# Challenges and Advances in Validating Enzyme Design Proposals: The Case of the Kemp Eliminase Catalysis†

**Maria P. Frushicheva‡**, **Jie Cao‡**, and **Arieh Warshel‡,***
‡Department of Chemistry, 418 SGM Building, University of Southern California, 3620 McClintock Avenue, Los Angeles, California 90089-1062, USA

## Abstract

One of the fundamental challenges in biotechnology and biochemistry is the ability to design effective enzymes. Despite recent progress, most of the advances on this front have been made by placing the reacting fragments in the proper places, rather than by optimizing the preorganization of the environment, which is the key factor in enzyme catalysis. Thus, rational improvement of the preorganization would require approaches capable of evaluating reliably the actual catalytic effect. This work considers the catalytic effects in different Kemp eliminases as a benchmark for a *computer aided enzyme design*. It is shown that the *empirical valence bond* provides a powerful screening tool, with significant advantage over current alternative strategies. The insights provided by the empirical valence bond calculations are discussed emphasizing the ability to analyze the difference between the linear free energy relationships obtained in solution to those found in the enzymes. We also point out the trade off between reliability and speed of the calculations and try to determine what it takes to obtain reliable computer aided screening.

Advances in rational enzyme design are expected to have a great potential in industrial application and eventually in medicine (1). In fact, the ability to design efficient enzymes can be considered as the best manifestation of a true understanding of enzyme catalysis. However, most attempts of rational enzyme design, and the resulting constructs have been much less effective than the corresponding natural enzymes (1). Moreover, despite the progress in directed evolution (e.g., (2)), we do not have unique rationales for the resulting rate enhancements.

In our view the best way to identify the problems with the current rational design approaches (for review (1)) is to examine the actual performance of simulations of catalytic effects. Unfortunately, most current studies have not been based on simulations of the relevant activation barriers. In this respect, it has been argued (3, 4), that the problems in generating highly effective designer enzymes are due (at least in part) to the incomplete modeling of the transition state (TS) and to the limited awareness to the key role of the reorganization energy.

Capturing the reorganization energy by simulation techniques is, in fact, routinely done by the empirical valence bond (EVB) approach, which is a semiempirical valence bond type,

**Supporting Information Available**
Sample EVB runs of the reaction in solution and in 1A53-2 designed protein, and PDB structures of the reactant state (KE70_RS.pdb) and transition state (KE70_TS.pdb) of the KE70 (wild type) designed variant with 5-nitrobenzisoxazole substrate are available in the section "Kemp Elimination Reaction" in our website http://futura.usc.edu. Computational and modeling information is available free of charge via the Internet at http://pubs.acs.org/.

quantum mechanical / molecular mechanics (QM/MM) approach (e.g., (4, 5)). More specifically, semiquantiative computational studies of the effect of mutations on enzyme catalysis date back to the EVB simulations of the anticatalytic effect of mutations of trypsin (6). Subsequent calculations of known and predicted mutational effects include EVB studies (e.g., (4, 5, 7, 8)) and more recent molecular orbital-combined quantum mechanical / molecular mechanics (MO-QM/MM) studies (e.g., (9–12)). The above studies, and in particular the quantitative one, have established the importance of the changes in reorganization energy upon mutations (4, 8). Recent study (3) has explored the ability of the EVB approach to be used in quantitative screening of design proposals. More recently we also demonstrated the ability of the EVB to reproduce the observed effect of directed evolution in refining Kemp eliminases (13).

Here we will focus again on the Kemp elimination reaction (14), since the design of enzymes that catalyze this reaction has been the center of recent excitements (15, 16). At this point it may be useful to clarify that the ability to screen computationally different design proposals is not a trivial feature that is shared by all current approaches. That is, while the accomplishments of constructing artificial enzymes from scratch (e.g., (2)) are truly impressive, it is not clear if the final screening is based on a reliable "scoring function". Here the best way to make a clear judgment is to take cases where we *know* the observed catalytic effects and see which approach can reproduce them reliably. More specifically, while mutation predictions would be very convincing, having a clear benchmark (with known mutational results) and reproducing it by a given computer program should be a clear condition for a proper validation. With this perspective in mind, we can start by noting that gas phase calculations (or calculations with a small cluster of very few protein residues) are very unlikely to reproduce the corresponding observed catalytic effects. Furthermore, even recent attempts to use MO-QM/MM approach (e.g., (17)) have not provided a reasonable estimate of the observed catalytic effect or the trend of the mutational effects in artificially design enzymes. The same is true even for recent attempts to improve the MO-QM/MM calculations (18) and for ONIOM calculations (19) (see concluding discussion section).

Some of the problems with current perspectives in the field are illustrated by a recent Current Opinions (20), where the focus has been placed on catalytic proposals that are known to be ineffective (e.g., dynamics and ground state destabilization), while overlooking computational studies that actually reproduced the observed catalytic effect. Here the issue is not the different opinions on what might be important in catalysis, but the which effect has been found by proper computer modeling to contribute even mildly to catalysis where the results are known (see (4, 21)). In other words, if we are inserted in *computer aided enzyme design* (CAED), we should ask what was found in computations that actually examined all possible catalytic factors, and reproduce the observed catalysis, before we venture to consider options that were found to give a minor (or no) effect. The issue here is not what might contribute to catalysis, but what *does* contribute, and the judgment (in particularly when it comes to CAED) should be based, at least at some level, on the ability to reproduce the observed effect of the active site. This will help to eliminate effects that are not capable to lead to catalysis.

With the above perspective in mind, we followed the preliminary study of ref (13), but focused here on a more extensive validation of the power of the EVB in calculating the catalytic effect in Kemp eliminases, on exploring the effect of directed evolution and on the requirement for improving the catalysis of the previously designed enzymes. We also focus here on the origin of the difference between the linear free energy relationships (LFER) for the Kemp elimination reaction in solution and in the designed enzymes, as a way to better understand the problems in obtaining effective catalysis. Furthermore, in view of the

uncertainty about the binding mode of the substrate in the system used in ref (13) we put significant effort on studies of catalytic antibodies where we have direct information on the binding mode.

Overall, we demonstrate that our approach allows one to explore the effect of different mutations and in principle to predict which mutations can lead to enhanced catalysis. We also clarify the difference between our strategies to other less quantitative approaches.

## Systems and Methods

### Systems

The enzymes chosen as a benchmark for the present study catalyze the Kemp elimination reactions (14) with different proton donors and proton acceptors (the reaction studied are described schematically in Fig. 1). In order to demonstrate the considerations in determining the rate constants for the reference reactions in these systems we quantified first the energetics of the reaction in water, with 5-nitrobenzisoxazole as a donor and a carboxylic acid as a base (4). Here one may try to use the rate constant of the uncatalyzed reaction in water ($1.2 \cdot 10^{-6}$ M$^{-1}$ s$^{-1}$ at pH =7.25 (16)), but this reaction involves a large contribution from the hydroxide ion OH$^-$ available at this pH (14). In this respect we would like to clarify that we have no problem modeling pH effect in solution reaction (see (22)), but the well understood effect of OH$^-$ will be different at different pHs, and thus is not so useful in establishing a reasonable reference state (since the enzyme rate constant does not depend on the OH$^-$ concentration). In fact, we are interested in a reaction with a well-defined base (and the effect of the base concentration has been well understood quite early by the Biochemical community (see discussion in (4)). Thus, as discussed repeatedly (e.g., (4)) it is much more useful and relevant to consider the "chemically filtered" reference reaction (4), which is defined as the solution reaction that follows exactly the same mechanism as the one in the enzyme. With this in mind we determined the corresponding rate constant by taking the estimate (23) of the rate of the reaction of 5-nitrobenzisoxazole in acetate buffer ($5 \cdot 10^{-5}$ M$^{-1}$s$^{-1}$) and extrapolating it to 55M. The resulting rate constant ($k_{cage}$) was around 0.0025 s$^{-1}$, for the case when the donor and acceptor (the substrate and a carboxylic acid) are placed at the same solvent cage. Similar conclusions were obtained by other estimates, including *ab initio* calculations (see SI) where we obtained a barrier of about 17–18 kcal/mol, which is in the same range as the observed barrier (considering the fact that we did not include entropic corrections). At any rate, the relevant activation barrier for the reference reaction is given in Table 1. Similar analysis (see also SI) can also be performed with other bases or with other leaving groups and, in fact the relevant results can be taken from (24) but with the 55 M cage correction. It must be emphasized that we are not talking on $k_{cat}/K_M$ relative to the uncatalyzed reaction, where the effect of the enzyme is larger, since this rate enhancement includes the well understood effect of binding to the active site (4) and the real challenge is to optimize $k_{cat}$. Interestingly, and significantly, the reference reaction in a solvent cage is not much slower than the reaction catalyzed by the originally designed enzyme ($k_{cat}$=0.02–0.3 s$^{-1}$; see (16)). This indicates that a major part of the catalytic effect is due to just placing the donor and acceptor in a contact distance (see also (15)).

The enzymes studied here are the catalytic antibodies 34E4 (15) and 13G5 (25), the designed template structures KE70 and KE59 (16), the designed proteins HG2 and 1A53-2 (26);, human (HSA, (27)) and bovine serum albumins (BSA, (28)). The X-ray coordinates of the studied systems were obtained from the Protein Data Bank and are listed in Table 1. A typical system (namely the 34E4 antibody) is depicted in Fig. 2. The studied enzyme systems include 5-nitrobenzisoxazole, 5,7-dinitrobenzisoxazole, unsubstituted benzisoxazole and 6-glutaramidebenzisoxazole substrates (Fig. 1(b)) with different catalytic bases (Glu/Asp, Lys and His).

## Methods

The free energy surface of the reference solution reaction of 5-nitrobenzisoxazole with a carboxylic acid as a base was estimated by *ab initio* calculations, in the same way described in our previous works (e.g., (29)), and the resulting surface is shown in Fig. S1 of (13). The *ab initio* effective charges of the RS, TS and PS are given in Fig. S2 (13). The same approach was used for the other reference reactions.

The reaction in the protein was modeled by the EVB method whose feature have been extensively described in many previous works (see (30, 31) for a recent review) and is not described here. We like to clarify however, that the reliability of the EVB has been establishes in many cases including in reproducing an initio gas phase results (see 30, 31, (32)). We note, however, that the EVB surfaces are calibrated by forcing them to reproduce the above mention *ab initio* results for the solution reactions. In doing so, we have to take into account the complex nature of the reaction with a concerted proton transfer and C-N bond breaking. Usually this requires three diabatic states, but for the present purpose we decided to use two states with modified charges that reproduce the *ab initio* TS charges (see SI and (13)). This treatment is justified, since we forced the EVB TS to reach the same geometry, and more importantly the same charge distribution, as the corresponding *ab initio* solution system, with the concerted path. Thus the whole issue of catalysis is related to the difference between the reactant state and transition state and not the exact path between them. In other words, in the present work we are looking for a fast yet accurate method and once we realized that the standard two state charge distribution does not give the correct TS charges, we have two options: either to use three states (which is more expensive), or to use the current strategy, which guarantee the correct TS features, but requires one to introduce a correction for the energetics of moving from the transition state to the product state (which is not crucial to the present study). Of course, the second option is more reasonable, considering what we are looking for.

At any rate, the EVB calculations were carried by MOLARIS simulation program (33) using the ENZYMIX force field. The EVB activation barriers were calculated at the configurations selected by the same free energy perturbation umbrella sampling (FEP/US) approach used in all our studies (e.g., (4, 5)) The simulation systems were solvated by the surface constrained all atom solvent (SCAAS) model (33) using a water sphere of 18 Å radius centered on the substrate and surrounded by 2 Å grid of Langevin dipoles and then by a bulk solvent, while long-range electrostatic effects were treated by the local reaction field (LRF) method (33). The EVB region includes the substrate and the functional group of the proton acceptor (e.g., the carboxylic group of the glutamic acid). Validation studies were done within 22 Å radius of inner sphere, where we repeated the calculations of the activation barrier and obtained practically the same results (of course treating the distanced ionized groups with a high dielectric macroscopic model). The FEP mapping was evaluated by 21 frames (20 ps each) for moving along the reaction coordinate using SCAAS model. All the simulations were done at 300 K with a time step of 1fs. The simulations were repeated 20 times in order to obtain reliable results (see **Results** section) with different initial conditions (obtained from arbitrary points of the relaxation trajectory). The average was determining by taking in each case the difference between the calculated minimum at the RS and the given TS. The mutant systems were generated from the native enzymes via 100 ps relaxation runs.

In addition to the above EVB approach, it is useful to have faster screening approaches. Here our initial screening has been based on evaluation of electrostatic group contributions that are defined as the effect of "mutating" all the residual charges of the given group to zero. In principle, we could perform such mutations and evaluate the PDLD/S-LRA binding energy for the given native and mutant proteins (34). But since we are dealing with charged and polar residues, it is reasonable to start with the faster screening approximation

introduced and described in (35). This approach estimates the electrostatic contribution of different residues to the activation barrier by an expression that can be formally simplified by grouping together the atomic charges of each residue and written:

$$\Delta\Delta g_{elec}^{\neq} \cong 332 \sum_{ij} \overline{q}_j \Delta Q_i / r_{ij} \overline{\varepsilon}_{ij}$$

(1)

where $\overline{q}_j$ is the effective "charge" of the $j^{th}$ residue (which may represent dipoles in cases of polar residues), the $\Delta Q_s$ are the changes in the substrate charges upon going from the RS to the TS and $\varepsilon_{ij}$ is the dielectric constant for the specific interaction. Now we can explore the trend of Eq. 1 for any possible mutations if we artificially assign to all protein residues a charge of +1.0 and then ask what charge change will lead to the most negative $\Delta\Delta g_{elec}^{\neq}$. This is obtained by expressing the step in the optimization procedure step as in a steepest decent optimization of the barrier of Eq. 1. This leads to:

$$(\Delta\overline{q}_j)_{opt} \cong -\alpha\partial\Delta\Delta g_{elec}^{\neq}/\partial\overline{q}_j = -\alpha\sum_i \Delta Q_i / r_{ij}\varepsilon_{ij}$$

(2)

where now the $\Delta\overline{q}_j$ are proportional to the electrostatic group contributions, when all the protein groups are positively charged. In this case if $\Delta\overline{q}_j$ is positive the given residue should be either negative or have its dipole orientation back from the potential due to $\Delta Q_s$.

The treatment of Eq. 2 can be further refined by demanding that the mutations will also retain or optimize the protein stability. This can be done by our recent focused dielectric approach (35), which gives an approximated expression:

$$\Delta G_{fold} \cong 332 \sum_{ih} q_j q_h / r_{ih} \varepsilon_{focus}$$

(3)

where $\varepsilon_{focus}$ is the optimal dielectric constant (35) and the $q_s$ reflects the given pH.

We can now optimize $q_j$ under the constraint of large folding energy, and this lead to

$$(\Delta q_j) = -\partial G_{fold}/\partial q_j + \alpha\partial q^{\neq}/\partial q_j$$

(4)

The use of the combined TS stabilization and the stability constraints is expected to provide a quite effective way for guiding the mutations of distanced ionized residues.

It should be noted that the group contributions can only provide very rough hints, since they do not reflect the reorganization energy consistently. In fact, obtaining more quantitative group contributions is possible (especially for the actual given sequence by using the linear response approximation that captures the reorganization effect).

## Results

### Modeling the Catalytic Effects in Different Kemp Eliminases

The most basic requirement for effective enzyme design is the ability to reproduce the observed catalytic effects of different design constructs and/or those of natural active sites. Thus we started with a systematic evaluation of the activation barriers for the different

systems considered in Table 1. These systems involve different proteins, different mutants and different bases. Our typical procedure of obtaining the average activation barrier is illustrated in Fig. 3 and Table 2, where we show the results of 20 EVB free energy profiles. The same averaging procedure has been applied to all the system studied and the results of the calculations are summarized in Fig 4. The main point that emerged from the Table 1 and Fig. 4 is the fact that our EVB approach provides a reliable and effective way for screening different design options. In particular we would like to point out that our results are much more reliable than those reported by alternative approaches (see concluding discussion).

After establishing the reliability of our EVB approach, we turn our focus to attempts to explore possible avenues for controlling and changing the catalytic power of Kemp eliminases. As discussed in our previous work ((13), see also ref (36)) it is extremely hard to design an effective Kemp eliminase since the change in charge distribution upon moving to the TS is not large and furthermore this change is not localized. In fact, as found in our previous work (13), the main catalytic effect that has been obtained in Kemp eliminases (beyond the trivial effect of bringing the donor and acceptor to the same cage) is the destabilization of the ground state charge distribution of the proton acceptor. Although this is not the regular avenue taken by native enzymes it is the simplest direction for improving Kemp eliminases. At any rate, since this effect is associated with the pKa of the proton acceptor, one may ask what can be gained by changing the type of the donor / acceptor, and the corresponding analysis is given in the next section.

### Examining LFER Trends

An interesting and potentially promising aspect of our study is the ability to explore the linear free energy relationship (LFER) between the catalysis and the nature of the donor and the acceptor (the base). In solution we have a clear LFER for the 5-nitrobenzisoxazole as a function of the acceptor $pK_a$, for a series of different catalytic bases (e.g., pyridine catalytic base with $pK_a^{exp}=5.13$ and $\Delta g^{\ddagger}_{exp}=20.3$ kcal/mol) with $\beta=0.74$. This correlation is easily reproduced by the EVB model (see Fig. 5) by changing of the gas phase proton affinity (which is reflected by the EVB gas phase shift). Furthermore, we have a similar correlation between the activation barrier and the $\Delta pK_a$ of the donor and acceptor in the reaction in water (Fig. 6) and this correlation is also reproduced by the EVB. It is tempting to assume that we may obtain a similar correlation in the enzyme. Unfortunately, however, in the protein the situation is more complex. That is, while the gas phase proton affinity is identical in the protein and in solution, the local environment changes the apparent $pK_a$ in the protein (the calculated values are given in Table S2). This is further complicated by the fact that the environment of the proton donor is also changed by the protein. In view of this complication we start by first exploring the correlation between the observed $pK_a$ of the proton acceptor and the activation barrier (Fig. 7).

As seen from the figures the correlation is very qualitative, with large deviations. This is significant since we obtained much better correlation between the total calculated and observed activation barriers (Fig. 4). Thus we must conclude that the correlation between the apparent $pK_a$ (or $\Delta pK_a$) and the activation barrier cannot provide a simple guide improving the catalytic power of Kemp eliminases. More specifically, the correlation between the activation barriers and the $pK_a$ of the proton acceptor in the protein (Fig. 7) does follow the same trend as the LFER in solution (Fig. 5). Furthermore, although the LFER in the protein for $\Delta pK_a$ has a slop in the correct direction, the overall effect on the activation barrier is small since the $\Delta pK_a$ in the protein is much smaller than the corresponding $\Delta pK_a$ in water (Fig. 8). Thus the effect on the activation barrier is small. A significant part of the problem reflects the fact that an increase in the $pK_a$ of the acceptor, due to decrease in polarity of its site, is usually accompanied by an increase in the $pK_a$ of the donor (Fig. 8), due to the reduce in the polarity at the donor site. Of course, one can try to generate different polarities

in both sites, but due to the complex nature of the change in the charge of the donor upon TS formation, it is hard to obtain effective active site mutations that will stabilize the developing donor charge (see (13)). In other words it is much simpler, for example, to reduce the number of water molecules in the active site than selectively stabilize the ionized form of the proton donor. With this insight in mind we can look for a way to change the apparent $pK_a$ of the acceptor by external charges. This is done in the following session.

### Optimizing Long Range Electrostatic Effects

The above studies have established our ability to conduct an effective screening but did not address the issue of enzyme design. This issue was discussed in our previous works (3, 13) and in the specific case of the Kemp elimination reaction it has been argued that the refinement of the local environment cannot be effective, due to the small change in charge distribution during the reaction. Thus we focused on the small improvements of the catalytic effect by distanced ionized residues. That is, our previous studies (e.g., (4, 37)) have indicated that the effect of ionized residues at a distance of more than 6 Å can be approximated by using a relatively large effective dielectric constant for charge–charge interaction. Here we can use the approach introduced in our previous study (13), and described in the method section, where we optimize the effect of distanced ionized residues on the difference in charge distribution between the RS and TS, while simultaneously retaining the protein stability. This is done by using Eq. 4.

Here we report a qualitative attempt to use Eq. 3 in refining the 13G5 system. The calculated group contributions are given in Fig. 9 and Table S4, where we focus on the predicted effect of distanced charged residues. The most effective sites for distanced charged residues are also depicted in Fig. 10. For example, we would predict from Table S4 a rate acceleration of about 5 (TS stabilization is about 0.9 kcal/mol) due to the $V^{H98}E$ mutation, but this is done without exploring the effect on stability (by the use of Eq. 4)). The stability consideration should also take into account the fact that we already have two negatively charged groups ($E^{H6}$ and $D^{H72}$) at the same region. It may be simpler to explore our predication by simply perfuming the $E^{H6}Q$ mutation (where we predict a rate reduction by a factor of about 4).

We focus here on predicting the effects of distanced charged residues in view of the fact that the use of Eq. 1 for predicting the effect of active site residues (in the neighborhood of the substrate) is not expected to be particularly effective. Of course, we can use the full EVB calculations in the same way used in section ***Modeling the Catalytic Effects in Different Kemp Eliminases*** to predict catalytic mutations, but as discussed in (13), it is hard to obtain a major catalytic effects in the case of Kemp elimination reaction and thus it would be more effective to try to exploit the simpler strategy of using Eq. 1 for distanced residues. In this respect it is useful to consider the $His^{H95}Asn$ mutant, where the mutation of His to Asn reduces the rate significantly. However, here it is possible that in the case of the native enzyme $His^{H95}$ is the proton acceptor and $Asp^{H35}$ stabilizes the protonated $His^{H95}$, while in the case of the $H^{H95}N$ mutant $Asp^{H35}$ is the proton acceptor. The exact role of the Asp-His pair will be the subject of a subsequent study. Here we mainly bring this point to illustrate the difficulty in proposing modification of ionizable residues in the immediate neighborhood of the reacting system. On the other hand, the effect of distanced residues is much simpler to predict and the prediction is likely to be reliable.

### Examination of the Inactive Designs

One of the questions that should be addressed in enzyme design is the ability to eliminate design constructs that lead to inactive enzymes. This issue is explored here by calculating the activity of several inactive enzymes. The results are summarized in Table 3. As seen from the table we obtained high activation barriers for the non-active systems examined.

Obviously a more extensive examination is needed, however, despite the interest in predicting non-active constructs and eliminating them from design experiments, we place much more importance on the ability to improve the catalytic activity of mutants that are already slightly active. In fact, we believe that this is the direction where would be the most beneficial to use our quantitative methods.

## Discussion

The rational design of enzymes with native activity requires the ability to predict the proper TS stabilization, and this involves the challenge of capturing the overall preorganization effect. Attempts to estimate the catalytic effect by using gas phase models, or even by looking at the electrostatic interaction between different residues and the TS, are unlikely to reproduce the correct catalytic effect since it is impossible to assess the preorganization effect without including the protein and simulating its reorganization during the reaction.

The challenge of evaluating the catalytic power of a given mutant is not different than that addressed in our early 1986 study of computer aided mutations (6). At this stage it seems to us that the potential of the EVB has been demonstrated in well defined cases (e.g., (3, 38)), where it was found to reproduce the large effects of mutations that destroy the catalytic effect of evolved enzymes. Thus our main current challenge is to use this approach in improving non-efficient enzymes.

It is also important to clarify that we appreciate the advances made in designing artificial enzymes (and clearly those done with catalytic antibodies), in terms of generating active sites that bind and fit the given reacting system. However, we do not believe that the current steps are sufficient for generating effective catalysts and a CAED must involve the ability to predict the catalysis in the given active site.

At this point it is useful to clarify the difference between our EVB approaches and current alternative approaches. The essential requirement from a proper screening approach is the ability to reproduce the observed catalytic effects. Obviously this major requirement cannot be accomplished by gas phase models (including even gas phase models with the substrate and very few residues) that were used for the initial screening in some cases (e.g., (16, 24)). Instructive MM and related simulations (19, 39) can tell us about the optimal donor / acceptor geometries and to help in generating proper scaffolds for the reacting systems, but are unlikely to be able to predict the catalytic trends in properly oriented systems. More relevant and instructive would be a comparison of the EVB to current MO-QM/MM studies of enzyme design. Here it would be useful to consider several recent studies of the Kemp eliminase and related systems: the semiempirical MO-QM/MM study of Jorgensen and coworkers have provided reliable results for the water reference reactions (40), but the predicted trend in the protein (17) is not encouraging. More specifically, the MO-QM/MM approach performs nicely in exploring the effect of changing the distance between the donor and acceptor (i.e., the Glu to Asp mutation (41)). However, the real challenge is to reproduce the effect of changing the environment (which occurs in directed evolution experiments and is usually responsible for the catalytic activity) and this challenge has not been yet met by the current MO-QM/MM studies of Kemp eliminases (which drastically underestimated the barrier in the enzyme). Interestingly Ref (18) argued that it could improve the MO-QM/MM results. However, the reported results (and the agreement with the corresponding experimental results) seem to overlook the energetics of forming the protonated water molecule that is assumed to be the proton donor. Perhaps the current difficulties with the MO-QM/MM are due to the fact that the reported studies kept the main chain fixed. Alternatively, Houk and coworkers (19) have attempted to use ONIOM truncated protein model but obtained a relatively poor agreement (with a spread of about 12

kcal/mol for experiment deviations of 2 kcal/mol (see Fig S3 in (19)). This work also presented views that might lead to some confusion. First, there are problems with the argument that calculations with an error of 1.5 kcal/mol may not be useful since the observed mutational effects are around 2 kcal/mol. In fact, predictability with 2 kcal/mol error range would be a fantastic tool in attempts to generate enzyme with large catalytic effects. Second, and potentially more problematic, is the idea that the predicted power requires a very high level QM method. This suggestion is risky (in terms of its possible impact on the experimental community) and unjustified. That is, predictive approaches like the EVB are not interested in predicting the absolute QM energy of the substrate, since what counts is the change in this energy upon moving from water to the protein active site. Thus the effort in developing predictive method must be spent on having a good convergence and a proper long rang treatment and not on getting the best basis set.

Overall, we have no doubt that MO-QM/MM approach with proper sampling (e.g., with our approach of using a reference potential (42) will be able to provide a proper screening tool (in particularly when used with an EVB as a reference potential). However, at present the EVB seems to provide the most effective way for obtaining reliable screening results perhaps because its ability to allow for sufficiently extensive sampling.

Some workers may see a great potential for using disolvation effects in enzyme and causing a catalysis by ground state destabilization (GSD) (see discussion in (4)) but native enzymes do not catalyze reaction by exploiting disolvation effects (4). Furthermore, even Kemp eliminases have not been able to exploit this effect significantly. One of the problems is that even if we could create a strong RS disolvation for the base it would lead to a very large $pK_a$ and this will not help at physiological pH. That is, if we try to destabilize the RS by destabilizing the ionized base (e.g., the ionized Asp), the base will be protonated by a bulk proton. Here the best option is to use the polar TS stabilization but unfortunately it is very hard to obtain for the Kemp TS (see (13)). Perhaps a part of the reason why enzymes do not use disolvation effects is the available pH range and the available amino acids (see discussion of ODCase in (43)).

Note that the GSD issue has been established with the reliable linear response (LRA) calculations of the ground state solvation free energy (13) and by detailed comparison to the related case in dehalogenase, where a similar situation is handled by a neutrally evolved enzyme in a completely different way, with ground state stabilization and with very large transition state stabilization (see ref. (13)). Our conclusion about the fact that the Kemp eliminases use RSD is not related to the exact structure of the TS, namely concerted or stepwise, but to the charge distribution of the TS (which has been treated here in a rigorous way by our special procedure).

Since the present work invested a major computational power in validating the EVB results, one may ask why we have not provided some predictions. The answer involves two points. First, we do provide several clear predictions with regards to the effects of mutating distanced residues predictions. This reflects our finding (see (13) that it is extremely hard to get large catalysis in Kemp eliminases by simple mutations of the active sit residues (this is why we enough turn our attention to other systems, where the changes in the substrate charges upon going to the TS are larger). Second, we already demonstrated our ability to have reasonable predictions of the Asn155 to Ala in subtilisin in ref (7). Thus the present paper is more about what does it take to get a reliable prediction than about actual predictions. More specifically, in our previous works (e.g., (3)) we examined the trade off between speeds and reliability in different approaches for enzyme design. At present we feel that our fast strategies (like using group contributions and reorganization energies) are not predictive enough, and that it is important to use the extensive averaging considered here.

However, with the current advances in computer power, this is not such a bad news. That is, as seen from Table 4, we can screen 14 mutants (twenty runs for each mutant) in 24 hours on 200 nodes and 70 mutants can be screened using 1000 nodes. We also used a more parallel approach, where the mapping is distributed on several processers, but this did not lead to a more efficient screening.

At present there are still many who believe in dynamical and other esoteric effects that are presumed to contribute to catalysis (for review see (21)). In many cases it is clearly suggested that improving such effects will be crucial for optimal enzyme design (e.g., (44)). However, it seems to us that by far the main factor that actually contributes to catalysis is the preorganization effect and thus we feel that there is no rational way for improving dynamics and related effects as these factors do not contribute to catalysis (21). Furthermore, we would like to clarify that TS stabilization by delocalization effects (16) is unlikely to provide a significant catalytic factor since the same effect exists in the reference solution reaction. Thus, the possible effect of $\pi$- stacking (which was considered in (16)) is not expected to lead to a significant rate enhancement above the simple effect of having nearby induced dipoles, which is much less effective than having preorganized polar environment. In fact, as realized by Hilvert and coworkers (15) the corresponding dispersion or more precisely inductive effect is small.

Our previous work (13) attempted to refine the electrostatic environment near O1. This effort can be considered by some as an extension of the idea of placing an acid near O1 (e.g., (2, 15)). However, the idea that such a base is needed is reminiscence from physical organic chemistry concepts that capture some of the electrostatic effect, but end up looking at factors that do not play major role in enzyme catalysis. In our view, the issue is not a charge transfer or a covalent bond to the substrate as might be concluded from gas phase calculations, since we are not dealing here with a bifunctional reaction with two steps (unless we have a new chemistry), but with a pure electrostatic stabilization. It is true that the attempts to focus on the base lead in some cases to what we consider the correct direction, like placing Lys or His near O1 (2), but this has little to do with the $pK_a$ of Lys, as one would assume from the traditional picture. It actually reflects the electrostatic free energy of interactions.

Obviously our strategy can be and will be improved in the future, but the main point is the ability to consider enzyme design by using energy based concepts in a rational way. In this respect it is useful to point out a special nontrivial advantage of our CAED approach. That is, in the case of attempts to improve a specific enzyme, there is an acquired advantage in the accumulated experience of modeling different mutants (as was the case here in the study of the evolved mutants). Even significant errors in predicting the first set of mutants can lead to improvement of the model (e.g., in selecting enzyme specific dielectric to compensate for convergence difficulties in the treatment of ionized residues) and to a better understanding of the specific enzymes and thus to a better predictive ability in further design rounds. Such an advantage is not expected from computational design approaches that are not base on capturing the physics of the catalytic process.

Finally, we would like to reclarify that we have demonstrated the ability to reproduce quantitatively the absolute catalytic effects and mutational effects in naturally evolved enzymes (4) and in designer enzymes (this work). This clearly indicates that the catalytic power of enzyme is not due to elusive effects (e.g., conformational dynamics), but to what is by now well understood; namely the electrostatic preorganization. Thus our difficulties in improving designer enzymes are not due to overlooking misunderstood factors, but to the difficulties in optimizing well understood factors. In other words, a method that reproduces the catalytic rate in known systems should be able to do so in any unknown sequence and the challenge is to find the unknown optimal sequence. At any rate, it seems to us that the

present study provided a useful analysis of the reasons for the less than perfect performance of current designer enzymes.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Abbreviations

| | |
|---|---|
| **TS** | transition state |
| **EVB** | empirical valence bond |
| **MO-QM/MM** | molecular orbital-combined quantum mechanical / molecular mechanics |
| **CAED** | computer aided enzyme design |
| **LFER** | linear free energy relationships |
| **RS** | reactant state |
| **PS** | product state |
| **FEP/US** | free energy perturbation umbrella sampling |
| **SCAAS** | surface constrained all atom solvent |
| **LRF** | local reaction field |

## Acknowledgments

## References

1. Toscano MD, Woycechowsky KJ, Hilvert D. Minimalist active-site redesign: Teaching old enzymes new tricks. Angewandte Chemie-International Edition. 2007; 46:4468–4470.

2. Khersonsky O, Rothlisberger D, Dym O, Albeck S, Jackson CJ, Baker D, Tawfik DS. Evolutionary optimization of computationally designed enzymes: Kemp eliminases of the KE07 series. J Mol Biol. 2010; 396:1025–1042. [PubMed: 20036254]

3. Roca M, Vardi-Kilshtain A, Warshel A. Toward accurate screening in computer-aided enzyme design. Biochemistry. 2009; 48:3046–3056. [PubMed: 19161327]

4. Warshel A, Sharma PK, Kato M, Xiang Y, Liu H, Olsson MH. Electrostatic basis for enzyme catalysis. Chem Rev. 2006; 106:3210–3235. [PubMed: 16895325]

5. Warshel, A. Computer Modeling of Chemical Reactions in Enzymes and Solutions. New York: Wiley Interscience; 1991.

6. Warshel A, Sussman F. Toward computer-aided site-directed mutagenesis of enzymes. Proc. Natl. Acad. Sci. USA. 1986; 83:3806. [PubMed: 3520555]

7. Hwang JK, Warshel A. Semiquantitative calculations of catalytic free energies in genetically modified enzymes. Biochemistry. 1987; 26:2669–2673. [PubMed: 3300766]

8. Liu H, Warshel A. The Catalytic Effect od Dihydrofolate Reductase and Its Mutants Is Determined by Reorganization Energies. Biochemistry. 2007; 46:6011–6025. [PubMed: 17469852]

9. Grigorenko BL, Nemukhin AV, Topol IA, Cachau RE, Burt SK. QM/MM Modeling the Ras-GAP Catalyzed Hydrolysis of Guanosine Triphosphate. Proteins. 2005; 60:495–503. [PubMed: 15906320]

10. Marti S, Andres J, Moliner V, Silla E, Tunon I, Bertran J. Computer-Aided Rational Design of Catalytic Antibodies: The 1F7 Case. Angewandte Chemie-International Edition. 2007; 46:286–290.

11. Marti S, Andres J, Moliner V, Silla E, Tunon I, Bertran J. Predicting an Improvement of Secondary Catalytic Activity of Promiscuos Isochorismate Pyruvate Lyase by Computational Design. J. Am. Chem. Soc. 2008; 130:2894–2895. [PubMed: 18281989]

12. Mulholland AJ. Computational enzymology: modelling the mechanisms of biological catalysts. Biochemical Society Transactions. 2008; 36:22–26. [PubMed: 18208378]

13. Frushicheva MP, Cao J, Chu ZT, Warshel A. Exploring challenges in rational enzyme design by simulating the catalysis in artificial kemp eliminase. Proc Natl Acad Sci U S A. 2010; 107:16869–16874. [PubMed: 20829491]

14. Kemp DS, Casey ML. Physical organic chemistry of benzisoxazoles II. Linearity of the bronsted free energy relationship for the base-catalyzed decomposition of benzisoxazoles. J. Am. Chem. Soc. 1973; 95:6670–6680.

15. Debler EW, Ito S, Seebeck FP, Heine A, Hilvert D, Wilson IA. Structural origins of efficient proton abstraction from carbon by a catalytic antibody. Proc Natl Acad Sci U S A. 2005; 102:4984–4989. [PubMed: 15788533]

16. Rothlisberger D, Khersonsky O, Wollacott AM, Jiang J, DeChancie J, Betker J, Gallaher JL, Althoff EA, Zanghellini A, Dym O, Albeck S, Houk KN, Tawfik DS, D B. Kemp Elimination Catalysts by Computational Enzyme Design. Nature. 2008; 453:190–195. [PubMed: 18354394]

17. Alexandrova AN, Rothlisberger D, Baker D, Jorgensen WL. Catalytic mechanism and performance of computationally designed enzymes for Kemp elimination. J Am Chem Soc. 2008; 130:15907–15915. [PubMed: 18975945]

18. Acevedo O. Role of water in the multifaceted catalytic antibody 4B2 for allylic isomerization and Kemp elimination reactions. J Phys Chem B. 2009; 113:15372–15381. [PubMed: 19860435]

19. Kiss G, Rothlisberger D, Baker D, Houk KN. Evaluation and ranking of enzyme designs. Protein Sci. 2010; 19:1760–1773. [PubMed: 20665693]

20. Lassila JK. Conformational diversity and computational enzyme design. Curr Opin Chem Biol. 2010; 14:676–682. [PubMed: 20829099]

21. Kamerlin SC, Warshel A. At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis? Proteins. 2010; 78:1339–1375. [PubMed: 20099310]

22. Kamerlin SC, Williams NH, Warshel A. Dineopentyl phosphate hydrolysis: evidence for stepwise water attack. J Org Chem. 2008; 73:6960–6969. [PubMed: 18729515]

23. Thorn SN, Daniels RG, Auditor MT, Hilvert D. Large rate accelerations in antibody catalysis by strategic use of haptenic charge. Nature. 1995; 373:228–230. [PubMed: 7816136]

24. Hu Y, Houk KN, Kikuchi K, Hotta K, Hilvert D. Nonspecific medium effects versus specific group positioning in the antibody and albumin catalysis of the base-promoted ring-opening reactions of benzisoxazoles. J Am Chem Soc. 2004; 126:8197–8205. [PubMed: 15225061]

25. Debler EW, Muller R, Hilvert D, Wilson IA. An aspartate and a water molecule mediate efficient acid-base catalysis in a tailored antibody pocket. Proc Natl Acad Sci U S A. 2009; 106:18539–18544. [PubMed: 19846764]

26. Privett, HK. CaltechETD:etd-05272009-091024. California Institute of Technology; 2009. An iterative approach to de novo computational enzyme design and the successful application to the Kemp elimination.

27. Zhu L, Yang F, Chen L, Meehan EJ, Huang M. A new drug binding subsite on human serum albumin and drug-drug interaction studied by X-ray crystallography. J Struct Biol. 2008; 162:40–49. [PubMed: 18258455]

28. Hollfelder F, Kirby AJ, Tawfik DS, Kikuchi K, Hilvert D. Characterization of Proton-Transfer Catalysis by Serum Albumins. J Am Chem Soc. 2000; 122:1022–1029.

29. Kamerlin SC, Warshel A. On the energetics of ATP hydrolysis in solution. J Phys Chem B. 2009; 113:15692–15698. [PubMed: 19888735]

30. Kamerlin SC, Warshel A. The EVB as a quantitative tool for formulating simulations and analyzing biological and chemical reactions. Faraday Discussions. 2010; 145:71–106.

31. Kamerlin SC, Warshel A. The empirical valence bond model: theory and applications. Wiley Interdisciplinary Reviews: Computational Molecular Science. 2011; 1:30–45.

32. Kamerlin SC, Cao J, Rosta E, Warshel A. On unjustifiably misrepresenting the EVB approach while simultaneously adopting it. J Phys Chem B. 2009; 113:10905–10915. [PubMed: 19606825]

33. Lee FS, Chu ZT, Warshel A. Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by the Polaris and Enzymix Programs. J. Comput. Chem. 1993; 14:161–185.

34. Singh N, Warshel A. Absolute binding free energy calculations: on the accuracy of computational scoring of protein-ligand interactions. Proteins. 2010; 78:1705–1723. [PubMed: 20186976]

35. Vicatos S, Roca M, Warshel A. Effective approach for calculations of absolute stability of proteins using focused dielectric constants. Proteins. 2009; 77:670–684. [PubMed: 19856460]

36. Hollfelder F, Kirby AJ, Tawfik DS. On the magnitude and specificity of medium effects in enzyme-like catalysts for proton transfer. J Org Chem. 2001; 66:5866–5874. [PubMed: 11511264]

37. Warshel A. What about protein polarity? Nature. 1987; 330:15–16. [PubMed: 3670392]

38. Shurki A, Warshel A. Why does the Ras switch "break" by oncogenic mutations? Proteins. 2004; 55:1–10. [PubMed: 14997535]

39. Baker D. An exciting but challenging road ahead for computational enzyme design. Protein Sci. 2010; 19:1817–1819. [PubMed: 20717908]

40. Acevedo O, Jorgensen WL. Solvent effects and mechanism for a nucleophilic aromatic substitution from QM/MM simulations. Org Lett. 2004; 6:2881–2884. [PubMed: 15330638]

41. Alexandrova AN, Jorgensen WL. Origin of the activity drop with the E50D variant of catalytic antibody 34E4 for Kemp elimination. J Phys Chem B. 2009; 113:497–504. [PubMed: 19132861]

42. Rosta E, Klahn M, Warshel A. Towards accurate ab initio QM/MM calculations of free-energy profiles of enzymatic reactions. J Phys Chem B. 2006; 110:2934–2941. [PubMed: 16471904]

43. Warshel A, Strajbl M, Villa J, Florian J. Remarkable rate enhancement of orotidine 5'-monophosphate decarboxylase is due to transition-state stabilization rather than to ground-state destabilization. Biochemistry. 2000; 39:14728–14738. [PubMed: 11101287]

44. Nagel ZD, Klinman JP. A 21st century revisionist's view at a turning point in enzymology. Nat Chem Biol. 2009; 5:543–550. [PubMed: 19620995]

## 1(a)



## 1(b)

**1**      **2**      **3**

**4**

**Figure 1.**
**(a).** A schematic description of the Kemp elimination reaction.

**(b).** The substrate structures used in the study: 5-nitrobenzisoxazole (**1**), 5,7-dinitrobenzisoxazole (**2**), unsubstituted benzisoxazole (**3**) and 6-glutaramidebenzisoxazole (**4**).
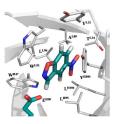
**Figure 2.**
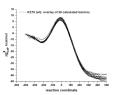The structure of active site of the 34E4 catalytic antibody with 5-nitrobenzioxazole substrate.

**Figure 3.**
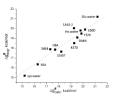Showing the results of 20 EVB mapping runs for the KE70 (wt) designed variant.

**Figure 4.**
Correlation between the calculated and observed activation barriers.
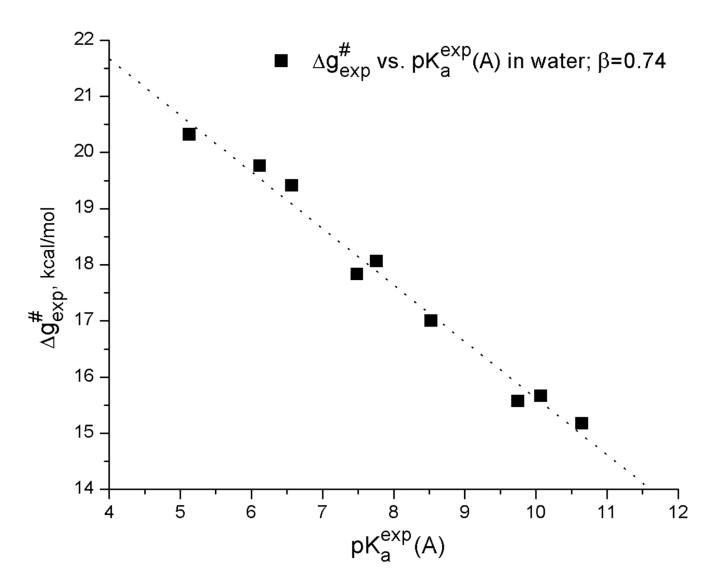
**Figure 5.**
Observed LFER of the 5-nitrobenzisoxazole in solution as a function of the acceptor $pK_a$ (the experimental values of the $pK_a(A)$ are taken from (14)).
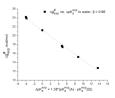
**Figure 6.**
Experimental LFER for the reaction in water as a function of the corresponding $\Delta pK_a$. The experimental $pK_a(D)$ are taken from (24) and $pK_a(A)$ is taken from (14) (more details can be found in Table S2).
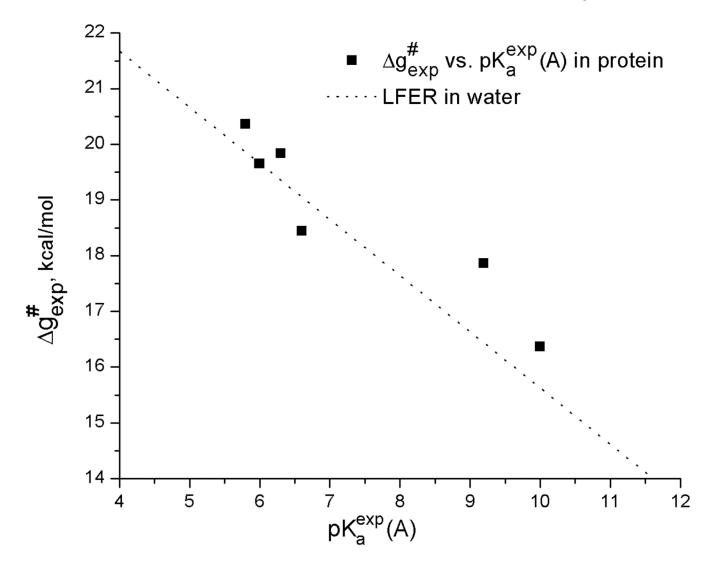
**Figure 7.**
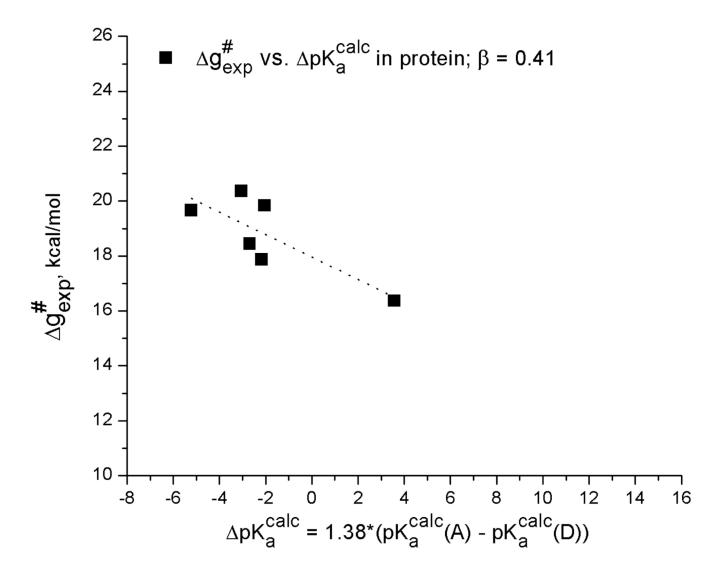The LFER as a function of the acceptor pK$_a$ for the reaction in the protein.

**Figure 8.**
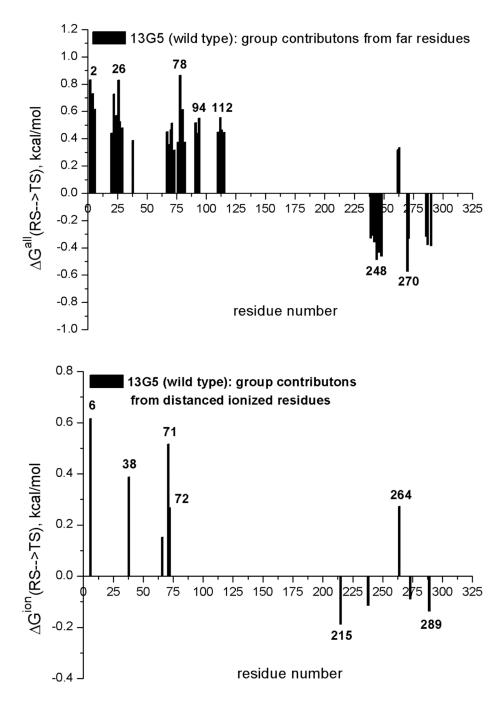The LFER as a function of $\Delta pK_a$ for the reaction in the protein.

**Figure 9.**
Group contributions from the all distanced (top picture) and ionized distanced (bottom picture) residues. The calculations were done for the 13G5 catalytic antibody. The group contributions are for changing the charge at the indicated site from zero to +1.0. The residue numbers from 2 to 215 and from 216 to 431 are located at the heavy and light chains of the 13G5 catalytic antibody respectively.
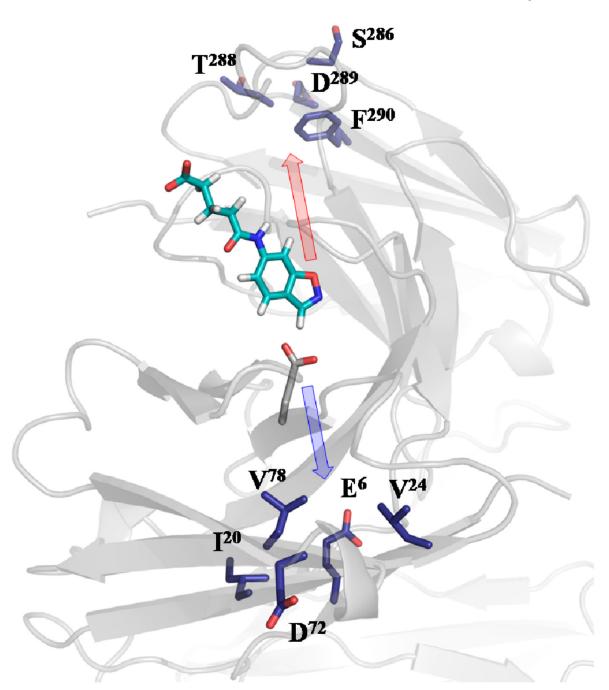
**Figure 10.**
Predicted sites of the most effective mutations for catalysis by distanced residues.

**Table 1**

The correlation between the observed and calculated $\Delta g^{\ddagger}_{cat}$ (in kcal/mol) for the Kemp Elimination reaction with different catalytic bases (acceptors) and substrates (donors).

| System:[1] | Base:[2] | Donor (benzisoxazole):[3] | $\Delta g^{\ddagger}_{obs}$, kcal/mol | $\Delta g^{\ddagger}_{calc}$ kcal/mol | SD |
|---|---|---|---|---|---|
| Water (cage) | Glu/Asp | 5-$NO_2$ | 21.2 | 21.2 | |
| 34E4 antibody (wt, PDB: 1Y0L) | $E^{H50}$ | 5-$NO_2$ | 17.9 | 17.3 | 0.47 |
| 34E4 mutant ($E^{H50}$D, PDB: 1Y18) | $D^{H50}$ | 5-$NO_2$ | 19.9 | 20.2 | 0.66 |
| 34E4 mutant ($Y^{L32}$K, from wt) | $E^{H50}$ | 5-$NO_2$ | 19.5 | 19.9 | 0.96 |
| HG2 mutant (S265T, from HG2 wt) | D126 | 5-$NO_2$ | 17.7 | 18.2 | 0.92 |
| 1A53-2 (X-ray from Prof. Mayo) | E178 | 5-$NO_2$ | 20.0 | 20.7 | 1.01 |
| Water (cage) | Glu/Asp | 5,7-$(NO_2)_2$ | 17.7 | 17.7 | |
| 34E4 antibody (wt, PDB: 1Y0L) | $E^{H50}$ | 5,7-$(NO_2)_2$ | 13.6 | 13.3 | 0.93 |
| Water (cage) | Glu/Asp | unsubstituted | 23.9 | 24.0 | |
| 34E4 antibody (wt, PDB: 1Y0L) | $E^{H50}$ | unsubstituted | 23.5 | 23.7 | 0.68 |
| Water (cage) | Glu | 6-glutaramide | 24.2 | 24.2 | |
| 13G5 antibody (wt, PDB: 3FO0) | $D^{H35}$ | 6-glutaramide | 19.7 | 19.7 | 0.37 |
| 13G5 mutant ($H^{H95}$N, from wt) | $D^{H35}$ | 6-glutaramide | 20.4 | 20.6 | 0.29 |
| 13G5 mutant ($E^{L34}$Q, PDB: 3FO2) | $D^{H35}$ | 6-glutaramide | 19.8 | 19.4 | 0.52 |
| 13G5 mutant ($E^{L34}$A, PDB: 3FO1) | $D^{H35}$ | 6-glutaramide | 18.4 | 18.3 | 0.38 |
| Water (cage) | His | 5-$NO_2$ | 19.8 | 19.8 | |
| KE70 designed variant (wt) | H16-D44 | 5-$NO_2$ | 18.5 | 19.3 | 0.67 |
| KE70 mutant (D44N, from wt) | H16-D44N | 5-$NO_2$ | 19.1 | 19.5 | 1.09 |
| Water (cage) | Lys | 5-$NO_2$ | 15.2 | 15.2 | |
| HSA (wt, PDB: 3B9L) | K199 | 5-$NO_2$ | 17.9 | 17.7 | 0.84 |
| BSA (wt, from HSA sequence) | K222 | 5-$NO_2$ | 16.4 | 16.3 | 1.05 |

[1] The system definition includes the name of the protein, its mutant and the X-ray PDB code.

[2] The base definition includes the name of the base and number in the protein X-ray structure.

[3] All donors are in the family of the benzisoxazoles with different group-substitutions: 5-NO$_2$ states for the 5-nitrobenzisoxazole; 5,7-(NO$_2$)2 is 5,7-dinitrobenzisoxazole, unsubstituted is unsubstituted benzisoxazole and 6-glutaramide is 6-glutaramidebenzisoxazole.

**Table 2**

The calculated reaction barriers for the KE70 designed variant for 20 conformations.

| conformation, ps | $\Delta g^{\ddagger}_{calc}$, kcal/mol | conformation, ps | $\Delta g^{\ddagger}_{calc}$, kcal/mol |
|---|---|---|---|
| 10 | 18.71 | 110 | 18.97 |
| 20 | 18.61 | 120 | 19.45 |
| 30 | 18.9 | 130 | 19.07 |
| 40 | 18.44 | 140 | 19.67 |
| 50 | 19.28 | 150 | 19.82 |
| 60 | 20.15 | 160 | 18.99 |
| 70 | 18.51 | 170 | 20.11 |
| 80 | 20.67 | 180 | 19.08 |
| 90 | 18.19 | 190 | 20.08 |
| 100 | 19.03 | 200 | 19.78 |
| average $\Delta g^{\ddagger}_{calc}$, kcal/mol | | 19.28 | |
| $\Delta g^{\ddagger}_{obs}$, kcal/mol | | 18.51 | |
| SD | | 0.67 | |

**Table 3**

The calculated $\Delta g^{\ddagger}_{cat}$ (in kcal/mol) for the Kemp Elimination reaction of inactive enzymes.

| System:[1] | Base:[2] | Donor (benzisoxazole):[3] | $\Delta g^{\ddagger}_{obs}$, kcal/mol | $\Delta g^{\ddagger}_{calc}$, kcal/mol |
|---|---|---|---|---|
| Water (cage) | Glu/Asp | 5-NO$_2$ | 21.2 | 21.2 |
| KE59 designed variant | E231 | 5-NO$_2$ | inactive | 24.3 |
| HG2 (wt, X-ray from Prof. Mayo) | D126 | 5-NO$_2$ | inactive | 31.7 |
| HG2 mutant (S265A, from wt) | D126 | 5-NO$_2$ | inactive | 26.7 |

[1] The system definition includes the name of the protein, its mutant and the X-ray PDB code.

[2] The base definition includes the name of the base and number in the protein X-ray structure.

[3] All donors are in the family of the benzisoxazoles with different group-substitutions: 5-NO$_2$ states for the 5-nitrobenzisoxazole; 5,7-(NO$_2$)$_2$ is 5,7-dinitrobenzisoxazole, unsubstituted is unsubstituted benzisoxazole and 6-glutaramide is 6-glutaramidebenzisoxazole.

**Table 4**

Estimating the efficiency of the EVB screening calculations[1].

| | computer time per mutant (runs, processor) | No. of mutants per 24h per 200 processors[2] | No. of mutants per 24h per 1000 processors[2] |
|---|---|---|---|
| $\Delta g^{\ddagger}_{cat}$ using EVB | 17 h (1, 1) | 14 | 70 |

[1] The calculations were conducted on the University of Southern California HPCC (High Performance Computing and Communication) Linux computer, using Dual Intel Xeon(64-bit) 3.2 GHz 2GB Memory nodes.

[2] Twenty runs per mutant.