

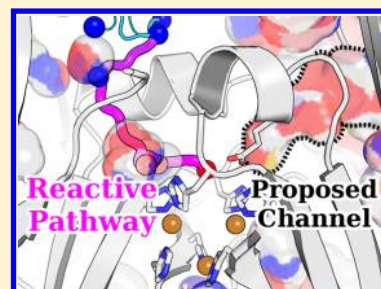
# The Pathway for O<sub>2</sub> Diffusion inside CotA Laccase and Possible Implications on the Multicopper Oxidases Family

João M. Damas,\* António M. Baptista, and Cláudio M. Soares\*

Instituto de Tecnologia Química Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal

## S Supporting Information

**ABSTRACT:** Laccases and multicopper oxidases (MCOs) oxidize a wide range of organic compounds while reducing O<sub>2</sub> to water, enabling numerous biotechnological applications. It is still unknown how O<sub>2</sub> reaches the internalized catalytic center of MCOs where it gets reduced, despite a proposed channel inferred from X-ray crystallography structures. Herein, an alternative new pathway is found through the use of a combination of free energy calculations (implicit ligand sampling), landscape analysis, and Markov modeling. The reported pathway is shown to be the one mostly contributing to O<sub>2</sub> reaching the catalytic center. This pathway is considered in light of the whole MCO family, and a relation to the protonation state of a structurally conserved acidic residue right above the center is advanced.

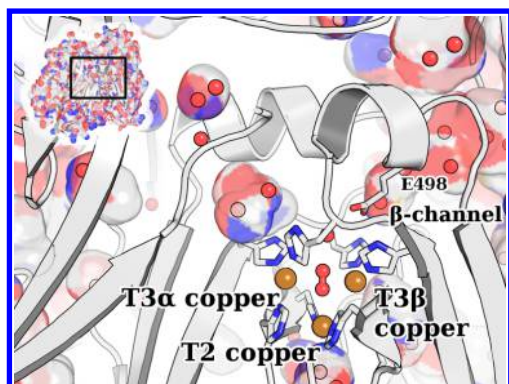


## INTRODUCTION

With the emergence of green chemistry, laccases have become a major focus of research and development due to their biotechnological potential.<sup>1</sup> Laccases are multicopper oxidases (MCOs) capable of nonspecifically oxidizing a wide variety of organic compounds while reducing dioxygen (O<sub>2</sub>) to water.<sup>2</sup> Near their surface lies a T1 copper center which is believed to shuttle electrons from the oxidized substrates to an internalized T2–T3 copper center (Figure 1), where the reduction process occurs. Characterization of the structure of laccases and the kinetics of its catalysis is essential for understanding the underlying molecular details, which may lead to enzyme engineering and optimization toward specific applications. Through rational and semirational (rational combined with

directed evolution) approaches, it was shown possible, for example, to tune the T1 center redox potential<sup>3</sup> or the reducing substrate specificity.<sup>4</sup> However, these engineering efforts have dealt mostly with the T1 center and electron uptake. The reactive properties at the T2–T3 center seem to be much harder to engineer without diminishing its activity, probably due to the fine-tuning of the center toward irreversible O<sub>2</sub> catalysis,<sup>5</sup> comprising second sphere carboxylate residues that stabilize the coordinatively unsaturated center<sup>6</sup> and participate on protonation events during catalysis.<sup>7–13</sup> In fact, the catalytic/turnover cycle of O<sub>2</sub> at the T2–T3 has been intensively characterized by spectroscopic and theoretical techniques,<sup>14–16</sup> as well as X-ray crystallography studies.<sup>17</sup> Still, enzymatic O<sub>2</sub> catalysis depends not only on its turnover at the center ( $k_{\text{cat}}$ ) but also on its diffusion and binding affinity ( $k_{\text{cat}}/K_{\text{m}}$ ), and regarding this last component, typical kinetic experiments on this kind of enzyme do not distinguish the diffusion events between the solvent and active center, from the (un)binding events at that internalized active center.<sup>18</sup>

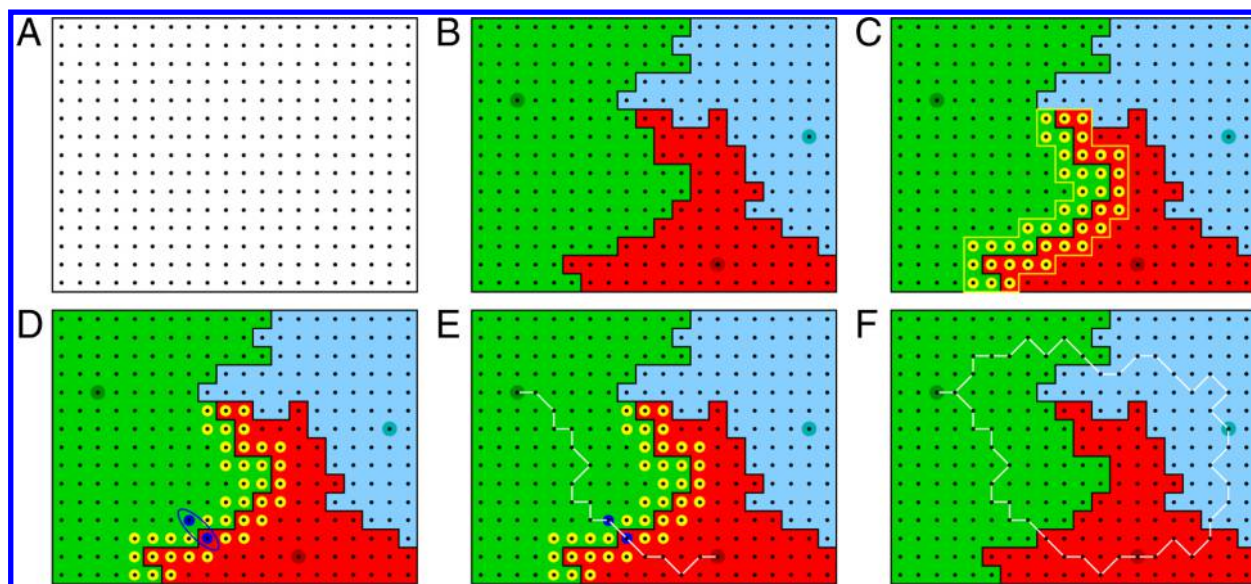
A hydrophilic water-filled channel from the solvent toward the T3 $\beta$  copper ( $\beta$ -channel, Figure 1) has been generally assumed as the probable pathway for O<sub>2</sub> entry to the T2–T3 center of laccases,<sup>19</sup> but further evidence has been lacking. A study with xenon (Xe) as an O<sub>2</sub> probe has addressed this issue on MaL laccase,<sup>20</sup> whose  $\beta$ -channel is blocked by its own C-terminal region, but despite the finding of hydrophobic binding sites, the proposed final step for diffusion to the center requires crossing through a two-stranded  $\beta$ -sheet. The lack of studies on O<sub>2</sub> diffusion inside laccases may be explained by the difficulty of addressing it from the experimental point of view. However, in these past few years, biomolecular modeling techniques have



**Figure 1.** Close-up view of the internalized T2–T3 center of O<sub>2</sub>-bound fully oxidized CotA laccase (PDB: 1W6L). The T3 coppers and T2 copper (brown spheres) are coordinated by three and two histidines, respectively. Waters are shown as red spheres and cavities are highlighted by the molecular surface of the protein. The inset localizes this center within the whole protein.

Received: March 7, 2014

Published: June 5, 2014



**Figure 2.** Two-dimensional scheme for the identification of lowest-energy pathways. Grid points are represented as black dots. In this scheme, three basins are identified (green, blue, and red), with their corresponding energy minima in matching colored circles, and delimited by black lines. The objects necessary for finding the lowest-energy pathway (in white) between the green and red basins are exemplified, with the border in yellow circles and the saddle pair in blue circles.

proven themselves as an invaluable approach to study these phenomena.<sup>18,21–26</sup> There are several examples of proteins whose function crucially depends on the ability of small ligands to reach internalized regions, and molecular dynamics (MD) or related techniques have successfully been used to observe and study these processes microscopically, allowing the calculation of free energy barriers and pathways of the diffusion,<sup>22–24</sup> or kinetic data comparable to experimental data.<sup>18,25,26</sup> Even simpler computational techniques have been shown to give fundamental insights leading to enzyme semirational optimization.<sup>27</sup>

Herein, we have applied a combination of MD simulations, free energy calculations (implicit ligand sampling, ILS), energy landscape analysis, and Markov modeling to the diffusion of O<sub>2</sub> inside a bacterial model laccase, CotA, in order to shed light on the mode of entry of O<sub>2</sub> in laccases and, more generally, MCOs. The obtained results bring a new perspective to the field of laccases and MCOs.

## METHODOLOGY

**System Preparation.** Based on previous calculations of the fully reduced CotA laccase,<sup>10</sup> where Glu498 was found to be protonated about half the time around pH 7, two solvated systems were prepared to account for the protonated (E498<sup>−</sup>) and unprotonated (E498<sup>0</sup>) states of Glu498. The protonation states of the other residues were the same for both systems and were also determined through a combination of Monte Carlo and Continuum Electrostatics<sup>28,29</sup> (more details in the Supporting Information).

**MD Simulations.** The GROMOS 54A7 force field<sup>30</sup> and the SPC water model<sup>31</sup> were used to describe nearly all of the system (details about the copper centers in the Supporting Information), with GROMACS 4.5.4<sup>32</sup> being used to perform 15 MD simulations of each system (E498<sup>0</sup> and E498<sup>−</sup>) for 20 ns. These simulations were performed using particle mesh Ewald (PME)<sup>33</sup> in the NPT ensemble with Berendsen baths<sup>34</sup> at 298.15 K and 1 atm (details on initialization and other simulation parameters in the Supporting Information). Then, in

order to study O<sub>2</sub> diffusion, we devised a protocol<sup>35</sup> that uses different starting conformations of the protein, which should improve the sampling of O<sub>2</sub> diffusion. For each system (E498<sup>0</sup> and E498<sup>−</sup>), three configurations were sampled from each of the first three MD simulations of CotA laccase in water (at 12, 16, and 20 ns). To each of these 18 conformations (9 for each system), 68 molecules of O<sub>2</sub> were added in the solvent region at random positions and with random velocities sampled from a Maxwell–Boltzmann distribution (a different set of 68 random positions and velocities for each conformation). Water molecules within a 2 Å distance from the added O<sub>2</sub> molecules were removed, and 10 ps of MD simulation with the O<sub>2</sub> molecules position restrained at 1000 kJ·mol<sup>−1</sup>·nm<sup>−2</sup> were run to let the solvent adapt to the added O<sub>2</sub>. After this initialization protocol, each of these 18 conformations was used as the starting point for nine MD simulations of each system for 50 ns (450 ns total for each system). In addition to the parametrization described above, a previously parametrized model<sup>36</sup> was used for O<sub>2</sub>, which was parametrized to successfully reproduce its free energy of solvation.

**ILS Calculations.** In addition to the explicit O<sub>2</sub> diffusion studies (see above), the implicit ligand sampling (ILS) method<sup>22</sup> was used to calculate the free energy of having O<sub>2</sub> anywhere inside CotA and its environment. From this method, the potential of mean force of having a diatomic ligand at a position **r** is given by

$$W(\mathbf{r}) = -\beta^{-1} \ln \sum_{m=1}^M \sum_{k=1}^C \frac{e^{-\beta \Delta E(\mathbf{r}, \mathbf{q}_m, \boldsymbol{\Omega}_k)}}{MC} \quad (1)$$

where *M* is the number of sampled protein+solvent configurations, *C* is the number of equally probable orientations of the ligand,  $\beta = 1/k_B T$ , where *k<sub>B</sub>* is the Boltzmann constant and *T* is the absolute temperature, and  $\Delta E(\mathbf{r}, \mathbf{q}_m, \boldsymbol{\Omega}_k)$  is the interaction energy between the protein+solvent in configuration **q<sub>m</sub>** and the diatomic ligand located at **r** with an orientation **Ω<sub>k</sub>**. The interaction energy  $\Delta E(\mathbf{r}, \mathbf{q}_m, \boldsymbol{\Omega}_k)$  accounts for the nonbonded interactions, both electrostatic and van der



Waals; in the case of the O<sub>2</sub> model used herein (see above), only van der Waals interactions are present. For these calculations, we modified the GROMACS 4.5.4 Widom TPI algorithm to perform ILS (details in the Supporting Information). The last 2000 configurations (2 ns) of each of the 30 MD trajectories of CotA laccase in water without dioxygen (15 for each state, see above) were used for ILS calculations ( $M = 2000 \times 30 = 60\,000$  in eq 1), with the configurations being fitted to the protein X-ray structure using the C $\alpha$  atoms. A grid of  $62 \times 52 \times 55$  dimensions, with Cartesian coordinates for the minimum vertex of  $x = 38.39$  Å,  $y = 44.565$  Å, and  $z = 10.875$  Å and for the maximum vertex of  $x = 100.39$  Å,  $y = 96.565$  Å, and  $z = 65.875$  Å, was used in the calculations. These calculations were performed on each configuration using the modified GROMACS 4.5.4 with ILS (version 1.0) described above, with 400 insertions ( $C = 400$  in eq 1) per grid cube. This high amount of calculations is made possible by distributing the configurations over different CPUs and ensemble averaging them (see eq 1) afterward, over all configurations of both systems together and each system (E498<sup>0</sup> and E498<sup>-</sup>) in separate, resulting in three discretized scalar fields (3D landscapes). These 3D landscapes describe the Gibbs free-energy of moving O<sub>2</sub> from vacuum to a given position in the system (either E498<sup>0</sup>, E498<sup>-</sup>, or a mixture of both),  $\Delta G_{\text{vac} \rightarrow \text{prot}}(\text{O}_2)$ . In the end, these energy landscapes were converted into the  $\Delta G_{\text{wat} \rightarrow \text{prot}}(\text{O}_2)$  landscapes of interest (details in the Supporting Information).

**Energy Landscape Analysis.** We have devised a method for analyzing discretized 3D energy landscapes (generalizable to higher dimensions) that is an extension of a previously developed approach<sup>37</sup> (where the energy minima and corresponding low-energy basins are determined). The objective is to identify the saddle points between basins and, therefore, the lowest-energy pathways between their minima. For that, we start by determining the discretized *steepest-descent path* of each grid point, obtained by iteratively linking the neighbor-to-neighbor jumps of maximum energy decrease, with the neighborhood of a point defined as the adjacent 26 grid points forming a  $3 \times 3 \times 3$  cube around it (in the 2D scheme of Figure 2, it is rather defined as the adjacent eight grid points forming a  $3 \times 3$  square around it); the end points of these steepest-descent paths are the minima on the landscape. A *basin* is then defined as the set of all grid points in 3D space whose steepest-descent path along the landscape leads to a particular minimum, i.e., all grid points that “fall” to the same minimum belong to the same basin. After finding all energy minima and classifying the entire landscape in terms of basins (the process from A to B in Figure 2), a *border pair* was defined as a pair of neighbor grid points in which each point belongs to a different basin. A *border* between two neighbor basins was then defined as the set of all border pairs whose points belong to either of the two basins (the border between the green and red basins is highlighted in yellow in Figure 2C), thereby forming two “half-borders,” one belonging to each basin (that would coalesce to a surface in the continuous limit). Each border pair consists of a maximum and a minimum energy grid point, and each grid point may belong to more than one border pair or even to more than one border. After identifying a border (yellow circles in Figure 2C), the corresponding *saddle pair* (blue circles highlighted in Figure 2D) was defined as the border pair whose maximum energy is the lowest among all border pairs composing that border, and when this condition was verified by more than one border pair, the one with lowest minimum

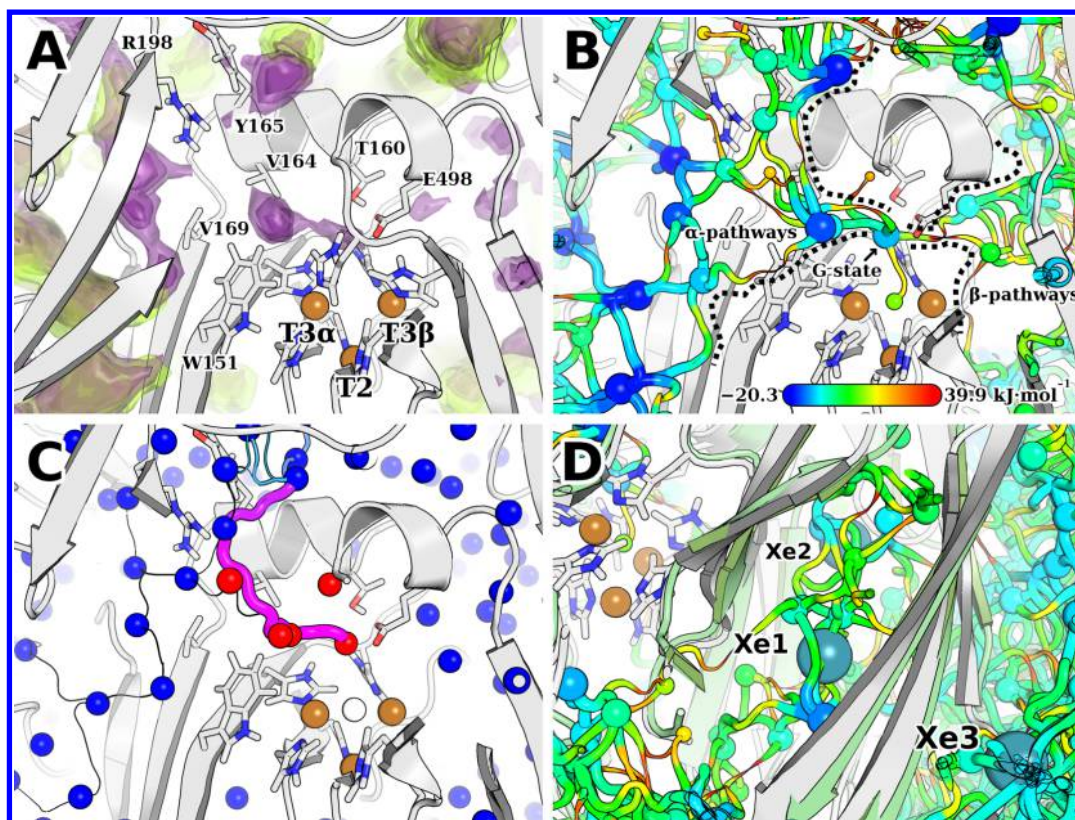
energy was chosen (the saddle pair would tend to a saddle point in the continuous limit). The *lowest-energy pathway* between the minima of two neighbor basins was then constructed by connecting the two steepest-descent paths from each of the saddle pair points (white line in Figure 2E). Once this is done for all neighbor basins, all landscape minima get connected among themselves by lowest-energy pathways, as illustrated in Figure 2F.

**O<sub>2</sub> Diffusion Kinetics Modeling.** The ILS calculations (see above) provide an exhaustive sampling over the full space of the system, even for the high-energy regions, allowing us to treat the energy landscape as a model for O<sub>2</sub> diffusion in CotA laccase. The energy landscape classification into basins provides an obvious division of states—a steepest-descent tessellation. Considering these basins as belonging to the state space of O<sub>2</sub> diffusion inside CotA laccase, we may use them to construct a simplified model for that process. This model is a Markov stochastic process, which will describe the time-discrete evolution of the system in the state space, i.e. the diffusion of O<sub>2</sub> inside the protein. To construct such a model, we need to construct a transition probability matrix where each element  $T_{ij}(\Delta t)$  corresponds to the probability of transition to basin/state  $j$  after a time  $\Delta t$ , when being in basin/state  $i$  at an arbitrary time. This matrix is usually inferred from the statistics of the dynamics in the state space, but as the ILS calculations do not provide these statistics, the matrix had to be inferred from the energy landscape. Using the Metropolis sampling for jumps between neighbor grid points and assuming that the grid-point probability distribution within any state  $i$  at time  $t$  can be approximated by the steady-state probability of state  $i$  (Kramers’ assumption), the transition probability from state  $i$  to state  $j$  (when  $i \neq j$ ) can be calculated using the method described by Mann and Klemm<sup>38</sup> as

$$T_{ij}(\Delta t) = \frac{1}{Z_i(3^D - 1)} \sum_{x \in i} \sum_{\substack{y \in j \\ y \sim x}} \min(e^{-\beta E(x)}, e^{-\beta E(y)}) \quad (2)$$

where  $Z_i$  is the partition function of state  $i$  given by  $\sum_{x \in i} e^{-\beta E(x)}$ ,  $D$  is the number of dimensions of the landscape,  $x$  and  $y$  are neighbor grid points (denoted as  $y \sim x$ ) that belong to different states (i.e., that belong to the border; see above),  $\beta = 1/k_B T$ , where  $k_B$  is the Boltzmann constant and  $T$  is the absolute temperature, and  $E(x)$  is the energy at grid point  $x$ . Given the previous analysis of the landscape (see above), the transition probability matrix can be easily calculated from it. The self-transition probabilities  $T_{ii}(\Delta t)$  were calculated as  $1 - \sum_{j \neq i} T_{ij}(\Delta t)$ . Using this formalism, we constructed a Markov model from the obtained ILS 3D energy landscape using a cutoff for transitions with a saddle pair with an energy greater than 40 kJ·mol<sup>-1</sup>, meaning that only transitions with a lower energy barrier (higher probability) and only the states that correspond to those transitions were included in the model. Hence, this model excludes low-probability transitions and also “inaccessible” states (with only low-probability transitions leading to them). We further simplified this model by coarse-graining the multiple solvent basins into a single one (see details in the Supporting Information). Then, if we denote the probability of state  $i$  at time  $t$  as  $p_i(t)$ , the time-discrete evolution for the Markov chain is given by

$$p_j(t + \Delta t) = \sum_i p_i(t) T_{ij}(\Delta t) \quad (3)$$



**Figure 3.** (A)  $O_2$  energy landscape obtained by explicit MD simulations (green) and ILS calculations (purple). For each method, three isosurfaces are nested, with  $\Delta G_{\text{wat} \rightarrow \text{prot}}$  levels of 0,  $-5$ , and  $-10 \text{ kJ}\cdot\text{mol}^{-1}$  (going inward). (B)  $O_2$  lowest-energy pathways obtained from the ILS energy landscape. Pathways (tubes) connect minima (spheres), with thickness and radius proportional to the energy (thinner and smaller for higher energy). Energy at the minima and along the pathways follows the color scale in  $\text{kJ}\cdot\text{mol}^{-1}$ . (C) Reactive probability flux of  $O_2$  from the solvent toward the G state. The forward committor (see Methodology) for each state is colored on the minima (spheres) from blue ( $q^+ = 0$ ) to red ( $q^+ = 1$ ). The flux is mapped on each pathway (tubes), with thickness proportional to it (nonexistent for very low fluxes), as well as indicated using a color scale from cyan (low) to magenta (high). (D) Comparison of the Xe-bound MaL structure (PDB code: 3QPK<sup>20</sup>) with the  $O_2$  pathways for CotA laccase. The pathways are represented as in panel B. The  $C^\alpha$ -aligned MaL structure is represented in green cartoon, with the bound xenons represented in teal spheres (labeled as in the original reference). In all panels, the  $C^\alpha$  central structure<sup>37</sup> of CotA is represented as in Figure 1.

The iteration of this Markov chain for  $t \rightarrow \infty$  gives the equilibrium or stationary probability distribution  $\pi_i = p_i(\infty)$ . This stationary probability distribution obeys the invariance relation  $\pi_j = \sum_i \pi_i T_{ij}(\Delta t)$ , and the iteration process from any given starting probability distribution  $p_i(0) \neq \pi_i$  corresponds to a relaxation process toward  $\pi_i$ . In this case, where  $T_{ij}(\Delta t)$  is calculated from eq 2, since the border is the same for any given states pair  $ij$ , the detailed balance relation  $\pi_i T_{ij}(\Delta t) = \pi_j T_{ji}(\Delta t)$  is also verified. In this work, the iteration of the Markov chain (eq 3) was performed from a probability distribution of

$$p_i(0) = \begin{cases} 1 & i = \text{solvent} \\ 0 & i \neq \text{solvent} \end{cases} \quad (4)$$

**Reactive Flux Analysis.** The transition path theory<sup>39,40</sup> was applied to our Markov model in order to statistically characterize the transition pathways and calculate the reactive flux between the solvent state and the G state. Our Markov model (see above) represents the full dynamics of  $O_2$  in solvated CotA laccase. Among the ensemble of trajectories that belong to the full dynamics, there is a subset that leaves the solvent and continues until reaching the catalytic center (or G state), the members of which are called reactive trajectories. Transition path theory gives us the mathematical framework to statistically characterize the reactive trajectories. The probability distribution of the reactive trajectories may be calculated using

the so-called committors (forward and backward). In our case, the forward committor  $q_i^+$  is defined as the probability that the process starting from a state  $i$  will reach first the G state rather than the solvent state,<sup>39</sup> while the backward committor  $q_i^-$  is defined as the probability that the process arriving at state  $i$  came last from the solvent state rather than the G state.<sup>39</sup> The average flux of reactive trajectories  $f_{ij}$  is then given by

$$f_{ij} = \begin{cases} \pi_i q_i^- T_{ij} q_j^+ & i \neq j \\ 0 & i = j \end{cases} \quad (5)$$

The reactive or effective flux  $f_{ij}^+$ , which gives the net average number of reactive trajectories per time unit making a transition from state  $i$  to state  $j$  on their way from the solvent state to G state,<sup>39</sup> is given by

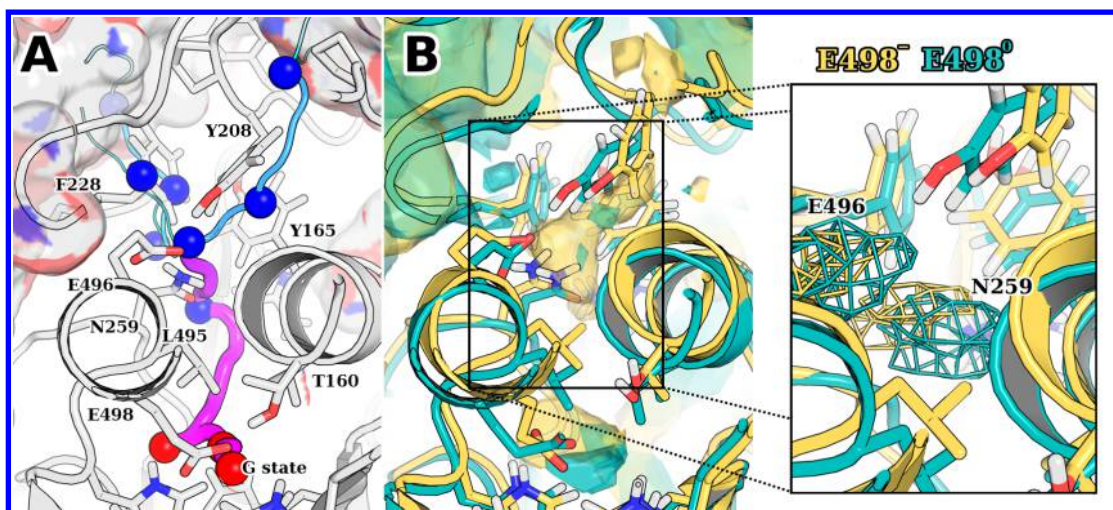
$$f_{ij}^+ = \max(f_{ij} - f_{ji}, 0) \quad (6)$$

The committors and fluxes were calculated using the EMMA software.<sup>41</sup>

## RESULTS AND DISCUSSION

Protonation equilibrium simulations have already been able to predict and provide insight into important residues for protonation events during catalysis by CotA.<sup>10,12</sup> Namely, it was shown that the oxidation state of the T2–T3 center may





**Figure 4.** (A) Highest-flux pathways of O<sub>2</sub> (as in Figure 3C, different perspective). (B) O<sub>2</sub> energy isosurfaces obtained by ILS calculations for Glu498 protonated (E498<sup>0</sup>, teal) and deprotonated (E498<sup>-</sup>, yellow), with a 0 kJ·mol<sup>-1</sup>  $\Delta G_{\text{wat} \rightarrow \text{prot}}$  level for both. The represented structures (cartoon in the same color scheme as the isosurfaces) are the central structures considering the RMSD of Asn259. On the right (inset), Asn259 and Glu496 side-chains are replaced by meshes representing their probability density (probability 0.02).

influence the protonation state of Glu498 (a surrounding carboxylate residue, see above and Figure 1) in CotA laccase,<sup>10</sup> showing a half-protonated state for the O<sub>2</sub>-ready fully reduced state. With this in mind, and in order to better model this oxidation state of CotA, all simulations and analyses were run considering both protonation states of Glu498 (see Methodology), unless stated otherwise.

Standard MD simulations of O<sub>2</sub> diffusion failed to show this molecule to reach the T2–T3 center (data not shown) but revealed some O<sub>2</sub> high-affinity regions inside the protein (green energy isosurfaces, Figure 3A). There is a good correlation between these high-affinity regions and some of the cavities found in the X-ray crystallography structure (Figure 1), except for the  $\beta$ -channel. Since O<sub>2</sub> binds to the T2–T3 center and these diffusion MD simulations lack enough time to sample that event, we turned to ILS calculations,<sup>22</sup> which should provide a more complete characterization of the energy landscape for O<sub>2</sub> diffusion inside CotA laccase. Indeed, more than corroborating the regions found in standard MD, this approach shows high-affinity regions for O<sub>2</sub> close to the T2–T3 center, namely above the T3 coppers where O<sub>2</sub> binds (compare purple to green energy isosurfaces, Figure 3A). Other lower-affinity regions appear on the  $\beta$ -channel, with the Glu498–Thr160 pair placed between them and the center, while some hydrophobic residues, such as Trp151, Val164, Tyr165, and Val169, delimit and separate the higher-affinity regions near T3 $\alpha$  (Figure 3A). Furthermore, regions of high-probability for H<sub>2</sub>O also separate these high-affinity regions for O<sub>2</sub> (compare with Figure S4), hinting at H<sub>2</sub>O playing an actual structural role on O<sub>2</sub> dynamics inside the protein.

By analyzing the topology of the ILS O<sub>2</sub> energy landscape (see Methodology), we find many energy minima and lowest-energy pathways that connect these minima from the solvent to the T2–T3 center. These pathways funnel together as we approach the center in two distinct groups that we designate as  $\alpha$ -pathways and  $\beta$ -pathways, coming from the side of T3 $\alpha$  or T3 $\beta$ , respectively (Figure 3B). Both groups of pathways lead to a minimum right above the center (G state, Figure 3B), which is the only one with a low-energy pathway to the O<sub>2</sub>-binding minimum (between T3 coppers). This pathway from the G

state to the center, although indicative, corresponds to a chemical binding process which cannot be fully grasped by the molecular mechanics force-field employed. Therefore, we ignored this single pathway and assumed that reaching the G state is a necessary (although possibly not sufficient) condition for catalysis, by the fact that there are no other low-energy pathways. The tuning of the T2–T3 center toward irreversible catalysis<sup>5</sup> also supports this idea of a reaction driven very favorably toward O<sub>2</sub> binding. With this reasonable assumption, a kinetic Markov model for O<sub>2</sub> diffusion was constructed from the analyzed energy landscape (see Methodology). Unfortunately, it is hard to validate this kinetic model for CotA laccase as there are no kinetic data for O<sub>2</sub> binding solely. Nonetheless, it is possible to test the consistency of the kinetic model internally. After  $2 \times 10^7$  steps, the Markov chain had already converged ( $\sum_i dp_i/dt \approx 0$ ) and the probability distribution at this time  $p_i(\infty) \approx \pi_i(\text{Markov})$  had perfect correlation with the probability distribution of the ILS energy landscape itself  $\pi_i(\text{ILS})$  (Figure S5). This shows that the calculated transition probabilities are able to reproduce the equilibrium distribution of O<sub>2</sub> inside CotA, serving as an accuracy assessment of the energy barriers themselves. In fact, not only the barriers but also the entire ILS energy landscape appears to be fairly converged (Figure S6 and S7, Movie S1), especially when analyzed in a time-ordered way (Figure S6). Furthermore, it is clear that the use of multiple MD trajectories is instrumental for a more efficient definition of the energy landscape—e.g., the resemblance with the final landscape is already much higher in panel 04 of Figure S6 (with 30 trajectories of around 133 ps) than in panel 04 of Figure S7 (with two trajectories of 2000 ps).

Validation and convergence aside, this kinetic model reveals that most (99.98%) of the reactive probability flux of O<sub>2</sub> from the solvent toward the G state passes through  $\alpha$ -pathways and, among them, mainly (93.71%) siphoning through a group of minima above the packed  $\alpha$ -helices ( $\alpha 3$  and  $\alpha 10$ ) and coalescing into a single pathway as it approaches the G state (Figure 3C). This is, to our knowledge, a new pathway for O<sub>2</sub> diffusion in MCOs, and its higher hydrophobic nature makes it more reasonable when compared with the generally assumed hydrophilic  $\beta$ -channel. Furthermore, it hints at a pathway

capable of explaining the arrival of O<sub>2</sub> on laccases with the  $\beta$ -channel blocked by the C-terminal region, such as MaL (Figure S8), even though the specific pathway (energies and fluxes) would require similar simulations on such laccases. On the basis of the Xe-bound MaL laccase structure, it was proposed that O<sub>2</sub> could reach the T2–T3 center following the Xe3→Xe1→Xe2 path (Xe positions can be verified in Figure 3D).<sup>20</sup> The O<sub>2</sub> pathways determined for CotA laccase show that O<sub>2</sub> can reach the Xe1–Xe2 region directly from the solvent, with no low-energy pathways leading from there to the center (Figure 3D). While Xe3 lays at the protein surface (seemingly stabilized within a loop involved in crystal contacts), Xe1 and Xe2 are  $\beta$ -sandwiched deep within the protein, and their positions agree with our calculations, with some O<sub>2</sub> energy minima only slightly shifted from them, probably due to the structural differences between the proteins in that zone (Figure S9). This is predicted even if the starting CotA structure does not present cavities in that region and provides a good comparison with experimentally derived data.

Focusing again on the newfound pathway (Figure 4A), it has a Y shape, with a distance of  $\sim 36.2$  Å or  $\sim 40.7$  Å from the solvent to the G state, as calculated from the steepest-descent pathways. Its bifurcated entry is surrounded mainly by the hydrophobic aromatic residues Tyr165, Tyr208, and Phe228, with the multiple minima found there being probably due to the torsional diversity of Phe228. These residues provide some hydrophobic shielding from the solvent, but the residues following along the pathway are, in contrast, hydrophilic residues Arg198 and Asn259, which are solvated (Figure S4) and probably have a high influence on O<sub>2</sub> arriving to the G state. Additionally, the gating provided by these residues seems to be dependent on the protonation state of Glu498, with the energy landscapes for each protonation state showing clear differences in that zone (Figure 4B). The change in Glu498 protonation state affects the hydrogen bonding of Glu496 with Asn259 and Tyr208 (Figure S10), as well as the position and dynamics of the Glu496–Asn259 pair (central positions in Figure 4B and the meshes of the inset, respectively). Furthermore, the binding zone of the G state is similarly affected by Glu498 (Figure 4B).

## CONCLUDING REMARKS

The newfound pathway for O<sub>2</sub> in CotA laccase, together with the dependence of O<sub>2</sub> diffusion on the protonation state of Glu498, brings new questions for MCOs. More than explaining O<sub>2</sub> diffusion on C-terminal plugged laccases, is this pathway conserved within the MCO family? And since Glu498 is mostly structurally conserved among MCOs (or at least conserved on the region above the T2–T3 center), does it have a general role on the diffusion and binding of O<sub>2</sub> on MCOs? Glu498 (or equivalents) may indeed have multiple roles on the catalytic cycle of MCOs. Furthermore, the combination of the methodologies used herein opens new doors for tackling, for example, the H<sub>2</sub>O dynamics within the protein, both as a reaction product and as a structural barrier for O<sub>2</sub> diffusion. As suggested quite recently,<sup>42</sup> methodologies similar to the one presented herein can indeed be also applied to the study of halides and hydroxides on the inhibition of MCOs. This, together with extending these methodologies to the use of constant-pH MD<sup>43</sup> or even QM/MM, may in the future bring further insights on the molecular details of catalysis on MCOs, possibly leading to improvements for biotechnological applications.

## ASSOCIATED CONTENT

### Supporting Information

Preparation of the systems, copper centers parametrization, MD simulations details, implementation of ILS on GROMACS, conversion to the  $\Delta G_{\text{wat} \rightarrow \text{prot}}(\text{O}_2)$  landscape, coarse-graining of the solvent states, supplementary results and references, and an mpg file. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: [jmdamas@itqb.unl.pt](mailto:jmdamas@itqb.unl.pt)

\*E-mail: [claudio@itqb.unl.pt](mailto:claudio@itqb.unl.pt)

### Author Contributions

The manuscript was written through contributions of all authors.

### Funding

This work was supported by Fundação para a Ciência e Tecnologia (FCT, grant PEst-OE/EQB/LA0004/2011) and by the European Union (grant FP6-2004-NMP-NI-4/026456, BIORENEW). J.M.D. is also supported by FCT through a Ph.D. fellowship (SFRH/BD/41316/2007).

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Dr. Sara R. R. Campos for assistance with programming and Dr. Bruno L. Victor for helpful discussions on O<sub>2</sub> parametrization and diffusion. Dr. Lígia O. Martins and Dr. Isabel Bento are also acknowledged for helpful discussions.

## ABBREVIATIONS

MCO, multicopper oxidase; O<sub>2</sub>, dioxygen; MD, molecular dynamics; ILS, implicit ligand sampling; E498<sup>−</sup>, CotA with unprotonated Glu498; E498<sup>0</sup>, CotA with protonated Glu498

## REFERENCES

- (1) Kunamneni, A.; Plou, F. J.; Ballesteros, A.; Alcalde, M. Laccases and their applications: a patent review. *Recent Pat. Biotechnol.* **2008**, *2*, 10–24.
- (2) Solomon, E. I.; Sundaram, U. M.; Machonkin, T. E. Multicopper oxidases and oxygenases. *Chem. Rev. (Washington, DC, U. S.)* **1996**, *96*, 2563–2605.
- (3) Marshall, N. M.; Garner, D. K.; Wilson, T. D.; Gao, Y.-G.; Robinson, H.; Nilges, M. J.; Lu, Y. Rationally tuning the reduction potential of a single cupredoxin beyond the natural range. *Nature* **2009**, *462*, 113–U127.
- (4) Gupta, N.; Farinas, E. T. Narrowing Laccase Substrate Specificity Using Active Site Saturation Mutagenesis. *Comb. Chem. High Throughput Screening* **2009**, *12*, 269–274.
- (5) Yoon, J.; Fujii, S.; Solomon, E. I. Geometric and electronic structure differences between the type 3 copper sites of the multicopper oxidases and hemocyanin/tyrosinase. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 6585–6590.
- (6) Quintanar, L.; Yoon, J. J.; Aznar, C. P.; Palmer, A. E.; Andersson, K. K.; Britt, R. D.; Solomon, E. I. Spectroscopic and electronic structure studies of the trinuclear Cu cluster active site of the multicopper oxidase laccase: Nature of its coordination unsaturation. *J. Am. Chem. Soc.* **2005**, *127*, 13832–13845.
- (7) Quintanar, L.; Stoj, C.; Wang, T. P.; Kosman, D. J.; Solomon, E. J. Role of aspartate 94 in the decay of the peroxide intermediate in the multicopper oxidase Fet3p. *Biochemistry* **2005**, *44*, 6081–6091.
- (8) Ueki, Y.; Inoue, M.; Kurose, S.; Kataoka, K.; Sakurai, T. Mutations at Asp112 adjacent to the trinuclear Cu center in CueO as



the proton donor in the four-electron reduction of dioxygen. *FEBS Lett.* **2006**, *580*, 4069–4072.

(9) Augustine, A. J.; Quintanar, L.; Stoj, C. S.; Kosman, D. J.; Solomon, E. I. Spectroscopic and kinetic studies of perturbed trinuclear copper clusters: The role of protons in reductive cleavage of the O-O bond in the multicopper oxidase Fet3p. *J. Am. Chem. Soc.* **2007**, *129*, 13118–13126.

(10) Bento, I.; Silva, C. S.; Chen, Z.; Martins, L. O.; Lindley, P. F.; Soares, C. M., Mechanisms underlying dioxygen reduction in laccases. Structural and modelling studies focusing on proton transfer. *BMC Struct. Biol.* **2010**, *10*.

(11) Chen, Z. J.; Durao, P.; Silva, C. S.; Pereira, M. M.; Todorovic, S.; Hildebrandt, P.; Bento, I.; Lindley, P. F.; Martins, L. O. The role of Glu(498) in the dioxygen reactivity of CotA-laccase from *Bacillus subtilis*. *Dalton Trans.* **2010**, *39*, 2875–2882.

(12) Silva, C. S.; Damas, J. M.; Chen, Z. J.; Brissos, V.; Martins, L. O.; Soares, C. M.; Lindley, P. F.; Bento, I. The role of Asp116 in the reductive cleavage of dioxygen to water in CotA laccase: assistance during the proton-transfer mechanism. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 186–193.

(13) Brissos, V.; Chen, Z.; Martins, L. O. The kinetic role of carboxylate residues in the proximity of the trinuclear centre in the O-2 reactivity of CotA-laccase. *Dalton Trans.* **2012**, *41*, 6247–6255.

(14) Rulisek, L.; Solomon, E. I.; Ryde, U. A combined quantum and molecular mechanical study of the O-2 reductive cleavage in the catalytic cycle of multicopper oxidases. *Inorg. Chem.* **2005**, *44*, 5612–5628.

(15) Solomon, E. I.; Augustine, A. J.; Yoon, J. O(2) Reduction to H(2)O by the multicopper oxidases. *Dalton Trans.* **2008**, 3921–3932.

(16) Rulisek, L.; Ryde, U. Theoretical studies of the active-site structure, spectroscopic and thermodynamic properties, and reaction mechanism of multicopper oxidases. *Coord. Chem. Rev.* **2013**, *257*, 445–458.

(17) Bento, I.; Carrondo, M. A.; Lindley, P. F. Reduction of dioxygen by enzymes containing copper. *J. Biol. Inorg. Chem.* **2006**, *11*, 539–547.

(18) Wang, P.-h.; Bruschi, M.; De Gioia, L.; Blumberger, J. Uncovering a Dynamically Formed Substrate Access Tunnel in Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase. *J. Am. Chem. Soc.* **2013**, *135*, 9493–9502.

(19) Zhukhlistova, N. E.; Zhukova, Y. N.; Lyashenko, A. V.; Zaitsev, V. N.; Mikhailov, A. M. Three-dimensional organization of three-domain copper oxidases: A review. *Crystallogr. Rep.* **2008**, *53*, 92–109.

(20) Kallio, J. P.; Rouvinen, J.; Kruus, K.; Hakulinen, N. Probing the Dioxygen Route in *Melanocarpus albomyces* Laccase with Pressurized Xenon Gas. *Biochemistry* **2011**, *50*, 4396–4398.

(21) Baron, R.; McCammon, J. A.; Mattevi, A. The oxygen-binding vs. oxygen-consuming paradigm in biocatalysis: structural biology and biomolecular simulation. *Curr. Opin. Struct. Biol.* **2009**, *19*, 672–679.

(22) Cohen, J.; Arkhipov, A.; Braun, R.; Schulten, K. Imaging the migration pathways for O<sub>2</sub>, CO, NO, and Xe inside myoglobin. *Biophys. J.* **2006**, *91*, 1844–1857.

(23) Saam, J.; Ivanov, I.; Walther, M.; Holzhuetter, H.-G.; Kuhn, H. Molecular dioxygen enters the active site of 12/15-lipoxygenase via dynamic oxygen access channels. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 13319–13324.

(24) Maragliano, L.; Cottone, G.; Ciccotti, G.; Vanden-Eijnden, E. Mapping the Network of Pathways of CO Diffusion in Myoglobin. *J. Am. Chem. Soc.* **2010**, *132*, 1010–1017.

(25) Wang, P.-h.; Best, R. B.; Blumberger, J. Multiscale Simulation Reveals Multiple Pathways for H-2 and O-2 Transport in a NiFe-Hydrogenase. *J. Am. Chem. Soc.* **2011**, *133*, 3548–3556.

(26) Wang, P.-H.; Blumberger, J. Mechanistic insight into the blocking of CO diffusion in NiFe-hydrogenase mutants through multiscale simulation. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6399–6404.

(27) Pavlova, M.; Klvana, M.; Prokop, Z.; Chaloupkova, R.; Banas, P.; Otyepka, M.; Wade, R. C.; Tsuda, M.; Nagata, Y.; Damborsky, J.

Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate. *Nat. Chem. Biol.* **2009**, *5*, 727–733.

(28) Baptista, A. M.; Soares, C. M. Some theoretical and computational aspects of the inclusion of proton isomerism in the protonation equilibrium of proteins. *J. Phys. Chem. B* **2001**, *105*, 293–309.

(29) Teixeira, V. H.; Soares, C. M.; Baptista, A. M. Studies of the reduction and protonation behavior of tetraheme cytochromes using atomic detail. *J. Biol. Inorg. Chem.* **2002**, *7*, 200–216.

(30) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J. Biophys. Lett.* **2011**, *40*, 843–856.

(31) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces*, Pullman, B., Ed.; D. Reidel: Dordrecht, Holland, 1981; pp 331–342.

(32) Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845–854.

(33) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G.; Smooth, A. Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(34) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(35) Volkers, G.; Damas, J. M.; Palm, G. J.; Panjikar, S.; Soares, C. M.; Hinrichs, W. Putative dioxygen-binding sites and recognition of tigecycline and minocycline in the tetracycline-degrading monooxygenase TetX. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1758–1767.

(36) Victor, B.; Baptista, A.; Soares, C. Dioxygen and nitric oxide pathways and affinity to the catalytic site of rubredoxin:oxygen oxidoreductase from *Desulfovibrio gigas*. *J. Biol. Inorg. Chem.* **2009**, *14*, 853–862.

(37) Campos, S. R. R.; Baptista, A. M. Conformational Analysis in a Multidimensional Energy Landscape: Study of an Arginylglutamate Repeat. *J. Phys. Chem. B* **2009**, *113*, 15989–16001.

(38) Mann, M.; Klemm, K. Efficient exploration of discrete energy landscapes. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2011**, *83*, 011113.

(39) Metzner, P.; Schuette, C.; Vanden-Eijnden, E. Transition Path Theory for Markov Jump Processes. *Multiscale Model. Simul.* **2009**, *7*, 1192–1219.

(40) E, W.; Vanden-Eijnden, E. Towards a theory of transition paths. *J. Stat. Phys.* **2006**, *123*, 503–523.

(41) Senne, M.; Trendelkamp-Schroer, B.; Mey, A. S. J. S.; Schuette, C.; Noe, F. EMMA: A Software Package for Markov Model Building and Analysis. *J. Chem. Theory Comput.* **2012**, *8*, 2223–2238.

(42) Mate, D. M.; Gonzalez-Perez, D.; Falk, M.; Kittl, R.; Pita, M.; De Lacey, A. L.; Ludwig, R.; Shleev, S.; Alcalde, M. Blood Tolerant Laccase by Directed Evolution. *Chem. Biol.* **2013**, *20*, 223–231.

(43) Baptista, A. M.; Teixeira, V. H.; Soares, C. M. Constant-pH molecular dynamics using stochastic titration. *J. Chem. Phys.* **2002**, *117*, 4184–4200.