THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY



pubs.acs.org/JPCL
Letter

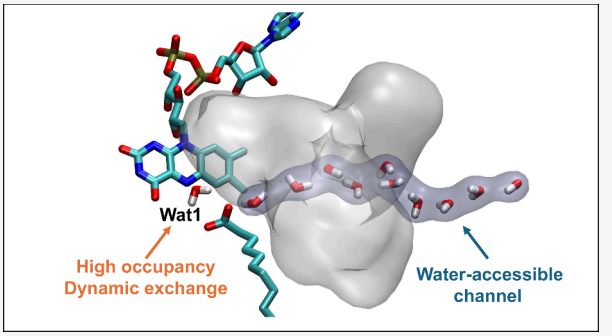
# Insights into Substrate Protonation and Solvent Accessibility in the Active Site of Fatty Acid Photodecarboxylase

Gianluca Dell'Orletta, Nico Di Fonte, Matteo Farina, and Isabella Daidone*

Cite This: J. Phys. Chem. Lett. 2025, 16, 12538−12544



**ABSTRACT:** Fatty acid photodecarboxylase catalyzes the light-driven decarboxylation of fatty acids into hydrocarbons via electron transfer (ET) from the substrate to the flavin adenine dinucleotide cofactor, proceeding through either proton-coupled ET or hydrogen-atom transfer. Through quantum mechanics/molecular mechanics calculations, we show that only the deprotonated fatty acid supports the charge-transfer states required to initiate catalysis. Molecular dynamics simulations combined with graph-theory-based analysis reveal that the crystallographic water site, Wat1, is consistently occupied by a stable, yet dynamically exchanging, population of water molecules. We also identify transient, solvent-accessible water channels connecting the active site to the bulk solvent, potentially facilitating proton transfer and water exchange. These findings support the notion that beyond preserving the structural integrity of the active site, this water population may also enable the flexible modulation of electron and proton transfer through an adaptive hydrogen-bonding network.





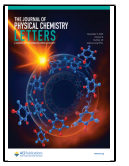

Photoenzymes have recently emerged as attractive catalysts in green chemistry due to their ability to harness light energy to drive selective chemical transformations.[1−4] To date, only three distinct types of natural photoenzymes have been identified. Among them, fatty acid photodecarboxylase (FAP) stands out as the only known flavoprotein that utilizes its photoexcited state to oxidatively generate substrate-centered radicals.[5] In its native role, FAP converts fatty acids into hydrocarbons and $CO_2$ with remarkable efficiency, achieving a quantum yield exceeding 80%.[6] Since its discovery in 2017,[6] FAP has become a focus of intense research aimed at elucidating its reaction mechanism[7−11] and applying this knowledge toward the rational design of biomimetic photocatalysts for sustainable chemistry.[12] A detailed understanding of the reaction mechanism is crucial for the development of next-generation FAP-inspired photocatalysts.

In the crystal structure of the enzyme−substrate complex of *Chlorella variabilis* FAP (*Cv*FAP), the fatty acid carboxylate group is located near the isoalloxazine chromophore of the FAD cofactor (see Figure 1).[6] Upon light absorption, oxidized FAD is excited to its first singlet state ($^1$FAD), which decays in around 300 ps to form the flavosemiquinone radical $FAD^{\bullet-}$ via a forward electron transfer (fET) from the substrate, as confirmed by transient fluorescence and absorption spectroscopy.[6,8,9] The resulting carboxyl radical ($RCOO^{\bullet}$) rapidly decarboxylates to generate alkyl radical $R^{\bullet}$ and $CO_2$. The $CO_2$ may be directly released or converted into $HCO_3^-$ before leaving the active site.[8] The catalytic cycle is then completed either by a back-electron transfer from the FAD cofactor coupled with a proton transfer (PCET) or via a hydrogen atom transfer (HAT), ultimately yielding the alkane product (RH). The reaction mechanism is summarized in Figure 1A.

It has been proposed that in the case of the PCET mechanism, this proton may originate from a nearby water molecule, which is subsequently reprotonated via a proton relay involving R451 and ultimately a proton from the bulk solvent.[8] An alternative hypothesis involves a HAT mechanism in which a hydrogen atom is directly donated by an amino acid side chain, such as C432. In this case, C432 is then reduced by an ET from the FAD cofactor and is later reprotonated, possibly from the bulk solvent.[7] In both scenarios, accessibility of the active site to the solvent is essential for the regeneration of the proton (or hydrogen) donor.

Two crystallographic water molecules (Wat1 and Wat2) have been identified near the fatty acid substrate and the FAD cofactor (see Figure 1B,C) and have been proposed to participate in catalysis.[8] Water may also play a role in inhibiting deactivation processes, such as preventing back electron transfer from $FAD^{\bullet-}$ to $RCOO^{\bullet}$ (through stabilization of $FAD^{\bullet-}$ via, for example, a rearrangement of Wat1[11]) or undesired electron transfer from $FAD^{\bullet-}$ to arginine R451 (for example, by shielding the positively charged R451 and the

Received: August 8, 2025
Revised: October 21, 2025
Accepted: October 30, 2025
Published: November 26, 2025







https://doi.org/10.1021/acs.jpclett.5c02464
J. Phys. Chem. Lett. 2025, 16, 12538−12544

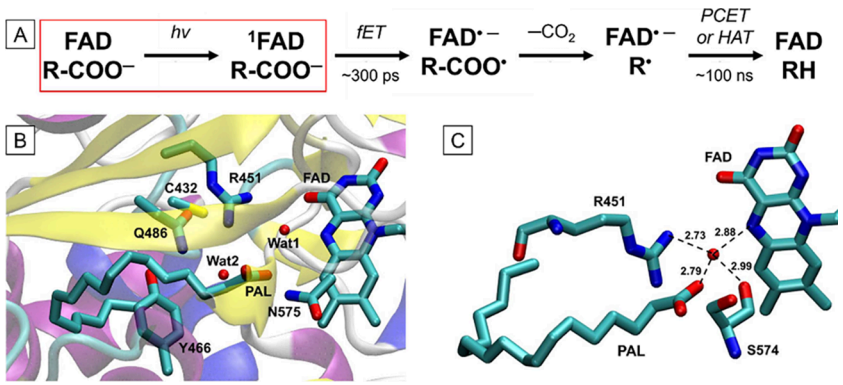



**Figure 1.** A) Schematic representation of the proposed *Cv*FAP photocatalytic cycle. The red rectangle highlights the first reaction step corresponding to the photoexcitation of the FAD cofactor. B) Close-up view of the catalytic site from the crystal structure obtained at 100 K in the dark (RCSB PDB entry: 6YRU).[8] Hydrogen atoms are not shown, as they were not resolved in X-ray crystallography. C) Close-up view of the residues surrounding Wat1, which may act as hydrogen bond donors or acceptors, stabilizing this water molecule in the active site.

negatively charged $FAD^{\bullet-}$).[13] However, these hypotheses are based on the crystal structure, whereas the structure and dynamics of the active site under physiological conditions may differ significantly and remain largely unexplored to date.

Despite the significant progress in elucidating the *Cv*FAP mechanism, two key questions remain unresolved. First, based on an analysis of interatomic distances in the crystal structure, the fatty acid could adopt either a protonated or deprotonated state, raising doubts about which form is catalytically competent.[9] Second, the mechanism by which water molecules reach the active site and potentially enable proton transfer from the bulk solvent remains unclear. Moreover, while crystallographic data suggest a potential catalytic or structural role for specific water molecules, e.g., Wat1, little is known about their stability and dynamics under physiological conditions.

In this work, we used quantum mechanical and molecular mechanical (QM/MM) calculations to elucidate the reactive protonation state of the fatty acid within the *Cv*FAP active site. Subsequently, we employed classical molecular dynamics (MD) simulations to investigate the structural and dynamic roles of water molecules in the enzyme, with a particular focus on Wat1. Finally, a graph-theory-based approach was used to identify potential water-accessible channels connecting the active site to the bulk solvent.

In previous computational investigations of the *Cv*FAP only the deprotonated fatty acid was considered.[8,10,11] In this interpretation of the crystallographic data at 100 K, the fatty acid is deprotonated and its carboxyl group is stabilized by hydrogen bonds with two water molecules (i.e., Wat1 and Wat2) and the side chains of R451 and N575.[8] However, visual inspection of the structure suggests that the fatty acid may also adopt a protonated state (see Figure 1C and Figures S1 and S2 in the Supporting Information), with the H of the carboxyl group forming a hydrogen bond with Wat1, which is stabilized by hydrogen bonds with N5 of the FAD cofactor and the oxygen of the backbone of S574, while the other oxygen is still stabilized by the interactions with R451 and N575.

We investigated whether both the deprotonated and protonated forms of the C16 fatty acid (palmitic acid, PAL) could participate in electron transfer to the FAD cofactor upon photoexcitation of *Cv*FAP. To this aim, we carried out two types of calculations considering both protonation states of PAL: (i) an electronic-structure analysis to determine the accessibility of a charge-transfer (CT) state involving electron transfer from the fatty acid to FAD, which triggers the fET initiating the photoreaction cycle; and (ii) evaluation of the $pK_a$ of the PAL substrate within the protein environment.

We first performed geometry optimizations of both forms at the Density Functional Theory (DFT) level using the QM/MM method (see Section 1.1 and Figure S3). A comparison between the X-ray crystal structure and the QM/MM-optimized structures of *Cv*FAP containing the deprotonated and protonated fatty acid is shown in Figure S4. Overall, the two optimized structures are highly similar, with no significant differences observed between the deprotonated and protonated states. In both cases, however, a slight rotation of the PAL headgroup is observed relative to the crystallographic structure. These data suggest that the protonation state of the fatty acid has minimal impact on the positioning of heavy atoms within the active site of the optimized crystallographic structure of *Cv*FAP. We then computed the 25 lowest-energy electronic states of the optimized geometries within the QM/MM framework at the time-dependent density-functional theory (TD-DFT) level[14] (see Section 1.1). The energies, oscillator strengths, nature of these excited states, and the CT character are reported in Tables S1 and S2. The nature of the excited states was characterized using Natural Transition Orbitals (NTOs)[15] and further confirmed by transition density matrix analysis performed with the TheoDORE code.[16,17] This analysis revealed that, in both structures, the first excited state is a locally excited (LE) bright state characterized by a $\pi \rightarrow \pi^*$ transition (see Figures 2A and 2B). The deprotonated structure exhibits three pure CT states from the fatty acid to FAD, with the lowest CT state at 4.56 eV. In contrast, the protonated structure shows such a CT state at 6.70 eV. To verify whether the large difference in CT energy also persists within the conformational ensemble sampled at 300 K, we extracted two representative structures from the MD simulations of the system containing the deprotonated and protonated forms of PAL (see Section 1.2 of the Supporting Information for the technical details) and computed their excited states. The results, reported in Table S3, confirm that the structure derived from the deprotonated system exhibits a low-energy CT state at 4.31 eV, whereas the protonated fatty acid does not show any CT from PAL to FAD up to 7.13 eV. The very high energy of the CT state in the protonated form suggests that this species is not catalytically competent in the


https://doi.org/10.1021/acs.jpclett.5c02464
J. Phys. Chem. Lett. 2025, 16, 12538−12544

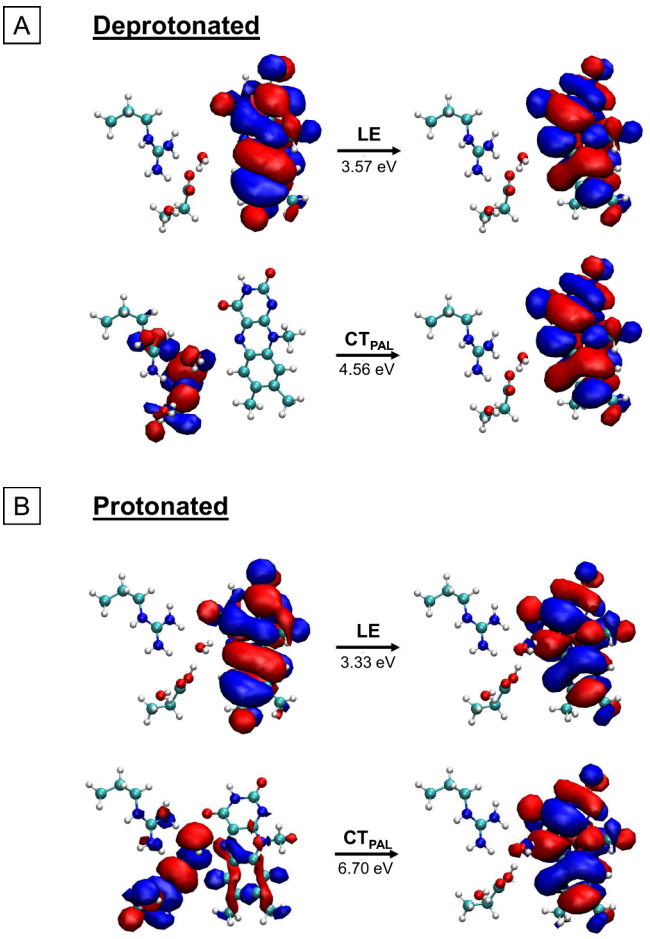



**Figure 2.** Natural Transition Orbitals (NTOs) of representative excited states of the active site region (which includes the side chain of R451, the isoalloxazine moiety of the FAD cofactor, the $C_2-COO^-$ fragment of the PAL within the active site, and the water molecules Wat1 and Wat2) within the *Cv*FAP enzyme with the PAL in two different protonation states: (A) deprotonated and (B) protonated. The excited states were computed at the QM/MM level based on the optimized structures derived from the experimental X-ray crystal structure. Excitation energies are indicated for each state.

*Cv*FAP-catalyzed reaction. Collectively, these findings indicate that both the protonation state and the conformational environment critically determine the accessibility of the photoinduced electron transfer step.

To further assess the protonation state of PAL within the active site, we performed p$K_a$ calculations using the PROPKA3 tool.[18,19] The analysis was carried out on the X-ray crystal structure, on ten representative structures extracted from the MD simulation of the deprotonated fatty acid, and on five structures extracted from the simulation of the protonated fatty acid to account for the different chemical environments surrounding the fatty acid in its deprotonated and protonated forms. The corresponding p$K_a$ values, summarized in Table S4, yielded an average value of 4.8 ± 0.2. Although this value should be interpreted with caution, taken together with the evidence provided by the excited-state calculations, it indicates that under the experimental (and simulated) working conditions (i.e., pH = 8.5), the substrate can be considered largely deprotonated.

Given the above results, all subsequent MD simulations and structural/dynamical analyses were therefore performed using the deprotonated fatty acid (see Section 1.2). After initial tests using the Amber ff14SB force field[20] for *Cv*FAP and the General Amber Force Field (GAFF)[21] for both the FAD cofactor and the fatty acid, we opted to employ the CHARMM36 force field[22] for the protein, the flavin-specific force field[23] for the oxidized FAD cofactor, and CHARMM General Force Field (CGenFF) parameters[24,25] for the C16 fatty acids in all MD simulations. This choice was motivated by the ability of the CHARMM-based force fields to more accurately reproduce the experimentally observed butterfly bending angle of the FAD cofactor at room temperature (13.7° from low-dose X-ray crystallography and 14.3°/11.7° from serial femtosecond crystallography, SFX)[8] (see Figure S5, *vide infra*). To evaluate the overall stability of the system during the production MD simulations, we monitored the time evolution of the root-mean-square deviation (RMSD) of (i) the protein heavy atoms with respect to the X-ray structure and (ii) the heavy atoms of the isoalloxazine moiety of the FAD cofactor. The corresponding RMSD plots are reported in Figures S6 and S7, showing that the structure of both the protein and the FAD cofactor remains stable along the trajectories.

To assess the stability of the fatty acid within the active site, we monitored two key parameters throughout the MD simulations: the distance between the N5 atom of the FAD isoalloxazine moiety and the carbon atom of the fatty acid carboxyl group (referred to as the N5−C distance) and the FAD butterfly bending angle defined by the C4−N5−N10−C9 dihedral. The time evolution of these parameters is shown in Figures S8 and S9. The distribution of the N5−C distance is shown in Figure 3A. The overall distribution displays a main peak at 4.7 Å, indicating that the fatty acid remains in close proximity to the FAD cofactor for most of the simulation time. A secondary peak at 7.4 Å arises from the displacement of the fatty acid in the 2PAL-2 and 2PAL-3 trajectories (Figure S8). In this alternative binding pose, which accounts for approximately 16% of all simulated configurations, the fatty acid forms electrostatic interactions with the guanidinium group of R451 and hydrogen bonds with Q486, Y466, and two water molecules. The vacant space near the FAD is subsequently occupied by additional water molecules (Figure S10). To further characterize this alternative configuration, we performed excited-state calculations on a representative structure extracted from this MD basin. The results, reported in Table S5, show that the CT state from PAL to FAD is located at a significantly higher energy (5.93 eV) than when PAL is positioned in proximity to FAD, indicating that electron transfer is unlikely in this binding pose.

Next, we investigated the occurrence of bidentate interactions between the fatty acid and the side chain of R451 (Figure S11). To this end, we monitored, across the production MD simulations, the number of hydrogen bonds formed between the carboxyl group of the fatty acid and the guanidinium group of R451. A hydrogen bond was considered present when the N−O distance was ≤ 3.5 Å and the N−H−O angle ranged between 150° and 180°. According to these criteria, bidentate interactions were observed in an average of 1.3% of the total simulation frames with a standard error of the mean (SEM) of 1.0% (the corresponding percentages for the single trajectories are reported in Table S6). Excited-state calculations on a representative bidentate structure (Table S5) revealed the CT state from PAL to FAD at 5.43 eV, again indicating that electron transfer is unlikely in this binding pose, as suggested in previous studies.[8,11] Overall, these results


https://doi.org/10.1021/acs.jpclett.5c02464
J. Phys. Chem. Lett. 2025, 16, 12538−12544

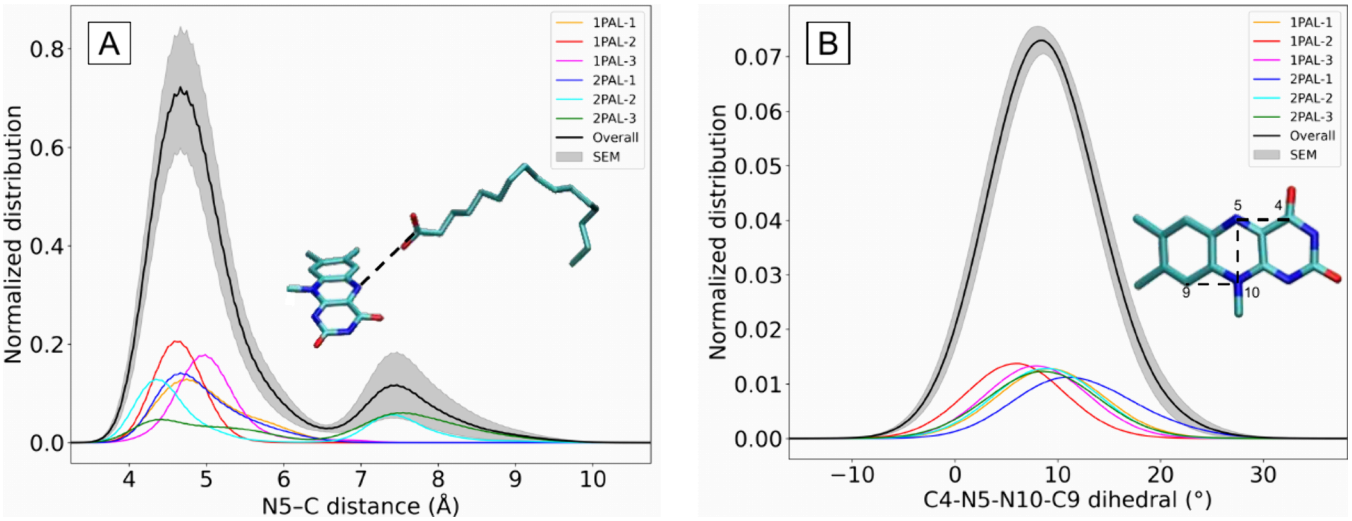



**Figure 3.** A) Distribution of the distance between the N5 atom of the FAD isoalloxazine ring and the carboxylic carbon of the fatty acid (N5−C distance), computed across the production MD simulations. The black line represents the overall average distribution, and the shaded area corresponds to the standard error of the mean (SEM). The main peak is located at 4.7 Å and the second one at 7.4 Å. B) Distribution of the C4−N5−N10−C9 dihedral angle of the FAD cofactor, computed across the production MD simulations. The black line represents the overall average distribution, and the shaded area indicates the SEM. For both analyses, the individual contributions from each simulation are also shown: 1PAL-1 (orange), 1PAL-2 (red), 1PAL-3 (magenta), 2PAL-1 (blue), 2PAL-2 (cyan), and 2PAL-3 (green). The inset shows the definition of the C4−N5−N10−C9 dihedral angle.

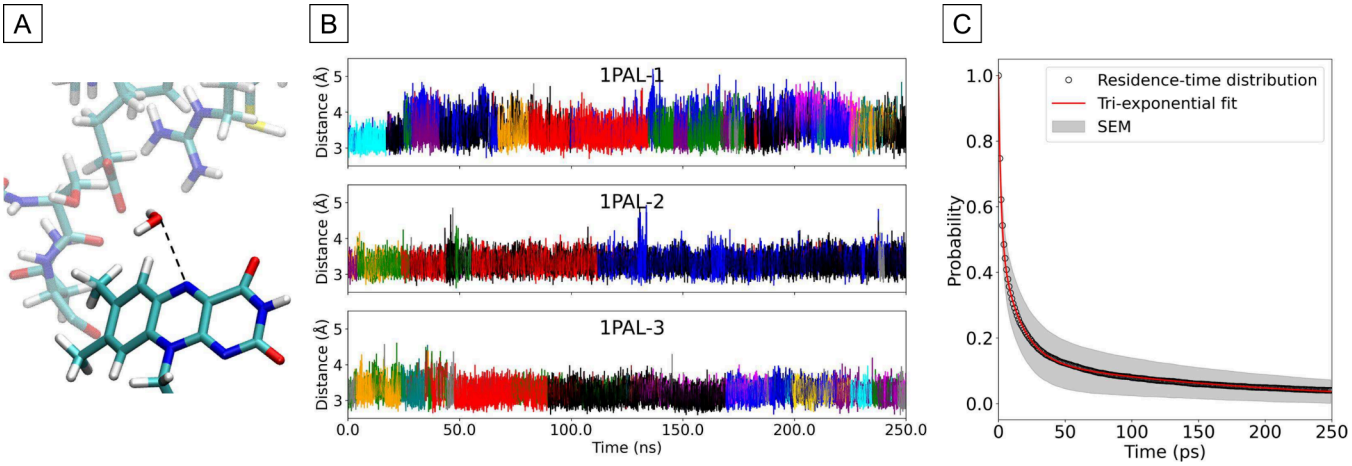



**Figure 4.** A) Molecular scheme showing the monitored distance between the water molecule occupying the Wat1 site and the N5 atom of the FAD cofactor. B) Time evolution of the distance between the N5 atom of FAD and the oxygen atom of the nearest water molecule in the three MD simulations with one PAL molecule (1PAL-1, 1PAL-2, and 1PAL-3). Different colors represent distinct water molecules that occupied the Wat1 site for more than 1.5% of the simulation time; water molecules with occupancies below 1.5% are shown in gray. C) Overall residence-time distribution (black line) averaged over all six MD simulations and the triexponential curve used to fit the data (red line). The shaded area indicates the SEM of the simulated points. The fit yields a fast component ($\tau_1$ = 1.8 ± 0.2 ps) and two slower components ($\tau_2$ = 15 ± 4 ps and $\tau_3$ = 189 ± 49 ps). An explanation for the high error on the slowest component is reported in Section 1.3.1.

suggest that both the displaced binding pose and the structures in which PAL forms a bidentate interaction with R451 are not catalytically competent within the photocycle. Their occurrence (together accounting for approximately 17% of the sampled configurations) may help rationalize why the quantum yield exceeds 80% but does not reach 100%.

The distribution of the C4−N5−N10−C9 dihedral angle across the MD simulations is reported in Figure 3B. The distributions are broad, with peak maxima ranging from 7° to 12°. These values are in close agreement with the low-dose crystallography experiment at room temperature, which reports a bending angle of 13.7°,[8] and SFX data for the dark-state structure of *Cv*FAP at room temperature, which reveal angles of 14.3° and 11.7° for molecules A and B. These correspond to the two copies present in the asymmetric unit of the *Cv*FAP microcrystals.[8] However, our values are approximately 10° smaller than those reported in the Supporting Information of the same study, suggesting that in our simulations FAD adopts a slightly more planar conformation compared to both the low-temperature crystallographic structure and the MD simulations presented in that work.

Given Wat1's key role in the catalytic cycle, particularly in modulating the CT state energy and potentially assisting bicarbonate formation following the PCET step,[8] we examined the stability of water molecules occupying this site. Figure 4B and Figure S12 show the time evolution of the distance


https://doi.org/10.1021/acs.jpclett.5c02464
J. Phys. Chem. Lett. 2025, 16, 12538−12544

between Wat1 candidates and the N5 atom of FAD for the 1PAL and 2PAL simulations, respectively. Across each of the production MD trajectories, no more than twenty-two distinct water molecules occupied the Wat1 site for a significant duration, indicating overall high occupancy of individual molecules. On average, a hydrogen bond between Wat1 and the N5 atom of the FAD cofactor was observed in 7% (SEM = 2%) of all the configurations (see Table S7, which reports this percentage for each individual MD simulation). From a dynamic point of view, fast exchange events are observed, revealing a high mobility of water molecules at this position (see the following for a quantitative analysis). A similar dynamic behavior of water molecules, yet characterized by relatively stable occupancy, has also been observed in other enzymatic systems.[26−30] These findings indicate that even at room temperature, the water molecule at the Wat1 site retains a structural role in the dark state of *Cv*FAP, while also exhibiting dynamic exchange behavior. As observed in other biological systems,[31−33] such dynamical water molecules may modulate the electron transfer from the fatty acid to the cofactor. Moreover, their ability to reconfigure hydrogen-bond networks in response to the catalytic cycle[34,35] may enable efficient proton transfer and facilitate PCET reactions.

To quantify the dynamical behavior, we next calculated the residence time of water molecules at the Wat1 site (see Figure 4C and Figure S13), following the procedure described in Section 1.3 of the Supporting Information. The overall residence-time distribution obtained averaging over all production MD simulations was fitted with a triexponential function, as previously applied to spectroscopic data on solvation/desolvation dynamics in a flavodoxin[36] (see Table S8 for the fitting parameters of the individual MD trajectories). The fit yielded a fast component ($\tau_1$ = 1.8 ± 0.2 ps) and two slower components ($\tau_2$ = 15 ± 4 ps and $\tau_3$ = 189 ± 49 ps), which closely match the experimental time scales observed in the flavodoxin active site. In this study, the ultrafast 1−3 ps phase was attributed to local water-network relaxation, while the 20−40 ps and hundreds of picosecond phases were linked to coupled fluctuations of the local water and protein environment.[36] It should be noted that the fastest and slowest components, $\tau_1$ and $\tau_3$, lie near the temporal resolution of our MD trajectories (1 ps) and the length of the subtrajectories used to evaluate residence times (250 ps), respectively. These limitations affect the confidence in their precise quantification. Nonetheless, the estimated order of magnitude remains reliable.

Finally, to identify potential water channels connecting the active site of *Cv*FAP to the bulk solvent, we developed a Python-based tool, which we named SProtein, that applies graph theory to networks of water molecules sampled during MD simulations. Details of the method are given in Supporting Information, Section 1.4.1. The code is publicly available on GitHub (https://github.com/NDiFonte/SProtein). While several studies have explored the connectivity of water molecules between protein residues,[37−41] and tools such as Bridges[38] and Cgraphs[39] and have been developed to characterize these water-mediated motifs using graph-theoretical approaches, there remains a lack of methods that systematically identify solvent-accessible pathways connecting the protein interior to the bulk solvent from MD-derived water networks. One such method is WatFinder,[42] which, however, is not specifically designed to identify water pathways connecting a given point within the protein to the bulk solvent, making it less straightforward to use for this purpose. A brief comparison between WatFinder and our approach is provided in the Supporting Information, Section 1.4.2, and Figure S14.

The SProtein analysis was performed with a time resolution of 100 ps, resulting in a total of 2501 frames for each trajectory. For each analyzed frame of the MD trajectories, we constructed an undirected graph in which nodes represent oxygen atoms of water molecules. Degree centrality (see Section 1.4.1 and Figure S15) was used to determine when a channel reached the bulk solvent: water molecules within 10 Å of the protein surface were considered and classified as "bulk" if they exhibited a degree ≥ 7 when considering only neighboring water molecules. The starting node for channel identification was defined as follows: if the distance between the N5 atom of the FAD cofactor and the carboxylic carbon of the fatty acid was ≤ 6 Å, the O1 atom of the fatty acid was selected as the starting node; otherwise, the N5 atom of the cofactor was used to ensure that the identified paths effectively connected to the protein's active site. The Dijkstra's algorithm[43] was then used to compute the shortest paths from the starting node to water molecules defined as "bulk". Once the water channels were identified, the protein residues lining each channel were determined by calculating the distance between each water molecule in the channel and nearby amino acids. Residues whose centers of mass lay within 7 Å of any water molecule along the pathway were selected as lining residues.

The most frequently sampled channel is shown in Figure 5. On average, this channel was observed in 14% of the total MD frames, with a SEM of 7% (the percentage for each MD simulation is reported in Table S9). It can be noted that the channel was observed much more frequently when the N5−C distance between PAL and the FAD cofactor was at least 6 Å

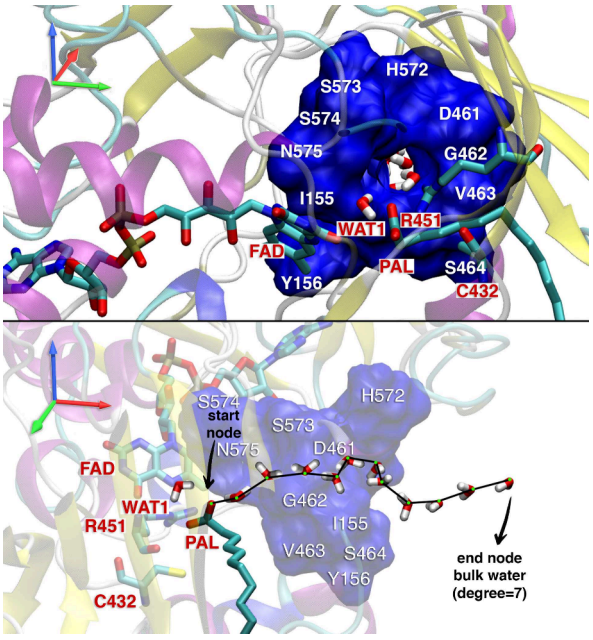



**Figure 5.** Water-accessible channel most frequently sampled along the MD trajectories as identified in our SProtein analysis. Two different perspectives are shown. The blue surface highlights the amino acid residues defining the channel, i.e., those contributing to its formation in at least 5% of the frames in which the channel was observed: I155, Y156, D461, G462, V463, S464, H572, S573, S574, and N575.



https://doi.org/10.1021/acs.jpclett.5c02464
J. Phys. Chem. Lett. 2025, 16, 12538−12544

(48%) than when PAL was close to FAD (6%). This channel is defined by residues I155, Y156, D461, G462, V463, S464, H572, S573, S574, and N575 (Figure 5). The frequencies of occurrence of these residues across all MD frames in which the channel is present in each MD trajectory and the corresponding weighted averages are reported in Figures S16 and S17, respectively. The more buried residues located near the active site, namely, S573, S574, G462, V463, and N575, exhibit the highest occurrence. All other channels contributed to less than 1% of the frames in the individual MD simulations. An example is shown in Figure S18.

The identification of water channels in *Cv*FAP, reported here for the first time, is significant because they allow water molecules to access the active site and may facilitate proton transfer, in line with previous studies on water-mediated proton transport and internal water dynamics in proteins.[44−46] Although our current simulations capture only the dark state of the enzyme and therefore cannot definitively predict the behavior of the water channels throughout the entire catalytic cycle, our results indicate that water molecules can access the active site and occupy the Wat1 position. These indicate that such access is achievable through the observed channels, while we cannot exclude the possibility that water may also enter via alternative pathways.

Using a combination of QM/MM calculations and molecular dynamics simulations, we elucidated the structural and dynamic factors underlying the resting state and, to a partial extent, the initial forward electron transfer in the photocatalytic mechanism of *Cv*FAP. Our QM/MM results and p$K_a$ estimates confirm that only the deprotonated form of the fatty acid can initiate the photoreaction. We then examined the structural and dynamical properties of water molecules within the active site. MD simulations revealed that the Wat1 site is consistently occupied by a high yet dynamically exchanging population of water molecules. This supports a dual role for Wat1 in maintaining active site structural integrity under physiological dark-state conditions while exhibiting significant dynamical flexibility. Additionally, using a graph-theory-based approach, we identified dynamic solvent-accessible water channels connecting the active site to bulk solvent, with one such channel consistently observed across simulations. These findings suggest a mechanism by which water molecules can access the active site, potentially facilitating reprotonation of catalytic residues and sustaining turnover during the photocatalytic cycle.

Finally, we estimated the residence time of water molecules at the Wat1 site. The observed multiexponential residence-time distribution, spanning from a few to several hundred picoseconds, closely matches experimental time scales reported for water dynamics in other flavoproteins.[36] This agreement not only validates the accuracy of our simulations but also reinforces the notion that the solvation of enzyme active sites is tightly coupled to protein dynamics.

## ■ ASSOCIATED CONTENT

### Data Availability Statement

The python code is publicly available on GitHub: https://github.com/NDiFonte/SProtein.

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.5c02464.

Details of the computational methods together with additional figures and tables (PDF)

Transparent Peer Review report available (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

**Isabella Daidone** − *Department of Physical and Chemical Sciences, University of L'Aquila, 67100 L'Aquila, Italy;* orcid.org/0000-0001-8970-8408; Email: isabella.daidone@univaq.it

### Authors

**Gianluca Dell'Orletta** − *Department of Physical and Chemical Sciences, University of L'Aquila, 67100 L'Aquila, Italy;* orcid.org/0009-0009-4144-0903

**Nico Di Fonte** − *Department of Physical and Chemical Sciences, University of L'Aquila, 67100 L'Aquila, Italy*

**Matteo Farina** − *Department of Chemistry, University of Rome La Sapienza, 00185 Rome, Italy*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.5c02464


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors acknowledge the CINECA award under the ISCRA initiative for the high-performance computing resources. The authors also acknowledge the use of ChatGPT (https://chat.openai.com/) to support writing style improvements. I.D. acknowledges funding from the European Union − NextGenerationEU under the Italian Ministry of University and Research (MUR), National Innovation Ecosystem grant (ECS00000041 VITALITY, CUP E13C22001060006). G.D. and I.D. acknowledge support from a PhD fellowship funded under the Italian National Recovery and Resilience Plan (PNRR), CUP E11I23000150001. I.D. acknowledges funding from the European Union − NextGenerationEU under the MUR PRIN2022 program (project P2022MC742PNRR, CUP E53D23018440001). The authors thank Laura Zanetti-Polzi for helpful feedback on the manuscript.

## ■ REFERENCES

(1) Taylor, A.; Heyes, D. J.; Scrutton, N. S. Catalysis by Nature's photoenzymes. *Curr. Opin. Struct. Biol.* **2022**, *77*, 102491.

(2) Harrison, W.; Huang, X.; Zhao, H. Photobiocatalysis for abiological transformations. *Acc. Chem. Res.* **2022**, *55*, 1087−1096.

(3) Schmermund, L.; Jurkas, V.; Özgen, F. F.; Barone, G. D.; Büchsenschütz, H. C.; Winkler, C. K.; Schmidt, S.; Kourist, R.; Kroutil, W. Photo-biocatalysis: biotransformations in the presence of light. *ACS Catal.* **2019**, *9*, 4115−4144.

(4) Emmanuel, M. A.; Bender, S. G.; Bilodeau, C.; Carceller, J. M.; DeHovitz, J. S.; Fu, H.; Liu, Y.; Nicholls, B. T.; Ouyang, Y.; Page, C. G.; et al. Photobiocatalytic strategies for organic synthesis. *Chem. Rev.* **2023**, *123*, 5459−5520.

(5) Moulin, S. L.; Beyly-Adriano, A.; Cuiné, S.; Blangy, S.; Légeret, B.; Floriani, M.; Burlacot, A.; Sorigué, D.; Samire, P.-P.; Li-Beisson, Y.; et al. Fatty acid photodecarboxylase is an ancient photoenzyme that forms hydrocarbons in the thylakoids of algae. *Plant Physiol.* **2021**, *186*, 1455−1472.

(6) Sorigué, D.; Légeret, B.; Cuiné, S.; Blangy, S.; Moulin, S.; Billon, E.; Richaud, P.; Brugière, S.; Couté, Y.; Nurizzo, D.; et al. An algal photoenzyme converts fatty acids to hydrocarbons. *Science* **2017**, *357*, 903−907.

(7) Heyes, D. J.; Lakavath, B.; Hardman, S. J.; Sakuma, M.; Hedison, T. M.; Scrutton, N. S. Photochemical mechanism of light-driven fatty acid photodecarboxylase. *ACS Catal.* **2020**, *10*, 6691−6696.
(8) Sorigué, D.; Hadjidemetriou, K.; Blangy, S.; Gotthard, G.; Bonvalet, A.; Coquelle, N.; Samire, P.; Aleksandrov, A.; Antonucci, L.; Benachir, A.; et al. Mechanism and dynamics of fatty acid photodecarboxylase. *Science* **2021**, *372*, No. eabd5687.
(9) Wu, R.; Li, X.; Wang, L.; Zhong, D. Ultrafast dynamics and catalytic mechanism of fatty acid photodecarboxylase. *Angew. Chem., Int. Ed.* **2022**, *61*, No. e202209180.
(10) Aleksandrov, A.; Bonvalet, A.; Müller, P.; Sorigué, D.; Beisson, F.; Antonucci, L.; Solinas, X.; Joffre, M.; Vos, M. H. Catalytic mechanism of fatty acid photodecarboxylase: on the detection and stability of the initial carbonyloxy radical intermediate. *Angew. Chem., Int. Ed.* **2024**, *63*, No. e202401376.
(11) Londi, G.; Salvadori, G.; Mazzeo, P.; Cupellini, L.; Mennucci, B. Protein-Driven Electron-Transfer Process in a Fatty Acid Photodecarboxylase. *JACS Au* **2025**, *5*, 158−168.
(12) Guo, X.; Xia, A.; Zhang, W.; Huang, Y.; Zhu, X.; Zhu, X.; Liao, Q. Photoenzymatic decarboxylation: A promising way to produce sustainable aviation fuels and fine chemicals. *Bioresour. Technol.* **2023**, *367*, 128232.
(13) Wu, Y.; Paul, C. E.; Hollmann, F. Stabilisation of the fatty acid decarboxylase from Chlorella variabilis by caprylic acid. *ChemBioChem.* **2021**, *22*, 2420−2423.
(14) Hirata, S.; Head-Gordon, M. Time-dependent density functional theory within the Tamm−Dancoff approximation. *Chem. Phys. Lett.* **1999**, *314*, 291−299.
(15) Martin, R. L. Natural transition orbitals. *J. Chem. Phys.* **2003**, *118*, 4775−4777.
(16) Plasser, F. TheoDORE: A toolbox for a detailed and automated analysis of electronic excited state computations. *J. Chem. Phys.* **2020**, *152*. DOI: 10.1063/1.5143076
(17) Plasser, F.; Lischka, H. Analysis of excitonic and charge transfer interactions from quantum chemical calculations. *J. Chem. Theory Comput.* **2012**, *8*, 2777−2789.
(18) Søndergaard, C. R.; Olsson, M. H.; Rostkowski, M.; Jensen, J. H. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of p K a values. *J. Chem. Theory Comput.* **2011**, *7*, 2284−2295.
(19) Olsson, M. H.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: consistent treatment of internal and surface residues in empirical p K a predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.
(20) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(21) 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.* **2004**, *25*, 1157−1174.
(22) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; De Groot, B. L.; Grubmüller, H.; MacKerell, A. D., Jr CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71−73.
(23) Aleksandrov, A. A molecular mechanics model for flavins. *J. Comput. Chem.* **2019**, *40*, 2834−2842.
(24) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671−690.
(25) Soteras Gutiérrez, I.; Lin, F.-Y.; Vanommeslaeghe, K.; Lemkul, J. A.; Armacost, K. A.; Brooks III, C. L.; MacKerell, A. D., Jr Parametrization of halogen bonds in the CHARMM general force field: Improved treatment of ligand−protein interactions. *Bioorg. Med. Chem.* **2016**, *24*, 4812−4825.
(26) Knight, J. D.; Hamelberg, D.; McCammon, J. A.; Kothary, R. The role of conserved water molecules in the catalytic domain of protein kinases. *Proteins: Struct., Funct., Bioinf.* **2009**, *76*, 527−535.
(27) Hanoian, P.; Hammes-Schiffer, S. Water in the active site of ketosteroid isomerase. *Biochemistry* **2011**, *50*, 6689−6700.
(28) Rydberg, P.; Rod, T. H.; Olsen, L.; Ryde, U. Dynamics of water molecules in the active-site cavity of human cytochromes P450. *J. Phys. Chem. B* **2007**, *111*, 5445−5457.
(29) Chakrabarti, B.; Bairagya, H. R.; Mukhopadhyay, B. P.; Sekar, K. New biochemical insight of conserved water molecules at catalytic and structural Zn2+ ions in human matrix metalloproteinase-I: a study by MD-simulation. *J. Mol. Model.* **2017**, *23*, 57.
(30) Barman, A.; Smitherman, C.; Souffrant, M.; Gadda, G.; Hamelberg, D. Conserved hydration sites in Pin1 reveal a distinctive water recognition motif in proteins. *J. Chem. Inf. Model.* **2016**, *56*, 139−147.
(31) Migliore, A.; Corni, S.; Di Felice, R.; Molinari, E. Water-mediated electron transfer between protein redox centers. *J. Phys. Chem. B* **2007**, *111*, 3774−3781.
(32) Bortolotti, C. A.; Amadei, A.; Aschi, M.; Borsari, M.; Corni, S.; Sola, M.; Daidone, I. The reversible opening of water channels in cytochrome c modulates the heme iron reduction potential. *J. Am. Chem. Soc.* **2012**, *134*, 13670−13678.
(33) Sirohiwal, A.; Neese, F.; Pantazis, D. A. Microsolvation of the redox-active tyrosine-D in photosystem II: correlation of energetics with EPR spectroscopy and oxidation-induced proton transfer. *J. Am. Chem. Soc.* **2019**, *141*, 3217−3231.
(34) Sharma, V.; Enkavi, G.; Vattulainen, I.; Róg, T.; Wikström, M. Proton-coupled electron transfer and the role of water molecules in proton pumping by cytochrome c oxidase. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 2040−2045.
(35) Sirohiwal, A.; Pantazis, D. A. Functional water networks in fully hydrated photosystem II. *J. Am. Chem. Soc.* **2022**, *144*, 22035−22050.
(36) Chang, C.-W.; He, T.-F.; Guo, L.; Stevens, J. A.; Li, T.; Wang, L.; Zhong, D. Mapping solvation dynamics at the function site of flavodoxin in three redox states. *J. Am. Chem. Soc.* **2010**, *132*, 12741−12747.
(37) Shinobu, A.; Agmon, N. Proton wire dynamics in the green fluorescent protein. *J. Chem. Theory Comput.* **2017**, *13*, 353−369.
(38) Siemers, M.; Lazaratos, M.; Karathanou, K.; Guerra, F.; Brown, L. S.; Bondar, A.-N. Bridge: A graph-based algorithm to analyze dynamic H-bond networks in membrane proteins. *J. Chem. Theory Comput.* **2019**, *15*, 6781−6798.
(39) Bondar, A.-N. Graphs of hydrogen-bond networks to dissect protein conformational dynamics. *J. Phys. Chem. B* **2022**, *126*, 3973−3984.
(40) Gorriz, R. F.; Volkenandt, S.; Imhof, P. Protonation-state dependence of hydration and interactions in the two proton-conducting channels of cytochrome c oxidase. *Int. J. Mol. Sci.* **2023**, *24*, 10464.
(41) Jain, H.; Karathanou, K.; Bondar, A.-N. Graph-Based Analyses of Dynamic Water-Mediated Hydrogen-Bond Networks in Phosphatidylserine: Cholesterol Membranes. *Biomolecules* **2023**, *13*, 1238.
(42) Krieger, J. M.; Doljanin, F.; Bogetti, A. T.; Zhang, F.; Manivarma, T.; Bahar, I.; Mikulska-Ruminska, K. WatFinder: a ProDy tool for protein−water interactions. *Bioinformatics* **2024**, *40*, btae516.
(43) Cherkassky, B. V.; Goldberg, A. V.; Radzik, T. Shortest paths algorithms: Theory and experimental evaluation. *Math. Program.* **1996**, *73*, 129−174.
(44) Reidelbach, M.; Betz, F.; Mäusle, R. M.; Imhof, P. Proton transfer pathways in an aspartate-water cluster sampled by a network of discrete states. *Chem. Phys. Lett.* **2016**, *659*, 169−175.
(45) Guerra, F.; Siemers, M.; Mielack, C.; Bondar, A.-N. Dynamics of long-distance hydrogen-bond networks in photosystem II. *J. Phys. Chem. B* **2018**, *122*, 4625−4641.
(46) Reidelbach, M.; Imhof, P. Proton transfer in the D-channel of cytochrome c oxidase modeled by a transition network approach. *Biochim. Biophys. Acta* **2020**, *1864*, 129614.