J|A|C|S
ARTICLES

Published on Web 11/01/2008

Catalytic Mechanism and Performance of Computationally Designed Enzymes for Kemp Elimination

Anastassia N. Alexandrova,[†] Daniela Röthlisberger,[‡] David Baker,[‡,§] and William L. Jorgensen*,[†]

Sterling Chemistry Laboratory, Department of Chemistry, Yale University, New Haven, Connecticut 06520 and Department of Biochemistry and Howard Hughes Medical Institute, University of Washington, Seattle, Washington 981951

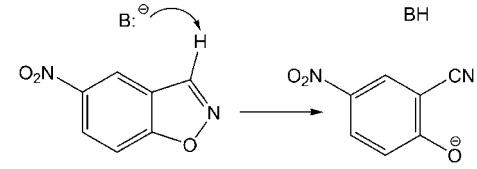
Received May 29, 2008; E-mail: william.jorgensen@yale.edu

Abstract: A series of enzymes for Kemp elimination of 5-nitrobenzisoxazole has been recently designed and tested. In conjunction with the design process, extensive computational analyses were carried out to evaluate the potential performance of four of the designs, as presented here. The enzyme-catalyzed reactions were modeled using mixed quantum and molecular mechanics (QM/MM) calculations in the context of Monte Carlo (MC) statistical mechanics simulations. Free-energy perturbation (FEP) calculations were used to characterize the free-energy surfaces for the catalyzed reactions as well as for reference processes in water. The simulations yielded detailed information about the catalytic mechanisms, activation barriers, and structural evolution of the active sites over the course of the reactions. The catalytic mechanism for the designed enzymes KE07, KE10(V131N), and KE15 was found to be concerted with proton transfer, generally more advanced in the transition state than breaking of the isoxazolyl N−O bond. On the basis of the free-energy results, all three enzymes were anticipated to be active. Ideas for further improvement of the enzyme designs also emerged. On the technical side, the synergy of parallel QM/MM and experimental efforts in the design of artificial enzymes is well illustrated.

Introduction

Theoretical and experimental efforts are being directed toward the design of artificial enzymes.[1−11] The challenge is great in trying to approach the performance of biological catalysts in view of their specificity and high activity under mild conditions.[12] It is, therefore, inviting to mimic strategies that Nature employs and, especially, apply them in the design of catalysts for important organic reactions that Nature seems to have overlooked. In the course of this pursuit, an efficient enzyme design protocol has been recently reported,[11] and it was employed to design artificial enzymes for the Kemp elimination[13] of 5-nitrobenzisoxazole (Scheme 1).[1] Designed enzymes were synthesized and tested experimentally and yielded k_{cat}/k_{uncat} values in the 10^2-10^5 range; subsequent in vitro evolution of one of the designs provided a greater than 200-fold increase in k_{cat}/K_M and a maximal k_{cat}/k_{uncat} of 1.2×10^6. The simplicity of the Kemp elimination and its sensitivity to medium effects have made it the target for numerous prior efforts at catalyst discovery, as summarized elsewhere.[14a] Thus far, catalytic antibodies have been the only alternatives to achieve rate accelerations in the million-fold range.[14b]

Scheme 1. Kemp Elimination of 5-Nitrobenzisoxazole



Desirable features in the design of catalysts for the Kemp elimination have been elucidated previously with the aid of ab

[†] Yale University.
[‡] Department of Biochemistry, University of Washington.
[§] Howard Hughes Medical Institute, University of Washington.

(1) Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. *Nature* **2008**, *453*, 190–195.
(2) Bolon, D. N.; Mayon, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14274–14279.
(3) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F., III.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387–1391.
(4) Vasileiou, C.; Vaezeslami, S.; Crist, R. M.; Rabado-Smith, M.; Geiger, J. H.; Borhan, B. *J. Am. Chem. Soc.* **2007**, *129*, 6140–6148.
(5) Pinto, A. L.; Hellinga, H. W.; Caradonna, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5562–5567.
(6) Dwyer, M. A.; Looger, L. L.; Hellinga, H. W. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11255–11260.
(7) Kaplan, J.; DeGrado, W. F. *Proc. Natl. Acad. Sci.* **2004**, *101*, 11566–11570.
(8) Hilvert, D. *Annu. Rev. Biochem.* **2000**, *69*, 751–793.
(9) Lerner, R. A.; Benkovic, S. J.; Schultz, P. G. *Science* **1991**, *252*, 659–667.
(10) Dahiyat, B. I.; Mayo, S. L. *Protein Sci.* **1996**, *5*, 895–903.
(11) Zanghellini, A.; Jiang, L.; Wollacott, A. M.; Cheng, G.; Meiler, J.; Althoff, E. A.; Röthlisberger, D.; Baker, D. *Protein Sci.* **2006**, *15*, 2785–2794.
(12) (a) Radzicka, A.; Wolfenden, R. A. *Science* **1995**, *267*, 90–93. (b) Frey, P. A.; Hegeman, A. D. *Enzymatic Reaction Mechanisms*; Oxford University Press: New York, 2007.
(13) (a) Casey, M. L.; Kemp, D. S.; Paul, K. G.; Cox, D. D. *J. Org. Chem.* **1973**, *38*, 2294–2301. (b) Kemp, D. S.; Casey, M. L. *J. Am. Chem. Soc.* **1973**, *95*, 6670–6680.

initio quantum mechanical calculations.[15] The primary requirement is a catalytic base for proton abstraction. For this role, Röthlisberger et al. chose Asp, Glu, or an unprotonated His;[1] the His can also be coupled with an Asp or a Glu in a catalytic dyad, which allows tuning of the pK_a and, thus, the protonation state of the His. An additional feature explored in some enzyme designs was inclusion of residues with side-chain groups capable of hydrogen bonding to the isoxazolyl oxygen atom of the substrate, specifically, Ser, Lys, and Tyr. The intention was to stabilize the evolving phenolate ion or even protonate it in a relatively hydrophobic active site. Finally, residues were added to form π-stacking interactions with the substrate. Their intended tasks were to orient properly the substrate and catalytic base and stabilize the evolving negative charge through π delocalization or polarization. The remainder of the enzyme needs to provide an orientational match for projection of the side chains in the designed active sites; TIM barrels emerged as the preferred scaffolds.[1]

Thus far, direct information on the catalytic mechanism for these enzymes is lacking. Base-catalyzed Kemp eliminations in aqueous solution are well established as E2 reactions with concerted proton transfer and N−O bond cleavage.[13] The product 2-cyano-4-nitrophenolate ion is stable in water with a pK_a of 4.1 for its conjugate acid.[13a] Fundamental mechanistic questions are whether this concerted pathway is also followed in the designed enzymes and the details of the involvement of active-site residues. A mechanistic spectrum can be imagined depending on the order of proton transfer, N−O bond cleavage, and possible protonation of the phenolate or isoxazole. Formation of an intermediate from the initial proton transfer is certainly a possibility.[13] Indeed, variation of the catalytic residues could lead to multiple mechanisms ranging from base catalyzed to acid catalyzed.[15] It is also not known how the positions and conformations of the residues in the active sites evolve during the reactions and whether they actually perform the desired catalytic functions. The locations and possible participation of water molecules are additional points to address.

The present quantum and molecular mechanics (QM/MM) investigations are directed at elucidating these issues, providing deeper understanding of the enzyme structure−function relationships and giving guidance for further rational enzyme design and optimization of catalytic performance. Specifically, Kemp eliminations of 5-nitrobenzisoxazole catalyzed by four designed enzymes in explicit water were investigated using QM/MM Monte Carlo (MC) and free-energy perturbation (FEP) methodologies. The mechanistic details and energetics of the enzymatic reactions are elucidated, and comparison is made to results for the reference reaction in water.

Methods

Initial coordinates for all proteins were obtained from the model building of Röthlisberger et al. using the *Rosetta* program.[1] In preparation for the MC simulations, the complexes with 5-nitrobenzisoxazole were truncated such that only protein residues with any atom within ca. 15 Å of the substrate were retained. For the four cases, 188−253 residues were retained out of a total of 247 or 253. The overall charge was set to zero via neutralization of residues remote from the active site. The substrate was initially positioned with the hydrogen on C3 oriented toward the putative base, as in the original design.[1] A short, 50-step conjugate gradient relaxation was performed for each system. Subsequently, a 20 Å cap containing TIP4P water molecules[16] was centered on the substrate, and ca. 500 water molecules remained after removal of those with their oxygen atom within 2.5 Å of any non-hydrogen atom of the complex. Figure 1 illustrates a system prepared in this way.

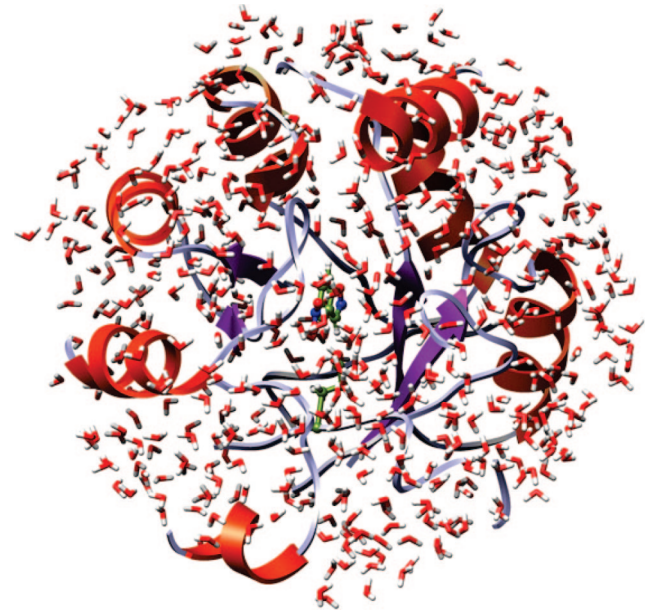


Figure 1. Truncated protein−substrate complex with ca. 500 TIP4P water molecules; the QM region (dark green) consists of the side chain of the catalytic base and the substrate. The nearest ca. 200 residues of the protein and solvent form the MM region.

Initial simulations for the systems were carried out using QM/MM/MC statistical mechanics calculations, as implemented in the *MCPRO* program.[17] Technically similar studies of reactions catalyzed by macrophomate synthase and fatty acid amide hydrolase have been previously reported.[18,19a] The simulations consisted of 5×10^6 configurations of solvent equilibration, 10×10^6 configurations of full equilibration, and 35×10^6 configurations of averaging at 25 °C. The protein backbone was held fixed, while the side chains, substrate, and water were sampled. All degrees of freedom for the substrate were variable, bond angles and dihedral angles were sampled for the side chains, and the TIP4P water molecules could translate and rotate. The final configurations were then used as starting points for the mechanistic investigations. Some MC simulations were carried out with the protein backbone sampled using a concerted rotation algorithm;[20] however, this resulted in rmsd's of only ca. 0.15 Å from the starting backbone structure.

Potential of mean force (PMF) calculations were carried out for the Kemp eliminations by perturbing along reaction coordinates using MC/FEP calculations. Two distances were chosen as reaction coordinates in all cases, and two-dimensional free-energy maps were created. The increment between simulations was 0.04 Å for preliminary scanning and 0.02 Å for refinement near stationary points. Double-wide sampling was used so two free-energy changes were obtained for each window by incrementing in both the forward

(14) (a) Hu, Y.; Houk, K. N.; Kikuchi, K.; Hotta, K.; Hilvert, D. *J. Am. Chem. Soc.* **2004**, *126*, 8197–8205. (b) Thorn, S. N.; Daniels, R. G.; Auditor, M.-T. M.; Hilvert, D. *Nature* **1995**, *373*, 228–230.
(15) Na, J.; Houk, K. N.; Hilvert, D. *J. Am. Chem. Soc.* **1996**, *118*, 6462–6471.
(16) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(17) Jorgensen, W. L.; Tirado-Rives, J. *J. Comput. Chem.* **2005**, *26*, 1689–1700.
(18) Guimarães, C. R. W.; Udier-Blagovic, M.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2005**, *127*, 3577–3588.
(19) (a) Tubert-Brohman, I.; Acevedo, O.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2006**, *128*, 16904–16913. (b) Alexandrova, A. N.; Jorgensen, W. L. *J. Phys. Chem. B* **2007**, *111*, 720–730.
(20) Ulmschneider, J. P.; Jorgensen, W. L. *J. Chem. Phys.* **2003**, *118*, 4261–4271.

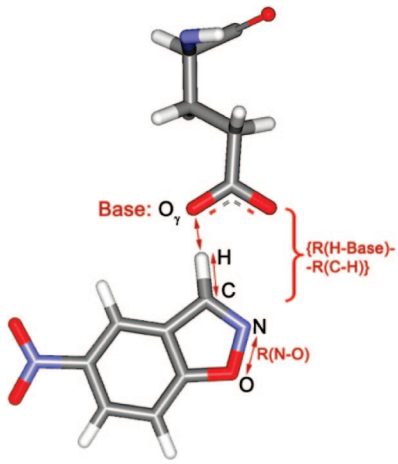


Figure 2. Choices for the reaction coordinates: R(N−O) and {R(H-Base) − R(C−H)}. An oxygen atom of the ionized Glu plays the role of the base in this case.

and the reverse directions. One reaction coordinate was the length of the breaking C−N bond. The reaction coordinate for the hydrogen transfer was chosen as the difference between the lengths of the forming and breaking bonds to the hydrogen, {R(H-Base) − R(C−H)}, where the sum {R(H-Base) + R(C−H)} was fixed at the value obtained in the initial QM/MM MC simulation (Figure 2). The combined reaction coordinate for hydrogen transfer reduces the dimensionality for creation of the free-energy surfaces from three to two. The approach has been previously validated; it was shown that the error for the barrier to hydrogen transfer thus obtained is insignificant.[19] In addition, for the hydrogen transfers the quadrature method of Tubert-Brohman et al.[19a] was also followed whereby the free energy changes are fit by cubic polynomials using free-energy derivatives obtained from only five windows of double-wide sampling (i.e., 10 free energy changes in all). Analytical integration then yields the free-energy profile. This approach was shown to introduce errors in the activation barrier of no greater than 0.5 kcal/mol versus conventional results using 25−30 windows.[19a] In the present case, this treatment for hydrogen transfers reduced the computational effort from roughly 36 000 to 7200 h for construction of each two-dimensional free-energy surface on 3 GHz Pentium processors.

In the QM/MM implementation, the energy of the QM region was computed using the PDDG/PM3 semiempirical molecular orbital method.[21] PDDG/PM3 has been extensively tested for gas-phase structures and energetics,[21] and it has performed well in QM/MM/MC studies for numerous pericyclic,[22] substitution,[23] and elimination reactions in solution.[19b,24,25] Notably, the previous study of Kemp eliminations of benzisoxazole-3-carboxylates yielded free energies of activation within 2−3 kcal/mol of experimental data in water and methanol and revealed the importance of ion pairing in aprotic media.[25] In the present study the QM region consisted of the substrate, 5-nitrobenzisoxazole, and the side chain of the putative base, Glu or Asp (Figure 1). The linking hydrogen atom for the side chain was made coincident with C_α of the basic residue.[18,19a] The energy of the QM region was computed for every attempted move of the solute, i.e., every 10 configurations. The rest of the system was treated using the OPSL-AA force field.[26] To compute the electrostatic part of the interaction energy between the QM and MM regions, CM3P atomic charges were used for the QM atoms and MM charges for the MM atoms.[27] The CM3P charges were unscaled as the QM region had a net charge of −1 for all cases examined here.[28] Interactions were truncated using a 10 Å residue-based cutoff based on distances between non-hydrogen atoms. Each FEP window entailed 5×10^6 configurations of solvent equilibration, 10×10^6 configurations of full equilibration, and 25×10^6 configurations of averaging. At each point on the free-energy surface configurations from the Monte Carlo runs were saved. They are used to analyze and illustrate the structural variations along the reaction paths.

Results

The naming of the enzymes is adopted from Röthlisberger et al.;[1] KE07, KE10(V131N), KE15, and KE16 are the artificial enzymes considered here. The experimental kinetic parameters and choices for the catalytically important residues are listed in Table 1. KE10 in ref 1 is the Asn131Val variant of the enzyme studied here; it is more active than KE10(V131N) with a k_{cat}/K_M of 51.6 ± 4.0 s^{-1} M^{-1}.[1] In all cases, Glu or Asp is the catalytic base and Trp or Tyr is used for π interactions; three of the designs include Lys or Asn to hydrogen bond with the isoxazolyl oxygen of the substrate and possibly stabilize the evolving phenolate ion. The present computations were completed before the experimental kinetic parameters were available.

KE07. KE07 is one of the earliest designs that was found to be active. Though the present calculations were initiated from the structure generated with *Rosetta*, a crystal structure, 2rkx in the RCSB Protein Data Bank, was subsequently obtained for it and found to be very close to the computational design.[1] KE07 is based on a TIM-barrel protein fold; the natural enzyme is an imidazoleglycerolphosphate synthase from *Thermotoga maritima*, and a crystal structure at 1.45 Å resolution is available in the Protein Data Bank (code 1thf).[29] Glu was incorporated in the binding cavity of the 1thf structure as residue 101 to serve as the catalytic base. Lys222 was provided as the potential hydrogen-bond donor for the isoxazolyl oxygen, and Trp50 was included for π stacking with the substrate.[1]

The complex of the protein with the substrate was relaxed in the QM/MM/MC simulations in TIP4P water. Further, using R(N−O) and {R(H−$O_{\gamma Glu}$) − R(C−H)} as the two reaction coordinates, a free-energy surface for the catalyzed reaction was obtained from FEP calculations. The results are summarized in Figure 3. The two-dimensional map is shown in Figure 3A. Representative configurations corresponding to the stationary points on the surface are shown in Figure 3B−E. From the free-energy map the reaction is found to follow a single-step, concerted mechanism. However, the reaction pathway is asym-

(21) (a) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 1601–1622. (b) Tubert-Brohman, I.; Guimarães, C. R. W.; Repasky, M. P.; Jorgensen, W. L. *J. Comput. Chem.* **2003**, *25*, 138–150. (c) Tubert-Brohman, I.; Guimarães, C. R. W.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2005**, *1*, 817–823.
(22) (a) Repasky, M. P.; Guimaraes, C. R. W.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2003**, *125*, 6663–6672. (b) Guimaraes, C. R. W.; Repasky, M. P.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2003**, *125*, 6892–6899. (c) Acevedo, O.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2006**, *128*, 6141–6146. (d) Acevedo, O.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2007**, *3*, 1412–1419.
(23) (a) Vayner, G.; Houk, K. N.; Jorgensen, W. L.; Brauman, J. I. *J. Am. Chem. Soc.* **2004**, *126*, 9054–9058. (b) Acevedo, O.; Jorgensen, W. L. *Org. Lett.* **2004**, *6*, 2881–2884.
(24) Acevedo, O.; Jorgensen, W. L. *J. Org. Chem.* **2006**, *71*, 4896–4902.
(25) Acevedo, O.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2005**, *127*, 8829–8834.
(26) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
(27) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **2003**, *24*, 1291–1304.
(28) Udier-Blagovic, M.; Morales de Tirado, P.; Pearlman, S. A.; Jorgensen, W. L. *J. Comput. Chem.* **2004**, *25*, 1322–1332.
(29) Lang, D. A.; Obmolova, G.; Thoma, R.; Kirschner, K.; Sterner, R.; Wilmanns, M. *Science* **2000**, *298*, 1546–1550.

Table 1. Observed Kinetic Parameters and Computed Geometrical Parameters (Å) and Free Energy of Activation (kcal/mol) for the Four Designed Enzymes

design	PDB code	key residues[a]	k_{cat},[b] s^{-1}	k_{cat}/K_M,[b] s^{-1} M^{-1}	{*R*(H-base) + *R*(C−H)}	*R*(N−O) in TS	$\Delta G^{\ddagger}$
KE07	1thf	E101, K222, W50	0.018 ± 0.001	12.2 ± 0.1	2.86	1.74	8.1
KE10 (V131N)	1a53	E178, N131, W210	ND[c]	3.68 ± 0.32	2.78	1.71	13.5
KE16	1thf	D48, K201, Y126	0.006 ± 0.001	13.4 ± 1.4	2.80	ND[c]	ND[c]
KE15	1thf	D48, None, Y126	0.022 ± 0.003	35.1 ± 4.8	2.94	1.78	12.3

[a] The key residues are specified in the following order: catalytic base, hydrogen-bond donor, π-stacking residue. [b] From ref 1. [c] Not determined; see text.

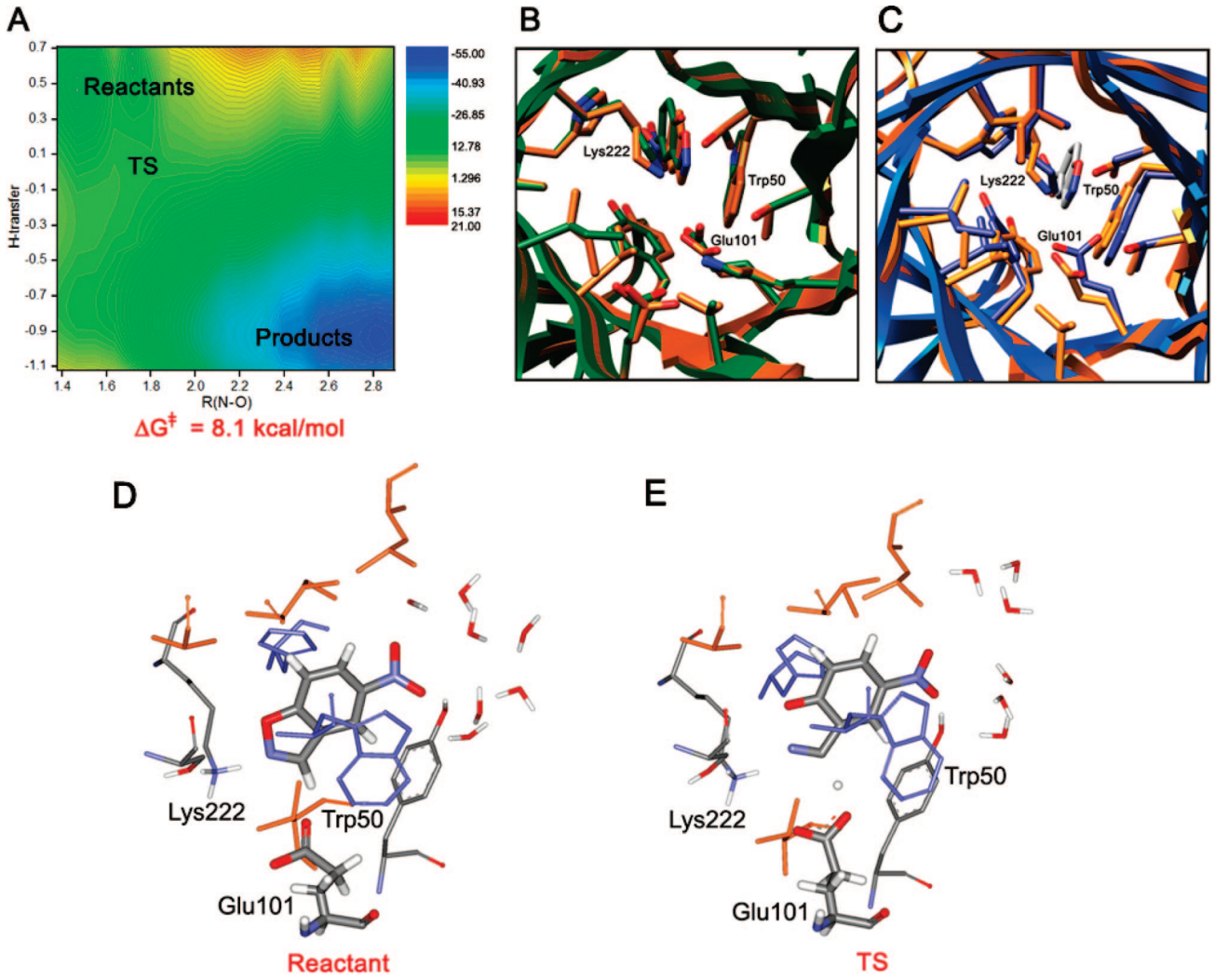


Figure 3. KE07: (A) Computed two-dimensional free-energy map for the catalyzed Kemp elimination. (B) Overlay of the *Rosetta* designed structure (green) and the QM/MM/MC relaxed structure (orange) for the enzyme−substrate complex. (C) Overlay of the X-ray structure (blue) and QM/MM/MC relaxed structure (orange). (D) Configuration of the binding site with the substrate bound at the end of the QM/MM/MC run for the reactant state. (E) Configuration of the transition state from the QM/MM/MC FEP simulations.

metric; the hydrogen transfer from the substrate to Glu101 is significantly advanced before the N−O bond opening is triggered. At the transition state, **TS**, the N−O bond has lengthened to 1.74 Å (Table 1); the corresponding computed value for the Kemp decarboxylation of benzisoxazole-3-carboxylate in water was 1.75 Å.[25] The prior gas-phase quantum mechanical studies of the elimination reaction of 5-nitrobenzisoxazole with acetate ion also found a relatively early transition state with *R*(N−O) = 1.81 Å at the B3LYP/6-31+G(d) level.[15]

Figure 3B−D shows the structure of the initial complex of KE07 with 5-nitrobenzisoxazole, as obtained from the QM/MM/MC relaxation of the structure from *Rosetta*. In Figure 3B and 3C it is overlaid with the starting structure from *Rosetta* and with the crystal structure of KE07,[1] which lacks the substrate. The positioning and conformations of the ligand and side chains are very similar in both cases. The agreement between the designed and crystallographic results for the backbone is striking; it validates the enzyme design protocol, which is based on point mutations in existing protein folds.[1] Overall, the structural consistency in Figure 3B and 3C supports the utility of both the QM/MM/MC and *Rosetta* methodologies.

Figure 3D further illustrates the active site in the reactant state at the end of the QM/MM/MC equilibration. In this initial complex the substrate is in close contact with the catalytic base, Glu101, with $R(H-O_{Glu101})$ = 1.65 Å and {*R*(H−O) +

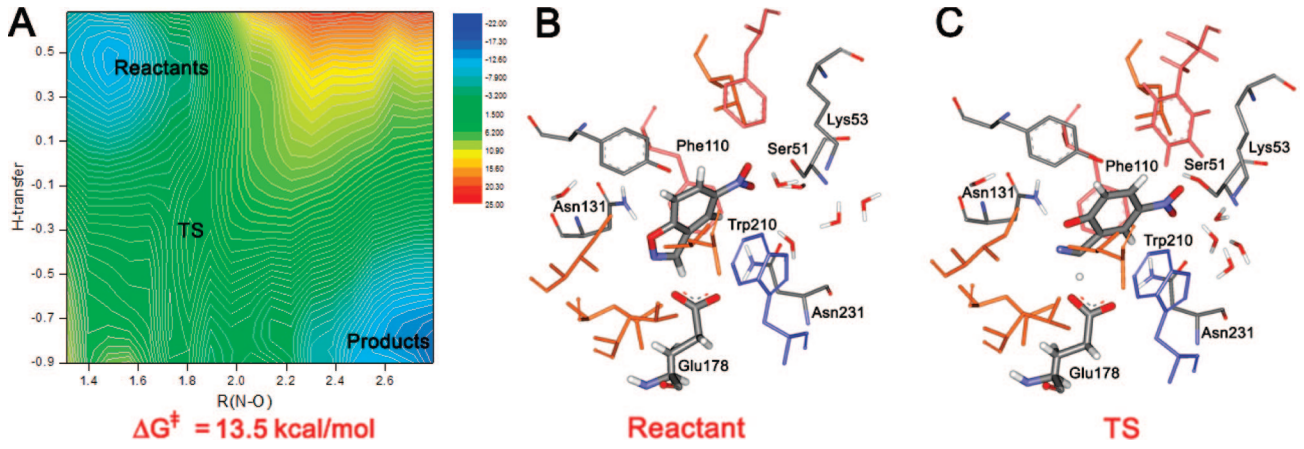


Figure 4. KE10(V131N): (A) Computed two-dimensional free-energy map for the catalyzed Kemp elimination. (B) Configuration of the binding site with the substrate bound at the end of the QM/MM/MC run for the reactant state. (C) Configuration of the transition state from the QM/MM/MC FEP simulations.

$R\{C-H)\} = 2.86$ Å. However, Lys222 forms a hydrogen bond with the isoxazolyl nitrogen rather than the oxygen, $R(H_{Lys}-N) = 2.26$ Å. The orientation of Lys222 is supported by the nearby Ser48, as Lys222 also participates in a hydrogen bond with the oxygen atom of its hydroxyl group. The π-stacking interaction with Trp50 is present, though Trp50 is somewhat tilted away from the substrate and interacts primarily with the substrate's six-membered ring. No water molecules are found in the active site in the MC simulations. However, the nitro group of the substrate points outward toward the solvent and forms several hydrogen bonds with water molecules. Overall, the interactions found in the binding site at the reactant geometry from the MC simulations are consistent with the desired design[1] with the exception of the detail for Lys222.

Similarly, Figure 3E illustrates a typical configuration from the FEP calculations near the transition state. The transferring H is approximately halfway between the donor and the acceptor, and the N−O bond is stretched to 1.74 Å. Trp50 is π stacking well with the six-membered ring portion of the substrate. The hydrogen-bond length between an ammonium hydrogen of Lys222 and the isoxazolyl N, 2.25 Å, is still shorter than the distance to the isoxazolyl oxygen, 3.46 Å. This configuration is promoted by the persistent hydrogen bond between Lys222 and the hydroxyl group of Ser48. The contact between Lys222 and the isoxazolyl N may contribute to the observed asymmetry of the reaction mechanism in which the hydrogen transfer progresses significantly before the N−O bond lengthens. N−O bond cleavage would presumably be better promoted by stabilization of the forming charge on the oxygen. Water remains excluded from the active site in spite of the charge redistribution. In comparing Figure 3D and 3E it is also noted that catalysis might be improved in a design in which the carboxylate oxygen in the reactant state was better oriented to receive the proton on a syn lone pair to achieve the stable *Z*-carboxylic acid conformation.[15] Furthermore, it would be preferable to have the carboxylate group rotated such as to minimize the electrostatic repulsion between the oxygen that does not receive the proton and the isoxazolyl nitrogen (cf. Figure 3 in ref 15).

For the structure of the product (not illustrated) the positioning of the side chains of KE07 is very similar to that for the transition state according to the FEP calculations. The N−O bond opening and hydrogen transfer are complete, and finally the hydrogen bond with the ammonium group of Lys222 is shifted from the isoxazolyl N to O in response to the charge concentration on the oxygen.

The difference in free energy between the transition state and reactant geometries corresponds to the free energy of activation, $\Delta G^{\ddagger}$, which is directly related to the rate of the catalyzed reaction, k_{cat}. From the two-dimensional map in Figure 3A the computed $\Delta G^{\ddagger}$ for KE07 is 8.1 kcal/mol; the estimated uncertainty is less than ±1 kcal/mol based on the fluctuations in the averages for the individual FEP windows. As presented below, the corresponding, computed $\Delta G^{\ddagger}$ for the reference reaction in water with hydroxide ion as the base is 19.8 kcal/mol. Thus, the enzymatic environment is predicted to provide strong catalysis. This assumes that formation of the enzyme−substrate complex is not highly unfavorable, which seems unlikely based on the computed structures, as in Figure 3B−D. The catalytic efficiency, k_{cat}/K_M, has not been computed owing to the technical challenges for K_M, which requires computation of the absolute free energy of binding for the substrate. However, the computed $\Delta G^{\ddagger}$ and the one that can be derived from the observed k_{cat} can be compared. Assuming first-order kinetics for conversion of the enzyme−substrate complex to the product, the observed k_{cat} of 0.018 s^{-1} (Table 1) corresponds to a $\Delta G^{\ddagger}$ of 17.1 kcal/mol. Thus, the computations overestimated the catalytic prowess of the enzyme, though they clearly made the correct qualitative prediction that KE07 would be a good catalyst for the Kemp elimination. The observed k_{cat}/k_{uncat} turned out to be 1.6×10^4.[1] There may be multiple contributors to the discrepancy between the computed $\Delta G^{\ddagger}$ and the one estimated from the observed k_{cat}; candidates include the lack of backbone sampling, which may cause the initial state to be too constrained and standard issues with the QM/MM calculations such as the accuracy of the semiempirical QM and the lack of instantaneous polarization in the MM. Though the present approach has worked well for several solution-phase elimination reactions,[19b,24,25] the enzymatic environment provides greater challenges for configurational sampling and proper representation and treatment of the electrostatic interactions.

KE10(V131N). Results are collected for the KE10 variant in Table 1 and Figure 4. The modeling again started from the structure generated with *Rosetta*. No crystal structures are available for comparison in this and the remaining cases. The design is based on the 1a53 crystal structure for indole-3-glycerolphosphate synthase from *Sulfolobus solfataricus*, which

also belongs to the TIM-barrel family.[30] In this design, Glu178 is the intended base. It is paired with Asn231 through a hydrogen bond, which promotes proper positioning with respect to the binding cavity and the substrate, though it is expected to weaken the basicity of Glu178. From the initial QM/MM/MC relaxation the hydrogen-bond contact between the substrate and Glu178 is tight, $\{R(\mathrm{H-O}) + R(\mathrm{C-H})\} = 2.78$ Å (Figure 4B), and the orientation of the carboxylate group relative to the substrate is ideal.[15] Additionally, a long hydrogen bond is found between the isoxazolyl oxygen of the substrate and the amino group of Asn131 with $R(\mathrm{O-HN}) = 2.80$ Å; we consider strong hydrogen bonds to have $R(\mathrm{O-H})$ or $R(\mathrm{N-H})$ below 2.5 Å. Thus, Asn131 provides some hydrogen-bond-donating potential for the evolving phenolate oxygen. Trp210 was included to π stack with the substrate, and it is positioned in front in Figure 4B. In addition, Phe110 is seen to sandwich the substrate from the back. Both residues are more proximal to the six-membered ring of the substrate than the isoxazole ring. Thus, Phe110 and Trp210 form a slot for the substrate and the catalytic base is well positioned at the opening at one end of the slot. This critical region of the binding site is densely packed, and no water molecules enter it during the MC runs. At the opposite end of the substrate the nitro group is less solvent exposed than for KE07; two hydrogen bonds are found between a nitro oxygen and the side-chain functional groups of Ser51 and Lys53 as well as a hydrogen bond with a water molecule.

The free-energy map from the MC/FEP simulations for the Kemp elimination catalyzed by KE10(V131N) is shown in Figure 4A. Since the catalytic strategy is similar for KE10 and KE07, it is not surprising that the mechanistic details and free-energy maps are also similar. The reaction proceeds via a concerted mechanism, which is more synchronous than for KE07. As illustrated in Figure 4C, all residues play the desired roles. The Glu178-Asn231 catalytic pair functions properly: the abstracted hydrogen smoothly transfers from the substrate to Glu178 and the hydrogen bond between Glu178 and Asn231 lengthens simultaneously; $R(\mathrm{O-HN})$ averages 1.75, 1.85, and 2.10 Å for the reactant, **TS**, and product, respectively. Furthermore, the hydrogen bond between the isoxazolyl oxygen and the NH_2 in the side chain of Asn131 shortens as the negative charge increases on the oxygen in progressing to the transition state; $R(\mathrm{O-HN})$ is 2.05 Å in the reactants and 1.92 Å in the **TS**. Notably, this contact becomes less tight in the product, $R(\mathrm{O-HN}) = 2.93$ Å, owing to geometrical constraints in the binding site that emerge upon fully opening the N−O bond. Thus, the stabilization provided by Asn131 appears to be greatest for the transition state. This effect, the lack of a proximal ion-pairing partner for the product state, and stabilization of Glu178 by Asn231 in the reactant state likely contribute to the fact that the computed overall ΔG for the catalytic step with KE10(V131N) is significantly less exergonic than for KE07 (−4.5 vs −36.3 kcal/mol). In view of this, it would not be surprising if product inhibition is reduced with KE10(V131N). Furthermore, based on the proton-transfer coordinates, the transition state occurs significantly earlier for KE07 than the KE10 variant (Figures 3A and 4A), which may again reflect the greater exergonicity with KE07 and lessened basicity of Glu178.

The computed $\Delta G^{\ddagger}$ for the reaction with KE10(V131N) from the FEP calculations is 13.5 kcal/mol, which is 5.4 kcal/mol higher than for KE07, but still comfortably lower than for the reference hydroxide-catalyzed reaction. The experimental value of k_{cat} for KE10(V131N) was not determined; however, k_{cat}/K_M was found to be a factor of 3.3 lower than for KE07 (Table 1). The present QM/MM MC simulations predict the k_{cat} ratio to be even smaller, so it is possible that K_M is smaller with KE10(V131N), i.e., it has better substrate binding. This seems plausible as the substrate in KE10 is less solvent exposed and more engulfed by the protein including the Phe110-substrate-Trp210 stacking. In any event, the MC/FEP results were correct in predicting that KE10(V131N) would be a successful catalyst for the Kemp elimination, though likely less active than KE07.

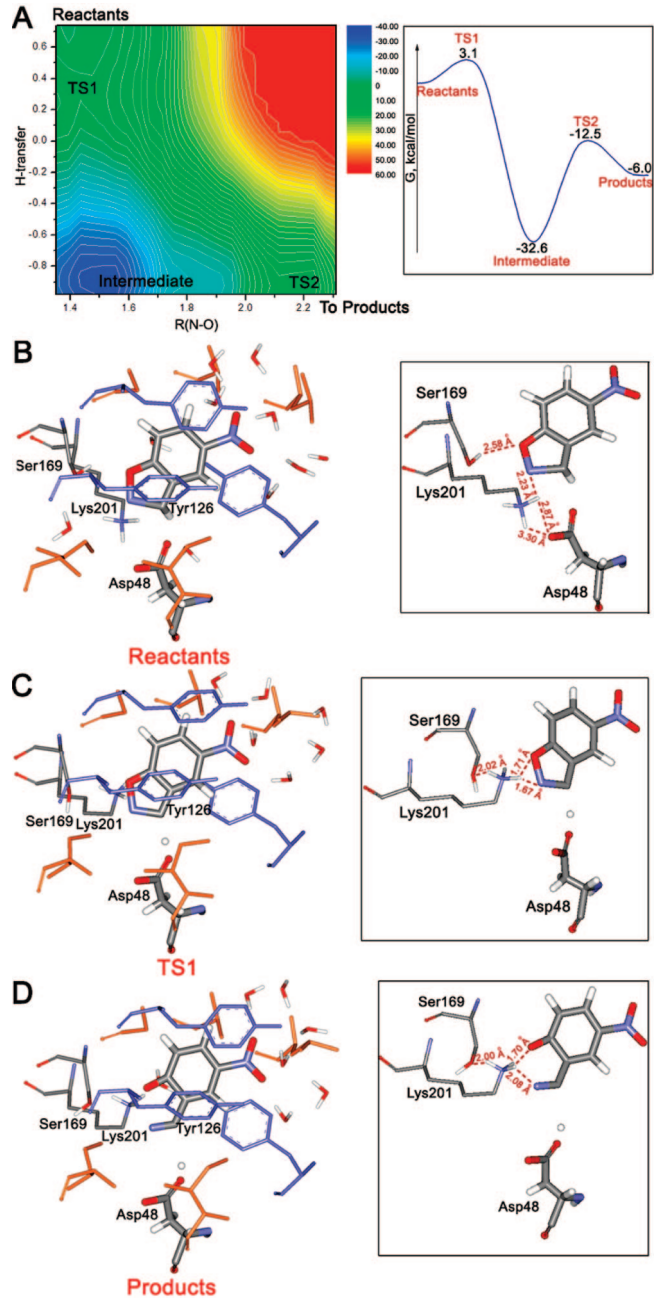


Figure 5. KE16: (A) computed free-energy data for the catalyzed Kemp elimination, and illustrative QM/MM/MC structures for the (B) reactants, (C) transition state TS1, and (D) products.

KE16. Figure 5 summarizes the results of QM/MM MC simulations for KE16, another artificial enzyme based on the 1thf TIM-barrel protein fold. In comparison to KE07, which used Glu101 as the catalytic base, the base for KE16, Asp48,

(30) Hennig, M.; Darimont, B. D.; Jansonius, J. N.; Kirschner, K. *J. Mol. Biol.* **2002**, *319*, 757–766.

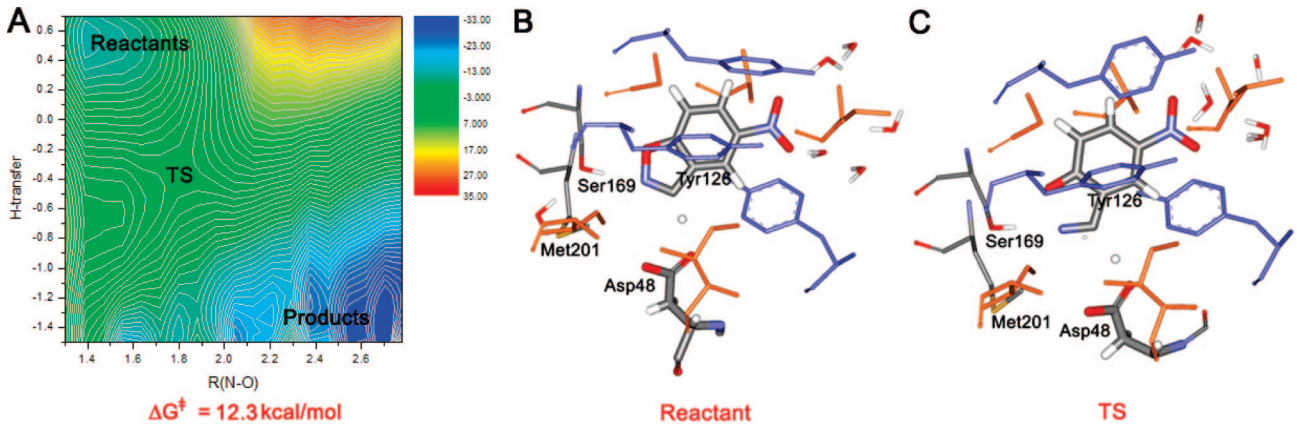


Figure 6. KE15: (A) Computed two-dimensional free-energy map for the catalyzed Kemp elimination. (B) Configuration of the binding site with the substrate bound at the end of the QM/MM/MC run for the reactant state. (C) Configuration of the transition state from the QM/MM/MC FEP simulations.

was installed earlier in the sequence. Lys201 is now introduced to form a hydrogen bond to the oxazolyl oxygen of the substrate. Ser169 is another important residue, which can form a hydrogen bond to the ammonium group of Lys201, and Tyr126 is the residue included for π stacking with the substrate (Figure 5B). There are also additional tyrosines in the binding site, forming parallel and perpendicular π interactions with Tyr126. The initial QM/MM/MC relaxation of KE16 revealed a design issue, as highlighted in Figure 5B and the inset. Specifically, Lys201 appears to not be functioning as desired; in the reactant state it participates in a hydrogen bond with the carboxyl group of the catalytic base, Asp48, as well as one with the isoxazolyl nitrogen of the substrate. When the substrate is removed from the binding site and a full QM/MM/MC relaxation is performed on the apo protein, a tight salt bridge between Asp48 and Lys201 arises. This is troublesome since, upon binding, the substrate has to compete with Lys201 for hydrogen bonding to Asp48. Furthermore, depending on the hydrophobicity of the binding site, full hydrogen transfer might occur, yielding the catalytically disabled protonated Asp48 and deprotonated Lys201. Thus, there is uncertainty concerning the protonation states of Asp48 and Lys201 for both the apo protein and substrate-bound complex.

In the event, the QM/MM FEP MC simulations were carried out for KE16 assuming that Asp48 is deprotonated and Lys201 is protonated in the reactant state. Figure 5A shows the FEP map for the Kemp elimination. A different surface is found; instead of a concerted mechanism as in the other cases, the reaction follows a stepwise pathway. The first step is proton abstraction from the substrate with a low barrier, $\Delta G^{\ddagger} = 3.1$ kcal/mol. Deprotonated Asp48 can be viewed as 'hot', and the transition state (Figure 5C) and resultant isoxazolyl anion are stabilized by ion pairing and hydrogen bonds between the ammonium group of Lys201 and both the isoxazolyl nitrogen and oxygen. Further stabilization comes from Lys201 also acting as a hydrogen-bond donor to the hydroxyl group of Ser169. The intermediate is in a deep free-energy well in which the proton has transferred but the N−O bond is still intact. The reaction is completed via ring opening of the isoxazolyl anion through **TS2** to yield the product. Overall, the mechanism can be classified as a conjugate E1cB process. The FEP calculations for the second step just used the N−O distance as the reaction coordinate. The two-dimensional free-energy surface in Figure 5A only extends to an N−O distance of 2.3 Å; however, the overall profile is projected in one dimension in the inset. In the product state (Figure 5D), Lys210 has shifted so that its ammonium group is hydrogen bonded to the phenolate oxygen.

The facile proton transfer from the substrate to Asp48 is the rate-determining step; the elimination step has a high barrier in view of the stability of the intermediate, but **TS2** still ends up below the reactant state in free energy. The intermediate is the most stable species in this reaction route, and the enzyme efficiency could be adversely affected by trapping in this state. Additional uncertainty arises from again noting that the starting point was deprotonated Asp48 and protonated Lys201. It may be that the true rate-determining step is achievement of this state from the corresponding doubly unionized one. Though in principle this could be addressed by FEP calculations for the Asp to Lys proton transfer, in practice, the results for such processes where two ions are created from two neutral species are prone to large uncertainties associated with the accompanying large changes in the electrostatics, their sensitivity to choices made for the protonation states of all other ionizable residues, and treatment of long-range interactions. Thus, given the uncertainties, comparisons between the computed activation barrier for the Kemp elimination step and the observed k_{cat} are unwarranted.

In summary, the computational results indicated that pursuit of KE16 was risky. Experimentally, it turned out that catalytic activity was observed for this design;[1] however, the k_{cat} and k_{cat}/k_{uncat} were only 0.006 s^{-1} and 5.2 $\times$ 10^3. The computations suggest that improvement could be sought by repositioning the hydrogen-bond donor so that it is less likely to stabilize the base, Asp48, and more likely to hydrogen bond with the incipient phenolate oxygen.

KE15. The repositioning of the hydrogen-bond donor was partially reflected in KE15 (Figure 6). This 1thf-based design is very similar to KE16; Asp48 is retained as the base, but Lys201 is replaced with Met. This eliminates the potentially disadvantageous Asp-Lys interaction, but Met201 does not allow for electrostatic stabilization of the emerging phenolate ion. Met201 is buried in the wall of the binding pocket, and it is expected to play no direct role in the reaction (Figure 6B). Stabilization of the reactant complex by π interactions is achieved by a cluster of tyrosine residues, especially Tyr126 in this design. The QM/MM/MC simulations indicate that Ser169 now plays a more active role. Though it was not designated as the hydrogen-bond donor,[1] the simulations find that it does donate a hydrogen bond to the substrate in the reactant (Figure 6B), transition (Figure 6C), and product states.

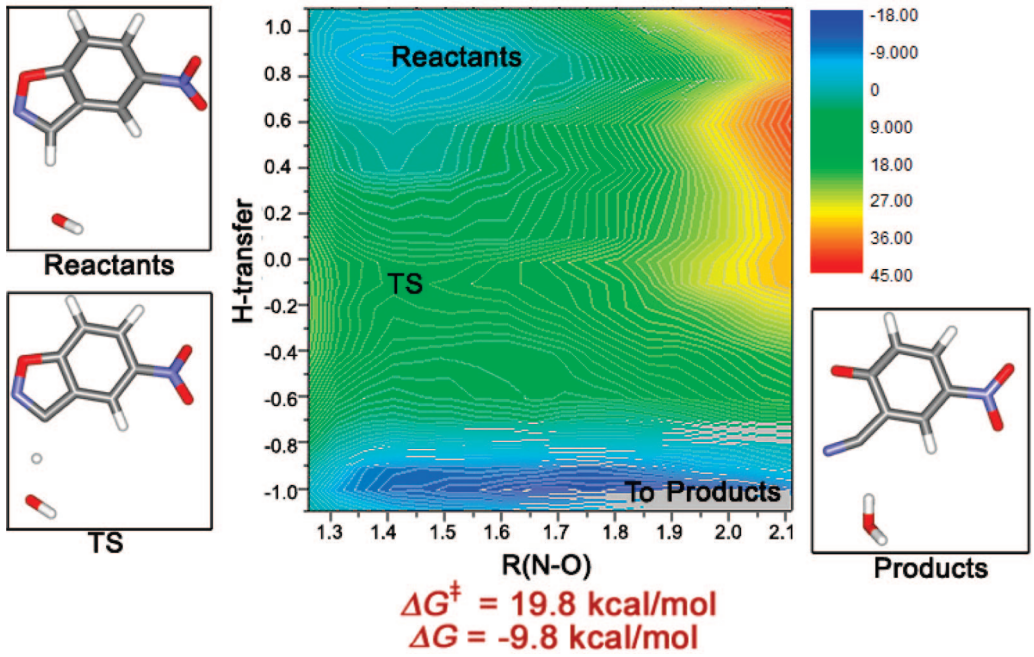


Figure 7. Computed two-dimensional free-energy map for the Kemp elimination catalyzed by hydroxide ion in aqueous solution. The QM region consisted of hydroxide ion plus the substrate; 395 TIP4P water molecules comprised the MM region.

The free-energy map for the Kemp elimination catalyzed by KE15 is similar to those for KE07 and KE10 (Figure 6A). The reaction again proceeds via a concerted mechanism with relatively synchronous proton transfer and N−O bond opening as for KE10. In Figure 6B and 6C the key residues are seen to deliver the desired interactions in the binding site. The orientation of the carboxylate group of Asp48 with the substrate is good, and Ser169 forms a bifurcated hydrogen bond with both the isoxazolyl N and O of the substrate in the reactant and transition states. It shifts to hydrogen bond only to the phenolate oxygen in the product (not illustrated).

The computed free energy of activation is 12.3 kcal/mol, while the observed k_{cat} of 0.022 s^{-1} (Table 1) corresponds to a $\Delta G^{\ddagger}$ of 17.0 kcal/mol. Since the FEP calculations revealed the normal concerted mechanism for KE15 with an activation free energy significantly below the reference value of 19.8 kcal/mol, it was anticipated that KE15 should be a viable catalyst for the Kemp elimination. With a k_{cat}/k_{uncat} of 1.9×10^4 and k_{cat}/K_M of 35.1 M^{-1} s^{-1}, its proficiency is very similar to that of KE07. In view of the complexity of the systems and limitations of the QM/MM methodology, the agreement between the computed and observed activation barriers is encouraging. However, it is emphasized that the comparison of the computed results for the catalyzed and uncatalyzed processes is what is most important for catalyst design.

Reference Reaction in Water. In order to assess the catalytic effect, the Kemp elimination reaction was also modeled in aqueous solution. An immediate issue arose concerning the choice of base. Kemp and Casey did provide experimental free energies of activation for reaction of 5-nitrobenzisoxazole with hydroxide ion, trimethylamine, and water as the base.[13b] Initially, it was decided to make the comparisons with the fastest process, hydroxide catalysis. Thus, the same procedure was followed as for the enzymatic reactions with the substrate and base being equilibrated in contact. The simulations were performed for the reacting system, OH^- plus the substrate, in a periodic box containing 395 TIP4P water molecules in the isothermal−isobaric ensemble at 25 °C and 1 atm. The key results of the FEP calculations are summarized in Figure 7. The mechanism of the hydroxide-catalyzed reaction in water is not significantly different from the one in the designed enzymes. The process remains concerted. It is relatively asynchoronous with proton transfer more advanced in the transition state than N−O bond lengthening; however, the free-energy map is quite flat in the vicinity of the transition state. The process is certainly a classic E2 elimination, as demonstrated by Kemp and co-workers.[13] The computed free energy of activation is 19.8 kcal/mol, again with an estimated uncertainty of less than ±1 kcal/mol. This result is in respectable agreement with the experimental value of 16.0 kcal/mol at 30 °C.[13b] The reaction is also computed to be significantly exergonic with a ΔG of −9.8 kcal/mol, which is qualitatively consistent with the observed exothermicities of Kemp eliminations in water.[13] In this case, the {R(H-base) + R(C−H)} distance was fixed at 3.12 Å after the equilibration period. This is 0.2−0.3 Å greater than for the enzymatic reactions (Table 1). Though simple correlation of this distance and the barrier heights is not found here, it is reasonable that compression of this distance by the enzymes assists catalysis.

The salient point in the present context is that the computed activation barriers near 10 kcal/mol for the reactions catalyzed by the enzymes KE07, KE10(V103N), and KE15 are significantly smaller than the barrier of 19.8 kcal/mol computed for the reference reaction catalyzed by hydroxide ion. The difference is particularly impressive given that hydroxide ion is a much stronger base in water than carboxylate ions. Thus, confidence was high that the designed enzymes would be catalytic. The situation for KE16 was unclear in view of the uncertainties associated with the Asp48-Lys201 pair. Finally, as an alternative comparison point, the reaction in water between free glutamate as the base and 5-nitrobenzisoxazole was also examined with the FEP calculations. The process was found to be concerted and asynchronous with a free energy of activation of 30.2 kcal/

mol, and $\{R(\text{H-base}) + R(\text{C}-\text{H})\}$ increased to 3.20 Å. Clearly, the carboxylate ion is much better stabilized in water compared to the active sites of the designed enzymes, making it a much poorer base.

Conclusions

QM/MM/MC/FEP investigations for Kemp eliminations of 5-nitrobenzisoxazole catalyzed by four computationally designed enzymes were carried out. The catalytic mechanisms and functionality of the key residues in the active sites were elucidated. In comparison to results for reference reactions in water, it was expected that three of the four enzymes should be catalytic and a confident prediction could not be made for the fourth design, KE16. All four designed enzymes were catalytic with k_{cat}/K_M values falling in a relatively narrow range (Table 1) and k_{cat}/k_{uncat} values of 10^3-10^4 for KE07, KE15, and KE16.[1] The catalytic mechanism for KE07, KE10(V103N), and KE15 was computed to be concerted with proton transfer generally more advanced in the transition state than breaking of the isoxazolyl N−O bond. The same is true for the reference reactions in water catalyzed by hydroxide ion or glutamate, so no new intermediates are formed in the enzymatic processes. Hence, no new chemistry is introduced in the designed enzymes in contrast to many highly proficient enzymes;[31] the catalytic effect mainly comes from the specific microenvironment in the binding sites, especially desolvation of the catalytic base.[14,15] If the mechanism cannot be altered, then improved enzyme designs for the Kemp elimination should focus on further increasing the basicity of the catalytic base, optimizing the mutual orientation of the base and substrate for the proton transfer, and better positioning of hydrogen-bond-donating groups to interact with the isoxazolyl N and O in the transition state.

Acknowledgment. Gratitude is expressed to the Defense Advanced Research Projects Agency, Howard Hughes Medical Institute, and the National Institutes of Health (GM032136) for support of this work and to Drs. Kendall N. Houk, Julian Tirado-Rives, and Ivan Tubert-Brohman for assistance and discussions.

JA804040S

(31) (a) Gao, J.; Ma, S.; Major, D. T.; Nam, K.; Pu, J.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3188–3209. (b) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106*, 3210–3235.