Calculations of Enzymatic Reactions: Calculations of pK_a, Proton Transfer Reactions, and General Acid Catalysis Reactions in Enzymes†

Arieh Warshel

ABSTRACT: A method that allows one to correlate available X-ray data with activation free energies of enzymatic reactions is presented. This method is based on the empirical valence bond approach which uses experimental information to evaluate the energies of the valence bond resonance forms involved in a reaction and then calculates the environment-dependent stabilizations of the ionic resonance forms in the enzyme and in solution and correlates them with the rate acceleration by the enzyme. The method is reliable since it is based on calibration of potential surfaces by solution experiments and on transfer of the calibrated surfaces to the enzyme active site, using only simple calculations of electrostatic interactions. The close relation between the method and the intuitive valence bond description of bond-breaking bond-making reactions provides a new insight into enzymatic reactions, describing them as crossings between covalent and ionic valence bond resonance forms. Such a description correlates the stabilization of the ionic resonance forms by the enzyme active site with the enzyme catalytic activity. The paper considers the energetics of several enzymatic processes, including ionization of acidic groups in enzyme active sites, stability of ion pairs in enzymes and in solutions, proton transfer reactions, and general acid catalysis reactions. The calculations support the idea that enzymes can be viewed as "supersolvents" that stabilize (solvate) ionic transition states more effectively than do aqueous solutions.

One of the basic problems in biochemistry is understanding the molecular origin of enzyme catalysis. Although this complicated question must be related to the exact three-dimensional structure of the enzyme–substrate complex, it was not resolved even when X-ray studies provided detailed information about the active sites of several enzymes. Basically, the rate of an enzyme reaction is determined by the interaction energy between the active site and the substrate, but available experimental information does not discriminate between the various components of these interactions and their relative importance. Thus, it is of great importance to develop theoretical approaches that allow one to use available X-ray structures to elucidate the energetics of enzyme catalysis.

Before choosing a theoretical approach, it is important to define the problem of enzyme catalysis in a clear way, relating the rate of the enzymatic reaction to free energies. A typical enzymatic reaction involves the steps described in Figure 1. The overall reaction rate is determined by two free energy contributions, the binding free energy ΔG_{bind} that determines the concentration of the enzyme–substrate complex (ES)

$$[E][S]/[ES] = \exp(-\Delta G_{bind}/RT) = K_M \qquad (1)$$

and the activation free energy $\Delta G^{\ddagger}_{cat}$ that determines the rate-limiting step

$$k_{cat} \simeq (k_B T/\hbar) \exp(-\Delta G^{\ddagger}_{cat}/RT) \qquad (2)$$

where k_B and $\hbar$ are the Boltzmann and Planck constants, respectively. The interplay between ΔG_{bind} and $\Delta G^{\ddagger}_{cat}$ in determining the overall rate of the enzymatic reaction is discussed elsewhere (Fersht, 1977). In most enzymatic reactions, ΔG_{bind} is negative and the maximum rate for [S] $\gg$ [E] is given approximately by k_{cat}, which depends only on $\Delta G^{\ddagger}_{cat}$. In reactions in solution (Figure 1), the equivalent of ΔG_{bind} is ΔG_{cage}, the free energy needed to bring the reactants into the same solvent cage. In many cases, ΔG_{cage} is positive, and the overall rate is given approximately by $(k_B T/\hbar) \exp[-(\Delta G_{cage} + \Delta G^{\ddagger}_{cage})/RT]$. A direct comparison of the reactions in solution and in enzyme requires comparison of ΔG_{cage} and ΔG_{bind}. ΔG_{cage} includes the entropy contribution of bringing the reactants together (Page & Jencks, 1971), which cannot be calculated at present for polar reactants in solution. Although the factors that affect ΔG_{cage} are of great interest, we believe that the key question in terms of structure–function correlation is the difference between $\Delta G^{\ddagger}_{cat}$ and $\Delta G^{\ddagger}_{cage}$. The reasons for this are as follows: (i) The catalytic role of enzymes is frequently described in terms of stabilization of ES‡. This effect is included in $\Delta G^{\ddagger}_{cat}$ and *not* in ΔG_{bind}. (ii) $\Delta G^{\ddagger}_{cat}$ determines the maximum rate of most enzymatic reactions. (iii) When comparing the rates of reactions of different substrates in the same active sites, one finds that in many cases the main difference is in k_{cat} rather than K_M [e.g., pepsin and elastase (Fersht, 1977)]. Furthermore, even when the increase in k_{cat} for different substrates is similar to the corresponding decrease in K_M [e.g., chymotrypsin hydrolysis of *N*-acetyl amino acid methyl esters (Fersht, 1977)], all the differences in the catalytic activity are included in the changes of $\Delta G^{\ddagger}_{cat}$. That is, for such cases, $\Delta\Delta G_{bind} \simeq -\Delta\Delta G^{\ddagger}_{cat}$ and the differences in catalysis are due to ground-state destabilization effects which are *included* in $\Delta G^{\ddagger}_{cat}$ ($\Delta\Delta G^{\ddagger}_{cat}$ is the difference between the ground- and transition-state stabilization). In all these cases, the crucial and as yet unanswered question is how the structure of the enzyme–substrate complex determines k_{cat}.

A method that allows one to estimate $\Delta G^{\ddagger}_{cat}$ from X-ray results should be simple and flexible enough to allow interpolation of free energy parameters from reactions in solutions. Such an approach is the empirical valence bond (EVB) approach developed in the previous paper (Warshel & Weiss, 1980). This method uses the valence bond concept of ionic–covalent resonance to obtain the gas-phase Hamiltonian and then treats the molecules in solution by adding solvation free energies to the energies of their ionic resonance forms. The resulting potential surfaces are calibrated by using pK_a measurements and other information about the reaction in solution. The calibrated potential surface for the reaction in solution is then adapted to the enzymatic reaction by replacing the solvation free energies of the ionic resonance forms by the

† From the Department of Chemistry, University of Southern California, Los Angeles, California 90007. *Received July 25, 1980.* This work was supported by Grant GM 24492 from the National Institutes of Health and by the Alfred P. Sloan Foundation. For the first part of this work see Warshel & Levitt (1976). A.W. is an Alfred P. Sloan Fellow.

0006-2960/81/0420-3167$01.25/0

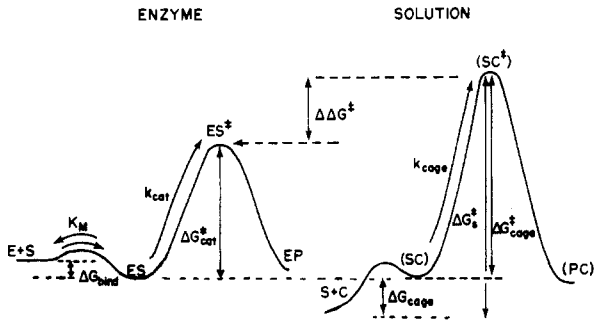


FIGURE 1: Comparison of the energetics of enzymatic and solution reactions. E, S, and P are respectively the enzyme, the substrate, and the product. C in the solution reactions is the catalyst (e.g., an acid in general acid catalysis). (SC), (SC‡), and (PC) are the indicated components inside a solvent cage. The rates k_{cat} and k_{cage} are determined by $\Delta G^{\ddagger}_{cat}$ and $\Delta G^{\ddagger}_{cage}$, respectively.

interactions with the enzyme active site. The use of this approach provides a simple and reliable way of analyzing the relation between the structures of proteins and their activities.

Under Materials and Methods, we give a brief outline of the EVB approach for studies of reactions in solutions and in enzymes. Under Results, we described the use of this approach in studying the energetics of several processes, including (i) dissociation of ionic groups of enzymes and the relation between enzyme structure and local pK_a, (ii) stabilization of ion pairs by enzymes, (iii) proton-transfer reactions in solutions and in enzymes, and (iv) general acid catalysis in solutions and in enzymes. These studies demonstrate that the transition states of many reactions possess significant ionic character and that the electrostatic interaction between the enzyme charges and dipoles and the transtion state plays a crucial role in enzyme catalysis.

Materials and Methods

(1) Calculations of Potential Surfaces. The calculations of $\Delta G^{\ddagger}_{cat}$ are based on the empirical valence bond (EVB) method for calculating potential surfaces of reactions in solutions (Warshel & Weiss, 1980). This method, which is closely related to the intuitive description of bonding, provides a simple way of transferring potential surfaces between different environments (solution to enzyme, enzyme to another enzyme). The evaluation of the EVB potential surface involves three steps:

(a) Selection of Resonance Forms and Evaluation of Approximate Potential Surface. The user of the EVB must choose, on the basis of experiment and intuition, a set of bonding arrangements or " resonance forms" which describe the reacting systems throughout the course of the reaction. Each can involve covalent, ionic, and mixed covalent/ionic bonding schemes frequently encountered in organic chemistry text books. As an example, the ionic bond cleavage reaction

$$X\text{–}Y \rightarrow X^- + Y^+ \qquad (3)$$

can be described by considering the resonance forms

$$\psi_1 = (X\text{–}Y) \qquad \psi_2 = (X^-Y^+) \qquad \psi_3 = (X^+Y^-) \qquad (4)$$

If X is more electronegative than Y, then the most important resonance forms are ψ_1 and ψ_2. After choosing the proper resonance forms, we describe their energies on the basis of experimental information. Thus, for example, the energy of the purely covalent bond, ψ_1, is approximated by a Morse-type potential function

$$E_1 = \bar{M}(r) = \bar{D}[\exp[-2\bar{a}(r - r_e)] - 2\exp[-\bar{a}(r - r_e)]] \qquad (5)$$

where r is the XY bond length, the parameter r_e is taken as the gas-phase equlibrium bond length of the X–Y bond, and $\bar{a}$ is taken from the corresponding stretching vibrational frequency. The parameter $\bar{D}$ is approximated by the geometric mean formula

$$\bar{D} = (D_{XX}D_{YY})^{1/2} \qquad (6)$$

where D_{XX} and D_{YY} are the gas-phase dissociation energies of the X–X and Y–Y bonds, respectively.

The energy of the ionic resonance form, ψ_2, depends on the environment; for the isolated (gas-phase) molecule, this energy is given by eq 7

$$E_2^g = I_Y - EA_X - e^2/r + V_{nb} \qquad (7)$$

where I and EA are ionization potential and electron affinity and the nonbonded potential V_{nb} is chosen so that the minimum of $[-e^2/r + V_{nb}(r)]$ is at the sum of the ionic radii of X and Y. For the molecule in solution, E_2 is given by eq 8

$$E_2^s = E_2^g + G_{sol}^s \qquad (8)$$

where G_{sol}^s is the solvation free energy of the X^-Y^+ ion pair. For a reaction in protein, E_2 is given by eq 9

$$E_2^p = E_2^g + G_{sol}^p = E_2^s + (G_{sol}^p - G_{sol}^s) \qquad (9)$$

where G_{sol}^p is the electrostatic interaction between the ionic state and the protein (referred to here as the "solvation" energy by the protein). As seen from eq 9, it is possible to compare enzyme and solution energies without evaluating the corresponding gas-phase energies. The crucial quantity $(G_{sol}^p - G_{sol}^s$ is evaluated by the approach described under Calculations of Solvation Energies.

After obtaining the energies of the various resonance forms, we have to consider their mixing. This is done by solving the proper secular equation which is given in the example of the above bond cleavage reaction by

$$\begin{pmatrix} E_1 - E & H_{12} \\ H_{12} & E_2 - E \end{pmatrix} = 0 \qquad (10)$$

where H_{12} is an empirical interaction term evaluated following Coulson & Danielsson (1954), by forcing the ground-state energy of the isolated (gas-phase) molecule to reproduce its observed bond energy using the relation

$$H_{12} = [(E_1 - M_{XY})(E_2^g - M_{XY})]^{1/2} \qquad (11)$$

where M_{XY} is the Morse potential of the real XY bond (rather than the purely covalent bond). This potential is fitted to reproduce the observed dissociation energy and equilibrium bond length of the X–Y bond.

Using E_1, E_2, and H_{12}, we obtain, from eq 10, the potential surface for a bond cleavage reaction in solution (Figure 2). The figure illustrates how the very large solvation energy makes the high energy ionic excited state a ground state for large r. More importantly, it demonstrates that the transition state possesses very significant character of a solvated ion pair. Thus, the activation energy is strongly dependent on the solvation energy.

The same type of approach can be applied to more complicated reactions, evaluating the valence bond matrix elements semiempirically [see Warshel & Weiss (1980)] while incorporating the calculated solvation energies in the diagonal matrix elements.

(b) Calibrating the Potential Surface. Having the approximate valence bond surface, we calibrate it by fitting calculated energies at asymptotic points ($r = \infty$) to the corresponding observed values in solution reactions. For example, in the ionic bond cleavage reaction considered above, we force

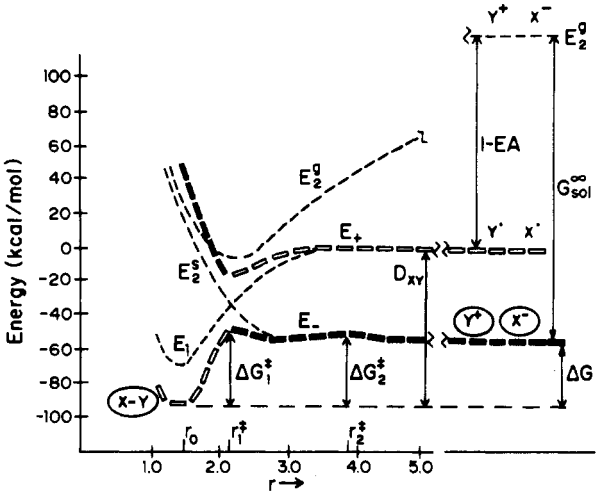


FIGURE 2: Potential surface for ionic bond dissociation in polar solvents. The surface is based on calculations of the reaction R–O–R′ → R^+ + $R'O^-$ (Warshel & Weiss, 1980). E_1, E_2^g, and E_2^s are respectively the energetics of the covalent, gaseous ionic, and aqueous ionic resonance forms. E_- and E_+ are the ground- and excited-state energies in solution. ΔG is the total reaction free energy, $\Delta G^{\ddagger}_2 - \Delta G$ is the barrier for forming the ion pair in the solvent cage, and $\Delta G^{\ddagger}_1$ is the activation barrier at $r^{\ddagger}_1$ where E_1 and E_2^s intersect. The regions of the surfaces that correspond to ionic states are shown as closed rectangles; those that correspond to covalent state as open rectangles.

the calculated ΔG (Figure 2) to reproduce the observed ΔG of reaction by using the relation

$$E_2^s(\infty) - E_-^s(r_0) = \Delta G_{obsd} \quad (12)$$

Similarly, in studying proton-transfer reactions of the form

$$AH + B \rightarrow A^- + BH^+ \quad (13)$$

we calibrate the asymptotic energy of the ionic state (the calculated energy for proton transfer in solution, ΔG_{PT}^s) by using the observed pK_a values of A and B (eq 14). It is

$$\Delta G_{PT}^s = 2.3RT[pK_a(AH) - pK_a(BH^+)] \quad (14)$$

possible in some cases to calibrate the calculated $\Delta G^{\ddagger}$ by using the observed activation free energy (Warshel & Weiss, 1980). Calibration guarantees the correct asymptotic behavior of the potential surfaces and provides a reasonable estimate of the ionic character of the transition state.

(*c*) *Comparing Enzymatic and Solution Reactions.* Evaluating a calibrated potential surface for the reaction in solution provides a unique way of constructing the potential surface for the corresponding enzymatic reaction, simply by replacing the solvation free energy by the interaction with the enzyme active site. That is, in order to evaluate the potential surface for the above ionic bond cleavage inside an enzyme active site, it is only necessary to replace G_{sol}^s in E_2^s by the corresponding electrostatic interaction between the enzyme and the ions, G_{sol}^p (see eq 9). This procedure is formally correct since we deal with indentical substrates and the only difference is in their interaction with the medium, which is accounted for by the solvation free energy. As will be demonstrated below, this approach provides a powerful tool for comparative studies of chemical reactions in solution and in enzymes.

(*2*) *Calculations of Solvation Energies.* Reliable evaluation of the solvation energies G_{sol} of the ionic resonance forms in polar solvents and in protein active sites is essential for meaningful comparison of solution and enzymatic reactions. We evaluate G_{sol} by the previously developed microscopic models (Warshel & Levitt, 1976; Warshel, 1979a). Here we outline the main features of these models, discuss some modification needed for the EVB approach, and describe the incorporation of water molecules in the protein calculations.

(*a*) *Solvation of Ionic Resonance Forms in Aqueous Solutions.* In previous works, we estimated the solvation energy of charged systems in aqueous solution by the surface constrained soft sphere dipoles (SCSSD) approach (Warshel, 1979a) which represents the water molecules as point dipoles attached to the centers of soft spheres and minimizes the solute–solvent and solvent–solvent energies with respect to the orientations and positions of these dipoles. The solvation energy is then given by the difference between the minimum energy of the solute–solvent system and the energy of the solvent. Because the calculations are limited to a few solvation shells, these are surrounded by a surface of dipoles in the positions of the bulk solvent. The model was calibrated to reproduce solvation enthalpies of ions at 300 K.

We must adapt the SCSSD model to the EVB method in order to be consistent with the Born–Oppenheimer approximation; for a given reactant geometry, the solvent nuclei must be fixed while the solvation energies of all the resonance forms involved are calculated. This is accomplished by replacing the permanent dipoles of the SCSSD model by a combination of permanent and induced dipoles, fixing the permanent dipoles in orientations that minimize the energy of the ground electronic state, and allowing the induced dipoles to be polarized differently for each electronic resonance form. This is consistent with both intuition and the Born–Oppenheimer approximation since the positions and orientations of the permanent dipoles (which represent the solvent nuclear coordinates) cannot simultaneously have several values, while the orientations and magnitude of the induced dipoles represent the solvent electronic distribution, which is different for different electronic wave functions of the solute–solvent system. The practical implementation of this modification is described by Warshel & Weiss (1980). The SCSSD model in this modified form is calibrated to give solvation free energies of ions at 300 K.

(*b*) *Stabilization of Ionic Resonance Forms in Proteins.* The stabilization ("solvation") energies, $G_{sol}^{(i)}$ of the ionic resonance forms in enzyme active sites are calculated by evaluating the interaction between the charges of these resonance forms and the partial charges and induced dipoles of the enzyme atoms. Since the enzyme atoms remain fixed in their X-ray determined positions, this procedure is consistent with the Born–Oppenheimer approximation. The solvation energy is made up of the following contributions. The first one is the charge–charge interactions which are given (in kcal/mol) by

$$V_{Qq}^{(i)} = 332\sum_{j,j'} Q_j^{(i)} q_{j'}/r_{jj'} \quad (15)$$

where j runs over reacting and j' over protein atoms, $r_{jj'}$ is the distance between the j and j' atoms, q are the point charges of the protein, and $Q^{(i)}$ are the charges of the substrate in the ith resonance form. Whenever the charges of the protein can be divided into electroneutral groups and form a set of permanant dipoles, we replace the notation V_{Qq} by $V_{Q\mu}$. The second is the inductive interactions between the charges of the protein and the polarizable electrons of the protein atoms. These contributions are evaluated by assigning polarizabilities to each of the protein atoms and calculating self-consistently the magnitudes and directions of the resulting induced dipoles, μ_k, in the presence of the protein charges and each other. The interaction of the induced dipoles with the reactant and protein charges is given (kcal/mol) by eq 16. The details of this

$$V_{Q\alpha}^{(i)} = -166\sum_{j,k} Q_j^{(i)}\mu_k r_{jk}/r_{jk}^3 - 166\sum_{j',k} q_{j'}\mu_k r_{j'k}/r_{j'k}^3 \quad (16)$$

Table I: Calculations of pK_a's[a]

acid	water G_{sol}^w (AH)	water G_{sol}^w (A^-)	water ΔG_{sol}^w	water $\Delta G_{sol} + V_{QQ}$	water Δ	water pK_a^w	lysozyme $\Delta V_{Q\alpha}$	lysozyme $\Delta V_{Q\mu}$	lysozyme ΔG_{sol}^p	lysozyme $\Delta G_{sol}^p + V_{QQ}$	lysozyme pK_a^p
Asp-52	−10	−80	−71	−71	77	4 (4)	−40	−30	−70	−70	5 (4)
Glu-35	−10	−80	−71	−71	77	4 (4)	−45	−20	−65	−65	9 (7)
Glu-35 (in the presence of ionized Asp-52)	−90	−212	−121	−69	77	6 (5)	−90	−12	−102	−50	19[b] (8)

[a] All acids are simulated by HCOOH. G_{sol} (A^-) for Asp-52 includes the Asp^-–Glu system and for Glu-35 includes the Asp–Glu^- system. The Asp–Glu system in water is kept in the same geometry as in the enzyme. Δ is the energy of forming the ionized acid in the gas phase and transferring the proton to water [$\Delta = \Delta H_{PT}^g(AH + H_2O \rightarrow A^- + H_3O^+) + \Delta G_{sol}^w(H_2O \rightarrow H_3O^+)$] [see Warshel (1978a)]. V_{QQ} is the interaction between Asp^- and Glu^- in the gas phase. $\Delta G_{sol} = G_{sol}(A^-) - G_{sol}(AH)$; $\Delta V = V(A^-) - V(AH)$; $V_{Q\alpha}$ and $V_{Q\mu}$ are the interactions between the indicated systems and the induced and permanent dipoles of the protein, respectively. The pK_a's are given by $pK_a = (\Delta + \Delta G_{sol}(AH \rightarrow A^-))/(2.3RT)$. Observed pK_a's are given in parentheses. [b] This calculation corresponds to the pK_a in the presence of a bound substrate, while the experimental value probably reflects the penetration of water to the active sites (see discussion in text).

procedure and useful simplifications are given by Warshel & Levitt (1976).

(c) *Incorporation of Water Molecules in the Protein Calculations.* In many cases, the protein contains bound water molecules which play an important role in its action. The bound water molecules are located in our model by building a cubic grid of 1-Å spacing around the protein and deleting iteratively any point which is within 2.6 Å of any protein atom or within 3 Å of another grid point. The grid sites which are near or at the reacting region are occupied by "real" water molecules capable of forming hydrogen bonds. These water molecules are represented by charges of 0.3 and −0.6 (in fractions of an electron) for H and O, respectively, and van der Waals parameters from Warshel & Levitt (1976) (with zero van der Waals radii for hydrogen-bond interactions with O and N). The bound water molecules are oriented in a search procedure to give the best interaction with the protein and the substrate charges. The grid sites which are not in the immediate neighborhood of the reacting region are occupied with simplified water molecules which are represented by the Langevin type point dipoles of Warshel & Levitt (1976). The regions which are outside the first few solvation shells of the protein are treated as a dielectric continuum. The interaction of the "real" water molecules, the simplified point dipoles, and the continuum with the charges of the protein–substrate complex is included in the calculated solvation energies.

Results

(1) *Calculations of pK_a of Acids in Enzymes.* Before trying to evaluate potential surfaces of complicated enzymatic reactions, we examined the ionization of acidic groups in enzymes. The energetics of such a reaction can be considered by using the cycle described in Figure 3. As shown in the figure, the free energy of proton transfer from the enzyme to water is given by eq 17, where p and w designate protein and

$$\Delta G = 2.3RT(pK_a^p - pH) = [G_{sol}^w(AH) - G_{sol}^p(AH)] + [G_{sol}^p(A^-) - G_{sol}^w(A^-)] + 2.3RT(pK_a^w - pH) = \Delta G_{sol}^p(AH \rightarrow A^-) - \Delta G_{sol}^w(AH \rightarrow A^-) + 2.3RT(pK_a^w - pH) \quad (17)$$

water environments, respectively, $G_{sol}(AH \rightarrow A^-) = G_{sol}(A^-) - G_{sol}(AH)$, and ΔG is defined for 50% dissociation of the acid. Thus, calculations of acidity of groups in enzymes involve only evaluation of differences in solvation energies. These quantities can be calculated according to the models described under Materials and Methods.

As an example of such a calculation, we evaluated the pK_a of the functional groups of Asp-52 and Glu-35 of lysozyme. The calculations, which are summarized in Table I, show that the enzyme stabilizes (solvates) ionized acids by combining

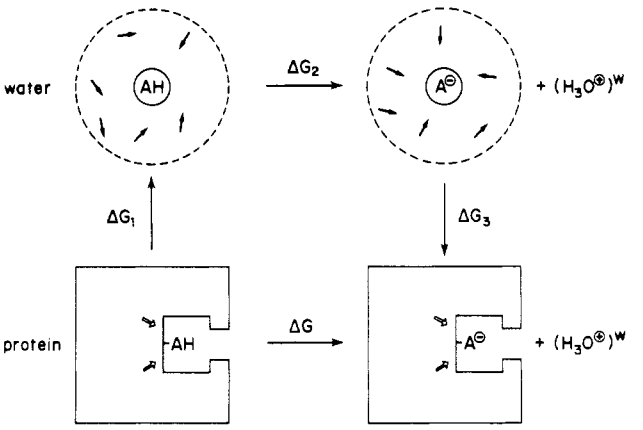


FIGURE 3: Describing the thermodynamic cycle used to estimate the energetic of dissociation of an acidic group of a protein. The ΔG_i are given by $\Delta G_1 = G_{sol}^w(AH) - G_{sol}^p(AH)$, $\Delta G_2 = 2.3RT(pK_a^w - pH)$, and $\Delta G_3 = G_{sol}^p(A^-) - G_{sol}^w(A^-)$.

two contributions, (i) the interaction between the acid charge and the permanent dipoles of the enzyme, $V_{Q\mu}$, and (ii) the interaction between the acid's charges and the induced dipoles of the enzyme, $V_{Q\alpha}$. The magnitude of both contributions can be appreciated by noting that the solvation energy of a charge Q in a solvent cavity of radius $\bar{R}$ is given roughly by eq 18

$$G_{sol} = -166(Q^2/\bar{R})(1 - 1/\epsilon_0) \quad (18)$$

(Born, 1920), where the solvation energy G_{sol} is given in kcal/mol and ϵ_0 is the bulk dielectric constant. The value of ϵ_0 in a nonpolar environment (which contains no permanent dipoles) is ~2, and according to eq 18, the solvation energy is only about half of the corresponding solvation in water (ϵ_0 = 80). Since the solvation free energy of a typical carboxylic acid in water is about −70 kcal/mol (Warshel, 1979a), the solvation energy in a nonpolar environment of an enzyme should be about −70/2 kcal/mol, and the acid pK_a should be larger than 20. Thus, in any case that the pK_a of a given group of an enzyme is similar to the corresponding pK_a in solution ($pK_a = 4 \pm 3$), there must be significant stabilization of the ionized group by neighboring permanent dipoles of the enzyme (significant $V_{Q\mu}$). In such cases, as illustrated in Figure 3 and 4, the local enzyme environment resembles a cryptate rather than the nonpolar environment of an "oil drop". In some cases, the permanent dipoles which stabilize the ionized group are those of *bound* water molecules (which are an integral part of the folded protein) and not those of amino acids of the enzyme. This is, for example, the situation in the case of Glu-35 where one bound water molecule is found (by the approach of part 2 under Materials and Methods) between Glu-35, Try-111, and the O_6 of the sugar E residue. However,

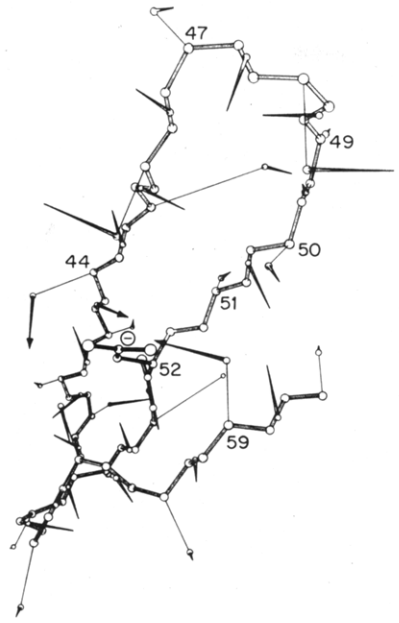


FIGURE 4: Permanent dipoles of the protein near Asp-52 in lysozyme. Each main-chain dipole includes the combined CO and NH peptide dipole. The dipoles of the side chains are drawn from their charge scaled "center of gravity".

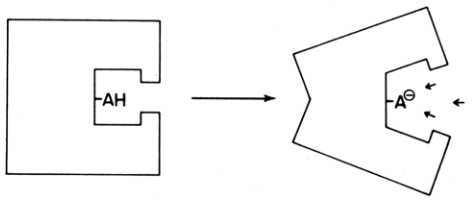


FIGURE 5: Showing schematically the effect of dissociation of a protein acidic group inside a nonpolar region of the protein. Since the ionized A^- is much more stable in water than in a nonpolar environment, the protein will unfold and allow water penetration.

in all cases, $V_{Q\mu}$ must be between −30 and −40 kcal/mol to account for the similarity between the pK_a of the groups in the enzyme and in aqueous solutions. In the absence of permanent dipoles around ionizable groups, they become unstable upon an increase of the external pH, leading eventually to partial unfolding of the protein and penetration of water to the neighborhood of the ionized group (Figure 5) [see also Perutz (1978)]. The equilibrium condition for this effect is given by eq 19, where ΔG_{unfold} is the unfolding free energy

$$2.3RT(pK_a^p - pK_a^w) \leq \Delta G_{unfold} \qquad (19)$$

of the protein and in most cases ≤ 10 kcal/mol. This gives an upper limit for the influence of the proteins on pK_a under equilibrium conditions. An interesting related example is presented by Glu-35 of lysozyme. The calculated pK_a of Glu-35 in the presence of a nonionized Asp-52 is ∼9, in reasonable agreement (considering the fact that we deal with large values of ΔG_{sol}) with the observed value of ∼7 obtained for the ester derivative of Asp-52 (Parsons & Raftery, 1972). However, the calculation of the pK_a of Glu-35 in the presence of the ionized Asp-52 gives the unrealistic value of $pK_a^p \simeq$ 19 (see Table I) as compared to the experimentally observed value (Parsons & Raftery, 1972) of $pK_a^p \sim 8$ and an experimental estimate of $pK_a^w \sim 5$ for glutamic acid in the (Asp^-, Glu) configuration in water (Table I). This means that the noncatalytic configuration (Asp^-, S, Glu^-) (where the substrate, S, is nonprotonated) is about $2.3RT(19 - 5) \simeq 19$ kcal/mol less stable than the corresponding configuration in water. Since our calculations used the observed X-ray structure and gave the correct pK_a for Glu-35 in the absence of the ionized Asp-52, the above inconsistency reflects most probably the penetration of water molecules to subsite D with partial displacement of the bound substrate. This mechanism of penetration of water molecules is reasonable since the binding energy of the substrate to subsite D is close to zero (Schindler et al., 1977), while the calculated difference between the stabilization of Glu-35 in water and in the ES complex is ∼19 kcal/mol. The above argument should not be confused with the consideration of the stabilization of the ionized Glu-35 in the catalytic configuration (Asp^-, SH^+, Glu^-) where the protonated substrate stabilizes the negatively charged Glu-35 (see below).

It is important to note that calculations which consider the protein as a continuous medium of low dielectric constant by using the theory of Tanford & Kirkwood (1957) or recent modifications (e.g., Matthew et al., 1979) cannot provide correct pK_a's for groups in the protein interior since they neglect the crucial effect of the protein permanent dipoles, $V_{Q\mu}$ (this will give an error of ∼20 pK_a units). Similarly, calculations that include only the protein permanent dipoles without considering the induced dipoles (e.g., Van Duijnen et al., 1979) cannot provide reliable pK_a's since they neglect the stabilization by the protein-induced dipoles, $V_{Q\alpha}$ (this will also give an error of ∼20 pK_a units).

(*2*) *Stabilization of Ion Pairs by Enzymes.* Transition states of bond-breaking reactions in solutions and in enzymes possess significant character of solvated ion pairs (Figure 2). The same is true for transition states in proton transfer reactions (AH + B → A^- + BH^+) (Warshel & Weiss, 1980). As shown in Figure 2, the free energy of an ion-pair type transition state is given relative to the corresponding nonpolar ground state by eq 20, where $\Delta = D_{XY} + I_Y - EA_X$ is the gas-phase energy

$$\Delta G^{\ddagger} \simeq \Delta + [G_{sol}(X^-Y^+)_{r=r^{\ddagger}_1} - G_{sol}(X\text{–}Y)_{r=r_0} + V_{QQ}] \qquad (20)$$

of forming X^-Y^+ at $r = \infty$ from X–Y, $Q_{QQ} \simeq -e^2/r^{\ddagger}_1$ is the gas-phase charge–charge energy of bringing the ions to $r = r^{\ddagger}_1$, and $G_{sol}(X^-Y^+)$ is the solvation of the ion pair at $r^{\ddagger}_1$. The simplest way an enzyme can modify $\Delta G^{\ddagger}$ is by changing the solvation free energy.

The large change in $\Delta G^{\ddagger}$ that the enzyme is capable of producing can be appreciated by noting that the solvation free energy of an electric dipole of magnitude μ in a dielectric medium is given roughly by eq 21 (Onsager, 1936), where $\bar{R}$

$$G_{sol} \simeq -166(\mu^2/\bar{R}^3)(2\epsilon_0 - 2)/(2\epsilon_0 + 1) \qquad (21)$$

is the radius of the solvent cavity around the solute. G_{sol} is given in kcal/mol for $\bar{R}$ in Å and μ in charge Å. The experimental estimate of the solvation energy of a dipole involving COO^- and NH_3^+ (or other charge groups of similar size) 3 ∼ apart in aqueous solution ($\epsilon_0 \sim 80$) is ∼−70 kcal/mol (Warshel, 1979a). A completely hydrophobic environment ($\epsilon \sim 2$) would give (according to eq 21) about half this stabilization, and an ionic transition state with $r^{\ddagger} \simeq 3$ Å would be ∼70/2 kcal/mol less stable in the nonpolar solvent than in water. Thus, the enzyme could, in principle, slow reactions drastically by surrounding the substrate with a nonpolar environment. This point is demonstrated schematically in Figure 6.

Since enzymes are designed to accelerate reactions rather than to slow them down, we conclude that an important feature of an enzyme is a polar active site that stabilizes an ionic transition state better than can water. Previous work (Warshel, 1978) has demonstrated that this is possible by comparing the energetics of charge separation in water and in a hypothetical enzyme. In this comparison, we adjust the assumed enzyme

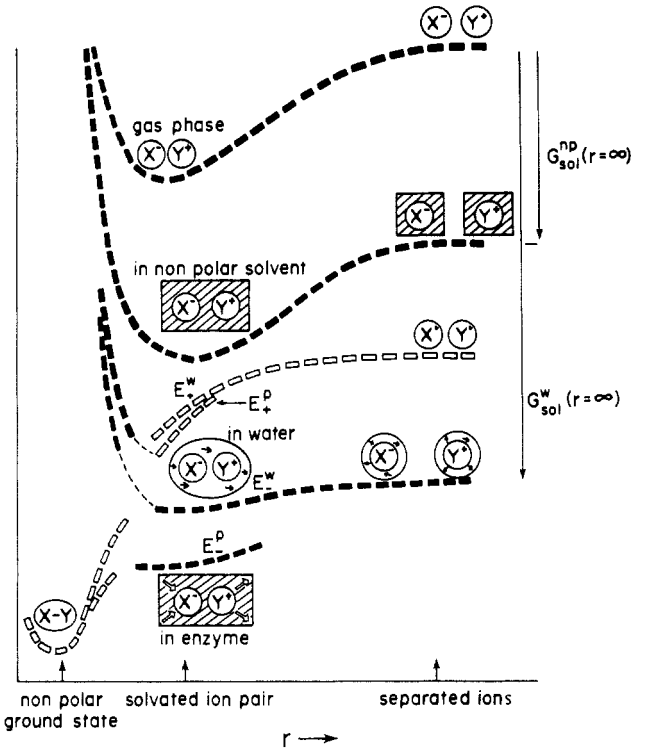


FIGURE 6: Relative stabilization of ionic transition states in different environments. Notation is as in Figure 2; G_{sol} is different for different environments. $E_{\pm}^{w}$ and $E_{\pm}^{p}$ are the ground and excited states in water and enzyme, respectively. G_{sol}^{np} and G_{sol}^{w} are the solvation free energies in nonpolar solvent and water, respectively. The ionic and covalent regions of the surfaces are indicated by closed and open rectangles, respectively.

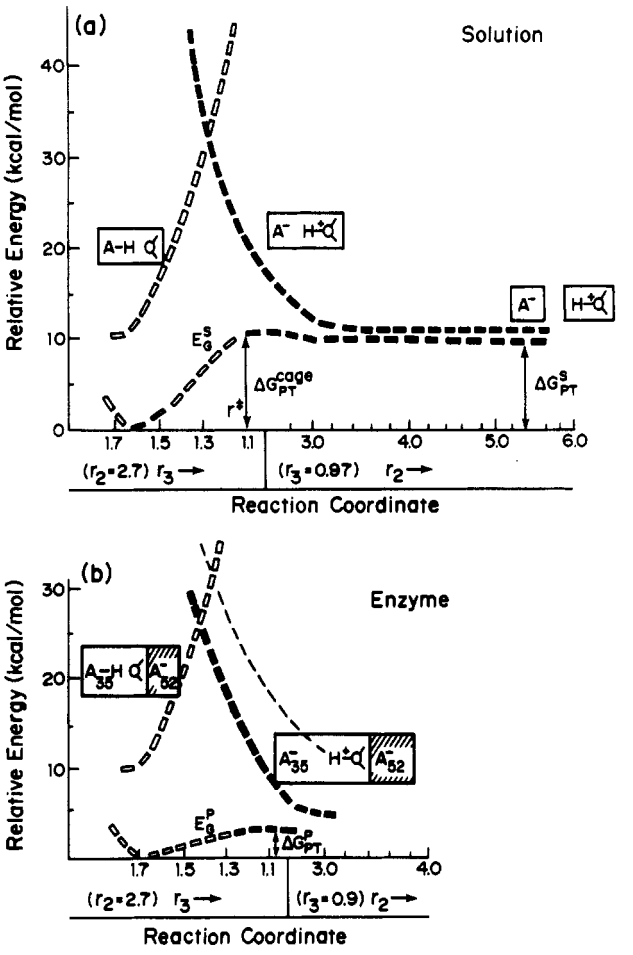


FIGURE 7: Potential surfaces for a proton-transfer reaction in solution and in an enzyme. (a) Potential surface for proton transfer between an acid (A) and a disaccharide molecule (R–O_4–R′) in solution. The surface is a simplified version of the calculations of Warshel & Weiss (1980). The figure presents the energies of the covalent (A–H O) and ionic (A^- H–O^+) resonance forms and the ground-state energy, E_G^s, obtained as a result of mixing of these resonance forms. The surface region where the ionic resonance form is the most important configuration is indicated by closed rectangles. r_2 and r_3 are the O–O_4 and H–O_4 distances (see Figure 9). ΔG_{PT}^s is the free energy for proton transfer at infinite separation of A and O and is calibrated to reproduce the corresponding observed value. ΔG_{PT}^{cage} is discussed in the text. The figure demonstrates that a proton transfer reaction can be described as a transfer from a covalentic to an ionic resonance form. (b) Potential surface for proton transfer from Glu-35 (A_{35}) to the O_4 of the substrate (see Figure 9) in the active site of lysozyme. The "solvation" of the ionic resonance from the enzyme includes the interaction with the ionized Asp-52 (A_{52}^-). E_G^p is the ground-state energy in the protein obtained as a result of mixing of the ionic and covalent resonance forms. The figure also presents for comparison the energy of the (A^- H–O^+) ionic resonance form in solution. It is clear from the figure that the difference between ΔG_{PT}^p and ΔG_{PT}^s is given approximately by the difference between the energies of the ionic resonance form in the enzyme and in solutions, which is dependent only on the difference in solvation energy of this resonance form in the enzyme and in solution.

permanent dipoles so that the solvation energy of the ions at infinite separation is equal to the corresponding solvation energy in water, thus implying $pK_a^p = pK_a^w$. Under this assumption, it was found that the solvated ion pair may be more stable in enzyme than in water. A more conclusive verification by careful experimental and theoretical studies is clearly needed.

(3) Proton Transfer Reactions in Solution and in Enzymes. *(a) Potential Surfaces for Proton Transfer Reactions.* Many classes of biological reactions involve proton transfer between different groups. These include the action of hydrolytic enzymes such as the serine proteases and lysozyme and proton pump systems such as bacteriorhodopsin (Warshel, 1979b). In this section, we will demonstrate how to obtain potential surfaces for such reactions.

A proton transfer reaction ($AH + B \rightarrow A^- + BH^+$) can be treated by the EVB approach by considering the three resonance forms in eq 22. The expressions for the energies and

$$\psi_1 = (A{-}H\ \ B) \qquad \psi_2 = (A^-\ \ H^+\ \ B)$$
$$\psi_3 = (A^-\ \ H{-}B^+) \tag{22}$$

interaction terms of these resonance forms are given in Warshel & Weiss (1980) where detailed evaluation of ground-state potential surfaces has indicated that the most important resonance forms are ψ_1 and ψ_3. Thus, for simplification of the present discussion, we consider ψ_1 and ψ_3 only and include their interaction with ψ_2 in their effective energies. In this case, the A–H bond in (A–H B) is represented as a real (partially polar) bond rather than a purely covalent bond, while $H{-}B^+$ in (A^- $H{-}B^+$) represents a $(BH)^+$ ion [with the charge distribution of $(BH)^+$ at infinite separation from A^-] rather than the pure $H{-}B^+$ resonance form. With the above simplification, the effective energy, E_3^s, of the ionic resonance form in solution is calibrated by requiring that its value at infinite separation of A^- and $(BH)^+$ be given [relative to the minimum of the ground-state energy of (A–H B)] by the observed ΔG_{PT}^s of eq 14. For consistency [see footnote b of Table I of Warshel & Weiss (1980)], the calculation of solution energy of (A–H) (B) at infinite separation includes (A–H OH_2) rather than (A–H).

Figure 7a represents a typical calibrated potential surface based on a simplification of the surface of Warshel & Weiss (1980). As shown in the figure, a proton-transfer reaction can be described as a simple crossing from a covalent to an ionic

Table II: Comparison of General Acid Catalysis in Solution and in Lysozyme[a]

	water cage			enzyme				
system	ΔG^{gas}	ΔG_{sol}	ΔG_{tot}	ΔG^{gas}	$\Delta V_{Q\alpha}$	ΔV_{Qq}	ΔG_{sol}^{p}	ΔG_{tot}
$[A^- ROH^+R]$	57	−47	$\Delta G_{PT}^{cage} \simeq 10$	57	−15	−40	−55	$\Delta G_{PT}^{p} \simeq 3$
$[A^-\text{--}R^+\text{-}OHR]$	78	−46	$\Delta G^{\ddagger}_{cage} \simeq 28$	78 + 5[b]	−17	−44	−63	$\Delta G^{\ddagger}_{cat} \simeq 19$

[a] All energies are given in kcal/mol. ΔG^{gas} is the gas-phase free energy for formation of the indicated systems from the equilibrium [AH–ROR] system. This value is obtained from the calibrated EVB potential surface by setting the solvation energy contributions to zero. ΔG^{gas} of $[A^-\text{--}R^+\text{--}OHR]$ includes an estimated contribution of −5 kcal/mol for the entropy associated with the motion of the fragments of the $R^+\text{--}OHR$ system in $[A^-\text{--}R^+\text{--}OHR]$. ΔG_{sol} is the calculated solvation energy of the ground-state charge distribution of the indicated system. The calculations simulated Glu-35 by a formic acid and the substrate by a disaccharide. $\Delta V_{Q\alpha}$ and ΔV_{Qq} are respectively the changes in inductive interactions and charge–charge interactions (relative to [AH--R–O–R]) between the indicated system and the enzyme active site which includes the ionized Asp-52. ΔG_{sol}^{p} is the sum of $\Delta V_{Q\alpha}$ and ΔV_{Qq}. ΔG_{tot} is the calculated energy of the EVB ground state which is given to a good approximation by $\Delta G^{gas} + \Delta G_{sol}$. [b] ΔG^{gas} in the water cage includes an estimate of −5 kcal/mol due to motion of the $[R^+\text{--}OHR]$ fragments. This contribution is not included in ΔG^{gas} in the enzyme case since there the motion of the fragments is probably restricted (Warshel, 1978).

resonance form. The energy of the ionic form is much more sensitive to the environment (solvent or enzyme) than the energy of the covalent form. This is demonstrated here for the catalytic reaction of lysozyme. In this reaction, a proton is transferred from glutamic acid 35 (Glu-35) to O_4 of the sugar residue in subsite D (Figure 9). The calibrated potential surface of the corresponding proton transfer reaction in solution is presented in Figure 7a. The potential surface for the enzymatic reaction is obtained from the EVB Hamiltonian by replacing the solvation energy term, G_{sol}^{s}, of the ionic resonance form in solution by the "solvation" energy, G_{sol}^{p}, of this resonance form in the protein (the interaction of this resonance form with the induced dipoles and residual charges of the protein atoms, including the negative charge of the ionized aspartic acid 52). The resulting potential surface for proton transfer in lysozyme is presented in Figure 7b. The corresponding comparison of ΔG_{PT} in the enzyme and the solvent cage is given in the first row of Table II. As demonstrated by Figure 7, the region of complete proton transfer is described to a good approximation by the ionic resonance form. The enzyme lowers the energy of this configuration relative to its energy in solution and reduces the free energy required for proton transfer.

As is demonstrated in Figure 7, the lowering of the activation energy, $\Delta G_{PT}(r^*)$, for proton transfer is almost linearly proportional to the stabilization of the ionic resonance form. Thus, studying the relation between the structure of the protein and the energetics of proton transfer reactions should focus on the structural basis for the electrostatic stabilization of the ionic resonance form. It appears that frequently asked questions about the degree of mixing between the ionic and covalent resonance forms at the transition state are only of secondary importance.

(b) "Back of the Envelope" Considerations of Proton Transfer Reactions. The analysis of the energetics of proton transfer reactions can be simplified to a level where it does not require any complicated calculations. This can be understood by following the argument below. The calculations of Warshel & Weiss (1980) indicated that for complete proton transfer (H–B bond of about 1 Å) the system can be described to a good approximation by the ionic resonance form ψ_3. Thus, it is reasonable to approximate the energy of proton transfer in a solvent cage by the sum of the free energy, ΔG_{PT}^{s}, of proton transfer at infinite separation between A and B and the free energy of bringing the ion pair into the same solvent cage. This gives

$$\Delta G_{PT}^{cage}(r^*) \simeq \Delta G_{PT}^{s} + V_{QQ}(A^- BH^+) + \Delta G_{sol}(A^- BH^+) \quad (23)$$

where V_{QQ} is the electrostatic interaction between the A^- and

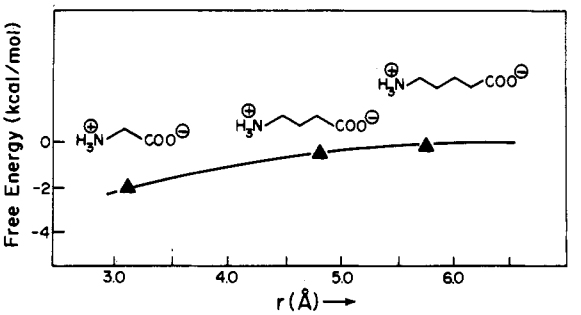


FIGURE 8: Experimental estimate of the energetics of charge separation in aqueous solutions. The figure uses ΔpK_a of zwitterions of amino acids to estimate the effective free energy of interaction between charged groups [see Warshel (1979a) for more details]. The figure demonstrates that the observed electrostatic interaction in solution is less than 3 kcal/mol for $r > 3$ Å.

BH^+ ions in the gas phase and ΔG_{sol} is the change in solvation free energy upon bringing the A^- BH^+ ions from infinite separation to r^*. It is useful to note that eq 23 can be written as

$$\Delta G_{PT}^{s} + 3 \geq \Delta G_{PT}^{cage}(r^*) \geq \Delta G_{PT}^{s} + V_{QQ}(r^*)/\epsilon(r) \geq \Delta G_{PT}^{s} - 3 \quad (24)$$

where $\epsilon(r)$ is the dielectric constant and the term $V_{QQ}(r)/\epsilon(r)$ is the effective free energy of electrostatic interaction [$(V_{QQ} + \Delta G_{sol})$ of eq 23] which is *experimentally* between 0 to −3 kcal/mol in solution, as is demonstrated in Figure 8. The upper limit of $\Delta G_{PT}^{s} + 3$ represents the possibility of an activation barrier at $r = r^*$ (this barrier should be less than the diffusion-controlled barrier). It is also useful to note that in view of the small electrostatic interaction for $r > 3$ Å, as is evident from Figure 8, the potential surface of proton transfer for $r > r^*$ is nearly a horizontal line. Since the stabilization of ψ_3 determines the overall energetics of proton transfer reactions, we can obtain a good approximation for the free energy of proton transfer in a protein active site, ΔG_{PT}^{p}, by the relation

$$\Delta G_{PT}^{p}(r^*) \simeq \Delta G_{PT}^{cage}(r^*) + [G_{sol}^{p}(r^*) - G_{sol}^{cage}(r^*)] \quad (25)$$

where G_{sol} is the solvation free energy of the $A^-\text{--}BH^+$ ion pair, "p" and "cage" designate the protein active site and a solvent cage, respectively, and the electrostatic interaction between the ion pair and the protein is viewed formally as a solvation energy. The first term in this equation is the energy of forming the ion pair in solution, and the second is the energy of transferring the ion pair from solution to the enzyme active site. Equation 25 can be approximated, by using eq 14 and 24, as

$$\Delta G_{PT}^{p} \simeq 2.3RT[pK_a^{w}(AH) - pK_a^{w}(BH^+)] + [G_{sol}^{p}(r^*) - G_{sol}^{cage}(r^*)] \pm 3 \quad (26)$$

Since the first term is available experimentally, it is only necessary to calculate $G_{sol}^{p} - G_{sol}^{cage}$. Equations 25 and 26 provide a good approximation for the free energy for proton transfer in an enzyme. For example, in the case of lysozyme, we obtain ΔG_{PT}^{p} of 3.0 and 4 kcal/mol by using eq 25 and 26, respectively, as compared to 3.5 kcal/mol obtained from the complete EVB calculations (Figure 7).

The above considerations indicate that the activation barrier for proton transfer reactions in enzymes can be studied by using only simple electrostatic calculations of G_{sol} without the need for complicated quantum mechanical procedures. This offers a simple and reliable way for examining such "textbook" reactions as the catalytic reaction of chymotrypsin (Blow et al., 1969) where one of the interesting questions is related to the energy of proton transfer from Ser-195 to His-57 in the presence of ionized Asp-102. This question can be studied by estimating the free energy, ΔG_{PT}^{p}, for proton transfer from the serine residue to the imidazole side chain (Im) by using eq 26 where ΔG_{PT}^{p} is given by the sum of the energy of forming the ion pair in solution (which is about 12 kcal/mol using pK_a's of 16 and 7 for serine and protonated imidazole) and the difference in solvation energies of the ion pair (Im-H^+--Ser^-) between enzyme and solution. This type of treatment (which is now in progress in our laboratory) involves little quantum mechanical calculation (quantum mechanical calculations are necessary only to obtain the isolated molecule charges) and is, at present, the most reliable way of studying the energetics of proton transfer in chymotrypsin.

Consideration of proton transfer reactions in proteins suggests a general concept in bioenergetics that related storage of energy to changes in pK_a^{p} (Warshel, 1979b). Thus, for example, proton transfer in lysozyme can be described as a transition between two configurations, $C^{(1)} = (A_{35}H,S,A_{52}^-)$ and $C^{(2)} = (A_{35}^-,SH^+,A_{52}^-)$, where S is the substrate and A_{35} and A_{52} are Glu-35 and Asp-52, respectively. According to eq 5 of Warshel (1979b), the free energy of proton transfer in the protein can be expressed as

$$\Delta G_{PT}^{p} = \Delta G^{(2)} - \Delta G^{(1)} = -\Delta G_{H^+}^{(2)}(A_{35}) + \Delta G_{H^+}^{(2)}(SH^+) \quad (27)$$

where $\Delta G^{(m)}$ is the free energy of forming the mth configuration and the $\Delta G_{H^+}^{(2)}$ are the free energies of bringing protons from water to the indicated sites of the protein in its $C^{(2)}$ configuration. Each of the two $\Delta G_{H^+}^{(2)}$'s includes half of the electrostatic interaction between A_{35}^- and SH^+ [see eq 3 of Warshel (1979b)]. Equation 27 can be expressed as

$$\Delta G_{PT}^{p} \cong 2.3RT[pK_a^{p(2)}(A_{35}) - pK_a^{p(2)}(SH^+)] \quad (28)$$

where the $pK_a^{p(2)}$'s include the interaction between A_{35}^- and SH^+. In the $C^{(2)}$ configuration, the protein decreases drastically the pK_a^{p} of A_{35} due to stabilization of A_{35}^- by SH^+ and increases the pK_a^{p} of the SH^+ due to its stabilization by A_{35}^-. It seems (see also next section) that the role of the enzyme is to minimize the pK_a^{p} difference between the proton donor and proton acceptor.

(4) Calculations on the Catalytic Reaction of Lysozyme. The accepted mechanism of the rate-limiting step of the catalytic reaction of lysosyme (Figure 9) consists of proton transfer from Glu-35 to O_4, cleavage of the protonated C–O_4 bond, and stabilization of the carbonium ion transition state by the ionized Asp-52. The rate k_{cat} of general acid catalysis by lysozyme is 1.7 s^{-1} at pH 5.25 (Chipman, 1971). This

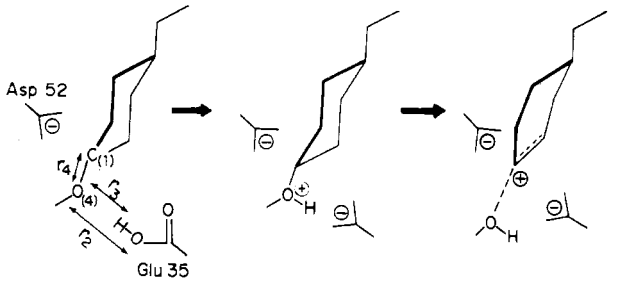


FIGURE 9: Catalytic reaction of lysozyme. The reaction involves proton transfer from Glu-35 to O_4 of the sugar residue in subsite D and cleavage of the protonated C_1–O_4 bond.

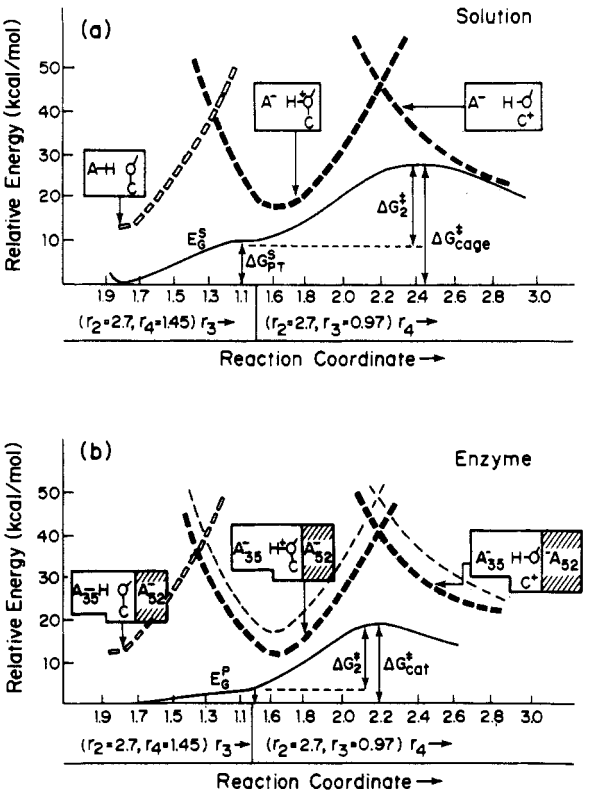
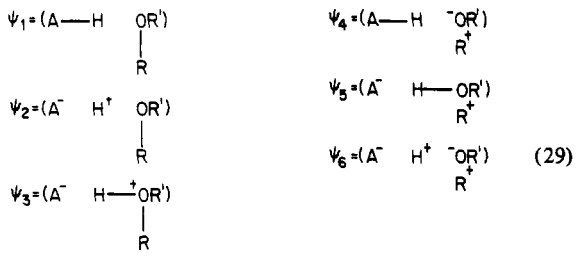


FIGURE 10: Potential surface for general acid catalysis in solution and in enzyme. (a) Potential surface for general acid catalysis of disaccharide cleavage in solution. The surface is a simplified version of the calculations of Warshel & Weiss (1980). The figure presents the energies of the three indicated resonance forms and the ground state, E_G^s, obtained from their mixing. The reaction coordinate involves the distances r_2, r_3, and r_4, which are defined in Figure 9. The change of r_3 for fixed r_2 and r_4 corresponds to proton transfer from A to O while the change of r_4 corresponds to cleavage of the C–O bond. The figure demonstrates that general acid catalysis reactions can be described as a transfer between resonance forms. (b) The energies of the important resonance forms and the ground state (E_G^s) obtained from their mixing, for general acid catalysis in the active site of lysozyme. The energies of the ionic resonance forms in solution are indicated by the closed rectangles and included for comparison. The figure illustrates that the difference in stabilization of the ionic resonance forms between the enzyme and solution is the reason for the difference between $\Delta G^{\ddagger}_{cat}$ and $\Delta G^{\ddagger}_{cage}$.

corresponds to an activation free energy, $\Delta G^{\ddagger}_{cat}$, of 18 kcal/mol calculated by using eq 2. The activation free energy of the reference reaction in the water cage, $\Delta G^{\ddagger}_{cage}$, is 26 ± 2 kcal/mol according to the estimate of a previous paper (Warshel & Weiss, 1980). In order to evaluate the difference between $\Delta G^{\ddagger}_{cat}$ in the enzyme and the corresponding $G^{\ddagger}_{cage}$ of the reference reaction in a solvent cage, we performed the calculations summarized in Figure 10 and Table II. In the calculations, we evaluated and calibrated the EVB Hamilto-

nian for the following configurations:



(29)

where A is a carboxylic acid and R and R′ are sugar residues. The calculations for the solution reaction are summarized in Figure 10a. For simplicity of the present discussion, we represent the main features of the potential surface by the three resonance forms ψ_1, ψ_3, and ψ_5 (the effect of the other resonance forms is incorporated into the effective energies of ψ_1, ψ_3, and ψ_5). It is apparent from the figure that the reaction can be described in terms of two steps: (1) a proton transfer step, which can be represented as crossing from the covalent resonance form ψ_1 to the ionic resonance form ψ_3 (the free energy of this step, ΔG_{PT}, can be estimated from pK_a measurements), and (2) cleavage of the protonated C–O bond, which can be described as a crossing from ψ_3 to ψ_5 [the activation free energy of this step, $\Delta G^{\ddagger}_2$, can be estimated from the observed rates of specific acid catalysis reactions (Warshel & Weiss, 1980)].

Using the calibrated EVB solution Hamiltonian, we evaluate the potential surface for the enzymatic reaction by replacing the solvation energy contributions, ΔG_{sol}, of the substrate–glutamic acid system in aqueous solution by its "solvation" by the remainder of the enzyme, including the ionized Asp-52. As shown in Figure 10b, the "solvation" by the enzyme stabilizes the ionic resonance forms ψ_3 and ψ_5 more than does aqueous solution; this results in a reduction of $\Delta G^{\ddagger}_{cat}$ relative to $\Delta G^{\ddagger}_{cage}$. Figure 10b demonstrates that the changes in the energies of the ionic resonance forms and the change in the activation energy of the reaction are strongly correlated and that simple considerations of the energies of the ionic configurations can account in this case for the energetics of enzyme catalysis. This seems to be true regardless of the controversial question of whether the ground state of the reacting system has 100% or 50% ionic character in the transition-state region.

If enzymes accelerate reactions by stabilizing ionic configurations, then their active sites must be designed to create an electrostatic potential that stabilizes the charge distribution of the relevant ionic configurations. This is clearly true for the reaction of lysozyme, where te ionized Asp-52 is located in an optimal place to stabilize ψ_3 and ψ_5. A similar effect is predicted in other enzymatic reactions that involve large rate accelerations. For example, a qualitative examination of the catalytic reaction of chymotrypsin (Warshel & Weiss, 1981) demonstrates that the negatively charged Asp-102 and the dipoles of the "oxyanion hole" are designed to stabilize the ionic resonance forms at the transition state.

The present analysis gives a new insight into the energetics of general acid catalysis in enzymes. The electrostatic potential from the protein at the active site can easily change ΔG_{PT} since the proton-transfer process involves a drastic change in the substrate charge distribution. On the other hand, $\Delta G^{\ddagger}_2$ is influenced much less by the protein electrostatic potential since the transfer from ψ_3 to ψ_5 upon partial cleavage of the protonated C_1–O_4 bond leads only to a shift of some positive charge from O_4 to C_1. This suggests an intriguing relation

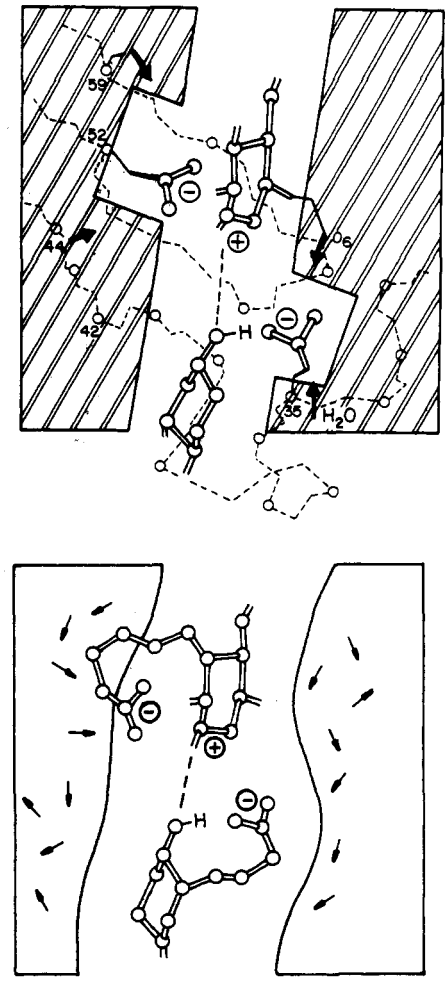


FIGURE 11: Comparing the reaction of lysozyme to that of an hypothetical model compound in a solvent cage [see also Warshel (1978)].

between the catalytic energy ($\Delta G^{\ddagger}_{cat} - \Delta G^{\ddagger}_{cage}$) and the change of p$K_a$ in enzymes. For example, using eq 28, we find that

$$\Delta G^{\ddagger}_{cat} - \Delta G^{\ddagger}_{cage} \simeq \Delta G_{PT}{}^{p} - \Delta G_{PT}{}^{s} \simeq 2.3RT[pK_a{}^{p(2)}(A_{35}) - pK_a{}^{w}(A_{35}) - (pK_a{}^{p(2)}(SH^+) - pK_a{}^{w}(SH^+))] \quad (30)$$

where SH^+ indicates the protonated O_4 of the substrate. Further studies of the relations between changes of pK_a and catalysis are clearly needed.

Additional insight into the energetics of general acid catalysis in lysozyme may be obtained if we compare the enzymatic reaction to a reference reaction of a hypothetical model compound with all the enzyme active groups, including Asp-52, in the same solvent cage (Figure 11). In this way, the difference between $\Delta G^{\ddagger}_{cage}$ and $\Delta G^{\ddagger}_{cat}$ is given approximately by the difference between the solvation of the [(–) (+) (–)] transition state in solution and in the enzyme active site. The relative stabilization of this transition state by the enzyme is calculated to be larger than in water [for details, see Warshel (1978)]. This means that the enzyme dipoles stabilize this system more than the water dipoles. The actual interaction of the enzyme dipoles with the [(–) (+) (–)] transition state is shown in Figure 11. A similar constellation seems to play a crucial role in the catalytic reaction of chymotrypsin.

One of the most interesting problems with regard to enzyme catalysis is the reason why different substrates react with very different k_{cat}. A possible clue to the nature of such effects may be obtained by comparing the calculated activation free energy by hydrolysis of *N*-acetylaminogluconolactone (NAG) and deoxy-NAG (where the O_6H is missing) (Osawa, 1968). The

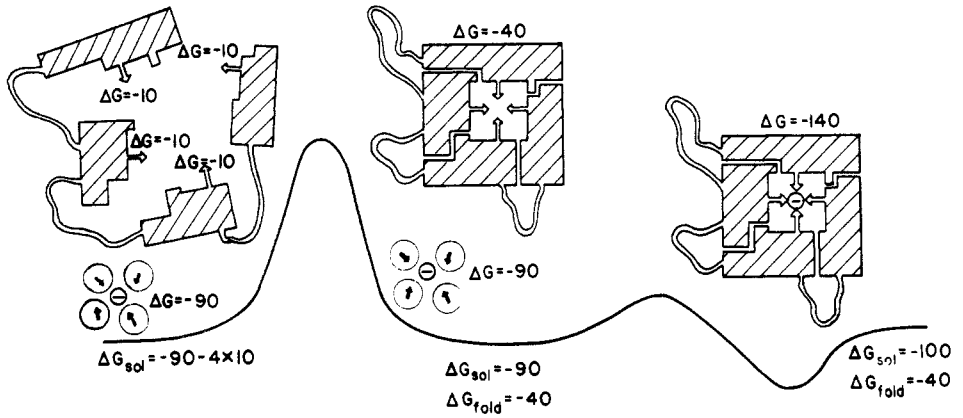


FIGURE 12: Relation between the folding of a protein and its role as a catalyst. A negatively charged ion is used to represent the transition state. The initial system (left) consists of an unfolded protein with dipoles solvated by the surrounding solution and the ion embedded in solution. Folding of the protein, driven by the interaction between the protein components, but opposed by the dipole–dipole interaction energy (middle), aligns the dipoles in an orientation able to stabilize the negative ion (right) better than water.

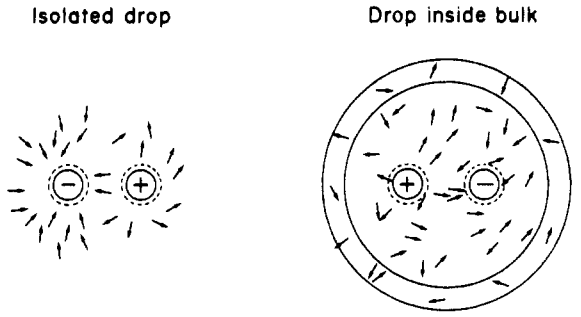


FIGURE 13: Demonstrating why bulk water is not such a good solvent; the polarization of water molecules toward the solute charges is reduced when the first solvation shells are surrounded by additional solvation shells.

calculations demonstrate that the absence of the O_6H dipole reduces the electrostatic stabilization of the ionized Glu-35 in the transition state by ~5 kcal/mol (see Figure 11). This results in about a 5 kcal/mol reduction in $\Delta G^{\ddagger}_{cat}$ as compared to an observed reduction of ~3 kcal/mol.

Concluding Remarks

This work presents a simple and reliable method for using X-ray structures of enzyme–substrate complexes to evaluate the activation energies of the enzymatic reactions. This method involves calibrating the potential surface of the reaction in solution and then adapting it to the enzyme reaction by replacing the solvation energies of the ionic resonance forms by the interaction with the enzyme active site. Our approach is demonstrated by studying several problems, including pK_a's of acidic groups, energetics of ion-pair-type transition states, proton transfer reactions, and general acid catalysis reactions. The calculations demonstrate that complicated enzymatic reactions can be studied by simply considering the change in stabilization of the ionic resonance form between the enzyme active site and a solvent cage without complicated quantum mechanical calculations. This point is best demonstrated in Figure 10b of this paper.

In addition to the development of a practical calculation scheme, we present in this paper further support to the suggestion (Warshel, 1978) that an enzyme can be viewed as "supersolvent" which can accelerate rates of reactions by providing optimal solvation to the ionic resonance forms at the transition state. As shown in Figure 11, this solvation is obtained by folding the enzyme dipoles in a cryptate-like arrangement. The relation between the folding of the protein and the storage of its catalytic energy is shown schematically in Figure 12. The figure demonstrates how the folding energy of the protein is used to arrange its dipoles in a relatively high energy configuration which provides optimal stabilization to an ionic transition state. This is analogous to the process whereby chemical energy is invested in synthesis of a cryptate from isolated fragments, and when the cryptate molecule is formed, it stabilizes the ion better than water does.

The argument that an enzyme can stabilize ionic transition states better than water can implies that the water dipoles cannot provide the "ultimate solvation". The reason for this might be due in part to the effect described in Figure 13 [see Warshel (1979a) for related calculations]. While the water molecules of small droplets of water tend to be oriented toward the charge solute, the presence of additional solvation shells reduces this stabilization; the solute can be solvated more effectively only at the expense of unfavorable interactions in the solvent. When an enzyme active site does not contain enough permanent dipoles, it can still stabilize ionic transition states by enveloping a hydrated substrate in a hydrophobic pocket and creating the equivalent of small droplets of water.

Previous work (Warshel, 1979b) has demonstrated that the formal consideration of an enzyme as a solvent provides a relation between the pK_a of the enzyme groups and its total energy. The present work indicates that this relation might be relevant to enzyme catalysis. For example, in the general acid catalysis reaction studied here, the difference between $\Delta G^{\ddagger}_{cat}$ and $\Delta G^{\ddagger}_{cage}$ is mainly due to the change in the pK_a's of the enzyme groups (eq 30). It appears that the enzyme change mainly the free energy of the proton transfer step, leaving the free energy of cleavage of the protonated bond the same as the corresponding free energy ($\Delta G^{\ddagger}_2$) in specific acid catalysis in solution. Further studies of the relation between the change of $\Delta G^{\ddagger}_{cat}$ by enzymes and changes of pK_a's are under way.

References

Blow, D. M., Birktoft, J. J., & Hartley, B. S. (1969) *Nature (London) 221*, 337.
Born, M. (1920) *Z. Phys. 1*, 45.
Chipman, D. M. (1971) *Biochemistry 10*, 1714.
Coulson, C. A., & Danielsson, U. (1954) *Ark. Fys. 8*, 239.
Fersht, A. (1977) in *Enzyme Structure and Mechanism*, W. H. Freeman, San Francisco, CA.
Matthew, J. B., Hanania, G. I. H., & Gurd, F. R. N. (1979) *Biochemistry 18*, 1919.
Onsager, L. (1936) *J. Am. Chem. Soc. 58*, 1486.

Osawa, T. (1968) *Carbohydr. Res. 7*, 217.
Page, M. I., & Jencks, W. P. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 1678.
Parsons, S. M., & Raftery, M. A. (1972) *Biochemistry 11*, 1623.
Perutz, M. F. (1978) *Science (Washington, D.C.) 201*, 1187.
Schindler, M., Assaf, Y., Sharon, N., & Chipman, D. M. (1977) *Biochemistry 16*, 423.
Tanford, C., & Kirkwood, J. G. (1957) *J. Am. Chem. Soc. 97*, 5333.
Van Duijnen, P. T., Thole, B. T., & Hol, W. G. J. (1979) *Biophys. Chem. 9*, 273.
Warshel, A. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5250.
Warshel, A. (1979a) *J. Phys. Chem. 83*, 1640.
Warshel, A. (1979b) *Photochem. Photobiol. 30*, 285.
Warshel, A., & Levitt, M. (1976) *J. Mol. Biol. 103*, 227.
Warshel, A., & Weiss, R. (1980) *J. Am. Chem. Soc. 102*, 6218.
Warshel, A., & Weiss, R. (1981) *Ann. N.Y. Acad. Sci.* (in press).

Enzyme-Catalyzed Condensation Reactions Which Initiate Rapid Peptic Cleavage of Substrates. 1. How the Structure of an Activating Peptide Determines Its Efficiency†

Marc S. Silver* and Susan L. T. James

ABSTRACT: The addition of a small peptide can significantly increase the rate at which pepsin cleaves a substrate at pH 4.5. Why? In order to find out, we have determined spectrophotometrically the relative ability of over a dozen peptides to speed the initial rate of disappearance of Phe-Trp-NH_2 and Leu-Trp-Met-Arg. Here are some of the criteria which establish the reliability of the acquired kinetic data: (1) rates depend linearly on $[E]_0$ and, to a good approximation, on [activator]; (2) measurements with both substrates yield the same ranking for the activators tested; (3) high-pressure liquid chromatographic investigations independently confirm conclusions derived from the spectrophotometric studies. The best activators found were Z-Ala-Phe and Z-Ala-Leu. At 3.2 mM they are respectively 60 and 30 times more effective than an equal concentration of Z-$(Ala)_2$. The two-step mechanism given below (for Phe-Trp-NH_2) best explains the structural specificity found, as well as other observations on the nature of these activated cleavages. It assumes that reaction commences when pepsin catalyzes synthesis of a peptide bond between activator and substrate. The polypeptide so formed subsequently undergoes scission at a different bond. The modified activator liberated, here designated Z-AA_2-AA_1-Phe, can eventually provide a variety of reaction products, as the succeeding paper demonstrates:

Z-AA_2-AA_1 + Phe-Trp-NH_2 ⇌ Z-AA_2-AA_1-Phe-Trp-NH_2

Z-AA_2-AA_1-Phe-Trp-NH_2 → Trp-NH_2 + Z-AA_2-AA_1-Phe

While the proteolytic enzyme pepsin[1] simply hydrolyzes some peptide substrates, it converts others to more complex product mixtures. Reactions of the latter type have puzzled chemists for over 20 years (Neumann et al., 1959). They are characterized by the formation of new peptide bonds at the expense of those present in the original substrate and are classified as either "acyl transfers" or "amino transfers". The transfer reactions are intriguing both because of their synthetic potential and because they appear to offer valuable insights into how pepsin functions mechanistically. Nevertheless a satisfactory resolution of the peptic mechanistic enigma has escaped chemists.

Recent reports on pepsin's acyl-transfer activity, primarily by T. Hofmann and his collaborators [e.g., Wang & Hofmann (1976a,b)] are most germane to the present work. These authors found that peptic cleavage of Leu-Tyr-NH_2 or Leu-Trp-Met-Arg at pH 3.4–4.7 afforded significant quantities of Leu_2 and Leu_3 in addition to leucine. Furthermore, introduction of millimolar concentrations of "activators", small peptides[2] such as Z-Leu-Met or Z-$(Leu)_2$, hastened the disappearance of the substrates and selectively favored their conversion into Leu_2 and Leu_3.

We set out to understand the basis for the activator effect with the hope that, in the process, we would learn something about pepsin's mechanism of action. A brief discussion of two alternative explanations for the activator effect will clarify why we pursued the particular experiments subsequently described.

One interpretation, advocated by Wang and Hofmann, starts with the assumption that the formation of acyl-transfer products in a reaction signals that acyl-pepsin intermediates have intervened. Equation 1 illustrates how such a mechanism

E + Leu-Trp-Met-Arg → Trp-Met-Arg + Leu-E (1a)

Leu-E + Leu-Trp-Met-Arg → $(Leu)_2$-Trp-Met-Arg → Leu_2 + Trp-Met-Arg (1b)

Leu-E + $(Leu)_2$-Trp-Met-Arg → $(Leu)_3$-Trp-Met-Arg → Leu_3 + Trp-Met-Arg (1c)

might explain the generation of Leu_2 and Leu_3 from Leu-Trp-Met-Arg; Leu-E represents the crucial acyl-enzyme intermediate. An activator such as Z-Ala-Leu (which has proven of great use in our work) produces its effect through secondary

† From the Department of Chemistry, Amherst College, Amherst, Massachusetts 01002. *Received September 18, 1980.* Supported by Grant PCM 78-17312 from the National Science Foundation.

[1] Only hog pepsin will be discussed, and all amino acid derivatives possess the L configuration.

[2] Abbreviations used: Pla, L-β-phenyllactyl; Phe(NO_2), L-*p*-nitrophenylalanyl; OP4P, 3-(4-pyridyl)propyl-1-oxy; APM, γ-aminopropylmorpholinyl; HPLC, reverse-phase high-pressure liquid chromatography; Z, benzyloxycarbonyl.

0006-2960/81/0420-3177$01.25/0