

# Oxygen Entry through Multiple Pathways in T-State Human Hemoglobin

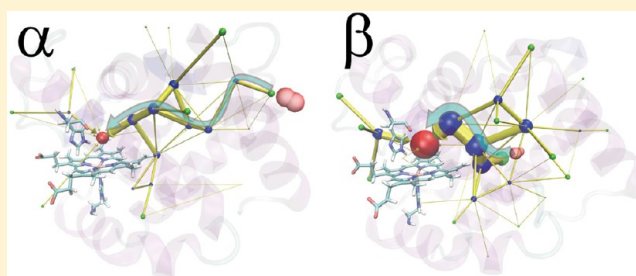
Masayoshi Takayanagi,<sup>†,‡,§</sup> Ikuo Kurisaki,<sup>‡,§</sup> and Masataka Nagaoka<sup>\*,‡,§</sup>

<sup>†</sup>Venture Business Laboratory and <sup>‡</sup>Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan

<sup>§</sup>Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Honmachi, Kawaguchi 332-0012, Japan

**S** Supporting Information

**ABSTRACT:** The heme oxygen ( $O_2$ ) binding site of human hemoglobin (HbA) is buried in the interior of the protein, and there is a debate over the  $O_2$  entry pathways from solvent to the binding site. As a first step to understand HbA  $O_2$  binding process at the atomic level, we detected all significant multiple  $O_2$  entry pathways from solvent to the binding site in the  $\alpha$  and  $\beta$  subunits of the T-state tetramer HbA by utilizing ensemble molecular dynamics (MD) simulation. By executing 128 independent 8 ns MD trajectories in  $O_2$ -rich aqueous solvent, we simulated the  $O_2$  entry processes and obtained 141 and 425  $O_2$  entry events in the  $\alpha$  and  $\beta$  subunits of HbA, respectively. We developed the intrinsic pathway identification by clustering method to achieve a persuasive visualization of the multiple entry pathways including both the shapes and relative importance of each pathway. The rate constants of  $O_2$  entry estimated from the MD simulations correspond to the experimentally observed values, suggesting that  $O_2$  ligands enter the binding site through multiple pathways. The obtained multiple pathway map can be utilized for future detailed analysis of HbA  $O_2$  binding process.



## 1. INTRODUCTION

Human hemoglobin (HbA) is an oxygen ( $O_2$ ) transport heme protein and is composed of four subunits, two  $\alpha$  and two  $\beta$  subunits. Each subunit has a heme group and reversibly binds an  $O_2$  molecule as a sixth ligand. Efficient  $O_2$  transport by HbA is achieved by allosteric regulation, which regulates the affinity of  $O_2$  ligand molecule to the heme. Understanding the HbA function of  $O_2$  transport at the atomic level is a great challenge for decades.<sup>1</sup>

The heme  $O_2$  binding site is buried in the interior of the subunits, and thus we should first investigate the  $O_2$  migration pathways from outer solvent to the binding site to comprehensively understand the  $O_2$  binding and dissociation processes at the atomic level. Although there are several internal cavities (for example, cavities identified as Xe sites<sup>2</sup>) inside the  $\alpha$  and  $\beta$  subunits in the crystal structures,<sup>3</sup> there is no tunnel through which diatomic ligand molecules (such as  $O_2$ , CO, NO) can access the binding site. Therefore, it is indispensable to investigate transient tunnels that are temporally created by the thermal structural fluctuations.

Several experimental works of hemoglobin (Hb) and myoglobin (Mb) have revealed that ligand molecules dissociated from the heme migrate into the internal cavities. It was revealed by time-resolved X-ray crystallography of photolyzed MbCO that dissociated CO ligands transfer to Xe cavities before escaping to solvent.<sup>4–6</sup> In the crystallographic structures of globular proteins at cryogenic temperature,

dissociated CO ligands are trapped in the inside cavities.<sup>7–10</sup> Multiple CO migration pathways of Mb through inside cavities were also analyzed by the molecular volume changes observed by transient grating technique.<sup>11</sup> However, it is only the ensemble-averaged distributions of ligands inside cavities that these experiments can show, and thus the ligand portals through which ligand molecules enter or escape between the outer solvent and internal cavities have not been directly observed. Therefore, there remains debate regarding the ligand entrance and exit portals.

Since an early X-ray crystal structure analysis of Hb,<sup>12</sup> it has long been assumed that ligands directly enter the binding site from solvent through “histidine gate” (His-gate), which is located at the distal histidine ( $\alpha$ His58 and  $\beta$ His63 in HbA or His64 in Mb) in globular proteins such as Mb and Hb. In fact, the His gate can be opened by the outward movement of the distal histidine side chain.<sup>13</sup> There is some experimental evidence that the main ligand portal in Mb and HbA is the His gate. Exhaustive mutagenesis mapping with geminate ligand recombination experiments has been applied to Mb and HbA, and only the mutations of His gate or in the active site residues had significant effects on the rates of ligand entry and escape.<sup>14–18</sup> A time-resolved X-ray crystallography of photo-

**Received:** February 9, 2013

**Revised:** April 24, 2013

lyzed L29W mutant Mb exhibited the extremely long existence time  $\sim 1$  ms in the Xe1 site, meaning that there is no ligand escape portals in the Xe cavities and thus the His gate should be the main portal.<sup>6</sup>

To identify the ligand portals, molecular dynamics (MD) simulation is a powerful computational technique because it can analyze each ligand migration event at the atomic level. In fact, computational studies with MD simulations have been conducted to elucidate ligand escape and entry processes inside globular proteins.<sup>19–29</sup> For HbA, there are several computational studies to investigate ligand escape processes. CO escape pathways in T- and R-state HbA were analyzed by protein energy landscape exploration (PELE) analysis,<sup>19</sup> and O<sub>2</sub> escape pathways in the isolated  $\alpha$  subunit were analyzed by ensemble MD approach.<sup>20</sup> Recently, temperature-controlled locally enhanced sampling (TLES) method was employed to obtain O<sub>2</sub> escape pathways in T-state HbA.<sup>21</sup> These reported the multiple ligand escape pathways inside HbA through multiple portals: the His-gate, a short pathway around the CD-corner, a long pathway extending to the space among A, G, and H-helix, and so on. The computationally obtained multiple portals contrast with the experimentally obtained single main portal: His gate.

Moreover, it is also assumed that structural perturbations by the existence of ligands inside protein cavities should also play important roles for the gate dynamics between cavities. In Mb, it has been shown that the CO dissociation from the heme brings about the tertiary structural changes and excess vibrational energy relaxation in the entire protein<sup>13,30–34</sup> and CO migrations through transient pathways involve collective motions correlated with local gating motions.<sup>25</sup> In addition, the existence of CO molecule inside cavities brings about the structural fluctuation, referred to as “breathing motion”, and promotes ligand transfer by opening the gates between cavities.<sup>8,9</sup> Therefore, it is expected that ligand molecules should be explicitly treated for computational analysis.

Furthermore, to understand ligand migration processes through multiple pathways, we need an efficient method for visualization of ligand migrations that includes both the shapes and the relative importance of each pathway. However, such visualization method has not yet been established, and simple visualizations have been performed. For example, pathways were visualized by the superimposition of ligand locations of trajectories from start to end (for example, from the binding site to the outer solvent)<sup>19,21,22</sup> and by the drawing of temporal cavities, which change in conjunction with the thermal fluctuations of proteins.<sup>23</sup> Although both have been used to illustrate the shapes of pathways, these simple visualizations themselves have no information regarding the relative importance of each pathway.

In this work, as a first step to understand HbA O<sub>2</sub> binding process at the atomic level, we detected all of the significant multiple O<sub>2</sub> entry pathways and portals from solvent to the binding site in the  $\alpha$  and  $\beta$  subunits of the T-state tetramer HbA. This is the first computational report of the detection of the multiple O<sub>2</sub> entry pathways in HbA. To correctly include the structural perturbations by the existence of O<sub>2</sub> inside cavities, we performed MD simulations with explicit O<sub>2</sub> molecules. We used ensemble MD approach<sup>13,30,31</sup> and executed 128 independent 8 ns MD simulations (total MD length is  $\sim 1$   $\mu$ s) at high O<sub>2</sub> concentrations to accelerate the O<sub>2</sub> entry process. The numbers of entry trajectories appeared in the ensemble MD calculations were 141 and 425 for the  $\alpha$  and

$\beta$  subunits, respectively, and are large enough to obtain all of the significant entry portals. We also developed the intrinsic pathway identification by clustering (IPIC) method, which achieved a persuasive visualization with which one can comprehend both the shapes and the relative importance of the multiple entry pathways. The rate constants of O<sub>2</sub> entry estimated from the MD simulations correspond to the experimentally observed values, suggesting that O<sub>2</sub> ligands enter the binding site through multiple pathways.

## 2. METHODS

**System Modeling.** The AMBER parm99 force field and TIP3P water<sup>35</sup> were used for amino acid residues and water molecules, respectively. For the deoxy heme parameters, we used bond, angle, and dihedral parameters, as developed by Henry et al.,<sup>36</sup> while for the electrostatic and van der Waals parameters we used those determined by Giammona<sup>37</sup> given in the contrib directory in the AMBER9 package.<sup>38</sup> For the O<sub>2</sub> molecule parameters, the atomic charges of both oxygen atoms were zero, the Lennard-Jones potential parameters were ( $\sigma$ ,  $\epsilon$ ) = (1.60 Å, 0.2 kcal/mol), and the O–O bond parameters were ( $K_r$ ,  $r_{eq}$ ) = (848.0 kcal/mol·Å<sup>2</sup>, 1.25 Å). Here we ignored the quadrupole moment of O<sub>2</sub> (−0.4 D Å) because it was reported that there was little difference in the CO dynamics inside Mb heme pocket by including the quadrupole moment of CO (−2.4 D Å), which is larger than that of O<sub>2</sub>.<sup>39</sup>

As an initial structure of HbA, we used the crystal structure of T-state deoxy HbA (PDB ID: 2DN2).<sup>3</sup> Crystal water molecules in the binding site of  $\alpha_1$  and  $\alpha_2$  subunits were removed because our preliminary simulations revealed that these water molecules strongly block the O<sub>2</sub> entry into the  $\alpha$  subunit binding site (only one O<sub>2</sub> entry occurred in 5 ns 32 MD trajectories), and our current purpose is to obtain O<sub>2</sub> entry pathways. Missing hydrogen atoms were added by the LEaP module of AMBER9 program package. The protonation states of histidine residues at pH 7.0 were assigned according to the pK<sub>a</sub> values estimated by PROPKA3.0 program:<sup>40</sup>  $\alpha$ His50,  $\alpha$ His72,  $\beta$ His117, and  $\beta$ His146 were doubly protonated, the proximal histidines ( $\alpha$ His87 and  $\beta$ His92) were singly protonated in the  $\delta$  position, and the rest were singly protonated in the  $\epsilon$  position. Then, the tetramer T-state HbA was energy minimized for 200 cycles and solvated with 11 903 TIP3P water molecules, 2 chloride counterions, and 120 O<sub>2</sub> molecules, yielding a periodic box size of  $\sim 76.7 \times 76.7 \times 76.7$  Å<sup>3</sup> after equilibration under the condition used in this work (temperature 310 K and pressure 1 atm). The concentration of O<sub>2</sub> molecule was  $\sim 0.55$  mol/L, which is 440-fold higher than the saturating concentration 1.25 mmol/L at 297 K and 1 atm. This high O<sub>2</sub> concentration was necessary to accelerate the slow O<sub>2</sub> entry process of tens of microseconds even under the O<sub>2</sub> saturating concentration.<sup>18</sup>

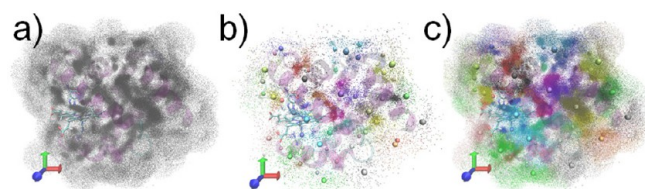
**Molecular Dynamics Simulation Procedure.** The PMEMD module of AMBER 9 package was used to execute all MD simulations. We used the particle mesh Ewald (PME) procedure to handle long-range electrostatic interactions under periodic boundary conditions. The SHAKE method was used to constrain the hydrogen-heavy atom bond distances, and the integration time step was 2 fs. For NPT (constant pressure and temperature) and NVT (constant volume and temperature) MD simulations, the weak coupling algorithm<sup>41</sup> was used to keep the pressure and temperature.

To prepare solution with randomly distributed O<sub>2</sub> molecules, a high-temperature (750 K) NVT MD for 300 ps and a

relaxation NPT MD for 200 ps at the target temperature and pressure (310 K and 1 atm) were executed with the restraints on the protein structure. Then, a production NPT MD at 1 atm and 310 K was executed for 8 ns without any restraints, saving snapshots every 1 ps. These high temperature, relaxation, and production MDs were repeated 128 times starting from the different initial velocities.

**Clustering O<sub>2</sub> Distribution in the Vicinity of Hb Subunits.** A hierarchical clustering algorithm with the Ward's minimum variance method<sup>42</sup> implemented in the *hclust* function in the R package (version 2.13.1)<sup>43</sup> was used to analyze the O<sub>2</sub> migration pathways inside Hb subunits.

The centers of mass (COM) distributions of O<sub>2</sub> in the vicinity of the  $\alpha$  and  $\beta$  subunits were analyzed by the clustering. In Figure 1, the clustering procedure of the  $\beta$  subunit



**Figure 1.** Clustering analysis procedure in the  $\beta$  subunit. (a) Distribution of the COMs of O<sub>2</sub> (drawn as gray dots) within a distance of 4.0 Å from any non-hydrogen atoms of the RMS fitted  $\beta$  subunits. Because of the limitation of the drawing program, we drew only 700 000 COMs of total  $\sim 37$  million COMs. (b) Result of the clustering analysis of the 31 454 representative COMs using the Ward's method with the number of cluster 65. The COMs and centers of clusters are drawn with dots and spheres, respectively, using different colors for each cluster. (c) Assignment of all of the COMs based on the clustered representative COMs in panel b.

distribution is shown. All of the images of protein structures were drawn with VMD 1.9.<sup>44</sup> To obtain the COMs in the vicinity of the  $\beta$  subunits, we executed an RMS fit of  $\beta_1$  subunit C $\alpha$  atoms to the crystal  $\beta_1$  subunit structure and that of  $\beta_2$  subunit to the same crystal  $\beta_1$  subunit structure. Then, we collected the COMs within a distance of 4.0 Å from any non-hydrogen atoms of the fitted  $\beta_1$  and  $\beta_2$  subunits. Applying this procedure to the 1 024 000 snapshots (8000 snapshots  $\times$  128 MD), we prepared  $\sim 37$  million COMs in the vicinity of the  $\beta$  subunits (Figure 1a).

However, this vast number of COMs makes it impossible to execute clustering with the Ward's method because of the huge computational cost and insufficient memory to store the distance matrix. Hence, by sampling the COMs at seven MD time points (5.0, 5.5, 6.0, 6.5, 7.0, 7.5, and 8.0 ns), we obtained 31 454 representative COMs and executed clustering with the Ward's method (Figure 1b). We set the number of clusters to group COMs as 65 because this is the smallest number with which the COM distribution in the heme binding site is isolated as a single cluster. Finally, all of the COMs were grouped by assigning their group to that of the nearest representative COM (Figure 1c). The same clustering procedure was also done in the  $\alpha$  subunit with  $\sim 32$  million COMs, 27 703 representative COMs, and the number of cluster 65. The cluster corresponding to the heme binding site is cluster 40 in the  $\alpha$  subunit and cluster 64 in the  $\beta$  subunit. All of the coordinates of cluster centers in the  $\alpha$  and  $\beta$  subunits are given as PDB format in the Supporting Information.

**Procedure of the Intrinsic Pathway Identification by Clustering (IPIC) Method.** We developed the IPIC method in which the clustering results are used to analyze and visualize the O<sub>2</sub> entry pathways inside Hb subunits. For convenience, in addition to the 65 cluster indices obtained in the above section (cluster 1–65), we define “cluster 0” as O<sub>2</sub> molecules whose shortest distance between its COM and non-hydrogen atoms of the subunit is larger than 4.0 Å. Thus the position of each O<sub>2</sub> molecule ( $x(t)$ ,  $y(t)$ ,  $z(t)$ ) can be described by the cluster indices  $C = \{0, 1, 2, \dots, 65\}$  as

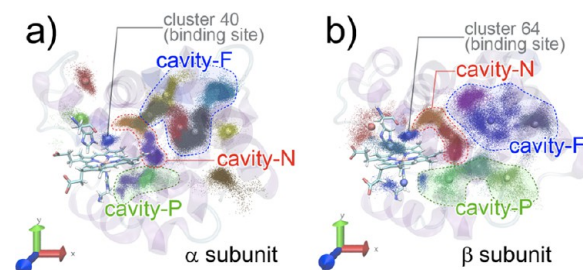
$$(x(t), y(t), z(t)) \rightarrow c(t) \in C \quad (1)$$

We classified the O<sub>2</sub> distribution into three groups: inside, surface, and outside. When the cluster of the O<sub>2</sub> is  $c = 0$ , the O<sub>2</sub> is outside. To distinguish between inside and surface clusters, we calculated transition frequencies of O<sub>2</sub> from cluster  $a$  to  $b$  ( $F_{ab}$ ) and transition probabilities from cluster  $a$  to  $b$  ( $p_{ab}$ ) from the 120 O<sub>2</sub> molecules, 128 MD trajectories, and 8000 MD snapshots per MD (1 snapshot/ps) as

$$F_{ab} = \sum_{n=1}^{120} \sum_{m=1}^{128} \sum_{t=0}^{7999} \begin{cases} 1 & c_m^n(t) = a \text{ and } c_m^n(t+1) = b \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$p_{ab} = \frac{F_{ab}}{\sum_{i \in C, i \neq b} F_{ib}} \quad (3)$$

where  $c_m^n(t)$  is the cluster index of the  $n$ th O<sub>2</sub> molecule at time  $t$  in the  $m$ th MD trajectory. When  $p_{0i} > 0.003$  (i.e., direct transition from outside occurs more than 0.3%), cluster  $i$  is defined as a surface cluster, and when  $p_{0i} < 0.003$ , cluster  $i$  is inside cluster. The locations of O<sub>2</sub> classified as inside clusters are shown in Figure 2.



**Figure 2.** COMs of O<sub>2</sub> molecules classified as inside clusters in the (a)  $\alpha$  subunit and (b)  $\beta$  subunit. The COMs and centers of clusters are drawn with dots and spheres, respectively, using different colors for each cluster.

We assume that O<sub>2</sub> molecules enter the binding site when they stay in the binding site cluster ( $c^{\text{bs}} = 40$  for the  $\alpha$  subunit and  $c^{\text{bs}} = 64$  for the  $\beta$  subunit) for five consecutive MD snapshots, or 5 ps. Then, the O<sub>2</sub> entry trajectories from solvent to the binding site ( $T_n$ ) can be retrieved starting from a surface cluster ( $c^{\text{surf}}$ ) at time  $t_1$  through inside clusters ( $c^{\text{inside}}$ ) to the binding site cluster ( $c^{\text{bs}}$ ) at time  $t_2$  as

$$T_n = \{c^{\text{surf}}(t_1), c^{\text{inside}}(t_1 + 1), c^{\text{inside}}(t_1 + 2), \dots, c^{\text{inside}}(t_2 - 1), c^{\text{bs}}(t_2), c^{\text{bs}}(t_2 + 1), \dots, c^{\text{bs}}(t_2 + 4)\} \quad (4)$$

In total, we retrieved 141 and 425 entry trajectories ( $\{T_1^\alpha, T_2^\alpha, \dots, T_{141}^\alpha\}$  and  $\{T_1^\beta, T_2^\beta, \dots, T_{425}^\beta\}$ ) for  $\alpha$  and  $\beta$  subunits, respectively. Using graph theory terminology,<sup>45</sup> we constructed a graph  $\{V, E\}$  with nodes (or vertices)  $V$  and edges  $E$  for both subunits as



$$V_c^\alpha = \sum_{n=1}^{141} \begin{cases} 1 & c \in T_n^\alpha \\ 0 & c \notin T_n^\alpha \end{cases} \quad (5)$$

$$E_{c_1 c_2}^\alpha = \sum_{n=1}^{141} \begin{cases} 1 & c_1 \rightarrow c_2 \in T_n^\alpha \\ 0 & c_1 \rightarrow c_2 \notin T_n^\alpha \end{cases} \quad (6)$$

$$V_c^\beta = \sum_{n=1}^{425} \begin{cases} 1 & c \in T_n^\beta \\ 0 & c \notin T_n^\beta \end{cases} \quad (7)$$

$$E_{c_1 c_2}^\beta = \sum_{n=1}^{425} \begin{cases} 1 & c_1 \rightarrow c_2 \in T_n^\beta \\ 0 & c_1 \rightarrow c_2 \notin T_n^\beta \end{cases} \quad (8)$$

### 3. RESULTS AND DISCUSSION

**Cavity Classification into Cavity-N, Cavity-P, and Cavity-F.** In this work, we executed 128 independent 8 ns MD simulations starting from the T-state deoxy HbA tetramer under O<sub>2</sub>-rich conditions. Because of the high O<sub>2</sub> concentration and total 1  $\mu$ s long MD trajectory, many O<sub>2</sub> entry events into the binding site occurred. To analyze complex entry pathways, we performed clustering of the O<sub>2</sub> COM distributions in HbA subunits and classified them into inside, surface, and outside. In Figure 2, centers of inside clusters and COMs classified as inside clusters are shown as spheres and points, respectively. The binding site is well-isolated as cluster 40 and 64 in the  $\alpha$  and  $\beta$  subunits, respectively. In both subunits, the location of the binding site cluster is consistent with the experimentally observed location above the heme pyrrole ring C, where CO density is observed after photodissociation of CO-liganded Hb at cryogenic temperature 25 K.<sup>7</sup>

The O<sub>2</sub> COM distribution of the inside clusters reflects cavity patterns inside HbA subunits. In particular, in the  $\beta$  subunit, the O<sub>2</sub> distribution (Figure 2b) can be classified into three groups. The first one is in “cavity-N” near the binding site, the second one is in “cavity-P” near the proximal histidine, and the third one is in “cavity-F” far from the heme. Although the boundaries between these three cavities in the  $\alpha$  subunit are more ill-defined than those of  $\beta$  subunit, we also classified the ligand distribution of the  $\alpha$  subunit into the same three groups for convenience (Figure 2a).

**Faster O<sub>2</sub> Entry into Binding Site in  $\beta$  Subunit.** To determine how often O<sub>2</sub> molecules reached the binding site during the 8 ns MD trajectories, we checked the probability of O<sub>2</sub> existence in the binding site cluster (Figure 3). In the tetramer HbA, there are two  $\alpha$  and two  $\beta$  subunits, and thus we have 256 samples from the 128 MD trajectories to calculate the probabilities.

Until 1.0 and 1.5 ns in the  $\alpha$  and  $\beta$  subunits, respectively, the probabilities were very small. After that, O<sub>2</sub> molecules began to enter the binding site and the probabilities continued to increase linearly during the 8 ns simulations. The slope of the regression line of the  $\beta$  subunit curve (4.3%/ns) is about two times larger than that of  $\alpha$  subunit (2.3%/ns), and the final probabilities at 8 ns are 46/256 = 18.0% and 83/256 = 32.4% in the  $\alpha$  and  $\beta$  subunits, respectively. This faster O<sub>2</sub> entry of the  $\beta$  subunit is consistent with the experimental observations; the rates of CO entry and NO binding in the  $\beta$  subunit are faster than those of the  $\alpha$  subunit.<sup>16</sup>

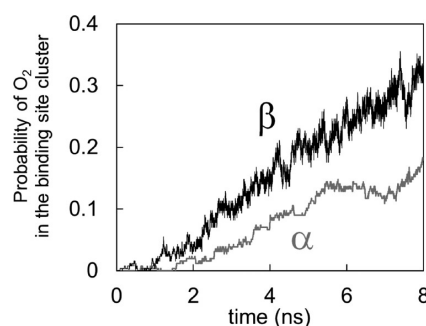


Figure 3. Probability of O<sub>2</sub> in the binding site in the  $\alpha$  and  $\beta$  subunits.

**Multiple O<sub>2</sub> Entry Pathways in HbA Subunits.** To quantitatively analyze and visualize the multiple O<sub>2</sub> entry pathways from solvent to the heme binding site, we applied the IPIC method and retrieved 141 and 425 entry trajectories  $T_n$  (eq 4) for  $\alpha$  and  $\beta$  subunits from surface clusters into the binding site cluster and constructed graphs with nodes and edges  $\{V, E\}$  for both subunits (eqs 5–8). From the definition of  $T_n$ , where only one surface cluster is included as a starting point, surface clusters are interpreted as O<sub>2</sub> portals to enter the inside cavities. To see the portal distributions in the  $\alpha$  and  $\beta$  subunits,  $V_c$  of surface clusters are listed in Tables 1 and 2,

Table 1. O<sub>2</sub> Entry Frequencies and Probabilities of Each Portal Cluster in the  $\alpha$  Subunit

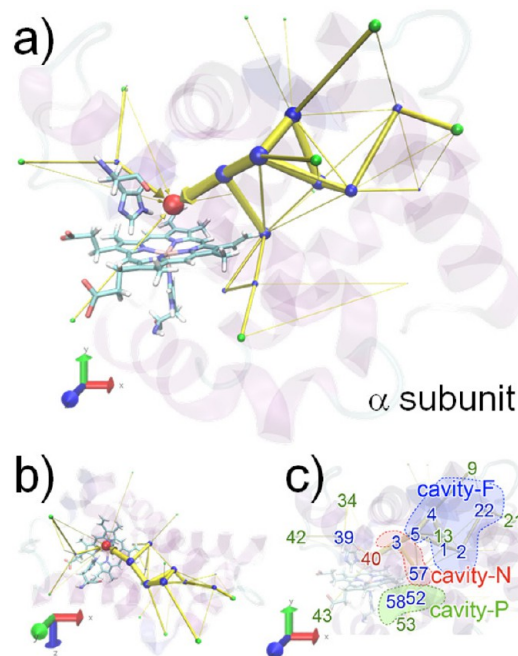
pathway	portal cluster $c$	frequency $V_c$	probability (%)
	21	33	23.4%
pathway-F	9	28	19.9%
cavity-F	13	21	14.9%
→cavity-N	6	1	0.7%
→binding site	7	1	0.7%
	sum	84	59.6%
pathway-C	34	18	12.8%
CD-corner	42	9	6.4%
→binding site	31	4	2.8%
	sum	31	22.0%
pathway-P	53	17	12.1%
cavity-P			
→cavity-N			
→binding site			
His-gate	43	7	5.0%
pathway-N	28	2	1.4%
cavity-N			
→binding site			
	total sum	141	100.0%

respectively. In Figures 4 and 5, the entry pathways in the  $\alpha$  and  $\beta$  subunits are visualized according to the graphs, respectively. Each node  $V_c$  is drawn as a sphere at the center of cluster  $c$ , whose radius is proportional to  $V_c^{0.75}$ . The sphere colors are red for the binding site cluster, blue for the inside clusters, and green for the surface clusters. Each edge  $E_{c_1 c_2}$  is drawn as a yellow cylinder connecting two centers of clusters  $c_1$  and  $c_2$ , whose radius is proportional to  $E_{c_1 c_2}^{0.75}$ .

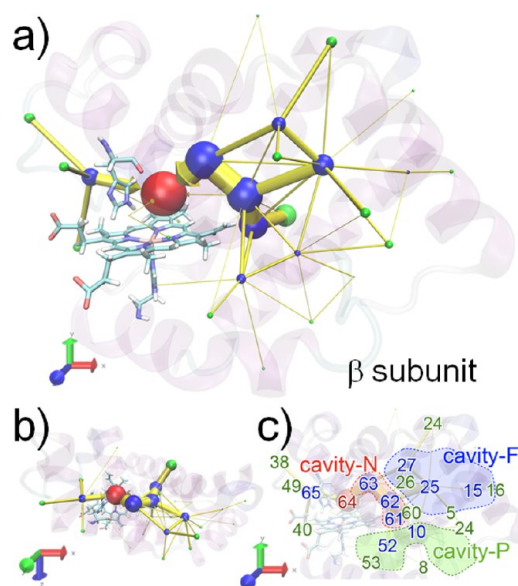
In the  $\alpha$  subunit, the main entry pathway starts at portals far from the binding site: portal clusters 21, 9, and 13 with probabilities 23.4, 19.9, and 14.9%, respectively. In this “pathway-F”, O<sub>2</sub> molecules first enter the cavity-F and then shuttle between cavity-F and cavity-N and finally reach the binding site through cluster 3. In Figure 6a,b, a typical entry

**Table 2.** O<sub>2</sub> Entry Frequencies and Probabilities of Each Portal Cluster in the  $\beta$  Subunit

pathway	portal cluster $c$	frequency $V_c$	probability (%)
pathway-N	60	183	43.1%
cavity-N	22	2	0.5%
→binding site	sum	185	43.5%
	40	44	10.4%
pathway-C	38	41	9.6%
CD-corner	49	28	6.6%
→binding site	36	4	0.9%
	sum	117	27.5%
	24	36	8.5%
	5	25	5.9%
pathway-F	4	20	4.7%
cavity-F	26	18	4.2%
→cavity-N	16	6	1.4%
→binding site	17	1	0.2%
	sum	106	24.9%
pathway-P	53	8	1.9%
cavity-P	8	6	1.4%
→cavity-N	9	2	0.5%
→binding site	sum	16	3.8%
His-gate	54	1	0.2%
	total sum	425	100.0%

**Figure 4.** Visualized O<sub>2</sub> entry pathways from solvent to the binding site in the  $\alpha$  subunit. The red sphere is the center of the binding site cluster, blue spheres are centers of inside clusters, and green spheres are those of surface clusters or O<sub>2</sub> entry portals. Yellow cylinders represent ligand migration between two clusters. Radii of spheres and cylinders indicate the frequencies of O<sub>2</sub> migration event. (a) Side view. (b) Top view. (c) Side view with the indices of major clusters.

trajectory of pathway-F ( $T_F^\alpha$ ) starting from the portal cluster 21 is shown. The O<sub>2</sub> molecule entered inside cluster 22 through the portal 21 at 2.222 ns and stayed in the large cavity composed of clusters 22, 2, 1, 4, and 5 (cavity-F). At 4.082 ns, it migrated to the adjacent cavity-N composed of clusters 5, 3, and 57. (Cluster 5 is shared by cavity-F and cavity-N.) Then, the O<sub>2</sub> shuttled between the cavity-F and cavity-N and finally

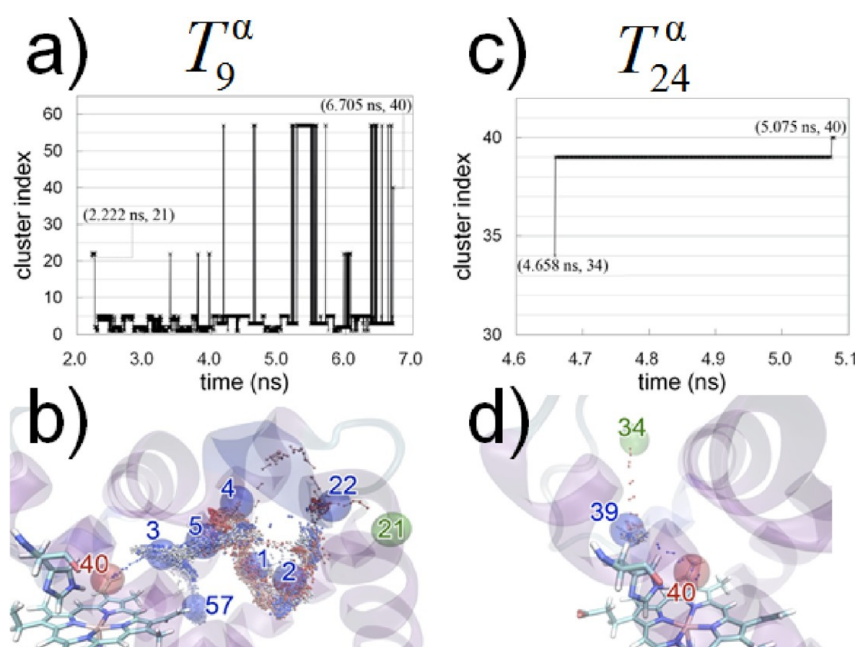
**Figure 5.** Visualized O<sub>2</sub> entry pathways from solvent to the binding site in the  $\beta$  subunit. (a) Side view. (b) Top view. (c) Side view with the indices of major clusters. Visualization details are identical to Figure 4.

reached the binding site cluster at 6.705 ns through cluster 3. Other entry trajectories starting at the main portal clusters 21, 9, and 13, including other minor portal clusters 6 and 7 (the probabilities are both 0.7%), are similar to this trajectory  $T_F^\alpha$ . Through this pathway-F, 59.6% O<sub>2</sub> molecules reached the binding site (Table 1).

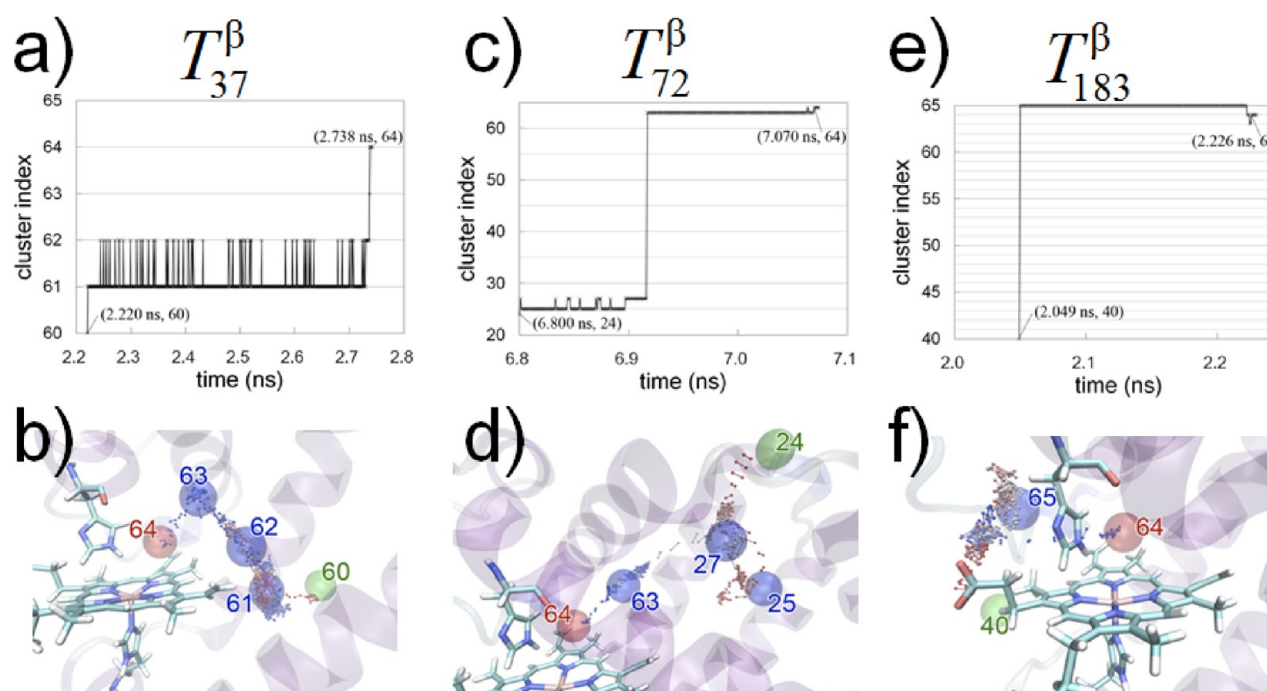
The second important pathway of the  $\alpha$  subunit is “pathway-C”, which starts around the CD-loop, and the portal clusters are 34, 42, and 31. In Figure 6c,d, a typical migration trajectory ( $T_C^\alpha$ ) from the portal at cluster 34, is shown. The O<sub>2</sub> molecule entered inside cluster 39 through the portal cluster 34 at 4.658 ns. Then, it reached the binding site at 5.075 ns. Through this pathway-C, 22.0% O<sub>2</sub> molecules reached the binding site (Table 1).

Other pathways are minor (Table 1). From surface cluster 53, O<sub>2</sub> molecules first enter cavity-P, which is near the proximal histidine and is composed of clusters 52 and 58. Then, the molecules transfer to cavity-N and finally reach the binding site. Through this “pathway-P”, 12.1% O<sub>2</sub> molecules reached the binding site. Furthermore, there are two other minor pathways, denoted as “His-gate” and “pathway-N”. In His-gate, ligands directly reach the binding site from portal cluster 43. It should be noted that in both  $\alpha$  and  $\beta$  subunits no gate opening by the conformational change of the distal histidine side chain was observed during the production MD simulations. In pathway-N, ligands first enter cavity-N and reach the binding site. The probabilities of these two pathways are small, 5.0 and 1.4% (Table 1).

In the  $\beta$  subunit, the entry pathways can be described similarly as in the  $\alpha$  subunit using the significant cavities, cavity-N, cavity-F, and cavity-P. Cavity-N is near the binding site and is composed of clusters 63, 62, and 61. Cavity-F is far from the binding site and is composed of clusters 27, 25, and 15. However, the probability through each pathway is different from that in the  $\alpha$  subunit. Although the main pathway in the  $\alpha$  subunit is pathway-F with probability 59.6% (Table 1), pathway-N is more significant than pathway-F in the  $\beta$  subunit (Table 2). In Figure 7a,b, a typical trajectory of pathway-N



**Figure 6.** Typical O<sub>2</sub> entry trajectories from solvent to the binding site in the  $\alpha$  subunit. (a,c) Cluster index changes of the migrating O<sub>2</sub> molecule in trajectory (a)  $T_9^\alpha$  through pathway-F and (c)  $T_{24}^\alpha$  through pathway-C. (b,d) Visualizations of trajectory (b)  $T_9^\alpha$  through pathway-F and (d)  $T_{24}^\alpha$  through pathway-C. The red sphere is the binding site, the blue spheres are inside clusters, and the green sphere is the surface cluster (portal) where the ligand first entered the  $\alpha$  subunit interior. The trace of the migrating O<sub>2</sub> is drawn as a small ball and stick model after time-window averaging over five frames (5 ps) to facilitate visualization. The color of O<sub>2</sub> represents MD simulation time with gradation from red to blue.



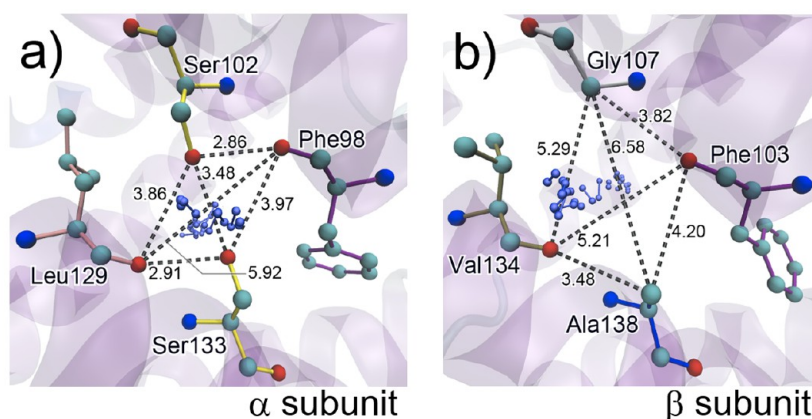
**Figure 7.** Typical O<sub>2</sub> entry trajectories from solvent to the binding site in the  $\beta$  subunit. (a,c,e) Cluster index changes of the migrating O<sub>2</sub> molecule in trajectory: (a)  $T_{37}^\beta$  through pathway-N, (c)  $T_{72}^\beta$  through pathway-F, and (e)  $T_{183}^\beta$  through pathway-C. (b,d,f) Visualization of trajectory (b)  $T_{37}^\beta$  through pathway-N, (d)  $T_{72}^\beta$  through pathway-F, and (f)  $T_{183}^\beta$  through pathway-C. Drawing details are identical to Figure 6b,d.

starting from portal cluster 60 ( $T_{37}^\beta$ ) is shown. The O<sub>2</sub> molecule directly entered cavity-N through portal cluster 60 at 2.220 ns and finally reached the binding site cluster 64 through cluster 63 at 2.738 ns. Through this pathway-N, 43.5% ligands reached the binding site (Table 2). Several O<sub>2</sub> molecules first entered cavity-F or through pathway-F from portal clusters 24, 5, 4, 26, 16, and 17. A typical trajectory  $T_{72}^\beta$  is shown in Figure 7c,d. The

probability through pathway-F is 24.9%, which is about half of that of pathway-N.

Pathway-C in the  $\beta$  subunit is as important as it is in the  $\alpha$  subunit. In Figure 7e,f, a typical trajectory through pathway-C ( $T_{183}^\beta$ ) is shown. The O<sub>2</sub> molecule first entered the cavity near the CD-corner (cluster 65) from portal cluster 40 and reached the binding site. The probability of this pathway-C in the  $\beta$





**Figure 8.** Residues composing the portal from solvent into cavity-N in the (a)  $\alpha$  and (b)  $\beta$  subunits. The colors of atoms are green, blue, and red for carbon, nitrogen, and oxygen atoms, respectively. A typical  $O_2$  entry trajectory is drawn with a blue ball-and-stick model after time-window averaging over five frames (5 ps) to facilitate visualization.

subunit was 27.5%, similar to that in the  $\alpha$  subunit 22.0%.  $O_2$  entry through pathway-P was also seen in the  $\beta$  subunit with a smaller probability 3.8%.

**Finding All Significant Entry Pathways.** As we have discussed in our previous ensemble MD work,<sup>31</sup> the amount of statistics, or the number of MD trajectories, should be determined according to the target precision. Here we briefly discuss the amount of statistics used in this multiple entry pathway analysis. Suppose that the true probability of  $O_2$  entry from a portal  $e$  is  $p_e$ , the probability of observing no  $O_2$  entry from the portal  $e$  in the independent  $N$   $O_2$  entry trajectories is  $(1 - p_e)^N$ , and that of observing at least one ligand entry is  $P_e^o \equiv 1 - (1 - p_e)^N$ . Thus the portal probability  $p_e$  whose  $P_e^o$  is more than 95% can be calculated as  $P_e^o = 1 - (1 - p_e)^N > 0.95$ , yielding  $p_e > 1 - (0.05)^{1/N}$  (or  $N > \ln 0.05 / \ln(1 - p_e)$ ). This means that portals whose  $p_e > 1 - (0.05)^{1/N}$  are observed with more than 95% probability in the  $N$  entry trajectories. We determined the target precision  $p_e = 3\%$  to obtain all of the significant entry portals and the required number of entry trajectories is  $N > 98.4$ .

In this work, each MD simulation can be calculated independently. Because the available computer resources for this work included 64 CPU cores (8 machines with 8 CPU cores), 64 MD simulations were calculated at the same time using the single-core version PMEMD MD solver program. We first executed 8 ns 64 MD simulations and obtained 62 and 203  $O_2$  entry trajectories in the  $\alpha$  and  $\beta$  subunits, respectively. In the  $\alpha$  subunit, the number of entry trajectories 62 was smaller than the required number  $N > 98.4$ , and thus we added 64 MD simulations. Finally, from the total 128 MD simulations, we obtained 141 and 425 entry trajectories, which is larger than the required  $N > 98.4$  in both subunits. As a result, portals whose  $p_e > 2.1\%$  and  $> 0.7\%$  were obtained with more than 95% probability in the  $\alpha$  and  $\beta$  subunits, respectively. Therefore, we can statistically conclude that with the size of current MD simulation data (8 ns 128 MD simulations) our entry pathway analysis could detect all of the significant  $O_2$  portals.

**Widely Open Portal in the  $\beta$  Subunit Causes Faster  $O_2$  Entry.** The present result with the IPIC method visualization (Figures 4 and 5) revealed that although the folded backbone structures of  $\alpha$  and  $\beta$  subunits are very similar, the main  $O_2$  entry pathway from outer solvent to the binding site is distinctly different between  $\alpha$  and  $\beta$  subunits: pathway-F is the main pathway in the  $\alpha$  subunit and pathway-N is that in the  $\beta$

subunit. The pathway difference is consistent with the previous computational analysis by implicit ligand sampling calculation, which revealed very diverse ligand migration pathways in globular proteins with very strong conservation of their secondary structure globin fold.<sup>29</sup>

To understand this difference at the atomic level, we focus on the portal through which  $O_2$  enters cavity-N. Figure 8 shows the structures of the portals between clusters 28 and 57 in the  $\alpha$  subunit and that between clusters 60 and 61 in the  $\beta$  subunit in the crystal structure. The portals are actually composed of four atoms in both subunits. By considering the vdW radius and the distances among gate atoms, we can determine whether the portal can be easily opened or not.

In the  $\alpha$  subunit (Figure 8a), the portal is composed of the side-chain oxygen atoms of Ser102 and Ser133 (in the forcefield used in this MD simulation, vdW radius is 1.721 Å) and the backbone oxygen atoms of Leu129 and Phe98 (1.6612 Å). The shorter diagonal of the quadrilateral formed by the four gate atoms represents the width of the gate. The shorter diagonal is between Ser oxygen atoms and is 3.48 Å. This is almost the same as the sum of the two oxygen atoms radius  $1.721 \times 2 = 3.442$  Å, and thus the gate is completely closed. Therefore, when an  $O_2$  ligand with vdW radius 1.6 Å goes through the gate, a large structural fluctuation with almost 1.6 Å is necessary. This large fluctuation rarely occurs, and the probability of  $O_2$  entry through this portal is very small: 1.4% in the  $\alpha$  subunit (Table 1).

In the  $\beta$  subunit (Figure 8b), the gate is composed of the side-chain carbon atom of Ala138 (1.908 Å), the backbone carbon atom of Gly107 (1.908 Å), and the backbone oxygen atoms of Phe103 and Val134 (1.6612 Å). The shorter diagonal is between Phe103 and Val134 oxygen atoms and is 5.21 Å. Thus the gate width is  $5.21 - 1.6612 \times 2 = 1.8876$  Å. This is wide enough for  $O_2$  ligands, resulting in the very frequent  $O_2$  entry through this gate 43.1% in the  $\beta$  subunit (Table 2). It should be noted that in the crystal structure this gate is blocked by a crystal water molecule that is located between the two oxygen atoms, and thus this large portal would not be detected by a static cavity analysis with the existence of the fixed crystal water molecules. This open portal should contribute to the faster  $O_2$  entry into the  $\beta$  subunit binding site (Figure 2).

**Correspondence of Calculated and Experimental  $O_2$  Entry Rates Suggests Multiple  $O_2$  Entry Pathways in HbA.** The current ensemble MD approach straightforwardly

Table 3. Rate Constants of O<sub>2</sub> Entry ( $k_{\text{entry}}$ ) and NO Binding ( $k_{\text{NO}}$ ) for Each Subunit and the Ratio of  $k^{\beta}/k^{\alpha}$ 

		HbA state	$\alpha$ ( $\mu\text{mol/L}$ ) <sup>-1</sup> s <sup>-1</sup>	$\beta$ ( $\mu\text{mol/L}$ ) <sup>-1</sup> s <sup>-1</sup>	$k^{\beta}/k^{\alpha}$ ratio
this work at 310 K	$k_{\text{entry}}$	T-state tetramer	45	99	2.20
from ref 16 after temperature correction	$k_{\text{entry}}$	isolated subunit	69	81	1.17
		isolated subunit	60	131	2.18
	$k_{\text{NO}}$	R-state tetramer	48	127	2.65

<sup>a</sup>To correct temperature difference between our 310 K simulations and the 293 K experiments,<sup>16</sup> we corrected the experimental values by multiplying by 1.92, assuming the activation energy of O<sub>2</sub> entry 6.9 kcal/mol.<sup>46</sup>

simulated the O<sub>2</sub> entry processes, and thus we can estimate the rate constants of O<sub>2</sub> entry into the binding site ( $k_{\text{entry}}^{\alpha}$  and  $k_{\text{entry}}^{\beta}$ ) and compare with the experimentally reported values. We assume a simple model of O<sub>2</sub> entry into the binding site

ignoring the escape process as  $\text{Hb}^x + \text{O}_2 \xrightarrow{k_{\text{entry}}^x} \text{Hb}^x\text{:O}_2$ , where  $x$  denotes the subunit ( $x = \alpha$  or  $\beta$ ) and  $\text{Hb}^x\text{:O}_2$  is the situation of O<sub>2</sub> in the binding site of subunit  $x$ , and the rate of O<sub>2</sub> entry is  $k_{\text{entry}}^x[\text{O}_2][\text{Hb}^x]$ . Then, the probability of O<sub>2</sub> in the binding site can be calculated as  $1 - \exp(-k_{\text{entry}}^x[\text{O}_2]t)$ . The probabilities obtained from our current MD simulations at  $t = 8$  ns with  $[\text{O}_2] = 0.55$  mol/L ns are 0.180 and 0.352 for  $\alpha$  and  $\beta$  subunits, respectively, resulting in  $k_{\text{entry}}^{\alpha} = 45$  and  $k_{\text{entry}}^{\beta} = 99$  ( $\mu\text{mol/L}$ )<sup>-1</sup> s<sup>-1</sup> (Table 3).

Birukou et al. experimentally measured the ligand entry rate constants ( $k_{\text{entry}}$ ) at 20 °C (293 K) into the isolated HbA subunits and, in addition, NO binding rate constants ( $k_{\text{NO}}$ ) to the isolated HbA subunits and R-state tetramer HbA.<sup>16</sup> Here  $k_{\text{NO}}$  can be considered as the NO entry rate constant because when NO enters the binding site the NO immediately binds to the heme iron atom.<sup>14</sup> To compare the experimental  $k_{\text{entry}}$  and  $k_{\text{NO}}$  with the current MD simulation values, we should correct temperature difference between our 310 K and their 293 K. Under the assumption that the Arrhenius equation can be applied in the O<sub>2</sub> entrance process, rate constants are described as  $k = A \exp(-E_a/RT)$ , where  $A$  is the pre-exponential factor,  $E_a$  is the activation energy, and  $R$  is the gas constant. Then, the temperature dependency can be obtained as

$$\ln \frac{k(T')}{k(T)} = \frac{E_a}{R} \left( \frac{1}{T} - \frac{1}{T'} \right)$$

From this formula with the experimentally estimated activation energy of O<sub>2</sub> entry  $E_a = 6.9$  kcal/mol,<sup>46</sup> rate constants at 310 K are faster than those at 293 K by a factor of 1.92. In Table 3, the experimentally observed rate constants are shown after multiplying 1.92 to correct temperature difference between 293 and 310 K. For example, the reported  $k_{\text{entry}}$  of the  $\alpha$  subunit,  $36$  ( $\mu\text{mol/L}$ )<sup>-1</sup> s<sup>-1</sup>, is corrected to be  $36 \times 1.92 = 69$  ( $\mu\text{mol/L}$ )<sup>-1</sup> s<sup>-1</sup> in Table 3. Although the quaternary structure condition is different between our T-state HbA and the R-state HbA or isolated subunits, the  $k_{\text{entry}}$  calculated from the current MD simulation corresponds with the experimental values. In addition, the  $k^{\beta}/k^{\alpha}$  ratio 2.20 is also comparable to the experimentally reported values 1.17 to 2.65.

There is some debate over the ligand migration pathways in Mb and HbA. Several experimental works revealed that the His gate is the main pathway,<sup>14,16–18</sup> but computational works revealed multiple migration pathways.<sup>19–29</sup> Here the good correspondence of the  $k_{\text{entry}}$  and  $k^{\beta}/k^{\alpha}$  ratio suggests that the ligand entry into the binding site mainly occurs through the multiple pathways, not through the His gate, because almost all of the O<sub>2</sub> molecules enter the binding site through multiple pathways in the current MD simulations (Table 1, 2).

Moreover, the faster O<sub>2</sub> entry in the  $\beta$  subunit can be explained by the wide portal into the cavity-N (Figure 8).

Another interesting result of this work is that it was necessary to remove the crystal water molecule in the  $\alpha$  subunit binding site to obtain the  $k^{\beta}/k^{\alpha}$  ratio corresponding to the experimental values. According to our preliminary MD simulations with the crystal water (only one O<sub>2</sub> entry into the  $\alpha$  subunit binding site occurred in 5 ns 32 MD trajectories), it is estimated that  $k^{\beta}/k^{\alpha}$  ratio would be  $>10$  and far larger than the experimental value. This suggests that, at least in the CO geminate recombination experiments by Birukou et al.,<sup>16</sup> the crystal water molecule does not strongly block the ligand entry. One of the possible explanations for this counterintuitive result is that the rate of crystal water entry is slow after CO photolysis and the  $\alpha$  subunit binding site is vacant during the  $\sim 100$   $\mu\text{s}$  CO geminate recombination experiments. Another possible explanation is that the crystal water frequently enters and escapes the binding site faster than the ligand entry process and hence hydrophobic ligand molecules can easily switch the water molecule. In fact, in Mb, the experimentally observed entry and escape rate constants of water in the binding site ( $k_{\text{in}}^{\text{H}_2\text{O}}[\text{H}_2\text{O}] = 9.0$   $\mu\text{s}^{-1}$  and  $k_{\text{out}}^{\text{H}_2\text{O}} = 1.7$   $\mu\text{s}^{-1}$ )<sup>47</sup> are larger than the O<sub>2</sub> entry rate even under the O<sub>2</sub> saturating condition ( $k_{\text{entry}}[\text{O}_2] = 0.045$   $\mu\text{s}^{-1}$  with  $k_{\text{entry}} = 34$  ( $\mu\text{mol/L}$ )<sup>-1</sup> s<sup>-1</sup><sup>14</sup> and  $[\text{O}_2] = 1.25$  mmol/L). Therefore, we anticipate that the latter explanation would be also plausible for HbA  $\alpha$  subunit.

#### 4. CONCLUSIONS

In this study, the multiple O<sub>2</sub> entry pathways from solvent to the binding site in the T-state HbA were quantitatively elucidated by using the ensemble MD approach with 128 independent 8 ns MD simulations in O<sub>2</sub>-rich aqueous solvent. According to the distribution of the O<sub>2</sub> locations from the total of 1.024  $\mu\text{s}$  trajectories, we obtained the shapes of the cavities inside HbA subunits and classified them into three groups: cavity-F, cavity-N, and cavity-P. Then, we obtained 141 and 425 O<sub>2</sub> entry trajectories from solvent to the binding site in the  $\alpha$  and  $\beta$  subunits, respectively. The numbers of entry trajectories were large enough to detect all of the significant O<sub>2</sub> entry portals, and we can quantitatively compare each portal. We also developed the IPIC method to realize persuasive visualization with which one can comprehend both the shapes and the relative importance of the multiple pathways. In the  $\alpha$  subunit, the main portal is pathway-F (from cavity-F through cavity-N to the binding site) and the second one is pathway-C (from cavity around the CD-corner to the binding site). Meanwhile, in the  $\beta$  subunit, the main portal is pathway-N (from cavity-N to the binding site) and the second one is pathway-C. The  $\beta$  subunit main portal is widely open compared with that of  $\alpha$  subunit, resulting in the faster O<sub>2</sub> entry in the  $\beta$  subunit. The rate constants of O<sub>2</sub> entry ( $k_{\text{entry}}$ ) estimated from the current MD simulations correspond to the experimentally observed values, suggesting that O<sub>2</sub> molecules enter the binding site



through multiple pathways, not through the His gate. The obtained multiple O<sub>2</sub> migration pathway map can be utilized for future detailed analysis.

The current procedure with the ensemble MD calculations and IPIC method is very useful to comprehensively analyze multiple and complex ligand migration pathways. In particular, the essential point in this procedure is the classification of ligand distributions by clustering, which is a fundamental method in data mining. Manual classification is often difficult because of the complex shapes of protein cavities and ill-defined cavity boundaries, and thus the manual classification is a complicated task, and the results would be analyst-dependent. Meanwhile, as shown in this work, efficient classification of ligand distributions can be achieved automatically by clustering with many ligand location data obtained by the ensemble MD calculations. Therefore, we believe that the current procedure can be a standard procedure to analyze complex ligand migration processes in large biomacromolecules composed of many subunits.

## ■ ASSOCIATED CONTENT

### ■ Supporting Information

Complementary discussion describing the contribution of the diffusion and migration processes to the O<sub>2</sub> entry rate and comparison between the IPIC method with previously applied methods is given. All coordinates of cluster centers are given as PDB format file with the structure of each subunit. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Tel/Fax: +81-52-789-5623. E-mail: [mnagaoka@is.nagoya-u.ac.jp](mailto:mnagaoka@is.nagoya-u.ac.jp).

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was partially supported by a Grant-in-Aid for Science Research from the Ministry of Education, Culture, Sport, Science and Technology in Japan and the Core Research for Evolutional Science and Technology (CREST) "High Performance Computing for Multi-scale and Multi-physics Phenomena" from the Japan Science and Technology Agency (JST).

## ■ REFERENCES

- (1) Bellelli, A. Hemoglobin and Cooperativity: Experiments and Theories. *Curr. Protein Pept. Sci.* **2010**, *11*, 2–36.
- (2) Savino, C.; Miele, A. E.; Draghi, F.; Johnson, K. A.; Sciara, G.; Brunori, M.; Vallone, B. Pattern of Cavities in Globins: the Case of Human Hemoglobin. *Biopolymers* **2009**, *91*, 1097–1107.
- (3) Park, S.-Y.; Yokoyama, T.; Shibayama, N.; Shiro, Y.; Tame, J. R. H. 1.25 Å Resolution Crystal Structures of Human Haemoglobin in the Oxy, Deoxy and Carbonmonoxy Forms. *J. Mol. Biol.* **2006**, *360*, 690–701.
- (4) Schotte, F.; Lim, M.; Jackson, T. A.; Smirnov, A. V.; Soman, J.; Olson, J. S.; Phillips, G. N., Jr.; Wulff, M.; Anfinrud, P. A. Watching a Protein as it Functions with 150-ps Time-Resolved X-ray Crystallography. *Science* **2003**, *300*, 1944–1947.
- (5) Schotte, F.; Soman, J.; Olson, J. S.; Wulff, M.; Anfinrud, P. A. Picosecond Time-Resolved X-ray Crystallography: Probing Protein Function in Real Time. *J. Struct. Biol.* **2004**, *147*, 235–246.
- (6) Schmidt, M.; Nienhaus, K.; Pahl, R.; Krasselt, A.; Anderson, S.; Parak, F.; Nienhaus, G. U.; Srajer, V. Ligand Migration Pathway and Protein Dynamics in Myoglobin: A Time-Resolved Crystallographic Study on L29W MbCO. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11704–11709.
- (7) Adachi, S.; Park, S.-Y.; Tame, J. R. H.; Shiro, Y.; Shibayama, N. Direct Observation of Photolysis-Induced Tertiary Structural Changes in Hemoglobin. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7039–7044.
- (8) Tomita, A.; Sato, T.; Ichiyanagi, K.; Nozawa, S.; Ichikawa, H.; Chollet, M.; Kawai, F.; Park, S.-Y.; Tsuduki, T.; Yamato, T.; et al. Visualizing Breathing Motion of Internal Cavities in Concert with Ligand Migration in Myoglobin. *Proc. Natl. Acad. Sci.* **2009**, *106*, 2612–2616.
- (9) Tsuduki, T.; Tomita, A.; Koshihara, S.; Adachi, S.; Yamato, T. Ligand Migration in Myoglobin: A Combined Study of Computer Simulation and X-ray Crystallography. *J. Chem. Phys.* **2012**, *136*, 165101.
- (10) Tomita, A.; Kreutzer, U.; Adachi, S.; Koshihara, S.; Jue, T. 'It's Hollow': the Function of Pores within Myoglobin. *J. Exp. Biol.* **2010**, *213*, 2748–2754.
- (11) Nishihara, Y.; Sakakura, M.; Kimura, Y.; Terazima, M. The Escape Process of Carbon Monoxide from Myoglobin to Solution at Physiological Temperature. *J. Am. Chem. Soc.* **2004**, *126*, 11877–11888.
- (12) Perutz, M. F.; Mathews, F. S. An X-ray Study of Azide Methaemoglobin. *J. Mol. Biol.* **1966**, *21*, 199–202.
- (13) Takayanagi, M.; Iwahashi, C.; Nagaoka, M. Structural Dynamics of Clamshell Rotation during the Incipient Relaxation Process of Photodissociated Carbonmonoxy Myoglobin: Statistical Analysis by the Perturbation Ensemble Method. *J. Phys. Chem. B* **2010**, *114*, 12340–12348.
- (14) Scott, E. E.; Gibson, Q. H.; Olson, J. S. Mapping the Pathways for O<sub>2</sub> Entry Into and Exit from Myoglobin. *J. Biol. Chem.* **2001**, *276*, 5177–5188.
- (15) Salter, M. D.; Blouin, G. C.; Soman, J.; Singleton, E. W.; Dewilde, S.; Moens, L.; Pesce, A.; Nardini, M.; Bolognesi, M.; Olson, J. S. Determination of Ligand Pathways in Globins: Apolar Tunnels Versus Polar Gates. *J. Biol. Chem.* **2012**, *287*, 33163–33178.
- (16) Birukou, I.; Schweers, R. L.; Olson, J. S. Distal Histidine Stabilizes Bound O<sub>2</sub> and Acts as a Gate for Ligand Entry in Both Subunits of Adult Human Hemoglobin. *J. Biol. Chem.* **2010**, *285*, 8840–8854.
- (17) Birukou, I.; Soman, J.; Olson, J. S. Blocking the Gate to Ligand Entry in Human Hemoglobin. *J. Biol. Chem.* **2011**, *286*, 10515–10529.
- (18) Birukou, I.; Maillett, D. H.; Birukova, A.; Olson, J. S. Modulating Distal Cavities in the  $\alpha$  and  $\beta$  Subunits of Human HbA Reveals the Primary Ligand Migration Pathway. *Biochemistry* **2011**, *50*, 7361–7374.
- (19) Lucas, M. F.; Guallar, V. An Atomistic View on Human Hemoglobin Carbon Monoxide Migration Processes. *Biophys. J.* **2012**, *102*, 887–896.
- (20) Lepeshkevich, S. V.; Biziuk, S. A.; Lemeza, A. M.; Dzharagov, B. M. The Kinetics of Molecular Oxygen Migration in the Isolated  $\alpha$  Chains of Human Hemoglobin as Revealed by Molecular Dynamics Simulations and Laser Kinetic Spectroscopy. *Biochim. Biophys. Acta* **2011**, *1814*, 1279–1288.
- (21) Shadrina, M. S.; English, A. M.; Peslherbe, G. H. Effective Simulations of Gas Diffusion Through Kinetically Accessible Tunnels in Multisubunit Proteins: O<sub>2</sub> Pathways and Escape Routes in T-state Deoxyhemoglobin. *J. Am. Chem. Soc.* **2012**, *134*, 11177–11184.
- (22) Ruscio, J. Z.; Kumar, D.; Shukla, M.; Prisant, M. G.; Murali, T. M.; Onufriev, A. V. Atomic Level Computational Identification of Ligand Migration Pathways Between Solvent and Binding Site in Myoglobin. *Proc. Natl. Acad. Sci.* **2008**, *105*, 9204–9209.
- (23) Mouawad, L.; Maréchal, J.-D.; Perahia, D. Internal Cavities and Ligand Passageways in Human Hemoglobin Characterized by Molecular Dynamics Simulations. *Biochim. Biophys. Acta* **2005**, *1724*, 385–393.
- (24) Nishihara, Y.; Hayashi, S.; Kato, S. A Search for Ligand Diffusion Pathway in Myoglobin Using a Metadynamics Simulation. *Chem. Phys. Lett.* **2008**, *464*, 220–225.

- (25) Nishihara, Y.; Kato, S.; Hayashi, S. Protein Collective Motions Coupled to Ligand Migration in Myoglobin. *Biophys. J.* **2010**, *98*, 1649–1657.
- (26) Elber, R.; Gibson, Q. H. Toward Quantitative Simulations of Carbon Monoxide Escape Pathways in Myoglobin. *J. Phys. Chem. B* **2008**, *112*, 6147–6154.
- (27) Elber, R. Ligand Diffusion in Globins: Simulations Versus Experiment. *Curr. Opin. Struct. Biol.* **2010**, *20*, 162–167.
- (28) Cohen, J.; Arkhipov, A.; Braun, R.; Schulten, K. Imaging the Migration Pathways for O<sub>2</sub>, CO, NO, and Xe Inside Myoglobin. *Biophys. J.* **2006**, *91*, 1844–1857.
- (29) Cohen, J.; Schulten, K. O<sub>2</sub> Migration Pathways are not Conserved across Proteins of a Similar Fold. *Biophys. J.* **2007**, *93*, 3591–3600.
- (30) Takayanagi, M.; Okumura, H.; Nagaoka, M. Anisotropic Structural Relaxation and its Correlation with the Excess Energy Diffusion in the Incipient Process of Photodissociated MbCO: High-Resolution Analysis via Ensemble Perturbation Method. *J. Phys. Chem. B* **2007**, *111*, 864–869.
- (31) Takayanagi, M.; Nagaoka, M. Incipient Structural and Vibrational Relaxation Process of Photolyzed Carbonmonoxy Myoglobin: Statistical Analysis by Perturbation Ensemble Molecular Dynamics Method. *Theor. Chem. Acc.* **2011**, *130*, 1115–1129.
- (32) Sagnella, D.; Straub, J. Directed Energy “Funneling” Mechanism for Heme Cooling Following Ligand Photolysis or Direct Excitation in Solvated Carbonmonoxy Myoglobin. *J. Phys. Chem. B* **2001**, *105*, 7057–7063.
- (33) Leitner, D. M.; Straub, J. E. *Proteins: Energy, Heat and Signal Flow (Computation in Chemistry)*; CRC Press: Boca Raton, FL, 2009.
- (34) Leitner, D. M. Energy Flow in Proteins. *Annu. Rev. Phys. Chem.* **2008**, *59*, 233–259.
- (35) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (36) Henry, E. R.; Levitt, M.; Eaton, W. A. Molecular Dynamics Simulation of Photodissociation of Carbon Monoxide from Hemoglobin. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2034–2038.
- (37) Giammona, D. A. Ph.D. Thesis, University of California, Davis, 1984.
- (38) Case, D. A.; Darden, T. A.; Cheatham, T. E. III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; et al. *AMBER 9*; University of California: San Francisco, 2006.
- (39) Meller, J.; Elber, R. Computer Simulations of Carbon Monoxide Photodissociation in Myoglobin: Structural Interpretation of the B States. *Biophys. J.* **1998**, *74*, 789–802.
- (40) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK<sub>a</sub> Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537.
- (41) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (42) Ward, J. H., Jr. Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* **1963**, *58*, 236–244.
- (43) R Development Core Team. *R: a Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2012.
- (44) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (45) Chartrand, G. *Introductory Graph Theory*; Dover Publications: New York, 1985.
- (46) Feitelson, J.; McLendon, G. Migration of Small Molecules through the Structure of Hemoglobin: Evidence for Gating in a Protein Electron-Transfer Reaction. *Biochemistry* **1991**, *30*, 5051–5055.
- (47) Goldbeck, R. A.; Bhaskaran, S.; Ortega, C.; Mendoza, J. L.; Olson, J. S.; Soman, J.; Kliger, D. S.; Esquerra, R. M. Water and Ligand Entry in Myoglobin: Assessing the Speed and Extent of Heme Pocket Hydration after CO Photodissociation. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1254–1259.