Semiquantitative Calculations of Catalytic Free Energies in Genetically Modified Enzymes†

J.-K. Hwang and A. Warshel*

Department of Chemistry, University of Southern California, Los Angeles, California 90089-0482

Received February 2, 1987; Revised Manuscript Received March 11, 1987

ABSTRACT: The catalytic free energy and binding free energies of the native and the Asn-155 → Thr, Asn-155 → Leu, and Asn-155 → Ala mutants of subtilisin are calculated by the empirical valence bond method and a free energy perturbation method. Two simple procedures are used; one "mutates" the substrate, and the other "mutates" the enzyme. The calculated changes in free energies ($\Delta\Delta G^{\ddagger}_{cat}$ and $\Delta\Delta G_{bind}$) between the mutant and native enzymes are within 1 kcal/mol of the corresponding observed values. This indicates that we are approaching a quantitative structure–function correlation. The calculated changes in catalytic free energies are almost entirely due to the electrostatic interaction between the enzyme–water system and the charges of the reacting system. This supports the idea that the electrostatic free energy associated with the changes of charges of the reacting system is the key factor in enzyme catalysis.

The "missing link" needed to correlate the structure and function of proteins is the ability to evaluate the corresponding activation free energies. The challenge of estimating activation free energies (Warshel, 1981a) has been emphasized recently by the emergence of genetic engineering. Now it is possible to assess various methods aimed at correlating structure and function by examining their performance in reproducing the effects of site-directed mutagenesis on enzyme catalysis. Such a study is the subject of this work, which takes as a test case the observed correlation between sequence and function in the catalytic reaction of subtilisin (Wells et al., 1986; Bryan et al., 1986).

Our calculations are based on the empirical valence bond (EVB) method (Warshel & Weiss, 1980; Warshel, 1981a; Warshel & Russell, 1986) and on a free energy perturbation method (Valleau & Torrie, 1977; Warshel, 1982, 1984a,b; Warshel & Russell, 1984; Warshel & Sussman, 1986; Mezei et al., 1985; Berendsen et al., 1985). The details of the method and a preliminary calculation are described elsewhere (Warshel, 1984a,b; Warshel & Sussman, 1986; Hwang & Warshel, 1987) (some essential points are given below). What is new and quite exciting is that the present calculations reproduce in an almost quantitative way the observed effects of several substitution experiments. This indicates that our results do not represent an accidental success and that the storage of catalytic free energies can be examined in a reliable way by computer simulation approaches.

Results and Discussion

The first test case is taken as the effect of the Asn-155 → Thr mutation on the hydrolysis of amides by subtilisin (Wells et al., 1986). The rate-determining step of this reaction is the formation of the tetrahedral intermediate [R′OC(O⁻)(R)X]

$$\text{Im} + \text{R}'\text{—OH} + \text{X—C(=O)R} \rightarrow \text{Im}^+\text{—H} + \text{R}'\text{—O}^- + \text{X—C(=O)R} \rightarrow \text{Im}^+\text{—H} + \text{R}'\text{OC(O}^-\text{)(R)X} \quad (1)$$

where Im and R′—OH designate His-64 and Ser-221. This rate-limiting step can be described as a transfer between three resonance structures:

$$\psi_1 = [\text{Im} \quad \text{H—O} \quad \text{C=O}]$$
$$\psi_2 = [\text{Im}^+\text{—H} \quad \text{O}^- \quad \text{C=O}] \quad (2)$$
$$\psi_3 = [\text{Im}^+\text{—H} \quad \text{O—C—O}^-]$$

where Im, O—H, and C=O are His-64, the hydroxyl group of Ser-221, and the carbonyl group of the substrate, respectively [for details, see Warshel and Russell (1986)]. To evaluate the changes in activation free energy in this reaction, we use a mapping potential of the form (Warshel, 1982, 1984a,b; Warshel & Russell, 1984)

$$V_m = \lambda_1^m\epsilon_1 + \lambda_2^m\epsilon_2 + \lambda_3^m\epsilon_3 \quad (3)$$

where the ϵ's are the force fields of the three resonance structures of eq 2. The λ's are the mapping parameters that change V in small increments from the reactant state to the product state. For each specified value of $\lambda_m = (\lambda_1^m, \lambda_2^m, \lambda_3^m)$ we run trajectories on the corresponding V_m and evaluate the difference in potential energy between V_m and the potential $V_{m'}$ that corresponds to a subsequent value of the vector λ. This difference is used to evaluate the change in free energy associated with moving from V_m to $V_{m'}$. The powerful relationship (Valleau & Torrie, 1977) used for this purpose is

$$\delta G(\lambda_m \rightarrow \lambda_{m'}) = -(1/\beta)\ \ln\ [\langle \exp[-(V_{m'} - V_m)\beta]\rangle_m]$$
$$\Delta G(\lambda_0 \rightarrow \lambda_N) = \textstyle\sum_m \delta G(\lambda_m \rightarrow \lambda_{m'}) \quad (4)$$

where $\beta = 1/k_BT$, k_B is the Boltzmann constant, λ_m and $\lambda_{m'}$ are successive values of the vector λ, and $\langle\ \rangle_m$ denotes the average obtained by running trajectories on V_m. Note that our mapping procedure is basically a process by which the substrate is transformed from one form to another ($\psi_1 \rightarrow \psi_3$).

The free energy function $\Delta G(\lambda)$ can be converted to the free energy surface associated with the true ground state (Warshel, 1984a). However, the corresponding procedure can be simplified considerably if one is interested in free energy differences at the transition state. Our simplified procedure will be described below.

The actual EVB parameters are given in Warshel and Russell (1986); the protein force field is given in Warshel et al. (1986). The solvent molecules are simulated by the surface-constrained all atom solvent method (SCAAS) (Warshel & King, 1985), where the solvent is confined to a sphere of 10 Å around the active site. This powerful and economical

† This work was supported by NIH Grant GM-24492, and part of the calculation was performed at the NSF San Diego supercomputer center, the use of which was supported by NSF Grant CHE-8519194.

0006-2960/87/0426-2669$01.50/0

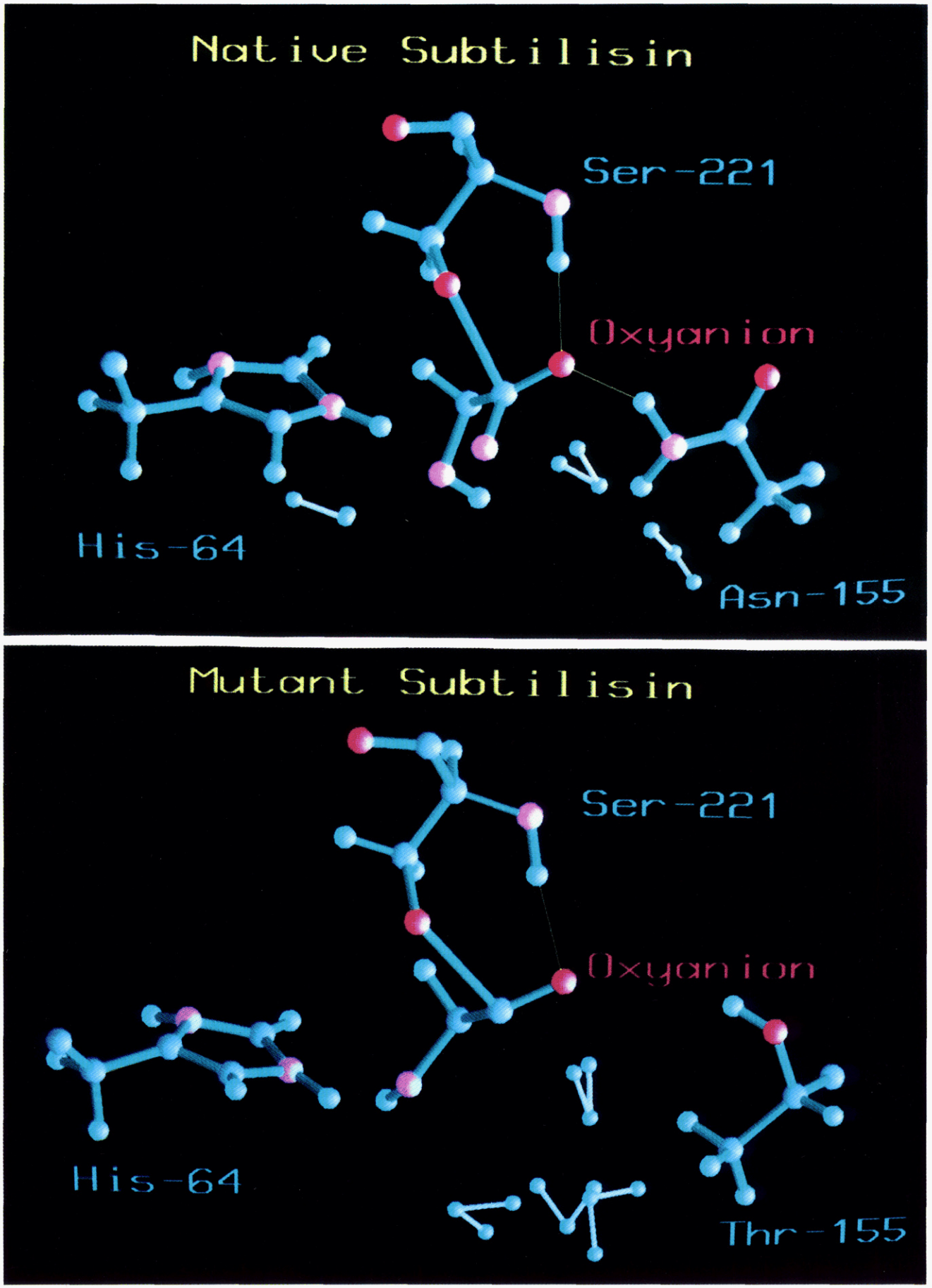

FIGURE 1: Typical configurations generated by simulation for native and Thr-155 mutant systems. The figure demonstrates that the electrostatic stabilization of the negatively charged oxyanion is smaller in the mutant case since one hydrogen bond is almost completely broken. The calculations were done by using a mapping potential that corresponds to the transition state region.

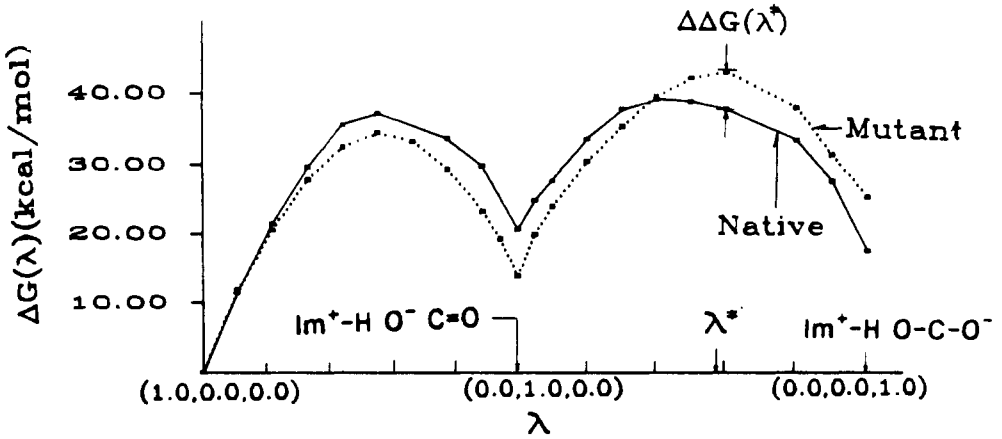

FIGURE 2: Calculated free energy profile for the mapping process that takes the system from the ground-state configuration to the oxyanion configuration (ϵ_3). The calculations were done for the native (solid line) and the Thr-155 mutant (dotted line) systems. The calculated free energy surfaces *do not* represent the energies of the real ground states obtained by mixing the various ϵ's [see Hwang and Warshel (1987) for a clear discussion]. To locate the transition state of the real ground state, we examined the trajectories obtained with different λ's and identified the $\lambda \sim \lambda^*$ that led to the most frequent intersection of ϵ_2 and ϵ_3. This gave $\lambda^* = (0.0, 0.4, 0.6)$.

approach is described in detail in Warshel et al. (1986). The coordinates of the enzyme were taken from Robertus et al. (1972), and the initial coordinates of the substrate were obtained by a model-building program [trying to reproduce the information given in Robertus et al. (1972)]. The enzyme–substrate–solvent system was allowed to equilibrate for 4 ps in the reactant state ($\lambda_1 = 1$, $\lambda_2 = 0$, and $\lambda_3 = 0$). The subsequent mapping process involves a gradual change of the λ's at the order indicated in Figure 2 (with 2 ps for each increment of 0.1 in λ).

Typical configurations generated during the simulation, using a mapping potential that corresponds to the transition state region, are shown in Figure 1. As seen from the figure, the native enzyme stabilizes the oxyanion configuration (ψ_3) by an hydrogen bond from Asn-155. This electrostatic interaction is practically missing in the mutant enzyme (the hydrogen-bond distance is much longer in the mutant). The possible importance of this stabilization was recognized in the original crystallographic study (Robertus et al., 1972). However, the crystallographic information cannot tell us about the actual catalytic contribution of the given structure element. This problem can be quite complicated even in the simple case of a substitution of a group that forms a hydrogen bond with the charged oxyanion. The hydrogen bond can be replaced by a bound water molecule or by an hydrophobic group, and the resulting changes in stabilization could be quite different in the two cases. The problem is much more complicated when the relation between the substitution and the energetics of the transition state is not obvious.

To convert structure to catalytic free energy, we use the free energy mapping calculations described in Figure 2. As seen from the figure, we obtain a change in catalytic free energy ($\Delta\Delta G^{\ddagger}_{cat}$) of 5.0 ± 0.6 kcal/mol, in excellent agreement with the corresponding observed value (4.7 kcal/mol) (Wells et al., 1986). It is important to note that the $\Delta\Delta G^{\ddagger}_{cat}$ obtained from Figure 1 is almost exactly reproduced by considering only the electrostatic energy contribution to $\Delta\Delta G^{\ddagger}$ (the change of electrostatic energy at the transition state is 4.8 ± 0.3 kcal/mol). It is also instructive to note that (see Figure 2) the catalytic difference between the mutant and native enzymes is manifested only at the transition state, where the carbonyl oxygen is significantly charged. Trying to evaluate this difference at other points along the reaction coordinate gives incorrect results. That is, the calculated $\Delta\Delta G^{\ddagger}$ is rather sensitive to the charge distribution of the transition state. For example, as seen from Figure 2, one can get a negative $\Delta\Delta G^{\ddagger}$ (instead of ~5 kcal/mol) for $\lambda^* \sim (0.0, 0.8, 0.2)$. This means that λ^* and the corresponding charge distribution should be evaluated in a consistent way. Finding the correct transition state of the system is quite complicated [the clearest explanation is given in Hwang and Warshel (1987)]. The problem is to locate the saddle points on the ground-state surface obtained by mixing ϵ_2 and ϵ_3. Since we are interested here only in the location of the transition state and not in the entire free energy surface, we use a simplified version of the EVB method. This is done by examining the trajectories associated with different λ's and finding the $\lambda \sim \lambda^*$ that leads to the most frequent intersection of ϵ_2 and ϵ_3; if the off-diagonal coupling between ϵ_2 and ϵ_3 changes slowly with λ, then the saddle points of the ground-state surface occur where $\langle \epsilon_2 - \epsilon_3 \rangle \sim 0$.

One of the major advantages of the EVB method is its ability to evaluate consistently the charge distribution of the reacting system in the given environment. Approaches that use gas-phase calculations to evaluate the charge distribution of the transition state [e.g., Chandrasekhar et al. (1984)] might lead to major errors, since the polarization of the reacting substrate by its surrounding environment is likely to change the charge distribution of the transition state. This problem, however, is less crucial in charge-transfer reactions (e.g., $\psi_2 \rightarrow \psi_3$) than in charge-separation processes (e.g., $\psi_1 \rightarrow \psi_2$).

Our second test case is the Asn-155 → Leu mutation (Bryan et al., 1986). Here we examine a variant of our procedure; that is, since the position of the transition state and its charge distribution were already evaluated in the calculations reported above, we can map the free energy by changing the protein rather than the reacting system. Now we use a mapping potential of the form

$$V_\theta(\lambda^*) = \theta_1 V_1(\lambda^*) + \theta_2 V_2(\lambda^*) \qquad (5)$$

where V_1 and V_2 are the force fields of the native and the mutant systems, respectively, at the transition state (λ^*) of the reacting system. Here the θ's are the mapping parameters that "mutate" the protein. These parameters play the same role as the λ's in eq 3, but different notation is used to distinguish between transforming the substrate and mutating the protein.

The V_θ potential gives (as shown in Figure 3) the free energy associated with changing Asn-155 to Leu, where the reacting substrate is kept at its transition state. Similar calculations

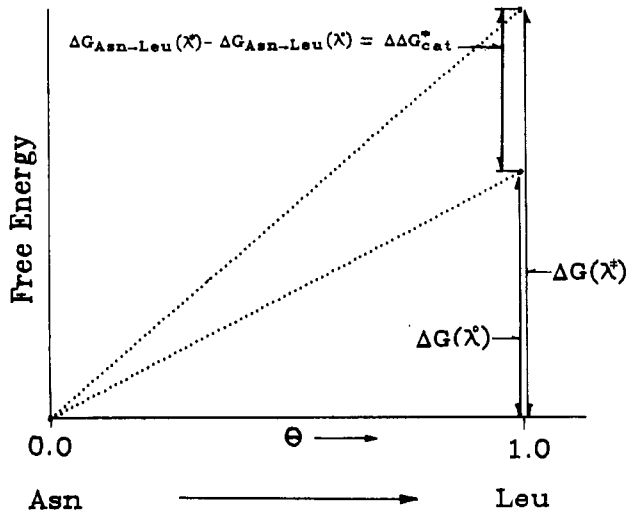

FIGURE 3: Schematic description of the free energy cycle involved in obtaining $\Delta\Delta G^{\ddagger}_{cat}$ by changing Asn-155 to Leu-155.

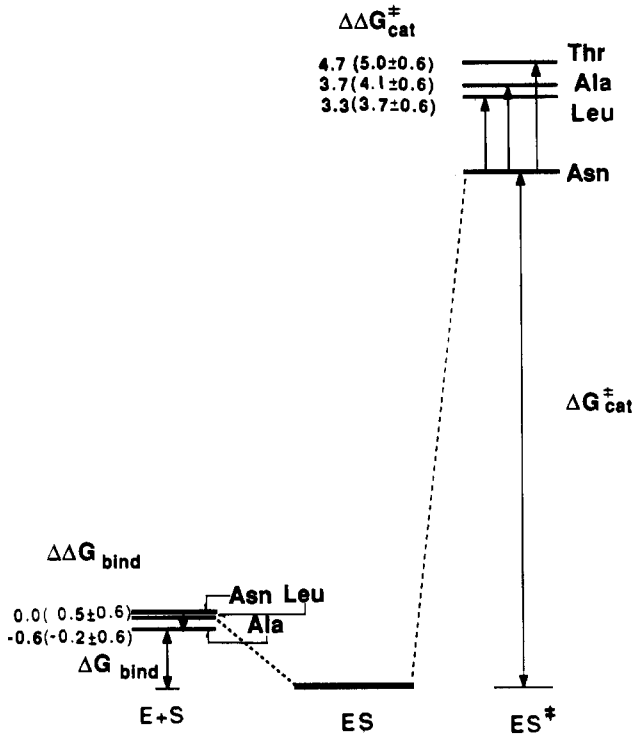

FIGURE 4: Calculated and observed free energy changes for the Asn-155 → Thr, the Asn-155 → Leu, and the Asn-155 → Ala mutations. The calculated free energies are given in parentheses near the corresponding observed values.

for the ground state complete the thermodynamic cycle needed to evaluate $\Delta\Delta G^{\ddagger}$. The same calculations for the mutation without the substrate give the change in binding free energy, $\Delta\Delta G_{bind}$. Here again we reproduce quantitatively the observed experimental results (Figure 4). It is interesting to note that also in this case we obtain basically the same results by only considering the electrostatic contribution to V_{θ}. Evaluating eq 2 by considering only electrostatic terms converges much faster than calculations that evaluate the full V_{θ}. This finding supports the view that electrostatic energies are the key factor in enzyme catalysis (Warshel & Russell, 1984) and explains why our early EVB calculations that used simpler solvent models (Warshel & Weiss, 1980; Warshel, 1981a; Warshel & Russell, 1984, 1986) gave reasonable estimates of catalytic energies.

Finally, we took as a test case the Asn-155 → Ala mutation. This calculation (Figure 4), which was completed *before* we knew the actual observed result, also reproduced the experimentally observed trend (J. A. Wells and D. A. Estell, personal communication) ($\Delta\Delta G^{obsd}_{cat}$ = 3.7 kcal/mol; $\Delta\Delta G^{obsd}_{bind}$ = −0.6 kcal/mol).

Conclusions

Our early examination of the current method (Warshel, 1984a) gave an estimated error of about 5 kcal/mol. Using longer simulation time in a study of the catalytic reaction of trypsin (Warshel & Sussman, 1986) reproduced the observed change in $\Delta G^{\ddagger}$, but this could reflect an accidental success. The present study, however, has examined additional test cases and reproduced the observed change in $\Delta G^{\ddagger}$ to within 1 kcal/mol. This might mean that we are entering a stage of quantitative structure–function correlation in macromolecules.

This work indicates that the changes in catalytic free energies can be estimated by calculating the electrostatic energy (solvation energy) of the *charges* of the relevant resonance structures. This reinforces the idea that enzymes can be viewed as "supersolvents" for the charges during the reaction and that the evaluation of the solvation energy of the reacting substrate by the enzyme–water system is the key for understanding enzyme catalysis (Warshel, 1981b).

Recent studies of the oxyanion hole and related systems have been discussed in terms of transition-state stabilization (Bryan et al., 1986). This general concept (Pauling, 1946) does not resolve several key problems. Obviously, catalysis is determined by the difference between the energies of the ground state and the transition state. The question is, however, how is this energy difference reduced? It appears that catalysis can be accomplished in a very effective way by a macrocyclelike environment with preoriented dipoles (Warshel, 1978). The resulting storage of catalytic energy is not just in the stabilization of the transition-state charge distribution but in the fact that the dipoles are already partially oriented in the proper direction. Thus the system does not have to invest a large free energy in polarizing its dipoles. Apparently, the catalytic free energy is stored in the unfavorable folding energy needed to prepolarize the protein dipoles. This point, which has been predicted in Figure 12 of Warshel (1981b), is now confirmed by our calculations. That is, the calculated free energy of changing Asn-155 to Leu in the E + S state is only 5.6 kcal/mol, while the corresponding change for the two groups in water is 12.5 kcal/mol. This means that the Asn dipoles are in an unfavorable environment in the folded native enzyme. Binding the charged transition state converts the unfavorable folding energy to a favorable dipole-charge energy [see Figure 12 of Warshel (1981b) for a clear illustration]. Our prepolarization concept is related to the concept of solvent "reorganization energy", which has been established as a key factor for reactions in polar solvents (Marcus, 1964; Hwang & Warshel, 1987).

Acknowledgments

We are grateful to Drs. J. A. Wells and D. A. Estell for communicating their results prior to publication.

References

Berendsen, H. J. C., Postma, J. P. M., & van Gusteren, W. (1985) in *Molecular Dynamics and Protein Structure* (Hermans, J., Ed.) pp 43–46, Polycrystal Book Service, Western Spring, IL.

Bryan, P., Pantoliano, M. W., Quill, S. G., Hsiao, H.-Y., &

Poulos, T. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3743–3745.
Chandrasekhar, J., Smith, S. F., & Jorgensen, W. L. (1984) *J. Am. Chem. Soc. 106*, 3049–3050.
Hwang, J.-K., & Warshel, A. (1987) *J. Am. Chem. Soc. 109*, 715–720.
Marcus, R. A. (1964) *Annu. Rev. Phys. Chem. 15*, 155–196.
Mezei, M., Mehrotra, P. K., & Beveridge, D. L. (1985) *J. Am. Chem. Soc. 107*, 2239–2245.
Pauling, L. (1946) *Chem. Eng. News 263*, 294–297.
Robertus, J. D., Kraut, J., Richard, A. A., & Birktoft, J. J. (1972) *Biochemistry 11*, 4293–4303.
Valleau, J. P., & Torrie, G. M. (1977) in *Modern Theoretical Chemistry* (Berne, B. J., Ed.) Vol. 5, pp 169–194, Plenum, New York.
Warshel, A. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5250–5254.
Warshel, A. (1981a) *Acc. Chem. Res. 14*, 284–290.
Warshel, A. (1981b) *Biochemistry 20*, 3167–3177.
Warshel, A. (1982) *J. Phys. Chem. 86*, 2218–2224.
Warshel, A. (1984a) *Pontif. Acad. Sci. Scr. Varia 55*, 59–81.
Warshel, A. (1984b) *Proc. Natl. Acad Sci. U.S.A. 81*, 444–448.
Warshel, A., & Weiss, R. M. (1980) *J. Am. Chem. Soc. 102*, 6218–6226.
Warshel, A., & Russell, S. T. (1984) *Q. Rev. Biophys. 17*, 283–442.
Warshel, A., & King, G. (1985) *Chem. Phys. Lett. 121*, 124–129.
Warshel, A., & Russell, S. T. (1986) *J. Am. Chem. Soc. 108*, 6569–6579.
Warshel, A., & Sussman, F. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3806–3810.
Warshel, A., Sussman, F., & King, G. (1986) *Biochemistry 25*, 8368–8372.
Wells, J. A., Cunningham, B. C., Craycar, T. P., & Estell, D. A. (1986) *Philos. Trans. R. Soc. London, Ser. A No. 317*, 415–423.