

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/26818861>

Locally Enhanced Sampling Study of Dioxxygen Diffusion Pathways in Homoprotocatechuate 2,3-Dioxygenase

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · SEPTEMBER 2009

Impact Factor: 3.3 · DOI: 10.1021/jp902597t · Source: PubMed

CITATIONS

3

READS

38

4 AUTHORS:



Liang Xu

Dalian University of Technology

25 PUBLICATIONS 129 CITATIONS

[SEE PROFILE](#)



Xin Liu

Dalian University of Technology

62 PUBLICATIONS 716 CITATIONS

[SEE PROFILE](#)



Weijie Zhao

Dalian University of Technology

26 PUBLICATIONS 45 CITATIONS

[SEE PROFILE](#)



Xicheng Wang

Dalian University of Technology

58 PUBLICATIONS 760 CITATIONS

[SEE PROFILE](#)

Locally Enhanced Sampling Study of Dioxygen Diffusion Pathways in Homoprotocatechuate 2,3-Dioxygenase

Liang Xu,^{†,‡} Xin Liu,[‡] Weijie Zhao,[§] and Xicheng Wang*,[†]

Department of Engineering Mechanics, State Key Laboratory of Structural Analyses for Industrial Equipment, and Department of Chemistry, Dalian University of Technology, Dalian 116023, China, and School of Chemical Engineering, State Key Laboratory of Fine Chemicals, Dalian University of Technology, Dalian 116012, China

Received: March 23, 2009; Revised Manuscript Received: July 31, 2009

Extradiol dioxygenases are characterized by activating dioxygen and incorporating both oxygen atoms into their substrates. Both experimental and theoretical investigations have been focused on the detection of intermediates in the reaction cycle in order to develop a general chemical mechanism of O₂ activation and insertion. However, little is known about the mechanism of how O₂ reaches the reaction sites of the related enzymes, which raises the question whether the rate of catalysis is limited by O₂ access to the active site. In this paper, two locally enhanced sampling molecular dynamics simulations were performed to determine the potential O₂ pathways inside a recently solved X-ray structure of homoprotocatechuate 2,3-dioxygenase. It is found that nominally identical subunits of the single homotetrameric structure contain distinct O₂ affinity diffusion pathways, which partly correlates with the observation of the simultaneous presence of three different reaction intermediates in four independent active sites. Residues that are critical for O₂ diffusion are also examined and discussed. In particular, we find that the breathing motion of the internal cavity defined by these residues results in O₂ migration process.

1. Introduction

Dioxygenases have long been of interest to chemists and biochemists due to their importance in the biodegradation of aromatic molecules in the environment. Both intradiol and extradiol dioxygenases are non-heme iron-containing enzymes that play an important role in the biodegradation of catechol and its derivatives by catalyzing the cleavage of aromatic rings. Generally, the intradiol dioxygenases require Fe³⁺ to cleave the C–C bond between the phenolic hydroxy groups to produce *cis,cis*-muconic acid, while the extradiol dioxygenases use Fe²⁺ (or Mn²⁺) as a cofactor to cleave the C–C bond adjacent to the phenolic hydroxy groups to yield 2-hydroxymuconaldehyde.^{1,2} As a result, an O₂ is activated and both oxygen atoms are inserted into the products. Pertinent recent reviews on the catalytic mechanisms of dioxygenases include those focusing on geometric and electronic structures,^{3,4} biomimetic modeling and intermediates,^{5,6} and reaction mechanism.^{7–9} The mechanism of the non-heme iron catechol dioxygenase has also been theoretically studied with density functional theory^{2,10,11} and has continued to be an active topic of research. The experimental determination of the structure of Fe²⁺-containing homoprotocatechuate 2,3-dioxygenase (2,3-HPCD) with X-ray crystallography showed that three different intermediates reside in different subunits of a single homotetrameric enzyme molecule.¹² This observation indicates that the four nominally identical subunits could take on slightly different structures.^{12,13} According to previous reports, the rate of inhibition of hydrogenases is limited not only by its reaction at the active site but also by O₂

access to the active site.¹⁴ But whether each subunit of 2,3-HPCD contains different O₂ pathways in a single homotetrameric enzyme is unknown. The precise, atomic-resolution pathways for O₂ migration in the protein, along with predicted relative significant parts of the pathways, should help to rationalize the selectivity of specific intermediate at the active site of each subunit and greatly facilitate the selection of specific site mutations for such studies.

The existence of gas channels that connect the active site to the solvent in proteins has been proposed on the basis of experimental and computational investigations of various enzymes, such as O₂ migration pathways for myoglobin,^{15–21} hemoglobin,^{22–24} copper amine oxidase,²⁵ [Fe]-hydrogenase,²⁶ human ferritin protein,²⁷ monooxygenase and oxidase flavoenzymes,²⁸ and H₂ pathways for [Fe]-hydrogenase,²⁶ [Ni–Fe]-hydrogenase,^{14,29} and [Fe–Fe]-hydrogenase.³⁰ In contrast, gas pathways in non-heme iron enzymes have only been characterized by molecular dynamics (MD) simulations for a small number of proteins so far, including O₂ channels in quercetin 2,3-dioxygenase³¹ and 12/15-lipoxygenase³² and NO diffusion in nitrile hydratase.³³ These studies have consistently suggested that gas pathways primarily originate from localized thermal fluctuations of protein structure. Hence, these pathways are transient and permanent gas tunnels cannot be detected in static crystal structures in the majority of cases.

Computational study on the process of ligand escape from the protein is challenging for two reasons:²⁰ First, the actual diffusion of gases may take hundreds of nanoseconds, a time scale that is hard to reach by straightforward molecular dynamics (MD) simulations; and second, the process does not take place at equilibrium state. One approximate computational method, named locally enhanced sampling (LES)³⁴ approach, is dynamics in nature. LES simulation of ligand diffusion in myoglobin mutants showed encouraging correlations with numerous experiments on a picosecond time scale.²⁰ Recent advances in

* Corresponding author: phone +86-411-84706223; fax +86-411-84706223; e-mail guixum@dlut.edu.cn.

[†] Department of Engineering Mechanics, State Key Laboratory of Structural Analyses for Industrial Equipment.

[‡] Department of Chemistry.

[§] School of Chemical Engineering, State Key Laboratory of Fine Chemicals.

computational power and methodology have made it possible to perform long time scale all-atom MD simulations of myoglobin (>90 ns) at room temperature in explicit solvent.¹⁸ However, in the case of 2,3-HPCD mentioned above, the system contains more than 100 000 atoms when the dioxygenase, substrate, and ligands are embedded in a periodic water box. The O₂ pathways and their relationship to function are still difficult to address. It seems reasonable to consider the homotetrameric enzyme as a whole structure because it has been deduced that the crystal packing defects may play important roles in regulating gas access to the buried active sites.²⁹

To investigate the possible O₂ pathways in 2,3-HPCD, the LES method was used to study O₂ diffusion in the protein. Energy of pair interaction between O₂ and residues lining the optimal pathways was calculated in order to elucidate the different reactivity of the subunits of the enzyme. Our results provide clear evidence that the rate of dioxygenase-catalyzed reaction is generally correlated with the high-affinity channel where the chance of finding O₂ is much higher.

2. Methods

2.1. Force Field Parameters for the Active-Site Residue.

The initial coordinates for MD simulation were obtained from the 1.9 Å resolution refined X-ray crystal structure (PDB entry 2IGA).^{12,35} This protein from *Brevibacterium fuscum* (EC 1.13.11.15) is homotetrameric, with each subunit containing an active site. In particular, four subunits (A–D) were bound by product, alkylperoxy, superoxo, and alkylperoxy intermediates, respectively. The active center is the five-coordinate Fe²⁺,^{4,36} bound by the conserved 2-His-1-carboxylate motif (His155, His214, and Glu267) and one doubly deprotonated 4-nitrocatechol (4NC).¹² Following previous experimental and computational work, the His, His, and Glu motif of the active site was modeled with two imidazoles and one acetate,^{5,8,10} with their initial orientations taken from the crystal structure (see Supporting Information). The initial position of 4NC, relative to His and Glu, was chosen as the result of superposition of 4NC with the three intermediates found in the crystal structure.

The CHARMM27 all-atom force field parameter set of the model structure, as well as the partial atomic charges, were calculated by using the PARATOOL program, a plugin included with the VMD visualization program that can be obtained free of charge.³⁷ Technical details of computations are provided in Supporting Information. The resulting atomic charges are given in Table S1 in Supporting Information. The nonbonded distances of interacting atoms at the active site of X-ray and MD averaged structures are also given in Table S2 in Supporting Information. These distances are almost the same as those observed in the X-ray structure (PDB entry 2IGA), indicating the method used here can reproduce the geometry of the active site observed in the crystal structure. Therefore, MD parameters for the conserved residues may be applicable for modeling the diversity of such enzymes.

2.2. LES Simulation.

The ligands in the original crystal structure were replaced by four 4NC dianions due to the fact that extradiol-cleaving enzymes utilize an ordered mechanism with substrate binding prior to O₂ activation.⁵ Two methods, structure superimposition and Autodock 4.0,³⁸ were used to determine the binding mode of 4NC in the active site of each subunit. Little difference could be discerned for the two approaches. Hydrogen atoms were added by using the PSFGEN plugin of VMD.³⁷ All His residues were protonated at N^e except His155 and His214 in each chain since they bind with the Fe²⁺ via N^e. MD simulation was performed by use of the program

NAMD 2.6 package³⁹ with CHARMM27 force field and LES algorithm as implemented in this code.³⁴ Enhanced space in this approach contained coordinates of 15 copies of O₂ in each subunit. Inspection of the alkylperoxy intermediate location in one subunit helped to manually insert an O₂ molecule into such a prepared 2,3-HPCD model. The structure with 4NC and O₂ bound in each subunit, taken together with crystal water molecules, was solvated to form a 100 × 118 × 123 Å simulation water box by use of the Solvate plugin of VMD. The resulting solvated system was neutralized by randomly adding Na⁺ ions in the bulk water by using the Autoionize plugin of VMD. The total number of atoms was 136 339, including 37 851 water molecules and 60 dioxygen molecules.

The system was first energy-minimized for 10 000 steps, while protein, 4NC, O₂, and crystal water molecules were kept fixed. Subsequently, the system was heated from 0 to 300 K in seven stages of 20 ps each, with heavy atoms restrained with 20 kcal mol⁻¹ Å⁻² spring constant and switched to 10, 5, 2.5, 1.2, 0.6, and 0.1 kcal mol⁻¹ Å⁻² for each stage. Finally, the restraint was completely removed, followed by an equilibration at 300 K for 100 ps. Langevin dynamics and Langevin piston method were used to maintain the temperature at 300 K and pressure at 1 bar. Particle mesh Ewald (PME) was used for long-range electrostatics. The multiple time steps of 2, 2, and 4 ps were used for bonded, nonbonded, and long-range electrostatic interactions, respectively. The production LES simulation (later referred to as LES_S1) was run for another 10 ns, while after 9 ns it was observed that all copies of O₂ had been escaped to the exterior of the protein. Moreover, to avoid the use of protein conformation too close to the crystal structure, the trajectory of 10-ns classical MD simulation of the substrate (4NC)–protein–water complex was first obtained, and then the final relaxed structure was used to perform a new LES MD simulation (with 15 O₂ in each subunit) following the same protocol described above (later referred to as LES_S2). It was found that 5 ns was required to observe all O₂ molecules to diffuse from the active site to the solvent.

LES is a useful technique to speed up the diffusion rate of gases and improve the simulation statistics,³⁴ which has been successfully used to investigate the transport of O₂ in hemoglobin,¹⁹ human cytoglobin,²⁴ human ferritin,²⁷ [Fe–Fe]-hydrogenase,²⁹ quercetin 2,3-dioxygenase,³¹ and photoactive protein nitrile hydratase.³³ As a mean field-based theory, LES is also used to increase the range of conformational sampling and has been applied to study protein-assisted RNA folding.^{40,41} Within this method, *N* (15 in this study) noninteracting copies of O₂ molecule were built. The nonbonding (van der Waals and electrostatic) potential of each copy with the rest of the system is scaled by a factor of 1/*N*. In this way, enhanced O₂ can experience *N* different trajectories and substantially increase sampling conformational space because the barriers are reduced. It has been believed that the LES trajectory with 10 copies may in principle give information equivalent to data obtained from 1 order of magnitude longer classical MD simulations.^{19,24} Thus, LES enables simulate 15 O₂ diffusion events simultaneously and gives a more reasonable picture of the O₂ movement inside each subunit of the protein.

3. Results and Discussion

Comparison of the structure of the four subunits after minimization and equilibration provides direct evidence on their similarity in conformation, suggesting that ligand induces little conformational rearrangement of the four independent subunits (see Figure S2 in Supporting Information). In two LES

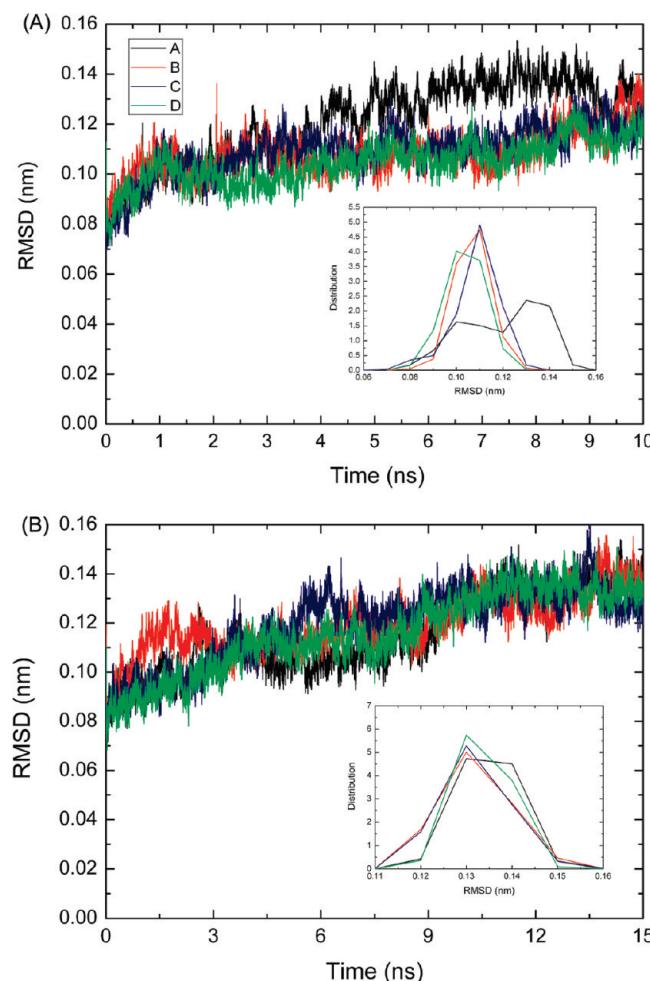


Figure 1. Root-mean-square deviations (rmsd) of C_α atoms of protein with respect to the X-ray structure (PDB entry 2IGA) for (A) LES_S1 and (B) LES_S2. Insets show the rmsd distribution histograms; only the last 5 ns trajectory was used for LES_S2.

simulations, the C_α root-mean-square deviation (rmsd) with respect to the crystal structure did not exceed 1.6 Å, as shown in Figure 1. Thus, the model used in the work is reasonably stable. In LES_S1, the deviations of subunit A are slightly larger than the other subunits, implying that subunit A is more flexible during the simulation. Moreover, it is evident from the rmsd distribution histogram (Figure 1A, inset) that subunit A explores more conformers with the largest peak located at a position different from the other subunits. However, after 10 ns of relaxation, the four subunits undergo almost the same conformational space (see Figure 1B). The main questions we want to address in this paper are as follows: What are the possible diffusion pathways for O₂ inside 2,3-HPCD? Does the homotetrameric protein possess the same or different O₂ pathways for each subunit? Is there a correlation between O₂ affinity and reaction rate in the active site of each subunit?

The collisions between O₂ and residues of each subunit are summarized in Table 1. A collision occurs if either oxygen atom of O₂ is found within a distance of 3.4 Å from any atom of a protein residue. The collision frequency is defined as the number of collisions between O₂ and one residue divided by the total number of collisions sampled from the 15 O₂ trajectories in each subunit. For clarity, only residues with a collision frequency exceeding 1% are listed in Table 1. However, to facilitate comparison, corresponding residues with a collision frequency less than 1% in this work are also listed in this table. To improve the statistics of collisions, only collisions during the process of exiting were counted, unless otherwise noted. Analysis of all

TABLE 1: Comparisons of Collision Frequencies of Different Residues with O₂ in LES_S1/LES_S2.^a

residue index	collision frequencies (%)			
	subunit A	subunit B	subunit C	subunit D
His155	1.91<0.01	1.73/2.11	10.21/4.73	0.50/1.29
Phe156	0.46/NA	0.37<0.01	2.07/0.49	0.11/0.12
Asn157	8.59/5.65	14.21/8.23	10.42/17.46	8.99/7.69
Val159	1.29/2.27	0.63/0.03	1.17/0.22	1.24/1.32
Ile181	NA/NA	<0.01/NA	NA/<0.01	5.99/NA
Ala190	<0.01<0.01	2.93<0.01	0.23/3.32	9.32<0.01
Ala191	NA/NA	0.65/NA	0.12/1.02	1.02<0.01
Trp192	2.86/0.81	11.41/3.01	6.69/15.86	10.49/6.76
His200	2.37/0.01	3.41/2.02	8.68/4.99	0.66/1.69
Thr202	0.26/NA	0.24/<0.01	2.11/1.15	0.10/<0.01
Ala203	2.45/0.07	8.38/<0.01	8.63/15.13	5.33/5.27
Leu204	NA/NA	1.02/<0.01	0.12/1.27	1.63/NA
Thr205	14.92/16.65	10.92/15.50	9.70/4.90	10.55/14.41
His213	13.81/13.20	6.57/10.81	8.81/1.17	3.67/9.83
His214	5.32/4.55	5.35/4.67	2.45/5.58	1.18/3.45
Arg292	0.02/NA	0.02/<0.01	<0.01/0.02	2.73/0.02
Arg293	9.13/6.27	7.94/9.92	5.25/4.98	11.30/8.91
Trp304	13.89/12.95	7.55/11.10	7.96/4.22	3.73/9.52
Tyr305	11.67/15.02	4.58/12.28	7.55/1.06	3.77/14.59
Glu329	1.03/0.74	0.56/0.14	0.26/0.11	4.20/2.25
4NC	5.78/4.03	9.30/2.47	4.11/10.52	3.35/5.14

^a NA indicates there is no contact within 3.4 Å in both LES simulations.

MD trajectories and O₂ collisions shows that there is one dominant O₂ diffusion pathway in each subunit. It should be

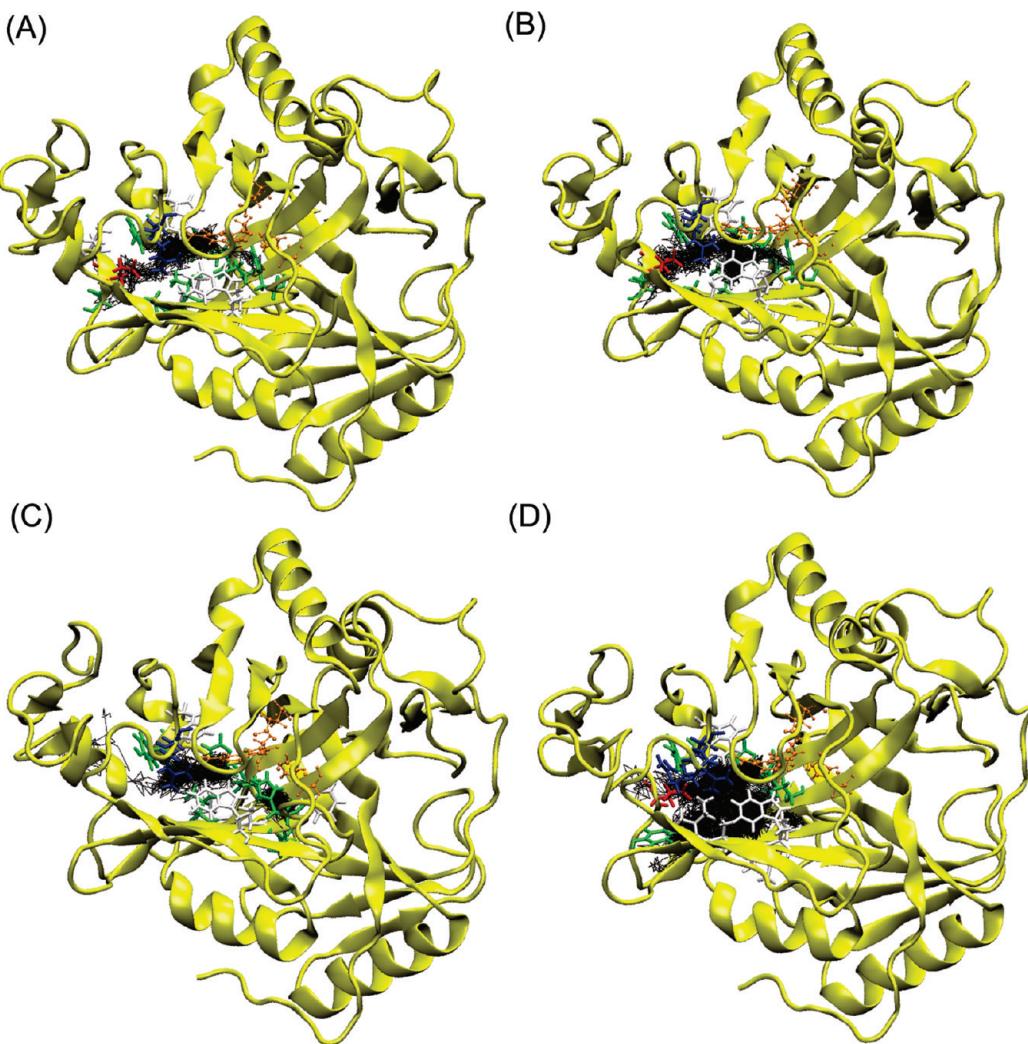


Figure 2. Typical trajectories of O₂ diffusion in four subunits (A–D) of 2,3-HPCD. Residues with a collision frequency greater than 1% are shown in Licorice representation. The conserved residues His155, His214, and Glu267 and the ligand 4NC are shown in CPK representation. Residues in white, blue, red, and green indicate nonpolar, basic, acidic, and polar.

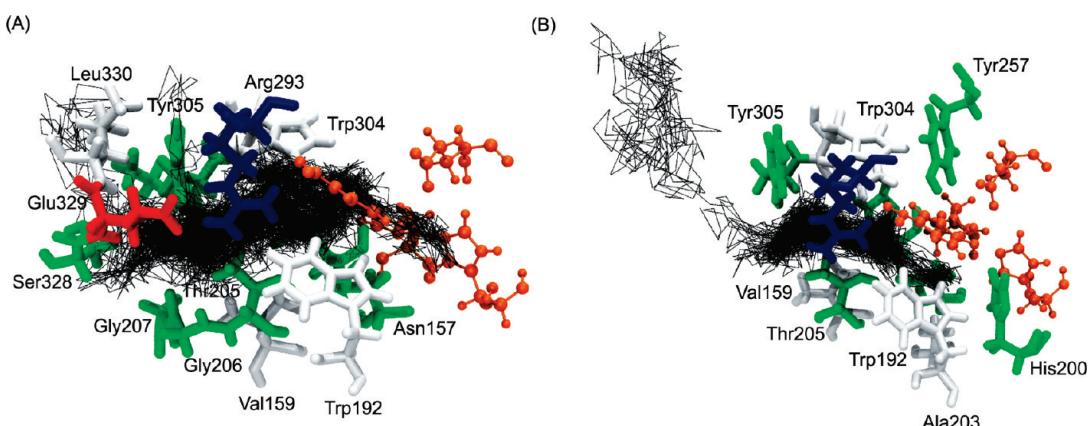


Figure 3. Trajectories of one copy of O₂ (A) leaving the protein and (B) returning to the active site in LES_S1. The residue representation is the same as in Figure 2.

noted that, in LES_S1, one copy of O₂ in subunit A was found to leave the protein and then return to the active site. Such a rare event was not detected in LES_S2. Figure 2 gives the traces of the center of mass of one chosen O₂ during MD simulation until leaving the protein matrix, as well as residues lining the energy preferred channel in each subunit (see Figure 3).

From Table 1 and Figure 2, it is clear that O₂ moved from the conserved active site of subunit A through Asn157, Val159; then

crossed a pocket composed of Thr205, His213, Arg293, Trp304, and Tyr305; and finally left the protein via Glu329. Residues that may affect O₂ diffusion significantly could be Thr205, His213, Arg293, Trp304, and Tyr305, with collision frequencies of 14.92%, 13.81%, 9.13%, 13.89%, and 11.67% in LES_S1 and 16.65%, 13.20%, 6.27%, 12.95%, and 15.02% in LES_S2, respectively.

As mentioned above, in LES_S1, one copy of O₂ in subunit A that left the protein returned to the active site and remained

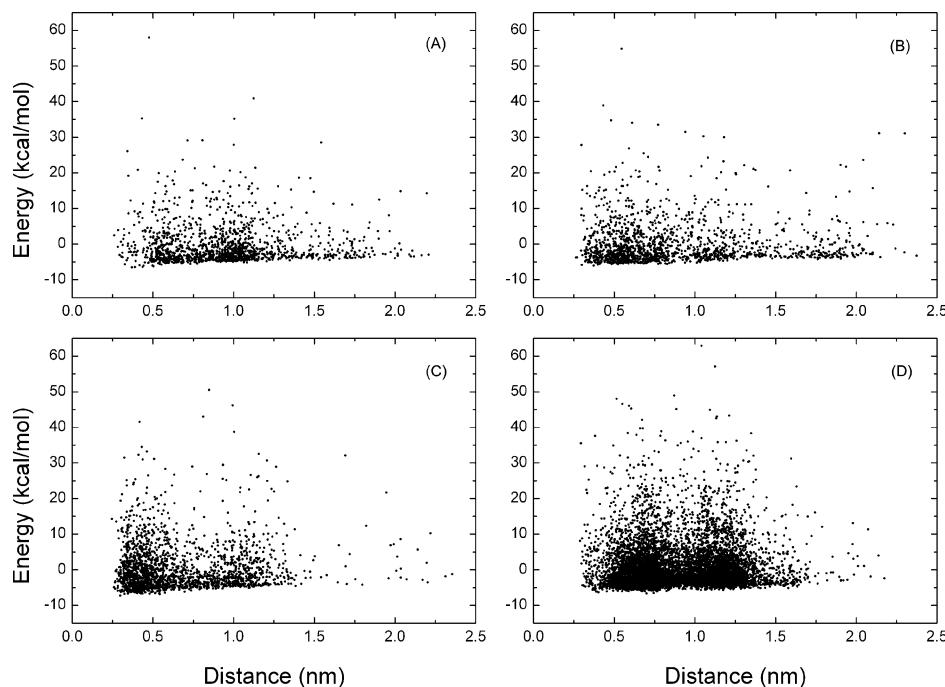


Figure 4. Distribution of pair interaction energies between O_2 and the protein along the optimal pathway of each subunit (A–D) in LES_S1.

there on the time scale of nanoseconds. Schematic routes and all determined residues with a collision frequency greater than 1% are represented in Figure 3. Note that the O_2 diffusion pathway is almost identical, but O_2 collides with a larger number of amino acids when it escapes from the protein (Figure 3A). We also notice that the time (~ 5.34 ns) spent by this O_2 in subunit A is much longer than the other copies of O_2 . Meanwhile, as mentioned above and shown in Figure 1A, subunit A explores alternative conformations after ~ 5 ns and becomes more mobile. Thus, it is possible for O_2 to move back to the active site along the same pathway.

For subunit B in LES_S1, the most significant difference from subunit A is that O_2 collides with Trp192 more frequently. This bulky hydrophobic residue along O_2 route favors O_2 localization, and as a result, the exit events are somewhat delayed in the current LES_S1 simulation. But in LES_S2, the time required for O_2 to leave subunits A and B is nearly identical, perhaps because the initial structure has been relaxed for 10 ns and the difference in conformation of Trp192 is not so obvious.

Both LES simulations indicate that O_2 are localized in the active site of subunit C for most of the time, colliding with His155, Asn157, Trp192, His200, and Ala203. Therefore, it could be more difficult for O_2 to move toward the active site. It is also expected that the common exit path for O_2 is formed by residues Thr205, His213, Arg293, Trp304, and Tyr305.

It is also surprising that the space of O_2 diffusion is significantly wider in subunit D (Figure 2D), suggesting that O_2 is trapped in the pocket composed by more hydrophobic residues, such as Asn157, Val159, Trp192, and Ala203. Analyses of O_2 trajectories in both LES simulations give the same overall picture as shown in Figure 2D. For example, the same residues of Thr205, His213, Arg293, Trp304, and Tyr305 are involved in the formation of the common exit path for O_2 as in other subunits.

Another measure of the extent of finding O_2 in the region around the active site is obtained by examination of pair interaction energy along each O_2 pathway. It should be noted that in the LES method the energy barrier between O_2 and the other part of the system is reduced. In this way, 15 copies of

O_2 could achieve more efficient sampling of the conformational space than just one. Thus, it is of interest to search for the optimal route with lowest energy barrier. If the pair interaction energy is relatively low for a certain region of the protein, then the chance of finding O_2 in that region could be higher. NAMD supports the calculation of interaction energy between two groups of atoms.³⁹ The 10-ns trajectory of LES_S1 and 5-ns trajectory of LES_S2 are used to calculate the interaction energy (van der Waals and electrostatic contributions within a cutoff of 12 Å) between each copy of O_2 and the protein. Then the energy-preferred O_2 pathway leading from the active site to the solvent was identified for each subunit. In this way, the effect of subtle structural change induced by crystal packing forces or ligand migration can be detected.

The distributions of pair interaction energy along each low-barrier O_2 pathway in LES_S1 and LES_S2 are shown in Figures 4 and 5. The energy profiles indicate that the O_2 pathways in the nominally identical subunits are characterized by different O_2 binding potentials, that is, A > B > C > D for LES_S1 and A > B > D > C for LES_S2, which relates well to the fact that the catalytic product is found in subunit A, the unstable alkylperoxo intermediate in subunit B and D, and another unstable superoxide intermediate in subunit C.¹² However, it should be noted that the alkylperoxo intermediate was found to be simultaneously presented in subunit B and D.¹² Thus, the rate of catalysis depends not only on reactant O_2 access to the active site but also on the enzyme capacity to catalyze the reaction at the active site, as another study suggested.¹⁴ As a result, characterization of O_2 pathways alone could not be expected to explain all of the experimental outcomes. Recent experiments have shown that the use of an alternative substrate, 4-sulfonylcatechol, and the Glu323Leu variant of HPCD has led to formation of a new intermediate that appears to occur between the alkylperoxo intermediate and the product complexes in the reaction cycle.⁴²

The energy distributions for O_2 movement along the four channels are consistent with the trajectories of O_2 as shown in Figure 2, which means local increase of O_2 collisions causes high-energy barriers to be overcome along the putative path-

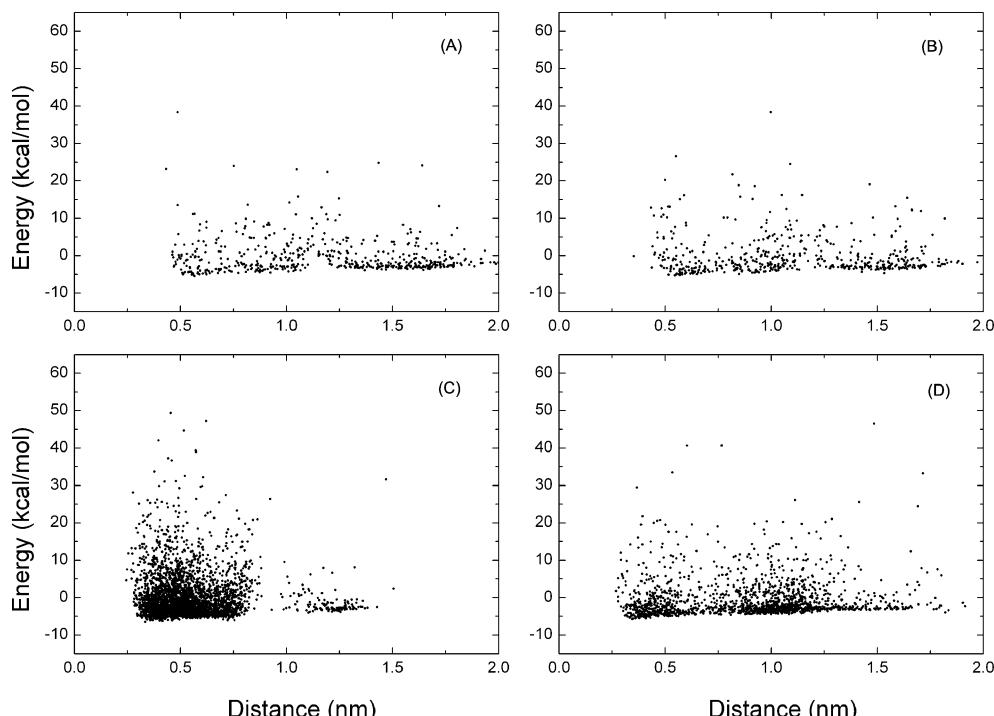


Figure 5. Distribution of pair interaction energies between O₂ and the protein along the optimal pathway of each subunit (A–D) in LES_S2.

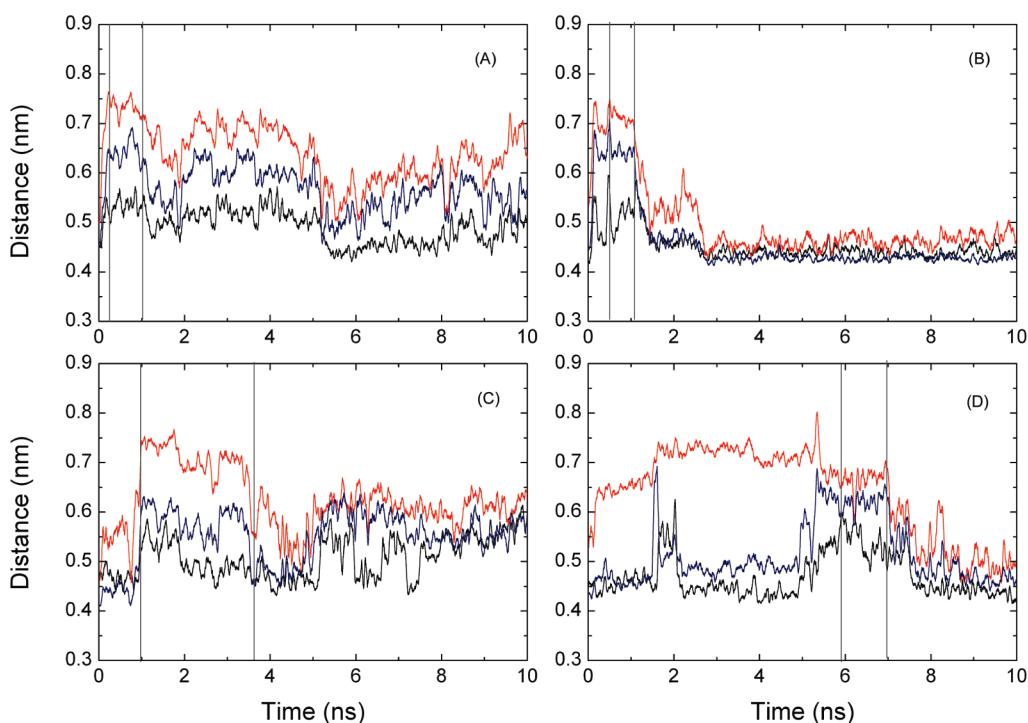


Figure 6. Distances between Thr205 C_γ₂ and His213 C_δ₂ (black), Thr205 C_γ₂ and Trp304 C_h₂ (red), and His213 C_δ₂ and Trp304 C_h₂ (blue), observed in LES_S1. Gray vertical lines indicate time spent by O₂ in the cavity.

ways. A distribution peak located at 10–11 Å away from the origin (the position of iron) can be seen from Figures 4 and 5 for all channels. Trajectories of amino acids show that this region is between Thr205, His213, Arg293, Trp304, and Tyr305, just below the protein surface, suggesting they are critical for O₂ migration.

To assess the effect of side chain flexibility of relevant residues on the formation of cavities, the x_1 (N—C_α—C_β—C_γ) and x_2 (C_α—C_β—C_γ—C_δ_A) dihedral angles of these amino acids were examined. In LES_S1, it was found that except for Arg293,

the side-chain conformations of other amino acids were very similar. In particular, the x_2 angles of Arg293 were found to be extraordinarily distinct (data not shown). However, such a difference is not observed in LES_S2. The dynamic behavior of cavity and gating motions of surrounding amino acids were also examined by calculating the distance between different residues for both LES simulations (Figures 6 and 7). These results consistently reveal that the change in the volume of the cavity defined by Thr205, His213, Arg293, Trp304, and Tyr305 correlates with the time dependence of the number of O₂

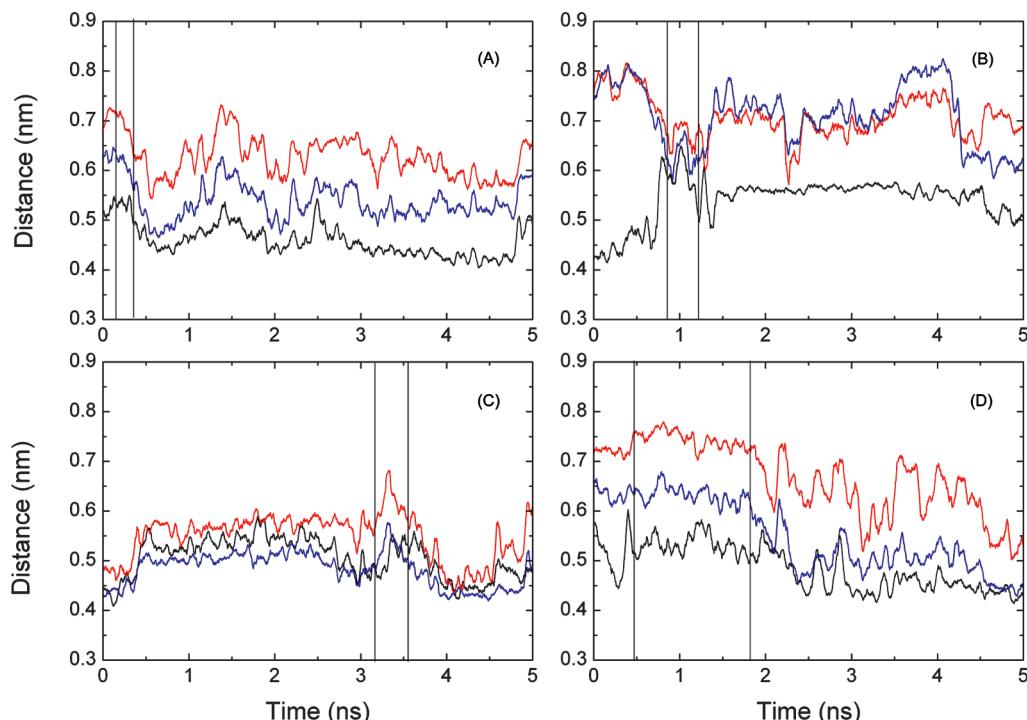


Figure 7. Distances between Thr205 $C_{\gamma 2}$ and His213 $C_{\delta 2}$ (black), Thr205 $C_{\gamma 2}$ and Trp304 $C_{\text{h}2}$ (red), and His213 $C_{\delta 2}$ and Trp304 $C_{\text{h}2}$ (blue), observed in LES_S2. Gray vertical lines indicate time spent by O_2 in the cavity.

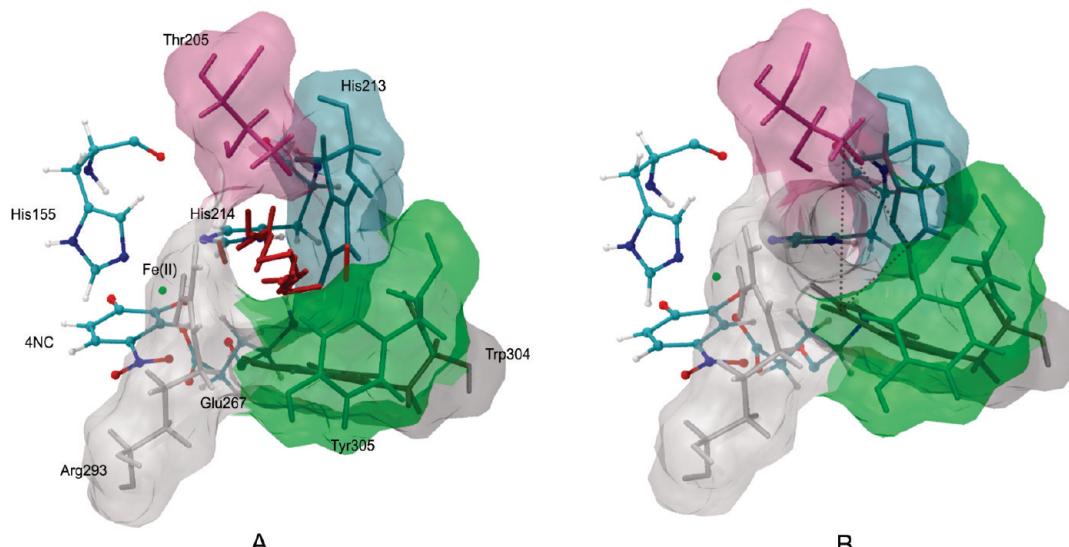


Figure 8. Snapshots of (A) open and (B) closed state of the cavity defined by Thr 205, Arg293, His213, Trp304 and Tyr305. Dotted line in panel B indicates the distance between corresponding atom of residue used in Figure 7. Fe-bound ligands are shown in CPK representation, and residues surrounding the cavity are shown as sticks (in Licorice representation) and transparent surfaces. O_2 molecules are shown as red sticks.

molecules in the cavity as identified by inspection of the O_2 trajectory. When O_2 molecules diffuse into the cavity, and the distance from Thr205 to His213, Thr205 to Trp304, and His213 to Trp304 increases concomitantly, the expansion of the cavity results in an open channel that enables O_2 passage. The transition from the open state to the closed one is associated with a decrease in the distance between these residues and corresponding cavity volume, as shown in Figure 8. Although different starting structures were used in both LES simulations, the present results support the mechanism that the breathing motion of internal cavities plays a crucial role in ligand migration in a protein matrix.²¹ In addition, it is worth noting that both LES simulations suggest different motions of these residues, resulting in different time spent by O_2 to leave the protein. These data

may help to design mutants of 2,3-HPCD and to check the nature of the O_2 diffusion pathway experimentally.

4. Conclusions

The present investigation focuses on the possible O_2 diffusion pathways in 2,3-HPCD, which is complementary to experimental observation of the simultaneous presence of different intermediates in nominally identical subunits of a single enzyme.¹² The different reaction rates in four independent active sites were ascribed to the crystal packing forces that cause the subunits to take on slightly different structures.^{12,13} In an attempt to elucidate how these subtle differences can influence the enzyme to exert precise control over the O_2 binding, activation, and insertion

steps of the catalytic cycle, we performed two LES MD simulations on the homotetrameric protein solvated in a water box. Our results clearly show that O₂ diffusion pathways in the four subunits are different. Recent experimental studies of Ni–Fe hydrogenase suggest that the rate of dioxygenase-catalyzed reaction is generally correlated with the rate of O₂ access to the active center.¹⁴ The comparison of collisions and pair interaction energy indicates that the pathway in subunit A exhibits a higher probability for finding O₂ than the other subunits, consistent with experimental observation that the reaction product was found in it. In contrast, the average time required for O₂ to escape from subunit B is longer than subunit A (see Figures 6 and 7); meanwhile, the energy barriers along this route of O₂ diffusion are also higher (see Figures 4 and 5). Similarly, the chance of finding O₂ in the region around the active site of subunits C and D is much less, providing one reasonable explanation for the fact that the reaction intermediates reside in them.

Residues that lead to the breathing motion of the cavity are identified as the main bottleneck of the channel for O₂ diffusion to the active site. The opening of the tunnel stems from collectively concerted motions of residues of Thr205, His213, and Trp304. The dynamics of these amino acids display that the diffusion of O₂ in 2,3-HPCD is clearly restricted to specific regions close to the conserved active-site domain. The existence of specific O₂ transport pathways offers important clues on various mutation studies of 2,3-HPCD. To this end, these observations will be valuable to engineer non-heme iron dioxygenases in order to find intermediates by characterizing O₂ diffusion routes.

Acknowledgment. This work was supported by the Youth Foundation of DLUT (893103), National Natural Science Foundation (10772042), the National Basic Research Program of China (Grants 2004CB518901 and 2009CB918501), and the High Science and Technology Project (2006AA01A124) of China.

Supporting Information Available: Computational details of the force field parameters and tables of atomic partial charges, nonbonded distances between Fe and its coordinated residues, and additional force field parameters for 4NC dianion. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) Bugg, T. D. H.; Lin, G. *Chem. Commun.* **2001**, 941.
- (2) Siegbahn, P. E. M.; Haeffner, F. *J. Am. Chem. Soc.* **2004**, 126, 8919.
- (3) Wallar, B. J.; Lipscomb, J. D. *Chem. Rev.* **1996**, 96, 2625.
- (4) Solomon, E. I.; Brunold, T. C.; Davis, M. I.; Kemsley, J. N.; Lee, S. K.; Lehnert, N.; Neese, F.; Skulan, A. J.; Yang, Y. S.; Zhou, J. *Chem. Rev.* **2000**, 100, 235.
- (5) Costas, M.; Mehn, M. P.; Jensen, M. P.; Que, L., Jr. *Chem. Rev.* **2004**, 104, 939.
- (6) Kryatov, S. V.; Rybak-Akimova, E. V. *Chem. Rev.* **2005**, 105, 2175.
- (7) Abu-Omar, M. M.; Loaiza, A.; Hontzeas, N. *Chem. Rev.* **2005**, 105, 2227.
- (8) Bugg, T. D. H. *Tetrahedron* **2003**, 59, 7075.
- (9) Lipscomb, J. D. *Curr. Opin. Struct. Biol.* **2008**, 16, 644.
- (10) Deeth, R. J.; Bugg, T. D. H. *J. Biol. Inorg. Chem.* **2003**, 8, 409.
- (11) Georgiev, V.; Borowski, T.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2008**, 13, 929.
- (12) Kovaleva, E. G.; Lipscomb, J. D. *Science* **2007**, 316, 453.
- (13) Kovaleva, E. G.; Neibergall, M. B.; Chakrabarty, S.; Lipscomb, J. D. *Acc. Chem. Res.* **2007**, 40, 475.
- (14) Leroux, F.; Dementin, S.; Burlat, B.; Cournac, L.; Volbeda, A.; Champ, S.; Martin, L.; Guigliarelli, B.; Bertrand, P.; Fontecilla-Camps, J.; Roussel, M.; Léger, C. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, 105, 11188.
- (15) Teeter, M. M. *Protein Sci.* **2004**, 13, 313.
- (16) Cohen, J.; Arkhipov, A.; Braun, R.; Schulten, K. *Biophys. J.* **2006**, 91, 1844.
- (17) Cohen, J.; Schulten, K. *Biophys. J.* **2007**, 93, 3591.
- (18) Ruscio, J. Z.; Kumar, D.; Shukla, M.; Prisant, M. G.; Murali, T. M.; Onufriev, A. V. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, 105, 9204.
- (19) Orlowski, S.; Nowak, W. *Theor. Chem. Acc.* **2007**, 117, 253.
- (20) Elber, R.; Gibson, Q. H. *J. Phys. Chem. B* **2008**, 112, 6147.
- (21) Tomita, A.; Sato, T.; Ichiyanagi, K.; Nozawa, S.; Ichikawa, H.; Chollet, M.; Kawai, F.; Park, S. Y.; Tsuduki, T.; Yamato, T.; Koshihara, S.; Adachi, S. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, 106, 2612.
- (22) Daigle, R.; Guertin, M.; Lagüe, P. *Proteins: Struct., Funct., Bioinf.* **2009**, 75, 735.
- (23) Boechi, L.; Martí, M. A.; Milani, M.; Bolognesi, M.; Luque, F. J.; Estrin, D. A. *Proteins: Struct., Funct., Bioinf.* **2008**, 73, 372.
- (24) Orlowski, S.; Nowak, W. *J. Mol. Model.* **2007**, 13, 715.
- (25) Johnson, B. J.; Cohen, J.; Welford, R. W.; Pearson, A. R.; Schulten, K.; Klinman, J. P.; Wilmut, C. M. *J. Biol. Chem.* **2007**, 282, 17767.
- (26) Cohen, J.; Kim, K.; Posewitz, M.; Ghirardi, M. L.; Schulten, K.; Seibert, M.; King, P. *Biochem. Soc. Trans.* **2005**, 33, 80.
- (27) Ciacchi, L. C.; Payne, M. C. *Chem. Phys. Lett.* **2004**, 390, 491.
- (28) Baron, R.; Riley, C.; Chenprakhon, P.; Thotsaporn, K.; Winter, R. T.; Alfieri, A.; Forneris, F.; van Berkel, W. J. H.; Chaiyen, P.; Fraaije, M. W.; Mattevi, A.; McCammon, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, 106, 10603.
- (29) Cohen, J.; Kim, K.; King, P.; Seibert, M.; Schulten, K. *Structure* **2005**, 13, 1321.
- (30) Teixeira, V. H.; Baptista, A. M.; Soares, C. M. *Biophys. J.* **2006**, 91, 2035.
- (31) Fiorucci, S.; Golebiowski, J.; Cabrol-Bass, D.; Antonczak, S. *Proteins: Struct., Funct., Bioinf.* **2006**, 64, 845.
- (32) Saam, J.; Ivanov, I.; Walther, M.; Holzhütter, H. G.; Kuhn, H. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, 104, 13319.
- (33) Kubiak, K.; Nowak, W. *Biophys. J.* **2008**, 94, 3824.
- (34) Elber, R.; Karplus, M. *J. Am. Chem. Soc.* **1990**, 112, 9161.
- (35) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, 28, 235.
- (36) Solomon, E. I.; Decker, A.; Lehnert, N. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, 100, 3589.
- (37) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph. Model.* **1996**, 14, 33.
- (38) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, 19, 1639.
- (39) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, 26, 1781.
- (40) Cojocaru, V.; Klement, R.; Jovin, T. M. *Nucleic Acids Res.* **2005**, 33, 3435.
- (41) Cojocaru, V.; Nottrott, S.; Klement, R.; Jovin, T. M. *RNA* **2005**, 11, 197.
- (42) Kovaleva, E. G.; Lipscomb, J. D. *Biochemistry* **2008**, 47, 11168.