ARTICLE

Reaction Mechanism of *N*-Cyclopropylglycine Oxidation by Monomeric Sarcosine Oxidase

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Mitsuo Shoji,[*ab] Yukihiro Abe,*[c] Mauro Boero,[d] Yasuteru Shigeta[a] and Yoshiaki Nishiya[e]

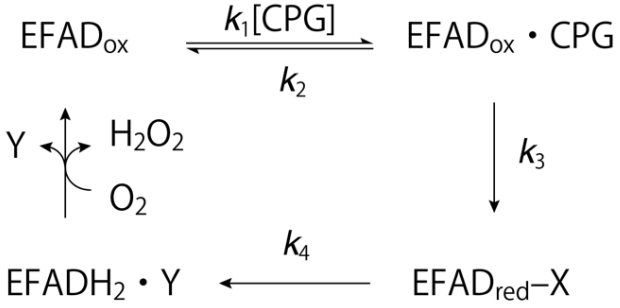
Monomeric sarcosine oxidase (MSOX) is a fundamental - yet one of the simplest - member of a family of flavoenzymes able to catalyze the oxidation of sarcosine (*N*-methylglycine) and other secondary amines. MSOX is one of the best characterized members of the amine oxidoreductases (AOs), however, its reaction mechanism is still controversial. A single electron transfer (SET) process was suggested on the basis of studies with *N*-cyclopropylglycine (CPG), although a hydride transfer mechanism would be more consistent in general for AOs. To shed some light on the detailed reaction mechanisms of CPG in MSOX, we performed hybrid quantum mechanical/molecular mechanical (QM/MM) simulations. We found that the polar mechanism is energetically the most favorable. The free energy profile indicates that the first rate-limiting step is the CPG binding to the flavin ring which simultaneously proceeds with the ring-opening of the CPG cyclopropyl group. This reaction step of the CPG adduct formation corresponds to the nucleophilic attack of the cyclopropyl group (C3 atom) to the flavin ring (C4a atom), whereas the expected radical species formation in the SET mechanism was not observed. The following inactivated species, which accumulates during the CPG oxidation in MSOX, can be ascribed to an imine state, and not an enamine state, on the basis of the computed UV/Vis spectra. The conformation of CPG was found to be crucial for reactions following the CPG adduct formation.

Introduction

Amine oxidation is an essential process for the metabolism of a wealth of microorganisms. All the amine oxidoreductases (AOs) in both the prokaryotic and eukaryotic forms contain covalently bound flavins. This common character seems to be due to intrinsic properties of flavin, which possesses a redox potential sufficiently high to oxidize a wide range of biological molecules.[1] The main oxidation reaction operated by the AOs is the conversion of a CN single bond of the substrates (mostly amino acids and amines) into a double bond. After hydrolysis of the CN double bond species (Schiff bases), the release of ammonium and the cleavage of the amines generate primary and secondary amines, respectively. In particular, monomeric sarcosine oxidase (MSOX) catalyzes the oxidation of sarcosine (*N*-methylglycine) and secondary amino acids.[2] MSOX is widely used in clinical diagnostics with creatininase and creatinase to test the renal function.[3] The reaction mechanism of the MSOX has been extensively investigated by spectroscopic measurements, crystal structural analyses and molecular modeling approaches.[4,5] Closely related members of the AO family, as, for instance, monoamine oxidase (MAO) and *D*-amino acid oxidase (DAAO), were also the focus of intense research activities.[6-12] Nonetheless, the reaction mechanisms of the AOs have not been fully understood and multiple reaction mechanisms have been proposed as plausible.[9,13-15]

To date, at least four different mechanisms have been proposed: Two variants of a single electron transfer (SET) mechanism, a polar mechanism and a hydride transfer one.[9] For each possible mechanism, different intermediates are expected. The SET mechanism proceeds through a one-electron transfer from the substrate amine to the flavin. Support has been provided to this SET mechanism by the reaction with the *N*-(cyclopropyl) glycine (CPG),[4,16] motivated by the fact the SET is the more straightforward way to explain the formation of a covalent flavin-substrate adduct in the case of the CPG substrate.

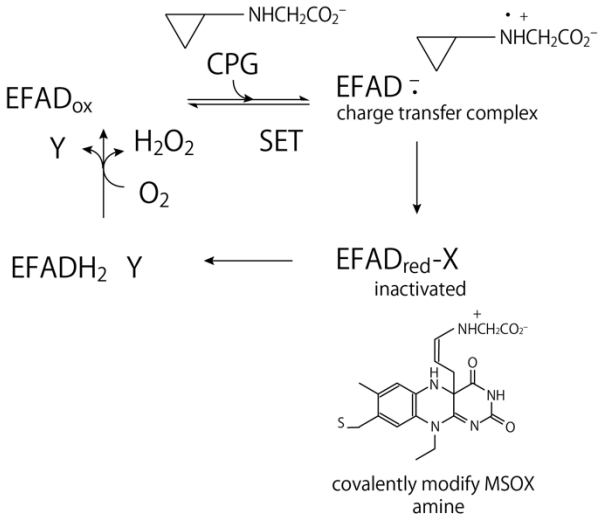
Scheme 1. Catalytic cycle of MSOX for CPG.[15]



[a] Center for Computational Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan. E-mail: mshoji@ccs.tsukuba.ac.jp
[b] JST-PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan.
[c] HPC Systems Inc., 646 Nijyohanjikicho, Ayanokojisagaru Karasumadori, Shimogyo-ku, Kyoto 600-8412, Japan. E-mail: yu-abe@hpc.co.jp
[d] University of Strasbourg, Institut de Physique et Chimie des Matériaux de Strasbourg, CNRS, UMR 7504, 23 rue du Loess, F-67034 France.
[e] Department of Life Science, Faculty of Science and Engineering, Setsunan University, 17-8 Ikedanaka-machi, Neyagawa, Osaka 572-8508, Japan.
†Electronic Supplementary Information (ESI) available: See DOI: 10.1039/x0xx00000x

The afore mentioned two variants of the SET mechanism differ in the subsequent one-proton and one-electron transfer processes occurring either simultaneously or in a stepwise way. The polar mechanism, instead, involves the formation of a covalent flavin-substrate intermediate (4a-flavin adduct), whereas the polar mechanism is triggered by thiol substrates forming 4a-flavin adducts,[17] although the monothiol adduct does not promote the formation of a two-electron-reduced flavin. The hydride transfer mechanism proceeds through the direct hydride transfer from a substrate to the flavin N5. The majority of former experimental and theoretical results suggested a hydride transfer mechanism for most of the AOs.[9,18] The cyclopropyl group of CPG is expected to be easily converted into radical intermediates, since the cyclopropylmethyl has been used to trap radical species in various enzymatic reactions.[19] Former studies indicated that the hydride transfer mechanism is unlikely to occur in CPG, because the C1 proton (H1) cannot be easily donated. The polar mechanism was also expected to be difficult to realize, because the CPG adducted intermediate binding at the N atom of CPG is unstable.[16] Given this scenario, the SET mechanism remains as the more viable route. J. M. Kim *et al.* reported that both the polar and SET mechanisms are possible for the MAOs.[20]

Scheme 2. Proposed CPG states in MSOX.[16]



In a previous work, we have shown that the reaction mechanism of MSOX with sarcosine is a hydrogen-atom-coupled electron transfer (HACET),[5] which can be roughly categorized into a hydride transfer, meaning that the proton and electron simultaneously transfer from the substrate to the flavin. We termed this process HACET because the electron transfer via the π orbital interaction between the substrate and flavin is essential for an efficient reaction and, in this respect, "hydride" is not an accurate definition of the actual proton state. A single electron transfer and 4a-flavin adduct states were found to be unstable and characterized by activation barriers higher than the ones for the HACET mechanism.

The study presented here is a thorough analysis of the reaction mechanism of MSOX with CPG. To this aim we resort to well-assessed computational approaches to investigate and unravel all the possible reaction pathways and intermediate states. The quantum mechanics/molecular mechanics (QM/MM) method used here is essential, since the inclusion of the explicit electronic structure is crucial to rationalize the electronic structure, especially when charge transfer processes occur, and to search for unforeseen protonation states and conformations, which can affect both the relative energetic stabilities of the intermediate states and the activation barriers. UV/Vis spectra calculations provide additional support to validate or discard the intermediate states.

Computational details

QM/MM simulations

A high-resolution (1.80 Å resolution) X-ray structure of sarcosine oxidase (PDB ID: 1EL5) was used as the starting structure in our simulations.[21] Crystallographic unavailable hydrogen atoms were added using the Reduce program.[22] Unsolved residues located before and after Gly3 and Glu386 at the N and C terminals were capped with acetyl and *N*-methylamine groups, respectively. CPG was initially replaced by *N*,*N*-dimethylglycine in the crystal structure. The carboxy group of the CPG was assumed to be in the deprotonated form, according to results reported for a sarcosine substrate.[23]

All the MM and MD simulations reported in this work were done using the AMBER 11 program package.[24] The parameters of the force fields Amber ff99SB, GAFF and TIP3P were used for the enzyme, the CPG and the solvating water, respectively. The atomic charges of the CPG and Cys315-FAD adduct were calculated at the DFT level (B3LYP/6-31G*) using the RESP method as implemented in the Amber Tools 1.4 and Gaussian package.[25] The monomer of MSOX was placed at the center of an orthorhombic simulation box of size equal to 88 x 89 x 109 Å^3, and the solvating water at ordinary liquid conditions completed the system. Periodic boundary conditions were applied to the simulation box for all simulations. The charge neutrality of the whole system was ensured by the addition of 21 sodium counter cations. Direct Coulomb interactions were computed using a 10 Å cut-off, and the remaining long-range interactions were evaluated via the PME method.

To equilibrate the system, an initial energy minimization of the system treated as fully classical (MM) was performed, followed by a 50-ps heating phase from T = 0 K to 300 K via MD in the canonical NVT ensemble. Subsequently, a 50-ps NPT simulation was done at room temperature (T = 300 K) and at a constant pressure (P = 1 bar). In these simulations, positional constraints were applied to the solute molecules. A 20 ns MD simulation was performed in the same NPT ensemble upon activation of the SHAKE algorithm. This equilibration phase was then followed by a 0.5 ns MD simulation in which the temperature was gradually decreased from T = 300 K to 0 K. A spherical model of 45 Å radius from the N10 atom of FAD was then extracted from the final configuration, comprehensive of the protein and all the sodium atoms. This spherical structure was the one used

to start the hybrid QM/MM simulations (Figure 1(A)), done with the NWChem 6.5 program suite.[26]

The QM subsystem, shown in Figure 1(B), includes the CPG substrate, two water molecules and the side chains of the residues FAD, Cys315, Lys265, Lys347. The thiol link of Cyr315 was explicitly included in the QM region, since the covalent bond has a direct effect in the increase of the reduction potential of FAD by ΔE ~ 120mV.[27] The selection of the most appropriate QM region has been done by considering various sizes including or excluding additional atoms and side chains residues. This benchmark analysis is reported in ESI†.

The total charge of this QM region is equal to zero. The QM driver is the DFT functional in which the exchange-correlation interaction is computed at the B3LYP level complemented by the Grimme's D3 dispersion (UB3LYP-D3).[28] A 6-311G** basis set (TZVP) is used for all the atoms. The force field for the MM subsystem was the Amber ff99 one. The electronic embedding scheme was adapted for describing the QM-MM nonbonded interactions and the Coulomb interactions were explicitly computed without cut-off in the total energy calculations. To saturate the chemical bonds crossing the QM-MM interface, hydrogen link atoms were used, and the QM subsystem was optimized at the selected DFT level while keeping fixed the MM embedding region. Reaction pathways were sampled with the nudged elastic band (NEB) algorithm, and the transition states were further refined by a saddle point search point search to verify if single imaginary vibrational mode exists. Free energies were evaluated upon inclusion of the enthalpy corrections and entropy evaluation estimated from the frequency calculations in the optimized intermediate and transition states. For the UV/Vis absorption calculation, time dependent DFT (TDDFT) calculations were performed at the B3LYP/6-311++G** level (B3LYP/6-311++G**|Amber ff99) using a reduced QM region in which the thiol link (CH_2-S-CH_2) bridging the Cys315 and the FAD was kept in the MM embedding part to reduce the computational cost. The reported absorption spectra for the MSOX wild type and C315Ala mutant are rather similar with the absorption peak centered at ~454 nm.[27,16] The B3LYP functional has been shown to be the most reliable one in terms of accuracy of the absorption spectra of riboflavin with respect to both MPWB1K and PBE0. The TD-B3LYP/6-311+G(2d, 2p) approach still suffers from deviations from the experimental values by 40-50 nm.[29] Yet, this computational set-up has been extensively used to compute with appreciable accuracy the spectra of the flavin chromophores.[30,31] All the molecular structures shown in the figures were drawn using the visual molecular dynamics (VMD) program.[32]

Results and discussion

Reaction Mechanism

All the intermediate states identified along the reaction pathways and their related free energy profiles are shown in Figures 2, 3 and 4. The optimized structures of the transition states are sketched in Figure S1 (ESI†). The atomic coordinates of all the intermediates and transition states are given in ESI†. In the following sections, the intermediate states (**1**-**5**) are thoroughly discussed, and in the following section the UV/Vis spectra of the intermediate states are analyzed. Finally, the reaction pathway is discussed and its atomic-level mechanism in MSOX is unravelled.

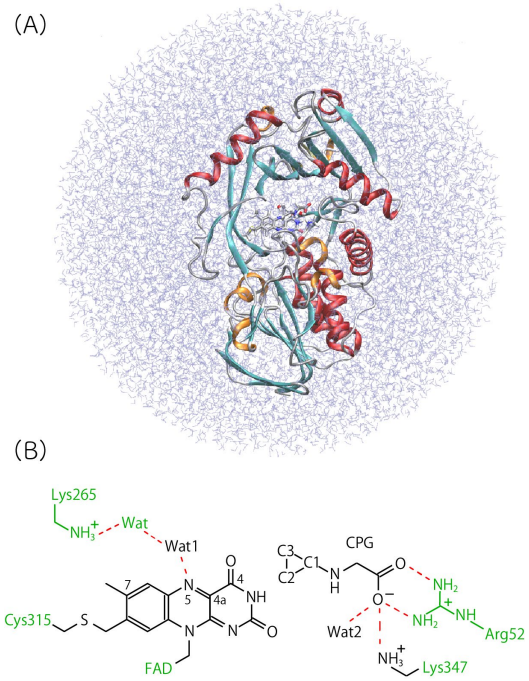


Figure 1. (A) QM/MM system used in the present study. (B) Schematic illustration of the active site of the monomeric sarcosine oxidase (MSOX). Atoms in the QM region are shown in black, and the green parts represent fragments of the embedding MM system. Hydrogen bonds in the initial state (state 1) are drawn in red broken lines.

MSOX-CPG complex (state 1)

The first state on which we focused is an enzyme-substrate complex. In this state, hereafter indicated as state **1**, the CPG is located in proximity of the flavin ring of FAD. The carboxy group of CPG is tightly anchored by two base residues, Arg52 and Lys347. Compared to the binding position of the sarcosine substrate, analyzed at an analogous QM/MM level,[5] the hydrogen-bond (H-bond) between the carboxy group of the substrate and the guanidine group of Arg52 are different. Namely, at the CPG binding, the carboxy group of CPG rigidly interacts with Arg52 via bidentate H-bonds, and the position of CPG shifts toward the N3 side of the flavin ring. Thus, the N atom of the CPG amine and the C3 (C2) atom of the CPG cyclopropane turn out to be the closest ones to the flavin C4 and C4a atoms, respectively, and their distances in the state **1a** are $R(N^{CPG}, C4^{FAD})$= 2.81 Å and $R(C3^{CPG}, C4a^{FAD})$=3.68 Å. On the other hand, at the sarcosine binding state, the N atom of the sarcosine amine and the C_M atom of the sarcosine are closest ones to the C4a and N5 atoms of flavin, respectively, as the carboxy group of the CPG interacts with the side chain of Arg52 via a monodentate hydrogen bond, and the position of the sarcosine shifts to the C7 atom of the flavin. These different positions depend on the substrates relative to the FAD flavin and are considered to be crucial to trigger different amine

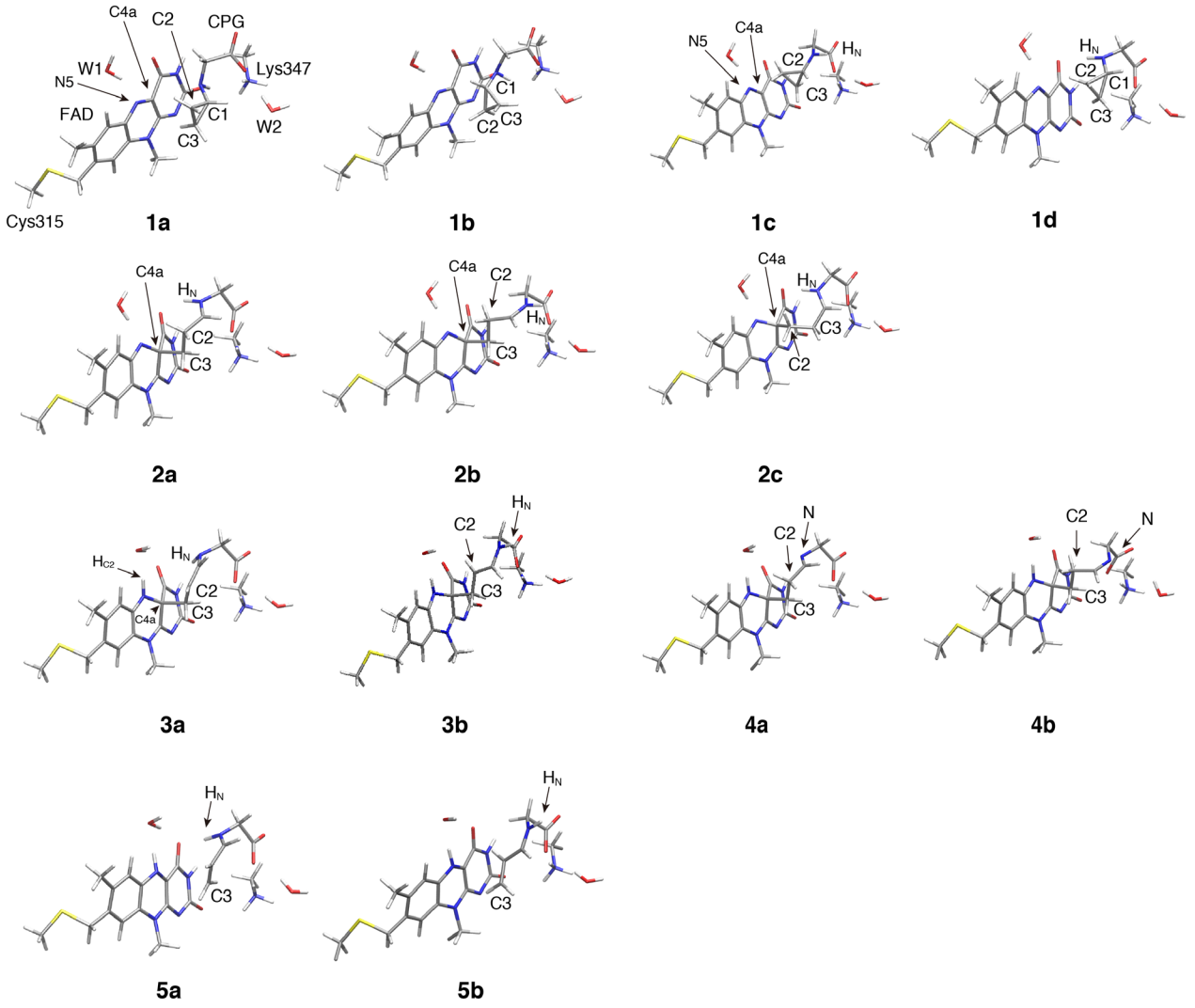


Figure 2. QM/MM optimized structures of the intermediate states. The letter following each number in states indicates a specific conformer. For the sake of clarity, only the atoms of the QM subsystem are shown.

oxidation (H- elimination) reactions; for instance, in sarcosine, the C1 hydrogen is transferred to the FAD, whereas in CPG, the C2 hydrogen is the one transferred to the FAD.

The search for stable CPG conformers in the state **1** allowed us to identify four different conformers,[33] hereafter denoted as the states **1a**-**1d**. The most stable conformer is **1a** (Figure 2), followed by a conformer with the rotated cyclopropane head of CPG (state **1b**). Their relative stability locates the **1b** above **1a** by ΔG =1.4 kcal mol^{-1}. The other two conformers of state 1, **1c** and **1d**, present a rotated amine group pointing toward the opposite and same side, respectively, of the O4 atom of FAD. Both **1c** and **1d** are energetically located above **1a** by ΔG = 6~9 kcal mol^{-1}. However, in these amine conformations (**1c** and **1d**), the distances of ($C2^{CPG}$, $N5^{FAD}$) and ($C3^{CPG}$, $C4a^{FAD}$) shrink to about 3 Å (Table S1, in ESI†). As discussed in detail in the following sections, such a parallel orientation of the $C2^{CPG}$-$C3^{CPG}$ bond and $N5^{FAD}$-$C4a^{FAD}$ bond in the **1c** and **1d** states play a central role in the oxidation of CPG promoted by FAD. From a

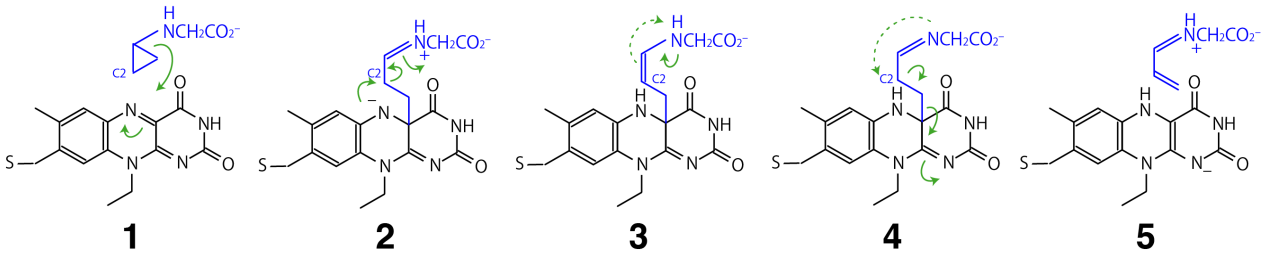


Figure 3. Schematic illustration of the intermediate states. CPG moieties are colored blue for easy identification. The green curved arrows show the movement of electrons. The dashed arrows showing the direct proton transfer between C2 and N in the CPG moiety are only helpful guides to connect the states.

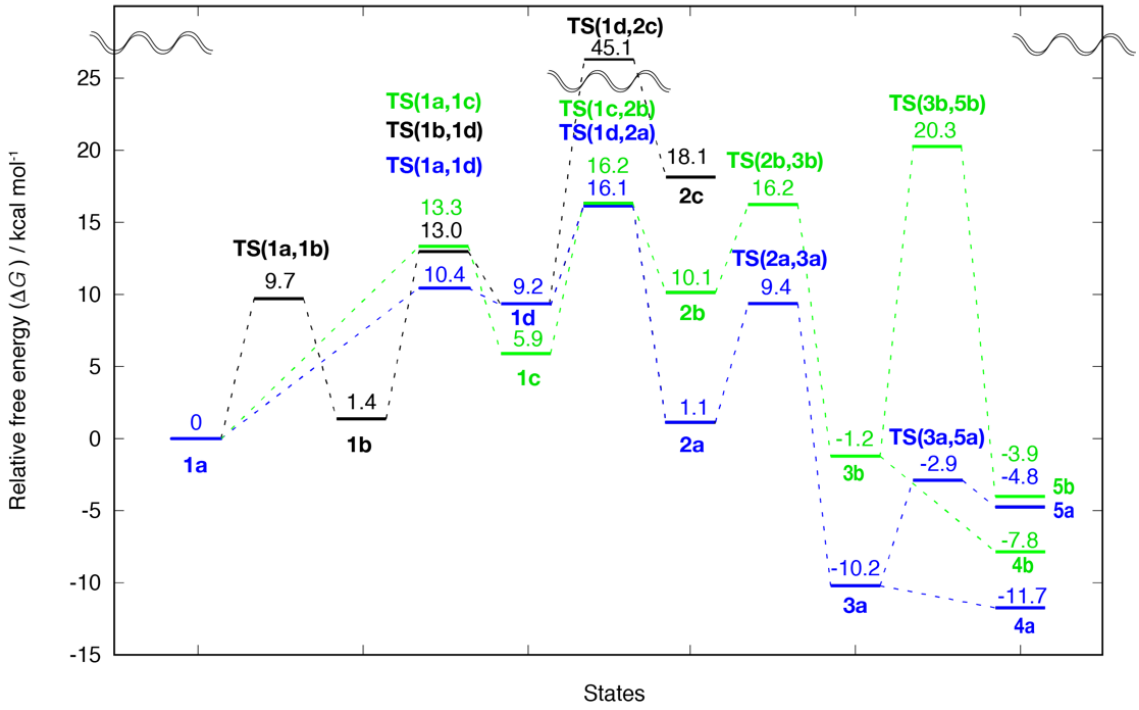


Figure 4. Free energy profile calculated at the QM/MM (UB3LYP-D3/TZVP|AMBER99) theoretical level. The green and blue colors refer to the different amine conformation in CPG along the reaction pathway.

geometric standpoint, in **1a**, only the $C3^{CPG}$ atom is close to the $N5^{FAD}$ and $C4a^{FAD}$ atoms, whereas in state **1b**, only the $C2^{CPG}$ atom is close to the $N5^{FAD}$ and $C4a^{FAD}$ atoms (Table S1). The activation barriers of all the states **1**, such as ΔG(**TS(1a,1b)**) and ΔG(**TS(1a,1c)**), are characterized by values $\Delta G^{\neq} \leq 13.3$ kcal mol^{-1}. Such a relative low activation barrier is a clear indication of the fact that the conformers of state **1** (**1a**-**1d**) can be easily interconverted.

CPG bound states (states 2)

By searching the possible reactions from the state **1**,[33] a C4a adduct formation of CPG to FAD was found to be a plausible candidate. This pathway includes the cleavage of a C3-C1 bond in CPG, accompanied by the formation of a new C3-C4a bond between CPG and FAD. Two unique features characterize this reaction pathway. The first one is the fact that only the **1c** and **1d** states can give rise to C4a adducts (states **2b** and **2a**) with a low activation barrier. The free energy barriers to the states **2a** and **2b** result rather similar and sufficiently low (ΔG(**TS(1c,2b)**) ≈ ΔG(**TS(1d,2a)**) ≈ 16kcal mol^{-1}) to make this reaction pathway realizable.

The second feature is that the $C3^{CPG}$-$C1^{CPG}$ bond cleavage and the $C3^{CPG}$-$C4a^{FAD}$ bond formation occur at the different faces, in a way similar to what is observed in an S_N2 reaction. If the reaction proceeds in the same face as an S_N1 reaction, then $C3^{CPG}$-$C1^{CPG}$ bond cleavage occurs before the formation of the $C3^{CPG}$-$C4a^{FAD}$ bond. At the transition state, the $C3^{CPG}$ radical is formed, and the radical species is expected to undergo a rotation bringing the system from the $C1^{CPG}$ atom direction to the $C4a^{FAD}$ one. The activation barrier for this pathway turns out to be ΔG ≈28 kcal mol^{-1}, and this relatively large value can be ascribed to the fact that a $C3^{CPG}$ radical has to be generated to rotate the methylene terminal (-$C3H_2$). Conversely, in the S_N2-like pathway, the $C3^{CPG}$ radical rotation is not necessary and the motion of the $C3H_2$ moiety in CPG is negligible, since the $C3^{CPG}$-$C1^{CPG}$ bond elongates, and the $C3^{CPG}$-$C4a^{FAD}$ bond shrinks. Thus, this reaction channel (**1d** -> **2a** and **1c** -> **2b**) is the one realizing the minimum energy pathway not to form a $C3^{CPG}$ radical species during the C4a adduct formation.

We also checked whether or not an alternative C atom ($C2^{CPG}$) of the CPG moiety can react with $C4a^{FAD}$. This inspection has shown that the formed product is the state **2c**. Nonetheless, this state **2c** and its associated transition state are energetically less stable (ΔG(**2c**) = 18 kcal mol^{-1} and ΔG(**TS(1d,2c)**) = 45 kcal mol^{-1}), hence the $C2^{CPG}$ atom is unlikely to become an alternative reactive side with respect to $C4a^{FAD}$.

Among the C4a adduct states of states **2**, the conformation identified as state **2a** is the most stable. In the state **2b**, the amine NH of CPG points away from the FAD flavin ring, and this state **2b** is energetically higher than **2a** by 9 kcal mol^{-1}. In Figure 4, the green and blue colors identify the amine conformation in CPG along the reaction pathway. Among the states **1**, the one labeled as **1d** is energetically higher than **1c** by ΔG = 3 kcal mol^{-1}, and this same state can revert easily to **2a** as the direction of the CPG amine is identical in both cases. Similarly, state **1c** can be converted into **2b** by the CPG binding. Since in this case the energetic order of the states 1 and 2 is reversed, the relative energy at which the transition state is located turns out to be similar, i.e. ΔG(**TS(1c,2b)**) ≈ ΔG(**TS(1d,2a)**) ≈ 16 kcal mol^{-1}.

We did also a thorough inspection of the states realized by the SET mechanism. Since the SET mechanism can realize a diradical structure, namely a CPG radical form and a flavin radical, the QM subsystem assumes a triplet spin state electronic configuration. Our calculations show that SET in state **1** is still unstable, being characterized by ΔG (**1_{SET}**) = 16 kcal mol^{-1}, and the second one expected as stable, presenting a ring-opened cyclopropane group (2_{SET}) is equally unlikely in view of its higher energy ΔG (**2_{SET}**) = 18 kcal mol^{-1}. We remark that this energy is when higher than the one characterizing the **TS(1d,2a)** in the radical formation step. The transition state for the C3-C1 bond cleavage in the SET mechanism is energetically slightly higher (ΔG (**TS(1_{SET}**, **2_{SET}**)) = 20 kcal mol^{-1}). The molecular structures of these states in the SET mechanism are reported in Figure S2 (ESI†). This energetically demanding pathway indicates that the SET mechanism is unfavorable compared to the polar mechanism via the nonradical nucleophilic attack by the C3^{CPG} atom.

CPG H transferred states (states 3)

In the hydrogen transfer process, we observed that one proton (H^+) binding to the C2^{CPG} atom in state **2** (H_{C2}) is transferred to the N5^{FAD} atom in state **3**. The relative stability in terms of energy difference between states **2a** and **2b** (ΔG(**2b**) – ΔG(**2a**) = 9 kcal mol^{-1}) is barely affected by the process, since upon overcoming the transition states (ΔG(**TS(2b, 3b)**) – ΔG(**TS(2a,3a)**) = 7 kcal mol^{-1}) the realized products are still separated by an identical energy difference (ΔG(**3b**) – ΔG(**3a**) = 9 kcal mol^{-1}). In these products, the conformers a and b differ only in the trans- and cis-configuration of the C1^{CPG}–N^{CPG} bond. The state **3a** results much more stable in comparison with states **1** and **2** (ΔG(**3a**) = –10 kcal mol^{-1}). The transition state **TS(2a,3a)** is, in turn, less energetically demanding than the **TS(1d, 2a)**, although the transition state **TS(2b, 3b)** is still high in energy, and, as such, comparable to **TS(1d,2a)**. Hence, a reaction channel passing across **TS(2b, 3b)** is relatively unfavorable compared to the reaction via **TS(2a,3a)**.

Imine intermediates (states 4)

The state **3** can realize an alternative protonation state, the imine state, in which the proton of the CPG amine is transferred to the FAD N5. This imine state is slightly more stable than the state **3a** (ΔG(**4a**) – ΔG(**3a**) = –1.5 kcal mol^{-1}). We can infer that water molecules are prone to access freely the CPG. In this respect, the water molecules indicated as Wat1 and Wat2 in Figure 1(B) can play an active role in the mechanism, if they can be positioned in the proper active places. This can occur, for instance, as a consequence of the incoming of additional water molecules forcing at least two of them to form the needed H-bond network. Water molecules in proximity of the CPG might mediate the proton transfer to the state **4**, although we could not clearly determine such a reaction pathway with the present model without including additional water molecules. In the state **3b**, the corresponding imine state (**4b**) is the more stable one (ΔG(**4b**) – ΔG(**3b**) = –6 kcal mol^{-1}). In the reduction step of L-proline in MSOX, a deprotonation of the amino group is required.[34] This suggests that a proton transfer is possible for the amine group of the substrate upon binding to the active site. Therefore, the enamine-imine tautomerization of **3** -> **4** is a viable process at the active site of MSOX.

Product states (states 5)

After the C4a^{CPG}-C3^{CPG} bond cleavage occurring in the state **3a**, a secondary amino acid product, ethenyl amino acetic acid, is formed in the state **5a**. Compared to the initial substrate of CPG, this amino acid product is two-electron oxidized as the H_{C2}^- moiety was transferred to the FAD flavin from CPG. The state **5a** turns out to be less stable in comparison with states **3a** and **4a** (ΔG(5a)–ΔG(4a) = 5 kcal mol^{-1}), and the activation barrier for accessing the **5a** state from the **3a** is rather modest (ΔG(**5a**) – ΔG(**3a**) = 7 kcal mol^{-1}). These two facts, taken together, suggest that the reaction from **3a** to **5a** does not occur spontaneously because the reverse reaction (**5a** -> **3a**) is faster compared to the forward reaction (**3a** -> **5a**). It should be noted that the product amino acid is still bound to the active site. The conclusion that can be drawn is that a dissociation of the product amino acid is essential to revert to the initial oxidized flavin ($EFAD_{ox}$) by reaction with O_2. The transition state **TS(3b, 5b)** is highly unfavored energetically (ΔG(**TS(2b, 3b)**) = 20 kcal mol^{-1}) mainly because the terminal $C3H_2$ moiety in CPG should undergo a large rotation of approximately 90º in the **3b** -> **5b** reaction.

UV/Vis spectra

Accurate spectroscopic measurements have been reported for MSOX[16,35] and these can provide a solid benchmark on which theoretical approaches can be tested. In the absence of CPG, the MSOX exhibits an intense absorption band at λ_{max}=454 nm ascribed to an unmodified oxidized flavin.[16] Upon CPG addition, a charge transfer band was observed in the 500-700 nm range,[35] accompanied by a shift of the absorption to λ_{max} = 422 nm. Under aerobic conditions, the species with λ_{max} = 422 nm accumulates and acts as a mechanism-based inhibitor for MSOX. This inactivated species is assigned to a CPG-modified flavin.[16] Under anaerobic conditions, MSOX carrying the CPG substrate show that the absorption spectra changes to broadband over the 400-500 nm range, which is assigned to the formation of a 1,5-dihydro-FAD ($EFADH_2$). Thus, the MSOX catalytic cycle becomes trapped at the final oxidation step which, in turn, regenerates the unmodified oxidized state (EFADox).

In order to verify the experimental UV/Vis spectra assignments for our simulated model, TDDFT calculations were done for all the intermediates identified in the present work. The results for the major states **0**, **1a**, **3a**, **4a** and **5a** are shown in Figure 5. For the sake of completeness, the UV/Vis spectra for all the intermediate states are reported in Figure S3 (ESI†). The state

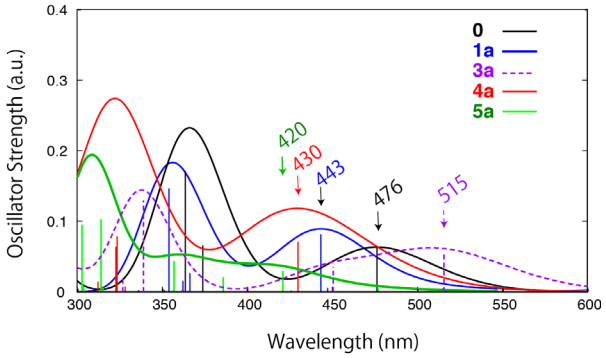


Figure 5. Computed UV/Vis spectra as provided by TDDFT (B3LYP/6-311++G**) calculations.

labeled as **0** is the unmodified MSOX without the CPG substrate. The calculated absorption peak for **0** is λ(**0**) = 476 nm, which is overestimated by 22 nm with respect to the experimental value, a fact not entirely surprising in TDDFT approaches. In the CPG binding MSOX, the state **1a**, the computed absorption peak is located at λ(**1a**) = 443 nm, qualitatively consistent with the peak shift toward shorter wavelengths. Inactive species candidates with adsorption spectra peaked at λ_{max} = 422nm are two states **3a** and **4a**. The calculated absorption peaks for **3a** are λ(**3a**) = 515 nm and 450 nm. We remark that also the **3b** state is characterized by adsorption peaks in the same range, namely λ(**3b**) = 586 nm and 451 nm. These wavelength absorptions are anyhow longer than λ(**1a**) = 443 nm for **1a** and, as such, inconsistent with the experimental results. Nonetheless, our computed absorption peaks for the states **3a**, within the limitations of the TDDFT approach, are perfectly reasonable and corroborated by the observation that strong π orbital interactions between the FAD flavin and CPG unsaturated part (-CH=CH-NH-) have been found for the HOMO and LUMO orbitals in the states **3a** and **3b**. These interactions can reduce the HOMO-LUMO gap and thus shift the transitions between these two states to lower frequencies. The calculated absorption peak for the state **4a**, the protonated tautomer of the state **3a**, is located at λ(**4a**) = 430 nm. From a free energy standpoint, the state **4a** is more stable than **3a**. In an attempt at getting a comprehensive scenario, we considered also a protonated state of **4a** (**4a+H⁺**), but in this case we obtained a substantially longer wavelength absorption peak (λ(**4a+H⁺**) = 524 nm). State **4b,** a conformer of **4a**, is also characterized by a similar absorption band peaked at λ(**4b**) = 418 nm. Therefore, we can identify the state **4a**, the imine state, as the experimentally observed inactivated species.

In state **5a** (green line in Figure 5) the absorption spectrum shows a long tail spreading in the wavelength range up to 500 nm. This feature is very similar to the experimental one for final state before the FAD oxidation. A broad absorption spectrum is found also for a model lacking the CPG oxidized product (**EFADH⁻** state). One additionally protonated state, the $EFADH_2$ state without the CPG oxidized product, exhibits a low absorption peak centered at λ(**$EFADH_2$**) = 482 nm. These calculated UV/Vis spectra in the absence of a substrate are reported in Figure S4 (ESI†). On these grounds, the $EFADH_2$ state can be ruled out as a possible final metastable state. The energy profile of Figure 4 shows that state **5a** can be rapidly converted into the inactive state **4a**, being the activation barrier for **5a** -> **TS(3a, 5a)** -> **3a** as low as ΔG = 2 kcal mol^{-1}. Therefore, we expect that the final metastable state before the FAD oxidation ($EFADH_2$ state in Schemes 1 and 2) can be assigned to the reduced FAD in the absence of the CPG substrate (**EFADH⁻** state), where O_2 can easily react with flavin to produce hydrogen peroxide (H_2O_2) and an oxidized flavin ($EFAD_{ox}$).

Reaction mechanism of the CPG oxidation by MSOX

The insight into the reaction mechanism of CPG in MSOX provided by the present study is the major focus of this section. The first step of this mechanism is the binding of CPG to the FAD flavin ring, a process that proceeds through the nucleophilic attack of the CPG C3 atom occurring simultaneously with the C3–C2 bond cleavage. This reaction can also be interpreted as a bond switch between the pristine C1-C3 bond in CPG and a newly formed CPG C3– FAD C4a bond. This reaction step, which passes through the most unstable state **TS(1d, 2a)**, is deeply different from the previously proposed single electron transfer process, in which a one-electron transfer occurs to CPG from the FAD flavin, and the cyclopropyl group opens forming a radical state. We have already shown that the radical intermediate states in SET are higher in energy than the **TS(1d, 2a)**. We recall also that a rotation of the amino group in CPG via state **1d** is crucial and has to be realized before the CPG adduct production in the state **2a**. Based on the reaction rate kinetics, the first slow step in Scheme 1 gives k_3 = 0.10 min^{-1},[16] corresponding to $\Delta G^{\neq}$ = 21.3 kcal mol^{-1} in terms of free energy. This activation barrier is comparable to the one calculated for the **TS(1d,2a)**, namely ΔG(**TS(1d,2a)**) = 16.2 kcal mol^{-1}.

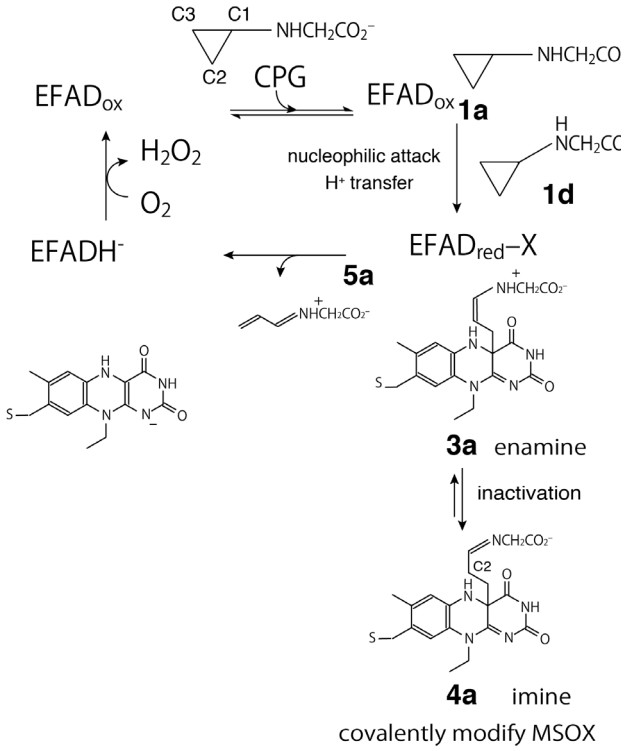
The second feature worthy of note is that the inactive species spectroscopically observed and having a spectrum peaked at λ_{max}=422nm is assigned to an imine state **4a**, and not to the amine state **3a**. The states **4a** and **3a** are related to the imine-enamine tautomers, and only the position of one proton is different. Their UV/Vis spectra are significantly different, and their adsorption peaks are centered on λ(**3a**) = 515 nm and λ(**4a**) = 430 nm. Hence, there is no room for doubt: The imine state **4a** is the inactive species. We remind that the intermediate state **4a** is the most stable among all the intermediate states **1**-**5** identified in this study, and the inactive species is positioned in such a way that it screens completely the active site (C4a and N5) which becomes inaccessible to O_2. The existence of this stable intermediate state of **4a** is consistent with the observation that the species **4a** is accumulated under aerobic conditions and the CPG acts as a mechanism-based inhibitor.

The third remarkable feature is that the following slow process $EFAD_{red}$-X -> $EFADH_2$ is expected to include not only the dissociation of the CPG oxidized product (**3a** -> **5a**), but also the release of the product from the active site (**5a** ->). As shown in the free energy profile of Figure 4, the bond cleavage (**3a** -> **TS(3a, 5a)** -> **5a**) does not correspond to the high activation barrier of $\Delta G^{\neq}$= 22.7 kcal mol^{-1}, expected from the reaction rate kinetics (k_4=0.01 min^{-1} in scheme1).[16] By looking at Figure S4, the UV/Vis spectra for the reduced flavin in states **5a**, **5b** and

EDADH$^-$ are not affected by the presence of the product in the active site. Thus, the last metastable state realized before the FAD oxidation process ($EFADH_2$ species in Schemes 1 and 2) is represented by the **EFADH**$^-$ state, where the oxidized CPG product is released from the active site. Since the product amino acid is tightly anchored to the surrounding base residues, Lys347 and Arg52, because of their strong electrostatic interactions, an additional activation barrier must be overcome in the product release process. The negatively charged flavin (**EFADH**$^-$) is well suited to release the substrate product with a negatively charged carboxy group.

The reaction mechanism we propose for the CPG oxidation in MSOX (**0** -> **1a** -> **1d** -> **2a** -> **3a** -> **4a** -> **3a** -> **5a** -> **EFADH**$^-$ -> **0**) is summarized in Scheme 3.

Scheme 3. Reaction mechanism of CPG in MSOX derived from the present QM/MM calculations.



The present QM/MM results clearly suggest that the reaction of MSOX with CPG can be categorized as a polar mechanism. The SET mechanism, for the reasons summarized in this work, can be ruled out the CPG oxidation reaction in MSOX. In our previous theoretical study,[5] we were able to show that sarcosine in MSOX proceeds through a HACET mechanism, which can be roughly categorized as a hydride transfer mechanism. In the HACET mechanism, sarcosine is needed to close to the flavin ring to form substantial orbital overlap between the amine π orbital and flavin aromatic π orbital. In particular, the approach of N_{amine} and C4a atoms is essential to low the activation barrier in the sarcosine oxidation. Based on these considerations, the HACET mechanism can be reinterpreted: The hydride transfer of the sarcosine oxidation in MSOX is stabilized by a polar mechanism through the N_{amine}-C4a interaction, although an amine adduct is not stably formed for the sarcosine substrate (ΔE = 16.3 kcal mol^{-1}, ΔE^{+} = 44.9 kcal mol^{-1}).[5] Based on the similarities between the reactions in sarcosine and CPG substrates, a polar mechanism is likely to play a key role also for other amine substrates in MSOX.

Conclusions

By resorting to accurate QM/MM molecular modelling, we unravelled the reaction mechanism of CPG in MSOX. The CPG substrate is widely used to trap radical species in enzymatic reactions, and the CPG has been identified as a prototype example of the SET mechanism in MSOX.

The amine group of CPG can take two conformations, resulting in the two alternative pathways. In the first one, the CPG NH amine points directly toward the O atom of flavin, whereas in the second one the CPG NH amine points away from this same O flavin atom. The former pathway has lower activation barriers for the transitions **2a** -> **3a** and **3a** -> **5a** compared to the latter one with higher activation barriers for the processes **2b** -> **3b** and **3b** ->**5b**. On these grounds, we can infer that the most favorable reaction pathway is the one realizing the subsequent steps **1a** -> **1d** -> **2a** -> **3a** -> **4a** -> **3a** -> **5a**. According to this mechanism, the rotations of the cyclopropyl and the CPG amine are crucial to trigger the **1a** -> **1d** step. From this point, the CPG-flavin interaction proceeds through a nucleophilic attack of the C3 atom belonging to the cyclopropyl group to the C4a atom of the flavin ring, thus realizing the **1d** -> **2a** step. The related activation barrier is ΔG(**TS(1d**, **2a)**) = 16.1 kcal mol^{-1}, namely the first slow step in the overall catalytic cycle ($\Delta G_1^{\neq}$=21.4 kcal mol^{-1}). The radical states are located at energetically higher values than the **TS(1d, 2a)**, which clearly suggests that the SET mechanism can be ruled out.

After a proton transfer (**2a** -> **3a**) and an enamine-imine transformation (**3a** -> **4a**), the imine state (**4a**) is formed, identified in our simulations as the most stable intermediate state. Such a high stability of the intermediate state can be ascribed to the inhibition of the catalytic activity. The imine species turns out to be the only possible one possessing the characteristic absorption peaks in a wavelength range (λ(**4a**) = 430 nm, λ(**4b**) = 418 nm) compatible with the experimental value λ_{max}= 422nm. This allows us to identify the inactivated species as the imine state (**4a**).

The second slow step after the formation of the inactivated species is a product release from the active site. This somehow escapes the capabilities of our QM/MM model. Nonetheless, the reaction steps **4a** -> **3a** -> **TS(3a, 5a)** -> **5a** could be traced and investigated. Indeed, we found that the associated activation barrier, $\Delta\Delta G$ = ΔG(**TS(3a,5a)**) – ΔG(**4a**) = 8.8 kcal mol^{-1}, is too low to be compared with the analogous barrier estimated from the kinetic rate ($\Delta G_2^{\neq}$ = 22.7 kcal mol^{-1}). This indicates that something is missing and that there should be an additional activation barrier to be overcome to access the product release. One possible explanation is that the carboxy group of the product is strongly anchored to the two base residues of Arg52 and Lys347; these strong interactions have to be destabilized somehow to make the release from the active

site feasible. The final intermediate state before the FAD oxidation has been identified as the **EFADH**$^{-}$ state in the absence of a substrate. This conclusion is supported by the computed UV/Vis spectra.

We have formerly shown that the key interaction in the sarcosine oxidation is the approach of the $N^{sarcosine}$ and $C4a^{flavin}$ atoms.[5] This is comparable to the nucleophilic attack of $C3^{CPG}$ to the $C4a^{flavin}$ atom evidenced in the present study. The explicit and implicit influences from the polar mechanism is expected to stimulate investigations in other amine oxidoreductases.

Acknowledgements

This research was supported by JSPS KAKENHI ground numbers 16KT0055, 17H04866, 18H05154, 19H05781 and JST, PRESTO Grant Number JPMJPR19G6, Japan. Numerical calculations were carried out under the support of (1) Multidisciplinary Cooperative Research Program in CCS, University of Tsukuba, (2) HPCI system research project (project ID: hp19110) using the computational resource of CX400 provided by the Information Technology Center in Nagoya University. M.B. thanks the HPC Mesocentera at the University of Strasbourg funded by the Equipex Equip@Meso project (Programme Investissements d'Avenir) and the CPER Alsacalcul/Big Data, and the Grand Equipement National de Calcul Intensif (GENCI) under allocation DARI-A6 A0060906092.

References

1 E. Romero, J. R. G. Castellanos, G. Gadda, M. W. Fraaije, A. Mattevi, *Chem. Rev.,* 2018, **118**, 1742-1769.
2 Y. Nishiya, S. Nakano, K. Kawamura, Y. Abe, *J. Anal. Bio-Sci.,* 2012, **35**, 426-430.
3 H. Suzuki, *Amino Acids,* 1994, **7**, 27-43.
4 R. B. Silverman, *Acc. Chem. Res.,* 1995, **28**, 335-342.
5 Y. Abe, M. Shoji, Y. Nishiya, H. Aiba, T. Kishimoto, K. Kitaura, *Phys. Chem. Chem. Phys.,* 2017, **19**, 9811-9822.
6 J. R. Miller, D. E., Edmondson, *Biochemistry,* 1999, **38**, 13670-13683.
7 E. Abad, R. K. Zenn, J. Kästner, *J. Phys. Chem. B,* 2013, **117**, 14238-14246.
8 R. B. Silverman, Y. Zelechonok. *J. Org. Chem.,* 1992, **57**, 6373-6374.
9 P. F. Fitzpatrick, *Arch. Biochem Biophys.,* 2010, **493**, 13-25.
10 A. Mattevi, M. A. Vanoni, F. Todone, M. Rizzi, A. Teplyakov, A. Coda, M. Bolognesi, B. Curti, *Proc. Natl. Acad. Sci. USA,* 1996, **93**, 7496-7501.
11 K. A. Kurtz, M. A. Rishavy, W. W. Cleland, P. F. Fitzptrick, *J. Am. Chem. Soc.,* 2000, **122**, 12986-12897.
12 S. S. Erdem, Ö. Karahan, I. Yildiz, K. Yelekci, *Org. Biomol. Chem.,* 2006, **4**, 646-658.
13 R. Harris, R. Meskys, M. J. Sutcliffe, N. S. Scrutton, *Biochemistry,* 2000, **39(6)**, 1189-1198.
14 G. Zhao, M. S. Jorns, *Biochemistry,* 2006, **45**, 5985-5992.
15 Z. Chen, G. Zhao, S. Martinovic, M. S. Jorns, F. S. Mathews, *Biochemistry,* 2005, **44**, 15444-15450.
16 G. Zhao, J. Qu, F. A. Davis, M. S. Jorns, *Biochemistry,* 2000, 39, 14341-14347.
17 J. M. Kim, S. E. Hoegy, P. S. Mariano, *J. Am. Chem. Soc.,* 1995, **117**, 100-105.
18 E. C. Ralph, J. S. Hirschi, M. A. Anderson, W. W. Cleland, D. A. Singleton, P. F. Fitzpatrick, *Biochemistry,* 2007, **46**, 7655-7664.
19 C. J. Suckling, *Angew. Chem. Int. Ed. Engl.,* 1988, **27**, 537-552.
20 J.-M. Kim, S. E. Hoegy, P. S. Mariano, *J. Am. Chem. Soc.,* 1993, **115(23)**, 10591-10595.
21 M. A. Wagner, P. Trickey, Z. Chen, F. S. Mathews, M. S. Jorns, *Biochemistry,* 2000, **39(30)**, 8813-8824.
22 J. M. Word, S. C. Lovell, J. S. Richardson, D. C. Richardson, *J. Biol. Biol.,* 1999, **285**, 1735-1747.
23 G. Zhao, M. S. Jorns, *Biochemistry,* 2005, 44, 16866-16874.
24 D.A. Case, T.A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvary, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P. A. Kollman, AMBER 11, University of California, San Francisco, 2010.
25 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, Gaussian 03, Gaussian, Inc., Wallingford CT, 2004.
26 M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J van. Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, W.A. de Jong, *Comput. Phys. Commun.,* 2010, **181**, 1477-1489.
27 A. H.- Abdalah, G. Zhao, M. S. Jorns, *Biochemistry,* 2006, **45**, 9454-9462.
28 S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104-19.
29 M. Wu, L. A. Eriksson, *J. Phys. Chem. A,* 2010, **114**, 10234-10242.
30 M. P. Kabir, Y. O- Gonzalez, S. Gozem, *Phys. Chem. Chem. Phys.,* 2019, **21**, 16526-16537.
31 K. Lincke, J. Langeland, A. O. Madsen, H. V. Kiefer, L. Skov, E. Gruber, K. V. Mikkelsen, L. H. Andersen, M. B. Nielsen, *Phys. Chem. Chem. Phys.,* 2018, **20**, 28678-28684.
32 W. Humphrey, A. Dalke,; K. Schulten, *J. Mole. Graphics,* 1996, **14**, 33-38.
33 M. Shoji, M. Kayanuma, Y. Shigeta, *Bull. Chem. Soc. Jap.,* 2018, **191(10)**, 1465-1473.
34 G. Zhao, M. S. Jorns, *Biochemistry,* 2002, **41**, 9747-9750.
35 G. Zhao, M. S. Jorns, *Biochemistry,* 2006, **45**, 5985-5992.