CO2 Migration Pathways in Oxalate Decarboxylase and Clues about Its Active Site

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · SEPTEMBER 2013

Impact Factor: 3.3 · DOI: 10.1021/jp4074834 · Source: PubMed

3 AUTHORS, INCLUDING:

Tarak Karamakar
Jawaharlal Nehru Centre for Advanced Scienti…
6 PUBLICATIONS **8** CITATIONS

SEE PROFILE

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

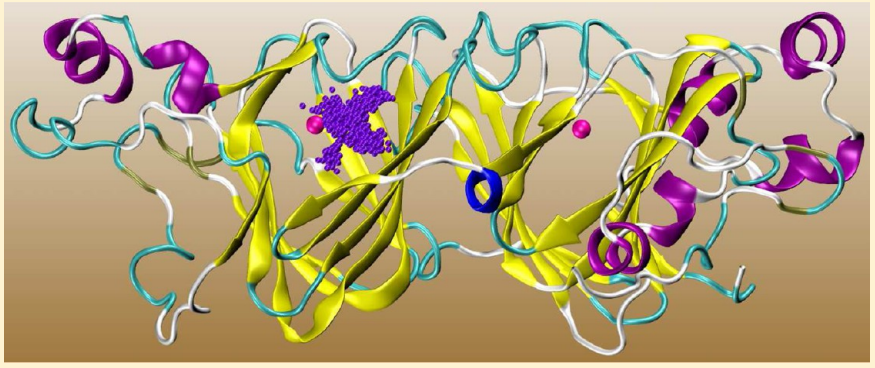
pubs.acs.org/JPCB

CO_2 Migration Pathways in Oxalate Decarboxylase and Clues about Its Active Site

Tarak Karmakar, Ganga Periyasamy, and Sundaram Balasubramanian*

Chemistry and Physics of Materials Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064, India

Supporting Information



ABSTRACT: Oxalate decarboxylase catalyzes the decarboxylation of oxalate to formate and CO_2 in the presence of molecular oxygen. This enzyme has two domains, each containing a Mn(II) ion coordinated with three histidine residues. The specific domain in which the decarboxylation process takes place is still a matter of investigation. Herein, the transport of the product, i.e., CO_2, from the reaction center to the surface of the enzyme is studied using atomistic molecular dynamics simulations. The specific pathway for the migration of the molecule as well as its microscopic interactions with the amino acid residues lining the path is delineated. Further, the transport of CO_2 is shown to occur in a facile manner from only domain I and not from domain II, indicating that the former is likely to be the active site of the enzyme.

INTRODUCTION

Oxalic acid is a product of cellular metabolism and accumulates in many plants such as spinach, bean, tomato, sunflower, etc. Excess intake of oxalate through these dietary sources leads to the formation of a calcium oxalate complex as kidney stones (urolithiasis), liver stones, etc.,[1] in humans and other vertebrates. *Oxalobactor fermigenis*, a bacterium that mainly colonizes the gastrointestinal tract of vertebrates, produces two enzymes, formyl-CoA transferase and oxalyl-CoA decarboxylase.[2] These enzymes indirectly degrade oxalate by performing consecutive enzymatic reactions. Oxalate decarboxylase (OXDC) is a Mn-containing enzyme that catalyzes the direct decarboxylation of oxalate to produce formate and CO_2.

There are several proteins, for example, carbonic anhydrase, that utilize CO_2 as a substrate and convert it into some other chemical species such as carbonic acid or carbonate.[3] On the other hand, there are some other enzymes[4] that produce CO_2 as a reaction product. OXDC falls in the latter category and releases CO_2 and formate (FMT). Moreover, the transport of CO_2 in a biological system is an important process that needs to be studied in its own respect and has attracted recent interest.[5−10]

OXDC has been well characterized in several fungi such as *Aspergillus niger*,[11] *Postia placenta*,[12] *Flammulina velutipes*,[13] and *Dichomitus squalens*.[14] The structure of the first bacterial OXDC, from *Bacillus subtilis*, was determined and studied by Anand et al. in 2002.[15] OXDC is a protein of the cupin superfamily and exists in the bicupin structure (see Figure 1). Under physiological conditions, this enzyme exists as a hexamer. Each monomer contains two bicupin domains, which are the N-terminus (domain I) and C-terminus (domain II). Both domains comprise mainly β-sheets and are connected by crossover loops. Each domain contains one Mn(II) ion[16] embedded within the protein interior. The distance between the two Mn(II) ions is nearly 26 Å. The oxidation state and the octahedral coordination environment for the Mn(II) center were proposed on the basis of electron paramagnetic resonance (EPR)[17] and X-ray absorption spectroscopy (XAS) spectroscopic studies.[18] Three histidine amino acids are coordinated to each Mn ion.[13] In domain I, the rest of the coordination environment is satisfied by one formate, one glutamic acid, and one water molecule. In domain II, one glutamic acid and two water molecules satisfy the Mn(II) valency. Further, experimental studies (EPR, X-ray diffraction, etc.) have confirmed the coordination of formate with Mn(II) through one oxygen; i.e., a monovalent ligation motif has been established.[13,17,18]

Received: July 27, 2013
Revised: September 3, 2013

   dx.doi.org/10.1021/jp4074834 | J. Phys. Chem. B XXXX, XXX, XXX−XXX

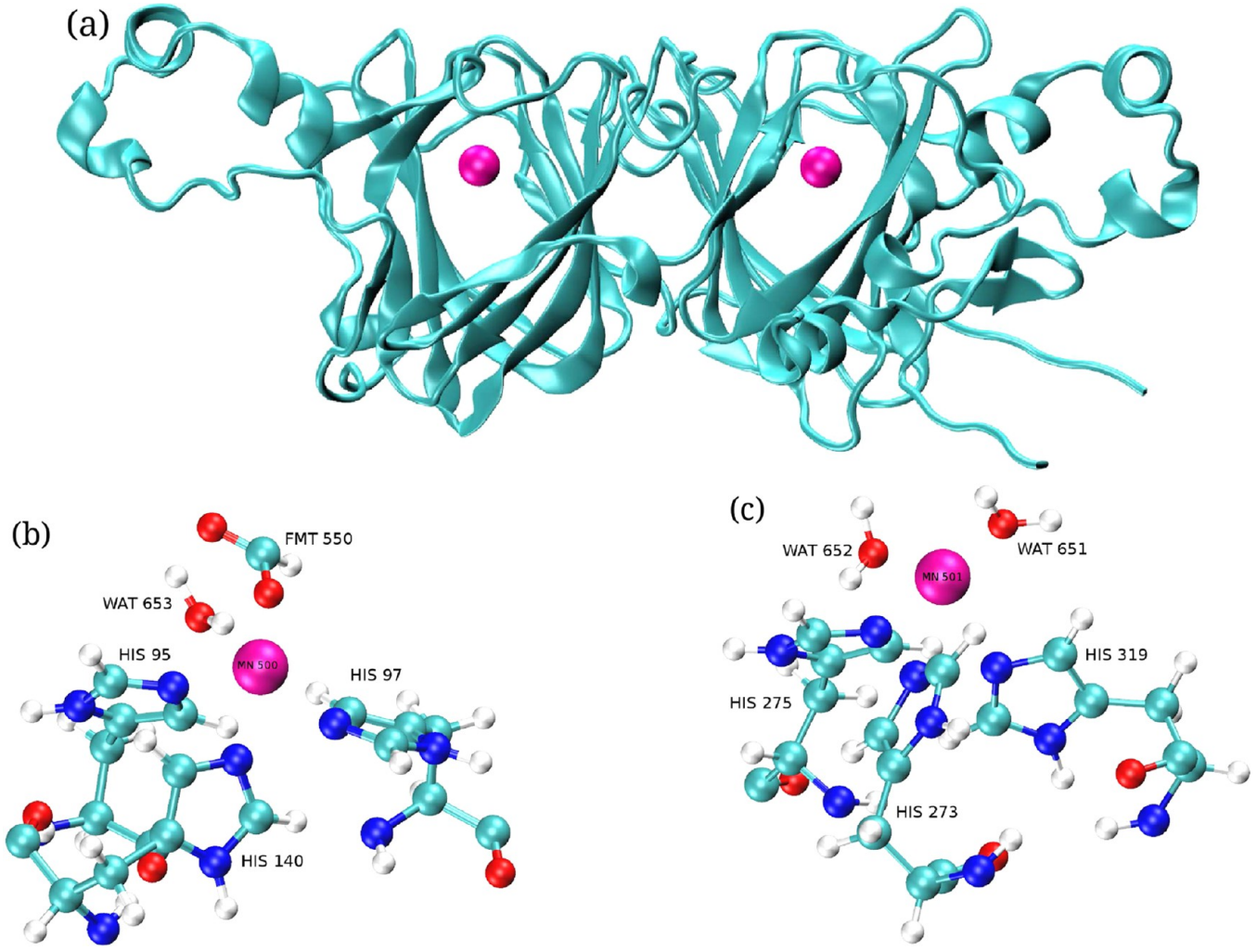


Figure 1. (a) OXDC protein [Protein Data Bank (PDB) entry 1L3J]. (b and c) Coordination environments of Mn in (b) domain I and (c) domain II. Water molecules of 1L3J are renamed as 891 to 651 and 895 to 652; WAT 653 is absent in domain I of 1L3J and thus added here on the basis of environments observed in this family of enzymes such as PDB entries 1J58, 2UY8, etc.).

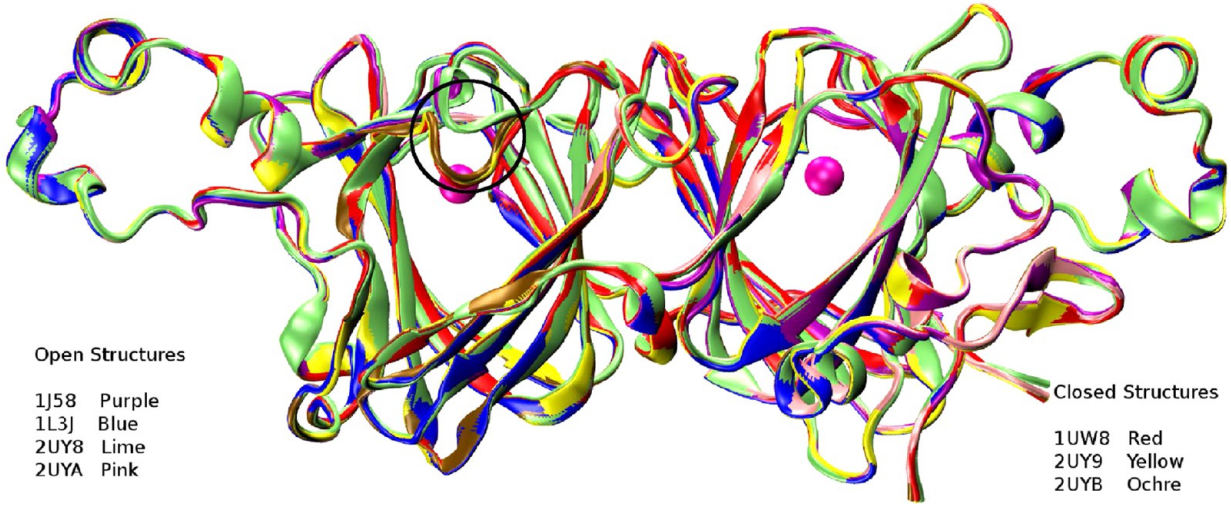


Figure 2. Overlay of available PDB structures of OXDC protein. Residues 161−165 are circled. The magenta sphere is a Mn^{2+} ion.

Crystallization of OXDC has been accomplished by several groups under various conditions.[15,19] The overlay of all the structures of OXDC (Figure 2) available from the Protein Data Bank (PDB) reveals many structural similarities between them, yet both domains can easily be differentiated by a careful examination of secondary structures of specific amino acid

dx.doi.org/10.1021/jp4074834 | J. Phys. Chem. B XXXX, XXX, XXX−XXX

sequences. The major difference is in the backbone of domain I where the flexible loop (residues 161−165) is open[19,20] in a few proteins (PDB entries 1L3J, 1J58, 1UY8, and 2UYA) and closed in the remaining structures (PDB entries 1UW8, 2UYB, and 2UY9). On the basis of the structural similarities between the two domains, we assume a small loop (residues 338−342) in domain II to be analogous to the flexible loop (residues 161−165) in domain I. The *B* factors of the former are found to be smaller (by a factor of 3) than the values for residues 161−165 present in domain I, suggesting that the loop in domain II is not flexible and is always closed.

Despite many studies of OXDC enzymes, the exact location of the decarboxylation process is a matter of debate; two schools of thought on this issue exist. Anand et al.,[15] for the first time, synthesized this enzyme with a formate anion located inside domain I. They proposed the existence of a narrow water channel from this site to the exterior of the protein. Further, they proposed that domain II be the catalytic site based on mutation studies. They concluded that the presence of Glu 333 as a proximate proton donor allows domain II to be the active site; however, such a proton donor is apparently absent in domain I. On the other hand, Just et al.,[19] who synthesized the same enzyme in both its open and closed forms, do not subscribe to the viewpoint of domain II being the active site. They suggested that in the closed form of the protein, the Glu 162 residue could come closer to domain I and can act as a proton donor to mediate the decarboxylation reaction[19,21] in that site. It has also been hypothesized that the water molecule present near the domain I Mn center may also act as proton donor,[22] but this proposal is yet to be established strongly from an experimental or theoretical viewpoint. CO_2 and formate can be released by the retreat of the loop to its open form. Thus, the movement of the loop is intimately connected to the activity of the enzyme through domain I. In PDB entry 1L3J, the loop is already open and hence the CO_2 release can be studied. Furthermore, in the open structure, a solvent accessible channel spanning from the protein core (near domain I) to its exterior that has been proposed to be the substrate uptake and/or product release pathway is observed.[19] Despite the immense interest in this enzyme, atomistic simulations have not been conducted on OXDC. If we limit ourselves to molecular dynamics simulations using empirical potentials, modeling the actual chemical reaction[23−26] is ruled out. Our focus herein is twofold: (i) to examine if product release can be used as an additional marker for distinguishing the active site and (ii) to map out the microscopic mechanism of transport of carbon dioxide from the protein interior to its exterior. Anticipating our work, we identify domain I to be the catalytically active site. CO_2 is transported through the water channel lined by a series of amino acid residues that interact with the gas molecule via either weak Lewis acid−base or hydrogen bonding interactions.

■ COMPUTATIONAL DETAILS

The enzyme PDB structure (1L3J, *B. subtilis*) at a resolution of 1.75 Å was obtained from RCSB Protein Data Bank. Coordinates of a few amino acid residues in the two terminii (1−7 and 380−385) that were missing in the PDB structure were constructed and joined to the coordinates of 1L3J using the GaussView[27] molecular graphics program. The enzyme has been shown to be active under acidic conditions, at a pH value of ≤6.[28] To mimic this condition, all the metal-coordinated histidine residues were kept in a neutral state while the remaining histidine residues were protonated to the +1 state. The protonation was accomplished through the pdb2gmx routine embedded within the GROMACS software. Crystal water molecules [except those coordinating the Mn(II) and Mg(II) ions] were deleted from the coordinate file. The protein was placed in the center of a cubic box with an edge length of 123.43 Å and solvated by 60140 TIP3P water molecules. The system was neutralized by adding seven Na^+ ions. The total number of atoms in the system was 186499. The system was then energy minimized by the steepest descent method and then warmed from 0 to 300 K over 300 ps in the constant-*NVT* ensemble with a time step of 0.5 fs. This thermally equilibrated structure was utilized for the long production simulation run in the constant-*NPT* ensemble. During this *NPT* simulation, the temperature of the system was maintained at 300 K by using a Nosé-Hoover thermostat with a coupling constant of 1 ps. A Parrinello-Rahman barostat with a coupling constant of 1 ps was employed to keep the system at a pressure of 1 bar. Long-range electrostatic interactions were calculated with the particle mesh Ewald method. All the production runs were performed with a time step of 1 fs, and their lengths varied from 10 to 15 ns.

The run lengths employed here may appear short relative to the duration of simulations being reported currently in the literature. However, longer simulations were not deemed to be necessary in our case as the release of the CO_2 molecule occurred within this time scale. The sampling is enhanced by multiple simulations initiated from many initial configurations.

The LINCS algorithm[29] was utilized to keep all the covalent bonds involving hydrogen atoms (C−H, N−H, and O−H) rigid. A flexible model of CO_2, generated from CGENFF,[30] was employed. The CHARMM27 force field[31] was used for the protein, and parameters of the ligands were assigned from the CGENFF force field. In our study, we have used a nonbonded metal−ligand model for the metal coordination shell. Nonbonded parameters of metal ions were adopted from previously performed molecular dynamics simulation investigations.[32] All the simulations were conducted in GROMACS version 4.5.5.[33−37] Visual Molecular Dynamics (VMD)[38] was used to visualize trajectories and prepare figures.

Additionally, one simulation (A4) was conducted with the Amber94 force field[39] for the protein and EPM2[40] model for the CO_2 molecule. Results from this run are presented in the Supporting Information (Figures S1 and S2); these are qualitatively similar to those presented here.

In 1L3J, the formate ion is bound to Mn(II) in domain I. A 10 ns trajectory (labeled F1P in Table 1) was generated for this system to establish its stability within the protocols employed. Subsequently, one CO_2 molecule was placed in domain I of 1L3J along with the crystal formate, while the other domain contained two water molecules bound to Mn(II) (labeled Set A in Table 1). Three independent simulations of this kind were performed, by varying the seed used in the generation of random initial velocities of atoms. Each of them lasted for 10 ns, and they are labeled runs A1−A3 (see Table 1).

To examine the feasibility of the release of product from domain II, another set of simulations was conducted. Because the crystal structure did not contain the formate ion in domain II, a simulation for 5 ns was conducted with a formate and a water molecule located in domain II to test its stability. This was followed by simulations in which a CO_2 molecule was further added to domain II. Three independent simulations of this kind, each lasting for 15 ns and labeled B1−B3 (see Table 1), were performed. The root-mean-square deviations (rmsd's)

dx.doi.org/10.1021/jp4074834 | J. Phys. Chem. B XXXX, XXX, XXX−XXX

Table 1. Summary of Simulations; F1P: Formate in Domain-I of Protein, F2P: Formate in Domain-II of Protein, SET A: Domain-I with FMT and CO_2, SET B: Domain-II with FMT and CO_2

domain I	domain II
F1P	
1 FMT	2 H_2O
1 H_2O	
Set A (A1–A3)	
1 CO_2	
1 FMT	2 H_2O
1 H_2O	
F2P	
2 H_2O	1 FMT
	1 H_2O
Set B (B1–B3)	
	1 CO_2
2 H_2O	1 FMT
	1 H_2O

relative to the structure of the enzyme post-energy minimization (and prior to the initiation of the MD run) were calculated using the grms code present in GROMACS.

The electrostatic potential map of 1L3J was generated as follows. PDB2PQR[41,42] was used to generate the "pqr" file containing information about the charge of the protein. This file was then imported into PYMOL[43] to calculate the electrostatic potential, based on Poisson–Boltzmann equations, with the APBS (Adaptive Poisson-Boltzmann Solver) plugin.[44] The electrostatic potential surface was visualized in PYMOL.

RESULTS AND DISCUSSION

The rmsd of the protein backbone from the F1P run shown in Figure 3 demonstrates the stability of the 1L3J structure and validates the protocols adopted in this work.

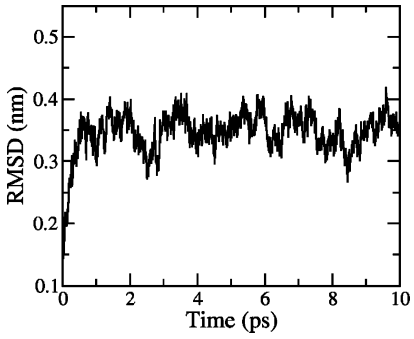


Figure 3. rmsd of unmodified protein 1L3J obtained from the F1P run.

Identifying the CO_2 Migration Pathway. Various initial configurations were generated to probe the active site by inserting both formate and CO_2 at different positions within domain I and domain II. We shall first discuss the configuration at domain I and its evolution during the simulation.

Domain I. As mentioned earlier, the formate ion is present in domain I of 1L3J. To test if this domain is the active site, we placed a CO_2 molecule near the Mn center of this domain and adjacent to the formate. The protein structure (including the carbon dioxide and formate locations) was then energy minimized.

The simulations show that the motion of CO_2 toward the protein exterior follows a well-defined cylindrical path (see Figure 4a) exhibiting the CO_2 migration path in one of the trajectories; results from other trajectories are presented in Figures S1 and S2 of the Supporting Information. The distance between Mn(II) (of domain I) and the carbon atom of the CO_2 molecule obtained from this run is plotted as a function of time in Figure 4b. CO_2 reaches the exterior of the protein in approximately 2–3 ns and stays there. The motion of CO_2 is not always directly to the exterior, and one can observe the molecule oscillating between the interior and the exterior in some trajectories. The backbone rmsd values in this run (Figure 4) are within acceptable limits, indicating the overall structural stability of the protein during the release of the CO_2 molecule from domain I. The observed cylindrical path of CO_2 motion is in good agreement with the shape of the water channel proposed by Just et al.,[19] based on crystallographic studies. The nature and type of interactions decorating the path will be discussed later. Results from two other independent trajectories are provided in the Supporting Information.

Domain II. An approach similar to that outlined above was employed to study the release of CO_2 from domain II, as well. In the experimentally determined structure, domain II lacked a formate ion, which, however, was present in domain I. Thus, if the proposition of domain II being the active site[15] needs to be tested, both formate and CO_2 have to be present there as reaction products. This was accomplished as follows. First, one water molecule (either water 651 or water 652) that was coordinated to Mn in domain II was replaced by a formate molecule. The resulting formate-bound structure was then simulated for 5 ns, post-energy minimization and post-warmup. In both the cases (replacement of either water molecule), the environment was disrupted after a few nanoseconds because of steric hindrance near the Mn center. In the first case (water 652 replaced with FMT), the formate ion was nearly bound to the metal ion through a divalent motif and the coordinated histidines moved away from the metal center. In the latter case (water 651 replaced with FMT), a few side chains (Tyr 340, Glu 333, and Arg 270) moved away from their original positions, thereby disrupting the crystallographic environment of this domain.

However, in the second case, the histidine side chains were at least coordinated with the metal. Thus, we decided to keep the formate ion in this location and orientation. Additionally, CO_2 was placed near the formate molecule in domain II, and the simulations were conducted for 5 ns.

We observed the movement of the carbon dioxide molecule inside site II during minimization and a short *NVT* simulation run. Post-insertion (after minimization and warmup), CO_2 moved toward the deep cavity inside this domain and rattled around the Mn(II) ion. Three simulations, in each of which the initial velocities of the CO_2 molecule were chosen randomly (with different random number seeds), were extended for an additional 10 ns each. In none of these was CO_2 seen to arrive at the exterior of the protein (see Figure 5). Further, the insertion of either formate alone or formate and CO_2 into domain II causes the rmsd (see Figure 5c) to increase without bounds. Results from other equivalent trajectories, presented in Figures S5 and S6 of the Supporting Information, are similar. Mutation studies with Glu 333 present in domain II led to a reduction in the activity of the enzyme. Anand et al.[15] thus suggested domain II to be the active site. However, as described by

dx.doi.org/10.1021/jp4074834 | J. Phys. Chem. B XXXX, XXX, XXX–XXX

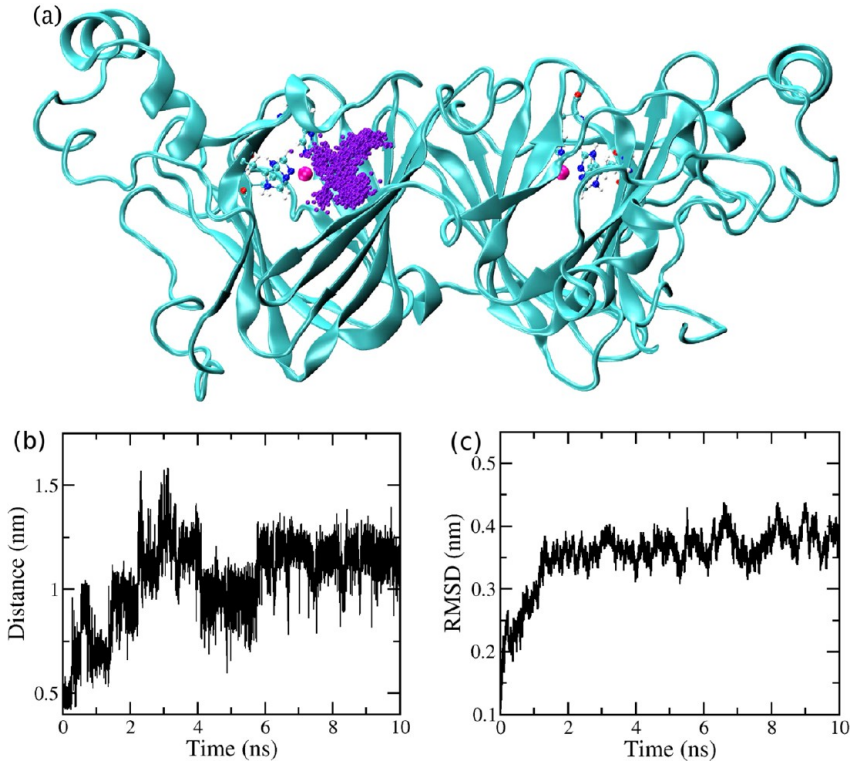


Figure 4. (a) Migration of CO_2 from domain I in run A1: magenta for Mn and violet for CO_2 positions over time. (b) Mn−CO_2 distance vs time, demonstrating the release of the gas molecule. (c) rmsd of the protein backbone in run A1.

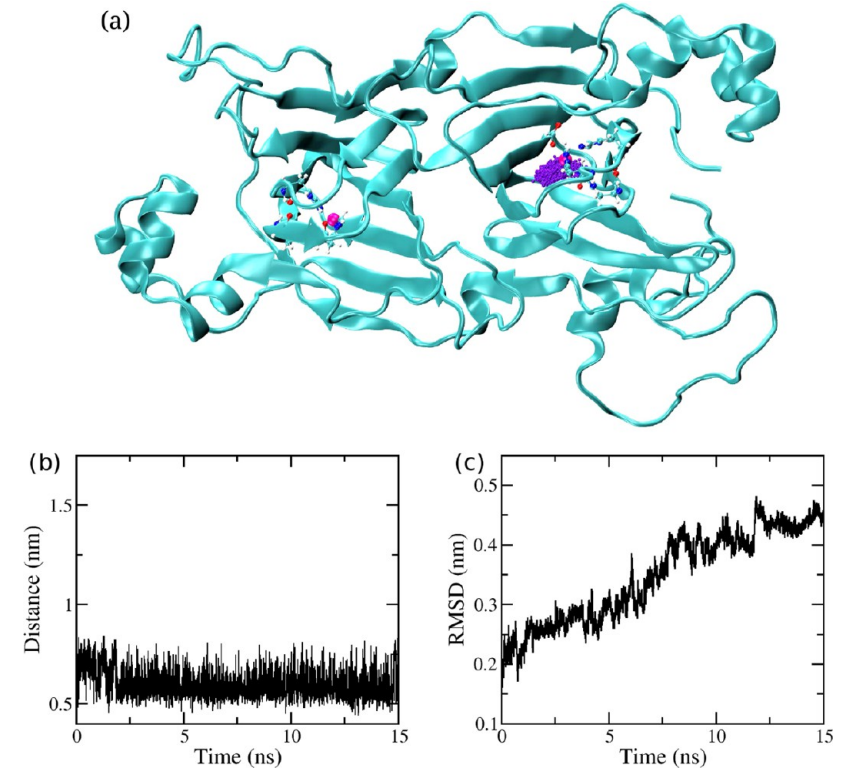


Figure 5. (a) Movement of CO_2 inside domain II. (b) Mn−CO_2 distance vs time, indicating the trapped rattling motion of CO_2. (c) rmsd of the protein backbone in run B1.

Just et al.,[19] the motion of residues in the two domains is coupled through common residues such as Trp 96 and Trp 274 and the mutation of Glu 333 could have led to changes in the protein's structure to reduce its activity. Our observations of increased rmsd values upon insertion of CO_2 and formate into domain II suggest the same.

dx.doi.org/10.1021/jp4074834 | J. Phys. Chem. B XXXX, XXX, XXX−XXX

This outcome reinforces the absence of a suitable "path" through which a CO_2 molecule could come out from the interior of domain II. Our observations clearly point to domain II being nonreceptive to hosting the formate, as well as being incapable of releasing the CO_2 molecule (Figure 6). These findings strongly suggest that it is unlikely to be the active site. In contrast, the binding of formate to Mn in domain I through a monovalent motif did not cause any disruption of the structure of the protein. Domain I was able to host the reaction products comfortably and release the CO_2, as well, indicating the strong likelihood that it is the active site for the decarboxylation of oxalate.

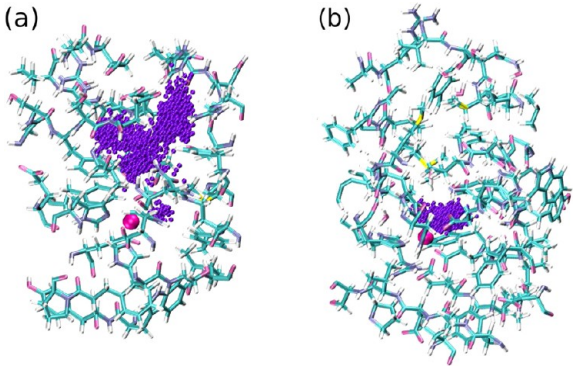


Figure 6. (a) Migration of CO_2 through the channel from domain I. (b) No path for the expulsion of CO_2 from domain II.

It is necessary to understand, from a molecular perspective, the nature of the transport of the CO_2 molecule from the active domain to the exterior of the protein. The specific interactions between CO_2 and various amino acids were analyzed and are discussed below.

Nature of the Channel and the Amino Acids Lining Its Interior. As depicted in Figure 6a, CO_2 follows a well-defined path for its migration from the active site domain (domain I) toward the surface. The stages of migration of CO_2 throughout the channel can be assessed by capturing snapshots within the simulation trajectory.

Hydrophobic Core. The residence of CO_2 in hydrophobic pockets has been described in several enzymes, one of which is human carbonic anhydrase.[45] In OXDC, just beside the coordination shell of the metal center, there exists a "pocket" (Figure 7) populated with predominantly hydrophobic residues (Val, Ile, Leu, Phe, and Met). During the minimization and equilibration runs, the gas molecule temporarily transits to this hydrophobic region, irrespective of its initial location. Arg 92 is also present near this region, which provides additional stabilization to CO_2.

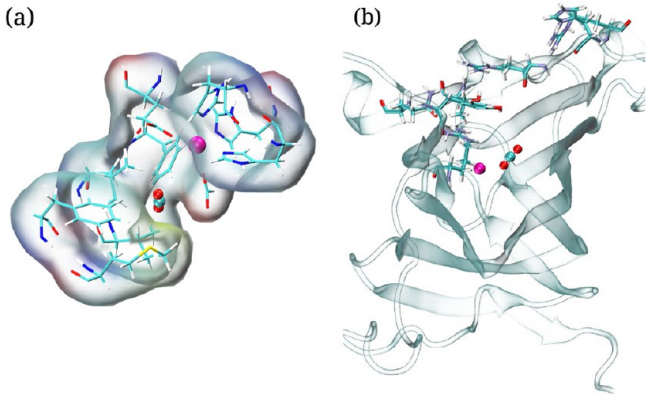


Figure 7. CO_2 in domain I: (a) initial hydrophobic region around CO_2 and (b) "CO_2-philic" amino acids lining the channel.

Hydrophilic and Hydrophobic Halves of the Channel. The channel is mostly cylindrical in shape and consists of two chemically distinct regions lined with two different types of amino acids. One side of the channel is largely populated with hydrophobic side chains such as Val, Leu, Ile, Phe, etc., while the other is decorated with hydrophilic polar amino acids such as Arg, Tyr, His, Glu, Gln, Thr, etc.

Polar Amino Acids Mediating CO_2 Transport. Although CO_2 is located initially inside the hydrophobic core, its movement away from the metal center occurs because of the surrounding highly negatively charged Glu residue and also the negatively charged product molecule, FMT. A polar amino acid group, Arg 92 has been proposed to be the residue that binds the substrate.[19] It further facilitates the transition of the gas molecule toward the middle of the channel. The amino acids in domain I that predominantly interact with CO_2 were identified by the following procedure. At first, the distances between the C of CO_2 and all the C_α atoms were calculated for each frame of a trajectory, and then the minimal C_α–C distance in each frame was selected to prompt the amino acids to interact with the CO_2. These are listed in Table 2.

With an increasing distance from the metal site, Phe, Gln, Thr, and Tyr are the residues located just after Arg 92. Interestingly, His 95, Arg 92, Tyr 200, Arg 58, Phe 155, and Phe 160 that line the channel in domain I exhibit $\pi-\pi$ stacking. This type of stacking has been well studied in the literature.[46−49] These key residues assist the migration of the gas molecule from the active site. To understand their specific role in CO_2 migration, we have examined the motifs of interaction of CO_2 with the side chains of these polar amino acid residues through gas phase calculations. The patterns of interaction with all these active amino acid side chains are depicted in Figure 8. During its transport, CO_2 interacts with the polar amino acids either through hydrogen bonds or by the formation of Lewis acid–base (Lab) pairs.[50−52] Both these interaction motifs are represented in Figure 8. Interestingly, Phe 155 and Phe 160 also form extremely weak hydrogen bonds through their ring hydrogens with the oxygen termini of CO_2. The two terminal oxygen atoms of the CO_2 molecule have partial negative charges that interact with the side chains of polar amino acid residues via hydrogen bonds. On the other hand, as the central carbon atom of CO_2 has a partial positive charge, it acts as a Lewis acid center. This Lewis acidic carbon center forms complexes with amino acid side chain atoms containing lone pairs of electrons. The patterns of interaction of CO_2 with these amino acid residues are in good agreement with earlier theoretical studies performed by several groups.[50,51]

Amino Acid–CO_2 Interaction. We have calculated the gas phase binding energy of CO_2–amino acid pairs, based either on quantum chemical or on force fields (Table 3). In the QM calculations, performed using Gaussian-09[53] at MP2/aug-cc-pVTZ and M06−2x/aug-cc-pVTZ levels of theory, the truncated side chains (up to the C_α atom) were considered, while in the classical calculations, the entire (including -NH_3 and -COO attached to the C_α atom) amino acid was used. Interactions of CO_2 with Tyr, Arg, and Phe side chains have been well studied in previous investigations.[50] Additionally,

dx.doi.org/10.1021/jp4074834 | J. Phys. Chem. B XXXX, XXX, XXX−XXX

Table 2. Amino Acid Residues That Line the Water Channel in Domain I

	Val	Ile	Leu	Ala	Asn	Phe	Ser	Tyr	Arg	Thr	Glu	Gln	His
position	82	84	153	65	57	155	81	200	58	165	67	167	56
						160	161		66		101		
									71		162		
									92				

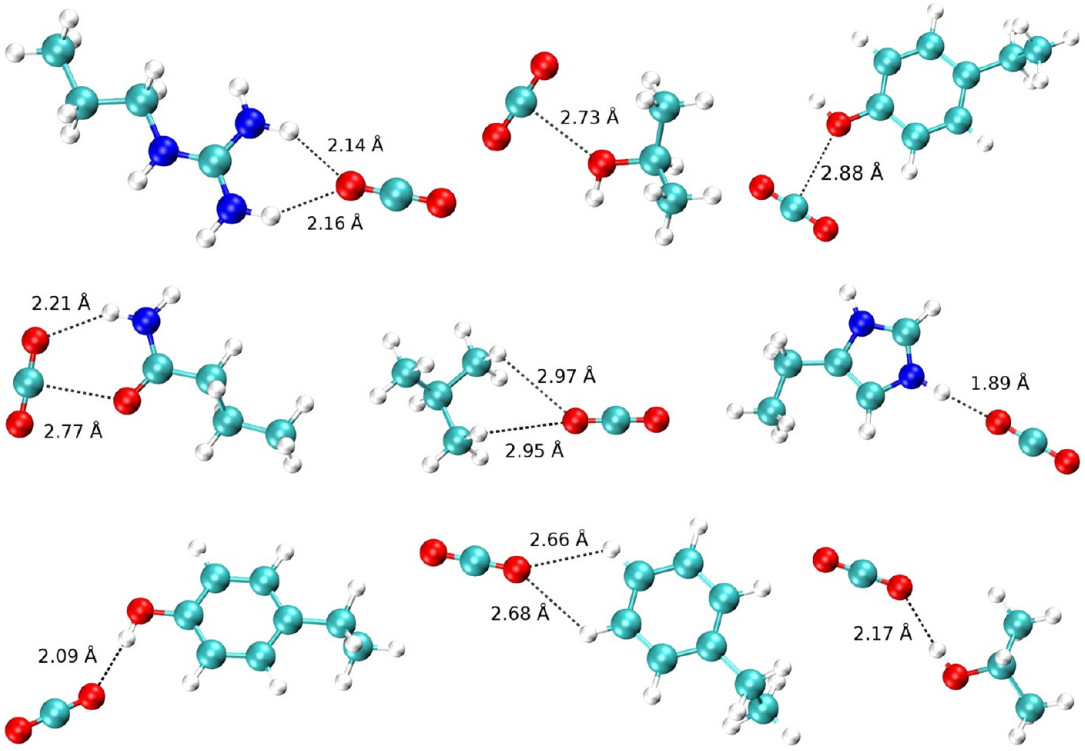


Figure 8. Optimized structures obtained using the MP2/aug-cc-pVTZ level of theory for CO_2–amino acid pairs in the gas phase: Arg, Thr-Lab, Tyr-Lab, Gln, Val, His, Tyr-Hyd, Phe, and Thr-Hyd (from Top to bottom and left to right, respectively). Here, Lab and Hyd mean Lewis acid–base and hydrogen bonding interactions, respectively. Tyr-Lab alone was optimized at the B3LYP/6-311+G(d,p) level of theory. Similar interaction motifs observed during the MD simulations of the CO_2–1L3J pair are presented in Figure S3 of the Supporting Information.

Table 3. Binding Energies for Amino Acid Side Chains (up to the C_α atom) and CO_2 Obtained Using MP2/aug-cc-pVTZ (BSSE corrected) and M06-2x/aug-cc-pVTZ Levels of Theory, Compared with That from the CHARMM-CGENFF Force Field

amino acid	MP2/aug-cc-pVTZ, BSSE (kcal/mol)	M06-2x/aug-cc-pVTZ (kcal/mol)	CHARMM-27, CGENFF (kcal/mol)
Arg	−6.51	−6.68	−4.40
His	−6.13	−6.08	−3.45
Gln	−4.36	−5.36	−3.03
Thr-Lab	−3.34	−4.14	−3.67
Tyr-Lab	−2.46[a]	−2.69[b]	−2.47
Tyr-Hyd	−2.36	−2.39	−2.88
Thr-Hyd	−1.80	−1.14[c]	−2.87
Phe	−0.94	−0.70	−0.80
Val	−0.65	−1.86	−1.00

[a]MP2/aug-cc-pVTZ//B3LYP/6-311+G(d,p) level of theory. [b]M06-2x/aug-cc-pVTZ//B3LYP/6-311+G(d,p) level of theory. [c]M06-2x/aug-cc-pVTZ//MP2/aug-cc-pVTZ level of theory.

CO_2 shows specific interactions with two more polar residues, glutamine (Gln) and threonine (Thr). It forms an adduct with Gln through both hydrogen bond and Lab interaction patterns, while with Thr, it can interact either through hydrogen bonding or by forming a Lewis acid–base pair (Figure 8).

Arg, His, and Glu residues show more affinity for CO_2 binding than other residues do. Tyr and Thr interact with CO_2 either through H-bonding or through Lewis acid–base interaction. Phe and Val also form adducts with CO_2 by the help of weak interactions. On the basis of the nature of interactions of these (Arg, His, and Tyr) amino acids with CO_2, these amino acids are called "CO_2-philic" residues.[54] On the basis of mutation studies, Arg 92 was earlier proposed[19] to bind the substrate (oxalate); our results suggest that it can also bind the product CO_2. A formate molecule was found to be bound to Arg 92 through strong hydrogen bonding interaction (see Figure S4 of the Supporting Information). CO_2 also interacts with this Arg residue initially before coming to the middle of the domain. Thus, this region can be proposed to be the "resting zone" of the reactant as well as product molecules.

Protein–CO_2 Interaction Energy. The potential energy of the CO_2 molecule with the protein was analyzed as a function of time for run A2, and the same is displayed in Figure 9. The distance from the gas molecule to the Mn reaction center in this trajectory is also displayed.

A clear correlation between the interaction energy and the location of the gas molecule is observed in Figure 9. The energy

dx.doi.org/10.1021/jp4074834 | J. Phys. Chem. B XXXX, XXX, XXX–XXX

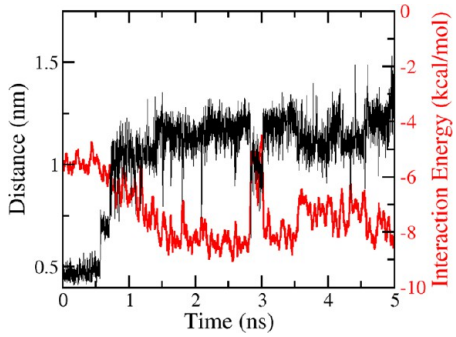


Figure 9. Protein−CO_2 interaction energy (red, axis on the right) and distance (black, axis on the left) from CO_2 to the metal site in domain I, as a function of time.

of interaction of CO_2 with the protein decreases as it migrates from the core of domain I to the external surface of the protein. The protein−CO_2 interaction has been previously studied by Drummond et al.,[54] who calculated the energy of interaction of CO_2 with the PEPCK protein. The values (6−9 kcal/mol) obtained by us fall within the range of values (0−16 kcal/mol) they observed.

The migration of the product gas molecule can also be rationalized on the basis of the electrostatic potential map of the protein, specifically focusing on the active domain. The same is displayed in Figure S7 of the Supporting Information. A charge gradient is observed between the inside core and the exterior surface of the protein. The presence of a negatively charged amino acid residue (Glu) and FMT in the core of domain I repels the CO_2 molecule and moves it initially to the hydrophobic region, and later toward the middle of the channel. Arg, Tyr, and Phe are so situated inside the channel that the transport is facilitated throughout the channel. Near the surface of this domain, there are positively charged regions that are very suitable for CO_2 binding. Here, CO_2 interacts with positively charged amino acids like Arg 71 and His 56 through the terminal oxygen atoms. Thus, the positioning of these residues together allows the migration of CO_2 to the exterior of the protein.

CONCLUSIONS

We have conducted empirical potential-based MD simulations of an oxalate decarboxylase enzyme, OXDC, with the primary purpose of identifying migration pathways of the CO_2 molecule, one of the reaction products. The simulations were performed from many initial conditions for CO_2, in each of the domains of this enzyme.

The release of the decarboxylation product, CO_2, from domain I is facile, while it is not so for release from domain II. This easy release of CO_2 aids us in identifying domain I as the likely active site. In domain II, the CO_2 molecule is trapped inside and does not exit the protein interior during the duration of the MD trajectory. The amino acids lining domain I throughout the CO_2 migration path can be variously categorized: the ones in the core are mostly hydrophobic, while the channel is decorated with polar, CO_2-philic residues. CO_2 does show specific interactions with Arg, Tyr, Gln, Thr, and Phe amino acids. Two kinds of interaction motifs have been observed for CO_2 with these amino acids, hydrogen bonding and Lewis acid−base interaction. Because the interior core of domain I is negatively charged (primarily arising from the presence of the formate ion), a gradient of potential from negative to positive from the core to the outer surface of the protein exists. CO_2 prefers to migrate out toward the positively charged region, which is situated almost at the surface of the protein.

Our simulations have also clearly demonstrated the likelihood of domain I being the active site of this enzyme, which is in agreement with the earlier proposal of Just et al.[19] The molecular perspective of the CO_2 transport channel is captured well via these simulations. This path of CO_2 migration may be further hypothesized to be the probable path of inclusion of substrate in the active site domain I in the catalytic cycle. Studies of transport of substrate and product gas molecules in similar enzymes are in progress in our laboratory.

ASSOCIATED CONTENT

Supporting Information

CHARMM general force field parameters for CO_2 implemented in GROMACS (Table S1), Mn−CO_2 distance versus time plots for runs A1−A4 (Figure S1), backbone rmsd of runs A1−A4 and F1P (Figure S2), interaction motifs of amino acid side chains with CO_2 as observed in a simulation trajectory (Figure S3), binding of formate by Arg 92 in domain I (Figure S4), Mn−CO_2 distance versus time plots for runs B1−B3. (Figure S5), backbone rmsd of runs B1−B3 (Figure S6), and electrostatic potential map of OXDC (Figure S7). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: bala@jncasr.ac.in.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge the Department of Science and Technology, India, for support.

REFERENCES

(1) Mäkelä, M.; Hildén, K.; Lundell, T. Oxalate decarboxylase: Biotechnological update and prevalence of the enzyme in filamentous fungi. *Appl. Microbiol. Biotechnol.* **2010**, *87*, 801−814.

(2) Baetz, A. L.; Allison, M. J. Purification and characterization of oxalyl-coenzyme A decarboxylase from *Oxalobacter formigenes*. *J. Bacteriol.* **1989**, *171*, 2605−2608.

(3) Liang, J. Y.; Lipscomb, W. N. Binding of substrate CO_2 to the active site of human carbonic anhydrase II: A molecular dynamics study. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3675−3679.

(4) Wolfenden, R.; Lewis, C. A.; Yuan, Y. Kinetic Challenges Facing Oxalate, Malonate, Acetoacetate, and Oxaloacetate Decarboxylases. *J. Am. Chem. Soc.* **2011**, *133*, 5683−5685.

(5) Chiarla, C.; Giovannini, I.; Giuliante, F.; Vellone, M.; Ardito, F.; Tenhunen, J.; Nuzzo, G. Significance of hemoglobin concentration in determining blood CO_2 binding capacity in critical illness. *Respir. Physiol. Neurobiol.* **2010**, *172*, 32−36.

(6) Endeward, V.; Cartron, J.-P.; Ripoche, P.; Gros, G. RhAG protein of the Rhesus complex is a CO_2 channel in the human red cell membrane. *FASEB J.* **2008**, *22*, 64−73.

(7) Soupene, E.; King, N.; Feild, E.; Liu, P.; Niyogi, K. K.; Huang, C.-H.; Kustu, S. Rhesus expression in a green alga is regulated by CO_2. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7769−7773.

(8) Soupene, E.; Inwood, W.; Kustu, S. Lack of the Rhesus protein Rh1 impairs growth of the green alga *Chlamydomonas reinhardtii* at high CO_2. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7787−7792.

(9) Drummond, M. L.; Wilson, A. K.; Cundari, T. R. Carbon Dioxide Migration Pathways in Proteins. *J. Phys. Chem. Lett.* **2012**, *3*, 830−833.

(10) Kustu, S.; Inwood, W. Biological gas channels for NH_3 and CO_2: Evidence that Rh (Rhesus) proteins are CO_2 channels. *Transfusion clinique et biologique: journal de la Société française de transfusion sanguine* **2006**, *13*, 103−110.
(11) Emiliani, E.; Bekes, P. Enzymatic oxalate decarboxylation in *Aspergillus niger*. *Arch. Biochem. Biophys.* **1964**, *105*, 488−493.
(12) Micales, J. A. Localization and induction of oxalate decarboxylase in the brown-rot wood decay fungus *Postia placenta*. *Int. Biodeterior. Biodegrad.* **1997**, *39*, 125−132.
(13) Chakraborty, S.; Chakraborty, N.; Jain, D.; Salunke, D. M.; Datta, A. Active site geometry of oxalate decarboxylase from *Flammulina velutipes*: Role of histidine-coordinated manganese in substrate recognition. *Protein Sci.* **2002**, *11*, 2138−2147.
(14) Mäkelä, M. R.; Hildén, K.; Hatakka, A.; Lundell, T. K. Oxalate decarboxylase of the white-rot fungus *Dichomitus squalens* demonstrates a novel enzyme primary structure and non-induced expression on wood and in liquid cultures. *Microbiology* **2009**, *155*, 2726−2738.
(15) Anand, R.; Dorrestein, P. C.; Kinsland, C.; Begley, T. P.; Ealick, S. E. Structure of Oxalate Decarboxylase from *Bacillus subtilis* at 1.75 Å Resolution. *Biochemistry* **2002**, *41*, 7659−7669.
(16) Moomaw, E. W.; Angerhofer, A.; Moussatche, P.; Ozarowski, A.; García-Rubio, I.; Richards, N. G. J. Metal Dependence of Oxalate Decarboxylase Activity. *Biochemistry* **2009**, *48*, 6116−6125.
(17) Angerhofer, A.; Moomaw, E. W.; García-Rubio, I.; Ozarowski, A.; Krzystek, J.; Weber, R. T.; Richards, N. G. J.; Multifrequency, E. P. R. Studies on the Mn(II) Centers of Oxalate Decarboxylase. *J. Phys. Chem. B* **2007**, *111*, 5043−5046.
(18) Chang, C. H.; Svedružić, D.; Ozarowski, A.; Walker, L.; Yeagle, G.; Britt, R. D.; Angerhofer, A.; Richards, N. G. J. EPR Spectroscopic Characterization of the Manganese Center and a Free Radical in the Oxalate Decarboxylase Reaction: Identification of a tyrosyl radical during turnover. *J. Biol. Chem.* **2004**, *279*, 52840−52849.
(19) Just, V. J.; Stevenson, C. E. M.; Bowater, L.; Tanner, A.; Lawson, D. M.; Bornemann, S. A Closed Conformation of *Bacillus subtilis* Oxalate Decarboxylase OxdC Provides Evidence for the True Identity of the Active Site. *J. Biol. Chem.* **2004**, *279*, 19867−19874.
(20) Just, V. J.; Burrell, M. R.; Bowater, L.; McRobbie, I.; Stevenson, C. E. M.; Lawson, D. M.; Bornemann, S. The identity of the active site of oxalate decarboxylase and the importance of the stability of active-site lid conformations. *Biochem. J.* **2007**, *407*, 397−406.
(21) Burrell, M. R.; Just, V. J.; Bowater, L.; Fairhurst, S. A.; Requena, L.; Lawson, D. M.; Bornemann, S. Oxalate Decarboxylase and Oxalate Oxidase Activities Can Be Interchanged with a Specificity Switch of up to 282000 by Mutating an Active Site Lid. *Biochemistry* **2007**, *46*, 12327−12336.
(22) Muthusamy, M.; Burrell, M. R.; Thorneley, R. N. F.; Bornemann, S. Real-Time Monitoring of the Oxalate Decarboxylase Reaction and Probing Hydron Exchange in the Product, Formate, Using Fourier Transform Infrared Spectroscopy. *Biochemistry* **2006**, *45*, 10667−10673.
(23) Saylor, B. T.; Reinhardt, L. A.; Lu, Z.; Shukla, M. S.; Nguyen, L.; Cleland, W. W.; Angerhofer, A.; Allen, K. N.; Richards, N. G. J. A Structural Element That Facilitates Proton-Coupled Electron Transfer in Oxalate Decarboxylase. *Biochemistry* **2012**, *51*, 2911−2920.
(24) Reinhardt, L. A.; Svedružić, D.; Chang, C. H.; Cleland, W. W.; Richards, N. G. J. Heavy Atom Isotope Effects on the Reaction Catalyzed by the Oxalate Decarboxylase from Bacillus subtilis. *J. Am. Chem. Soc.* **2003**, *125*, 1244−1252.
(25) Li, T.; Huo, L.; Pulley, C.; Liu, A. Decarboxylation mechanisms in biological system. *Bioorg. Chem.* **2012**, *43*, 2−14.
(26) Imaram, W.; Saylor, B. T.; Centonze, C. P.; Richards, N. G.; Angerhofer, A. EPR spin trapping of an oxalate-derived free radical in the oxalate decarboxylase reaction. *Free Radical Biol. Med.* **2011**, *50*, 1009−1015.
(27) Dennington, R.; Keith, T.; Millam, J. *GaussView*, version 5; Semichem Inc.: Shawnee Mission, KS, 2009.
(28) Tanner, A.; Bowater, L.; Fairhurst, S. A.; Bornemann, S. Oxalate Decarboxylase Requires Manganese and Dioxygen for Activity: Overexpression and characterization of *Bacillus subtilis* YvrK and YoaN. *J. Biol. Chem.* **2001**, *276*, 43627−43634.
(29) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(30) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. 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.
(31) Bjelkmar, P.; Larsson, P.; Cuendet, M. A.; Hess, B.; Lindahl, E. Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. *J. Chem. Theory Comput.* **2010**, *6*, 459−466.
(32) Ghitti, M.; Spitaleri, A.; Valentinis, B.; Mari, S.; Asperti, C.; Traversari, C.; Rizzardi, G. P.; Musco, G. Molecular Dynamics Reveal that isoDGR-Containing Cyclopeptides Are True $\alpha v\beta 3$ Antagonists Unable To Promote Integrin Allostery and Activation. *Angew. Chem., Int. Ed.* **2012**, *51*, 7702−7705.
(33) Berendsen, H.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(34) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(35) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(36) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845−854.
(37) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, *7*, 306−317.
(38) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(39) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(40) Harris, J. G.; Yung, K. H. Carbon Dioxide's Liquid-Vapor Coexistence Curve And Critical Properties as Predicted by a Simple Molecular Model. *J. Phys. Chem.* **1995**, *99*, 12021−12024.
(41) Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; Klebe, G.; Baker, N. A. PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **2007**, *35*, W522−W525.
(42) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **2004**, *32*, W665−W667.
(43) *The PyMOL Molecular Graphics System*, version 1.3r1; Schrödinger, LLC: Portland, OR, 2010.
(44) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037−10041.
(45) Domsic, J. F.; Avvaru, B. S.; Kim, C. U.; Gruner, S. M.; Agbandje-McKenna, M.; Silverman, D. N.; McKenna, R. Entrapment of Carbon Dioxide in the Active Site of Carbonic Anhydrase II. *J. Biol. Chem.* **2008**, *283*, 30766−30771.
(46) Gervasio, F. L.; Chelli, R.; Procacci, P.; Schettino, V. The nature of intermolecular interactions between aromatic amino acid residues. *Proteins: Struct., Funct., Bioinf.* **2002**, *48*, 117−125.
(47) Zhang, Z.; Xu, Z.; Yang, Z.; Liu, Y.; Wang, J.; Shao, Q.; Li, S.; Lu, Y.; Zhu, W. The Stabilization Effect of Dielectric Constant and

645 Acidic Amino Acids on Arginine-Arginine (Arg-Arg) Pairings:
646 Database Survey and Computational Studies. *J. Phys. Chem. B* **2013**,
647 *117*, 4827−4835.
648 (48) Lee, D.; Lee, J.; Seok, C. What stabilizes close arginine pairing in
649 proteins? *Phys. Chem. Chem. Phys.* **2013**, *15*, 5844−5853.
650 (49) Flocco, M. M.; Mowbray, S. L. Planar Stacking Interactions of
651 Arginine and Aromatic Side-Chains in Proteins. *J. Mol. Biol.* **1994**, *235*,
652 709−717.
653 (50) Cundari, T. R.; Wilson, A. K.; Drummond, M. L.; Gonzalez, H.
654 E.; Jorgensen, K. R.; Payne, S.; Braunfeld, J.; De Jesus, M.; Johnson, V.
655 M. CO_2-Formatics: How Do Proteins Bind Carbon Dioxide? *J. Chem.*
656 *Inf. Model.* **2009**, *49*, 2111−2115.
657 (51) Hussain, M. A.; Soujanya, Y.; Sastry, G. N. Evaluating the
658 Efficacy of Amino Acids as CO_2 Capturing Agents: A First Principles
659 Investigation. *Environ. Sci. Technol.* **2011**, *45*, 8582−8588.
660 (52) Saharay, M.; Balasubramanian, S. Electron Donor-Acceptor
661 Interactions in Ethanol-CO_2 Mixtures: An Ab Initio Molecular
662 Dynamics Study of Supercritical Carbon Dioxide. *J. Phys. Chem. B*
663 **2006**, *110*, 3782−3790.
664 (53) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B. *GAUSSIAN 09*,
665 revision A.1; Gaussian Inc.: Wallingford, CT, 2009.
666 (54) Drummond, M. L.; Wilson, A. K.; Cundari, T. R. Nature of
667 Protein-CO_2 Interactions as Elucidated via Molecular Dynamics. *J.*
668 *Phys. Chem. B* **2012**, *116*, 11578−11593.