Acc. Chem. Res. **2001,** *34,* 72−79

Elucidating the Nature of Enzyme Catalysis Utilizing a New Twist on an Old Methodology: Quantum Mechanical−Free Energy Calculations on Chemical Reactions in Enzymes and in Aqueous Solution

PETER A. KOLLMAN,* BERND KUHN,
OREOLA DONINI,[⊥] MIKAEL PERAKYLA,[†]
ROB STANTON,[‡] AND DIRK BAKOWIES[§]
Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143

Received February 22, 2000

ABSTRACT
How do enzymes achieve very large rate enhancements compared to corresponding uncatalyzed reactions in solution? We present a computational approach which combines high-level ab initio quantum mechanical calculations with classical free energy calculations to address this question. Our calculations lead to accurate estimates of $\Delta G^{\ddagger}$ for both trypsin and catechol *O*-methyltransferase-catalyzed and reference uncatalyzed reactions and give new insights into the nature of enzyme catalysis. The same methodology applied to steps in the catalytic mechanism of citrate synthase further supports the conclusion that one need not invoke special concepts such as “low-barrier hydrogen bonds” or “pK_a matching” to explain enzyme catalysis.

Introduction

One of the fundamental and most interesting questions in the action of enzymes is, how do enzymes achieve their catalytic rate enhancement relative to the uncatalyzed reaction? This question is particularly timely in the era of protein engineering, catalytic antibodies, and structural genomics, because answers to this question should aid in the design of more effective and more useful catalysts.

Among the first conceptual analyses of enzyme catalysis were those of Haldane[1] and Pauling,[2,3] who suggested that enzymes bind the transition state for catalysis more strongly than the ground state and this binding enables catalysis. But how does the enzyme achieve this differential stabilization?

The most likely two kinds of enzyme−substrate noncovalent interactions that can lead to stronger binding of the transition state than the ground state are electrostatic (including hydrogen bonding) and van der Waals (steric and dispersion) interactions. As shown by Warshel and Levitt,[4] electrostatic interactions are likely to be the dominant factor.

However, this is not all that must be considered if one is comparing an enzyme-catalyzed reaction with a corresponding uncatalyzed reaction in solution. Page and Jencks[5] and others have emphasized entropic factors such as preorganization in bringing the reacting groups together to enable facile catalysis. Recent work by Bruice and Lightstone[6] has further analyzed such preorganization effects in intramolecular cyclization to form anhydrides.

Thus, excluding covalent catalysis, an enzyme can enable catalysis by binding more strongly to the transition state than the noncovalent enzyme−substrate complex and/or by preorganizing the catalytic groups and substrates upon binding, so that less of a free energy cost need be paid in proceeding with catalysis.

Are these two factors—noncovalent stabilization of the transition state and preorganization of the reacting groups by the enzyme—sufficient to explain enzyme catalysis? To answer this question, one must be able to simulate the enzyme-catalyzed reaction both in solution and in the enzyme and to quantitate the reaction free energies. This requires a combination of accurate quantum mechanical electronic structure calculations for those fragments where bonding is changing and accurate molecular mechanical models to represent the environmental effects of enzyme and solution.

Attempts to study enzyme catalysis using such a combination of quantum and molecular mechanics have been the focus of many theoretical studies, most notably in the work of Warshel and co-workers.[7] The main result of their work is that the key to enzyme catalysis is the stabilization provided to the transition state compared to the ground state, which is greater in the enzyme because the enzyme dipoles are preorganized in the enzyme structure, in contrast to the case in aqueous solution where one has to pay a free energy price to organize the water dipoles for transition state stabilization. Warshel based this conclusion on empirical valence bond (EVB)/free energy calculations. An EVB quantum mechanical model was calibrated to reproduce the reaction energetics in solution, and then the same parameters were used in the enzyme. The interaction of the reacting atoms with their environment was treated with classical free energy calculation methods.

A breakthrough in the theoretical study of organic reactions in solution using ab initio quantum mechanical methods occurred 15 years ago in the work of Jorgensen et al.[8] Their method involved studying the energetics of structures along a gas-phase ab initio reaction pathway using quantum mechanics and then calculating the relative solvation free energies of these structures by classical

Peter Kollman is Professor of Pharmaceutical Chemistry in the School of Pharmacy, University of California, San Francisco (UCSF), where he has been on the faculty for 29 years. Bernd Kuhn is a current postdoc; Oreola Donini, Rob Stanton, and Dirk Bakowies are former postdocs; and Mikael Perakyla is a former visiting faculty member at UCSF.

---

* To whom correspondence and reprint requests should be addressed.

[⊥] Current address: Kinetek Pharmaceuticals, Vancouver, B.C., Canada.

[†] Current address: Department of Chemistry, University of Kuopio, Finland.

[‡] Current address: Dupont, San Francisco, CA.

[§] Current address: Department of Physical Chemistry, ETH Zurich, Switzerland.

10.1021/ar000032r CCC: $20.00

free energy perturbation methods. In the interim, Severance and Jorgensen have studied many interesting organic reactions using this approach.[9]

We were attracted to both the ab initio aspect of Jorgensen's approach and the applicability to enzyme systems of Warshel's approach. Conceptually, the greatest difficulty in applying Jorgensen's approach to an enzymatic reaction is that, in organic reactions, the reactants typically involve few atoms, all of which can be treated quantum mechanically, and which can interact with the environment (solvent) in a nonbonded fashion. In contrast to organic reactions, there are usually covalent bonds connecting the chemically reacting parts with the rest of the protein in an enzyme-catalyzed reaction. We have been able to come up with a tentative solution to this "link atom" problem, and thus to generalize the Jorgensen approach to allow the ab initio study of both enzyme-catalyzed and solution reactions.[10] Some new and interesting insights on the nature of enzyme catalysis emerge as a result.

Methodology

The initial approach of Jorgensen and co-workers[11] was to apply eq 1 to the free energies ($\Delta G_{\mathrm{react}}$) along reaction potential surfaces, where ΔE(QM) is the relative gas-phase

$$\Delta G_{\mathrm{react}} = \Delta E(\mathrm{QM}) + \Delta G(\mathrm{FE}) \qquad (1)$$

ab initio energy of the chemical species undergoing the reaction (heretofore called the solute) and ΔG(FE) is the relative interaction free energy of different solute structures with the environment along the reaction pathway, calculated using classical free energy methodologies. Developments since that time in ab initio calculations have enabled the vibrational entropy along the reaction pathway of the solute to be calculated, turning ΔE(QM) into ΔG(QM), for which representative current software is described in ref 12. One could apply eq 1 to molecular systems using ΔE(QM) from semiempirical models,[13,14] but these are not likely to be as reliable or accurate as high-level ab initio or density functional theory (DFT) methods which include correlation; thus, one uses the highest level method one can to calculate ΔE(QM), which in recent years has often been a second-order Møller−Plesset perturbation theory (MP2) or DFT approach. However, caution must be applied to DFT-calculated transition state energies, which were underestimated by the B3LYP functional in a number of cases.[15]

The other main methodological question in applications of eq 1 is what to use for the nonbonded interaction model, most importantly the electrostatics of the solute, to calculate ΔG(FE). The initial applications[8,11] determined partial charges, q_i, along the reaction pathway by evaluating the interactions of a water molecule with the solute at various points along the reaction pathway and adjusting the solute charges to reproduce these energies, but recent work suggests that the ESP and RESP[16,17] approaches to derive the q_i of the solute are robust, straightforward, and accurate. Bakowies and Kollman have demonstrated this approach (using RESP or ESP charges) to be significantly more accurate than a simpler empirical approach for the formamide + OH^- reaction.[18]

What are the challenges in applying eq 1 to an enzyme-catalyzed reaction? There are three. The first is that one cannot simply use a coordinate approach on the gas-phase reaction surface to study a reaction of interest. For example, the formamide + $OH^- \rightarrow$ formate$^-$ + NH_3 reaction will, when following the gas-phase minimum energy pathway, form the species formamide$^-$ + H_2O rather than undergoing nucleophilic attack of the formamide C=O by OH^-. This problem can be partially solved by using recent developments in continuum solvation free energies added to the energies found in ab initio calculations,[19] in order to have the ab initio calculations follow a pathway that is relevant to the solution and the enzyme reaction. In the case of enzyme-catalyzed reactions, one needs to consider a set of structures of the reacting molecules that follows a pathway consistent with the thermally accessible structures of the active site. In our studies of trypsin,[10] catechol *O*-methyltransferase,[20] and citrate synthase[21] described below, we illustrate the use of active site constraints enforced by molecular mechanical methods to ensure that our calculated ab initio reaction pathway is sterically reasonable within the active site.

The second challenge of using the Jorgensen approach for enzyme-catalyzed reactions is simply computational—carrying out a sufficient number of ab initio calculations at a high enough level in order to accurately describe ΔE(QM) for all the atoms involved in the reaction. As we will see, one's ability to effectively use the smallest number of quantum mechanical atoms possible lets one meet this challenge and thereby use the most accurate model for ΔE(QM) with available computational resources.

The final and most difficult challenge is to deal with the "link atom" problem, the fact that the reacting atoms in an enzyme-catalyzed reaction are usually connected to the rest of the protein via covalent bonds. We have been able to deal with this problem using the RESP approach to derive the charges for the reacting atoms, with the Lagragian constraints inherent in this technology used to "splice" the charges of the reacting atoms into the framework of the rest of the protein. Such charges, if derived at the Hartree−Fock (HF) 6-31G* level, are internally consistent with the charges of the remainder of the protein. The above methodology enables us to carry out the calculation of ΔG(FE), including only the interaction free energy of the reacting atoms with the remainder of the protein and the aqueous solvent.

In the process of comparing solution and enzyme-catalyzed reactions, we realized[10] that eq 1 should be replaced with eq 2, where ΔG(cratic) is the free energy

$$\Delta G_{\mathrm{react}} = \Delta E(\mathrm{QM}) + \Delta G(\mathrm{FE}) + \Delta G(\mathrm{cratic}) \qquad (2)$$

required to align the reacting groups into a geometry enabling a facile reaction. This term describes the pre-organization free energy noted in the Introduction.

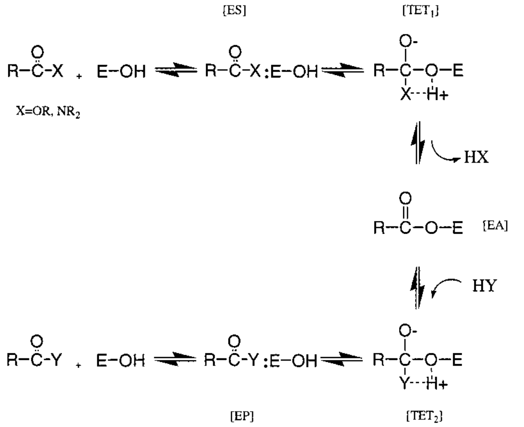

FIGURE 1. Schematic representation of the reaction pathway of serine protease catalysis of amide and ester bonds. E−OH represents the enzyme's hydroxyl group on Ser195. The following abbreviations are used: ES, Michaelis complex; TET1, first tetrahedral intermediate; EA, acyl enzyme intermediate; HY, a water molecule in our scheme; TET2, the second tetrahedral intermediate following water attack of the acyl−enzyme intermediate; and EP, the enzyme−product complex. For amides, the rate-limiting step is the formation of TET1.

ΔG(cratic) is not simply an entropy term; it includes the free energy to bring the molecules together and the free energy to orient them properly for reaction. Work by Hermans and Wang[22] describes how to calculate ΔG(cratic). Such a free energy contribution (ΔG(cratic)) will be critical for calculating the free energy barrier, $\Delta G^{\ddagger}$, for a solution reaction and should be significantly smaller for the enzyme-catalyzed reaction.

QM-FE Applied to Serine Protease-Catalyzed Hydrolysis

We first applied this new methodology to the hydrolysis of peptide bonds catalyzed by trypsin, a typical representative of the serine protease family.[10] The mechanism of amide hydrolysis is summarized in Figure 1 and is among the most studied enzyme-catalyzed reactions. Both experiment and theory[18] suggest that, for hydrolysis of an amide bond, the rate-limiting step is the formation of the first tetrahedral intermediate during the first step (TET1), which is acylation by the enzyme. Furthermore, there is much evidence that both the structure and free energy of the tetrahedral intermediate during acylation are very similar to those of the transition state for its formation from the initial noncovalent complex.[18,23] Given this, we carried out QM-FE calculations to estimate the $\Delta G_{\mathrm{react}}$ (eq 2) between the noncovalent (Michaelis) complex (MICO) and the tetrahedral intermediate (TET1) and related this to the observed $\Delta G^{\ddagger}$ for the trypsin-catalyzed reaction. The quantum mechanical (reactive) part of the system was methanol (representing serine 195), methylimidazole (representing histidine 57), and *N*-methyl acetamide (representing the scissile peptide bond) (Figure 2). The structures for the quantum mechanical calculation were derived by both molecular mechanical and quantum mechanical optimizations on MICO and TET1 models; in the case of the quantum mechanical optimizations, these were done with some dihedral restraints to ensure that the structures fit into the protein geometry. It is important to carry out the quantum mechanical optimizations for bond lengths and angles of the system, since these are critical for achieving the lowest possible ΔE(QM). Subsequently, we evaluated E(QM) (MICO vs TET1) at higher ab initio levels. Given the optimum structures for MICO and TET1, each of the sets of partial charges were fit with the RESP approach, using a united atom model for the carbons at the link positions (C_β for Ser and His and C_α and C_α' for the peptide bond) and Lagrangian constraints to ensure a net unit charge for each residue/fragment. Then, we used the GIBBS module of AMBER to calculate ΔG(FE) for mutating the system from MICO to TET. The results of these calculations are presented in Table 1.

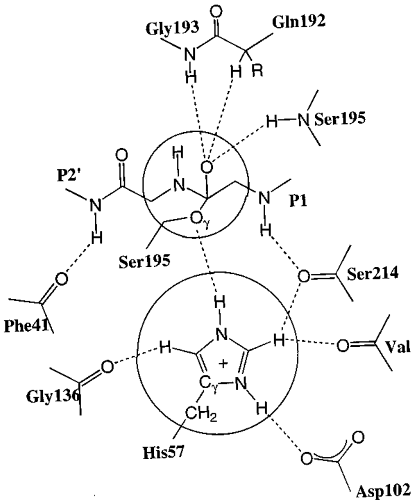

FIGURE 2. Schematic diagram of TET1, with the key residues noted. The backbone NH and side chain OH of Ser195 are not shown contiguously due to the 2D nature of this schematic. The atoms treated quantum mechanically are circled, with Ser195 and His57 including a methyl group attached to the Oγ and Cγ. In ref 16, the detailed free energy component analysis is presented for both a normal substrate and BPTI for the process MICO → TET.

The first point to note is that, with a normal substrate, $\Delta G_{\mathrm{react}}$ is calculated to be ∼16 kcal/mol, in good agreement with the experimental $\Delta G^{\ddagger}$. This ΔG comes about because, in the absence of environmental stabilization, it is quite unfavorable to go from the neutral MICO to the zwitterionic TET1 (ΔE(QM) $\approx$ 54 kcal/mol), while the environment significantly stabilizes TET1 (ΔG(FE) $\approx$ 38 kcal/mol). By calculating ΔG(FE) with the quantum mechanical residues in a box of TIP3P water molecules, we can estimate $\Delta G_{\mathrm{react}}$ for the corresponding solution reaction. Interestingly, ΔG(FE) is calculated to be only 6 kcal/mol less favorable for the solution than for the enzyme-catalyzed reaction, which is not consistent with the fact that the solution reaction at pH 7 is estimated to have a $\Delta G^{\ddagger}$ of ∼32 kcal/mol, about 16 kcal/mol larger than the enzyme-catalyzed value.

The difference comes in the free energy required to bring the reacting groups together in a productive geometry, something the enzyme has already done in forming its tertiary structure and binding the substrate. We suggest that the free energy price it has paid in binding the

Table 1. QM-FE Calculations on the Energetics (kcal/mol) of Amide Hydrolysis by Trypsin, MICO → TET

substrate	environment	ΔE(QM)	ΔG(FE)	ΔG(cratic)	$\Delta G_{\mathrm{react}}$
Ac-Ala-Phe-Arg-Ala-NH_2	enzyme[a]	53.8	−38.1	(0)	15.7
Ac-Ala-Phe-Arg-Ala-NH_2	solution[a]	53.8	−32.1	11.4	33.1
Ac-Pro-Cys-Lys-Ala-Arg-NH_2	enzyme[b]	53.8	−39.3	(0)	14.5
Ac-Pro-Cys-Lys-Ala-Arg-NH_2	enzyme[b]	53.8	−39.0	(0)	13.9
Ac-Pro-Cys-Lys-Ala-Arg-NH_2	enzyme (BPTI)[b]	53.8	−26.8	(0)	27.0

[a] Reference 4. [b] Reference 16.

substrate leads to a Michaelis constant, K_M, significantly larger than it would have been if one did not have to pay the entropy/desolvation free energy price for substrate binding. When ΔG(cratic) for the solution reaction is considered (~11 kcal/mol, Table 1), the calculated free energies for both solution and enzyme-catalyzed reactions are in good agreement with the experimental $\Delta G^{\ddagger}$.

In subsequent studies on trypsin catalysis, we have addressed two further interesting questions. BPTI, a trypsin inhibitor, binds to trypsin with the scissile bond of Lys15 poised to be hydrolyzed, yet BPTI acts as an inhibitor and is hydrolyzed $\geq 10^{10}$ more slowly than a typical trypsin substrate.[24] By evaluating ΔG(FE) for BPTI and comparing this with values for a substrate with the same sequence around the scissile peptide bond (Table 1), we can see that a significant fraction of this $\sim 10^{10}$ factor can be explained because BPTI loses a large amount of its favorable noncovalent interaction with trypsin on going from MICO to TET1, and thus $\Delta G^{\ddagger}$ for its hydrolysis is significantly increased. A further contribution to the reduced rate of catalysis of BPTI is likely to be its tendency to religate once the acyl−enzyme has been formed rather than to be hydrolyzed by water to form cleaved BPTI and regenerate the enzyme.

By evaluating the free energy contributions on a residue by residue basis, for both a typical substrate and BPTI, we have been able to delineate individual residue contributions to ΔG(FE) (MICO → TET1) in both cases.[24] Although the value of free energy component analysis has been a subject of much debate in the literature,[25,26] in this case, where the choice of pathway is rather unequivocal, we believe that significant useful insight is achieved.

Why are serine proteases not threonine proteases, when clear examples of threonine proteases exist?[27] Craik and Corey[28] showed that T195S trypsin had a k_{cat}/K_M of $\sim 10^{-6}$ that of wild-type, and free energy calculations by Perakyla and Kollman[29] show, indeed, that the ΔG(FE) for MICO → TET is ~8 kcal/mol larger for S195T than for wild-type trypsin. Free energy component analysis suggests that the most important reason why S195T is a less effective catalyst than wild-type is the unfavorable van der Waals interaction of the Thr 195 methyl with the sulfur of Cys 42. Perakyla and Kollman further showed that allothreonine, which involves a change in stereochemistry at the C_β of threonine, should be significantly more active than the normal stereoisomer.[29] We hope this prediction will be tested experimentally.

Catechol *O*-Methyltransferase

We have applied our QM-FE approach to the reaction catalyzed by catechol *O*-methyltransferase (COMT).[20] This reaction involves the transfer of a methyl cation from *S*-adenosylmethionine (SAM) to an aromatic O^- of a catechol molecule in a direct S_N2 process (Figure 3). COMT contains a Mg^{2+} ion that apparently shifts the pK_a of the two ortho-positioned OH groups of the catechol to stabilize the monoanion species (pK_a = 9.9) at pH 7, enabling the catalyzed reaction to be a simple methyl cation transfer to an anion, resulting in a neutral product (Figure 4).

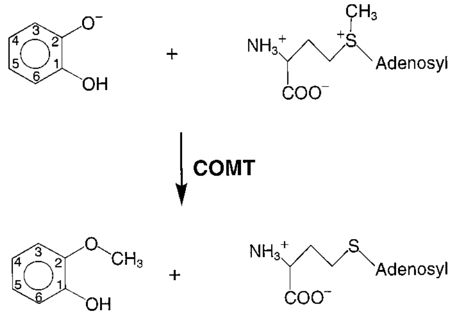

FIGURE 3. Reaction catalyzed by catechol *O*-methyltransferase (COMT).

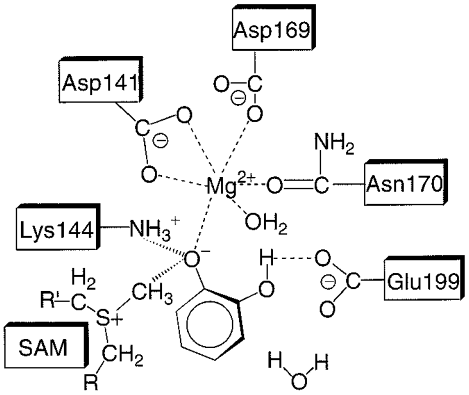

FIGURE 4. Schematic drawing of the active site of COMT showing the groups interacting in the structure just prior to reaction.

In this reaction, the only quantum mechanical atoms were the catechol and the $S(CH_3)_3^+$ fragment of SAM. After carrying out molecular mechanical optimizations on both the reactant and the product in the enzyme, one could restrain the reaction coordinate to interpolate between these structures. Along the same pathway, we carried out free energy calculations to obtain ΔG(FE).

Thus, in contrast to the case with trypsin, where we calculated only a single $\Delta G_{\mathrm{react}}$, corresponding to the free energy difference between the Michaelis complex and the tetrahedral intermediate,[10] in the case of the COMT reaction, we were able to calculate $\Delta G_{\mathrm{react}}$ along the reaction pathway. This led to a maximum in ΔE(QM) of about +10 kcal/mol relative to the reactants. On the other hand, ΔE(QM) for the products was ~ -45 kcal/mol, since in the absence of a polar environment neutral molecules tend to be much more stable than zwitterions.

When one considers the enzyme environment and ΔG(FE) along the reaction pathway, the reactants and products are of comparable free energy, and the maximum in $\Delta G_{\mathrm{react}}$ occurs at $\sim$24 kcal/mol. Considering both vibrational entropy effects and some deviations of the reaction surface from our linearly interpolated pathway leads to a $\Delta G_{\mathrm{react}}$ with a maximum at $\sim$21 kcal/mol, in respectable agreement with the experimental $\Delta G^{\ddagger}$ for COMT of 18 kcal/mol.

When one turns to the same reaction profile, evaluated in a box of water molecules rather than the enzyme environment, the maximum in $\Delta G_{\mathrm{react}}$ is only 5 kcal/mol larger than that in the enzyme. This is inconsistent with the difference in solution and enzyme-catalyzed rates of reaction of $\sim 10^{11}$, and agreement can only be achieved by considering ΔG(cratic) for the solution reaction, which is 9–13 kcal/mol, leading to good agreement between $\Delta G_{\mathrm{cat}}^{\ddagger}$ vs $\Delta G_{\mathrm{uncat}}^{\ddagger}$ determined experimentally and theoretically.

How does the enzyme stabilize the transition state for this reaction, compared to the solution reaction? In addition to the fact that it does not have to pay the cratic free energy "price", it can do so by not stabilizing the ionic reactants too much, by using a significant van der Waals component to bind the cofactor and substrate in an active site that, with the exception of Mg^{2+} and nearby charged groups that interact with the two aromatic OH functionalities, is quite nonpolar. Free energy component analysis also suggests that Met40 is a residue that stabilizes the transition state by a small ($\sim$1 kcal/mol) amount relative to the reactants and products. In net, compared to the gas-phase reaction, the enzyme does not stabilize the reactants relative to the transition state as much as water does and thus achieves $\sim$5 kcal/mol of lowering of $\Delta G^{\ddagger}$, independent of the cratic free energy.

Citrate Synthase

Citrate synthase (CS) is one of a number of enzymes that uses a weakly basic $-CO_2^-$ group to abstract a poorly acidic proton of a C–H group.[30] How these enzymes accomplish this has been the subject of much current interest and controversy. "Low-barrier hydrogen bonds" [31] and "matched pK_a's" [32–34] have been invoked to explain this, but in our view and that of a number of others, the enzyme is simply stabilizing the reaction with electrostatic interactions. In the case of triose phosphate isomerase (TIM)[35] or CS, the abstraction of a C–H proton leaves a relatively unstable enolate stabilized by a neutral histidine, whose pK_a is 12. Given the suggestion that this histidine "stabilizes" the enolate by transferring a proton to generate the enol, we evaluated the QM-FE energy for this process in CS, where His274 stabilizes the anion formed by abstraction of a terminal CH proton of acetyl CoA by Asp375. Consistent with earlier studies, which evaluated the QM energy and a simpler method to evaluate the enzyme environmental effect,[36] the $\Delta G_{\mathrm{react}}$ for enolate–enol was very positive ($\sim$10 kcal/mol). The free energy for forming the enolate was also much more favorable in the enzyme than in solution, further supporting the fact that CS (and, by implication, the other enzymes in this class) stabilizes the abstraction of a proton from C–H by electrostatic effects.

Summary of Methodological Issues

We have presented the application of a new/old methodology, QM-FE, to enzyme-catalyzed reactions and the corresponding solution reactions. There are a number of useful features of this approach compared to others in the literature.

First, this approach, like Warshel's EVB, allows a full simulation of a corresponding reaction in an enzyme and in solution. It has the advantage over EVB in that one does not have to calibrate the parameters for the solution reaction in order to study the enzymatic reaction—both can be studied from first principles.

Second, the RESP approach[16,17] to charge derivation seems to allow one to deal effectively with the link atom problem and use a minimal size of quantum mechanical atoms. This enables one to use high-level ab initio quantum mechanical calculations to evaluate ΔE(QM), which, we have shown,[10] is the largest potential source of error in calculating the free energy along the reaction profile, $\Delta G_{\mathrm{react}}$. As we have noted, the RESP approach with an HF 6-31G* model is also internally consistent and balanced with the force field model of the protein and the TIP3P water model, and this fact should lead to some error cancellation. The QM-FE approach also allows one to study many different phenomena (e.g., why BPTI is not a substrate, why Ser proteases are not Thr proteases[24,29]) without redoing the QM calculations, just the less demanding FE calculations.

There are more rigorous ways to study chemical reactions in condensed phases than the QM-FE approach presented here. One can use ab initio molecular dynamics,[37] but, currently, this is limited to relatively small numbers of atoms and short time scales. Alternatively, one can study either the entire enzyme or a significant fraction of it with efficient semiempirical molecular orbital approaches such as the divide-and-conquer strategy.[38] These models are not as accurate as high-level ab initio models but can be specifically reparametrized for some systems to give greater accuracy.[39] If one's goal is to study the mechanism of enzymatic reactions that involve radicals rather than ionic species, then a suitably truncated fully ab initio model appears to be adequate,[40,41] since ΔG(FE) is likely to be quite small for a radical reaction and ΔG(cratic) will also be quite small for a preorganized enzyme active site. There are also many other interesting studies which utilize QM/MM on enzyme systems.[42–45]

Nonetheless, the QM-FE approach has a number of novel features—first principle study of both enzyme and solution reaction and first quantitative inclusion of ΔG(cratic)—that make it a useful advance. The fact that

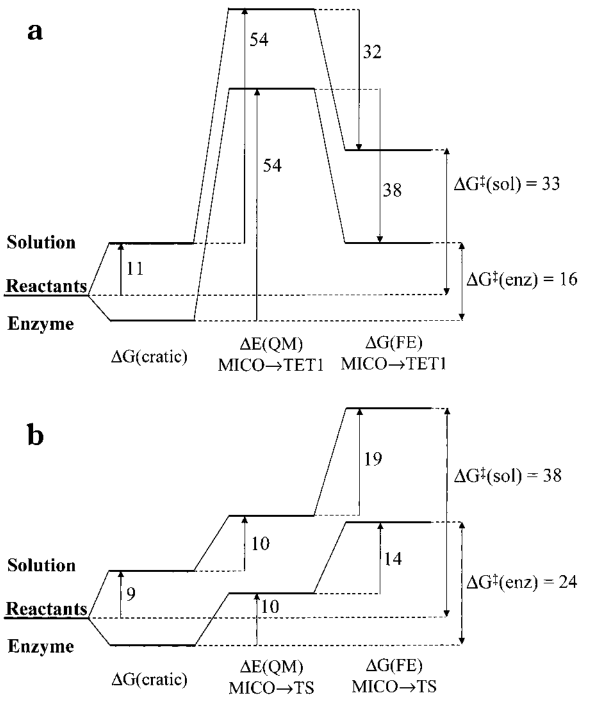

FIGURE 5. Schematic free energy diagrams for going from the noncovalent Michaelis complex (MICO) to the tetrahedral intermediate (TET1) in trypsin (a) and for going from the Michaelis complex to the transition state (TS) in COMT (b). Free energies are given in kilocalories per mole.

ΔG_{react} is in reasonable agreement with the experimental $\Delta G^{\ddagger}$ for both solution and enzyme-catalyzed reactions of trypsin and COMT suggests that errors in not using a quantum mechanical model for more of the enzyme are not large in these cases. Nonetheless, more enzymes have to be studied with this approach to ensure that the agreement between calculation and experiment is not fortuitous.

A reviewer of this paper has noted, “It is critical that [matching experimental free energies] not be the gold standard for calculational approaches. There are many examples in the literature of getting the right ΔG with the wrong atomic description. This is not surprising once it is recognized that ΔG is a single parameter from a very complex multi-variable reaction process: the ability to get one number correct, especially but not exclusively when that number is already known, is a necessary but inadequate test of a theoretical approach.” We agree with this reviewer that calculating a $\Delta G^{\ddagger}$ in agreement with experiment could be a fluke, but we must reemphasize that we have calculated $\Delta G^{\ddagger}$ in reasonable agreement with experiment for the enzyme-catalyzed and solution reactions of trypsin and COMT with no adjustable parameters.[46] However, we feel that calculating the free energy along reaction pathways *is* the gold standard. So far, our calculations have been mainly “postdictive” rather than predictive because we needed to assess how our QM-FE method would work. In the future, we hope to apply the method in many situations to assess the reasonableness of one mechanism over another, just as Mulholland and Richards[36] and we[21] have done in citrate synthase.

Summary of Conceptual Advances

The most important advance of the QM-FE approach is that we have been able to calculate $\Delta G^{\ddagger}$ for both solution and enzyme-catalyzed reactions in reasonable agreement with experiment and with no adjustable parameters. To achieve this, one has needed to include ΔG(cratic) for the solution reaction, while assuming it is zero for the enzyme-catalyzed reaction. Obviously, the assumption that $\Delta G(\text{cratic}) = 0$ for the enzymatic reaction is an extreme one. However, since we have also not explicitly considered the free energy associated with substrate binding in our estimate of ΔG_{react}, our implicit assumption is that any ΔG(cratic) is included in the substrate binding free energy which, as long as it corresponds to a negative free energy of binding, will not lead to an increased ΔG_{react}.

ΔG(cratic) is clearly significant for both the trypsin and COMT solution reactions ($\sim$10 kcal/mol) and thus is approximately two-thirds of the difference between the $\Delta G^{\ddagger}$ values for solution and enzyme-catalyzed reactions in both enzymes. In the case of trypsin, most of this free energy cost is due to having to orient three molecules (methanol, imidazole, and *N*-methylacetamide) for catalysis, whereas in COMT, a substantial fraction (about half) of ΔG(cratic) is due to the free energy cost of bringing the two reactants together, which is reasonable given that one is losing solvation free energy by associating a cation and anion.

Warshel has not considered the ΔG(cratic), arguing that it was appropriate to consider a preorganized “reference state” for both solution and enzyme-catalyzed reactions. We disagree, since to preorganize the solution reaction costs a significant free energy; in fact, the largest single component of the $\Delta\Delta G^{\ddagger}$ between solution and enzyme-catalyzed reactions for trypsin and COMT.

Warshel does summarize[7] as follows: “...the discussion of entropic factors might be very complicated and involves major semantic problems, such as the definition of the relevant reference state. Thus, it is essential to be able to *calculate* the actual entropic contribution to $\Delta G^{\ddagger}$ with well-defined potential surfaces.” We agree with the idea behind his last sentence, albeit we would substitute “cratic free energy” for “entropy”. It is the key paper by Hermans and Wang[22] that lets us accomplish this calculation.

Our results let us combine Warshel’s analyses that enzymes stabilize transition states electrostatically[7] better than solution reactions because of their preorganized charges and dipoles with Lan and Bruice’s NAC (near attack conformation) concept,[47] which focuses on how the enzyme preorganizes the reacting groups. Thus, the key to enzyme catalysis is related to the fact that the free energy for aligning the reactive groups is more favorable (less unfavorable) in the enzyme than in solution (ΔG(cratic)) and, second, that the enzyme can then stabilize the transition state for the reaction due to its preorganized environment (ΔG(FE)) better than the aqueous environment in the solution reaction. These two factors correspond to those suggested as the key to enzyme catalysis in the Introduction.

As to the "electrostatic stabilization",[7] one should emphasize, noting the COMT example, that for some reactions the enzyme interacts *less* strongly with its zwitterionic reactant and partially zwitterionic transition state than does water, and this offers a catalytic advantage. The gas-phase reaction has a calculated $\Delta G^\ddagger$ of ~10 kcal/mol, and this value is raised by putting the system into a polar environment. The enzyme achieves a smaller *increase* in $\Delta G^\ddagger$ by using a mainly hydrophobic active site to bind the substrate, with the exception of the Mg^{2+} to bind/orient the two OH groups of the catechol. Hydrophobic/van der Waals effects, which are known to be the predominant contribution to tight noncovalent binding of ligands to biological macromolecules,[48−50] enable the enzyme to use a less strong electrostatic interaction to bind SAM and catechol and thus destabilize the less polar transition state less relative to the zwitterionic ground state. Thus, COMT, where the transition state is less polar than the ground state, can use a different strategy than trypsin, where the transition state is more polar, in each case making ΔG(FE) less positive/more negative for the enzyme-catalyzed reaction than for the solution reaction. This is a generalization of the Warshel "preorganized dipole" [7] model of how the enzyme achieves stabilization of transition states in enzyme catalysis.

The recent solution of the crystal structure of orotidine 5′-phosphate decarboxylase (ODCase) highlighted the remarkable catalysis of enzyme vs solution in this case (10^{17}).[51,52] Since the reaction is a unimolecular decomposition, ΔG_{cratic} does not play a role. The major unresolved issue is whether the enzyme accomplishes its catalysis by "ground-state destabilization"[51] or by using a different mechanism in the enzyme than in solution.[53]

Finally, the fact that we have been able to calculate $\Delta G^\ddagger$ for both solution and enzyme-catalyzed reactions in two rather different cases, using a combination of high-level ab initio quantum mechanics and classical molecular dynamics/free energy calculations, further supports the idea that no special concepts such as "low-barrier hydrogen bonds" [31−34] need to be invoked to understand how enzymes achieve their enormous rate enhancement over solution reactions—the stabilization results from preorganization of the reactants (ΔG(cratic)) and preorganization of the enzyme active site (ΔG(FE)), utilizing "normal" noncovalent interactions. Now the continuing challenge will be to make use of the insights afforded by these theoretical studies in the design of more efficient catalysts and in understanding why the magnitude of the catalytic effect by catalytic antibodies is often significantly less than that for corresponding enzymes.[54,55]

P.A.K. is grateful to the NIH (GM-29072) for support of this work.

References

(1) Haldane, J. *Enzymes*; Longmans, Green and Co., MIT Press: New York, 1930; p 182.

(2) Pauling, L. *Chem. Eng. News* **1946**, *24,* 1375.

(3) Pauling, L. *Am. Sci.* **1948**, *36*, 51.

(4) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103,* 227−249.

(5) Page, M.; Jencks, W. Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68,* 1678−1683.

(6) Bruice, T.; Lightstone, F. Ground State and Transition State contributions to the Rates of Intramolecular and Enzymatic Reactions. *Acc. Chem. Res.* **1999**, *32,* 127−136.

(7) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Wiley: New York, Chicester, 1991. The key arguments about cratic free energies are summarized in section 9.4.

(8) Chandresekehar, J.; Smith, S. F.; Jorgensen, W. L. S_N2 Reaction Profiles in the Gas Phase and Aqueous Solution. *J. Am. Chem. Soc.* **1984**, *106*, 3049−3059.

(9) Severance, D. L.; Jorgensen, W. L. In *Structure and Reactivity in Aqueous Solution*; Cramer, C. J., Truhlar, D. G., Eds.; ACS Symposium Series 568; American Chemical Society: Washington, DC, 1994; pp 243−259.

(10) Stanton, R. V.; Perakyla, M.; Bakowies, D.; Kollman, P. A. Combined Ab Initio and Free Energy Calculations to Study Reactions in Enzymes and Solution: Amide Hydrolysis in Trypsin and Aqueous Solution. *J. Am. Chem. Soc.* **1998**, *120,* 3448−3457.

(11) Chandresekhar, J.; Smith, S. F.; Jorgensen, W. L. Theoretical Examination of the S_N2 Reaction Involving Chloride Ion and Methyl Chloride in the Gas Phase and Aqueous Solution. *J. Am. Chem. Soc.* **1985**, *107,* 154−163.

(12) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M.; Peng, C.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian Inc.: Pittsburgh, PA, 1998.

(13) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AMI: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107,* 3902−3909.

(14) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods. II. Applications. *J. Comput. Chem.* **1989**, *10,* 221−264.

(15) Glukhovtsev, M. N.; Bach, R. D.; Pross, A.; Radom, L. The Performance of B3LYP Density Functional Theory in Describing S_N2 Reactions at Saturated Carbon. *Chem. Phys. Lett.* **1996**, *260,* 558−564.

(16) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollrnan, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints For Deriving Atomic Point Changes—The RESP Model. *J. Phys Chem.* **1993**, *97,* 10269−10280.

(17) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. Application of RESP Charges to Calculate Conformational Energies, Hydrogen Bond Sinergies and Free Energies of Solvation. *J. Am. Chem. Soc.* **1993**, *115,* 9620−9631.

(18) Bakowies, D.; Kollman, P. Theoretical study of base-catalyzed amide hydrolysis: Gas- and aqueous-phase hydrolysis of formamide. *J. Am. Chem. Soc* **1999**, *121,* 5712−5726.

(19) Tomasi, J.; Cammi, R.; Mennucci, B. Medium effects on the properties of chemical systems: An overview of recent formulations in the polarizable continuum model (PCM). *Int. J. Quantum Chem.* **1999**, *75,* 783−803.

(20) Kuhn, B.; Kollman, P. A. QM-FE and Molecular Dynamics Calculations on Catechol O-methyltransferase: Free Energy of Activation in the Enzyme and in Aqueous Solution and Regioselectivity of the Enzyme Catalyzed Reaction. *J. Am. Chem. Soc.* **2000**, *122*, 2586−2596.

(21) Donini, O.; Darden, T.; Kollman, P. A. QM/FE Calculations of Aliphatic Hydrogen Abstraction in Citrate Synthase and in Solution: Reproduction of the Effect of Enzyme Catalysis and Demonstration that an Enolate rather than an Enol is Formed. *J. Am. Chem. Soc.,* submitted.

(22) Hermans, J.; Wang, L. Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme. *J. Am. Chem. Soc.* **1997**, *119,* 2707−2714.

(23) Weiner, S. J.; Singh, U. C.; Kollman, P. A. Simulation of Formamide Hydrolysis in the Gas Phase and in Aqueous Solution. *J. Am. Chem. Soc.* **1985**, *107,* 2219−2229.

(24) Perakyla, M.; Kollman, P. A. Why Does Trypsin Cleave BPTI so Slowly?" *J. Am. Chem. Soc.* **2000**, *122*, 3436−3444.

(25) Boresch, S.; Karplus, M. The Meaning of Component Analysis-Decomposition of the Free Energy in Terms of Specific Interactions. *J. Mol. Biol.* **1995**, *254,* 801−807.

(26) Smith, P.; van Gunsteren, W. When Are Free Energy Components Meaningful? *J. Phys. Chem.* **1994**, *98,* 13735−13740.

(27) Saarela, J.; Laine, M.; Tikkanen, R.; Oinonen, C. Activation and oligomerization of aspartylglucosaminidase. *J. Biol. Chem.* **1998**, *273*, 25320−25328.

(28) Craik, C. S.; Corey, D. R. An Investigation into the Minimum Requirements for Peptide Hydrolysis by Mutation of the Catalytic Triad of Trypsin. *J. Am. Chem. Soc.* **1992**, *114,* 1784−1790.

(29) Perakyla, M.; Kollman, P. A. Studies in progress.

(30) Kurz, L. C.; Nakra, T.; Stein, R.; Plungkhen, W. Effects of changes in three catalytic residues on the relative stabilities of some of the intermediates and transition states in the citrate synthase reaction. *Biochemistry* **1998**, *37,* 9724−9737.

(31) Frey, P. A.; Whitt, S.; Tobin, J. B. A Low-Barrier Hydrogen Bond in the Catalytic Triad of Serine Proteases. *Science* **1994**, *264,* 1927−1930.

(32) Gerlt, J. A.; Gassman, P. G. An Explanation for Rapid Enzyme-Catalyzed Proton Abstraction from Carbon Acids—Importance of Late Transition States in Concerted Mechanisms. *J. Am. Chem. Soc.* **1993**, *115,* 11552−11568.

(33) Gerlt, J. A.; Gassman, P. G. Understanding the Rates of Certain Enzyme-Catalyzed Reactions—Proton Abstraction from Carbon Acids, Acyl-Transfer Reactions, and Displacement Reactions of Phosphodiesters. *Biochemistry* **1993**, *32,* 11943−11952.

(34) Kozarich, J. W.; Gerlt, J. A.; Kenyon, G. L.; Gassman, P. G. Electrophilic Catalysis Can Explain the Unexpected Acidity of Carbon Acids in Enzyme-Catalyzed Reactions. *J. Am. Chem. Soc.* **1991**, *113,* 9667−9669.

(35) Alagona, G.; Ghio, C.; Kollman, P. A. Do Enzymes Stabilize Transition States by Electrostatic Interactions or PKa Balance—The Case of Triose Phosphate Isomerase (TIM). *J. Am. Chem. Soc* **1995**, *117,* 9855−9862.

(36) Mulholland, A. J.; Richards, W. G. Acetyl-CoA enolization in citrate synthase: A quantum mechanical molecular mechanical (QM/MM) study. *Proteins: Struct., Funct. Genet.* **1997**, *27,* 9−25.

(37) Kim, D.; Klein, M. Liquid hydrogen fluoride with an excess proton: Ab initio molecular dynamics study of a superacid. *J. Am. Chem. Soc.* **1991**, *121,* 11251−11252.

(38) Ermolaeva, M. D.; van der Vaart, A.; Merz, K. M. Implementation and testing of a frozen density matrix—Divide and conquer algorithm. *J. Phys. Chem.* **1999**, *103,* 1868−1875.

(39) Bash, P. A.; Ho, L. L.; Mackerell, A. D.; Levine, D. Progress Toward Chemical Accuracy in the Computer Simulation of Condensed Phase Reactions. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93,* 3698−3703.

(40) Wirstam, M.; Blomberg, M. R. A.; Siegbahn, P. E. M. Reaction mechanism of compound I formation in heme peroxidases: A density functional theory study. *J. Am. Chem. Soc.* **1999**, *121,* 10178−10185.

(41) Smith, D. M.; Golding, B. T.; Radom, L. Understanding the mechanism of B-12-dependent methylmalonyl-CoA mutase: Partial proton transfer in action. *J. Am. Chem. Soc.* **1999**, *121,* 9388−9399.

(42) Kolmodin, K.; Aqvist, J. Computational modeling of the rate-limiting step in low molecular weight protein tyrosine phosphatase. *FEBS Lett.* **1999**, *456*, 301−305.

(43) Alhambra, C.; Wu, L.; Zhang, Z. Y.; Gao, J. L. Walden-inversion-enforced transition-state stabilization in a protein tyrosine phosphatase. *J. Am. Chem. Soc.* **1998**, *120,* 3858−3866.

(44) Alhambra, C.; Gao, J. L.; Corchado, J. C.; Villa, J.; Truhlar, D. Quantum mechanical dynamical effects in an enzyme-catalyzed proton-transfer reaction. *J. Am. Chem. Soc.* **1999**, *121,* 2253−2258.

(45) Bala, P.; Grochowski, P.; Lesyng, B.; McCammon, J. A. Quantum dynamics of proton-transfer processes in enzymatic reactions. Simulations of phospholipase A(2). *Ber. Bunsen-Ges. Phys. Chem. Chem. Phys.* **1998**, *102,* 580−586.

(46) We note that there is some ambiguity over what $\Delta G^{\ddagger}$ actually is for the solution reactions. In trypsin, at pH 7, putting 1 M imidazole and methanol in the solution will not likely speed up the reaction over simple OH^- hydrolysis, which would be expected to have a $\Delta G^{\ddagger}$ of ~32 kcal/mol at pH 7 (22 kcal/mol at pH 14 with $[OH^-] = 1$ M). Thus, all one can definitively say is that the trypsin solution reaction should have a $\Delta G^{\ddagger} > 32$ kcal/mol at pH 7. For COMT, as noted in footnote 6 in Mihel et al. (Mihel, I.; Knipe, J. O.; Coward, J. K.; Schowen, R. L. α-Deuterium Isotope Effects and Transition State Structure in an Intramolecular Model System for Methyl-Transfer Enzymes. *J. Am. Chem. Soc.* **1979**, *101,* 4349−4351), the relative rates of reaction of the enzyme catalyzed and solution reaction for phenoxide reacting with $R_2S^+-CH_3$ is $\sim 10^{11}$. A value of 10^{16} is noted in this footnote, but this corresponds to using phenol, not phenoxide, acting as nucleophile. Our calculations[20] compare $\Delta G^{\ddagger}$ for phenoxide reacting with the $S-CH_3$ cation in both enzyme and solution and (Figure 5b) find this ratio to correspond to relative rates of $10^{10}-10^{13}$, in the same range as experiment.

(47) Lan, E. Y.; Bruice, T. C. Importance of correlated motions in forming highly reactive near attack conformations in catechol O-methyltransferase. *J. Am. Chem. Soc.* **1998**, *120,* 12387−12394.

(48) Miyamoto, S.; Kollman, P. A. Absolute and Relative Binding Free Energy Calculations of the Interaction of Biotin and its Analogues with Streptavidin Using Molecular Dynamics Free Energy Perturbation Approaches. *Proteins: Struct., Funct. Genet.* **1993**, *16,* 226−245.

(49) Miyamoto, S.; Kollman, P. A. What Determines the Strength of Noncovalent Association of Ligands to Proteins in Aqueous Solution. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90,* 8402−8406.

(50) Misra, V.; Honig, B. On the Magnitude of the Electrostatic Contribution to Ligand-Protein Interactions. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92,* 4691−4695.

(51) Wu, N.; Mo, X.; Gao, J.; Pai, F. Electrostatic Stress in Catalysis: Structure and Mechanism of the Enzyme Orotidine Monophosphate Decarboxylase. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97,* 2017−2022.

(52) Radzicka, A.; Wolfenden, R. A Proficient Enzyme. *Science* **1995**, *267,* 90−93.

(53) Lee, J.; Houk, K. A Proficient Enzyme Revisited: The Predicted Mechanism for Orotidine Monophosphate Decarboxylase. *Science* **1997**, *276,* 942−945.

(54) Thayer, M. M.; Olender, E. H.; Arvai, A. S.; Koike, C. K. Structural basis for amide hydrolysis catalyzed by the 43C9 antibody. *J. Mol. Biol.* **1999**, *291,* 329−345.

(55) Xu, J. A.; Deng, Q. L.; Chen, J. G.; Houk, K.; Bartek, J.; Hilvert, D.; Wilson, I. Evolution of Shape Complementarity and Catalytic Efficiency from a Primordial Antibody Template. *Science* **1999**, *286,* 2345−2348.

AR000032R