Theoretical Study of Primary Reaction of *Pseudozyma antarctica* Lipase B as the Starting Point To Understand Its Promiscuity

Katarzyna Świderek,[†,‡] Sergio Martí,[§] and Vicent Moliner*,[§]

[†]Departament de Química Física, Universitat de València, 46100 Burjassot, 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

Ⓢ Supporting Information

ABSTRACT: *Pseudozyma antarctica* lipase B (PALB) is a serine hydrolase that catalyzes the hydrolysis of carboxylic acid esters in aqueous medium, but it has also shown catalytic activity for a plethora of reactions. This promiscuous activity has found widespread applications. In the present paper, the primary reaction of PALB, its native hydrolytic activity, has been studied using hybrid quantum mechanical/molecular mechanical (QM/MM) potentials. Free energy surfaces, obtained from QM/MM Molecular Dynamics (MD) simulations, show that the reaction takes place by means of a multistep mechanism where the first step, the activation of the carbonyl group of the substrate and the nucleophilic attack of Ser105 to the carbonyl carbon atom, presents the highest energy transition state. Our results, which are in good agreement with kinetic experimental data, suggest that the origin of the catalytic activity of the enzyme is due to favorable interactions established between the residues of the active site that create an oxyanion hole, Gln106 and Thr40, as well as the Asp187 that is capable of modulating the pK_a of His224 to act as a base or an acid depending on the step of the catalytic process. Understanding the main features of PALB in catalyzing the hydrolysis reaction can be the starting point to design mutations that improve the efficiency of its secondary reactions. This is, to design an optimum biocatalyst capable of accelerating fundamental reactions in organic synthesis is the primary aim.

KEYWORDS: hydrolysis, enzyme promiscuity, Pseudozyma antarctica lipase B, Candida antarctica lipase B, kinetic isotope effects, enzyme catalysis
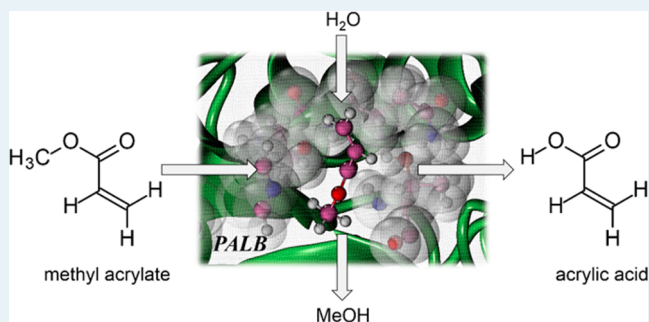
During the past decades, the possibility of modifying naturally existing enzymes has attracted the attention of academia and industry since it can open the door to obtaining new biocatalysts that present high efficiency, selectivity, fewer unwanted side products, the use of a one-pot procedure, or the possibility of working at aqueous mild conditions of temperature or pressure. In this regard, nowadays the use of computer-based methods, in combination with in vitro directed evolution and labor-intensive screening of microbial cultures for the desired activities appears to be a promising tool.[1] Different goals, apart from increasing or generating activity, have been achieved during the past years in the designing of new biocatalysts, such as altered substrate specificity, improved thermostability, organostability, regioselectivity, enantioselectivity, and so forth.[2]

Among all possible methods, apart from generation of the full protein machinery from scratch, or de novo protein design, that has shown outstanding but a limited number of successful challenges,[3−5] the redesign of a protein active site is one of the most promising strategies at present. In this technique, the cavity of an existing natural protein is modified by mutating a subset of its side-chain amino acids. The starting protein scaffold for the rational redesign of an active site can be a promiscuous enzyme,[6−8] an enzyme that presents the ability to catalyze, at lower efficiency, a reaction other than its cognate reaction that is maintained via selective pressure. This provides a raw starting point for the evolution of enzymes, as a new duplicated gene presenting low activity would be the foundational material for adaptive evolution.[6]

Pseudozyma antarctica (formerly *Candida antarctica*) lipase B (PALB) is a serine hydrolase belonging to the group of α/β-hydrolases that catalyze the hydrolysis of carboxylic acid esters in aqueous medium, but it also shows catalytic activity as a carboxylic acid esterase, thioesterase, peptidase, dehalogenase, epoxide hydrolase, halo peroxidase, and having the ability to cleave C−C bonds.[9,10] This promiscuous activity makes PALB a good model for developing new biocatalysts. As a serine hydrolase, it has been proposed that catalysis proceeds by activation of the carbonyl of the substrate ester through an
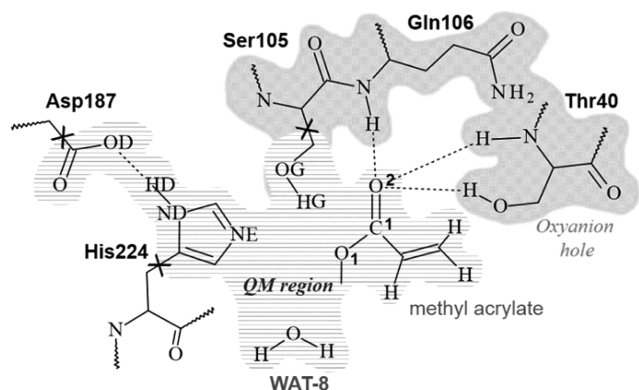
oxyanion hole made by two residues (Thr40 and Gln106), which would stabilize the oxyanion intermediate and the nucleophilic attack of a serine residue (see detail of active site in Scheme 1).[11] This serine would be coupled to another two

Scheme 1. Representation of the Active Site of PALB Together with Bound Methyl Acrylate Substrate[a]

[a]Light grey shaded region contains atoms treated quantum mechanically. Oxyanion hole are highlighted in dark grey. Quantum link atoms are represented as crosses.

conserved residues, thus forming a catalytic triad (Ser105, Asp187, and His224).[12] Nevertheless, despite the amount of studies devoted to the use modified hydrolases to develop new applications, a detailed description of the molecular mechanisms of the primary activity of PALB remains unclear, which hamper a quicker development of more efficient biocatalysts based on this promiscuous enzyme.

The aim of the present study is to get a complete picture of the molecular mechanism of PALB when working as a serine hydrolase by means of molecular dynamics (MD) simulations based on hybrid quantum mechanics/molecular mechanics (QM/MM)[13] potentials. Information derived from the study of the primary reaction will be used to know what residues can be kept and which ones should be mutated to improve its activity on other reactions.

■ THEORETICAL METHODS

System. The starting geometry was obtained from the *Candida antarctica* lipase B (PALB) PDB entry 1TCA,[14] which consists of a monomer. The PROPKA3 semiempirical program[15] was used to recalculate the pK_a values of the titratable protein residues to verify their protonation states at a pH of 7. According to the results, most residues were found at their standard protonation state, except for the Asp134 residue that should be protonated and His224, which was protonated in the δ-position. Additionally, the disulphate bridges between Cys22 and Cys64, Cys216 and Cys258, and Cys293 and Cys311 were defined. Since the total charge of the system was neutral, no counterions were required. After adding the hydrogen atoms to the structure, series of optimization algorithms were applied. In particular, after a series of steepest descent conjugated gradient minimization steps (until a gradient tolerance of 5 kJ·mol^{-1} Å^{-1}), L-BFGS-B was carried out until completely optimized positions of hydrogen atoms were reached (gradient tolerance of 0.1 kJ·mol^{-1} Å^{-1}). To avoid a denaturation of the protein structure, all the heavy atoms of the protein and the inhibitor were restrained by means

of a Cartesian harmonic umbrella with a force constant of 1000 kJ·mol^{-1}·Å^{-2}. Afterward, the system was fully relaxed with the peptidic backbone restrained with a lower constant of 100 kJ·mol^{-1}·Å^{-2}. Finally, the optimized protein was solvated using a orthorhombic box of water molecules which was $100 \times 80 \times 80$ Å. Water molecules with an oxygen atom lying within 2.8 Å of any heavy atom were removed. The remaining water molecules were then relaxed using optimization algorithms. The full system was formed by the 64652 atoms, containing the protein (4625 atoms), the methyl acrylate substrate (12 atoms), and 20 006 solvation water molecules (60 015 atoms).

During the simulations, the atoms of the substrate, together with the side chains of His 224, Asp187 and residue Ser105 were selected to be treated by QM, using the semiempirial AM1 Hamiltonian. The rest of the system (protein plus water molecules) were described using the OPLS-AA and TIP3P force fields, respectively, as implemented in the fDYNAMO library.[16] To saturate the valence of the QM/MM frontier atoms, link atoms between the Cα and Cβ atoms of the three key residues were used (see Scheme 1).

Due to the large amount of degrees of freedom, any residue 20 Å apart from any of the atoms of the substrate was kept frozen in the remaining calculations. Cutoffs for the non-bonding interactions were applied using a switching scheme, within a range radius from 14.5 to 16 Å. The system was equilibrated by means of 500 ps of QM/MM MD at a temperature of 300 K. According to the time-dependent evolution of the RMSD of position of atoms belonging to the protein backbone (see Figure S1 in Supporting Information), the system can be considered equilibrated after 250 ps of simulation (final RMSD equal to 0.35 Å).

The potential energy of our scheme is derived from the standard additive scheme of QM/MM formulation:

$$E_{QM/MM} = \langle \Psi | \hat{H}_o | \Psi \rangle + \left(\sum \left\langle \Psi \left| \frac{q_{MM}}{r_{e,MM}} \right| \Psi \right\rangle \right. $$
$$\left. + \sum \sum \frac{Z_{QM} q_{MM}}{r_{QM,MM}} \right) + E_{QM/MM}^{vdW} + E_{MM} \tag{1}$$

$$E_{QM/MM} = E_{vac} + E_{QM/MM}^{elect} + E_{QM/MM}^{vdW} + E_{MM} \tag{2}$$

where E_{MM} is the energy of the MM subsystem terms, $E_{QM/MM}^{vdW}$ the van der Waals interaction energy between the QM and MM subsystems and $E_{QM/MM}^{elect}$ includes both the Coulombic interaction of the QM nuclei (Z_{QM}) and the electrostatic interaction of the polarized electronic wave function (Ψ) with the charges of the protein (q_{MM}).

According to this scheme, the interaction energy between the substrate and the environment is evaluated as the difference between the QM/MM energy and the energies of the separated, noninteracting, QM and MM subsystems with the same geometry. Considering that the MM part is described using a nonpolarizable potential, the interaction energy contribution of each residue of the protein is given by the following expression:

$$E_{QM/MM}^{Int} = \sum \left\langle \Psi \left| \frac{q_{MM}}{r_{e,MM}} \right| \Psi \right\rangle + \sum \sum \frac{Z_{QM} q_{MM}}{r_{QM,MM}} + E_{QM/MM}^{vdW} \tag{3}$$

This interaction energy can be exactly decomposed in a sum over residues provided that the polarized wave function (Ψ) is

employed to evaluate this energy contribution. The global polarization effect can be obtained from the gas phase energy difference between the polarized, Ψ, and unpolarized, Ψ^o, wave functions.

Potential of Mean Force (PMF). In order to obtain the reaction energy profile in terms of free energies, we have traced several PMFs. All PMFs have been calculated using the weighted histogram analysis method (WHAM)[17] combined with the umbrella sampling approach[18] as implemented in fDYNAMO. The procedure for the PMF calculation is straightforward and requires a series of molecular dynamics simulations in which the distinguished reaction coordinate variable, ξ, is constrained around particular values. The values of the variables sampled during the simulations are then pieced together to construct a distribution function from which the PMF is obtained as a function of the distinguished reaction coordinate $(W(\xi))$. The PMF is related to the normalized probability of finding the system at a particular value of the chosen coordinate by eq 4:

$$W(\xi) = C - kT \ln \int \rho(r^N)\delta(\xi(r^N) - \xi)dr^{N-1} \quad (4)$$

The activation free energy can be then expressed as:

$$\Delta G^{\ddagger}(\xi) = W(\xi^{\ddagger}) - [W(\xi^R) + G_{\xi}(\xi^R)] \quad (5)$$

where the superscripts indicate the value of the reaction coordinate at the reactants (R), and at the TS $(\ddagger)$, and G_{ξ} (ξ^R) is the free energy associated with setting the reaction coordinate to a specific value at the reactant state. Normally this last term makes a small contribution and the activation free energy is directly estimated from the PMF change between the maximum of the profile and the reactant's minimum:

$$\Delta G^{\ddagger}(\xi) \approx W(\xi^{\ddagger}) - W(\xi^R) = \Delta W^{\ddagger}(\xi) \quad (6)$$

The selection of the reaction coordinate is usually trivial when the mechanism can be driven by a single internal coordinate or a simple combination (as the antisymmetric combination of two interatomic distances). However this is not the case for all possible steps of the reaction subject of study in this paper where many coordinates are participating. Instead we were compelled to obtain a much more computationally demanding 2D-PMF using two coordinates: ξ_1 and ξ_2. The 2D-PMF is related to the probability of finding the system at particular values of these two coordinates:

$$W(\xi) = C' - kT \ln \int \rho(r^N)\delta(\xi_1(r^N) - \xi_1)\delta(\xi_2(r^N) - \xi_2)dr^{N-2} \quad (7)$$

To estimate the activation free energy from this quantity, we recovered one-dimensional PMF changes tracing a maximum probability reaction path on the 2D-PMF surface and integrating over the perpendicular coordinate. Computational details of the procedure followed to get the 2D PMF required to explore every single step of the proposed mechanisms are reported in the Supporting Information.

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, a new energy function defined in terms of interpolated corrections was used.[19]

$$E = E_{AM1/MM} + S[\Delta E_{LL}^{HL}(\xi_1, \xi_2] \quad (8)$$

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 semi-empirical Hamiltonian was used as LL method, while a density functional theory (DFT)-based method was selected for the HL energy calculation. In particular, HL energy calculations were performed by means of the hybrid M06-2X[20] functional using the standard 6-31+G** basis set. These calculations were carried out using the Gaussian09 program.[21]

Kinetic Isotope Effects. Finally, average kinetic isotope effects (KIE) have been computed for isotopic substitutions of key atoms, from the transition states and the reactant complex localized at AM1/MM level of theory.

$$KIE = \frac{\left(\frac{Q_L}{Q_H}\right)_a}{\left(\frac{Q_L}{Q_H}\right)_b} = \frac{\left(\frac{A_L B_L C_L}{A_H B_H C_H}\right)_a^{1/2}}{\left(\frac{A_L B_L C_L}{A_H B_H C_H}\right)_b^{1/2}} \cdot \frac{\left(\frac{M_L}{M_H}\right)_a^{3/2}}{\left(\frac{M_L}{M_H}\right)_{ab}^{3/2}} \cdot \frac{ZPE_a}{ZPE_b} \cdot \frac{Exc(a)}{Exc(b)}$$

$$(9)$$

where H is heavy isotope, L is light isotope, M = mass of isotope, Q = partition function, A, B, and C are the principal moments of inertia, a = transition state structure coordinates, b = reactant complex structure coordinates, ZPE is the zero point energy factor, and Exc are Boltzman excitation factors, respectively, using the rigid-rotor/harmonic-oscillator approximation, without the scaling of vibrational frequencies, as explained and applied in previous papers.[22,23] The Hessian was computed for a subset of atoms of the full system, those of the QM region, consistent with the "cutoff rule" and the local nature of isotope effects.[24]

■ RESULTS AND DISCUSSION

The exploration of the potential energy surfaces (PES) of every single possible step of the hydrolysis of methyl acrylate catalyzed by PALB (see Figure S2 of Supporting Information) has been used to compute the free energy surfaces in terms of 2D PMFs at the M06-2X//OPLS/TIP3P level. The corresponding AM1/MM 2D PMFs are qualitatively equivalent to those presented in Figure 1 (see Figure S3 of Supporting Information). According to the surfaces presented in Figure 1, the reaction would take place through a multistep mechanism equivalent to the ping-pong bibi reaction mechanism initially proposed by Hult and Berglund.[12,25] Characterization of the stationary points located on the surfaces in Figure 1 has allowed proposing a stepwise mechanism, as presented in Figure 2.

Our results indicate that the hydrolysis of methyl acrylate catalyzed by PALB starts with the transfer of the hydroxyl proton of Ser105 to His224 and the concomitant attack to the carbonyl carbon atom of the substrate (see Figure 2). Interestingly, Frushicheva and Warshel, in a recent study of the enantioselectivity of *Candida antarctica* lipase A (CALA) based on empirical valence bond methods (EVB), observed that acylation took place in two steps: in a first step, the proton was transferred from serine to histidine, and later, the attack of the negatively charged serine on the carbonyl carbon of substrate took place to arrive at a tetrahedral intermediate equivalent to our intermediate 1.[26] Nevertheless, as mentioned by the authors, a previous study based on ab initio evaluation of the potential energy surface for general base-catalyzed methanolysis of formamide in solution, as a reference reaction in solution for serine proteases, suggested that a concerted path
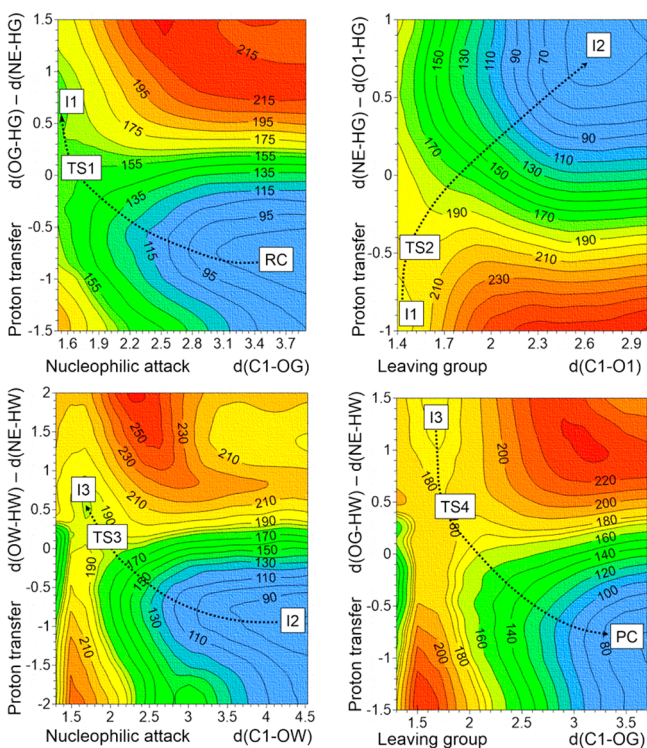
Figure 1. M06-2X//OPLS/TIP3P 2D-PMF surfaces for the reaction mechanism of hydrolysis of methyl acrylate catalyzed by PALB.

would give a similar barrier.[27] Following the analysis of our results, once the substrate is covalently bounded to the backbone of the protein, intermediate 1, His224 acts as an acid by protonating the ester oxygen of the substrate (O1), thus provoking the O1−C1 breaking bond and the methanol molecule release, intermediate 2. A conserved water molecule in the active site of PALB is responsible for the regeneration of the protonated His224 and the nucleophilic attack to the carbonyl carbon atom of the substrate, intermediate 3. Finally, a proton is transferred from the protonated His224 to the oxygen atom of Ser105 that forces the breaking of the substrate−protein O−C covalent bond. In this last step, the final product, acrylic acid, is obtained, and the protein is regenerated. As observed in the proposed mechanism, some residues of the active site of the protein, apart from the mentioned His224 and Ser106, emerge as crucial for the catalysis. Thus, Asp187 appears to form a dyad with His224 modulating the pK_a of this amino acid that behaves as a base or an acid depending on the step of the reaction. This is confirmed by a geometrical analysis of the distance between both residues during the full process (see AM1/MM distances listed in Table 1).

Distances between the two heavy atoms, OD−ND, is larger in reactants state, intermediate 2, and products state (structures where His224 presents a neutral electronic state) than in the other minimum energy structures. A significantly shorter distance is obtained when His224 has to reach its protonated state (intermediates 1 and 3), which favors stabilizing the positive charge developed in this residue, and in particular in TS1, where the two residues reach the closest position, thus favoring the proton transfer from Ser105.

The other key residues involved in the catalysis are those forming the oxyanion hole—Gln106 and Thr40. As previously proposed in the literature, our calculations show how interactions between the carbonyl oxygen of the substrate

and these two residues favor the reaction by polarizing the carbonyl bond. A deeper analysis can be carried out by computing the contributions of individual amino acids to substrate−protein interaction energies in RC, TS1, and TS2 (see Figure 3).

This figure, where negative values represent stabilizing effects, confirms the relevance of the residues forming the oxyanion hole in slightly stabilizing the substrate and, in particular, the negative charge developed in O2 in some of the steps of the full process. Hult and Berglund suggested that the oxyanion hole of PALB forms three hydrogen bonds to the oxygen of the substrate in the transition state, whereas, in the ground state, only two would be developed.[8] According to our results, one of the hydrogen bond interactions established with Thr40 is conserved along the reaction path, while the other one disappears when arriving at products complex. This could help the substrate to release in the last step. The third hydrogen bond interaction of the oxyanion hole, the one established with Gln106 appears to be much weaker. Interestingly, and in agreement with the proposal of Hult and Berglund, the largest value of the distance describing this interaction (HGln106−O2) is observed in reactants, while in the TS1 and during the intermediate steps, this distance is noticeably shortened (from 3.70 to 2.99 Å). Moreover, similar to the behavior of the HGln106−O2 hydrogen bond, once arriving to products, the distance is enlarged again. These distances present their minimum values in intermediates 1 and 3, where the double bond is transformed into a single bond and the maximum separation of charges is reached. This behavior is confirmed by inspection of not only averaged distances listed in Table 1 but the Mulliken atomic charges reported in Table 2. Charge on carbonyl oxygen atom of the substrate (O2) arrives to −1.04 and −1.09 au in I1 and I3, where the maximum positive values on C1 are also reached, 0.98 and 1.09 au, respectively.

From the energetics point of view, the reaction paths traced on the M06/MM and AM1/MM free energy surfaces (Figure 1 and S2, respectively), have been projected into a two-dimensional free energy profile that is depicted in Figure 4.

The first conclusion that can be derived is that a dramatic reduction of the free energy barriers of all steps of the process is obtained when improving the QM description of the system from a semiempirical to a density functional theory (DFT) Hamiltonian. In fact, it is noticeable how M06-2X/MM calculations predict a rather short half-life time for intermediates 1 and 3. This result would be close to the EVB-based results of Frushicheva and Warshel on CALA, suggesting a concerted mechanism for arriving at the tetrahedral covalent intermediate would give similar barrier to the stepwise mechanism.[26] On the other hand, intermediate 2 appears to be intensely stabilized by the oxyanion hole and the covalent bond the substrate establishes with Ser105. This result is obtained at both levels of theory, AM1/MM and M06-2X/MM. Consequently, we predict that this species could be isolated experimentally. According to the M06-2X/MM results, the highest energy TSs correspond to the first step describing the nucleophilic attack of the serine oxygen to the carbonyl carbon of the ester and the second step where the ester bond of the substrate is broken, with participation of His224. The M06-2X/MM energy of TS1 and TS2 relative to reactants is very close, 19.1 and 19.7 kcal·mol^{-1}, respectively. These values are in very good agreement with the 20.3 kcal·mol^{-1}, which is the free energy barrier that can be deduced from the experimental catalytic events measure by Berglund and co-workers,[28] and
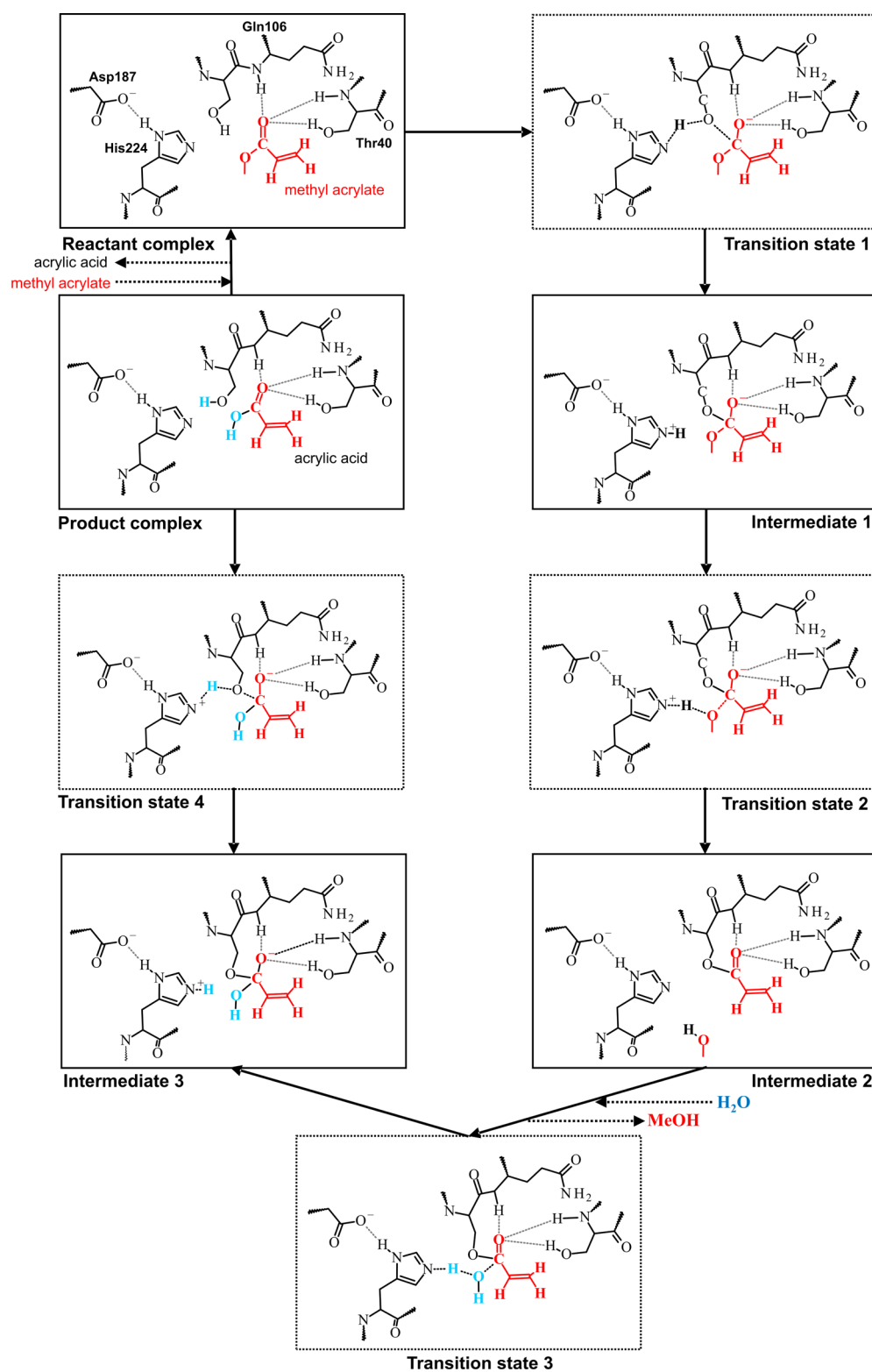
Figure 2. Multistep proposed reaction mechanism for hydrolysis of methyl acrylate to acrylic acid and methanol, catalyzed by PALB.

close to 18.3 kcal·mol⁻¹, which is the free energy barrier obtained by Frushicheva and Warshel for the concerted acylation step in the catalytic reaction of wild type CALA.[26] TS3, associated with the nucleophilic attack of the water molecule to the carbonyl carbon of the substrate, presents the highest energy barrier at this level of calculation. Nevertheless, it is important to note that energy of TS1 and TS2 are clearly

higher than TS3. The higher barrier of the third step is due to the strong stabilization of intermediate 2 and also to a possible overestimation due to the procedure followed to perform the energy corrections at the M06-2X level. In fact, as previously shown in our study of the Michael-type addition performed in gas phase by means of the AM1 and by DFT calculations within the B3LYP and M06-2X hybrid functionals, no displacement

Table 1. QM/MM Averaged Key Distances (in Å) Obtained in the States Located along the Reaction Path of the Hydrolysis of Methyl Acrylate Catalyzed by PALB

	RC	TS1	I1	TS2	I2	TS3	I3	TS4	PC
OG–HG	0.97 ± 0.02	1.32 ± 0.03	2.15 ± 0.03	2.24 ± 0.11	3.05 ± 0.21	4.90 ± 0.40	4.68 ± 0.21	6.00 ± 1.02	2.97 ± 0.33
HG–NE	2.25 ± 0.03	1.25 ± 0.03	1.02 ± 0.03	1.24 ± 0.03	2.58 ± 0.03	3.61 ± 0.35	3.44 ± 0.21	5.66 ± 1.13	2.74 ± 0.27
HG–O1	2.58 ± 0.26	2.26 ± 0.11	1.95 ± 0.12	1.32 ± 0.03	0.97 ± 0.03	0.97 ± 0.03	0.97 ± 0.03	0.97 ± 0.02	0.97 ± 0.02
O1–C1	1.37 ± 0.02	1.42 ± 0.03	1.46 ± 0.03	1.54 ± 0.02	3.08 ± 0.02	4.23 ± 0.28	4.09 ± 0.15	5.36 ± 0.55	3.39 ± 0.23
OG–C1	2.53 ± 0.02	1.53 ± 0.02	1.45 ± 0.02	1.42 ± 0.03	1.36 ± 0.02	1.43 ± 0.02	1.45 ± 0.02	1.55 ± 0.02	2.49 ± 0.02
O$^{(HOH)}$–C1	4.14 ± 0.32	4.13 ± 0.21	4.06 ± 0.26	4.01 ± 0.32	4.65 ± 0.77	1.49 ± 0.02	1.44 ± 0.03	1.41 ± 0.02	1.36 ± 0.02
H$^{(HOH)}$–NE	5.16 ± 0.83	4.98 ± 0.73	4.20 ± 0.36	5.03 ± 0.75	5.17 ± 0.71	1.22 ± 0.03	1.03 ± 0.03	1.24 ± 0.03	2.46 ± 0.03
H$^{(HOH)}$–OG	5.84 ± 0.75	4.73 ± 0.54	4.30 ± 0.31	5.14 ± 0.70	5.65 ± 0.77	2.34 ± 0.12	2.39 ± 0.18	1.31 ± 0.03	0.97 ± 0.03
C1–O2	1.24 ± 0.02	1.29 ± 0.02	1.30 ± 0.02	1.29 ± 0.02	1.24 ± 0.02	1.30 ± 0.02	1.31 ± 0.02	1.29 ± 0.02	1.24 ± 0.02
His224–Asp187									
OD–HD	2.37 ± 0.33	2.28 ± 0.27	2.22 ± 0.18	2.23 ± 0.20	2.37 ± 0.23	2.47 ± 0.25	2.31 ± 0.20	2.21 ± 0.18	2.35 ± 0.22
ND–HD	0.99 ± 0.02	1.00 ± 0.03	1.01 ± 0.03	1.00 ± 0.03	1.00 ± 0.03	1.00 ± 0.03	1.00 ± 0.03	1.00 ± 0.03	0.99 ± 0.02
OD–ND	3.11 ± 0.22	2.93 ± 0.16	2.98 ± 0.13	2.99 ± 0.14	3.13 ± 0.16	3.14 ± 0.16	3.05 ± 0.14	2.90 ± 0.12	3.10 ± 0.17
oxyanion hole									
H$^{(Gln106)}$–O2	3.70 ± 0.15	2.99 ± 0.12	2.95 ± 0.11	3.00 ± 0.11	3.01 ± 0.14	2.90 ± 0.10	2.87 ± 0.10	2.92 ± 0.13	3.46 ± 0.20
H$^{(N-Thr40)}$–O2	1.98 ± 0.14	1.83 ± 0.11	1.81 ± 0.10	1.81 ± 0.10	1.99 ± 0.14	1.80 ± 0.12	1.82 ± 0.10	1.79 ± 0.11	2.39 ± 0.21
H$^{(O-Thr40)}$–O2	1.98 ± 0.14	1.94 ± 0.13	1.90 ± 0.14	1.95 ± 0.13	1.97 ± 0.14	1.91 ± 0.12	1.86 ± 0.11	1.92 ± 0.13	1.92 ± 0.11
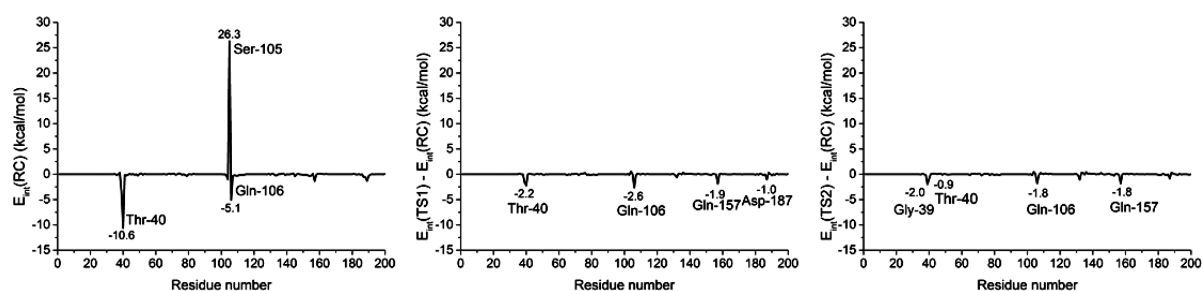
Figure 3. Averaged contributions of individual amino acids to substrate–protein interaction energies (in kcal·mol^{-1}), computed for RC, left panel, and the difference between TS1 and RC (center panel) and TS2 and RC (right panel).

Table 2. Averaged Atomic Charges (in a.u.) Obtained in the States Located along the Reaction Path of the Hydrolysis of Methyl Acrylate Catalyzed by PALB

|	RC	TS1	I1	TS2	I2	TS3	I3	TS4	PC
C1	0.920	0.914	0.984	0.837	0.798	1.038	1.091	1.010	0.892
O1	−0.358	−0.420	−0.605	−0.386	−0.455	−0.493	−0.533	−0.468	−0.609
O2	−0.777	−1.000	−1.041	−0.982	−0.753	−1.056	−1.095	−1.027	−0.778
Asp187−COO$^-$	−0.842	−0.839	−0.841	−0.849	−0.846	−0.848	−0.847	−0.840	−0.836
HD	0.344	0.320	0.305	0.302	0.302	0.335	0.341	0.328	0.326
ND	−0.302	−0.180	−0.081	−0.094	−0.172	−0.261	−0.230	−0.237	−0.288
OG	−0.352	−0.272	−0.512	−0.448	−0.420	−0.594	−0.663	−0.604	−0.608
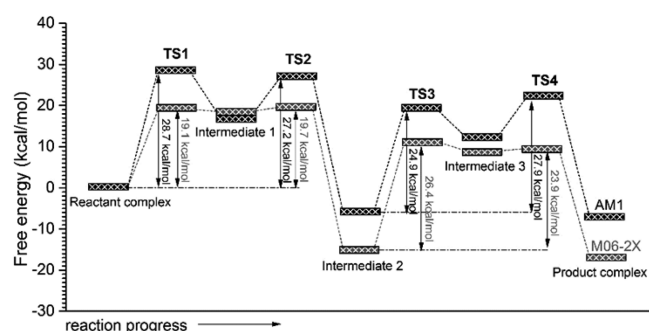
Figure 4. Free energy profile for the multistep reaction mechanism of hydrolysis of methyl acrylate catalyzed by PALB obtained at M06-2X/OPLS/TIP3P (in gray) and AM1/OPLS/TIP3P (in black).

was observed of the proton between the dyad formed by His224 and Asp187 at AM1 level when a positive charge is developed in the His224.[29] This is in agreement with the present QM/MM results, as deduced from Table 1. On the contrary, a spontaneous transfer of the proton to the aspartate can take place when using DFT methods, thus stabilizing the system. Since the corrected high level 2D PMFs surfaces have been obtained from the structures generated from the AM1/MM calculations, no possible exploration of a lower energy TS can be performed at M06-2X/MM.

In order to support the proposed mechanism as well as to confirm which is the rate-limiting step, kinetic isotope effects (KIE) could be measured. Thus, 10 different structures of TSs and reactants have been selected from the constrained MD simulations performed on these states in order to compute KIEs for substitution on key atoms C1, O1, and O2. Analysis of results listed in Table 3 shows that isotopic substitution of C1

Table 3. Averaged KIEs Computed at the AM1/MM Level at 300K for the Hydrolysis of Methyl Acrylate in the Active Site of PALB

|	TS1	TS2	TS3	TS4
$^{13}C1$	1.0081 ± 0.0019	1.0080 ± 0.0027	1.0010 ± 0.0040	1.0071 ± 0.0037
$^{18}O1$	0.9853 ± 0.0042	1.0056 ± 0.0064		
$^{18}O2$	0.9911 ± 0.0066	0.9805 ± 0.0075	0.9795 ± 0.0094	0.9815 ± 0.0091
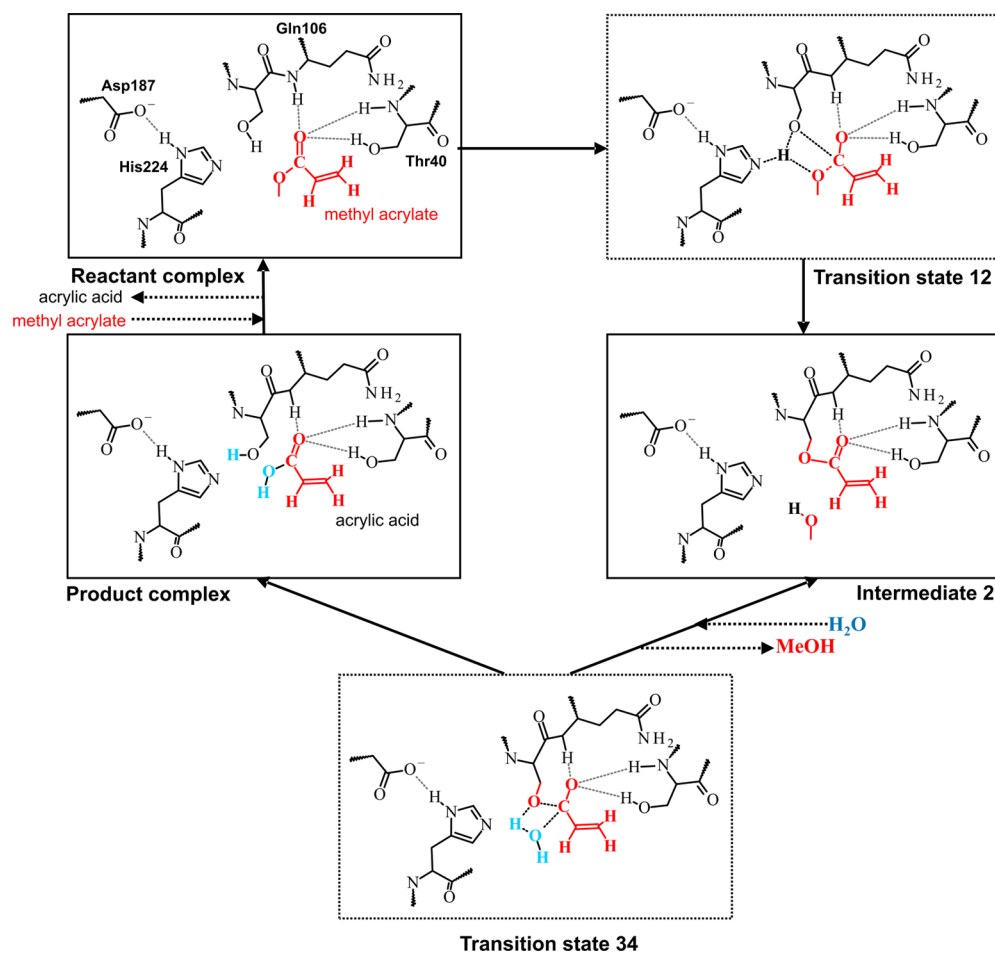
Figure 5. Two-step proposed reaction mechanism for hydrolysis of methyl acrylate catalyzed by PALB.

and O2 does not render differences between TS1 and TS3. Obviously, the change in orbital hybridization of these two atoms in step 1 and 3 is equivalent; C1 changes from sp2 to sp3, and O1 changes from double bond to single bond. Mass substitution on O1 could shed some light into the question. According to the four-step mechanism, and since KIEs are computed (and measured) as a relative effect between each of the TS and reactants (Michaelis complex), a very slightly inverse effect should be expected if TS1 were the rate-limiting step (0.9853 ± 0.0042). Nevertheless, a measurable normal effect would be expected in the case of TS2, when the O1–C1 bond is broken and a proton is transferred to O1 from His224 (1.0056 ± 0.0064). Since O1 belongs to the methanol molecule that is released in step 2, no calculations have been performed for steps 3 and 4.

Finally, an alternative reaction mechanism has been explored, where the reaction can proceed in two steps through an intermediate equivalent to the intermediate 2 located in the previously discussed mechanism. A schematic representation of the stationary point structures is presented in Figure 5, while

the free energy profile, computed at the AM1/MM and M06-2X/MM level, is shown in Figure 6. As observed, the attack of Ser105 and hydrolysis of substrate takes place in a single concerted step. The second step corresponds to the
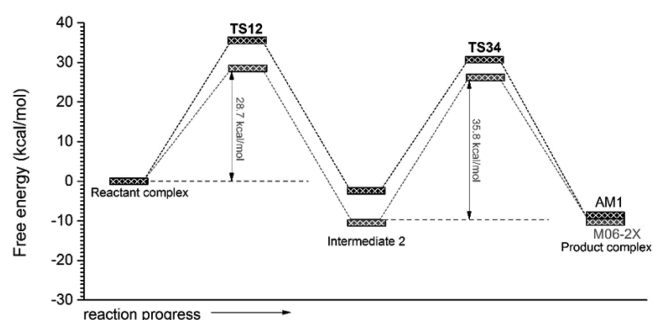
Figure 6. Free energy reaction pathway of the two steps mechanisms of hydrolysis of methyl acrylate catalyzed by PALB computed at AM1/MM (black line) and M06-2X/MM (gray line) level.

nucleophilic attack of the water molecule and the scission of the bond established between the substrate and Ser105 to release products and regenerate the protein. The results suggest that the four-step mechanism is kinetically more favorable than this two-step mechanism, considering that significantly lower energy barriers are obtained for the former.

CONCLUSIONS

The study of the primary reaction of promiscuous *Pseudozyma antarctica* (formerly *Candida antarctica*) lipase B (PALB), its native hydrolytic activity on the methyl acrylate molecule, has been studied using hybrid QM/MM MD simulations describing the QM region at two levels of theory, AM1 and M06-2X. PESs and free energy surfaces have been explored, and the results suggest that the catalytic reaction can take place by means of a four-step mechanism equivalent to the ping-pong bibi reaction mechanism initially proposed by Hult and Berglund.[12,25] A possible two-step mechanism has been also explored, but the resulting free energy barriers are much higher. According to our results, the first step, the activation of the carbonyl group of the substrate and the nucleophilic attack of Ser105 to the carbonyl carbon atom, presents the highest energy transition state, while the dramatic stabilization of intermediate 2, which correspond to the hydrolyzed acrylate molecule covalently bonded to the protein backbone, provokes that the third step, the nucleophilic attack of a conserved water molecule located in the active site, presents the highest energy barrier. It is important to point out the significant effect on the energetics of the reaction when improving the QM description of the system from semiempirical to DFT. Thus, a dramatic reduction of the free energy barriers of all steps of the process is observed at M06-2X/MM level by comparison with the AM1/MM results. Also, it is noticeable how M06-2X/MM calculations predict a rather short half-life time for intermediates 1 and 3. On the other hand, intermediate 2 appears to be a very stable intermediate stabilized by the oxyanion hole and the covalent bond the substrate establishes with Ser105. Consequently, we predict that this species could be isolated experimentally. According to the M06-2X/MM results, the highest energy TSs correspond to the first step describing the nucleophilic attack of the serine oxygen to the carbonyl carbon of the ester and the second step where the ester bond of the substrate is broken, with participation of His224. The M06-2X/MM energy of TS1 and TS2 relative (19.1 and 19.7 kcal·mol^{-1}, respectively) is very close to the free energy barrier that can be deduced from the experimental catalytic events measured by Berglund and co-workers (20.3 kcal·mol^{-1}),[26] and the free energy barrier obtained by means of EVB calculations by Frushicheva and Warshel for the acylation step in the catalytic reaction of wild type CALA (18.3 kcal·mol^{-1}).[26]

Our results suggest that the origin of the catalytic activity of the enzyme is due to favorable interactions established between the residues of the active site that create the oxyanion hole, Gln106 and Thr40, and the carbonyl oxygen atom of the substrate that develops a negative charge during the reaction and, in particular, in the rate-limiting TS. Nevertheless, there are other key residues for catalysis such as Ser105 that forms a covalent bond with the substrate stabilizing the intermediate states, or the Asp187 that is capable of modulating the pK_a of His224, favoring its role as a base or acid depending on the step of the process. KIEs have been computed for substitutions of some key atoms, revealing the possibility of using this tool to support the proposed mechanism and to distinguish the rate-limiting step.

A detailed knowledge of the molecular mechanism of the primary reaction of a promiscuous enzyme should be able to provide a useful reference if the protein is to be used as a template to engineer new catalytic properties. In particular, some of the features of PALB such as the oxyanion hole of the active site, capable of stabilizing a negatively charge that can be develop during the reaction, or the existence of a Asp-His dyad with a modulated pK_a can be used to redesign PALB into a catalysts for a different reaction.

ASSOCIATED CONTENT

Ⓢ Supporting Information
Computational details of generation of 2D PMFs. Time dependence of RMSD of atoms of backbone of the protein during equilibration (Figure S1), AM1/OPLS/TIP3P PESs (Figure S2), and AM1/OPLS/TIP3P 2D-PMFs (Figure S3) for the reaction mechanism of hydrolysis of methyl acrylate catalyzed by PALB. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: moliner@uji.es. Fax: (+34) 964-728066.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Arnold, F. H. *Nature* **2001**, *409*, 253−257.
(2) Bommarius, A. S.; Blum, J. K.; Abrahamson, M. J. *Curr. Opin. Chem. Biol.* **2011**, *15*, 194−200.
(3) 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.
(4) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St.Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; Houk, K. N.; Michael, F. E.; Baker, D. *Science* **2010**, *329*, 309−313.
(5) Świderek, K.; Tuñón, I.; Moliner, V. *WIREs Comput. Mol. Sci.* **2013**, DOI: 10.1002/wcms.1173.
(6) O'Brien, P. J.; Herschlag, D. *Chem. Biol.* **1999**, *6*, R91−R105.
(7) Aharoni, A.; Gaidukov, L.; Khersonsky, O.; McQ Gould, S.; Roodveldt, C.; Tawfik, D. S. *Nat. Genet.* **2005**, *37*, 73−76.
(8) Hult, K.; Berglund, P. *Trends Biotechnol.* **2007**, *25*, 231−238.
(9) Ollis, D. L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S. M.; Harel, M.; Remington, S. J.; Silman, I.; Schrag, J.; Sussman, J. L.; Verschueren, K. H. G.; Goldman, A. *Protein Eng.* **1992**, *5*, 197−211.
(10) Holmquist, M. *Curr. Protein Pept. Sci.* **2000**, *1*, 209−235.
(11) Uppenberg, J.; Öhrner, N.; Norin, M.; Hult, K.; Kleywegt, G. J.; Patkar, S.; Waagem, V.; Anthonsen, T.; Jones, T. A. *Biochemistry* **1995**, *34*, 16838−16851.
(12) Hult, K.; Berglund, P. *Trends Biotechnol.* **2007**, *25*, 231−238.
(13) Warshel, A.; Levitt, M. J. *Mol. Biol.* **1976**, *103*, 227−249.

(14) Uppenberg, J.; Hansen, M. T.; Patkar, S.; Jones, T. A. *Structure* **1994**, *2*, 293−308.

(15) (a) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins* **2005**, *61*, 704−721. (b) Bas, D. C.; Rogers, D. M.; Jensen, J. H. *Proteins* **2008**, *73*, 765−783. (c) Olsson, M. H. M.; Søndergard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theor. Comput.* **2011**, *7*, 525−537. (d) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theor. Comput.* **2011**, *7*, 2284−2295.

(16) Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21*, 1088−1100.

(17) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(18) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.

(19) (a) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S. *J. Phys. Chem. B* **2006**, *110*, 17663−17670. (b) Chuang, Y. Y.; Corchado, J. C.; Truhlar, D. G. *J. Phys. Chem. A* **1999**, *103*, 1140−1149. (c) Ferrer, S.; Martí, S.; Moliner, V.; Tuñón, I.; Bertran, J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 3482−3489.

(20) (a) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215−241. (b) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157−167.

(21) 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.

(22) Martí, S.; Moliner, V.; Tuñón, I.; Williams, I. H. *J. Phys. Chem. B* **2005**, *109*, 3707−3710.

(23) Martí, S.; Moliner, V.; Tuñón, I.; Williams, I. H. *Org. Biomol. Chem.* **2003**, *1*, 483−487.

(24) Ruggiero, G. D.; Guy, S. J.; Martí, S.; Moliner, V.; Williams, I. H. *J. Phys. Org. Chem.* **2004**, *17*, 592−601.

(25) Hult, K. In *Microbial Reagents in Organic Synthesis*; Servi, S., Ed.; Kluwer Academic Publishers: The Netherlands, 1992; pp 289−298.

(26) Frushicheva, M. P.; Warshel, A. *ChemBioChem* **2012**, *13*, 215−223.

(27) Strajbl, M.; Florián, J.; Warshel, A. *J. Am. Chem. Soc.* **2000**, *122*, 5354−5366.

(28) Svedendahl, M.; Jovanovic, B.; Fransson, L.; Berglund, P. *ChemCatChem* **2009**, *1*, 252−258.

(29) Świderek, K.; Pabis, A.; Moliner, V. *Org. Biomol. Chem.* **2012**, *10*, 5598−5605.