ACS Catalysis

Research Article

pubs.acs.org/acscatalysis

Protein Flexibility and Preorganization in the Design of Enzymes. The Kemp Elimination Catalyzed by HG3.17

Katarzyna Świderek,[†,‡] Iñaki Tuñón,[*,†] Vicent Moliner,[*,§] and Joan Bertran[∥]

[†]Departament de Química Física, Universitat de València, 46100 Burjasot, Spain

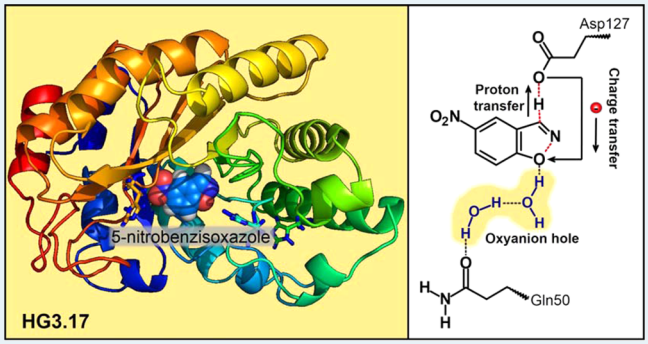
[‡]Institute of Applied Radiation Chemistry, Lodz University of Technology, 90-924 Lodz, Poland

[§]Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castellón, Spain

[∥]Departament de Química, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

Supporting Information

ABSTRACT: A recently designed enzyme, HG3.17, obtained by directed evolution, has shown a catalytic activity close to natural enzymes. Hybrid QM/MM molecular dynamics simulations for the Kemp elimination in this new enzyme have provided a deep insight into the origin of its catalytic efficiency. In this case, we have first demonstrated the presence of different conformations with significantly different reactivity. The larger reactivity is related with a better electrostatic preorganization of the environment that creates a more favorable electrostatic potential for the reaction to proceed. In HG3.17, efforts to improve the catalytic properties must be focused in possible mutations increasing the preorganization and decreasing the reorganization around the oxyanion hole. Mutations should be considered not only in the first shell of residues but in further shells since protein electrostatics is a long-range property. The present work stresses the fact that not all features of catalysis can be revealed of a single structure derived from X-ray diffraction.



KEYWORDS: *Kemp elimination, enzyme design, protein flexibility, QM/MM, molecular dynamics*

INTRODUCTION

Starting from a computationally designed catalyst for the Kemp elimination,[1] an artificial enzyme (HG3.17) has been evolved by Hilvert and co-workers reaching an activity that approaches the exceptional efficiency of highly optimized natural enzymes[2] an achievement that has been considered a breakthrough in protein design.[3] The Kemp elimination, that consists of the conversion of benzisoxazoles into salicylonitriles (see Figure 1), is an interesting reaction due to the fact that it implies a proton transfer from a carbon atom to a heteroatom.[4] Since no naturally occurring enzyme has been identified to catalyze this reaction, it has been used as a benchmark of different protocols to design new enzymes.

The first artificial enzyme designed to catalyze this reaction dates from the pioneering work of Houk, Tawfik, Baker, and co-workers.[5] In this study, the first step was to design reduced models of active sites (theozymes, as defined by Houk and co-workers[6]) by means of quantum mechanical calculations. Then, different proteins selected by means of Rosetta software[7] were used as starting scaffolds to graft the "theozyme" at each of the possible active-site positions, without clashing to the protein backbone and conserving the key interactions observed in the "theozyme".[8] After several cycles of sequence design and protein optimization, fifty-nine designs were experimentally characterized and eight of them resulted to be active. Later, some of these proteins were further improved by directed evolution rendering significant increased activity, ranging from 200-fold for KE07,[9] 400-fold in the case KE70,[10] and up to 2000-fold in KE59.[11] Their efficiencies were similar to the best designed catalytic antibodies, 34E[12] and 13G5.[13]

A different approach to develop more efficient catalysts has been based on the analysis of the origin of inactivity in failed designs of new enzymes. In particular, Houk, Mayo, and co-workers used an iterative approach starting from an inactive protein scaffold, HG1, that was designed to convert the xylan binding pocket of a thermostable xylanase into a Kemp eliminase.[1] Based on molecular dynamics (MD) simulations and X-ray crystallography, the authors found that the inactivity might be due to bound waters and high flexibility of residues in the active site. These findings guided the design of a more embedded binding pocket, moving the active site deeper into the protein, resulting in an active Kemp eliminase, HG2. MD simulation of this last structure led to an additional mutation that provided a better packing around the substrate by reducing

Received: November 30, 2014
Revised: March 8, 2015
Published: March 10, 2015

DOI: 10.1021/cs501904w
ACS Catal. 2015, 5, 2587−2595

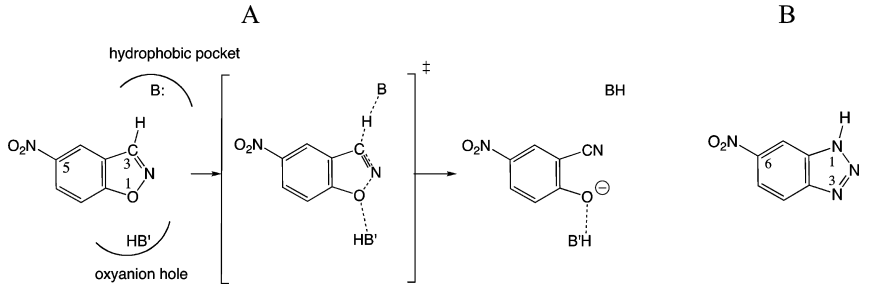


Figure 1. (A) Representation of the base-catalyzed Kemp elimination in 5-nitrobenzisoxazole. (B) Inhibitor 6-nitrobenzotriazole.

the active site conformational heterogeneity observed in HG2. The new designed protein, HG3, exhibited activity comparable to the best Rosetta designs[7−9] but was still far from the efficiency of natural enzymes. Hilvert and co-workers performed 17 rounds of mutagenesis and screening on this protein, giving rise to a new catalyst with an efficiency comparable to the one exhibited by many natural enzymes.[2] Based on the analysis of X-ray structures, three factors were suggested to be decisive in the improvement of the Kemp eliminase from HG3 to HG3.17: (1) the extraordinary high shape complementarity between the binding pocket of the protein and the substrate; (2) the ligand alignment with Asp127, the catalytic base, that was optimized by evolution resulting in an unusual short hydrogen bond; and (3) a new catalytic group, Gln50, that stabilized the development of a negative charge on the O1 atom of the substrate at transition state (TS).[2]

Keeping in mind the possible limitations associated with the analysis based on a single X-ray structure of the protein complexed with an inhibitor, the aim of the present study is to get a deeper insight into the origin of the high catalytic activity in HG3.17 considering the implications of protein flexibility. In particular, quantum mechanics/molecular mechanics (QM/MM) MD simulations have been performed to study the conformational diversity shown by the protein complexed with the real substrate in solution. The comparison of the reaction in the protein environment with the uncatalyzed counterpart reaction with acetate in aqueous solution will allow emphasizing the role of the electrostatically preorganized active site that can explain the rate enhancement observed in this new catalyst. In this paper, despite no specific mutations are proposed, we focus on strategies to guide future designs of new Kemp eliminases.

METHODS

The X-ray structure of the Kemp Eliminase HG3.17, which contains two surface mutations and complexed with the transition state analogue 6-nitrobenzotriazole (PDB ID 4BS0),[2] was used as a starting point in our simulations. This high resolution (1.09 Å) structure contains the coordinates hydrogen atoms, avoiding then ambiguities in the assignment of the protonation states of titratable residues. In the PDB structure 4BS0 the protein crystallizes as a dimer, containing two equivalent chains denoted as A and B. The monomers in this structure present small structural differences, including the number of crystal waters. In order to take into account the possible effects of conformational diversity, the two monomers were simulated independently. Both systems were placed in pre-equilibrated box of water molecules with size of 100 Å × 80 Å × 80 Å and neutralized by adding four sodium counterions. Water molecules with an oxygen atom lying within 2.8 Å of any heavy atom of the protein were removed. The transition state analogue was substituted by the substrate of the reaction, just changing the nature of nitrogen atoms 1 and 3 to carbon and oxygen atoms, respectively (see Figure 1). The system was simulated using a hybrid QM/MM potential, where the QM subsystem was composed by the substrate and the side chain of Asp127. The QM subsystem was described by means of the semiempirical AM1 Hamiltonian[14] during MD simulations and by the density functional method M06-2X/6-31+G*[15] during the exploration of the potential energy surface associated with the Kemp reaction. The rest of the system (protein, water molecules, and counterions) was described using the OPLS-AA[16] and TIP3P[17] force fields, as implemented in the fDYNAMO library.[18] To saturate the valence of the QM/MM frontier atoms, a link atom was placed between the Cα and Cβ atoms of the Asp127.[19] Cutoffs for nonbonding interactions were applied using a switching-force scheme, within a range radius from 14.5 to 16 Å. After minimization of the full system, those residues lying more than 20 Å apart of any of the substrate atoms was kept frozen in the remaining calculations. After thermalization, a QM/MM MD simulation of the system in the NVT ensemble was ran during 2 ns at a temperature of 300 K using the Langevin-Verlet algorithm using a time step of 1 fs. According to the time-dependent evolution of the RMSD of those atoms belonging to the protein backbone (see Figure S1 in the Supporting Information), the system can be considered equilibrated after 1 ns for both monomers A and B of simulation.

In order to obtain the free energy landscape, we have traced the potentials of mean force (PMFs) for both monomers A and B. The use of free energy landscapes in chemical reactivity has been explained and defined in the past.[20−22] The reaction is followed as a function of two distinguished reaction coordinates: the antisymmetric combination of distances defining the proton transfer from the substrate to Asp127 (ξ_1 = d(C1−H) − d(H−OD2Asp127)) and the distance defining the ring opening process (ξ_2 = d(N2−O3)). The procedure for the PMFs calculation requires a series of molecular dynamics simulations in which the distinguished reaction coordinates are constrained around particular values with the umbrella sampling procedure,[23] while the remaining degrees of freedom (including those of the protein environment) are conveniently sampled. The values of the variables sampled during the simulations are then pieced together to construct a distribution function using the weighted histogram analysis method (WHAM).[24]

Because of the large number of structures that must be evaluated during free energy calculations, QM/MM calculations are usually restricted to the use of semiempirical Hamiltonians.

DOI: 10.1021/cs501904w
ACS Catal. 2015, 5, 2587−2595

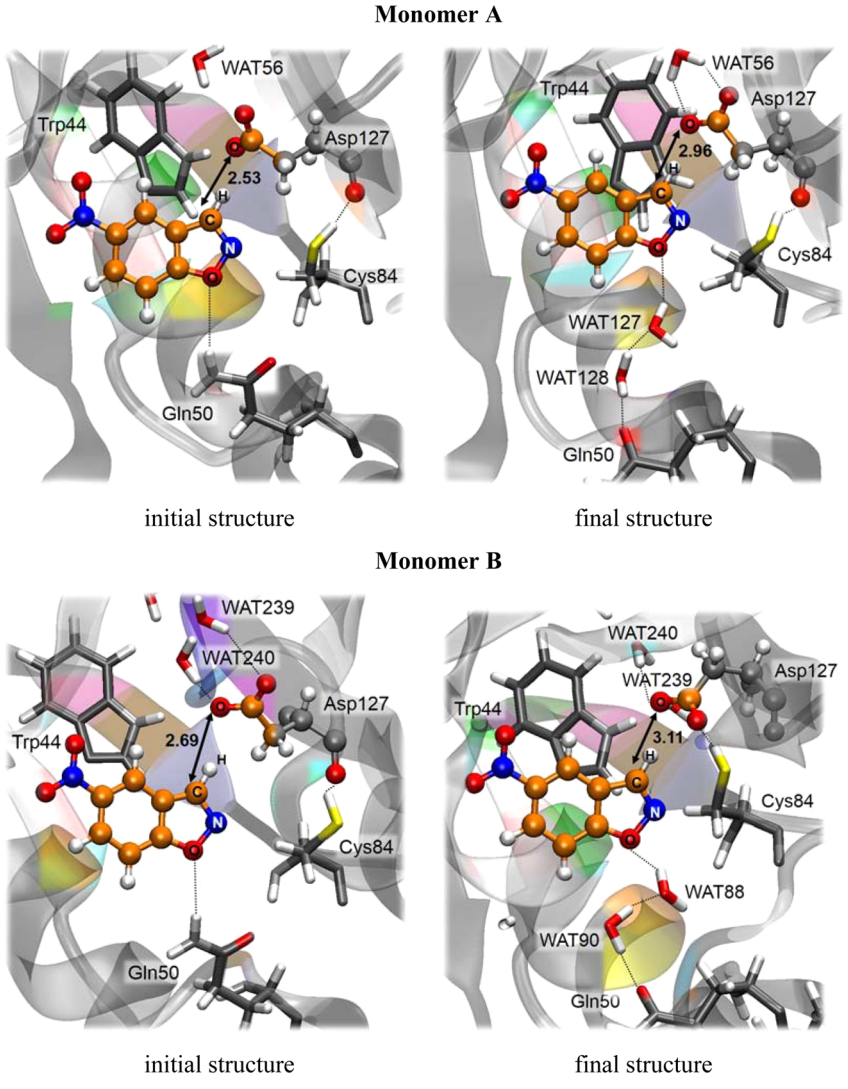


Figure 2. Initial structure (equivalent to the X-ray structure, despite different substrates are placed in the active site) and final structure (2 ns) of QM/MM MD simulations of monomer A and B.

In order to reduce the errors associated with the quantum level of theory employed in our simulations, following the work of Truhlar et al.[25−27] a spline under tension[28] is used to interpolate this correction term at any value of the reaction coordinates ξ_1 and ξ_2 selected to generate the free energy surfaces. In this way, we obtain a continuous function in a new energy function to obtain corrected PMFs:[29,30]

$$E = E_{\mathrm{AM1/MM}} + S[\Delta E_{\mathrm{LL}}^{\mathrm{HL}}(\xi_1, \xi_2)] \quad (1)$$

where S denotes a two-dimensional spline function, and its argument is a correction term evaluated from the single-point energy difference between a high-level (HL) and a low-level (LL) calculation of the QM subsystem. The AM1 semiempirical Hamiltonian was used as LL method, while the M06-2X/6-31+G** method was selected for the HL energy calculations. In this case, S is adjusted to a grid of 61 × 28 points obtained as HL single energy calculation corrections on geometries optimized at LL. The HL calculations were carried out using the Gaussian09 program.[31]

KEMP Elimination in Aqueous Solution. We studied the reaction between acetate and 5-nitrobenzisoxazole using both continuum and discrete representations of the solvent. The SMD continuum model[32] as implemented in Gaussian09 at the M06-2X/6-31+G** level was used for the former model. TS for the proton transfer through the anti and syn orbitals of the carboxylate group were localized and characterized by inspection of the frequencies in the presence of discrete TIP3P water molecules. Reactants for this bimolecular process were computed as fully separated acetate anion and 5-nitrobenzisoxazole obtained also with a discrete representation of the solvent. Simulations for the reactants and TS structures for the discrete model of the solvent were carried out using the same simulation protocol than for the protein.

RESULTS AND DISCUSSION

Our calculations start from the X-ray structure of the HG3.17 variant (PDB code 4BS0) that was crystallized as a dimer.[2] The two monomers, named as chains A and B in the PDB file, present the same sequence but different conformations. In particular the microenvironment where the base is located, the inner side of a hydrophobic binding pocket, contains different number of crystallographic water molecules. Differences in active site hydration, driven by protein conformational changes, can have important consequences on reactivity; as it has been already shown in the case of a ketosteroid isomerase.[33] In our

DOI: 10.1021/cs501904w
ACS Catal. 2015, 5, 2587−2595

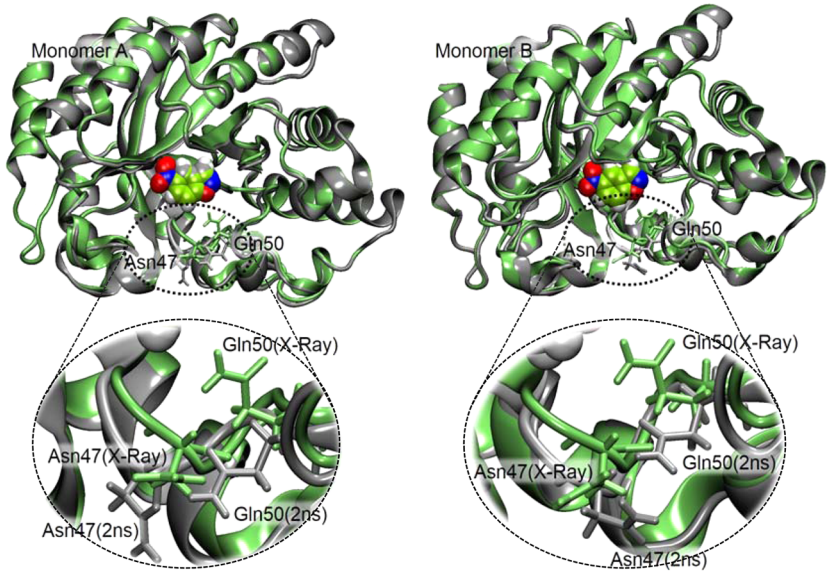


Figure 3. Overlapping of structures of the initial X-ray structure (green) and after 2 ns of QM/MM MD simulations (gray) in monomer A and monomer B. A close-up of the oxyanion hole is presented in the bottom panels.

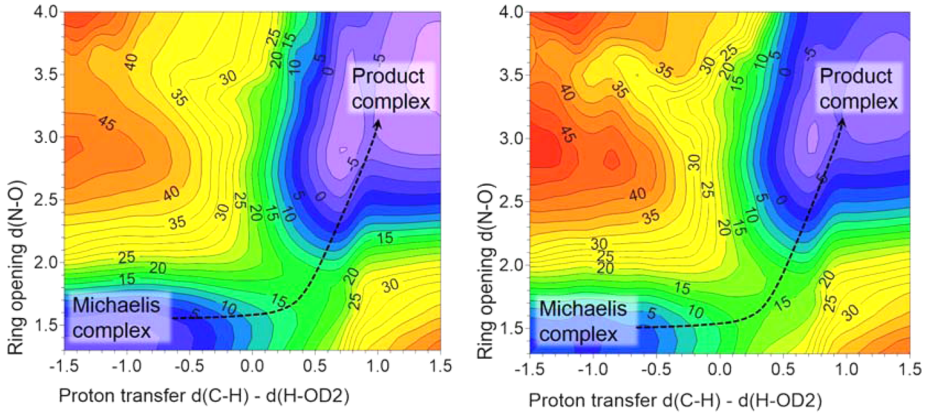


Figure 4. Free energy surfaces obtained at AM1/MM level with spline corrections at M06-2X/6-31+G(d,p)/MM level for monomer **A** and **B**. Energies are given in kilocalories per mole, and distances, in angstroms.

case, after preparing the monomers in solution, it appears that solvent water molecules from the bulk have different access to this active site in the two monomers. This reflects different hydrophobicity in the base region of the active site depending on the monomer. In the oxyanion hole of the active site, the phenoxide oxygen-leaving group is interacting with the amine group of Gln50 and a water molecule, a situation that is observed in both monomers. The detected differences motivated us to conduct the study using both monomer structures as starting point. After replacing the inhibitor 6-nitrobenzotriazole by the 5-nitrobenzisoxazole substrate (see Figure 1) in both monomers, 2 ns QM/MM MD simulations have been performed for each monomer in solution reaching an equilibrated situation (see Figure S1 and S2 of the Supporting Information). Snapshots of the initial and final structures of the active site of both monomers are presented in Figure 2.

Analysis of the MD simulations shows that X-ray structures do not correspond to the representative structures of the Michaelis complex in solution. First, the inhibitor and the real substrate present different electronic properties. And second, the ensemble of conformations explored by the protein in solution is different to the solid state structures of the protein-inhibitor complex. As observed in Figures 2 and 3, both active sites experience important conformational changes during the simulations. The distance between the substrate and the carboxylate oxygen atom of Asp127 is enlarged by more than 0.5 Å during the MD simulations in both monomers. This is basically due to the replacement of a nitrogen atom in the inhibitor by a carbon atom in the substrate that decreases the acidity of the hydrogen atom attached to this position 3 of the substrate (see Figure S3 of the Supporting Information for a comparison of the electronic population in the inhibitor and real substrate). As shown below, this distance will be reduced at the TS (ca. 2.7 Å) in order to facilitate the hydrogen transfer. In this sense, the inhibitor 6-nitrobenzotriazole appears to be a good transition state analogue.

The other important change observed in both monomers is found in the surroundings of the phenoxide leaving group, where two water molecules are placed between the O1 oxygen atom of the substrate and Gln50 during the simulations. These water molecules appear to be forming a bridge of hydrogen bond interactions between the oxygen atom of the substrate

DOI: 10.1021/cs501904w
ACS Catal. 2015, 5, 2587−2595

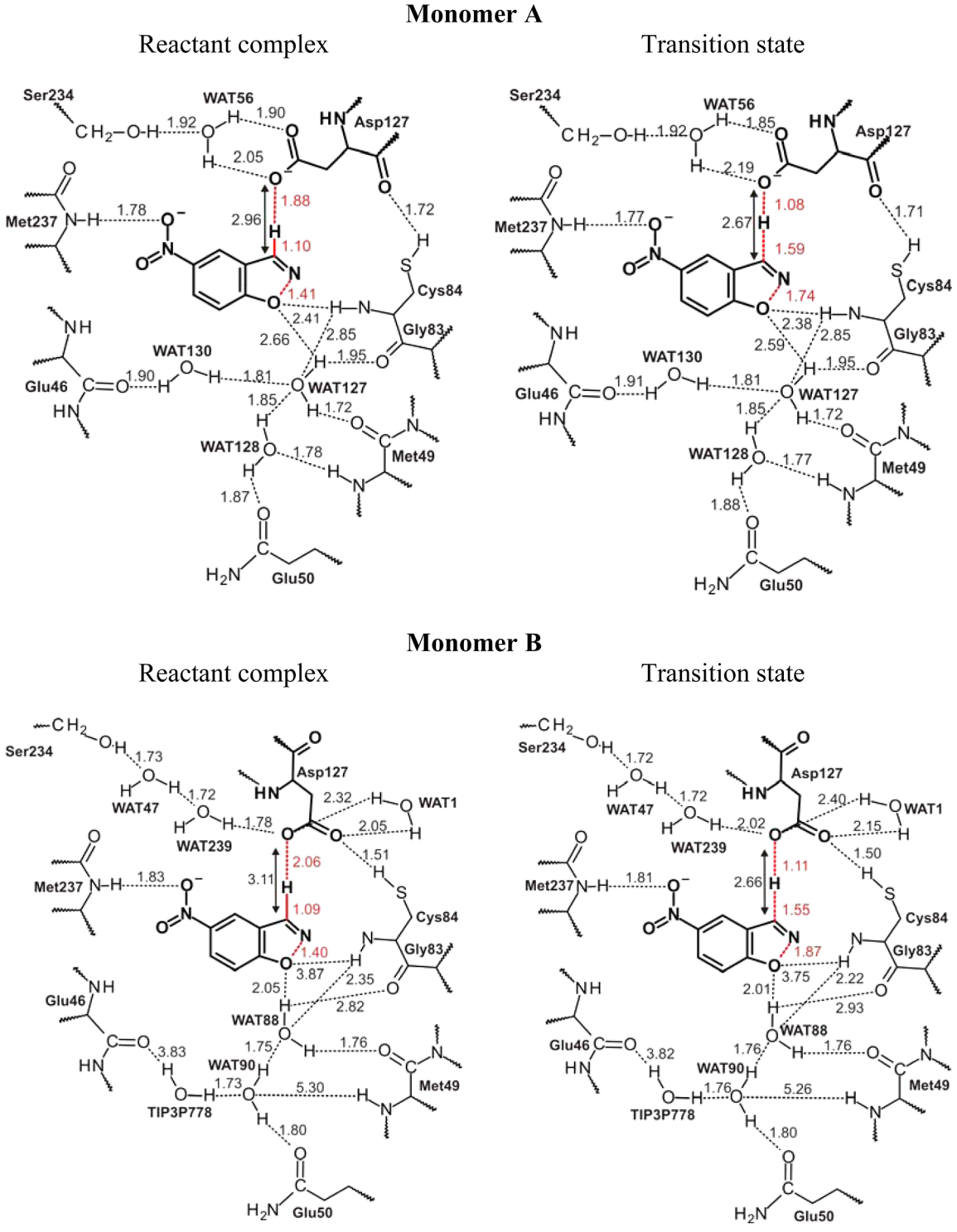


Figure 5. Schematic representation of TS and reactant complex structures in monomer A and monomer B obtained at the M06-2X/MM level. Distances are reported in angstroms.

and the oxygen atom of carbonyl group of Gln50. This result is in agreement with X-ray diffraction analysis of 13G5 antibody[34] and different variants of HG3,[35] suggesting that water molecules can be well-suited in the active site to donate a hydrogen bond to the leaving group during catalysis. Moreover, we observed that water molecules also displace the Gln50 residue when simulations were carried out with the transition state analogue 6-nitrobenzotriazole placed in the active site. Regarding future design processes, it is important to emphasize that the entrance of water molecules implies not only the rotation of the side chain of active site residues but also the displacement of the backbone of the protein loop placed in the bottom of the active site (see Figure 3).

The main difference between the final structures of monomer A and B is observed in the relative orientation of Asp127 with respect to the substrate. In the case of monomer B, the carboxylate group of Asp127 suffers a rotation and the new conformation is stabilized by means of a new hydrogen bond

DOI: 10.1021/cs501904w
ACS Catal. 2015, 5, 2587−2595

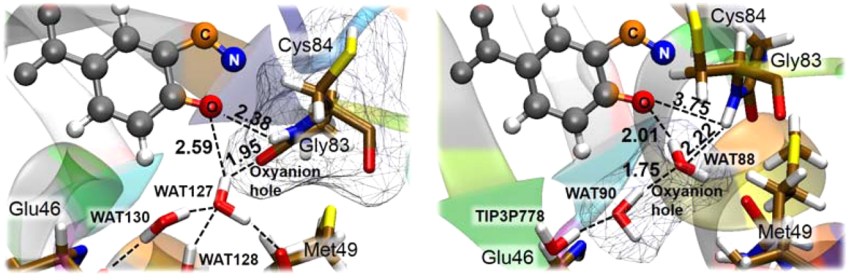


Figure 6. Detail of the oxyanion hole in TS structures located in monomers A and B.

established between this carboxylate group and the thiol moiety of Cys84. In contrast, the initial orientation of the carboxylate group of Asp127 is kept in monomer A. Thus, while Asp127 is interacting with the substrate through a syn orbital in monomer B, an interaction through an antiorbital is observed in monomer A. In the X-ray structure, this interaction involves the antiorbital in both monomers (see Figure 2). These observations suggest that different conformations can have different reactivity.

The last structures of the two MD simulations on monomer A and B were used to study the reactivity of the catalyzed Kemp elimination reaction. The N2−O1 distance and the antisymmetric combination of the distances defining the transfer of the hydrogen atom from the C3 carbon atom of the substrate to the carboxylate oxygen atom (OD2) of Asp127 were used as distinguished reaction coordinates to obtain free energy surfaces (see Figure 4) by means of simulations where the remaining degrees of freedom are sampled.

Free energy surfaces presented in Figure 4 show that both reaction paths proceed by concerted but asynchronous mechanisms. The free energy barriers deduced from the surfaces are 16.3 and 13.8 kcal·mol^{-1}, this last value in good agreement with the one derived from the experimental rate constant measured by Hilvert and co-workers (13.7 kcal·mol^{-1}) for the original HG3.17 version of the enzyme.[2] Thus, the conformation of monomer B seems to be significantly more reactive than monomer A. Inclusion of contributions derived from tunnelling effects and dynamical recrossing does not change the observed trend (see Table S1 in the Supporting Information for details). As observed in the results of the total transmission coefficients computed for monomer A and B (1.00 and 0.93, respectively), no significant differences are found between them in this regard. In fact, while in monomer A the recrossing transmission coefficients is slightly higher than in the monomer B (0.48 and 0.32, respectively), the opposite trend is observed in the tunneling transmission coefficient (2.08 and 2.90, respectively). In any case, the observed differences would affect the relative effective free energy barriers in less than 1/10 kcal·mol^{-1}. These results indicate that the differences in reactivity are not due to differences in dynamical behavior (in recrossing) or in tunneling.

In order to get a deeper insight into the origin of the observed different reactivity in monomer A and B, we have optimized TS and reactants structures in both monomers at M06-2X/MM level (see Figure 5). As observed, the TS of monomer A is described by a more advanced proton transfer and a less advanced isoxazolyl N−O bond breaking than the TS of monomer B. This difference can be understood by analysis of conformational changes occurring in the active sites. First of all, the less advanced hydrogen bond transfer observed in monomer B is due to the hydrogen bond interaction found between the thiol group of Cys84 and the carboxylate group of Asp127 that stabilizes its basic form. The donor−acceptor distance, that is considerably shortened from the Michaelis complex to the TS, is significantly longer in the reactant state of monomer B than in monomer A, because the larger stabilization of the base. This effect would increase the free energy barrier for the proton transfer in monomer B. Nevertheless, this is not the only remarkable difference between active sites of monomer A and B. The conformational change of Cys84 provokes differences not only in the surroundings of the base, but also in the oxyanion hole. The backbone N−H group of Cys84 interacts through a hydrogen bond with the phenoxide oxygen-leaving group in the TS of monomer A stabilizing the development of a negative charge on this atom. However, the backbone N−H group of Cys84 in monomer B activates a water molecule which is the species making a stronger hydrogen bond interaction with the oxygen atom (see Figure 6). Thus, the improvement in the design of the oxyanion hole seems to be the driving force explaining the different reactivity of both monomers. These conformational differences associated with the displacement of Cys84 are already observed in the reactants complexes, as shown in Figure 5. Interestingly, the structures of optimized reactants and TS at HL/MM (Figure 5) and LL/MM (see Figure S4 of the Supporting Information) are qualitatively in very good agreement, thus giving credit to the hybrid MD simulations performed with the AM1 semiempirical Hamiltonian and the correction scheme used a posteriori by means of eq 1.

Our TSs structures can be compared with previous structures located for similar reactions at different levels of theory. Thus, optimized TSs structures for the formate-catalyzed opening of benzisoxazole computed by Houk and co-workers in a gas phase model[36] show a similar stage in the proton transfer to the base but a less advanced N−O bond breaking. Computed transition states for the reaction of acetate and butylamine with 4-nitrobenzisoxazole with continuum models[37] presented the opposite trend, a similar description of the N−O bond breaking but a much more synchronous process for the proton transfer. Finally, Jorgensen and co-workers also localized different TSs for the Kemp elimination catalyzed by the designed enzymes KE07, KE10(V131N), and KE15 with QM/MM methods. In this case, the process appeared to be also concerted but with proton transfer generally more advanced in the transition state than the breaking of the isoxazolyl N−O bond.[38]

Analysis based on structural changes is supported by electrostatic arguments. The reaction can be described as a charge transfer from Asp127, negatively charged at the reactants

DOI: 10.1021/cs501904w
ACS Catal. 2015, 5, 2587−2595

Table 1. Averaged Electrostatic Potentials (kJ·mol^{-1}·e^{-1}) Created by the Environment, in Monomer A, Monomer B, and in Aqueous Solution, on the Key Atoms of the Reaction at the Reactants (R) and Transition States (TS) and Their Difference (Δ)

	monomer A			monomer B		
	reactants[a]	TS	Δ(TS − R)	reactants[a]	TS	Δ(TS − R)
OD1(Asp127)	345.8 ± 24.7	310.0 ± 21.7	−35.8	501.4 ± 29.5	431.4 ± 30.8	−70.0
OD2(Asp127)[c]	381.2 ± 22.9	357.7 ± 22.5	−23.5	585.8 ± 32.6	424.8 ± 52.4	−161.0
C3	200.1 ± 20.7	227.8 ± 20.1	27.7	256.9 ± 22.3	321.1 ± 29.3	64.2
N2	199.8 ± 21.2	243.7 ± 22.3	43.9	249.0 ± 27.0	330.8 ± 25.5	81.8
O1	141.1 ± 26.1	213.2 ± 30.9	90.1	203.8 ± 32.5	298.3 ± 29.6	94.5
	water anti orientation			water syn orientation		
	reactants[b]	TS	Δ(TS − R)	reactants[b]	TS	Δ(TS − R)
OD1(Asp127)	753.4 ± 55.0	530.4 ± 44.0	−223.0	753.4 ± 55.0	471.3 ± 44.8	−282.1
OD2(Asp127)[c]	719.7 ± 55.2	447.9 ± 47.4	−271.8	719.7 ± 55.2	386.6 ± 41.2	−333.1
C3	11.3 ± 33.9	312.8 ± 35.4	301.5	11.3 ± 33.9	374.3 ± 32.4	363.0
N2	93.3 ± 45.9	372.9 ± 41.8	279.6	93.3 ± 45.9	456.0 ± 38.7	362.7
O1	80.2 ± 45.6	376.5 ± 44.8	296.3	80.2 ± 45.6	397.8 ± 46.4	317.6

[a]Electrostatic potential computed for the Michaelis complex. [b]Separated reactants in water. [c]Oxygen acceptor atom of transferred proton.

state, to the phenoxide oxygen-leaving atom, as observed in the analysis of charges (see Table S2 of the Supporting Information). Then, the process can be favored if the electrostatic potential is increased around the leaving group and decreased around the base. Table 1 shows the average electrostatic potentials created by the protein in monomers A and B on key atoms of the reacting system. The electrostatic potentials on the oxygen atoms of the base (OD1 and OD2) are higher in monomer B than in A, stabilizing more the basic form of Asp127 and disfavoring the hydrogen transfer. On the other hand, the electrostatic potential on the O1 atom of the substrate, that is also larger in monomer B, favors the cleavage of the N−O bond and stabilizes the negative charge developed on this oxygen atom at the TS. From reactants to TS, the negative charge on this atom is increased by −0.31 au in monomer B, while in monomer A the increase is smaller, −0.22 au (see Table S2 of the Supporting Information).

This study demonstrates that differences up to 60-fold on the rate constants can be observed depending on the protein conformations. Similar and even larger differences in rate constants measured in single molecule experiments have been attributed to differences among protein conformations.[39] This is a conclusion that has also been derived from previous computer simulations in different enzyme systems.[40−43] Catalytic activity could be improved increasing the population of the most reactive protein conformer. Keeping in mind that this study has demonstrated the role of Cys84 modulating the transition between the two observed conformation, mutations in the proximity of this residue, favoring the rotamer observed in the simulations of monomer B, could render successful results. The present study stress the importance of performing large sampling of conformational space to get a realistic picture of the catalyst in solution. The X-ray structure, as mentioned above, does not necessarily correspond to the most reactive conformation. This is in agreement with previous computational studies performed on a wide range of different enzymatic systems,[44−47] showing that the crystal structure can represent only a minor conformation of those that can be observed in solution and, consequently, cannot always be considered as a structure relevant for the reaction. Moreover, the crystal structure is usually obtained with an inhibitor in the active site that does not correspond to the real TS of the reaction.

In order to fully understand the origin of the enhanced catalytic activity of the new enzyme, the process has to be compared with the counterpart noncatalyzed reaction in solution. In this case, acetate was selected as the base since this is the species closer to the aspartate residue that acts as a base in the cavity of HG3.17. Also, this is the system experimentally analyzed by Hilvert and co-workers.[10] Syn and anti conformations of TS structures in solution have been obtained following a similar computational protocol as the one used to study the reaction in the designed enzyme, including the interaction with MM solvent molecules. The exploration of the two-dimensional free energy surfaces shows that the reactant complex is not a minimum in solution. Since the process is a bimolecular reaction, the reactants of the reaction correspond then to the substrate and acetate as fully separated and solvated species. Then, a continuum model has been employed to study the reaction in solution, obtaining transition structures very close to those found with the discrete representation of the solvent (see Figure S5 in the Supporting Information). The free energy barriers derived from the continuum model calculations, measured from the separated reactants, are 26.2 and 25.6 kcal·mol^{-1} for the anti and syn conformation of the TS, respectively. These values are in good agreement with the experimental kinetic data in solution obtained by Hilvert and co-workers (23.4 kcal·mol^{-1}).[10] These results, confirmed by gas phase calculations, show that the intrinsic reactivity of the syn and anti orbitals is very close, in contradiction with previous proposal that suggested that the syn orbital is more basic than the anti orbital.[48]

The electrostatic potentials on key atoms of the reaction have been also obtained for the noncatalyzed process and are reported in Table 1. At the reactants state the potential created by the protein is less positive around the base and more positive around the leaving group than in aqueous solution. These results can be interpreted as the protein environment being better electrostatically preorganized for the reaction to proceed. Moreover, the differences observed in the electrostatic potential created by the environment when going from reactants to the TS (Δ in Table 1) are significantly larger (about 1 order of magnitude in some cases) for the reaction in aqueous solution than in the protein-catalyzed reaction. As a consequence a smaller energetic penalty has to be paid to reorganize the environment from reactants to TS, resulting in a

DOI: 10.1021/cs501904w
ACS Catal. 2015, 5, 2587−2595

smaller free energy barrier. This is a finding that has been previously emphasized by Warshel and co-workers.[49−51] and by Labas et al.[52] Comparing the electrostatic potentials calculated at the TS in the protein and in aqueous solution it seems possible to improve the efficiency of the catalysts optimizing the environment around the phenoxide leaving group, increasing the electrostatic potential on the O1 atom. Note that monomer B presents a more efficient oxy-anion hole from the electrostatic point of view, but this is essentially due to the water molecules present in the active site, which also suffer a larger reorganization. A larger reorganization is also observed in the potential calculated on OD2 in monomer B, because more water molecules are found around the base in this case (see Figure 5). A more efficient active site should be formed by protein residues electrostatically preorganized at the Michaelis complex.

CONCLUSIONS

The results presented in this study, based on hybrid QM/MM MD simulations for the Kemp elimination in a variant of the recently designed HG3.17 enzyme, have provided a deeper insight into the origin of its catalytic efficiency. The use of different monomers as starting point of our simulations, have demonstrated the significantly different reactivity that different protein conformations can exhibit. According to our results, the catalytic efficiency of the enzymatic conformations depend on the different degree of electrostatic preorganization achieved in each of them. In particular, this study demonstrates the key role played by the electrostatic properties of the oxy-anion hole present in the active site. The monomer with the lower free energy barrier, monomer B, has been shown to display the most adequate electrostatic potential around the leaving atom. The catalytic properties of the enzyme could then be favored maximizing the population of those conformers that are better electrostatically prepared for the reaction. Design optimization must consider not only a high shape complementarity between the protein and the polarized substrate in the TS conformation, but also the electrostatic complementarity between the active site and the reaction process. This implies not only the first shell residues because the electrostatic potential is a long-range property. Mutations should be guided to improve the active site from the electrostatic point of view and, in particular, in this case around the oxyanion hole. Nevertheless, a note of caution has to be introduced at this point since mutations that favor electrostatic interactions with the substrate in the TS, which a priori would improve the chemical step, can provoke structural changes that might affect other steps of the process such as substrate binding and/or products release.

The fact that the protein is preorganized to stabilize the TS does not mean that the protein provides a rigid environment for the reaction. As observed in our results, the differences observed between the reactant state and the transition state, not only in geometries but also in the changes of the electrostatic potentials created by the protein, indicate that a certain degree of reorganization is also required in the protein. This supports the need of a flexible environment model during the MD simulations. Protein flexibility is essential to maximize catalysis and simplified models where the protein is considered as a rigid body can hinder important aspects of the catalysts.[53−56] The main differences observed between the behavior of monomer A and B come from the fact that a flexible protein backbone has been used during the MD simulations, thus allowing to obtain different results for each initial conformer. A correct balance between preorganization and flexibility, both from a mechanical (short-range) and electrostatic (long-range) perspectives, is required to optimize the design of new enzymes. This can be only accomplished by means of computational simulations of the reaction process that consider the influence of a large and flexible environment. Finally, the present work stresses the fact that not all features of catalysis can be revealed of a single structure derived from X-ray diffraction and that complementary molecular simulations are required to achieved precision in the enzyme design process.

ASSOCIATED CONTENT

Supporting Information

The following file is available free of charge on the ACS Publications website at DOI: 10.1021/cs501904w.

Time dependent evolution of RMSD and key distances along the QM/MM MD simulations on the two monomer structures of initial X-ray structure and after the QM/MM MD simulations of monomer A and monomer B, atomic charges of substrate and inhibitor computed in gas phase, in solution and in the two monomers, geometries of optimized structures of reactants and TS of both monomers at AM1/MM level, TS geometries obtained in aqueous solution, and simulations methodological details of QM/MM calculations of geometries optimizations and recrossing and tunneling coefficients (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: moliner@uji.es (V.M.).
*E-mail: ignacio.tunon@uv.es (I.T.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Spanish Ministerio de Economía y Competitividad for project CTQ2012-36253-C03, Universitat Jaume I (project P1·1B2011-23), Generalitat Valenciana (PROMETEOII/2014/022 and ACOMP/2014/277 projects), Polish National Center for Science (NCN) (grant 2011/02/A/ST4/00246, 2012−2017), the Polish Ministry of Science and Higher Education ("Iuventus Plus" program project no. 0478/IP3/2015/73, 2015-2016), and the USA National Institute of Health (ref NIH R01 GM065368). Authors acknowledge computational resources from the Servei d'Informàtica of Universitat de València on the "Tirant" supercomputer and the Servei d'Informàtica of Universitat Jaume I.

REFERENCES

(1) Privett, H. K.; Kiss, G.; Lee, T. M.; Blomberg, R.; Chica, R. A.; Thomas, L. M.; Hilvert, D.; Houk, K. N.; Mayo, S. L. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 3790−3795.
(2) Blomberg, R.; Kries, H.; Pinkas, D. M.; Mittl, P. R. E.; Grütter, M. G.; Privett, H. K.; Mayo, S. L.; Hilvert, D. *Nature* **2013**, *503*, 418−421.
(3) Höhne, M.; Bornscheuer, U. T. *Angew. Chem., Int. Ed.* **2014**, *53*, 1200−1202.
(4) Casey, M. L.; Kemp, D. S.; Paul, K. G.; Cox, D. D. *J. Org. Chem.* **1973**, *38*, 2294−2301.
(5) Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.;

DOI: 10.1021/cs501904w
ACS Catal. 2015, 5, 2587−2595

Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. *Nature* **2008**, *453*, 190−195.
(6) Tantillo, D. J.; Chen, J.; Houk, K. N. *Curr. Opin. Chem. Biol.* **1998**, *2*, 743−750.
(7) Richter, F.; Leaver-Fay, A.; Khare, S. D.; Bjelic, S.; Baker, D. *PLoS One* **2011**, *6*, e19230.
(8) Zanghellini, A.; Jiang, L.; Wollacot, A. M.; Cheng, G.; Meiler, J.; Althoff, E. A.; Röthlisberger, D.; Baker, D. *Protein Sci.* **2006**, *15*, 2785−2794.
(9) Khersonsky, O.; Röthlisberger, D.; Dym, O.; Albeck, S.; Jackson, C. J.; Baker, D.; Tawfik, D. S. *J. Mol. Biol.* **2010**, *396*, 1025−1042.
(10) Khersonsky, O.; Röthlisberger, D.; Wollacott, A. M.; Murphy, P.; Dym, O.; Albeck, S.; Kiss, G.; Houk, K. N.; Baker, D.; Tawfik, D. S. *J. Mol. Biol.* **2011**, *407*, 391−412.
(11) Khersonsky, O.; Kiss, G.; Röthlisberger, D.; Dym, O.; Albeck, S.; Houk, K. N.; Baker, D.; Tawfik, D. S. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 10358−10363.
(12) Thorn, S. N.; Daniels, R. G.; Auditor, M. T. M.; Hilvert, D. *Nature* **1995**, *373*, 228−230.
(13) Müller, R.; Debler, E. W.; Steinmann, M.; Seebeck, F. P.; Wilson, I. A.; Hilvert, D. *J. Am. Chem. Soc.* **2007**, *129*, 460−461.
(14) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(15) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(16) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(17) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(18) Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21*, 1088−1100.
(19) Field, M. J.; Bash, P. A.; Karplus, M. A. *J. Comput. Chem.* **1990**, *11*, 700−733.
(20) Samuelson, S. O.; Martyna, G. J. *J. Chem. Phys.* **1998**, *109*, 11061−11073.
(21) Proust-De Martin, F.; Dumas, R.; Field, M. J. *J. Am. Chem. Soc.* **2000**, *122*, 7688−7697.
(22) Poulsen, T. D.; Garcia-Viloca, M.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. B* **2003**, *107*, 9567−9578.
(23) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(24) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(25) Corchado, J. C.; Coitiño, E. L.; Chuang, Y.; Fast, P. L.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 2424−2438.
(26) Nguyen, K. A.; Rossi, I.; Truhlar, D. G. *J. Chem. Phys.* **1995**, *103*, 5522−5530.
(27) Chuang, Y. Y.; Corchado, J. C.; Truhlar, D. G. *J. Phys. Chem. A* **1999**, *103*, 1140−1149.
(28) (a) Renka, R. J. *SIAM J. Stat. Comput.* **1987**, *8*, 393−415. (b) Renka, R. *J. ACM Trans. Math. Software* **1993**, *19*, 81−94.
(29) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S. *J. Phys. Chem. B* **2006**, *110*, 17663−17670.
(30) Ferrer, S.; Martí, S.; Moliner, V.; Tuñón, I.; Bertran, J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 3482−3489.
(31) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.
(32) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378−6396.
(33) van der Kamp, M. W.; Chaudret, R.; Mulholland, A. J. *FEBS J.* **2013**, *280*, 3120−3131.
(34) Debler, E. W.; Müller, R.; Hilvert, D.; Wilson, I. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 18539−18544.
(35) Hilvert, D. personal communication.
(36) Na, J.; Houk, K. N.; Hilvert, D. *J. Am. Chem. Soc.* **1996**, *118*, 6462−6471.
(37) Hu, Y.; Houk, K. N.; Kikuchi, K.; Hotta, K.; Hilvert, D. *J. Am. Chem. Soc.* **2004**, *126*, 8197−8205.
(38) Alexandrova, A. N.; Röthlisberger, D.; Baker, D.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2008**, *130*, 15907−15915.
(39) English, B. P.; Min, W.; van Oijen, A. M.; Lee, K. T.; Luo, G.; Sun, H.; Cherayil, B. J.; Kou, S. C.; Xie, X. S. *Nat. Chem. Biol.* **2006**, *2*, 87−94.
(40) Turner, A. J.; Moliner, V.; Williams, I. H. *Phys. Chem. Chem. Phys.* **1999**, *1*, 1323−1331.
(41) Pu, J. Z.; Gao, J. L.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3140−3169.
(42) Kanaan, N.; Ferrer, S.; Martí, S.; Garcia-Viloca, M.; Kohen, A.; Moliner, V. *J. Am. Chem. Soc.* **2011**, *133*, 6692−6702.
(43) Glowacki, D. R.; Harvey, J. N.; Mulholland, A. J. *Nat. Chem.* **2012**, *4*, 169−176.
(44) Garcia-Viloca, M.; Poulsen, T. D.; Truhlar, D. G.; Gao, J. *Protein Sci.* **2003**, *13*, 2341−2354.
(45) Pentikäinen, U.; Pentikäinen, O. T.; Mulholland, A. J. *Proteins: Struct., Funct., Genet.* **2008**, *70*, 498−508.
(46) Lodola, A.; Sirirak, J.; Fey, N.; Rivara, S.; Mor, M.; Mulholland, A. J. *J. Chem. Theory Comput.* **2010**, *6*, 2948−2960.
(47) Lodola, A.; Mor, M.; Zurek, J.; Tarzia, G.; Piomelli, D.; Harvey, J. N.; Mulholland, A. J. *Biophys. J.* **2007**, *92*, L20−22.
(48) Gandour, R. D. *Bioorg. Chem.* **1981**, *10*, 169−176.
(49) Frushicheva, M. P.; Cao, J.; Chu, Z. T.; Warshel, A. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16869−16874.
(50) Frushicheva, M. P.; Cao, J.; Warshel, A. *Biochemistry* **2011**, *50*, 3849−3858.
(51) Frushicheva, M. P.; Mills, M. J. L.; Schopfs, P.; Singh, M. K.; Prasad, R. B.; Warshel, A. *Curr. Opin. Chem. Biol.* **2014**, *21*, 56−62.
(52) Labas, A.; Szabo, E.; Mones, L.; Fuxreiter, M. *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834*, 908−917.
(53) Kurplus, M.; McCammon, J. A. *Annu. Rev. Biochem.* **1983**, *52*, 263−300.
(54) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186−195.
(55) Kamerlin, S. C. L.; Warshel, A. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1339−1375.
(56) García-Meseguer, R.; Martí, S.; Ruiz-Pernía, J. J.; Moliner, V.; Tuñón, I. *Nat. Chem.* **2013**, *5*, 566−571.

DOI: 10.1021/cs501904w
ACS Catal. 2015, 5, 2587−2595