Enzymatic Catalysts in Organic Synthesis

C.-H. WONG*

The synthetic value of enzymes is being increasingly recognized. With an understanding of enzyme-catalyzed reactions and the techniques now available for the low-cost production and rational alteration of enzymes and for the design of new enzymatic activities, biocatalytic synthesis should beome one of the most valuable synthetic methods. The fundamental as well as the practical issues of this rapidly growing area are described and its impact on the technology pertaining to synthesis is considered.

ENZYMES HAVE BEEN WIDELY USED AS CATALYSTS IN ORGANIC synthesis. The synthetic value of natural and also many unnatural enzymatic reactions has been demonstrated [for recent reviews, see (*1–4*)]. The phenomenal rate acceleration and stereoselectivity, together with the available techniques for the low-cost production of enzymes and the rational alteration of their properties, make enzymes attractive as synthetic catalysts. Environmental concerns and the regulatory constraints faced in the chemical and pharmaceutical industries have spurred the hope that biological methods may offer clean and mild synthetic processes. This field has also been stimulated by the exciting new discovery of catalytically active antibodies. This article reviews the fundamental as well as the practical aspects of enzymatic catalysis in synthetic organic chemistry, with particular focus on the use of cell-free enzymatic catalysts.

Transition-State Stabilization and Rate Acceleration

The origin of the catalytic power in enzymatic reactions has been an interesting subject for study [for general references, see (*5*) and (*6*)]. It is generally accepted that the rate acceleration is mainly due to transition-state stabilization by the enzyme. It seems fair to say that all types of enzymatic catalysis, such as acid-base catalysis, nucleophilic-electrophilic catalysis, and catalysis by approximation, strain, and distortion, are just the contributing factors that lead to reducing the transition-state energy (*7*). For example, a single substrate S can be converted to product P by both an enzyme-catalyzed and noncatalyzed reaction (Fig. 1). One can apply transition-state theory (*8*) to relate the first-order rate constants for the enzymatic (k_{cat}) and nonenzymatic (k) reactions to the corresponding equilibrium constants ($K^{\ddagger}_{cat}$ and $K^{\ddagger}$) for the formation of the transition-state complex, that is, $k_{cat}/k \simeq K^{\ddagger}_{cat}/K^{\ddagger}$. According to the thermodynamic cycle, these equilibrium constants are related to the dissociation constants for the transition state (K_T) and for the substrate (K_S), so that $K_S K^{\ddagger} = K_T K^{\ddagger}_{cat}$. This simple analysis concludes that the enzyme binds to the transition state $S^{\ddagger}$ more strongly than the ground state S by a factor approximately equal to the rate acceleration, that is, $k_{cat}/k \simeq K_S/K_T$.

The concept of transition-state binding has led to the development of transition-state analogs for use as enzyme inhibitors and for the identification of possible groups in transition-state binding (*9*). For example, replacement of the scissile amide linkage in a peptide substrate for the zinc protease thermolysin with a phosphonamidate linkage, a mimic of the tetrahedral transition state of the cleavage, generates a potent inhibitor of the enzyme (*10*). The x-ray crystal structure of the inhibitor-enzyme complex (*11*) suggests that the enzymatic functional groups interacting with the inhibitor are those involved in transition-state binding. Experimentally the inhibition constants of a series of structurally related phosphonamidate inhibitors correlate well with the second-order rate constants for the corresponding peptide substrates (*10*). Enzyme functional groups involved in binding and catalysis can also be identified by using site-directed mutagenesis together with kinetic analysis (*12*). Such developments have helped allow the rational design of drugs to inhibit specific enzymes associated with diseases and the rational alteration of enzyme specificity, stability, and catalysis.

The idea of transition-state binding also leads to an experimental approach toward the design and synthesis of immunogenic transition-state analogs for the hydrolysis of esters and peptides (*13*). For instance, the monoclonal antibodies prepared against the phosphonamidate antigens as described above should catalyze hydrolysis of the corresponding peptides. Catalytically active antibodies have been prepared with reported rate accelerations of four to six orders of magnitude for hydrolysis and pericyclic reactions (*14*). For applications in synthesis, the most interesting developments are the demonstrations of stereospecific transesterification, lipase-like hydrolysis, and aminolysis in water (*14*). Given that virtually any molecule can be used to elicit monoclonal antibodies, antibody catalysis may find use in any reaction, provided that a good and stable transition-state analog of that reaction can be designed, synthesized, and converted to an immunogenic substance. As indicated by Schultz (*15*), the interplay between chemistry and biology in the design of enzymatic catalysts is certainly one of the most fascinating developments in molecular science.

Selectivity

The major synthetic value of enzymes is their selectivity. Enzymes are selective not only in the type of reaction they catalyze, but also in their position of attack on a molecule. Chemoselective reaction of one of different functional groups in a molecule, regioselective

The author is a professor in the Department of Chemistry, Texas A&M University, College Station, TX 77843.

*Present address: Professor and Ernest W. Hahn Chair in Chemistry, Research Institute of Scripps Clinic, La Jolla, CA 92037.

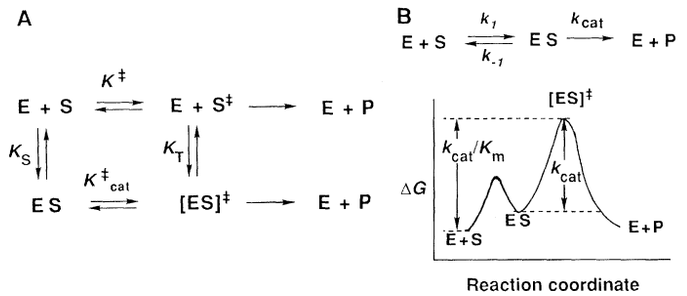


Fig. 1. (**A**) Thermodynamic cycle relating enzyme-catalyzed reaction of S to P; E, enzyme; ES, enzyme-substrate complex; $K^‡$, equilibrium constant for S and the transition state $S^‡$ in a noncatalyzed reaction; $K^‡_{cat}$, equilibrium constant for the ES complex and its transition state $[ES]^‡$ in the catalyzed reaction; and K_S and K_T, the respective dissociation constants for the ground-state and transition-state complexes. (**B**) The energy diagram of the enzymatic reaction; k_{cat} is the first-order rate constant for the chemical conversion of the Michaelis complex ES (or the turnover number of the enzyme), and k_{cat}/K_m is an apparent second-order rate constant for the enzyme reacting with the substrate. The upper limit of k_{cat}/K_m is k_1, the diffusive on rate of substance binding (10^8 to $10^9 M^{-1}$ s^{-1}). Under these conditions, k_{cat} is no longer rate limiting and the Michaelis-Menten kinetics changes to Briggs-Haldane kinetics, which may result in poor stereoselectivity.

reaction of one of the same or similar groups in a molecule, enantioselective reaction of one enantiomer of a racemic pair, and enantiotopic or diastereotopic group or face attack of a chiral or prochiral molecule are all well documented (*1–6*).

Virtually all such stereoselectivities originate from the energy difference in the enzyme–transition-state complex. For example, in an enantioselective transformation, it can be considered that two enantiomeric substrates (A and B), or two enantiotopic groups or faces, compete for the active site of the enzyme (Fig. 2). With the use of steady-state or Michaelis-Menten assumptions, the two reaction rates are $v_A = (k_{cat}/K_m)_A[E][A]$ and $v_B = (k_{cat}/K_m)_B[E][B]$ where K_m is the Michaelis constant; then $v_A/v_B = (k_{cat}/K_m)_A[A]/(k_{cat}/K_m)_B[B]$. This ratio of the specificity constants (k_{cat}/K_m) of two competing reactions is determined by the entantioselectivity of the reaction [see (*6*)]. The above kinetic constants can be related to the free-energy term $\Delta G^‡_A$, that is, $(k_{cat}/K_m)_A = \exp(\Delta G^‡_A/RT)$, and $(k_{cat}/K_m)_B = \exp(-\Delta G^‡_B/RT)$, where R is the gas constant and T is the absolute temperature. Thus $(k_{cat}/K_m)_A/(k_{cat}/K_m)_B = \exp(-\Delta\Delta G^‡/RT) = E$, where E is the enantioselectivity value. Enantioselection occurs because the enzyme and the substrates form diastereomeric transition states that differ in energy $\Delta\Delta G^‡$.

Equations and useful graphs for the quantitative treatment of enzyme-catalyzed enantioselective transformations have been reported (*16*). The expressions have been verified experimentally and possess predictive value in relating the parameters of the extent of conversion (C), the enantiomeric excess (ee), and the enantioselectivity value (E). Once the E value is determined, one can predict when to stop the reaction to maximize the ee of the product or the unreacted substrate. For example, in the irreversible kinetic resolution of a racemic substrate (A + B) with an enzyme, one can use Eq. 1 to predict the ee value of the unreacted substrate B or the recovered product P at certain degree of conversion, provided that the E value is known.

$$\frac{\ln[(1 - C)(1 - ee_B)]}{\ln[(1 - C)(1 + ee_B)]} = \frac{\ln[1 - C(1 + ee_P)]}{\ln[1 - C(1 - ee_P)]} = E \qquad (1)$$

The energy difference required to obtain a 99.9% ee product is relatively small, ~4.4 kcal/mol, and improvement of a given reaction with low enantioselectivity would not seem to be a difficult problem. Unfortunately, there is no simple solution to this problem. However, one can empirically change the reaction conditions (such as lowering or raising the temperature, or adding organic solvent) so that the energies of diastereomeric enzyme–transition-state complexes are altered favorably. This approach has been used in many esterase-catalyzed reactions (*17*). Another way to increase the magnitude of $\Delta\Delta G^‡$ is to modify the substrate by introducing different substituents (*18*).

One problem in enzymatic transformations, however, is the difficulty of predicting the stereochemistry of a reaction. Although the stereoselectivity in most enzymatic reactions is dictated by the particular tertiary structure of the catalyst, a changeover in stereoselectivity from one substrate to another is not uncommon (*19, 20*). For many synthetically useful but uncharacterized enzymes, this changeover represents a significant problem. The only solution at present is to develop a reliable active-site model for the enzyme. Based on the enzyme specificity results reported in the literature, one can use computer graphic analysis to develop such a model. Jones and co-workers have pioneered the development of horse liver alcohol dehydrogenase (*21*) and pig liver esterase (*22*) active-site models that are simple and reliable for use in the prediction of new reactions and in the rationalization of literature results.

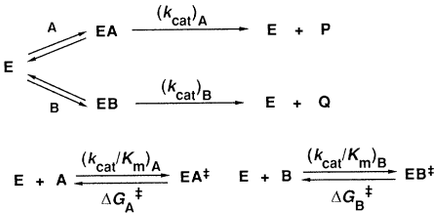
For enzymes with known x-ray structure, the use of site-directed mutagenesis and computer-assisted molecular modeling allows rational alteration of enzyme catalysis for specific use. For example, aspartate aminotransferase, a pyridoxal phosphate-dependent enzyme that catalyzes the transamination of Asp or Glu, was converted to lysine-arginine transaminase by replacing the active-site Arg with Asp acid (*23*). L-Lactate dehydrogenase was converted (by mutation of Gln^{102} to Arg) to L-malate dehydrogenase, which has double the enzyme activity of the natural malate dehydrogenase (*24*). Perhaps the most engineered enzyme is serine protease subtilisin. Almost every property of this enzyme, including catalysis, specificity, *p*H versus rate profile, and stability, has been altered (*22, 25*). These developments indicate that it is possible to engineer a known enzyme to act as a better catalyst for synthesis, because synthetic enzymatic reactions with unnatural substrates in unnatural environments are not optimized.

Although site-directed mutagenesis has become a useful tool for the modification of enzymes for specific use, traditional screening is still a very (and perhaps the most) economical and efficient way of finding the right enzyme for a desirable reaction. For example, a thermostable alcohol dehydrogenase from *Thermoanaerobium brockii* was found at a spring site in Yellowstone Park (*26*). A nitrile-hydrolyzing enzyme was found at an acrylnitrile plant. The enzyme from *Pseudomonas* is currently being used for the large-scale production of acrylamide from acrylnitrile and for the stereoselective hydrolysis of a variety of nitriles (*2*). A simple strategy for searching for a desired enzyme from microorganisms is to add the substrate of interest to the growth medium, and then select the microbial species capable of converting the substrate to the product of interest.

Enzyme Stability, Stabilization, and Reactor Configuration

Enzymes are intrinsically unstable in solution and can be inactivated by denaturation (caused by increased temperature or by an unfavorable *p*H or dielectric environment), dissociation of cofactors such as metals, and covalent changes such as oxidation and proteolysis. It is generally believed that the three-dimensional structure of a protein in a given environment is determined by its primary sequence (*27*) and is the structure with the lowest ΔG. Enzyme inactivation, however, could occur irreversibly through reversible

Fig. 2. Enzyme-catalyzed stereoselective transformations.



partial denaturation followed by irreversible changes. Thermal denaturation is the most studied mode of enzyme inactivation. Thermophilic enzymes usually differ from their mesophilic counterparts by only small changes in primary structures, whereas the three-dimensional structures of such enzymes are essentially the same (*28*). Mesophilic enzymes usually retain their native structures in aqueous solution at temperatures below 40°C, whereas for the thermophilic enzymes the limit is between 60° to 70°C. This difference corresponds to an increase in thermostability by 5 to 7 kcal/mol. Free-energy changes of this order can be derived from two to three additional salt bridges, hydrogen bonds, or aromatic-aromatic interactions (*29*). Mesophilic enzymes in principle can be made thermally stable by introducing additional binding forces. The techniques of site-directed mutagenesis, chemical cross-linking, and immobilization have been used to make enzymes more compact to prevent denaturation. Of course, such modifications should not perturb the active conformation of the enzyme. Immobilization is presently the most commonly used technique for enzyme stabilization (*30*). The procedures generally involve covalent or noncovalent attachment of enzymes to a support, and also the enzyme to be recovered for later use. Cross-linking and entrapment or encapsulation are also used. However, there is no general procedure available for immobilization of every enzyme, and substantial trial and error are required to find the best method. Functionalized ceramics, such as those treated with 3-aminopropyl triethoxysilane and the cross-linked copolymer of acrylamide and acryloxysuccinimide or epoxide, are commonly used as solid supports for covalent immobilization. Ion-exchange resin, glass beads, and XAD-8 are often used for adsorption of enzymes to be used in organic solvent or biphasic systems.

With regard to reactor configuration, batch and column hollow fibers are becoming attractive alternatives. In these systems, enzymes may be immobilized or membrane enclosed without attachment to a support.

Cofactor Regeneration

A number of synthetically useful enzymatic reactions require cofactors such as adenosine triphosphate (ATP), nicotinamide adenine dinucleotide phosphate ($NADP^+$), acetyl coenzyme A, *S*-adenosylmethionine, and 3′-phosphoadenosine-5′-phosphosulfate (PAPS). These cofactors are too expensive to be used as stoichiometric reagents. Regeneration of the cofactors from their reaction byproducts is thus required to make the process economical. Cofactor regeneration can also reduce the cost of the synthesis by (*31*): (i) influencing the position of equilibrium, that is, a thermodynamically unfavorable reaction can be driven by coupling with a favorable cofactor regeneration reaction; (ii) preventing the accumulation of cofactor by-product that may inhibit the forward process; (iii) eliminating the need for stoichiometric quantities of cofactors and thus simplifying the reaction workup; and (iv) increasing enantioselectivity relative to stoichiometric reactions (*32*).

The cofactor ATP and other nucleoside triphosphates have been used in selective enzymatic phosphorylations; the cofactor by-products in such reactions are nucleoside di- or monophosphates that must be converted to the corresponding triphosphates with acetyl phosphate or phosphoenol pyruvate. Large-scale enzymatic synthesis of organic phosphates was not realized until Whitesides *et al.* developed practical methods for the synthesis of phosphorylating reagents for use in regeneration of nucleoside triphosphates (*1*). This significant contribution has made possible the practical syntheses of sugars, nucleotides, oligosaccharides, and related substances (*33*). A general strategy useful for the synthesis of oligosaccharides is shown in Fig. 3.

The nicotinamide cofactors are involved in enzymatic oxidoreductions. Several practical methods are available for regeneration of NADH (*31*); these include formate–formate dehydrogenase, glucose–glucose dehydrogenase, and hydrogen-hydrogenase. For regeneration of NADPH from $NADP^+$, glucose–glucose dehydrogenase and isopropanol–*Thermoanaerobium brockii* alcohol dehydrogenase (TADH) are considered the most uselful. For regeneration of NAD^+ or $NADP^+$ from the reduced forms, the systems based on α-ketoglutarate–glutamic dehydrogenase and flavin mononucleotide (FMN)–FMN reductase are perhaps the most useful. Direct electrochemical regeneration of NAD^+ or $NADP^+$ through oxidation is promising; however, because of the problem of regioselectivity, electrochemical and other nonenzymatic regenerations of NADPH still require substantial development. When a nicotinamide cofactor–dependent enzyme is used for synthesis, regeneration of the cofactor can be catalyzed by a second enzyme or by the same enzyme used in the synthesis, provided that the overall equilibrium is favorable. The TADH-catalyzed reduction of ketones, coupled with the enzyme-catalyzed oxidation of isopropanol, is a typical example of the one-enzyme system (*26*). When a second

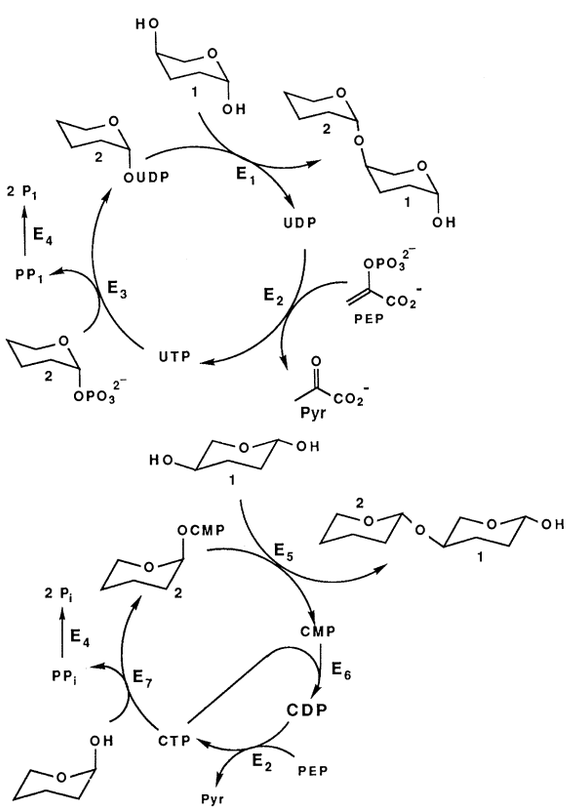
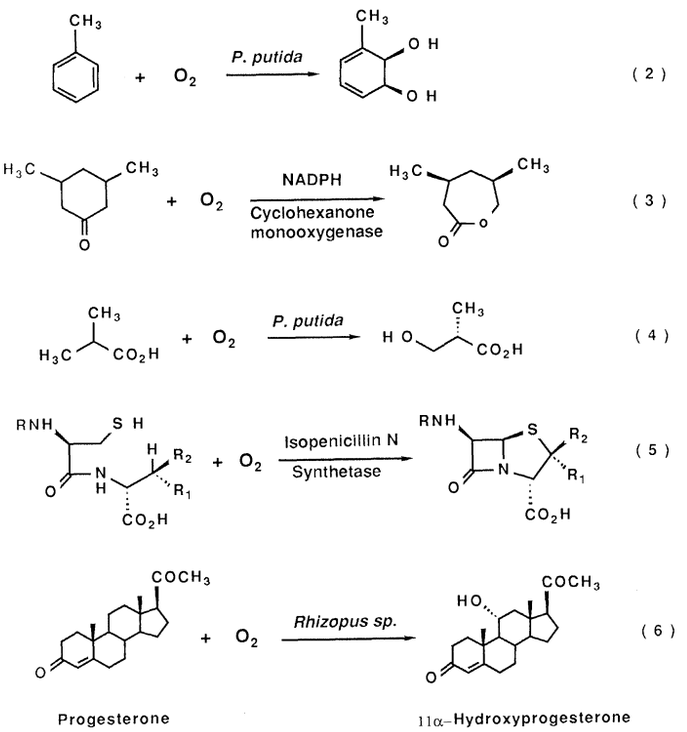


Fig. 3. Cofactor regeneration in controlled synthesis of oligosaccharides. For simplicity, only the glycosidic configurations are shown; E_1, glycosyl transferase specific for UDP sugars; E_2, pyruvate kinase; E_3, UDP-sugar pyrophosphorylase; E_4, pyrophosphatase; E_5, glycosyl transferase specific for CMP sugars; E_6, adenylate kinase; and E_7, CMP-sugar synthetase. UDP, uridine 5′-triphosphate; CMP, cytidine 5′-phosphate.

enzyme is used for cofactor regeneration, the regeneration system, in some instances, can be synthetically useful if the overall reaction is favorable and the products are easily separated. This idea was illustrated in a one-pot biphasic synthesis of a lactone and an amino acid product in which the NAD^+-dependent oxidation of a meso-diol was linked with the NADH-dependent reductive amination (Fig. 4). In the reaction, the lactone was extracted into the organic phase (hexane) while the amino acid product was retained in the aqueous phase, thus simplifying product separation (*34*). Furthermore, product inhibition of horse liver alcohol dehydrogenase caused by the lactone was minimized and thus the overall yield was increased.

The most highly developed large-scale process for NADH regeneration is that based on the use of formate dehydrogenase developed by Wandrey and Kula and co-workers (*35*) for production of amino acids in a membrane reactor. The cofactor NAD^+ was attached to polyethylene glycol to prevent leakage from the reactor.

Although nicotinamide cofactor–dependent enzymes are useful in synthesis, many other oxidoreductases use NADPH as an indirect reducing agent. Many mono- and dioxygenases represent this class of enzymes, which are able to activate molecular oxygen and insert oxygen stereoselectively into inactive molecules such as alkanes, aromatics, and olefins. For synthetic transformations, whole cells instead of free enzymes are used because of the instability of the enzymes. Some oxygenase-catalyzed transformations are illustrated in Eqs. 2 through 6 (*36*).



In Eq. 5, R = (*S*)-$HO_2CCH(NH_2)(CH_2)_3CO$-, $C_6H_5CH_2CO$-, or *m*-carboxyphenyl-acetyl, R_1 = methyl, and R_2 = methyl or methoxy.

Enzyme Catalysis in Organic Solvents

Enzymes were once thought to work only in aqueous solutions. Organic solvents were only used, if necessary, to change the solubility of substrates or products, or to influence the equilibrium. The enzymes used in water–water-miscible organic solvents, water–water-immiscible organic solvents, and reverse micelles still work within a pool of water. In 1966, Dastoli and Price first observed enzymatic activities of crystalline chymotrypsin and xanthine oxidase (*37*) in anhydrous nonpolar organic solvents. Since then, several enzymatic transformations in nonpolar organic solvents had been reported (*3, 4, 38*). Recently, Klibanov and Zaks have taken this a step further to examine the role of water in enzymatic reactions in a number of anhydrous polar and nonpolar organic media and concluded that in general enzymes only need a thin layer of water on the protein surface to retain their catalytically active conformation in anhydrous media (*3, 39*). The most adequate nonaqueous media are hydrophobic solvents that do not strip the essential water from enzymes. Water-immiscible solvents containing water below the solubility limit (about 0.02 to 10% by weight depending on the solvent used) are suitable for dry enzymes (crystalline or lyophilized powder) to be catalytically active. Within this range of water content, the enzymatic activity in an appropriate organic solvent can be optimized and, in some cases, is comparable to that in aqueous solution, and the catalysis follows Michaelis-Menten kinetics. Higher thermostability and different substrate specificities have been reported, presumably because enzymes are conformationally less flexible in nonaqueous media. Furthermore, a dramatic change of stereoselectivity in water versus hydrophobic organic solvents has also been observed. For example, the enantioselectivity of subtilisin-catalyzed hydrolysis of Ac-L-Phe-OEt versus Ac-D-Phe-OEt (Ac, acyl, and Et, ethyl) in water is about 2300 [determined by $(k_{cat}/K_m)_L/(K_{cat}/K_m)_D$]. In butyl ether, the enantioselectivity of the enzyme-catalyzed transesterification of the esters with *n*-propanol, however, was 7.3 (*40*). Klibanov and co-workers (*40*) explained this

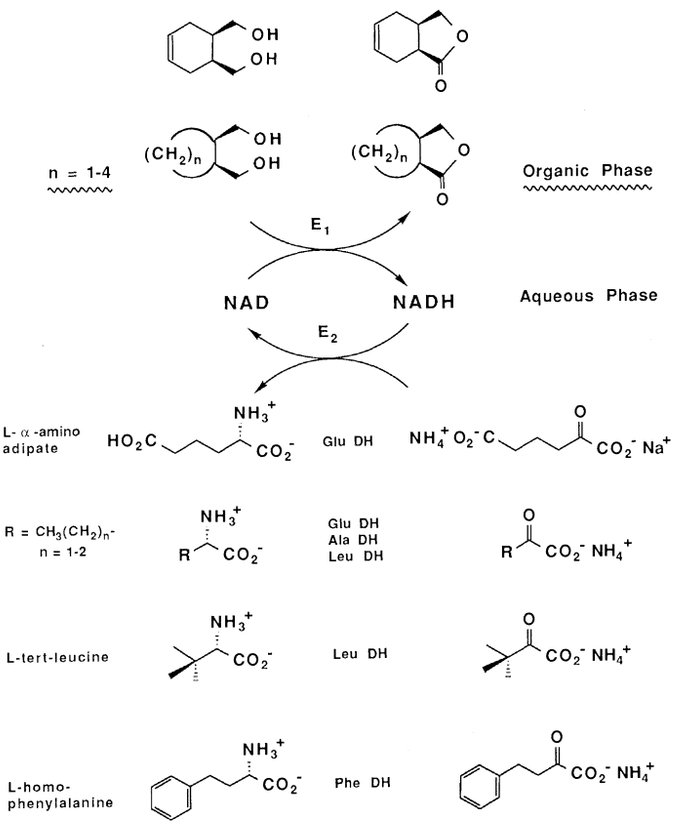


Fig. 4. Coupling of two nicotinamide cofactor–requiring enzymatic syntheses in a biphasic system; E_1, horse liver alcohol dehydrogenase; and E_2, glutamate dehydrogenase (Glu DH) or other amino acid dehydrogenases as indicated.

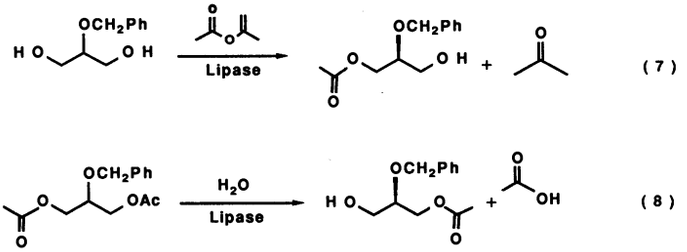
result by recognizing that when a substrate interacts with an enzyme, water must be excluded from the interface between them (*41*). The productive binding of the L-ester to the active center of subtilisin results in the release of more water molecules from the hydrophobic binding pocket of the enzyme than that of the D-isomer. This process of water release is less favorable in hydrophobic media compared with water. Thus the reactivity of the L-ester in hydrophobic media decreases substantially and the discrimination between the D- and L-esters is diminished. In any event, many reactions that are sensitive or thermodynamically impossible in water become possible in organic media. Enzyme-catalyzed dehydrations, transesterifications, aminolyses, and oxidoreductions in organic solvents are now common (*3, 4*). Novel enzymatic reactions in gases and supercritical fluids have also been exploited (*42*). Product or substrate inhibition can be lessened if the products or substrates are retained mostly in the nonaqueous phase during the reaction. In practice, enzymes are first optimized in aqueous solution by *p*H adjustment followed by lyophilization to a dry powder, which is then suspended in an organic solvent containing the substrates. The mixture is then shaken. After the reaction is complete, the enzyme can be recovered through filtration or centrifugation and used repeatedly. If necessary, immobilization can be used to improve enzyme stability and to facilitate its recovery.

Despite the advantages of enzymatic transformations in organic solvents, there are some disadvantages: (i) They may not be applicable to charged or polyfunctional species that are insoluble in organic solvents. (ii) Adjusting *p*H is difficult in large-scale processes. (iii) Enzymes can be unstable in certain organic solvents. (iv) Many organic solvents are not environmentally acceptable. (v) Loss of stereoselectivity may occur because of the effect of solvents as described above or the reversible nature of the reaction (see below). (vi) Severe substrate or product inhibition may occur for enzymatic reaction with hydrophilic substrates (see below).

The change in ΔG during enzyme-substrate (ES) complexation in water is considered to require, in addition to other changes, the disruption of a number of hydrogen bonds associated with the substrate and the enzyme active site (*43*). If it is assumed that the ES complex in organic solvents is water-free, the change in ΔG for the ES complexation upon a transition from water to the organic solvent is $-n\Delta G_{\text{H-bond}}$, where n is the number of hydrogen bonds lost and $\Delta G_{\text{H-bond}}$ is the free energy of a hydrogen bond. A transition from aqueous to organic solvents for the complexation of polar substrates with their natural enzymes may result in a tight binding of either the substrate or the product to the enzyme, and a severe substrate or product inhibition or both can occur. Indeed, several carbohydrate-converting enzymes such as hydrolases, transferases, isomerases, and sulfatases have been reported to be catalytically inactive in organic solvents (*43*).

When esterification or transesterification reactions are carried out in organic solvents, the process becomes slow and reversible, and the enantioselectivity and yield become difficult to optimize. For example, in the kinetic resolution of a DL-alcohol by transesterification, if the D-isomer is a better substrate than the L-isomer for the enzyme, then accumulation of the D-ester and the unreacted L-alcohol occurs in the forward reaction. In the reverse reaction, however, the D-ester is a better substrate and is converted to the D-alcohol according to the principle of microscopic reversibility. The enantiomeric excesses of both the D-ester and the L-alcohol decrease progressively as the extent of the reverse reaction increases. The same situation occurs in other stereoselective esterifications or transesterifications.

Product inhibition caused by the released alcohol is another concern. These problems can be eliminated by the use of enol esters as transesterification reagents (*20*). The leaving groups of enol esters tautomerize to ketones or aldehydes that are volatile and that do not undergo the reverse reaction. The combination of such a transesterification and hydrolysis provides a simple way for the preparation of both enantiomers. For example, both the *R* and *S* chiral glycerol derivatives shown in Eqs. 7 and 8 can be easily prepared from 2-*O*-benzylglycerol and the diacetate through lipase-catalyzed transesterification and hydrolysis, respectively. In each case, the enzyme possesses the same enantiotopic group selectively.



As mentioned, the technique of site-directed mutagenesis should, in principle, be very useful for the improvement of enzyme catalysis or stability or both in organic solvents as the environments are unnatural. Hydrophilic organic solvents such as dimethylformamide (DMF) tend to strip water from enzymes and inactivate them. Only two enzymes, subtilisin (*44*) and protease N (*45*), are known to be active in DMF. The half-life of subtilisin BPN′ in dry DMF is only 20 min (*46*). However, the powerful dissolving properties of DMF are useful for organic reactions. Recently, scientists at Genex have prepared a subtilisin mutant that is several hundred times more stable than the wild type in aqueous solution [(*47*), Fig. 5] and 50 times more stable in dry DMF (*46*). This mutant has been used in my research group for regioselective acylation of sugars and nucleosides, for enantioselective synthesis of chiral alcohols, acids and esters, and for synthesis of peptides (*46*).

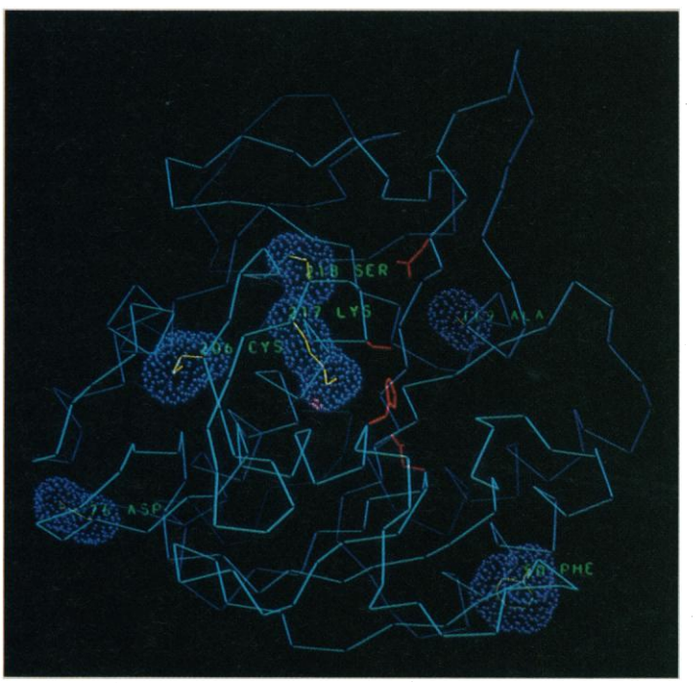


Fig. 5. Engineered substilisin 8350 produced by Genex (*47*) is catalytically active and stable in dimethylformamide. The six positions of mutation and their improved stabilizing interactions determined by x-ray structural analysis are: $Met^{50} \rightarrow$ Phe and $Gly^{169} \rightarrow$ Ala, hydrophobic; $Asn^{76} \rightarrow$ Asp, Ca^{2+} binding; $Gln^{206} \rightarrow$ Cys, van der Waals contact; and $Tyr^{217} \rightarrow$ Lys and $Asn^{218} \rightarrow$ Ser, hydrogen bonding.

Other Examples

The following examples cover some other recent developments in my laboratory regarding the synthesis of peptides and carbohydrates.

Peptide synthesis. An important goal in the area of peptide synthesis is to design a catalyst for the condensation of peptide segments containing natural or unnatural amino acids or both. Synthesis of peptides greater than 100 amino acids in length based on the chemical condensation strategy of segments prepared by solid-phase synthesis on styrene-bound oxime (*48*) represents a significant advance. The chemical coupling of peptide segments, however, often suffers from problems such as racemization, low yield, and poor solubility. For these reasons, proteases have been used as alternative reagents in peptide coupling because the reactions are catalytic, regio- and stereoselective, racemization-free, and required minimal side-chain protection. Both thermodynamically and kinetically controlled approaches have been used in enzymatic peptide synthesis. The former is a direct reversal of peptide hydrolysis, and the latter is the aminolysis of a protected peptide ester.

The kinetic approach is much more efficient, but the amidase activity of proteases tends to cleave sensitive peptide bonds in the substrates or products that result in miscoupling (transpeptidation) and undesirable peptide hydrolysis. Proteases perhaps can be modified so that they bind but do not cleave peptides. The modified protease could perhaps be used as a chiral template to catalyze the aminolysis reaction. Alternatively, proteases may be modified to have only esterase activities and yet retain their native binding properties. In this case, the amidase-deficient proteases may catalyze the aminolysis reaction through an acyl-enzyme intermediate. Indeed, thiolsubtilisin prepared by modification of the active-site Ser of subtilisin to Cys acts as a weak esterase and catalyzes the aminolysis of N-protected peptide *p*-chlorophenyl esters (*49*). Methylchymotrypsin prepared by selective methylation of α-chymotrypsin at Nε2 of the active-site His becomes a weak esterase without amidase activities and catalyzes the aminolysis of N-protected peptide methyl or cyanomethyl esters (Eq. 9; Cbz is the protecting group). The reactions may proceed through an acyl-enzyme intermediate (*50*). To form the acyl-enzyme intermediate, the imidazole group may undergo ring-flipping (*50*) so that the unmethylated N atom can act as a general base to abstract the Ser hydroxyl proton. Other serine proteases perhaps can be so modified to have such novel activities.

$$\text{Cbz-Tyr-Gly-Gly-Phe-OCH}_2\text{CN (1 eq) + Leu-NH}_2\text{ (2 eq)} \xrightarrow[\text{50\% DMF, } p\text{H 8.8}]{\text{Methylchymotrypsin}} \text{Cbz-Tyr-Gly-Gly-Phe-Leu-NH}_2\text{ (99\%)} \quad (9)$$

Serine and cysteine proteases, including trypsin, chymotrypsin, papain, and subtilisin, express insignificant amidase activities in the presence of certain concentrations (40 to 70%) of water-miscible organic solvents. The esterase activities, however, remain significant. The solvent-modified enzymes have been used in the aminolysis reactions to form a variety of usual and unusual peptides (*51*). The organic solvents apparently affect the amidase and esterase activities of proteases quite differently, perhaps because of a change in acidity and basicity of certain functional groups in the active site or a change of local conformation such that the amide-cleavage mechanism is selectively damaged (that is, the enzyme can no longer stabilize the transition state of amide cleavage).

These two approaches, active-site mutations and cosolvent-mediated modifications, represent interesting developments toward the design of peptide or protein ligases useful in peptide segment coupling. Antibody catalysis and in vitro site-directed mutagenesis may also find use in this area. The use of solvent-modified subtilisin in the synthesis of dermorphin and a chemotactic peptide through the coupling of peptide segments prepared in solution phase or by solid-phase methods with a styrene-bound oxime is illustrated in Eqs. 10 and 11.

$$\text{Boc-Tyr-D-Ala-Phe-Gly-O-N}\bigcirc + \text{Tyr-Pro-Ser(OBzl)-NH}_2\bullet\text{HCl (3 eq)} \xrightarrow{20\text{ min}}$$

$$\text{Boc-Tyr-D-Ala-Phe-Gly-Tyr-Pro-Ser(OBzl)–NH}_2 \quad (82\%\text{ yield}) \quad (10a)$$

$$\text{Boc-Tyr-D-Ala-Phe-Gly-OCH}_2\text{CH}_3 + \text{Tyr-Pro-Ser(OBzl)–NH}_2\bullet\text{HCl(3 eq)} \xrightarrow{30\text{ min}}$$

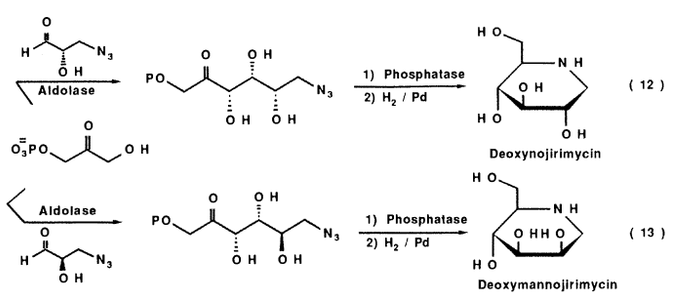
$$\text{Boc-Tyr-D-Ala-Phe-Gly-Tyr-Pro-Ser(OBzl)–NH}_2 \quad (80\%\text{ yield}) \quad (10b)$$

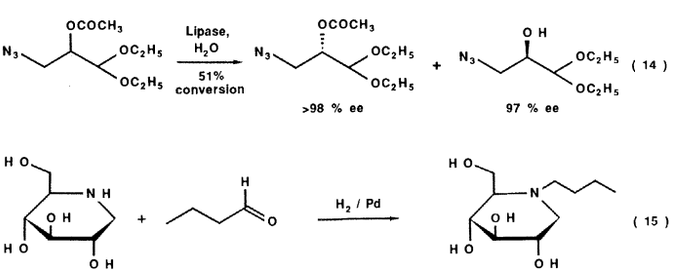
$$\text{Boc-Met-Leu-OCH}_3 + \text{Phe-NHCH}_2\text{Ph}\bullet\text{HCl(3 eq)} \xrightarrow{10\text{ min}}$$

$$\text{Boc-Met-Leu-Phe-NHCH}_2\text{Ph (95\% yield)} \quad (11)$$

The reactions are run in a 1:1 solution of water and DMF, *p*H 8.5 to 9 (triethylamine), in the presence of subtilisin.

Aldol condensation. Asymmetric C–C bond formation based on catalytic aldol addition reactions remains one of the most interesting and challenging subjects in synthetic organic chemistry. Many successful nonbiological strategies have been reported. Though enzymes often provide products with high optical purity, enzymatic aldol reactions have only been recently developed (*45, 52*). Of more than 15 aldolases isolated, the Schiff base–forming D-fructose-1,6-diphosphate (FDP) aldolase from rabbit muscle is the most often used. The Zn^{2+}-containing microbial FDP aldolase has been exploited synthetically to some extent. This enzyme from *Escherichia coli* exhibits excellent operational stability and, like that from rabbit muscle, accepts a broad range of aldehydes as second substrates, although both are very specific for dihydroxyacetone phosphate (DHAP) (*53*). In collaboration with A. Sinskey, my research group has constructed a 300-fold overproducing *E. coli* strain for use as the enzyme source. The combined enzymatic aldo condensation and catalytic reductive amination, as illustrated in Eqs. 12 and 13, are a





practical route to 1-deoxynojirimycin and 1-deoxymannojirimycin, both useful glycosidase inhibitors (*53*). The aldehyde substrates were prepared through *Pseudomonas* lipase-catalyzed resolution of 3-azido-2-hydroxypropanal diethyl acetal (Eq. 14). The donor DHAP can be replaced with a mixture of dihydroxyacetone and a small amount of inorganic arsenate (Fig. 6). Mechanistic studies (*54*)

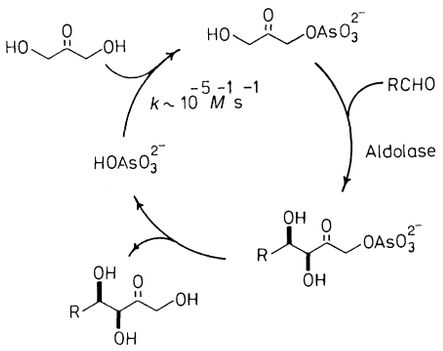


Fig. 6. Replacement of DHAP with dihydroxyacetone and a catalytic amount of arsenate in FDP aldolase–catalyzed reactions.

indicate that dihydroxyacetone reacts spontaneously with arsenate in aqueous solution to form dihydroxyacetone monoarsenate, which is a mimic of DHAP and thus is accepted by the aldolase as substrate. After the aldol reaction, the arsenate moiety dissociates from the product and is recycled. Many of the aldol products prepared so far are either biologically active or useful intermediates for synthesis of sugar-related substances. The *N*-butyl derivative of 1-deoxynojirimycin, for example, is very effective against the AIDS virus while lacking cytotoxicity (Eq. 15) (*55*). Given that many other aldolases exist and the enzymes generally accept a broad range of aldehyde substrates, enzymatic aldol reations should become one of the most valuable synthetic methods in carbohydrate chemistry. Research is being directed toward the design of new aldolase activities through chemical and biological means. The combined enzymatic aldol reaction–catalytic reductive amination as illustrated is obviously a useful strategy for the synthesis of other types of piperidines structurally related to sialic acid, fucose, rhamnose, and galactose.

Conclusion

Virtually all kinds of protein catalysts can be constructed from the 20 common amino acids. Although the relation of protein structure and catalysis to amino acid sequence is still impossible to predict, nature does provide a virtually unlimited number of protein catalysts that are either known or remain to be explored. New protein catalysts specific for a predetermined reaction are now available through the catalytic antibody approach. Known enzymes or proteins can be altered through site-directed mutagenesis or chemical or enzymatic modifications to provide additional new protein catalysts with novel catalytic properties. All of these approaches are available for the further exploitation of enzymes in synthesis.

REFERENCES AND NOTES

1. G. M. Whitesides and C.-H. Wong, *Angew. Chem. Int. Ed. Eng.* **24**, 617 (1985).
2. H. Yamada and S. Shimizu, *ibid.* **27**, 622 (1988).
3. A. M. Klibanov, *CHEMTECH* (1986), p. 354.
4. H. Simon *et al.*, *Angew. Chem. Int. Ed. Eng.* **24**, 539; R. Porter and S. Clark, Eds., *Enzymes in Organic Synthesis* (Pitman, London, 1985); M. Schneider, *Enzymes as Catalysts in Organic Synthesis* (Reidel, Dordrecht, 1986); C. Laane, J. Tramper, M. D. Lilly, *Biocatalysts in Organic Media* (Elsevier, Amsterdam, 1986); J. B. Jones, *Tetrahedron* **42**, 3351 (1986); S. Butt and S. M. Roberts, *Chem. Br.* **1987**, 127 (1987); A. Akitama *et al.*, *CHEMTECH* (1988), p. 627.
5. D. E. Koshland, Jr., and K. E. Neet, *Annu. Rev. Biochem.* **37**, 359 (1968); W. P. Jencks, *Catalysis in Chemistry and Enzymology* (McGraw-Hill, New York, 1969); R. D. Gandour and R. L. Schowen, *Transition States of Biochemical Processes* (Plenum, New York, 1978); C. T. Walsh, *Enzymatic Reaction Mechanisms* (Freeman, New York, 1979); M. I. Page and A. Williams, *Enzyme Mechanisms* (Royal Society of Chemistry, London, 1987).
6. A. R. Fersht, *Enzyme Structure and Mechanisms* (Freeman, New York, 1985).
7. For a comprehensive review and analysis in this regard, see J. Kraut, *Science* **242**, 533 (1988). The idea of transition-state stabilization was first suggested by L. Pauling and quantitatively analyzed by J. L. Kurz.
8. H. Eyring, *J. Chem. Phys.* **3**, 107 (1935).
9. W. P. Jencks, in *Current Aspects of Biochemical Energetics*, E. P. Kennedy, Ed. (Academic Press, New York, 1966), pp. 273–298; R. Wolfenden, *Annu. Rev. Biophys.* **5**, 271 (1976); G. E. Lienhard, *Science* **180**, 149 (1973); G. R. Stark and P. A. Bartlett, *Pharmacol. Ther.* **23**, 45 (1983).
10. P. A. Bartlett and C. K. Marlowe, *Biochemistry* **22**, 4618 (1983).
11. D. E. Tronrud, H. M. Holden, B. W. Matthews, *Science* **235**, 571 (1987).
12. A. R. Fersht, *Biochemistry* **26**, 8031 (1987); G. A. Gerlt, *Chem. Rev.* **87**, 1079 (1987); J. R. Knowles, *Science* **236**, 1252 (1987); S. J. Benkovic, C. A. Fierke, A. M. Naylor, *ibid.* **239**, 1105 (1988).
13. R. A. Lerner, in *Proceedings of the XVIIIth Solvay Conference on Chemistry*, G. Van Binst, Ed. (Springer-Verlag, Berlin, 1983), pp. 43–49.
14. For carbonate hydrolysis, see S. J. Pollack, J. W. Jacobs, P. G. Schultz, *Science* **234**, 1570 (1986); for ester hydrolysis, see A. Tramontano, K. D. Janda, R. A. Lerner, *ibid.*, p. 1566; A. Tramontano, A. A. Ammann, R. A. Lerner, *J. Am. Chem. Soc.* **110**, 2282 (1988); for amide hydrolysis, see K. D. Janda, D. Schloeder, S. J. Benkovic, R. A. Lerner, *Science* **241**, 1188 (1988); for the Claisen reaction, see D. Y. Jackson *et al.*, *J. Am. Chem. Soc.* **110**, 4841 (1988); D. Hilvert and K. D. Nared, *ibid.*, p. 5593; for photoreaction; see A. G. Cochran, R. Sugasawara, P. G. Schultz, *ibid.*, p. 7888; for redox photoreaction, see K. M. Shokat, C. J. Lenmann, R. Sugasawara, P. G. Schultz, *Angew. Chem. Int. Ed. Engl.* **27**, 1172 (1988); for transesterification, see A. D. Napper, S. J. Benkovic, A. Tramontano, R. A. Lerner, *Science* **237**, 1041 (1987); for aminolysis, see S. J. Benkovic, A. D. Napper, R. A. Lerner, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5355 (1988); K. D. Janda, R. A. Lerner, A. Tramontano, *J. Am. Chem. Soc.* **110**, 4835 (1988); for specific peptide cleavage, see B. L. Iverson and R. A. Lerner, *Science* **243**, 1184 (1989); K. D. Janda, S. J. Benkovic, R. A. Lerner, *ibid.* **244**, 437 (1989).
15. P. G. Schultz, *Science* **240**, 426 (1988). The design of other types of enzymatic catalysts is also described.
16. C.-S. Chen, Y. Fujimoto, G. Girdaukas, C. J. Sih, *J. Am. Chem. Soc.* **104**, 7294 (1982); C.-S. Chen, S.-H. Wu, G. Girdaukas, C. J. Sih, *ibid.* **109**, 2812 (1987).
17. L. K. P. Lam, R. A. H. F. Hui, J. B. Jones, *J. Org. Chem.* **51**, 2047 (1986).
18. W. E. Ladner and G. M. Whitesides, *J. Am. Chem. Soc.* **106**, 7250 (1984).
19. M. Ohno, S. Kobayashi, T. Iimori, Y. F. Wang, T. Izawa, *J. Am. Chem. Soc.* **103**, 2405 (1981); P. Mohr, N. Waespe-Sarcevic, C. Tamn, K. Gawronska, J. K. Gawronski, *Helv. Chim. Acta.* **66**, 2501 (1983); G. Sabbioni, M. L. Shea, J. B. Jones, *Chem. Commun.* **1984**, 236 (1984).
20. Y. F. Wang and C.-H. Wong, *J. Org. Chem.* **53**, 3127 (1988); Y. F. Wang *et al.*, *J. Am. Chem. Soc.* **110**, 7200 (1988).
21. J. B. Jones and I. J. Jakovac, *Can. J. Chem.* **60**, 19 (1982); I. J. Jakovac, H. B. Goodbrand, K. P. Lok, J. B. Jones, *J. Am. Chem. Soc.* **104**, 4659 (1982).
22. J. B. Jones, in *Biocatalysis*, D. A. Abramowicz, Ed. (Reinhold, New York, in press).
23. C. N. Cronin, B. A. Malcolm, J. F. Kirsch, *J. Am. Chem.* **109**, 2222 (1987).
24. H. M. Wilks *et al.*, *Science* **242**, 1541 (1988).
25. J. A. Wells and D. A. Estell, *Trends Biochem. Sci.* **291** (1988); M. W. Pantoliano *et al.*, *Biochemistry* **27**, 8311 (1988), and references therein.
26. R. J. Lamed and J. G. Zeikus, *Biochem J.* **195**, 183 (1981). For use in other syntheses, see E. Keinan, E. K. Hafel, K. K. Seth, R. Lamed, *J. Am. Chem. Soc.* **108**, 162 (1986); C.-H. Wong, D. G. Drueckhammer, H. M. Sweers, *ibid* **107**, 4028 (1985).
27. C. B. Anfinsen, *Science* **181**, 223 (1973); ——— and H. A. Scheraga, *Adv. Protein Chem.* **29**, 205 (1975).
28. V. V. Mozhaev and K. Martinek, *Enzyme Microb. Tech.* 6, 50 (1984).
29. S. K. Burley and G. A. Petsko, *Science* **229**, 23 (1985).
30. K. Mosbach, Ed., *Methods in Enzymology* (Academic Press, New York, 1987) vols. 135, 136, and 137.
31. For thorough reviews of cofactor regeneration, see H. K. Chenault and G. M. Whitesides, *App. Biochem. Biotech.* **14**, 147 (1987); H. K. Chenault, E. S. Simon, G. M. Whitesides, *Biotech. Genetic. Eng. Rev.* 6, 221 (1988); for hydrogenase, see A. Van Berkel-Arts *et al.*, *Biochemie* **68**, 201 (1986). For recent electrochemical methods, see P. M. Allen and W. R. Bowen, *Trends Biotech.* **3**, 145 (1985); for catalytic H_2 reduction, see S. Chao and M. S. Wrighton, *J. Am. Chem. Soc.* **109**, 5886 (1987).
32. C. J. Sih and C.-S. Chen, *Angew. Int. Ed. Engl.* **23**, 570 (1984).
33. C.-H. Wong, S. L. Haynie, G. M. Whitesides, *J. Org. Chem.* **47**, 5416 (1982); J. Thiem and W. Treder, *Angew. Chem. Int. Ed. Engl.* **25**, 1096 (1986); C. Auge, S. David, C. Mathieu, C. Gautheron, *Tetrahedron Lett.* **25**, 1467 (1984); E. S. Simon, M. D. Bednarski, G. M. Whitesides, *J. Am. Chem. Soc.* **110**, 7159 (1988).
34. J. R. Matos and C.-H. Wong, *J. Org. Chem.* **51**, 2388 (1986).
35. R. Wichmann, C. Wandrey, A. F. Buckmann, M. R. Kula, *Biotech. Bioeng.* **23**, 2789 (1981).
36. For oxidation of toluene, see D. T. Gibson, M. Hensley, H. Yoshioka, T. J. Mabry, *Biochemistry* **9**, 1626 (1970); for oxidation of the cyclohexanone, see M. J. Taschner and D. J. Black, *J. Am. Chem. Soc.* **110**, 6892 (1988); C. T. Walsh and Y.-C. J. Chen, *Angew. Chem. Int. Ed. Engl.* **27**, 333 (1988); for oxidation of 2-methylpropionic acid, see M. F. Zuger, F. Giovannini, D. Seebach, *ibid.* **22**, 1012 (1983); for isopenicillin N synthetase, see S. Wolfe, A. L. Demain, S. E. Jensen, D. W. S. Westlake, *Science* **226**, 1386 (1984); J. E. Baldwin *et al.*, *Chem. Commun.* **1986**, 975 (1986); for steroid hydroxylation, see Y. Ahronowitz and G. Cohen, *Sci. Am.* **245**, 141 (September 1981).
37. F. R. Dastol, N. A. Musto, S. Price, *Arch. Biochem. Biophys.* **115**, 44 (1966); F. R. Dastol and S. Price, *ibid.* **118**, 163 (1967).
38. K. Oyama, S. Nishimura, Y. Nonaka, K. I. Kihara, T. Hashimo, *J. Org. Chem.* **46**, 5241 (1981); A. R. Macrae, *J. Am. Oil Chem. Soc.* **60**, 291 (1983); S. Fukui and A. Tanaka, *Endeavour* **9**, 10 (1985); G. Kitchner, M. P. Scollar, A. M. Klibanov, *J. Chem. Soc.* **107**, 7072 (1985).

39. A. Zaks and A. M. Klibanov, *J. Biol. Chem.* **263**, 3194 (1988); *ibid.*, p. 8017, and references therein.
40. T. Sakurai, A. L. Margolin, A. J. Russell, A. M. Klibanov, *J. Am. Chem. Soc.* **110**, 7236 (1988).
41. M. J. S. Dewar, *Enzyme* **36**, 8 (1986); see also chap. 11 of (*6*).
42. T. W. Randolph, D. S. Clark, H. W. Blanch, J. M. Prausnitz, *Science* **239**, 387 (1988).
43. A. P. G. Kieboom, *Recl. Trav. Chim. Pays-Bas* **107**, 347 (1988).
44. S. Riva, J. Chopineau, A. P. G. Kieboom, A. M. Klibanov, *J. Am. Chem. Soc.* **110**, 584 (1988).
45. M.-J. Kim, W. J. Hennen, H. M. Sweers, C.-H. Wong, *ibid.*, p. 6481.
46. W. J. Hennen *et al.*, unpublished results.
47. P. N. Bryan *et al.*, *J. Cell. Biochem.* **13** (suppl. A), 66 (1989).
48. E. T. Kaiser *et al.*, *Science* **243**, 187 (1989).
49. T. Nakatsuka, T. Sasaki, E. T. Kaiser, *J. Am. Chem. Soc.* **109**, 3808 (1987); for other active-site mutations, see E. T. Kaiser, *Angew. Chem. Int. Ed. Engl.* 27, 913 (1988).
50. J. B. West *et al.*, *J. Am. Chem. Soc.* **110**, 3709 (1988). The acyl intermediate detected by nuclear magnetic resonance suggests the postulated ring-flipping mechanism, which is also supported by kinetic (J. D. Scholten, J. L. Hogg, F. M. Raushel, *ibid.*, p. 8246) and model studies (J. Rebek, Jr., *Struct. Chem.*, in press).
51. C. F. Barbas III, J. R. Matos, J. B. West, C.-H. Wong, *J. Am. Chem. Soc.* **110**, 5162 (1988).
52. M. D. Bednarski *et al.*, *ibid.* **111**, 627 (1989); J. R. Durrwachter and C.-H. Wong, *J. Org. Chem.* **53**, 4175 (1988); R. L. Pederson, M.-H. Kim, C.-H. Wong, *Tetrahedron Lett.* **29**, 4645 (1988); T. Ziegler, A. Straub, F. Effenberger, *Angew. Chem. Int. Ed. Engl.* 27, 716 (1988); S. David and C. Auge, *Pure Appl. Chem.* **59**, 1501 (1987); L. M. Reimer, D. L. Conley, D. L. Pompliano, J. W. Frost, *J. Am. Chem. Soc.* **108**, 8010 (1986); N. J. Turner and G. M. Whitesides, *ibid.* **111**, 624 (1989).
53. C. H. Von der Osten *et al.*, *J. Am. Chem. Soc.*, in press.
54. D. G. Drueckhammer *et al.*, *J. Org. Chem.* **54**, 70 (1989).
55. A. Karpas *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 9229 (1988).
56. I thank many co-workers for their essential contributions; their names are listed in the references. This research was supported by NSF, NIH, the Searle Scholars Program, and Dow Chemical Company.

The Role of Somatic Hypermutation in the Generation of Antibody Diversity

DEBORAH L. FRENCH, REUVEN LASKOV,* MATTHEW D. SCHARFF

The immune system is capable of establishing an enormous repertoire of antibodies before its first contact with antigen. Most antibodies that express germ-line sequences are of relatively low affinity. Once antigen enters the system, it stimulates a somatic mutational mechanism that generates antibodies of higher affinity and selects for the expression of those antibodies to produce a more effective immune response. The details of the mechanism and regulation of somatic hypermutation remain to be elucidated.

IN ORDER TO DEAL WITH THE MANY TOXIC SUBSTANCES AND pathogenic organisms in the environment, higher organisms have developed an immune system that can react with specificity and flexibility by producing antibodies to a seemingly infinite array of antigenic stimuli. Through a process of somatic assembly of germ-line genes, an individual can produce a sufficiently large and diverse repertoire of antibody molecules to react with all of the antigens in the environment. This repertoire, however, is not sufficient to protect an organism fully, because many of the germ line–encoded antibodies have relatively low affinities. Since antibodies play their major role in the circulation by neutralizing antigens such as toxins or virus molecules that have high-affinity cellular receptors, the ability of an organism to protect itself from environmental antigens often requires antibodies that are of higher affinity. Such antibodies are generated by hypermutation of the germ line–encoded antibody genes. In this article, we focus on current knowledge of the hypermutation process and describe those aspects of the process that remain to be elucidated.

D. L. French and M. D. Scharff are in the Department of Cell Biology and the Irvington House Institute for Medical Research, Albert Einstein College of Medicine, Bronx, NY 10461. R. L. Laskov is from the Hubert H. Humphrey Center for Experimental Medicine and Cancer Research, Hebrew University–Hadassah Medical School, Jerusalem, Israel.

*Present address: Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY 10461.

Establishing the Antibody Repertoire

The antibody molecule consists of two identical heavy (H) chains and light (L) chains that are disulfide-linked to form a bivalent H_2L_2 molecule (Fig. 1A). The amino terminal end of each chain is encoded by a variable (V)-region gene and the carboxyl terminal end is encoded by a constant (C)-region gene. The antigen binding site of the molecule is formed by the V-region domains and differs significantly among antibodies that react with different antigens. The V-region domain contains three areas of exceptional amino acid sequence variability called complementarity-determining regions, which are the sites of antigen contact. Surrounding the complementarity-determining regions are areas of less sequence variability called framework regions. The V-region domains of the protein are encoded by a series of smaller genetic elements designated variable (V_H), diversity (D), and joining (J) for the heavy chains and V_L and J_L for the light chains (Fig. 1A). Each genetic element is part of a larger linkage group, and those encoding the murine heavy and both light chain [kappa (κ) and lambda (λ)] V regions are located on different chromosomes. Each group of genetic elements is separate in the germ-line configuration (Fig. 1B), and for the heavy chain there are 100 to 1000 V_H (*1, 2*), approximately 12 D, and 4 J genetic elements. In the mouse, the κ light chain has more than 100