WILEY-VCH

FULL PAPER

Revealing the origin of the efficiency of the de novo designed Kemp eliminase HG-3.17 by comparison with the former developed HG-3.

Katarzyna Świderek,[a],[b] Iñaki Tuñón, [c],* Vicent Moliner,[a],[d],* and Joan Bertran[e]

Abstract: The design of new biocatalysts is a goal in biotechnology to improve the rate, selectivity and environmental impact of industrial chemical processes. In this regard, the use of computational techniques has provided valuable assistance in the design of new enzymes with remarkable catalytic activity. In this paper, hybrid QM/MM Molecular Dynamics simulations have allowed getting an insight into the origin of the limited efficiency of a computationally designed enzyme for the Kemp elimination; the HG-3. Comparison of results derived from this enzyme with those of a more evolved protein containing additional point mutations, HG-3.17, rendered important information to take into account in order to design new enzymes. For this Kemp eliminase reaction, larger reactivity has been demonstrated to be related with a better electrostatic preorganization of the environment that creates a more favourable electrostatic potential for the reaction to proceed. The limitations of HG-3 can be related with a lack of flexibility, a not well fitted active site, and a lack of protein electrostatic preorganization that decrease the reorganization around the oxyanion hole.

Introduction

Enzymes are biological catalysts that speed up chemical reactions making them compatible with life. Apart from their high catalytic efficiency, these catalysts show other advantages such as their chemo-, regio- and stereoselectivity, the ability to work under mild conditions of temperature and pressure, or the fewer unwanted side-products generated after the catalytic process. These features justify the increasing interest in understanding and applying biocatalysts in producing high value chemicals from the environmental point of view. Additionally, the use of biological catalysis reduces the economical costs since the number of required reaction steps can be reduced to a one-pot procedure, allowing the decrease of energy consuming steps such as separation and purification of intermediates. However, the number of enzymes that can be used for industrial applications is limited, what justify the efforts in designing new enzymes.

During the last decades, different protein scaffolds have been used as starting point to design new enzymes, including immune-globulins to produce catalytic antibodies (CAs), promiscuous proteins showing catalytic properties for a secondary reaction that are redesigned to improve this activity, or proteins without any known specific catalytic property. This last strategy is usually named as de novo design of new enzymes.

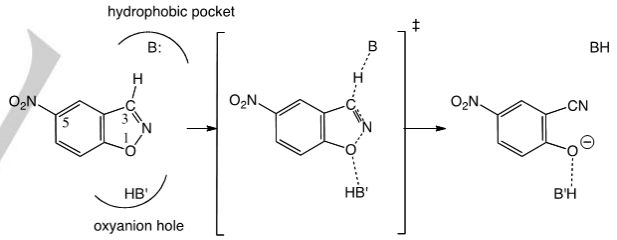


Figure 1. Schematic representation of the base-catalyzed Kemp elimination of 5-nitrobenzisoxazole.

The Kemp elimination (KE) reaction, that consists in the conversion of benzisoxazoles into salicylonitriles (see Fig. 1), is an interesting reaction in the field of enzyme design due to the fact that it implies a proton transfer from a carbon atom to a heteroatom.[1] In addition, since no naturally occurring enzymes have been identified to catalyse this reaction, it has been used as a benchmark of different protocols to design new enzymes.[2] The first de novo design of a Kemp eliminase (KE) was published in 2008 by Houk, Tawfik, Baker and co-workers.[3] In this work, that combined experimental and computational efforts, different proteins, selected by means of the Rosetta software,[4],[5] were used as starting scaffolds. Two different types of active site were explored: one employed a carboxylate group for deprotonation, while in the other a histidine side chain backed up by an aspartate or glutamate acts as a base. In addition, a hydrogen-bond donor was included in the active site in order to stabilize the developing

[a] Dr K. Świderek and Prof. V. Moliner
Departament de Química Física i Analítica; Universitat Jaume I, 12071 Castellón (Spain)
E-mail: moliner@uji.es
[b] Dr K. Świderek
Institute of Applied Radiation Chemistry, Lodz University of Technology, 90-924 Lodz, Poland.
[c] Prof. I. Tuñón
Departament de Química Física, Universitat de València, 46100 Burjassot, (Spain)
E-mail: tunon@uv.es
[d] Prof. V. Moliner
Department of Chemistry, University of Bath, Bath BA2 7AY, (United Kingdom)
[e] Prof. J. Bertran
Departament de Química; Universitat Autònoma de Barcelona, 08193 Bellaterra, (Spain)

Supporting information for this article is given via a link at the end of the document.

negative charge on the phenolic oxygen atom (see Fig. 1). Fifty-nine designs were characterized and eight of them presented activity as a KE. Some of these catalysts were further improved using different strategies. For instance, the design name as KE07 was used as a template for the incorporation of up to 8 mutations by directed evolution providing a rate acceleration of two orders of magnitude,[3] but with a significant decrease in its stability.[6] KE70 was used as template to incorporate computationally guided optimizations combined with further directed evolution. This approach rendered a 400-fold increase in the KE activity.[7] The catalyst with the highest activity, KE59, presented the inconvenience of being the least stable. In this case the strategy for the optimization of the activity included a combination of fold-stabilizing mutations and directed evolution cycles, resulting in more than 2,000-fold increase of the catalytic efficiency.[8]

Jorgensen and co-workers studied the reaction mechanism of four of the designed KE catalysts using hybrid QM/MM potentials based on Monte Carlo (MC) simulations.[9] According to their results, the catalytic mechanisms in KE07, KE10(V131N) and KE15 were found to be concerted with the proton transfer more advanced at the TS than the breaking of the isoxazolyl N–O bond. Instead, a step-wise mechanism where the proton transfer preceded the N-O bond breaking was found for KE16. However, the agreement between experimental and computed free energy barriers was not excellent,[9] which could be due to the approximations considered in the calculations: the protein backbone was considered fixed, the QM region was described at semiempirical level and the proton donor-acceptor distance was kept constant in the simulations. Houk and coworkers used density functional theory (DFT) methods to compute the activation energies of several designed KEs[10] in qualitative agreement with experimental trends. Warshel and coworkers[11],[12],[13] employed an Empirical Valence Bond (EVB) approach, to study several KEs (catalytic antibodies, albumines and de novo designs). In their analysis these authors stressed the role of enzyme preorganization (which is not simply a local property of the active site) and the difficulties of designing an environment adapted for the changes in the electronic distribution observed during the formation of the TS from the reactants. A well-preorganized environment requires smaller changes to reach the TS of the reaction, minimizing the energy cost associated to this reorganization, an aspect that was also pointed out by Labas et al.[14]

After the first *de novo* protein design by Houk, Tawfik, Baker and co-workers,[3] an iterative approach starting from an inactive protein, HG-1, was also used to convert the xylan binding pocket of a thermostable xylanase into a KE.[15] The study of the first generation of the new KE by means of X-ray and molecular dynamics (MD) simulations allowed identifying two possible limitations: the active site was an overly exposed to the solvent and, in addition, it appeared to be too flexible.[15] After a new design where the active site was deeper into the interior of the protein, HG-2, an additional mutation rendered the resulting HG-3 design, that catalyzes the KE of 5-nitrobenzisoxazole with a rate constant, kcat = 0.68 s^{-1}.[15] The new designed protein HG-3, exhibited activity comparable to the best Rosetta designs,[3],[6],[7],[8] but still far from the efficiency of natural enzymes. Starting from this computational designed catalyst, a highly active KE was later re-design by Hilvert and co-workers by means of an evolutionary strategy that included both global and local mutagenesis.[16] The most active variant, HG-3.17, emerged after 17 rounds of mutagenesis and screening (see Fig. 2), catalyzed the KE of 5-nitrobenzisoxazole with a rate constant, kcat = 700 ± 60 s^{-1},[16] an efficiency close to the natural enzymes.

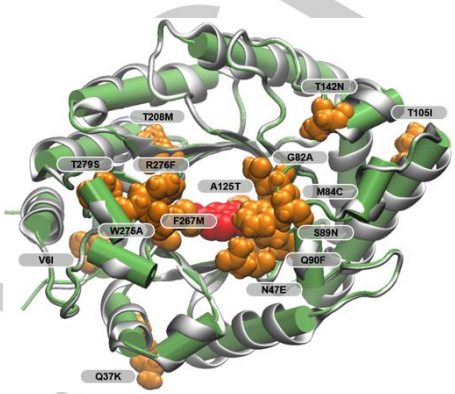


Figure 2. Overlay of HG-3 (HG-2 S265T, in silver) and HG-3.17 E47N,N300D (in cyan), both complexed with Transition State Analog 6-Nitrobenzotriazole. Mutations from HG-3 to HG-3.17 E47N,N300D are shown as blue balls while the inhibitor is displayed as red balls.

Based on the analysis of X-ray structures, three factors were suggested to be decisive in the improvement of the KE from HG3 to HG-3.17: 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).[16],[17] A deep analysis of the base-catalyzed KE of 5-nitrobenzisoxazole catalyzed by HG-3.17 has been recently studied in our laboratory by computational techniques.[18] Our theoretical study indicated that the high reactivity of HG-3.17 was related with a proper electrostatic preorganization of the environment that creates a favourable electrostatic potential for the reaction to proceed. It was concluded that the role played by the electrostatic properties of the oxy-anion hole present in the active site of some of the conformations of the new enzyme appears to be crucial.

In agreement with the opinion of Hilvert, Houk, Mayo and co-workers that proposed that valuable information can be lost if only successful protein designs are reported,[15] the present study is focused on revealing the origin of the limited efficiency of HG-3 KE,[15] by comparison with the further developed HG-3.17.[16] A deep comparative analysis of the reaction catalyzed by both proteins, based on molecular dynamics (MD) simulations performed with hybrid quantum mechanics/molecular mechanics (QM/MM) potentials, will allow obtaining conclusions highly valuable in the future design of new more efficient enzymes.

Results and Discussion

As explained in Computational Details section, our calculations on the KE reaction catalyzed by the HG-3 design started with a 2 ns QM/MM MD simulations to reach an equilibrated situation. Snapshots of the initial and final structures of the active site are presented in Fig. 3. Interestingly, two different conformations of the active site with significantly different reactivity were distinguished in our previous study.[18] The final structure after 2 ns of QM/MM MD simulations of monomer A and B of HG-3.17, as obtained in ref 18, are also presented in Figure 3. The comparative analysis the initial structure of HG-3 (panel A of Figure 3) and the structure obtained after the MD simulations (panel B of Figure 3) shows that X-ray structures do not exactly correspond to the representative structures of the Michaelis complex in solution after the equilibration MD simulations. This result is not unexpected keeping in mind that the inhibitor and the real substrate present different electronic properties and due to the fact that the ensemble of conformations explored by the protein in solution can be different to the solid state structures of the protein-inhibitor complex. This result is in agreement with the previous observations in the HG-3.17 theoretical study carried out by means of a similar computational protocol (see panel C and D of Figure 3). Nevertheless, as observed in Fig. 3, the active site does not experience important conformational changes during the simulations in the case of HG-3. In contrast to HG-3.17, where significant conformational changes took place during the MD simulations in the active site,[18] it appears that HG-3 presents a more rigid protein scaffold and the X-ray structure is quite representative of the reactants state.

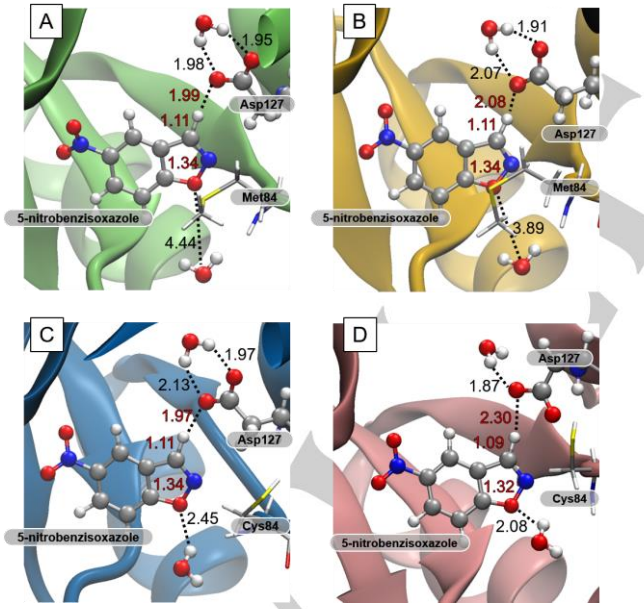


Figure 3. A) Initial structure of HG-3, equivalent to the X-ray structure, despite different substrates are placed in the active site, and B) final structure of the 2 ns QM/MM MD simulations. Panel C and D show the final structure after 2 ns of QM/MM MD simulations of monomer A and B of HG-3.17, respectively, as obtained in ref 18. Distances are in Å.

The distance between the C3 atom of the substrate and the carboxylate oxygen atom of Asp127 is almost maintained during the MD simulations (the distance is kept at 3.10 Å, after a very slight rotation of the substrate in the active site). Interestingly, this distance was enlarged by more than 0.5 Å from the X-ray structure in HG-3.17 after an equivalent QM/MM MD simulation.[18] Thus, despite the nitrogen atom at position 3 in the inhibitor was replaced by a carbon atom in the substrate, with the subsequent decreasing in the acidity of the hydrogen atom attached to this position 3 of the substrate, no significant changes are observed in the protein backbone around the hydrophobic pocket of the active site. As shown below, this distance will obviously be reduced at the TS (ca. 2.7 Å) in order to facilitate the proton transfer. Regarding the oxyanion hole, a larger mobility has been detected during this preliminary QM/MM MD simulation. Nevertheless, once again, the changes are not so dramatic as those detected in the HG-3.17.[18] As observed in Fig. 3, the water molecule located in the active site of HG-3 slightly approaches to the O1 atom of the substrate from 4.44 to 3.89 Å, with no other noticeable changes. It is important to point out that, due to the flexibility of the proteins, revealed when performing QM/MM MD simulations, just two decimal places are reported in the interatomic distances.

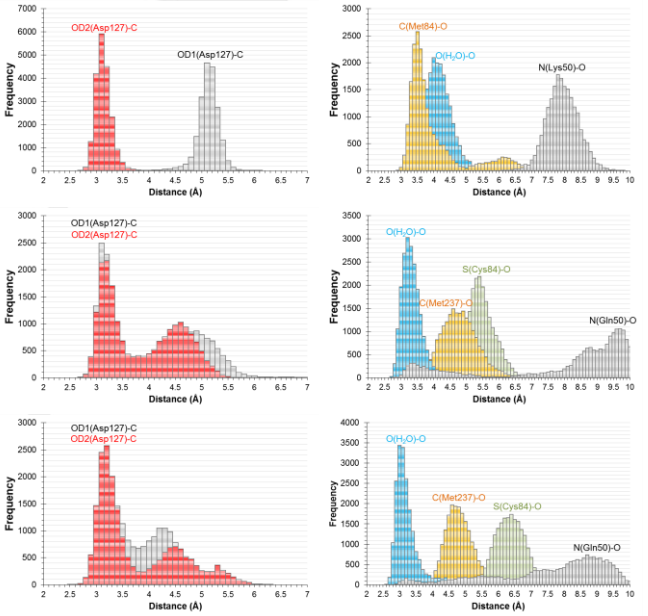


Figure 4. Distribution of distances between key atoms of the substrate and the relevant residues of the active site during the 2 ns QM/MM MD simulations on the HG-3 (top panels) and two conformers of HG-3.17 KE (center and bottom panels) designs.

The analysis of the time evolution of the geometry of the substrate in the active site of HG-3 has been complemented by the study of the H-bond interactions between the substrate and the residues of the active site during the 2 ns QM/MM MD simulation, and compared with the equivalent simulations on two conformers of the HG-3.17 KE. As deduced from the left panels of Fig. 4, while the Asp127 interact with the C-H of the substrate through any of the two carboxylate oxygen atoms in both of the HG-3.17 conformers, this is not the case in HG-3. The carboxylate group of Asp127 does not rotate in HG-3, and only one of the oxygen atoms will be capable of acting as the base for the reaction. Rotation of carboxylate group of Asp127 in HG-3.17 leads to a different rotameric state of this residue that was stabilized by means of a new hydrogen bond with the thiol moiety of Cys84. The equivalent residue in position 84 in HG-3 is a

methionine with a non-polar character that obviously cannot stabilize the carboxylate group of Asp127 as the polar cysteine in HG-3.17 does. It is interesting to point out that the new conformation observed in the active site of HG-3.17 after Asp127 rotation catalyzed the Kemp elimination reaction better than the original conformation.

The role of the oxyanion hole detected in the leaving group can be analyzed from the right panels of Fig. 4. Thus, while the active sites of both HG-3.17 KEs accommodate a water molecule that would stabilize the negative charge developed in the oxygen leaving group during catalysis, this is not the case in the HG-3. The bridge of hydrogen bond interactions between the O1 atom of the substrate and the oxygen atom of carbonyl group of Gln50 through two water molecules in HG-3.17,[18] capable of donating a hydrogen bond to the leaving group during catalysis, is not present in the HG-3. As observed in Fig. 4, the active site water molecule appears at a longer distance from the oxygen leaving group and a weak interaction is established between the leaving group and the methyl hydrogen atoms of Met84. In fact, it appears that the Met84 prevents the water molecule to do the role found in HG-3.17. Cys84 is not so close to O1 in HG-3.17 KEs, thus allowing waters to go inside the cavity. Gln50 does not appear to establish a direct interaction with the leaving group neither.

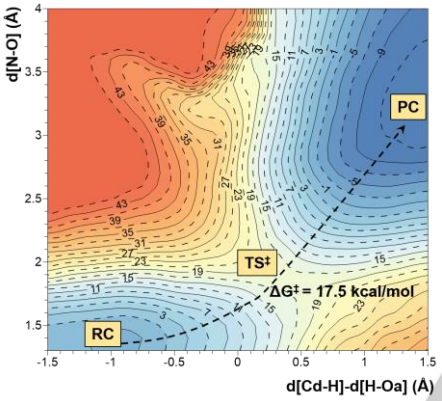


Figure 5. Free energy surface for the KE of 5-nitrobenzisoxazole catalized by HG-3 obtained at AM1/MM level with spline corrections at M06-2X/6-31+G(d,p)/MM level. Energies are given in kcal·mol^{-1}.

The last structure of the QM/MM MD simulation was used to study the reactivity of the KE catalyzed by HG-3. 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 obtained the free energy surface (see Fig. 5). The first conclusion that can be derived from the free energy surface presented in Fig. 5 is that, as in the case of HG-3.17,[18] the reaction path proceeds by a concerted but asynchronous mechanisms in HG-3. The activation free energy deduced from the surface is 17.5 kcal·mol^{-1}. This value is in very good agreement with the activation free energy derived from the experimental rate constants 0.68 ± 0.04 s^{-1} and 3.0 ± 0.1 s^{-1},[15] that would correspond to 17.7 – 16.8 kcal·mol^{-1}, applying the Transition State Theory. It is important to point out that the statistical error associated to activation free energies derived from the PMFs computed within the present methodology is considered to be below 1 kcal·mol^{-1}. Inclusion of contributions derived from tunnelling effects and dynamical recrossing by means of the Variational Transition State Theory did not change the free energies of activations obtained for the HG-3.17 in our previous study since the effect of the computed coefficients cancelled each other.[18]

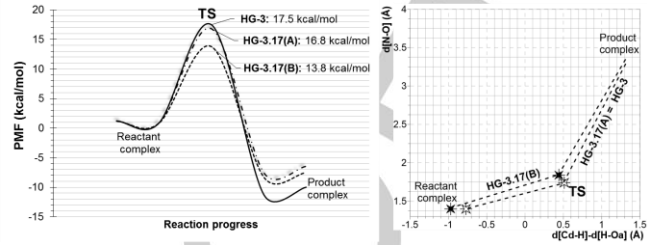


Figure 6. Representation of the free energy profiles obtained for the KE of 5-nitrobenzisoxazole catalyzed by HG-3 and by HG-3.17 conformer A and B, obtained at the AM1/MM level with spline corrections at M06-2X/6-31+G(d,p)/MM level. Results of the two conformers of HG-3.17 are taken from ref 18. Right panel shows the different position of the TSs on the free energy surfaces.

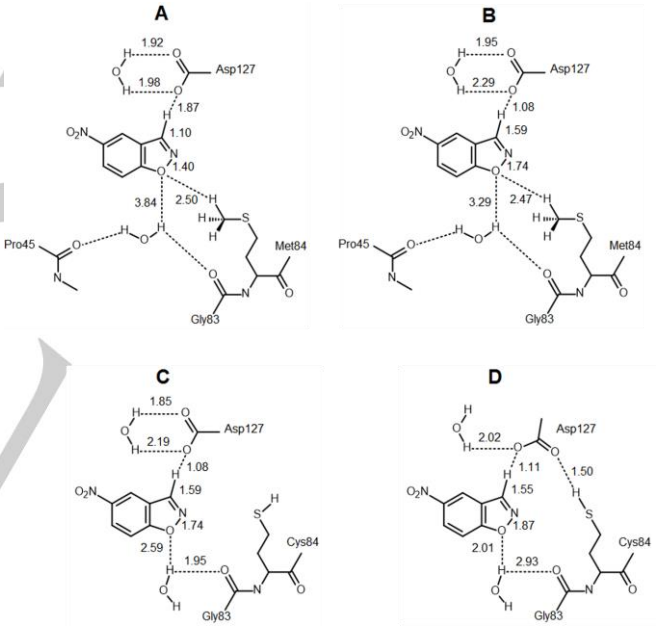


Figure 7. Schematic representation of reactant complex (panel A), and TS (panel B) in HG-3 obtained at M06-2X/MM level. Representation of TSs obtained in two different conformers of HG-3.17, conformer A (panel C) and conformer B (panel D), are taken from ref 18. Distances are reported in Å.

Deeper insights into the KE catalyzed by HG-3 can be based in the comparative geometrical analysis with the two different conformers of HG-3.17 that rendered free energies of activations for the conformers labelled as A and B of 16.3 and 13.8 kcal·mol^{-1}, respectively (see Fig. 6).[18] The structure corresponding to the TS and the reactant complex (Michaelis complex, MC) of HG-3, optimized at M06-2X/MM level, are shown in Fig. 7. As observed, the TS is described by a proton transfer in a very advanced stage of the reaction, while the isoxazolyl N-O bond distance can be considered in between the formed and completely broken bond. Thus, it appears that the TS of the reaction is mainly governed by the N-O breaking bond with a

proton practically transferred to the Asp127 base. These values are closer to those found in the TS of the less reactive conformation than those of the more reactive conformation of HG-3.17 (key inter-atomic distances of the TSs obtained in our previous study have been added into Fig. 7). The geometrical analysis of the proton transfer and the N-O breaking bond indicates that more asynchronous processes can be related with higher energy barriers. In all, according to these results, the improvement in the design of the oxyanion hole seems to be the driving force explaining the reactivity of the KE designs. The mutations introduced from HG-3 to HG-3.17 seems to alter the conformational landscape of the protein and the geometrical analysis shows how the reaction in the active site of the less reactive conformer of the evolved enzyme HG-3.17 proceeds through TS structures similar to the original design HG-3 while the most reactive conformations render more synchronous reaction paths.

Table 1. ESP charges (in a.u.) of key atoms computed in the active site of the protein at the M06-2X/6-31+g(d,p)/MM level using the ChelpG method for reactant complex (RC) and transition state (TS) in HG-3 and in HG-3.17 with two different conformations of the active site: A and B. Data of HG-3.17 obtained from ref. 18.

	HG-3.17				HG-3	
	conformer A		conformer B			
	RC	TS	RC	TS	RC	TS
OD1(D127)	-0.896	-0.775	-0.904	-0.753	-0.920	-0.776
OD2(D127)[a]	-0.899	-0.774	-1.051	-0.882	-0.978	-0.728
H	0.183	0.338	0.094	0.262	0.232	0.234
N2	-0.348	-0.298	-0.423	-0.331	-0.366	-0.335
O1	-0.222	-0.440	-0.199	-0.506	-0.169	-0.392
C3	0.189	-0.075	0.313	0.097	0.167	0.070

[a] Oxygen atom interacting with the C-H group of substrate.

Table 2. Averaged electrostatic potentials ($kJ \cdot mol^{-1} \cdot e^{-1}$) created by the environment on the key atoms of the reaction at the reactant complex (RC) and transition states (TS) and its difference (Δ), computed in conformer A and conformer B of HG-3.17 (data from ref. 18) and in HG-3.

HG-3.17		conformer A	
	RC	TS	Δ(TS-RC)
OD1(D127)	345.8 ± 24.7	310.0 ± 21.7	-35.8
OD2(D127)[a]	381.2 ± 22.9	357.7 ± 22.5	-23.5
C3	200.1 ± 20.7	227.8 ± 20.1	27.7
N2	199.8 ± 21.2	243.7 ± 22.3	43.9
O1	141.1 ± 26.1	213.2 ± 30.9	90.1
HG-3.17		conformer B	
	RC	TS	Δ(TS-RC)
OD1(D127)	501.4 ± 29.5	431.4 ± 30.8	-70.0
OD2(D127)[a]	585.8 ± 32.6	424.8 ± 52.4	-161.0
C3	256.9 ± 22.3	321.1 ± 29.3	64.2
N2	249.0 ± 27.0	330.8 ± 25.5	81.8
O1	203.8 ± 32.5	298.3 ± 29.6	94.5
HG-3			
	RC	TS	Δ(TS-RC)
OD1(D127)	346.4 ± 21.7	304.9 ± 24.6	-41.5
OD2(D127)[a]	291.2 ± 20.7	254.4 ± 22.2	-36.9
C3	160.7 ± 16.4	171.9 ± 17.6	11.2
N2	156.7 ± 18.8	173.6 ± 18.1	16.8
O1	124.9 ± 19.2	149.9 ± 18.6	25.0

[a] Oxygen atom interacting with the C-H group of substrate.

The analysis based on the geometries of RC and TS can be supplemented with the analysis of electrostatic charges of the key atoms involved in the reaction, as well as the electrostatic potential created by the protein on these atoms. From the electrostatic point of view, the reaction can be described as a charge transfer from Asp127, negatively charged at the RS, to the phenoxide oxygen-leaving atom in the products state. Then, the process can be favoured if the electrostatic potential is increased around the leaving group and decreased around the base. Table 1 shows the electrostatic charges on the key atoms involved in the reaction computed under the effect of the point charges of the protein and water molecules. The electrostatic potentials created by the MM region (protein and water molecules) on these key atoms of the reacting system are reported in Table 2. Standard deviations are not reported on charges of Table 1 since values are derived from a high level M06-2X/MM optimization. The values obtained for the two conformers of HG-3.17 obtained in our previous study[18] are also listed in Table 1 and 2. The first conclusion derived from the results on Table 1 is that the charges on the two oxygen atoms of Asp127 (OD1 and OD2), as well as the charge on the phenoxide oxygen-leaving atom (O1) are closer to the values obtained for the less reactive conformer A than to those of the more reactive conformer B of HG-3.17. Thus, the charge on the phenoxide oxygen-leaving atom of HG-3 in the TS (-0.392 a.u.) is lower than the value obtained for the less reactive A conformation of HG-3.17 (-0.440 a.u.) and significantly lower than the value presented in the more reactive B conformer of HG-3.17 (-0.506 a.u.). These results suggest a less electrostatically advanced stage of the reaction of the TS in HG-3 than in the HG-3.17 conformers. Nevertheless, on the contrary, when focusing on the base Asp127, the charge on O2 atom is lower in the HG-3 (-0.728 a.u.) than in conformers A and B of HG-3.17 (-0.774 and -0.882 a.u., respectively), revealing a more advanced stage of the reaction in the former. This result is in agreement with the description of the reaction based on geometrical analysis. As pointed out above, interestingly the free energy barrier of HG-3 was close to the value obtained for conformer A of HG-3.17, and significantly higher than the one obtained for conformer B. The evolution of the interatomic distances and charges on the atoms involved in the breaking and forming bonds can be also related with the electrostatic potential created by the protein in the active site. A trend is also detected when comparing the electrostatic potential on key atoms in the HG-3 and in the two conformers of HG-3.17. According to the values of Table 2, the electrostatic potential on OD2(Asp127) in the TS is higher in conformer B (424.8 ± 52.4 $kJ \cdot mol^{-1} \cdot e^{-1}$) than in conformer A ($357.7 \pm 22.5$ $kJ \cdot mol^{-1} \cdot e^{-1}$) of HG-3.17, and in HG-3 (254.4 ± 22.2 $kJ \cdot mol^{-1} \cdot e^{-1}$). This result means that the protein in the most reactive conformation of HG-3.17 stabilize the basic form of Asp127, disfavouring the proton transfer, while the less active HG-3, that presents the lowest value, would not stabilize the reactants complex state, being the best protein in favouring the proton transfer. Then, and in agreement with previous findings, it appears that the origin of catalysis is not so much about favouring the hydrogen transfer from the substrate to the base as about the O-N breaking bond. As explained above, the reaction coordinate at the TS is dominated by the N-O bond breaking (see Fig. 5). Indeed, the largest electrostatic potential on the O1 atom of the substrate is in the most reactive conformer B of HG-3.17 (298.3 ± 29.6 a.u.) that would favour the accumulation of negative charge on oxygen-leaving O atom and the cleavage of the N-O bond. On the contrary, HG-3 is the protein that shows the lowest value of the electrostatic potential on this O atom (149.9 ± 18.6 $kJ \cdot mol^{-1} \cdot e^{-1}$). Conformer A of HG-3.17 shows a value in between these two extremes, 213.2 ± 30.9 $kJ \cdot mol^{-1} \cdot e^{-1}$, in agreement with the trend on reactivity. Finally, according to the changes in the electrostatic

potentials generated by the different protein scaffolds on the key atoms, from reactants to TS, the smaller differences found on HG-3 supports the conclusion deduced from the analysis of the geometries that indicated a more rigid protein (see the Δ(TS-RC) columns on Table 2). It appears that the HG-3 suffers smaller changes from reactant to TS than any of the conformers of HG-3.17. A priori, this would be associated with a smaller energy of reorganization but, since the TS is not so well prepared to stabilize the TS, the resulting effect is a less efficient catalyst. The change in the electrostatic potential from RC to TS in conformer A and B of HG-3.17 are similar but, as mentioned, the protein conformation on the most reactive conformer B provides the most favourable electrostatic potential on the O1 atom, which is accumulating the largest negative charge at the TS (-0.501 a.u.), by comparison with the conformer A of HG-3.17 (-0.440) and the HG-3 (-0.392). This result further confirms the agreement between the electrostatic role of the protein in polarizing the substrate and the activation free energies.

Conclusions

We herein present a theoretical study of the Kemp eliminase reaction catalyzed by HG-3, a KE design by Hilvert, Houk, Mayo and co-workers.[15] Using a hybrid QM/MM scheme, the AM1 semiempirical Hamiltonian has been initially used to describe the QM region, and further energy corrections have been carried out based on the M06-2X DFT method. The results are compared with the those derived from our previous study[18] on a variant of the recently KE designed by Hilvert and co-workers, the HG-3.17, that presents a high catalytic activity, close to the efficiency of natural enzymes.[16] The free energy surface corresponding to the KE catalyzed by HG-3 renders an activation free energy (17.5 kcal·mol^{-1}) very close to the values derived from the experimental rate constants according to the TST (17.7 - 16.8 kcal·mol^{-1}),[15] and significantly higher than the most active conformation of HG-3.17 (13.8 kcal·mol^{-1}).[18] The comparison of the interatomic distances defining the proton transfer and the N-O breaking bond in the TS located in HG-3 with those obtained in the TSs of the two conformations of the HG-3.17 suggests that the synchronicity of the processes can be related with a lower energy barrier. In addition, a detailed analysis of the active site of the two designed KEs suggests that the catalytic efficiency is due to a combination of short and long distance interactions. Thus, the active site Met84 residue of HG-3, that was later mutated to a cysteine in HG-3.17, does not contribute to develop an optimum oxyanion hole of the active site to stabilize the negatively charge developed on the phenoxide oxygen-leaving atom in the TS. Our previous study indicated the key role played by the electrostatic properties of the oxy-anion hole present in the active site by means of a bridge of hydrogen bonds through well-placed water molecules. The original Met84 residue does not allow a proper orientation of water molecules to interact with the leaving group and, in fact, prevents the approach of water molecules to the O1 atom. Nevertheless, it appears that the improvements of long-range interactions from HG-3 to HG-3.17 are also responsible of the increase of the catalytic efficiency. Thus, when computing a more global magnitude as the electrostatic potential created by the protein on key atoms involved in the reaction, the result suggests that the higher the potential in the oxyanion hole, the lower the activation barrier. Keeping in mind that the oxygen leaving atom presents a larger negative charge on the TS that on the reactant complex (the Michaelis complex), higher (more positive) potentials on the leaving atom stabilize the TS and, consequently, favour the reaction to proceed. This conclusion is supported not only by the present calculations on HG-3 but by those on the different conformers of HG-3.17 previously studied in our group.[18] Thus, it appears that the trend in the free energies of activation are well correlated with the magnitudes of the electrostatic potential on the key atoms involved in the reaction that, in turn, polarizes the substrate increasing the negative charge on the O1 atom at the TS.

The existence of two different stable conformations found in the case of HG-3.17 with significant different reactivity indicates that the protein sequence itself does not completely determines the structure of the protein. Moreover, the KE design studied in this work, HG-3, presents electrostatic potentials, TS geometry and an activation free energy similar to those of the least reactive conformer of HG-3.17, indicating that, on the other side, different sequences can render similar protein structures. The sequence determines the conformational landscape of the protein and in this case the evolution from HG-3 to HG-3.17 seems to have favoured those conformations presenting higher reactivity. The relation sequence-structure is not biunivocal, as illustrated by the comparison between HG-3.17 conformers and HG-3, and the complexity of the protein conformational landscape should be considered for the optimization of new enzyme designs.

Thus, the comparative analysis of the Kemp elimination reaction catalyzed by different designed enzymes allows concluding that design optimization must be focused in finding a sequence providing the optimum conformational landscape for the protein, favouring those conformations with a larger preorganization for the reaction to be catalyzed. In this sense, it will be necessary to consider not only a high shape complementarity between the protein and the substrate but also the electrostatic complementarity between the active site and the reaction process. It appears that for this particular reaction, efforts should be focused on the oxyanion hole since, in the three studied cases, the improvements on the base that accepts the hydrogen does not entail a reduction in the activation free energy because the reaction TS is determined by the N-O bond breaking process. Our QM/MM MD simulations also suggest that the 17 mutations introduced from HG-3 to HG-3.17 contribute to an increase of the flexibility, which stress the idea of efficient catalysts as not completely rigid scaffolds.[19] The design of new biocatalysts must be focused on designing an adequate preorganized protein scaffold capable of electrostatically stabilize the TS of the desired reaction, with a certain degree of flexibility to accommodate the reactants-like structure of the substrate. Computational simulations can be used to guide the rational design of new enzymes based on the study of not only the successful designs but of failed designs, as suggested by Hilvert, Houk, Mayo and co-workers.[15]

Computational Details

The X-ray structure of Kemp Eliminase HG-2 complexed with the transition state analogue 5-Nitrobenzotriazole (PDB ID 3NYD)[15] was used as a starting point in our simulations. The coordinates of hydrogen atoms were added with fDYNAMO[20] after estimating the pKa of the titratable residues by means of PROPKA semiempirical program,[21],[22] to determine the proper protonation state at a pH equivalent to 7. The HG-3 structure was generated performing a S265T mutation on the HG-2, as described in the original paper.[15] In the PDB structure 3NYD the protein crystallizes as a dimer, but the active biological assemble is a monomer. The monomers in this structure present small structural differences, including the number of crystal waters, and consequently only one monomer was employed in the present study. The systems was placed in a pre-equilibrated box of water molecules with size of 100 × 80 × 80 Å^3 and neutralized by adding 4 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 nitrogen atoms at positions 1 and 3 (see Fig. 1) to oxygen and carbon atoms, respectively. 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[23] during Molecular Dynamics (MD) simulations and by the Density Functional method M06-2X[24] (with the standard 6-31+G** basis set) during the exploration of the Potential Energy Surface associated to the KE reaction. The rest of the system (protein, water molecules and counterions) was described using the OPLS-AA[25] and TIP3P[26] force fields, as implemented in the fDYNAMO library.[20] To saturate the valence of the QM/MM frontier atoms, a link atom was placed between the Cα and Cβ atoms of the Asp127.[27] Cut-offs 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 Fig. S1 in Supporting Information), the system can be considered equilibrated after 0.5-0.6 ns of simulation.

Potential of Mean Force (PMF). In order to obtain the free energy landscape, we have traced the Potential of Mean Force (PMF). 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 PMF calculation requires a series of molecular dynamics simulations in which the distinguished reaction coordinates are constrained around particular values with the umbrella sampling procedure.[28] 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).[29]

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. In order to reduce the errors associated with the quantum level of theory employed in our simulations, following the work of Truhlar and co-workers,[30],[31],[32] a new energy function defined in terms of interpolated corrections was used to obtain the PMFs:[33]

$$\boldsymbol{E} = \boldsymbol{E}_{\boldsymbol{LL/MM}} + \boldsymbol{S}[\Delta \boldsymbol{E}_{\boldsymbol{LL}}^{\boldsymbol{HL}}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)] \quad (1)$$

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. These calculations were carried out using the Gaussian09 program.[34]

Acknowledgements

This work was supported by the Spanish Ministerio de Economía y Competitividad (project CTQ2015-66223-C2), Universitat Jaume I (project P1•1B2014-26), Generalitat Valenciana (PROMETEOII/2014/022) and the Polish Ministry of Science and Higher Education ("Iuventus Plus" program project no. 0478/IP3/2015/73, 2015-2016). V.M. is grateful to the University of Bath for the award of a David Parkin Visiting Professorship. 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.

Keywords: enzyme design • Kemp elimination • QM/MM • electrostatic preorganization • protein flexibility

[1] M. L. Casey, D. S. Kemp, K. G. Paul, D. D. Cox, *J. Org. Chem*. **1973**, *38*, 2294–2301.

[2] K. Świderek, I. Tuñón, V. Moliner, J. Bertran, *Arch. Biochem. Biophys*. **2015**, *582*, 68−79.

[3] D. Röthlisberger, O. Khersonsky, A. M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J. L. Gallaher, E. A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K. N. Houk, D. S. Tawfik, D. Baker, *Nature* **2008**, *453*, 190–195.

[4] A. Zanghellini, L. Jiang, A. M. Wollacot, G. Cheng, J. Meiler, E. A. Althoff, D. Röthlisberger, D. Baker, *Protein Sci.* **2006**, *15*, 2785−2794.

[5] F. Richter, A. Leaver-Fay, S. D. Khare, S. Bjelic, D. Baker, *PLoS One* **2011**, *6*, e19230.

[6] O. Khersonsky, D. Röthlisberger, O. Dym, S. Albeck, C. J. Jackson, D. Baker, D. S. Tawfik, *J. Mol. Biol*. **2010**, *396*, 1025−1042.

[7] O. Khersonsky, D. Röthlisberger, A.M. Wollacott, P. Murphy, O. Dym, S. Albeck, G. Kiss, K.N. Houk, D. Baker, D.S. Tawfik, *J. Mol. Biol.* **2011**, *407*, 391–412.

[8] O. Khersonsky, G. Kiss, D. Röthlisberger, O. Dym, S. Albeck, K.N. Houk, D. Baker, D.S. Tawfik, *Proc. Nat. Acad. Sci. U.S.A.* **2012**, *109*, 10358–10363.

[9] A.N. Alexandrova, D. Röthlisberger, D. Baker, W.L. Jorgensen, *J. Am. Chem. Soc.* **2008**, *130*, 15907–15915.

[10] G. Kiss, D. Röthlisberger, D. Baker, K.N. Houk, *Protein Sci.* **2010**, *19*, 1760–1773.

[11] M.P. Frushicheva, M.J.L. Mills, P. Schopf, M.K. Singh, R.B. Prasad, A. Warshel, *Curr. Opin. Chem. Biol.* **2014**, *21*, 56–62.

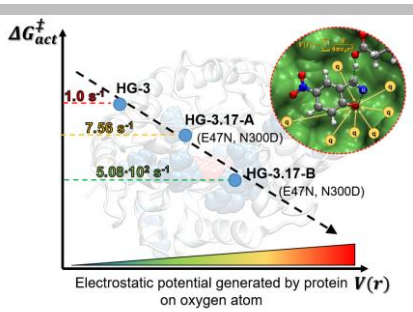
[12] M.P. Frushicheva, J. Cao, Z.T. Chu, A. Warshel, *Proc. Nat. Acad. Sci. U.S.A.* **2010**, *107*, 16869–16874.

[13] M.P. Frushicheva, J. Cao, A. Warshel, *Biochemistry* **2011**, *50*, 3849–3858.
[14] A. Labas, E. Szabo, L. Mones, M. Fuxreiter, *Biochim. Biophys. Acta*. **2013**, *1834*, 908–917.
[15] H. K. Privett, G. Kiss, T. M. Lee, R. Blomberg, R. A. Chica, L. M. Thomas, D. Hilvert, K. N. Houk, S. L. Mayo, *Proc. Nat. Acad. Sci. USA.* **2012**, *109*, 3790-3795.
[16] R. Blomberg, H. Kries, D. M. Pinkas, P. R. E. Mittl, M.G. Grütter, H. K. Privett, S.L. Mayo, D. Hilvert, *Nature* **2013**, *503*, 418-421.
[17] M. Höhne, U. T. Bornscheuer, *Angew. Chem. Int. Ed*. **2014**, *53*, 1200–1202.
[18] K. Świderek, I. Tuñón, V. Moliner, J. Bertran, *ACS Catal*. **2015**, *5*, 2587−2595.
[19] R. García-Meseguer, S. Martí, J. J. Ruiz-Pernía, V, Moliner, I. Tuñón, *Nat. Chem.* **2013**, *5*, 566-571.
[20] M. J. Field, M. Albe, C. Bret, F. Proust-De Martin, A. J. Thomas, *Comput. Chem.* **2000**, *21*, 1088−1100.
[21] H. Li, A. D. Robertson, J. H. Jensen, *Proteins* **2005**, *61*, 704−721.
[22] D. C. Bas, D. M. Rogers, J. H. Jensen, *Proteins* **2008**, *73*, 765−783.
[23] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, *J. Am. Chem. Soc.* **1985**, *107*, 3902-3909.
[24] Y. Zhao D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215−241.
[25] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225-11236.
[26] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926-935.
[27] M. J. Field, P. A. Bash, M. A. Karplus, *J. Comput. Chem.* **1990**, *11*, 700−733.
[28] G. M. Torrie, J. P. Valleau, *J. Comput. Phys*. **1977**, *23*, 187−199.
[29] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, J. M. Rosenberg, *J. Comput. Chem*. **1992**, *13*, 1011−1021.
[30] J. C. Corchado, E. L. Coitiño, Y. Chuang, P. L. Fast, D. G. Truhlar, *J. Phys. Chem. A* **1998**, *102*, 2424−2438.
[31] K. A.Nguyen, I. Rossi, D. G. Truhlar, *J. Chem. Phys*. **1995**, *103*, 5522−5530.
[32] Y. Y. Chuang, J. C. Corchado, D. G. Truhlar, *J. Phys. Chem. A* **1999**, *103*, 1140−1149.
[33] J. J. Ruiz-Pernía, E. Silla, I. Tuñón, S. Martí, V. Moliner, *J. Phys. Chem. B* **2004**, *108*, 8427−8433.
[34] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.

Entry for the Table of Contents

FULL PAPER

In this paper, hybrid QM/MM MD simulations have allowed revealing the origin of the limited efficiency of a computationally designed enzyme HG-3 by comparison with a more evolved protein containing additional point mutations, HG-3.17.



Katarzyna Świderek, Iñaki Tuñón, Vicent Moliner,* and Joan Bertran.*

Page No. – Page No.
Revealing the origin of the efficiency of the de novo designed Kemp eliminase HG-3.17 by comparison with the former developed HG-3.