

SCIENTIFIC REPORTS



OPEN

Computational insights for the hydride transfer and distinctive roles of key residues in cholesterol oxidase

Li-Juan Yu¹, Emily Golden¹, Nanhao Chen^{3,4}, Yuan Zhao², Alice Vrielink¹ & Amir Karton¹

Cholesterol oxidase (ChOx), a member of the glucose-methanol-choline (GMC) family, catalyzes the oxidation of the substrate *via* a hydride transfer mechanism and concomitant reduction of the FAD cofactor. Unlike other GMC enzymes, the conserved His447 is not the catalytic base that deprotonates the substrate in ChOx. Our QM/MM MD simulations indicate that the Glu361 residue acts as a catalytic base facilitating the hydride transfer from the substrate to the cofactor. We find that two rationally chosen point mutations (His447Gln and His447Asn) cause notable decreases in the catalytic activity. The binding free energy calculations show that the Glu361 and His447 residues are important in substrate binding. We also performed high-level double-hybrid density functional theory simulations using small model systems, which support the QM/MM MD results. Our work provides a basis for unraveling the substrate oxidation mechanism in GMC enzymes in which the conserved histidine does not act as a base.

Cholesterol oxidase (ChOx) is a flavoenzyme that catalyzes the oxidation of cholesterol to 5-cholesten-3-one and the subsequent isomerization to form the final 4-cholesten-3-one product. The enzyme from *Streptomyces* is a member of the glucose-methanol-choline (GMC) family and contains a single flavin adenine dinucleotide (FAD) cofactor that is non-covalently bound to the protein. Apart from its physiological functions in bacterial metabolism, pathogenesis, and macrolide biosynthesis¹, ChOx has also been employed as a useful biotechnological tool, for example, for the determination of serum cholesterol levels^{2,3}.

A number of experimental studies investigated the structure and catalytic mechanism of ChOx over the past two decades^{4–11}. The active site of ChOx is shown in Fig. 1a. With regard to the enzymatic mechanism, Sampson *et al.*^{7–11} identified three residues (Glu361, His447, and Asn485) that are significant for the enzymatic activity and Vrielink *et al.*^{4,5} revealed that Gly120 also plays an important role in the catalytic mechanism.

Mutations of the Glu361 residue in the active site to glutamine and aspartate result in 31-fold⁹ and 14-fold⁸ reductions, respectively, in the k_{cat} value for the oxidation of the substrate compared to the wild-type (WT) enzyme. This suggests that the carboxylic acid residue at this position may be important for the oxidative activity⁶. The His447 residue is located in the active site and the NE2 hydrogen forms a hydrogen bond with the hydroxyl group of the substrate. Mutations of His447 to glutamine and asparagine result in 140-fold and 4400-fold reductions in the k_{cat} value respectively⁹, relative to the WT enzyme, suggesting that His447 plays a significant role in the catalytic activity. However, the double mutant, His447Gln/Glu361Gln, exhibits a 500-fold decrease in the k_{cat} value relative to the WT enzyme¹⁰, which was 3-fold slower than that for the His447Gln single-mutant. These mutagenesis results indicate that factors affecting the oxidation rate of the two separate mutants, His447Gln and Glu361Gln, are not entirely additive and may be the result of structural perturbations rather than the absence of an active site base. We note that the double mutation does not completely prevent the hydride transfer implying that possibly the FAD is sufficiently electrophilic to oxidize the substrate without complete proton abstraction⁶. In addition, other structural studies have suggested that Asn485 forms an N–H••• π electrostatic interaction

¹School of Molecular Sciences, The University of Western Australia, Perth, WA, 6009, Australia. ²The Key Laboratory of Natural Medicine and Immuno-Engineering, Henan University, Kaifeng, 475004, China. ³School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou, 510006, China. ⁴Department of Chemistry, University of California, Davis, California, 95616, United States. Correspondence and requests for materials should be addressed to Y.Z. (email: zhaoyuan@henu.edu.cn) or A.K. (email: amir.karton@uwa.edu.au)

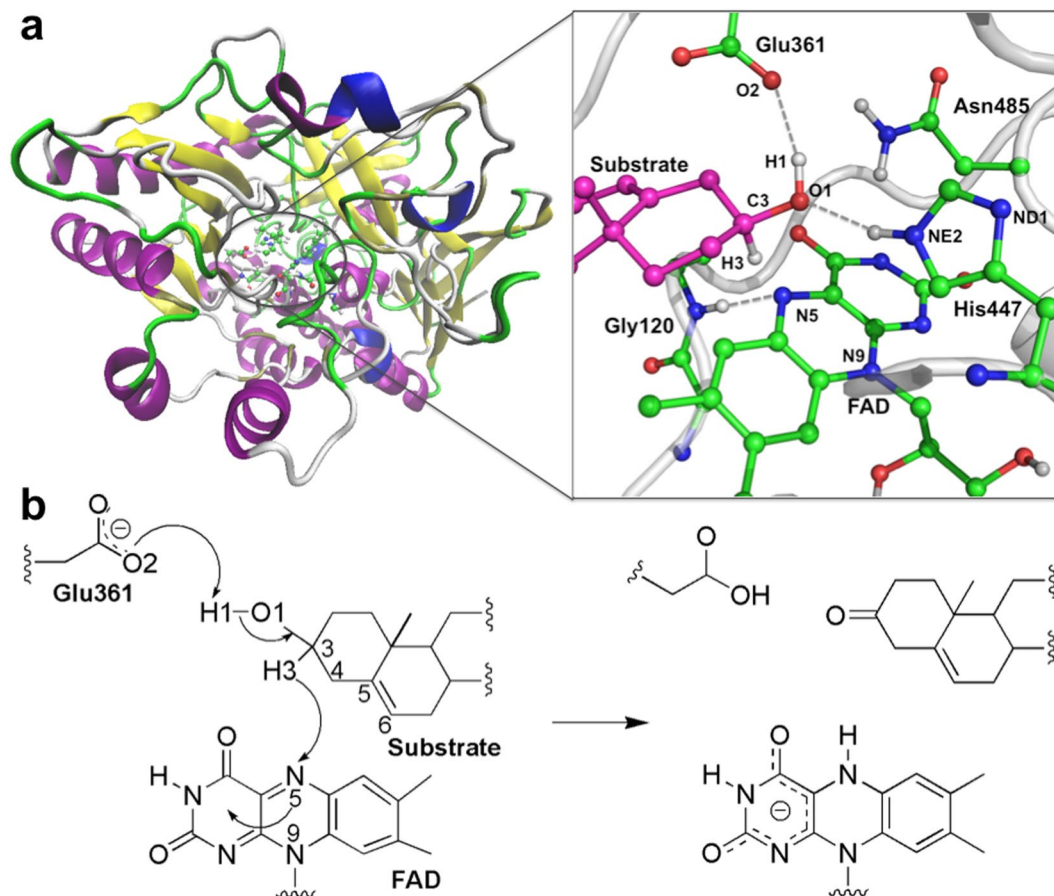


Figure 1. Key residues in the active site of ChOx and proposed reaction mechanism. (a) The substrate is colored in magenta. The grey dashed lines represent H-bond interactions. Atomic color scheme: H, white; C, green; O, red; N, blue. (b) Proposed mechanism for the substrate oxidation reaction catalyzed by ChOx.

with the flavin π -system to facilitate reduction of the FAD¹¹. Mutation of Asn485 to leucine decreases the k_{cat} value as well. It has been proposed that this residue modulates the electrostatic potential of the flavin to enhance substrate oxidation¹¹. The Gly120 residue, which is positioned below the isoalloxazine ring system and functions as a hydrogen bond donor to the N5 atom of FAD, was also proposed to be important in orienting the orbitals of the N5 atom of the FAD cofactor, thus priming it for the hydride transfer chemistry^{4,5}.

The recent high resolution X-ray structure of the oxidized ChOx in the presence of an isopropanol substrate (PDB ID: 4U2T)⁴ enables us to build a model of the ChOx active site composed of Glu361, the FAD cofactor, and the steroid substrate (see Fig. 1b). It was proposed that ChOx catalyzes the oxidation of the substrate *via* a hydride transfer from the steroid substrate C3 atom to the N5 atom of FAD and concomitant reduction of the FAD cofactor^{5,6}. Despite extensive experimental work on the structure of ChOx, a detailed theoretical simulation of the catalytic mechanism and roles of the key residues in the active site was not carried out. In this work, we performed extensive combined quantum mechanics and molecular mechanics molecular dynamics (QM/MM MD) simulations including the catalytic mechanism and substrate binding, and high-level double-hybrid density functional theory (DHDFT) calculations to explore the enzymatic catalysis. The results from the DHDFT calculations using small model systems support the QM/MM MD results. The combination of these two approaches gives an in-depth understanding of flavoenzyme catalysis by ChOx.

Results

QM/MM MD simulations on enzymatic mechanism. *Catalytic mechanism of ChOx.* 10 ns molecular dynamics (MD) simulations for the enzyme-substrate complex model were performed. The root-mean-square deviation (RMSD) of the protein backbone is indicated in Supplementary Fig. S1. The last snapshot of MD simulations was chosen as the initial QM/MM model by removing the water molecules beyond 30 Å from the N9 atom of the FAD cofactor (see Fig. 2). We considered two different reaction coordinates (RC1 and RC2, see Supplementary Fig. S2). The reaction coordinate RC1 is defined as the difference between the C3–H3 and N5–H3 bond distances that are involved in the hydride transfer: $\text{RC1} = d_{\text{C3-H3}} - d_{\text{N5-H3}}$. The reaction coordinate RC2 is defined as $\text{RC2} = d_{\text{C3-H3}} - d_{\text{N5-H3}} - d_{\text{O2-H1}}$, that is it describes the movement of both the hydride and the proton. The relative energy barrier heights for RC1 and RC2 are 130.6 and 125.4 kJ mol⁻¹, respectively. Another way of describing the combined movement of both the hydride and the proton is to fix the length of the O1–H1 distance at certain values and to scan along the reaction coordinate RC1. In particular, the O1–H1 distance was fixed at

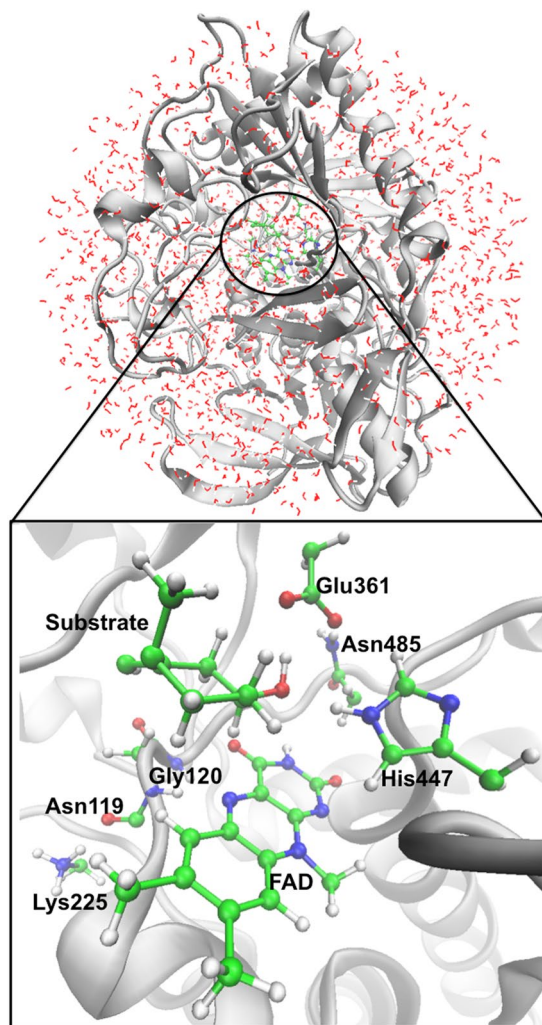


Figure 2. Overall view of the QM/MM model for ChOx and the active site. The active site includes the Lys225, Asn119, Gly120, Glu361, Asn485, and His447 residues as well as FAD and the substrate. The residues in the QM region are shown as ball and stick, the protein as ribbons, and the solvent as wireframes. The QM/MM boundary settings are specified in the subsection of “QM/MM MD simulation”.

1.01, 1.06, 1.11, 1.16, 1.21, 1.26, 1.31, 1.36, 1.41, 1.46, 1.51, 1.56, 1.61, 1.66, 1.71, and 1.76 Å. This is a cost-effective way to approximate a 2D scan of both the hydride and proton transfers. This approximated 2D scan is shown in the Supplementary Fig. S3a. It clearly shows that a first-order saddle point is obtained for the hydride transfer for each of the fixed O1–H1 bond distances. We find that the lowest reaction barrier height of $128.4 \text{ kJ mol}^{-1}$ is obtained at $r(\text{O1–H1}) \approx 1.56 \text{ Å}$ (Supplementary Fig. S3b). This result is consistent with our results for RC1 and RC2 above. Considering the computational cost of the potential of mean force (PMF) calculations, we chose RC1 as our final reaction coordinate since the reaction barrier height obtained along this coordinate is in good agreement with our best reaction barrier height obtained from the approximate 2D scan. By using QM/MM MD simulations and umbrella sampling, the free energy profile and corresponding configurations of reactant, transition state, and product involved in the enzymatic catalysis have been determined. In order to verify the efficiency of our sampling, the configurational overlap in the range of 5 ps to 20 ps is shown in Supplementary Fig. S4. This figure shows that the sampling covers the whole reaction coordinate with the current step size and biasing harmonic potential.

Figure 3 shows the free energy profile and the structures of the reactant (**React**), transition structure (**TS**), and product (**Prod**) obtained from the umbrella sampling technique. It is noteworthy that the deprotonation of the substrate by Glu361 occurs synchronously with the hydride transfer (see **TS** in Fig. 3). The O1–H1 bond distance increases from $1.01 \pm 0.03 \text{ Å}$ (**React**) to $1.90 \pm 0.17 \text{ Å}$ (**Prod**), while the O2–H1 bond distance decreases from $2.10 \pm 0.45 \text{ Å}$ (**React**) to $0.99 \pm 0.01 \text{ Å}$ (**Prod**). The free energy barrier of the hydride transfer is 99.3 kJ mol^{-1} , and the reaction energy is exergonic by 11.8 kJ mol^{-1} . In the QM/MM transition structure, the length of the C3–H3 bond that is being broken is $1.37 \pm 0.07 \text{ Å}$ and the length of the N5–H3 bond that is being formed is $1.35 \pm 0.09 \text{ Å}$. In this transition structure the O1–H1 and O2–H1 bond distances are 2.04 ± 0.30 and $1.00 \pm 0.03 \text{ Å}$, respectively. This indicates that the proton has transferred before the hydride transfer has been completed. We note that we

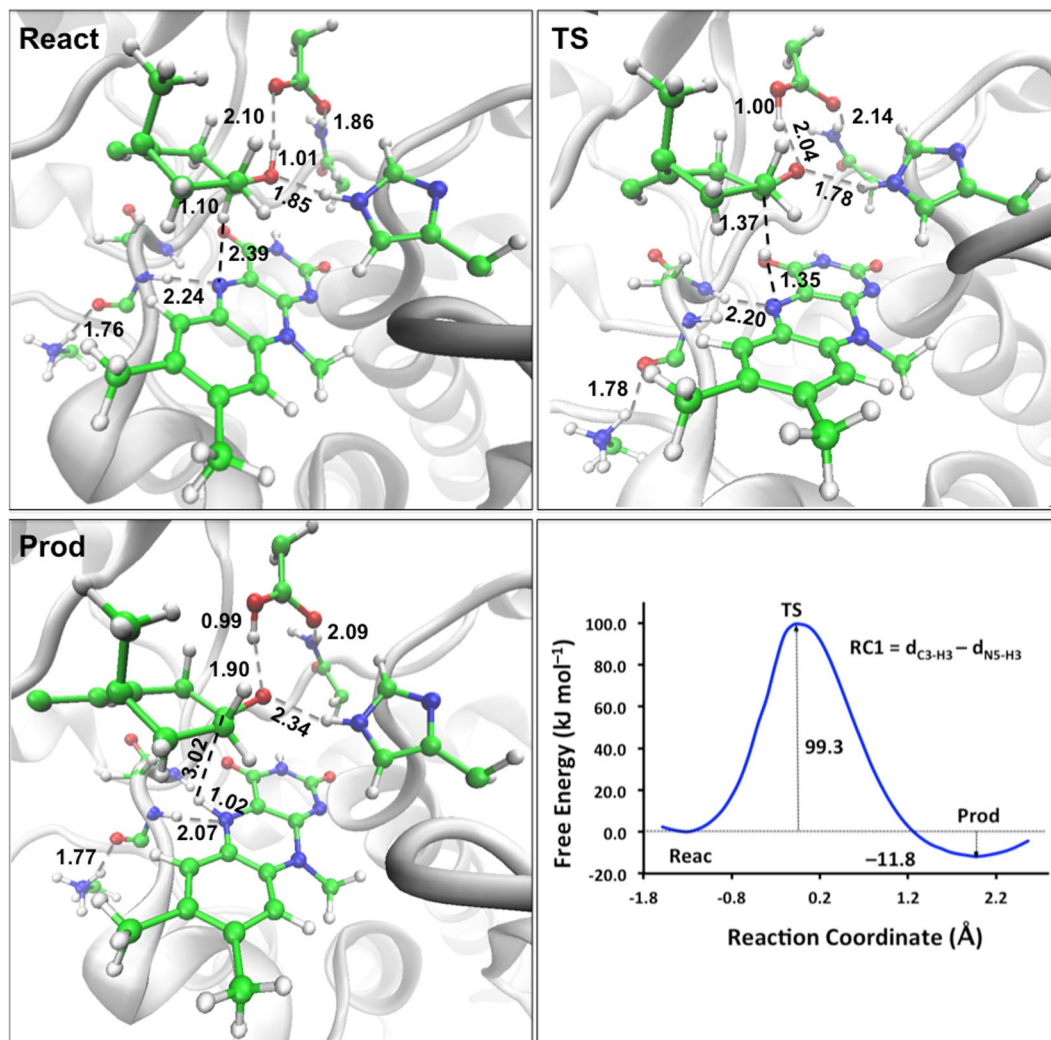


Figure 3. Structures of reactant (**React**), transition state (**TS**), product (**Prod**), and free energy profile of the hydride transfer in ChOx. Hydrogen bonds are shown as grey dashed lines. The bonds being broken and formed are shown as black dashed lines. The average distances (in Å) for selected bonds are taken from umbrella sampling.

cannot directly compare the barrier obtained from the QM/MM MD simulations with previous experimental kinetic data⁹ since the kinetic measurements do not reflect the rate of the hydride transfer alone. However, the experimental k_{cat} value obtained for the wild type ($44 \pm 2 \text{ s}^{-1}$) indicates that our QM/MM MD free energy barrier for the hydride transfer likely represents an overestimation. Nevertheless, it is important to stress that the present work is concerned largely with relative reaction barrier heights (that will benefit from some cancellation of errors) and are likely to be more accurate than the absolute reaction barrier heights.

In order to clarify the order of the proton and hydride transfers in the catalytic process, the entire snapshots along RC1 have been collected. Figure 4 plots the snapshots for the C3–H3, N5–H3, O1–H1, and O2–H1 distances along the reaction coordinate RC1 (between the interval $\text{RC1} = -1.59$ to 2.61 \AA). As shown in Fig. 4a, the length of the C3–H3 distance increases smoothly from ~ 1.1 to 3.8 \AA , while the length of the N5–H3 decreases from ~ 2.9 to 1.0 \AA along RC1. The cross-point of these two lines corresponds to the **TS** shown in Fig. 3. At this point, the C3–H3 and N5–H3 distances are equal to one another (i.e., $\text{RC1} = 0.0 \text{ \AA}$). Figure 4b shows that the proton transfer from the hydroxyl group of the substrate to the oxygen (O2) of Glu361 occurs at $\text{RC1} \approx -0.2 \text{ \AA}$. More specifically, the O1–H1 bond length remains at $\sim 1.0 \text{ \AA}$ and the O2•••H1 distance fluctuates at $\sim 1.8 \text{ \AA}$ along RC1 from around $\text{RC1} = -1.7$ to -0.2 \AA . The atom H1 jumps to O2 at $\text{RC1} \approx -0.2 \text{ \AA}$, after which the O2–H1 bond length stays at $\sim 1.0 \text{ \AA}$ and the O1•••H1 length fluctuates at $\sim 2.3 \text{ \AA}$. These results indicate that the proton transfer occurs spontaneously with the hydride transfer.

The bond lengths of His447NE2H–O1 and LysNH₃–O were also monitored during the hydride transfer (see Supplementary Fig. S5). The hydrogen bond distance between the NH₃ group of Lys225 and the O=C group of Asn119 (labeled as LysNH₃–O in Supplementary Fig. S5) fluctuates at $\sim 1.8 \text{ \AA}$ during the hydride transfer, which implies that this hydrogen bond has a negligible effect on the barrier height for the hydride transfer. As for the hydrogen bond distance between His447 and the hydroxyl group of the substrate, labeled as NE2H•••O1, it is

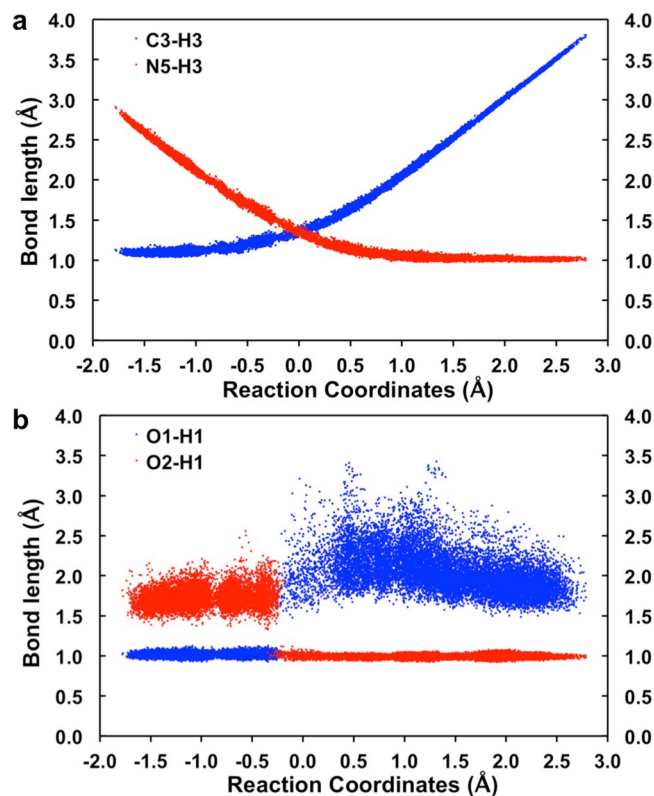


Figure 4. Statistics for bond lengths from the snapshots along RC1. (a) the C3–H3 and N5–H3 (b) the O1–H1 and O2–H1 bond lengths.

~ 1.8 Å before the proton transfer ($RC1 = -0.2$ Å) and increases to ~ 2.6 Å after the proton transfer. This indicates that the His447 plays a significant role during the oxidation reaction. The hydrogen bond formed between His447 and the hydroxyl group of the substrate orients the hydroxyl H atom towards the Glu361 base, which facilitates the proton transfer. This also results in an energetically favorable *trans* arrangement between the substrate hydroxyl H atom and the hydride.

The effects of mutating His447 on the active site environment. Previous experimental point mutations revealed that the His447 residue plays an important role in the oxidation reaction of the substrate in ChOx⁹. For example, mutations of the His447 residue to glutamine and asparagine result in 140-fold and 4400-fold reductions, respectively, in the k_{cat} value for the oxidation reaction⁹. However, the atomic resolution crystal structure showed that the NE2 atom of His447 was protonated⁶. This led to the suggestion that His447 is not a catalytic base. To further elucidate the role of His447, QM/MM calculations have been performed on systems in which the His447 residue has been mutated to glutamine and asparagine (denoted by His447Gln and His447Asn, respectively). In these computational experiments, we carried out QM/MM energy scans along the reaction coordinate of RC1 using the QM/MM optimized structures to obtain the minimum energy reaction path (see Supplementary Fig. S6). However for reasons of computational cost, the subsequent QM/MM free energy perturbation and 20 ps QM/MM MD simulations with umbrella sampling were not performed in this section. Figure 5 gives the TSs of the WT, His447Gln, and His447Asn systems. The energy barrier height for the hydride transfer in the mutated systems is higher by 32.8 (His447Gln) and 66.3 kJ mol⁻¹ (His447Asn) relative to that in the WT system (as shown in Supplementary Fig. S6). These QM/MM scan results are consistent with the mutagenesis experiments, where mutation of His447 to asparagine results in a larger reduction in the k_{cat} value compared to the glutamine mutant.

The selected bond distances for the TSs of the WT, His447Gln, and His447Asn systems are also labeled in Fig. 5. In the QM/MM-optimized structures the hydrogen bond distances between O1 of the substrate and NE2H of His447, His447Gln, and His447Asn are 1.76 (WT), 1.87 (His447Gln), and 3.72 Å (His447Asn). As we discussed earlier, the hydrogen bond formed between His447 and the hydroxyl group of the substrate facilitates the proton transfer. When the His447 residue is mutated to asparagine, due to the shorter side chain in asparagine relative to histidine, a hydrogen bond to the substrate is not formed (i.e., the distance between O1 and NE2H is 3.72 Å). These results indicate that the weak hydrogen bond or lack of hydrogen bond interactions between the substrate and the key residue (i.e. His447) may impede the proton transfer that facilitates the hydride transfer in the WT enzyme, and further increases the barrier height for the hydride transfer. This could be due to the fact that the OH group of the substrate is not as well fixed in the active site and the catalytic environment has changed in the mutagenesis systems (i.e. His447Gln and His447Asn).

In order to consider the protein/water environment for the His447Gln mutant system, we carried out 20 ns MD and QM/MM MD simulations followed by a potential energy scan. Firstly, 20 ns MD simulations were

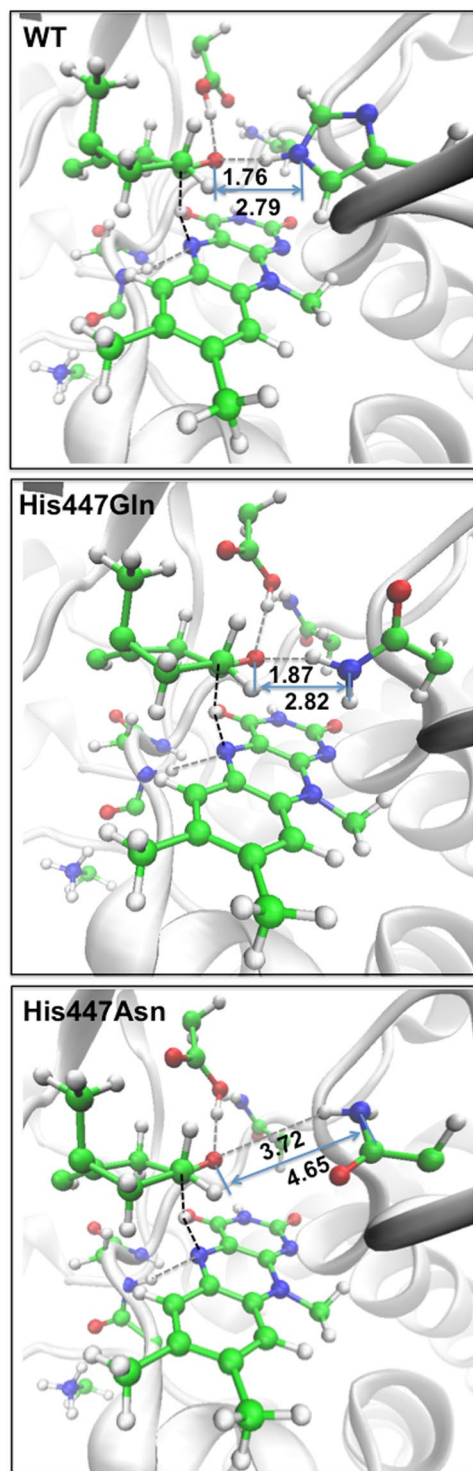


Figure 5. Transition structures for the WT, His447Gln, and His447Asn systems obtained by means of QM/MM scans. Hydrogen bonds are shown as grey dashed lines. The bonds being broken and formed are shown as black dashed lines. The distances for selected bonds are given in Å.

performed for the His447Gln mutant and a stable mutant was obtained. The RMSD of the protein backbone is shown in Supplementary Fig. S8a. Then the QM region was chosen similarly to that of the wild-type and optimized at B3LYP/6–31 G(d) level of theory. In order to obtain a stable protein/water system, before running the potential energy surface scan, we ran 5 ps QM/MM MD for the His447Gln mutant system. Then we carried out the relative energy scans using the different reaction coordinates (RC1 and RC2, *vide supra*). The scan plots are displayed in Supplementary Fig. S9. The relative energy barrier heights obtained from RC1 and RC2 are 158.7 and 162.8 kJ mol⁻¹ respectively, which are in good agreement with the one shown in Supplementary Fig. S6. We

System	ΔG_{bind}^a	$\Delta\Delta G_{\text{bind}}$
WT	-62.5	0.0
Glu361Ala	-27.0	-35.8 ± 6.3
His447Ala	-21.3	-41.3 ± 6.4
Asn485Ala	-57.7	-4.8 ± 1.5

Table 1. Gibbs binding free energies (ΔG_{bind} , in kJ mol^{-1}) for the WT, Glu361Ala, His447Ala, and Asn485Ala systems as well as the differences between the WT and mutant systems ($\Delta\Delta G_{\text{bind}}$) calculated with the MM-GBSA method. ^aThe Gibbs binding free energy (ΔG_{bind}) contains electrostatic, van der Waals, polar, non-polar, and entropy contributions.

believe that these relative energy barrier heights demonstrate the important role that His447 plays in the enzyme. We note that the main purpose of this work is to show that the predicted theoretical trend of the barrier height when mutating His447 to Gln is consistent with the experimental reduction of the k_{cat} value, rather than to reproduce the experimental k_{cat} value, which would require extensive additional PMF calculations.

Similar to our approach with the wild-type, we confirm that the chosen reaction coordinates (RC1 and RC2) give results that are consistent with the results of the more rigorous approximated 2D scan for the His447Gln system. In particular, the O1-H1 distance was fixed at 0.98, 1.03, 1.08, 1.13, 1.18, 1.23, 1.28, 1.33, 1.38, 1.43, 1.48, 1.53, 1.58, 1.63, 1.68, 1.73, 1.78, 1.83, 1.88, 1.93, 1.98, 2.03, 2.08, 2.13, and 2.18 Å for each scan of the hydride transfer reaction coordinate. The 2D scan is given as Supplementary Information (Supplementary Fig. S11a). This figure shows that a first-order saddle point is obtained for the hydride transfer for each of the fixed O1-H1 bond distances. We find that the lowest reaction barrier height of $156.8 \text{ kJ mol}^{-1}$ is obtained at $r(\text{O1-H1}) \approx 1.63 \text{ Å}$ (Supplementary Fig. S11b). This result is consistent with our results for RC1 and RC2 above. Thus, we conclude that both the RC1 and RC2 reaction coordinates can efficiently describe the hydride transfer as well as the proton transfer. Considering the computational cost of the PMF calculations and the main purpose of this work, we chose RC1 as our final reaction coordinate for the potential energy scan.

The role of key residues in substrate binding. It has been proposed that the hydrogen-bonding networks involving Gly120, Glu361, His447, Asn485, and the substrate play important roles in substrate binding^{4,8,9,11,12}. It is interesting to evaluate the binding free energy and the interaction between the substrate and the key residues. The ligand-residue interaction for residues from 101 to 499 in ChOx and binding free energy decomposition in the active site are displayed in Supplementary Fig. S7. We note that Gly120, Glu361, His447, and Asn485 play important roles in stabilizing the substrate in the active site (see Supplementary Fig. S7a). It is noteworthy that other residues also play significant roles in the binding free energy, e.g., Met114, Met122, Val250, Pro364, Tyr446, Pro448, and Phe487. For example, they are associated with binding free energies ranging from -5.2 (Pro364) to $-19.0 \text{ kJ mol}^{-1}$ (Pro448). These residues are involved in the surrounding hydrogen bond network of the FAD cofactor and stabilize FAD in the active site. However, they are not directly associated with stabilizing the substrate in the active site.

In order to get additional insights into individual contributions from the Glu361, His447, and Asn485 residues to the substrate binding, the binding free energies were calculated for the WT and the Glu361Ala, His447Ala, and Asn485Ala mutant systems using the MM-GBSA method^{13,14}. The Gibbs free binding energies for the WT and mutant systems as well as the differences between them are listed in Table 1. The Gibbs free binding energies for the mutants are lower than that for the WT by: 35.8 (Glu361Ala), 41.3 (His447Ala), and 4.8 kJ mol^{-1} (Asn485Ala). Mutations of these key residues result in destabilization of the substrate in the reaction center. Thus, the catalytic process becomes less efficient. This is another indication of the significant roles of these key residues in stabilizing the substrate in the active site. In order to obtain further insights into the roles of key residues in the enzymatic catalysis we also carried out double-hybrid density functional theory (DFT) calculations, which will be discussed in the following section.

Double-hybrid DFT calculations for probing the effects of hydrogen-bonding interactions on the barrier height of the hydride transfer.

Double-hybrid density functional theory (DHDFT) calculations were performed in order to obtain insights into the roles of the residues in the catalytic mechanism of the hydride transfer from the C3 atom of the substrate to the N5 atom of the FAD cofactor. These computational experiments were carried out on a number of model systems with increasing size in order to elucidate the catalytic roles of the Glu361 and Gly120 residues on the barrier height for the hydride transfer. In particular, we consider the following models. Model **A** consists of the FAD cofactor (modeled by an isoalloxazine ring system) and the steroid substrate (modeled by an HOCHMe₂ moiety). Model **B** consists of an additional Glu361, which forms a strong hydrogen bond with the hydroxyl group of the steroid substrate. Model **C** includes an additional Gly120 (modeled by a dimethylamine molecule), which forms a hydrogen bond to the N5 atom of FAD. The atoms involved in these three models are displayed in Supplementary Fig. S12.

Figure 6 depicts the Gibbs free energy profile (ΔG_{298}) for the hydride transfer in models **A** (red), **B** (black), and **C** (blue). The calculated reaction barriers (ΔH_{298}^\ddagger) and enthalpies (ΔH_{298}) are listed in Supplementary Table S1. Figure 7 shows the optimized transition structures (TSs) involved in models **A**–**C**. The corresponding reactant complexes (RCs) and product complexes (PCs) are shown in Supplementary Fig. S13.

Model **A**, which consists of the FAD cofactor and the substrate, is the simplest binary system for modeling the hydride transfer. The hydride transfer in this model is expected to be both kinetically and thermodynamically unfavorable, however we start with this simple picture since it provides a reference point for the effects from the

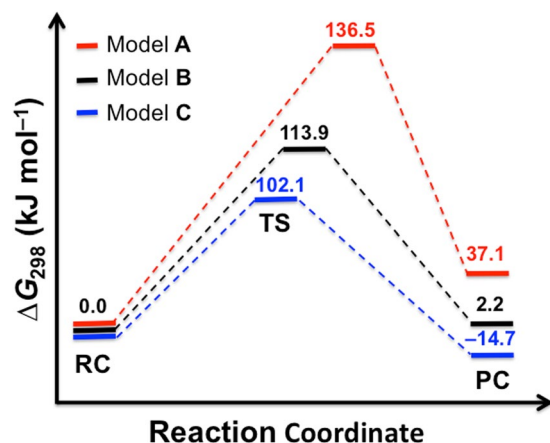


Figure 6. Gibbs free energy profiles for the hydride transfer in models A–C calculated at the RI-B2GP-PLYP-D3/Def2-QZVPP level of theory. Model A includes FAD and the substrate (red line); model B includes FAD, the substrate, and Glu361 (black line); and model C includes FAD, the substrate, Glu361, and Gly120 (blue line). The transition structures (TSs) involved in these reactions are shown in Fig. 7. The corresponding reactant complexes (RCs) and product complexes (PCs) are shown in Supplementary Fig. S13.

additional residues in models B and C. For the hydride transfer in model A, we obtain an activation Gibbs free energy of $\Delta G_{298}^{\ddagger} = 136.5 \text{ kJ mol}^{-1}$ and the reaction is endergonic by 37.1 kJ mol^{-1} . According to the Hammond–Leffler postulate the high reaction barrier and energy suggest a late TS¹⁵. Indeed, in the TS, the hydride transfer between the C3 and N5 atoms is almost complete, the $\text{C3}\cdots\text{H}^-$ and $^- \text{H}\cdots\text{N5}$ distances being 1.772 and 1.067 Å, respectively. For comparison, the $^- \text{H}\cdots\text{N5}$ bond length in the product complex is 1.035 Å (see Supplementary Fig. S13).

The enzymatic role of Glu361 is to act as a base to abstract the proton from the hydroxyl group of the substrate. Abstracting the hydroxyl proton has two effects on the reaction profile for the hydride transfer: (i) it activates the hydride transfer by pushing electron density to the C3–H3 bond and (ii) it stabilizes the product by converting the isopropanol carbocation to a neutral acetone. Model B includes the Glu361 residue. Indeed, the activation free energy in this model is lower by 22.6 kJ mol^{-1} relative to that in model A. In the TS of model B, the hydride transfer from C3 to N5 is concomitant with the proton transfer from the substrate to Glu361 (see model B in Fig. 7). In this TS, the hydride is located halfway between the C3 and N5 centers, in particular we obtain bond distances of 1.313 (C3 \cdots H⁻) and 1.343 Å ($^- \text{H}\cdots\text{N5}$). These bond distances indicate that the TS for the hydride transfer in model B is earlier than that in model A, and therefore the lower barrier is consistent with the Hammond–Leffler postulate¹⁵. The effect of Glu361 on the barrier height of the hydride transfer is also demonstrated by considering the atomic polar tensor (APT) charges on the hydride and the substrate in the TS^{16,17}. In model A, the overall charge on the substrate is +0.245 a.u. and the atomic charge on the hydride is -0.178 a.u. However, in model B, due to the proton transfer from the substrate to Glu361, the overall charge on the substrate becomes negative (namely it becomes -0.331 a.u.) and the hydride becomes more negatively charged as well with an atomic charge of -0.507 a.u. The partial negative charges on the hydride and the substrate in model B should facilitate an easier hydride transfer compared with model A where they have opposite charges. Abstracting the hydroxyl hydrogen in model B also stabilizes the product by converting the isopropanol carbocation to a neutral acetone. In this situation, the reaction energy is nearly thermo neutral, in particular it is reduced from +37.1 (model A) to +2.2 kJ mol⁻¹ (model B).

It has been previously suggested that the hydrogen bond between the main chain NH of Gly120 and the flavin redox center (N5) may facilitate the catalytic hydride transfer by re-arrangement of the lone-pair electrons on N5⁴. Model C includes a dimethylamine (HNMe₂) molecule that mimics the hydrogen bond between Gly120 and FAD. The optimized TS is shown in Fig. 7 and the optimized O2 \cdots H1 and O1 \cdots H1 distances in the RC, TS, and PC are listed in Supplementary Table S2. The Gibbs free energy barrier height of model C is $102.1 \text{ kJ mol}^{-1}$, i.e., it is 11.8 kJ mol^{-1} lower than that of model B (see Fig. 6), and the reaction energy of this model is exergonic by 14.7 kJ mol^{-1} . These values are in excellent agreement with the barrier obtained from the QM/MM MD simulations, namely: $\Delta G_{298}^{\ddagger} = 99.3$ and $\Delta G_{298} = -11.8 \text{ kJ mol}^{-1}$. A natural bond orbital (NBO) analysis¹⁸ performed at the B3LYP-D3/6-31+G(2df,p) level of theory revealed that lone-pair $\rightarrow \sigma_{\text{N-H}}$ interactions between the N5 center and N–H bond of Gly120 are associated with a stabilization energy of 20.6 kJ mol^{-1} in the reactant complex and 26.2 kJ mol^{-1} in the transition structure. The greater stabilization energy in the TS relative to the RC may partially account for the catalytic enhancement provided by the Gly120 residue. The above stabilization energies are approximated by the second-order perturbation energy ($E^{(2)}$) obtained from the NBO analysis¹⁹.

In summary, our high-level DHDFT modeling suggests that the Glu361 residue, which abstracts the proton from the hydroxyl group of the substrate, has a significant influence on the Gibbs free energy barrier for the enzymatic hydride transfer. Namely, the Glu361 residue lowers the barrier by 22.6 kJ mol^{-1} (comparison of models A and B). We also show that the Gly120 residue lowers the barrier further by 11.8 kJ mol^{-1} (comparison of models B and C).

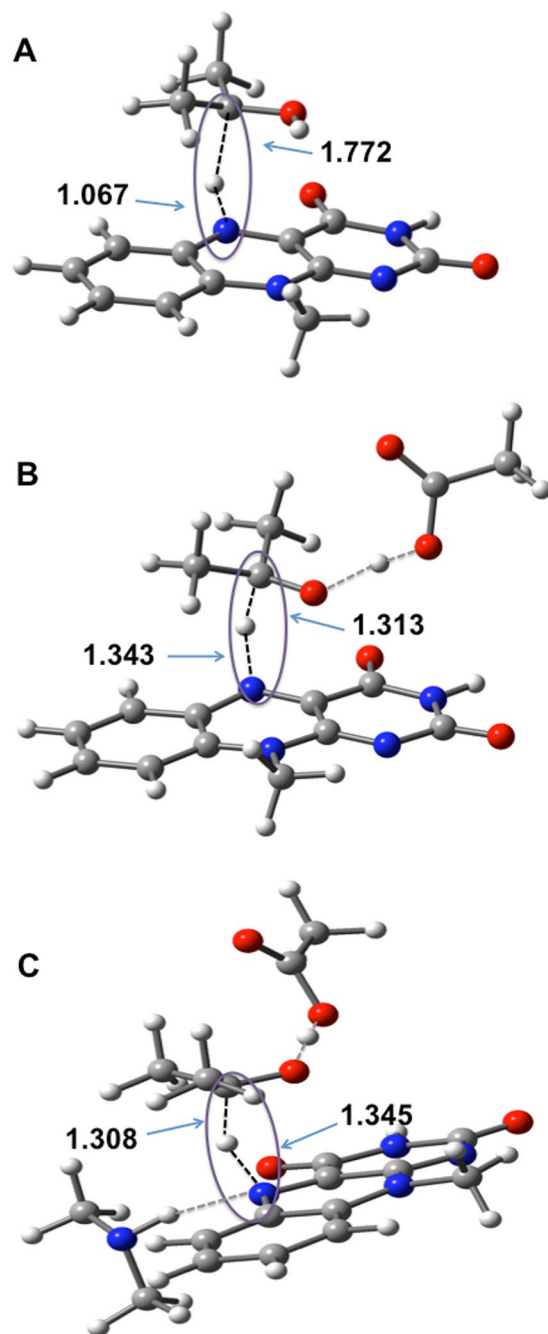


Figure 7. B3LYP-D3/6-31+G(2df,p) optimized TSs located on the potential energy profiles in models A, B, and C. Hydrogen bonds are shown as grey dashed lines. The bonds that are broken and formed in the TSs are shown in black dashed lines with purple circles. The C3–H3 and H3–N5 bond distances are given in Å. Atomic color scheme: H, white; C, grey; N, blue; O, red.

Discussion

Extensive QM/MM MD simulations and DHDFT calculations have been carried out in order to explore the mechanism of the oxidation reaction catalyzed by ChOx as well as the roles that the surrounding residues play for the hydride transfer from the substrate to the co-factor.

Using QM/MM MD simulations and umbrella sampling we explored the reaction mechanism for the hydride transfer. In addition, the roles of the key residues in the substrate binding are examined. Point mutations of His447 to glutamine and asparagine increase the relative energy barrier height of the hydride transfer by 32.8 and 66.3 kJ mol⁻¹, respectively. These results are in good quantitative agreement with the reduction in the experimental k_{cat} values upon mutation. Our simulations suggest that mutation of His447 disrupts the hydrogen-bonding network around the substrate, which in turn results in the substrate being in a less favorable position for the proton transfer from the substrate to the base Glu361. In addition, the MM-GBSA calculations show that mutations

of Glu361, His447, and Asn485 residues to alanine decrease the binding free energies of these systems by 35.8, 41.3, and 4.8 kJ mol⁻¹, respectively, relative to the WT system. This indicates that Glu361 and His447 play important roles in substrate binding. On the basis of per-residue type, the free energy decomposition shows that the residues of Gly120, Glu361, His447, and Asn485 in the active site also have notable contributions to the substrate binding, in which the electrostatic and van der Waals interactions play predominant roles in binding the substrate in the active site.

The DHDFT results show that Glu361 and Gly120 play significant roles in the enzymatic catalysis. Glu361 serves as a base for abstracting the proton from the hydroxyl group of the substrate and facilitates the hydride transfer. The Gibbs free energy barrier for the hydride transfer is lowered by 22.6 kJ mol⁻¹ upon inclusion of the Glu361 residue in our model. The hydrogen bond interaction between Gly120N-H and N5 of FAD also enhances the hydride transfer and lowers the barrier height of the hydride transfer further. We note that there is a good mechanistic agreement between the QM/MM MD simulations and our best DHDFT results.

In summary, our QM/MM MD simulations and high-level DHDFT using small model systems offer two complementary perspectives on the catalytic mechanism for the hydride transfer in ChOx. The combination of these two approaches gives an in-depth understanding of this flavoenzyme catalysis. More specifically, the QM/MM MD simulations show that the His447 and Glu361 residues play important roles in substrate binding. Our DHDFT models efficiently demonstrate the effects of the Glu361 and Gly120 residues on the reaction barrier height for the hydride transfer. The good agreement between the two approaches for the reaction barrier height of the hydride transfer increases our confidence in the results obtained from these models.

Methods

Preparation of the enzyme-substrate complex model. The initial coordinates used to build the model for the present study were based on the X-ray crystal structure of oxidized cholesterol oxidase (PDB ID: 4U2T)⁴. The steroid substrate dehydroepiandrosterone (DHA) was modeled in this structure in the proposed Michaelis state, based on the structure of the DHA/enzyme complex²⁰ and knowledge of the interactions between amino acid side chains and the substrate from previous crystallographic and mutagenesis results^{4-6,12,21}. Specifically the substrate hydroxyl group was positioned so as to make hydrogen bond contact to NE2 of His447 and the C3 hydrogen atom was positioned above the flavin N5 atom. The protonation states of the ionizable residues were determined at pH = 5.2 with the H++ program²² and the neighboring hydrogen bond networks. The pH value used here is taken from experiment⁴. The FAD cofactor and the substrate as well as the protein were described by the AMBER GAFF force field²³ and the AMBER14SB force field²⁴, respectively. The electrostatic potential (ESP) was calculated at the HF/6-31 G(d) level of theory using Gaussian 09 program suite²⁵. The computed ESP of FAD and the substrate were fit using restrained electrostatic potential (RESP) charge²⁶ fitting in *antechamber*. The whole system was solvated into an 88 × 84 × 103 Å rectangular box of TIP3P water²⁷ with a 12 Å buffer distance between the box edge and the nearest solute atoms. The system was neutralized by adding Cl⁻ ions. The protons were added automatically, and the topology parameters as well as initial coordinates were generated by the *tleap* AmberTool. After the energy minimization, the system was heated gradually from 0 to 300 K for 100 ps, and another 100 ps MD simulation was carried out to relax the system density to about 1.0 g cm⁻³. Finally, the standard 10 ns MD simulation under isothermal-isobaric (NPT) ensemble (NTP = 1) was employed with an integration time step of 2 fs *via* using the periodic boundary condition. The Berendsen barostat was used for pressure-control. The cutoff value was set to 10 Å for van der Waals and electrostatic interaction calculations. Long-range electrostatic interactions were dealt with by the Particle Mesh Ewald (PME) method^{28,29}. The Langevin thermostat (NTT = 3) was used to maintain the temperature at 300 K. All the bonds involving hydrogen atoms were constrained by using the SHAKE scheme³⁰. The RMSD was used to estimate the stability of the backbone of the enzyme, and the last 2 ns trajectories were utilized for the protein-ligand interaction decomposition of the binding free energy for both the WT and mutant systems (Glu361Ala, His447Ala, and Asn485Ala). Hou *et al.* reported that the MM-GBSA method with the generalized Born model performs well for binding free energy estimations^{31,32}. In the current study, the MMPBSA.py module³³ was then applied for the post-processing in which 1000 snapshots from an ensemble of conformations are used to calculate the free energy. This method has been also used to analyze the binding free energy for protein-ligand interactions for the WT and mutant systems. The per-residue decomposition was applied to estimate the contribution from each residue to the total binding free energy. The whole MD simulations were accomplished by applying the AMBER 14 software³⁴.

QM/MM MD simulations. Development and applications of combined QM/MM methods for enzymes have been extensively reviewed in the past³⁵⁻³⁷. In order to obtain the barrier height of the hydride transfer, QM/MM MD simulations have been carried out for this enzyme, considering the influence from the surrounding residues to the reaction site. The residues: Lys225, Asn119, Gly120, Glu361, His447, Asn485, the FAD cofactor, and the substrate are considered as the QM region (103 atoms in total, see Fig. 2). The QM/MM boundary settings were determined according to the pseudo-bond rule³⁸. These pseudo-bonds, C_{ps}(sp³)-C(sp³), C_{ps}(sp³)-C(sp², carbonyl) and C_{ps}(sp³)-N(sp³) with accurate parameterization are preferred to be used when cutting the QM and MM regions. The specified sets are listed below in detail: (1) Lys227: only the -NH₃⁺ and the adjacent -CH₂ groups included (C_{ps}(sp³)-C(sp³) type); (2) Asn119: only the -C=O group forming peptide bond with -NH group of Gly120 included (C_{ps}(sp³)-C(sp², carbonyl) type); (3) Gly120: the whole Gly120 residue included and the -NH group of Gly121 included (C_{ps}(sp³)-N(sp³) type); (4) Glu361: only the -COO⁻ and the adjacent -CH₂ groups included (C_{ps}(sp³)-C(sp³) type); (5) His447: the imidazole ring with one -CH₂ group included (C_{ps}(sp³)-N(sp³) type); (6) Asn485: the -NH₂ and two adjacent -C=O and -CH₂ groups included (C_{ps}(sp³)-C(sp³) type); (7) FAD: without the long chain (C_{ps}(sp³)-C(sp³) type); (8) DHA: only one six-membered ring with the -OH group was included as well as two adjacent short chains (C_{ps}(sp³)-C(sp², carbonyl) and C_{ps}(sp³)-C(sp³) types). All the atoms in QM region are treated using the B3LYP/6-31 G(d) level of theory and atoms in the MM region are described

by the AMBER99SB force field³⁹. This approach has been successfully applied to other enzymatic reactions in the past decade^{40,41}. The TIP3P model was used for the solvent water molecules²⁷. The QM/MM boundary was addressed by the pseudo-bond approach^{37,38} with improved parameters⁴². The spherical boundary condition was applied, in which all atoms within a radius of 25 Å from the spherical center atom were allowed to freely move. The 12 and 18 Å cutoffs were used for van der Waals and electrostatic interactions among MM atoms, respectively. Following the interactive minimization, the minimum reaction energy path was mapped out by using the reaction coordinate driving method (RCD). Then, a 200 ps MD simulation was used to equilibrate the MM part with the QM subsystem constrained. The obtained snapshot was used for the subsequent QM/MM MD simulations combined with the umbrella sampling⁴³. A total of 43 umbrella windows were employed to cover the reaction coordinate from -1.59 to 2.61 Å. For each window, 20 ps QM/MM MD simulations have been carried out with a time step of 1 fs by Beeman algorithm⁴⁴. Berendsen thermostat method⁴⁵ was used to control the system temperature at 300 K. The first 5 ps simulations were used for the equilibration and then another 15 ps simulations were used to obtain the PMF by the weighted histogram analysis method (WHAM)⁴⁶. The appropriate step size and harmonic force constant for each window were chosen to guarantee the sufficient sampling. All the QM/MM MD simulations were carried out by the modified Q-Chem 4.0⁴⁷ and Tinker 4.2 programs⁴⁸.

High-level double-hybrid DFT simulations. High-level DHDFT calculations using the B2GP-PLYP functional⁴⁹ were performed in order to obtain accurate reaction energies and barrier heights of the hydride transfer from the C3 atom of the substrate to the N5 atom of the FAD cofactor. DHDFT methods include non-local electron correlation from second-order Møller–Plesset (MP2) perturbation theory in addition to the regular ingredients of hybrid DFT⁵⁰. These methods overcome limitations of traditional DFT methods and display excellent performance for challenging chemical problems^{50–54}. Due to the slow basis set convergence of the MP2-type correlation component⁵⁵, the DHDFT calculations were carried out in conjunction with the quadruple-zeta Def2-QZVPP basis set⁵⁶.

The B3LYP-D3/6–31+G(2df,p) level of theory was applied to fully optimize the geometries in our models A, B, and C (see Fig. 7). Based on these optimized geometries, single-point energies at RI-B2GP-PLYP/Def2-QZVPP level were carried out using the ORCA program⁵⁷ (RI stands for resolution of identity approximation applied in the MP2 step)^{58,59}. Zero-point vibrational energy (ZPVE), enthalpic, and entropic corrections have been obtained from the B3LYP-D3/6–31+G(2df,p) harmonic frequency calculations for converting the electronic RI-B2GP-PLYP/Def2-QZVPP reaction energies and barrier heights into Gibbs free reaction energies and barrier heights at 298 K. Empirical dispersion corrections were included using the Becke–Johnson damping potential⁶⁰ as recommended by Grimme *et al.* (denoted by the suffix D3)^{61–63}. The equilibrium structures in the unconstrained optimizations were verified to have all real harmonic frequencies and the transition structures were confirmed to have only one imaginary frequency. The connectivities of the transition structures were confirmed by performing intrinsic reaction coordinate (IRC) calculations^{64,65}. The enzyme-like environment was simulated by a homogeneous polarizable continuum model with the dielectric constant of 4.0^{66–68} using the conductor-like polarizable continuum model (CPCM) at the HF/6–31+G(d) level of theory in conjunction with UAHF atomic radii, as recommended by Takano and Houk⁶⁹.

All the above geometry optimization, harmonic frequency, and enzyme-like environment solvent correction calculations were performed using the Gaussian 09 program suite²⁵.

References

- Kreit, J. & Sampson, N. S. Cholesterol oxidase: physiological functions. *FEBS J.* **276**, 6844–6856 (2009).
- Vrielink, A. & Ghisla, S. Cholesterol oxidase: biochemistry and structural features. *FEBS J.* **276**, 6826–6843 (2009).
- Pollegioni, L., Piubelli, L. & Molla, G. Cholesterol oxidase: biotechnological applications. *FEBS J.* **276**, 6857–6870 (2009).
- Golden, E., Karton, A. & Vrielink, A. High-resolution structures of cholesterol oxidase in the reduced state provide insights into redox stabilization. *Acta Crystallogr. Sect. D* **70**, 3155–3166 (2014).
- Golden, E. *et al.* An extended NH bond, driven by a conserved second-order interaction, orients the flavin N5 orbital in cholesterol oxidase. *Sci. Rep.* **7**, 40517 (2017).
- Lario, P. I., Sampson, N. & Vrielink, A. Sub-atomic resolution crystal structure of cholesterol oxidase: what atomic resolution crystallography reveals about enzyme mechanism and the role of the FAD cofactor in redox activity. *J. Mol. Bio.* **326**, 1635–1650 (2003).
- Sampson, N. S. & Kass, I. J. Isomerization, but not oxidation, is suppressed by a single point mutation, E361Q, in the reaction catalyzed by cholesterol oxidase. *J. Am. Chem. Soc.* **119**, 855–862 (1997).
- Kass, I. J. & Sampson, N. S. The importance of Glu361 position in the reaction catalyzed by cholesterol oxidase. *Bioorg. Med. Chem. Lett.* **8**, 2663–2668 (1998).
- Kass, I. J. & Sampson, N. S. Evaluation of the role of His447 in the reaction catalyzed by cholesterol oxidase. *Biochemistry* **37**, 17990–18000 (1998).
- Yin, Y., Liu, P., Anderson, R. G. & Sampson, N. S. Construction of a catalytically inactive cholesterol oxidase mutant: investigation of the interplay between active site-residues glutamate 361 and histidine 447. *Arch. Biochem. Biophys.* **402**, 235–242 (2002).
- Yin, Y., Sampson, N. S., Vrielink, A. & Lario, P. I. The presence of a hydrogen bond between asparagine 485 and the π system of FAD modulates the redox potential in the reaction catalyzed by cholesterol oxidase. *Biochemistry* **40**, 13779–13787 (2001).
- Yue, Q. K., Kass, I. J., Sampson, N. S. & Vrielink, A. Crystal structure determination of cholesterol oxidase from *Streptomyces* and structural characterization of key active site mutants. *Biochemistry* **38**, 4277–4286 (1999).
- Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A. & Case, D. A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate–DNA helices. *J. Am. Chem. Soc.* **120**, 9401–9409 (1998).
- Kollman, P. A. *et al.* Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **33**, 889–897 (2000).
- Hammond, G. S. A correlation of reaction rates. *J. Am. Chem. Soc.* **77**, 334–338 (1955).
- Cioslowski, J. A new population analysis based on atomic polar tensors. *J. Am. Chem. Soc.* **111**, 8333–8336 (1989).
- De Prof, F., Martin, J. M. & Geerlings, P. On the performance of density functional methods for describing atomic populations, dipole moments and infrared intensities. *Chem. Phys. Lett.* **250**, 393–401 (1996).
- Foster, J. & Weinhold, F. Natural hybrid orbitals. *J. Am. Chem. Soc.* **102**, 7211–7218 (1980).

19. Reed, A. E., Curtiss, L. A. & Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **88**, 899–926 (1988).
20. Li, J., Vrieland, A., Brick, P. & Blow, D. M. Crystal structure of cholesterol oxidase complexed with a steroid substrate: implications for flavin adenine dinucleotide dependent alcohol oxidases. *Biochemistry* **32**, 11507–11515 (1993).
21. Lyubimov, A. Y., Lario, P. I., Moustafa, I. & Vrieland, A. Atomic resolution crystallography reveals how changes in pH shape the protein microenvironment. *Nat. Chem. Biol.* **2**, 259–264 (2006).
22. Anandakrishnan, R., Aguilar, B. & Onufriev, A. V. H++3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **40**, W537–W541 (2012).
23. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
24. Maier, J. A. *et al.* ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
25. Frisch, M. *et al.* Gaussian 09, Revision D. 01. *Gaussian: Wallingford, CT, USA*, 2009.
26. Bayly, C. L., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269–10280 (1993).
27. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
28. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
29. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
30. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
31. Hou, T., Wang, J., Li, Y. & Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. I. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* **51**, 69–82 (2011).
32. Hou, T., Wang, J., Li, Y. & Wang, W. Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. *J. Comput. Chem.* **32**, 866–877 (2011).
33. Miller, B. R. III *et al.* MMPBSA. py: an efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **8**, 3314–3321 (2012).
34. Case, D. *et al.* AMBER 14, University of California, San Francisco, USA, 2014.
35. Warshel, A. Computer simulations of enzyme catalysis: methods, progress, and insights. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 425–443 (2003).
36. Gao, J. & Truhlar, D. G. Quantum mechanical methods for enzyme kinetics. *Annu. Rev. Phys. Chem.* **53**, 467–505 (2002).
37. Zhang, Y. Pseudobond ab initio QM/MM approach and its applications to enzyme reactions. *Theor. Chem. Acc.* **116**, 43–50 (2006).
38. Zhang, Y., Lee, T.-S. & Yang, W. A pseudobond approach to combining quantum mechanical and molecular mechanical methods. *J. Chem. Phys.* **110**, 46–54 (1999).
39. Hornak, V. *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725 (2006).
40. Wu, R., Lu, Z., Cao, Z. & Zhang, Y. Zinc chelation with hydroxamate in histone deacetylases modulated by water access to the linker binding channel. *J. Am. Chem. Soc.* **133**, 6110–6113 (2011).
41. Ke, Z., Smith, G. K., Zhang, Y. & Guo, H. Molecular mechanism for eliminolysis, a newly discovered post-translational modification. *J. Am. Chem. Soc.* **133**, 11103–11105 (2011).
42. Zhang, Y. Improved pseudobonds for combined ab initio quantum mechanical/molecular mechanical methods. *J. Chem. Phys.* **122**, 024114 (2005).
43. Torrie, G. M. & Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **23**, 187–199 (1977).
44. Beeman, D. Some multistep methods for use in molecular dynamics calculations. *J. Comput. Phys.* **20**, 130–139 (1976).
45. Berendsen, H. J., Postma, J. V., van Gunsteren, W. F., DiNola, A. & Haak, J. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
46. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **13**, 1011–1021 (1992).
47. Shao, Y. *et al.* Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **113**, 184–215 (2015).
48. Ponder, J. TINKER: Software Tools for Molecular Design, 4.2, Washington University School of Medicine, Saint Louis, MO (2004).
49. Karton, A., Tarnopolsky, A., Lamere, J.-F., Schatz, G. C. & Martin, J. M. Highly Accurate First-Principles Benchmark Data Sets for the Parametrization and Validation of Density Functional and Other Approximate Methods. Derivation of a Robust, Generally Applicable, Double-Hybrid Functional for Thermochemistry and Thermochemical Kinetics†. *J. Phys. Chem. A* **112**, 12868–12886 (2008).
50. Goerigk, L. & Grimme, S. Double-hybrid density functionals. *WIREs Comput. Mol. Sci.* **4**, 576–600 (2014).
51. Karton, A., O'Reilly, R. J. & Radom, L. Assessment of theoretical procedures for calculating barrier heights for a diverse set of water-catalyzed proton-transfer reactions. *J. Phys. Chem. A* **116**, 4211–4221 (2012).
52. Karton, A. & Goerigk, L. Accurate reaction barrier heights of pericyclic reactions: Surprisingly large deviations for the CBS-QB3 composite method and their consequences in DFT benchmark studies. *J. Comput. Chem.* **36**, 622–632 (2015).
53. Yu, L.-J., Sarrami, F., O'Reilly, R. J. & Karton, A. Reaction barrier heights for cycloreversion of heterocyclic rings: An Achilles' heel for DFT and standard ab initio procedures. *Chem. Phys.* **458**, 1–8 (2015).
54. Yu, L.-J., Sarrami, F., O'Reilly, R. J. & Karton, A. Can DFT and ab initio methods describe all aspects of the potential energy surface of cycloreversion reactions? *Mol. Phys.* **114**, 21–33 (2016).
55. Karton, A. & Martin, J. M. Basis set convergence of explicitly correlated double-hybrid density functional theory calculations. *J. Chem. Phys.* **135**, 144119 (2011).
56. Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005).
57. Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2**, 73–78 (2012).
58. Weigend, F. & Häser, M. RI-MP2: first derivatives and global consistency. *Theor. Chem. Acc.* **97**, 331–340 (1997).
59. Ren, X. *et al.* Resolution-of-identity approach to Hartree–Fock, hybrid density functionals, RPA, MP2 and GW with numeric atom-centered orbital basis functions. *New J. Phys.* **14**, 053020 (2012).
60. Becke, A. D. & Johnson, E. R. A density-functional model of the dispersion interaction. *J. Chem. Phys.* **123**(154101), 1–9 (2005).
61. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **132**(154104), 1–19 (2010).
62. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465 (2011).
63. Grimme, S. Density functional theory with London dispersion corrections. *WIREs Comput. Mol. Sci.* **1**, 211–228 (2011).

64. Gonzalez, C. & Schlegel, H. B. An improved algorithm for reaction path following. *J. Chem. Phys.* **90**, 2154–2161 (1989).
65. Gonzalez, C. & Schlegel, H. B. Reaction path following in mass-weighted internal coordinates. *J. Phys. Chem.* **94**, 5523–5527 (1990).
66. Himo, F. Quantum chemical modeling of enzyme active sites and reaction mechanisms. *Theor. Chem. Acc.* **116**, 232–240 (2006).
67. Blomberg, M. R., Borowski, T., Himo, F., Liao, R.-Z. & Siegbahn, P. E. Quantum chemical studies of mechanisms for metalloenzymes. *Chem. Rev.* **114**, 3601–3658 (2014).
68. Himo, F. & Siegbahn, P. E. Quantum chemical studies of radical-containing enzymes. *Chem. Rev.* **103**, 2421–2456 (2003).
69. Takano, Y. & Houk, K. Benchmarking the conductor-like polarizable continuum model (CPCM) for aqueous solvation free energies of neutral and ionic organic molecules. *J. Chem. Theory Comput.* **1**, 70–77 (2005).

Acknowledgements

We gratefully acknowledge the generous allocation of computing time from the National Computational Infrastructure (NCI) National Facility, and system administration support provided by the Faculty of Science at the University of Western Australia (UWA) to the Linux cluster of the Karton group, the provision of an Australian Postgraduate Award (to L.-J.Y.), and an Australian Research Council (ARC) Discovery Early Career Researcher Award (to A.K., project number: DE140100311). We also thank the National Science Foundation of China (to Y.Z., project number: 21603057), and Special Program for Applied Research on Super Computation of the NSFC-Guangdong Joint Fund (the second phase).

Author Contributions

Conceptualization: L.-J.Y., A.V. and A.K. Methodology: L.-J.Y., N.C., Y.Z. and A.K. Formal Analysis: L.-J.Y., N.C. and Y.Z. Investigation: L.-J.Y., N.C., Y.Z., A.V. and A.K. Writing Original Draft: L.-J.Y. Writing Review and Editing: L.-J.Y., E.G., N.C., Y.Z., A.V. and A.K. Figure Preparation: L.-J.Y. and E.G. Supervision: A.V. and A.K. Project administration: A.K.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-17503-x>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017