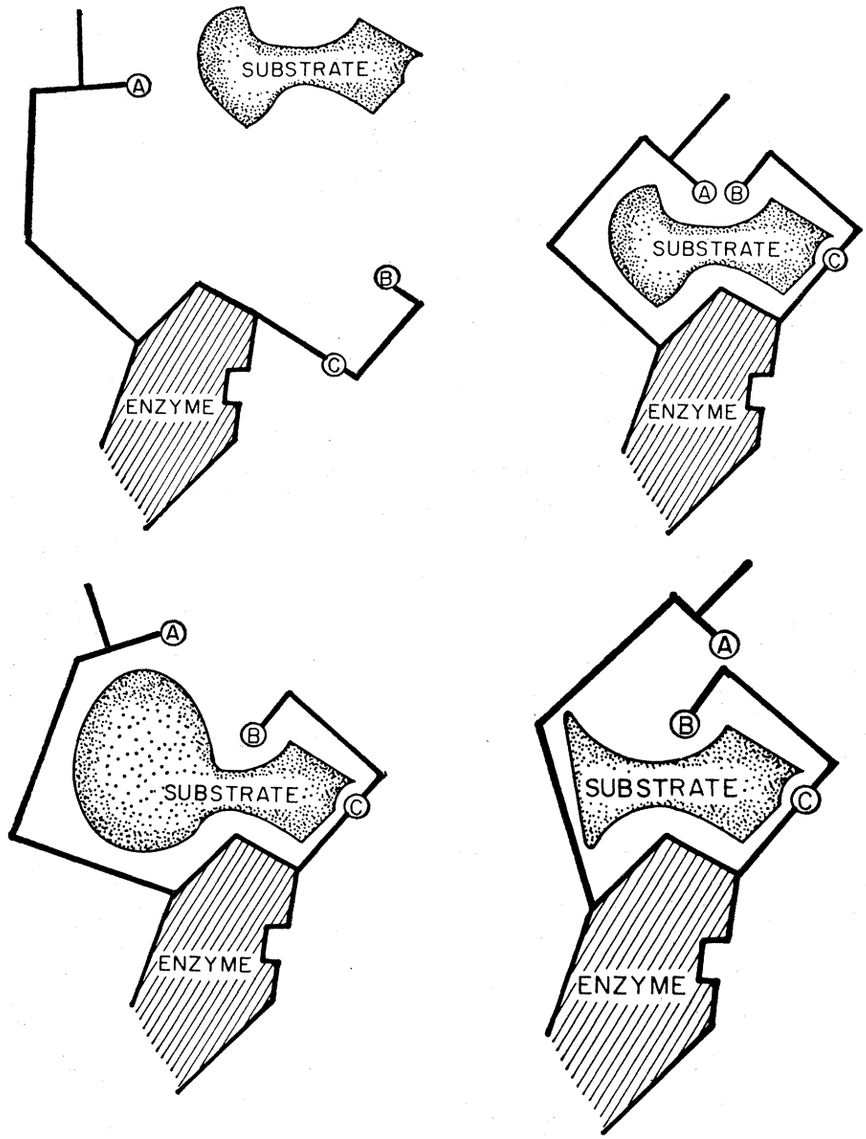


Fig. 5. Schematic representation of a flexible active site. Substrate binding induces proper alignment of catalytic groups *A* and *B* so that reaction ensues (top row). Compounds which are either too large or too small are bound, but fail to cause proper alignment of catalytic groups, hence fail to react (bottom row).

the structure of the substrate with the kinetic constants of the enzyme action. The concept of nonproductive complexes, for example, has been used effectively by Niemann (*47*) to explain some features of the specificity pattern of chymotrypsin. Although such studies do not yet make possible a detailed explanation in terms of protein structure, they provide fundamental data on which such a correlation must depend.

## Summary and Conclusions

This brief survey of some of the highlights in the correlation of enzyme structure and function indicates the considerable progress that has been made and also the important problems that are still unsolved. Protein reagents and x-ray crystallography are among the powerful tools which have changed our description of proteins from cumulus-cloud-type drawings to detailed pictures involving the precise positioning of atoms. Physical chemistry has also proceeded to a stage in which catalysis can be explained in terms of acids and bases, electron donors and acceptors. Finally, in the modification studies a correlation is beginning to be made between specific amino acid residues in the protein and their possible roles in polarizing electrons in the substrate molecules.

Of the conclusions already obtained, one of the most important is that the concept of a small "active site," which was deduced from simple geometric considerations, seems to be strongly supported by the modification studies. Not only are vast regions of the protein unnecessary for enzyme action but amino acid residues, one or two residues away from essential residues, can be modified with little or no change in the enzyme action. This conclusion has two important corollaries in regard to enzyme action.

1) A relatively few amino acids must be directly involved in the catalytic process. If all the amino acids in a region were essential, it would be possible to cluster these around the substrate so that a concerted action of many residues would be conceivable. Since residues distant from each other in the sequence must be brought together, the geometrical possibilities are far more limited. The fact that amino acid residues adjacent in sequence to these essential residues do not contribute makes it probable that the "contact" amino acids involved in catalysis must be very few in number—less than five would be a guess, but a reasonable one for a simple transfer reaction.

2) Electronic theories originating from the enigmatic properties of proteins and their high molecular weights seem mortally wounded by these modification studies. The observation that enzyme-catalyzed reactions have close similarities to well-known organic analogs supports this view and leads to the conclusion that enzymes must act through mechanisms paralleling the classical organic mechanisms of the chemical laboratory.

These accomplishments, however, lead to one of the major unsolved questions concerning enzyme action—how enzymes achieve such great catalytic powers. Although chemical models are available for each enzymatic reaction, the model reactions proceed far more slowly even under optimum conditions. We have theoretical equations for comparing these rates, and we find that the binding of the substrate to the enzyme surface and the aligning of substrates and catalytic groups can in some instances explain part of the increased efficiency of the enzyme. These same equations, however, indicate that in some cases we are still far short of explaining the enzymatic velocity (short by factors as great as $10^8$). The very enormity of this factor suggests that new chemical theories will be needed to explain enzyme action.

Although our detailed understanding of enzyme specificity is less far advanced than our understanding of catalysis, some important features of the specificity process have been determined. To explain the enzyme's ability to discriminate between closely similar compounds, a "fit" between the substrate and a portion of the enzyme surface seems essential. However, it appears probable that this fit is not a static one in which a rigid positive "substrate" fits on a rigid "negative" template but, rather, is a dynamic interaction in which the substrate induces a structural change in the enzyme molecule, as a hand changes the shape of a glove. Assumption of such a flexible active site allows us to explain many of the specificity properties of enzymes that are not explainable by the template hypothesis. It also provides a framework for explaining the kinetics of activation and inhibition of enzymes by compounds not directly involved in the enzyme action, a phenomenon which may be of crucial importance in the action of regulators of enzyme systems.

The relation of structure to function also makes the findings from genetic studies more comprehensible. Clearly, a mutation which changes an essential amino acid will produce an inert enzyme, and this will be a lethal or highly deficient mutant if the enzyme itself plays a key role in the organism. A change in one of the amino acid residues maintaining the three-dimensional structure, however, can produce either an inert or a partially active enzyme. Moreover, since the tertiary structure is maintained by interactions between many side chains, a "back mutation" may involve a second change, which nevertheless restores the original three-dimensional structure of the enzyme. Finally, changes which occur in the nonessential regions of the enzyme will have no effect, and may appear as genetically inactive regions of the chromosome.

It seems probable that, in the decade which lies ahead, the three-dimensional structure of several enzymes will be determined. Then, correlation of these structures with the specificity and catalytic properties of the enzyme should reveal the way in which nature's catalysts perform their functions. The progress made so far indicates that this "moment of truth" may not be too far off, and that it may then be possible to use this knowledge to design synthetic enzymes for chemical and medical purposes (*48*).

### References and Notes


1. A. Payen and J. F. Persoz, *Ann. Chim. (Phys.)* **53**, 73 (1833).
2. G. B. Beadle and E. L. Tatum, *Am. J. Botany* **32**, 678 (1945).
3. H. Brown, F. Sanger, R. Kitai, *Biochem. J.* **60**, 556 (1955).
4. J. C. Kendrew, *Brookhaven Symp. Biol.* **15**, 216 (1962).
5. J. Kraut, S. Sieker, D. F. High, S. J. Freer, *Proc. Natl. Acad. Sci. U.S.* **48**, 1417 (1962).
6. C. B. Anfinsen, E. Haber, M. Sela, F. H. White, Jr., *ibid.* **47**, 1309 (1961).
7. D. E. Koshland, Jr., *Advan. Enzymol.* **22**, 45 (1960).
8. C. H. W. Hirs, W. H. Stein, S. Moore, *J. Biol. Chem.* **235**, 633 (1960); D. H. Spackman, W. H. Stein, S. Moore, *ibid.*, p. 648; D. G. Smyth, W. H. Stein, S. Moore, *ibid.* **238**, 227 (1963).
9. D. Harker, *Advan. Biol. Med. Phys.* **4**, 1 (1956).
10. ———, *Ann. N.Y. Acad. Sci.* **69**, 321 (1957).
11. H. G. Gundlach, W. H. Stein, S. Moore, *J. Biol. Chem.* **234**, 1754 (1959).
12. W. A. Klee and F. M. Richards, *ibid.* **221**, 113 (1957).
13. W. D. Stein and E. A. Barnard, *J. Mol. Biol.* **1**, 350 (1959).
14. A. M. Crestfield, W. H. Stein, S. Moore, *J. Biol. Chem.*, in press.
15. C. H. W. Hirs, M. Halmann, J. H. Kycia, in *Biological Structure and Function* (Academic Press, New York, 1962), vol. 1, p. 41.
16. D. G. Herries, *Biochem. Biophys. Res. Commun.* **3**, 666 (1960).
17. P. J. Vithayathil and F. M. Richards, *J. Biol. Chem.* **236**, 1680 (1961).

18. ———, *ibid.* **235**, 2343 (1960).
19. C. B. Anfinsen, *ibid.* **221**, 405 (1956); J. T. Potts, Jr., and D. M. Young, *Federation Proc.* **22**, 418 (1963).
20. F. M. Richards, *Compt. Rend. Trav. Lab. Carlsberg* **29**, 329 (1955).
21. K. Hofmann, F. Finn, W. Haas, M. J. Smithery, Y. Wolman, N. Yanaihara, *J. Am. Chem. Soc.* **85**, 833 (1963).
22. B. S. Hartley, *Brookhaven Symp. Biol.* **15**, 85 (1962).
23. A. K. Balls and E. F. Jansen, *Advan. Enzymol.* **13**, 321 (1952).
24. B. S. Hartley and B. A. Kilby, *Biochem. J.* **56**, 288 (1954).
25. L. Weil, S. James, A. R. Buchert, *Arch. Biochem. Biophys.* **46**, 266 (1953).
26. W. J. Ray, Jr., H. G. Latham, M. Katsoulis, D. E. Koshland, Jr., *J. Am. Chem. Soc.* **82**, 4743 (1960).
27. D. E. Koshland, Jr., D. H. Strumeyer, W. J. Ray, Jr., *Brookhaven Symp. Biol.* **15**, 101 (1962).
28. G. Schoellman and E. Shaw, *Biochemistry* **2**, 252 (1963).
29. H. Gutfreund and J. M. Sturtevant, *Proc. Natl. Acad. Sci. U.S.* **43**, 719 (1956).
30. M. Bender, *J. Am. Chem. Soc.* **84**, 2582 (1962).
31. M. Caplow and W. P. Jencks, *J. Biol. Chem.* **238**, PC1907 (1963); S. A. Bernhard, W. C. Coles, J. F. Nowell, *J. Am. Chem. Soc.* **82**, 3043 (1960).
32. B. L. Vallee, J. F. Riordan, R. T. Simpson, *Biochemistry* **2**, 616 (1963).
33. E. L. Smith, *J. Biol. Chem.* **233**, 1392 (1958).
34. I. Krimsky and E. Racker, *Science* **122**, 319 (1955); J. I. Harris, B. P. Meriwether, J. H. Park, *Nature* **198**, 154 (1963).
35. J. C. Speck, P. T. Rowley, B. L. Horecker, *J. Am. Chem. Soc.* **85**, 1012 (1963).
36. R. L. Hill and E. L. Smith, *Biochim. Biophys. Acta* **19**, 376 (1956).
37. F. H. Westheimer, *Advan. Enzymol.* **24**, 441 (1962); M. Bender and R. Breslow, in *Comprehensive Biochemistry*, M. Florkin and E. Stotz, Eds. (Elsevier, New York, 1962), vol. 2, p. 1.
38. D. E. Koshland, Jr., J. A. Yankeelov, Jr., J. A. Thoma, *Federation Proc.* **21**, 1031 (1962); D. E. Koshland, Jr., *J. Theoret. Biol.* **2**, 75 (1962).
39. T. E. Bruice, *Brookhaven Symp. Biol.* **15**, 52 (1962); M. L. Bender, Y. Chow, F. Chloupek, *J. Am. Chem. Soc.* **80**, 5380 (1958).
40. M. Doudoroff, H. A. Barker, W. Z. Hassid, *J. Biol. Chem.* **168**, 725 (1947).
41. D. E. Koshland, Jr., in *Mechanism of Enzyme Action*, W. D. McElroy and B. Glass, Eds. (Johns Hopkins Press, Baltimore, 1954), p. 608.
42. E. Fischer, *Chem. Ber.* **27**, 2985 (1894).
43. D. E. Koshland, Jr., *Proc. Natl. Acad. Sci. U.S.* **44**, 98 (1958); ———, *J. Cellular Comp. Physiol.* **54**, suppl. 1, 245 (1959); ———, in *Horizons in Biochemistry*, B. Pullman and M. Kasha, Eds. (Academic Press, New York, 1962), p. 265; J. A. Thoma and D. E. Koshland, Jr., *J. Am. Chem. Soc.* **82**, 3329 (1960); D. E. Koshland, Jr., *Cold Spring Harbor Symp. Quant. Biol.*, in press.
44. G. M. Tomkins, K. L. Yielding, J. Curran, *Proc. Natl. Acad. Sci. U.S.* **47**, 270 (1961); A. J. Samuels, T. Nikey, L. Noda, *ibid.*, p. 1992; M. Citri and N. Garber, *Biochem. Biophys. Res. Commun.* **4**, 143 (1961).
45. J. C. Gerhart and A. B. Pardee, *J. Biol. Chem.* **237**, 891 (1962).
46. J. Monod, J. Changeux, F. Jacob, *J. Mol. Biol.* **6**, 306 (1963).
47. G. E. Hein and C. Niemann, *J. Am. Chem. Soc.* **84**, 4495 (1963).
48. Part of this work was supported by the U.S. Atomic Energy Commission and part, by the National Science Foundation (grant NSF G-19379).
49. N. K. Schaffer, L. Simet, S. Harshman, R. R. Engle, R. W. Drisko, *J. Biol. Chem.* **225**, 197 (1957); F. Turba and H. G. Gundlach, *Biochem. Z.* **327**, 196 (1955).
50. G. H. Dixon, D. L. Kauffman, H. Neurath, *J. Biol. Chem.* **233**, 1373 (1958).
51. J. A. Gladner and K. Laki, *J. Am. Chem. Soc.* **80**, 1263 (1958).
52. M. A. Naughton, F. Sanger, B. S. Hartley, D. C. Shaw, *Biochem. J.* **77**, 149 (1960).
53. H. S. Jansz, C. H. Porthumus, J. A. Cohen, *Biochim. Biophys. Acta* **33**, 387, 396 (1959).
54. H. S. Jansz, D. Brons, N. C. P. J. Warringa, *ibid.* **34**, 573 (1959).
55. J. H. Schwartz, A. M. Crestfield, F. Lipmann, *Proc. Natl. Acad. Sci. U.S.* **49**, 722 (1963); C. Milstein, *Biochim. Biophys. Acta* **67**, 171 (1963).
56. F. Sanger and D. C. Shaw, *Nature* **187**, 872 (1960).
57. C. Milstein and F. Sanger, *Biochem. J.* **79**, 456 (1961).

# Organic Photochemistry

## The study of photochemical reactions provides new information on the excited states of molecules.

George S. Hammond and Nicholas J. Turro

At least one organic photoreaction, photosynthesis, is obviously older than man, and reports of laboratory studies of organic photochemistry are to be found in the very early chemical literature. However, the field has not been numbered among the very active areas of organic chemistry until recently (*1–4*). Several factors have contributed to a dramatic surge of interest in the subject. First, it has become evident that, potentially, photochemistry provides short routes for the synthesis of systems that are only available otherwise through long and tedious synthetic programs. The promise of the method is due in part to the fact that photochemistry provides a means for selective injection of large "doses" of energy into individual molecules or specific parts of a molecule in any given system. Thus it becomes possible to achieve, in one step, reactions which would have activation energies high enough to lead to general disruption of the system if they were attempted by thermal means. A second important feature of modern photochemistry is the availability of spectroscopic techniques (such as flash photolysis) which make possible the direct study of transient species involved in photoreactions. That the existence of such tools would attract workers interested in reaction mechanisms was almost inevitable. Many of the concepts which are entering the literature of organic photochemistry have, in fact, been well known to spectroscopists for many years (*5, 6*), but some kinds of information about excited states which are not readily accessible by spectroscopic techniques are now being discovered.

In a sense, spectroscopists and photochemists are allied in their attempts to answer the following questions:

1) What happens to the excitation pumped into a molecule by absorption of visible or ultraviolet light?

2) What are the energies, electronic distributions, and geometric structures of various excited states?

3) What are the chemical properties of excited states?

### The Manifold of Excited States

Most polyatomic molecules have a number of metastable excited electronic states. Promotion of molecules to these states is accomplished by absorption of visible or ultraviolet light. Virtually all such transitions may be described approximately in terms of the excitation of a single electron from some orbital which is occupied in the ground state to an orbital which is vacant in the ground state. However, the first-formed excited state may undergo radiationless

The authors are, respectively, professor of organic chemistry and National Science Foundation fellow at the California Institute of Technology, Pasadena.