Vol. 12, pp. 131–138 (2015)
doi: 10.2142/biophysico.12.0_131

Biophysics and Physicobiology
https://www.jstage.jst.go.jp/browse/biophysico/

Regular Article

Analysis of water channels by molecular dynamics simulation of heterotetrameric sarcosine oxidase

Go Watanabe[1], Daisuke Nakajima[1], Akinori Hiroshima[1,2], Haruo Suzuki[1] and Shigetaka Yoneda[1]

[1]*School of Science, Kitasato University, Sagamihara, Kanagawa 252-0373, Japan*
[2]*Present address: Mitsui Knowledge Industry Co., Ltd., Minato-ku, Tokyo 105-6215, Japan*

Received October 1, 2015; accepted November 16, 2015

A precise 100-ns molecular dynamics simulation *in aquo* was performed for the heterotetrameric sarcosine oxidase bound with a substrate analogue, dimethylglycine. The spatial region including the protein was divided into small rectangular cells. The average number of the water molecules locating within each cell was calculated based on the simulation trajectory. The clusters of the cells filled with water molecules were used to determine the water channels. The narrowness of the channels, the average hydropathy indices of the residues of the channels, and the number of migration events of water molecules through the channels were consistent with the selective transport hypothesis whereby tunnel T3 is the pathway for the exit of the iminium intermediate of the enzyme reaction.

Key words: selective transport, enzyme reaction, flavo-protein, channelling

Heterotetrameric sarcosine oxidase (HSO) catalyzes sarcosine oxidation to generate glycine and hydrogen peroxide, and either formaldehyde or 5,10-methylene-tetrahydrofolate (5,10-CH_2-THF), depending on the availability of tetrahydrofolate (THF) [1]. HSO is a complex enzyme with molecular weight of approximately180,000, consisting of four non-identical subunits α, β, γ, and δ. HSO contains a nicotinamide adenine dinucleotide (NAD^+), Zn^{2+}, and two flavins, *i.e.*, flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN) [2]. Since the first X-ray analysis of HSO by Ida *et al.* [3], several X-ray structures have been reported for HSOs from *Corynebacterium* sp. U96 and *Pseudomonas maltophilia*, bound with various ligands, and in different redox conditions [4,5]. The X-ray structures have clarified that the α-subunit of HSO is composed of the N-terminal half domain (nicotine domain), the C-terminal half domain (folate domain), and a linker of about 10 residues connecting the N- and C-terminal halves, as shown in Figure 1. NAD^+, for which the function is not clear, is contained in the nicotine domain. The γ-subunit is located along the linker of the two domains of the α-subunit. The δ-subunit has a Cys_3His zinc finger motif to bind Zn^{2+}. The β-subunit is bound with an FAD noncovalently and an FMN covalently. Sarcosine is bound at the *re*-phase of the isoalloxazine ring of the cofactor FAD. It is proposed that the cofactor FAD is reduced by sarcosine, an electron is transferred from the reduced FAD to FMN, and the reduced FMN is oxidized by an oxygen molecule to yield an oxidized flavin. HSO contains a large spheroid-like cavity with volume of about 10 nm^3 near the substrate-binding site [3].

Seven tunnels have been proposed as the pathways for the migration of substrates and products in HSO [5] using the geometry-based tunnel prediction tool CAVER2 [6], as shown in Figure 1. The tunnels named T1, T2, T3, and T4 are connected to the large cavity. T1 and T2 are located near the boundary between the α- and β-subunits. T3 passed through the center of the folate domain of the α-subunit. Based on the structural and kinetics studies, the sarcosine oxidation in HSO has been postulated to accompany the selective transport of the substrate and products as described here on [5]. Sarcosine migrates into the large cavity from the surface of HSO through T1, T2, or both. Since the electrostatic surface potential of T1 and T2 are mostly negative, the zwitterionic sarcosine might migrate via T1 or T2, rather than via T3, which possesses neutral and hydrophobic regions. After the

Corresponding author: Shigetaka Yoneda, School of Science, Kitasato University, 1-15-1 Kitasato, Minami-ku, Sagamihara, Kanagawa 252-0373, Japan.
e-mail: syoneda@kitasato-u.ac.jp

substrate reaches the FAD active site, dehydrogenation occurs and an iminium intermediate (5-oxazolidinone) is generated. The intermediate goes through T3 to reach the THF binding site, and is converted to glycine and either 5,10-CH_2-THF or formaldehyde, depending on whether THF is bound to the site. The other three tunnels (T5, T6, and T7) from the FMN isoalloxazine ring site are considered channels for oxygen molecules and the hydrogen peroxide.

It has not been confirmed that the tunnels stably exist and the molecules can truly migrate via them since these tunnels has been geometrically predicted from one crystal structure. While the experimental studies hardly support the proposed selective transport of the reactants and the products involved in the HSO enzyme-catalyzed oxidation, the molecular dynamics (MD) simulation study can give a new insight to understand the dynamics of the enzyme. As the first step to investigate the migration route of the substrates and the products, we carried out a precise MD simulation of HSO and analyzed the pathways of water molecules using the trajectory of water molecules in the MD simulation. Then, we compared those pathways with the tunnels found in the previous study. We present herein the detailed description of our MD simulation of HSO and an analysis of the water channels, using the trajectory of water molecules in the simulation. We hereafter use "tunnel" to indicate a geometrical void in the protein and "channel" to indicate a region where water molecules practically reside.

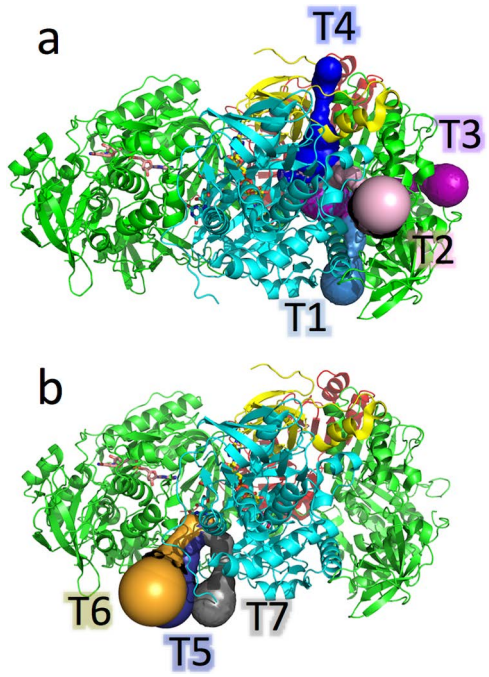

Figure 1 The X-ray structure of the HSO-DMG complex (PDB id; 1X31). The α-, β-, γ-, and δ-subunits are shown in green, cyan, red, and yellow, respectively. The γ-subunit is shown below the other subunits. The small molecules are drawn as sticks. Seven tunnels, T1, T2, T3, and T4 shown in (a), and T5, T6, and T7 shown in (b), are drawn as surfaces. Reproduced from the previous study [5].

Methods

The initial structure of our MD simulation was the X-ray structure (Fig. 1) of HSO bound with a substrate-analogue, dimethylglycine (DMG), from the Protein Data Bank (PDB). The numbers of residues were 963, 402, 195, and 91 for the α-, β-, γ-, and δ-subunits, respectively. The structure was initially positioned at the center of a rectangular periodic boundary box with size of $9.58 \times 12.031 \times 9.854\ nm^3$ filled with water molecules, as shown in Figure 2. The protonation states of amino acid residues and the hydrogen coordinates of HSO were determined using Protonate 3D [7] implemented in MOE 2013.08 [8]. The AMBER ff99SB-ILDN energy parameters [9] were used for HSO and Na^+, the general AMBER force-field parameters [10] for DMG, and TIP3P for water molecules. The partial charges of the atoms of DMG were calculated using the restrained electrostatic potential (RESP) method [11] based on quantum chemical calculations using the 6-31G* basis set of GAUSSIAN03 [12] (Table S1). The partial charges proposed by Calimet and Simonson were adopted for Zn^{2+} and the surrounding residues [13]. The total number of atoms was 114,858, including 24,997 atoms for the HSO-DMG complex, 89,799 atoms for surrounding water molecules, and 62 Na^+ ions that neutralize the total charge of the whole system.

The MD program GROMACS 4.5.5 [14] was used to perform MD simulations. After running the steepest descent energy minimization, the relaxation and equilibration runs were carried out at 200 K for 1 ns and at 300 K for 3 ns, respectively, under the NTV ensemble using the velocity-rescaling thermostat method [15] with a relaxation time of 0.1 ps and using positional restraints for all the non-hydrogen atoms of HSO. Finally, a production MD run with no positional restraints was performed for 100 ns at 300 K under the

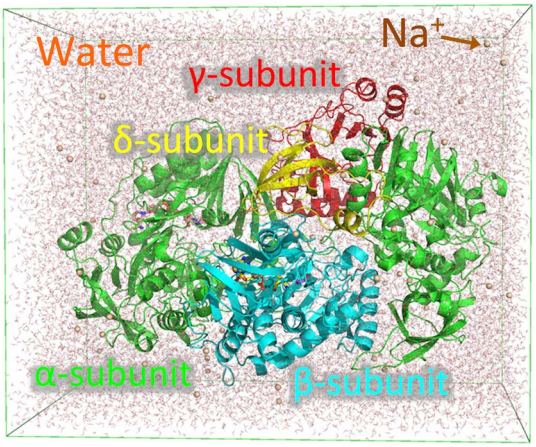

Figure 2 Snapshot of MD simulation at 100 ns. The periodic boundary box is drawn in green and the water molecules filling the box are drawn as small red lines. Small brown spheres show Na^+ ions. Four subunits of the HSO are represented with coloured ribbon: green for α-subunit, cyan for β-subunit, red for γ-subunit, and yellow for δ-subunit.

NTP ensemble using the Parrinello-Rahman method [16] with pressure of 1 bar. All bonds involving hydrogen atoms were constrained with LINCS [17] and the time step was set to 2 fs. The particle mesh Ewald (PME) [18] method with interpolation order of 6 and Fourier spacing of 0.15 nm was employed to treat the long-range electrostatic interactions. A cut-off of 1.2 nm was used for the Lennard-Jones interactions and the direct sum of electrostatic interactions.

To analyze water channels in the HSO molecule, the probability density distribution of water was determined from the trajectory of water molecules, as follows. The spatial region including the HSO molecule was divided into small rectangular cells of $0.3 \times 0.36 \times 0.3$ nm^3. This cell size is approximately equal to the size of water molecule. As the HSO molecule was rigid as described in the following sections, this cell size was expected to be effective for smoothing the small fluctuations of the protein atoms and the shape of water channels. The average number of water oxygen atoms within each cell over time was calculated. Before this calculation, the small rotation and translation of HSO in the simulation were removed by least squares (LS) fitting so that the HSO molecule was relatively immobile with respect to the cells. Cells with an average number of water oxygen atoms of greater than 0.5 were named "water cells." This criterion of 0.5 is equivalent to the probability density of water molecules of 15.4 nm^{-3} and is about half of the probability density of 33 nm^{-3} (1 g cm^{-3}), which is the typical value for bulk water. Other cells were assumed in the protein region. This cell division was not highly sensitive to the criterion. In order to classify the water cells into those within the protein (water channel) and those outside of the protein (outer water region), 26 lines were drawn from the center of each cell, as shown in Figure 3. If any atoms of the protein were located within 0.3 nm of a line, the line was determined to penetrate the protein. The water cells with less than 14 penetrating lines were assumed to compose the outer water region. The water cells with more than 13 penetrating lines were assumed to be within HSO to compose the water channels. Water channel cells located in small pockets on the HSO surface were removed manually using computer graphics. The resulting 844 water channel cells were classified into separate channels by a cluster analysis using the R software environment [19]. However, the large cavity that has been proposed in previous studies [3] was shared by several channels; therefore, clear identification of the cavity was not easy using only cluster analysis methods. Thus, after the simple cluster analysis with the farthest neighbor method was performed, the cavity cells were determined by using a rectangular box; when the channel cells were within the rectangular box, they were assumed to compose the cavity. The volume of the rectangular box was $1.2 \times 1.7 \times 1.4$ nm^3 and its center was set at the center of the cavity, *i.e.*, (–2.0661 nm, 6.9338 nm, 1.0942 nm) in the coordinate system of the X-ray structure 1X31. This procedure resulted in the clear assignment of the spheroid-like cavity as shown in the Results.

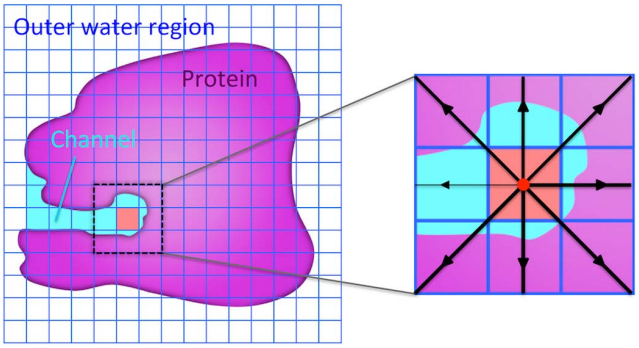

Figure 3 Two-dimensional schematic figure of the small rectangular cells used to estimate of the probability density of water molecules. The cells are depicted as quadrangles with blue edges. The cells with a small probability density of water compose the protein region (magenta). The cells with a large probability density are water cells. As shown in the inset, 26 lines were considered from the center of the cell (red circle), passing through the eight vertices, the twelve edge centers, and the six plane centers of the cell in the three-dimensional space. The cells with many penetrating lines (bold lines) form the water channel (cyan).

Results

Before the analysis of water channels, the precision of the MD simulation was examined based on the root mean square deviations (RMSD) of the Cα atoms of the simulation structure from the initial X-ray structure, as shown in Figure 4. The final equilibrated RMSDs were very small, approximately 0.14 nm. Furthermore, the fluctuation in position was large for atoms with large experimental B-factors (data not shown), indicating a good correspondence between fluctuations in the simulation and experiments. Thus, the present simulation was precise with respect to the structure and fluctuation of the entire HSO molecule. The simulation was also precise with respect to the local structures, including the substrate binding site. Figure 5 shows the initial and final structures of the substrate binding site. The two structures were very similar, accounting for the thermal fluctuations. In the X-ray structure, the carbonyl oxygen atoms of DMG are hydrogen bonded with Nη2 of Argβ69 and with Nζ of Lysβ358. As shown in Figures 6a and 6b, the hydrogen bonds were continuously formed in the simulation. The hydrogen bonds were often in the form of bifurcated hydrogen bonding. In contrast, it is natural that the methyl groups were rather loosely bound with the site, as shown in Figure 6c, because DMG is not a substrate of HSO [20].

The cavity and eleven channels determined by the water cells are shown in Figure 7 (see also Supplementary Materials). The number of cells of the cavity was 202 and those of the channels are provided in Table 1. The channels CH1, CH2, CH3, and CH4 were connected to the cavity and correspond to the tunnels identified in the previous study [5], T1, T2, T3, and T4, respectively (Fig. 1). CH6 and CH7 were also connected to the cavity, but are previously undetected. CH6

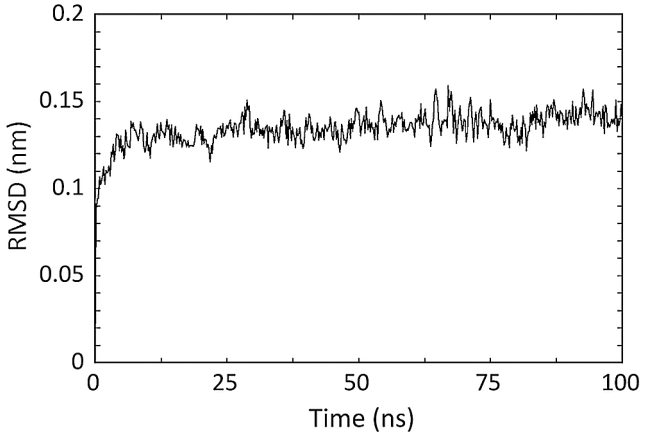

Figure 4 Time profile of RMSDs of the Cα atoms from the initial X-ray coordinates. The small rotation and translation were removed by the LS fitting. Without the LS fitting, the RMSDs were approximately 0.01 nm larger.

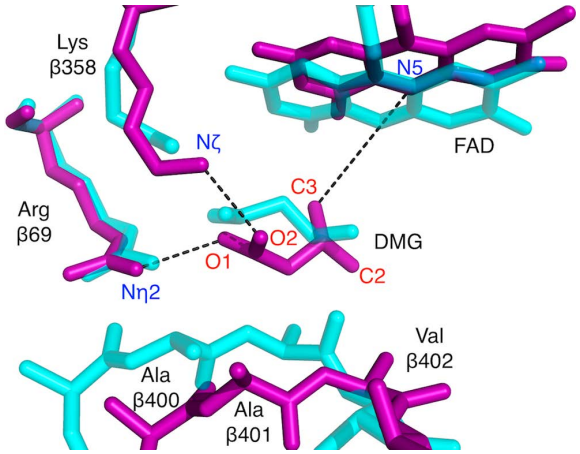

Figure 5 Structure of the DMG binding site. The initial X-ray structure and the structure at 100 ns of the MD simulation are shown in magenta and cyan, respectively. The atomic distances shown in Figure 6 are depicted as broken lines.

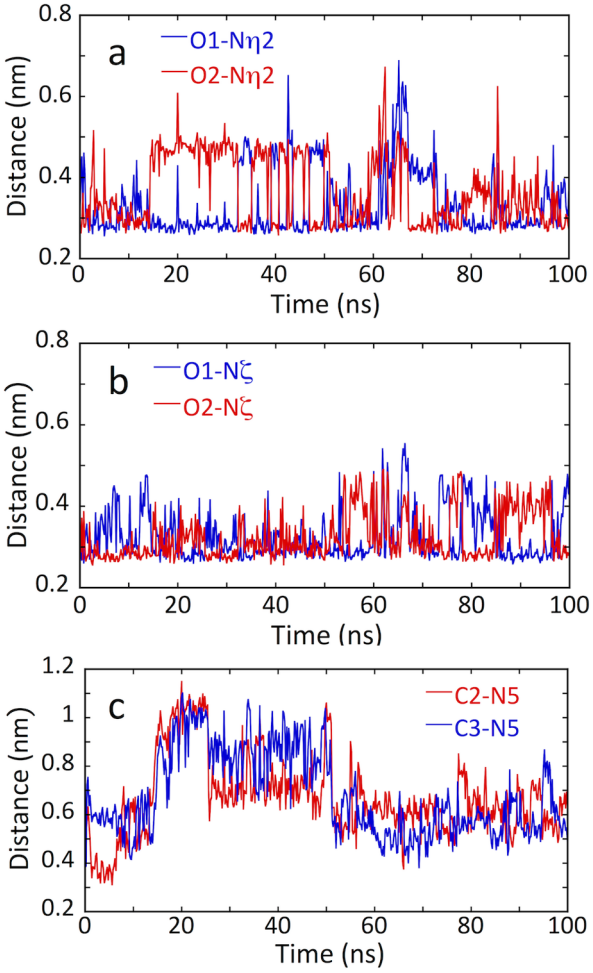

Figure 6 Time profiles of the atomic distance between (a) Nη2 of Argβ69 and the oxygen atoms, O1 or O2, of DMG; (b) the distance between Nζ of Lysβ358 and O1 or O2 of DMG; and (c) the distance between the N5 of the isoalloxazine ring of the FAD cofactor and the two carbon atoms, C2 and C3, of the methyl groups of DMG.

extended through the interface among γ-subunit, δ-subunit, and the C-terminal region of the α-subunit. CH7 extended through the interface among the α-, β-, and γ-subunits. CH8 and CH9 made a single combined channel from the cavity through the interface of the α-, β-, and δ-subunits. This combined channel has not previously been detected. CH8 was buried in the protein and CH9 was near the surface. These channels, except CH1, CH2, and CH3, were narrow, or had some narrow necks. CH5, CH10, and CH11 were not connected to the cavity. CH5 extended from FMN to the surface of HSO and corresponds to the combined region of T5, T6, and T7 in the previous study. Both CH10 and CH11 were connected to the NAD^+ binding site, and were also connected to each other to make a combined water channel. CH11 was open to the protein surface and was a pathway for the NAD^+ binding site. Although CH10 was nearly buried inside the protein, water molecules were found to migrate between CH10 and the outer water region in the simulation. Thus, for very small molecules like water, CH10 was another pathway to the NAD^+ binding site. The amino acid residues that made the channels are shown in Figure 8 and the average hydropathy index [21] of the residues are shown in Table 1. Only CH3 was hydrophobic with a positive average hydropathy index, unlike the other channels. This characteristics is in agreement with the electrostatic and structural properties of T3 determined in the previous study [5].

Figure 9 indicates typical trajectories of the water molecules migrating through each of CH1, CH2, CH3, and CH4. As can be seen in Figure 9, the water molecules mainly continued back-and-forth motions and showed sudden and long jumps at about several ten ps intervals. We calculated the number of migrations of water molecules between the cavity and the outer water region (Table 2). In this calculation, we

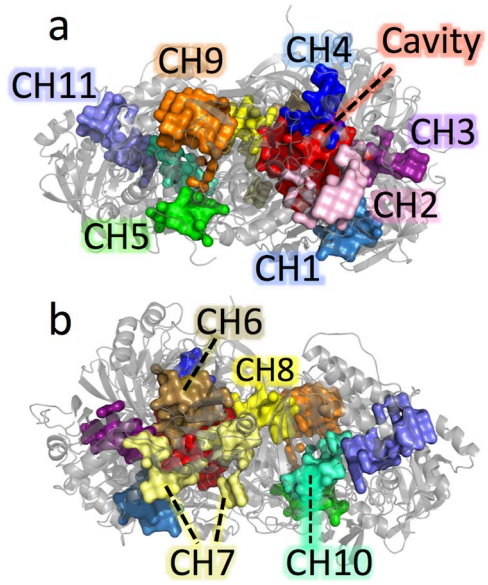

Figure 7 Eleven water channels determined. The channels, CH1, CH2, CH3, CH4, CH6, CH7, and the set of CH8 and CH9 connect the large cavity to the outer water region of HSO. Seen from the same direction of Figure 1 (a). Seen from behind (b).

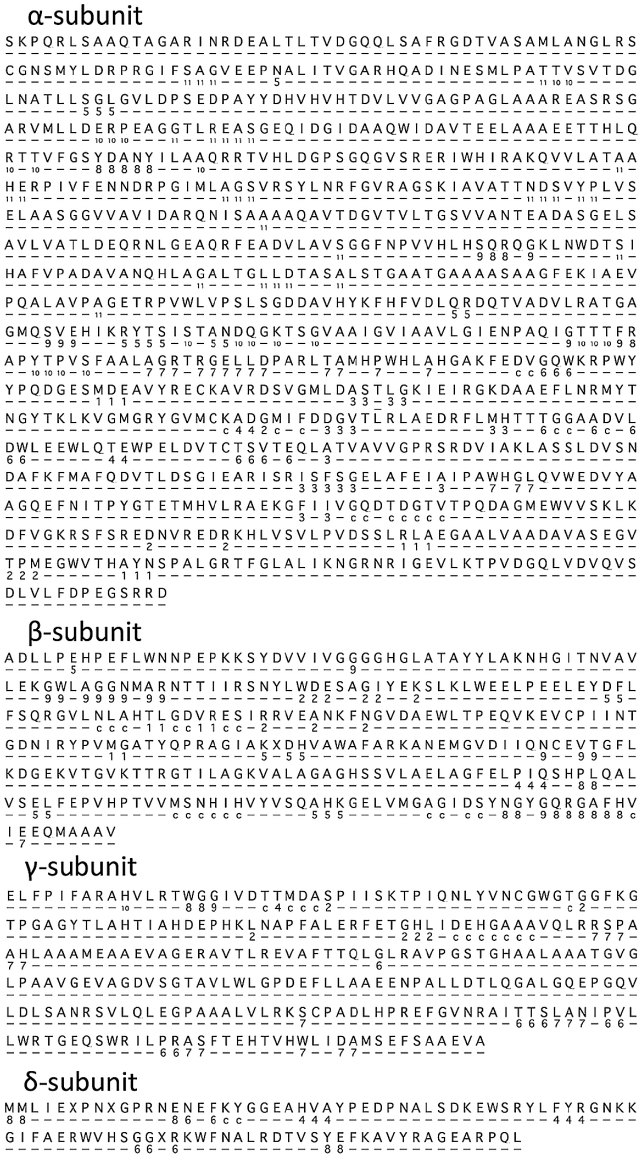

Figure 8 Amino acid residues that compose the inner walls of the channels and the cavity. These residues were determined by the criterion that the Cα atom is within 0.3 nm from the center of a cell of the channel.

excluded the cells of channels that were contiguous to the cavity in order to decrease the calculational noises made by fast and transient back-and-forth motions of water molecules between the cavity and the channels. The numbers of water molecules moving into and out of the cavity were approximately equal owing to the principle of detailed balance. The number of the water molecules migrating through the protein between the channels was also examined, but no permeating movement was detected except the migration of water molecules in the combined channel of CH8 and CH9 and in that of CH10 and CH11.

Discussion

The channels CH1, CH2, CH3, and CH4 correspond to T1, T2, T3, and T4, respectively. CH5 corresponds to the combined region of T5, T6, and T7. Thus, the five water channels determined in this study correspond to previous geometrical tunnel predictions. CH6 and CH7 were not found in the previous study [5] because they are narrow channels with few migrating water molecules, as shown in Table 2. Although migrations through CH4 and the combined channel of CH8 and CH9 are frequent (Table 2), the channels have a narrow neck through which molecules larger than water are difficult to migrate.

The surfaces of the channels and the protein are shown in Figure 10. The channel is drawn as the cluster of the cells using PyMOL. CH6 and CH7 connecting between the cavity and the outer region, CH10 buried in the protein, and the combined narrow channel of CH8 and CH9 are not obvi-

Table 1 The number of cells and the average hydropathy index[a]

	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8	CH9	CH10	CH11
Number of cells	46	65	46	32	49	63	83	35	105	50	68
Average hydropathy index	–0.39	–0.95	1.1	–0.53	–1.6	–0.69	–0.51	–1.4	–0.46	–1.5	–0.52

[a] CH8 and CH9 are connected to form a single combined channel as are CH10 and CH11.

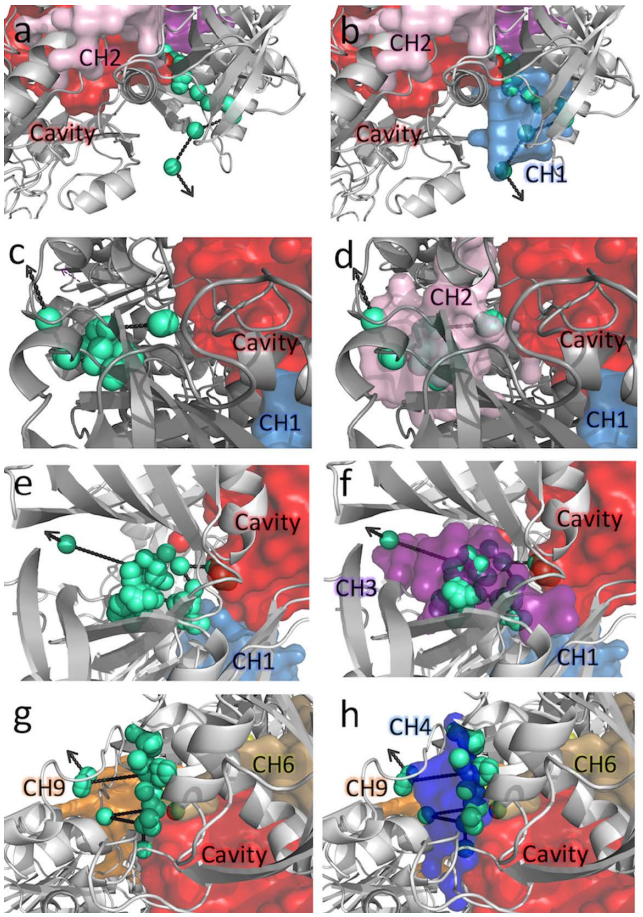

Figure 9 Trajectories of the selected water molecules from the cavity to the outer water region through CH1 (a) and (b), CH2 (c) and (d), CH3 (e) and (f), and CH4 (g) and (h). For clarity, the channels through which the water molecules migrated are not shown in (a), (c), (e), and (g). A gray ribbon represents the structure of the HSO at 100 ns of the MD simulation. Green spheres depict the positions of the oxygen atoms of water molecule at 40 ps intervals. Arrows show the direction of the migration of water molecule.

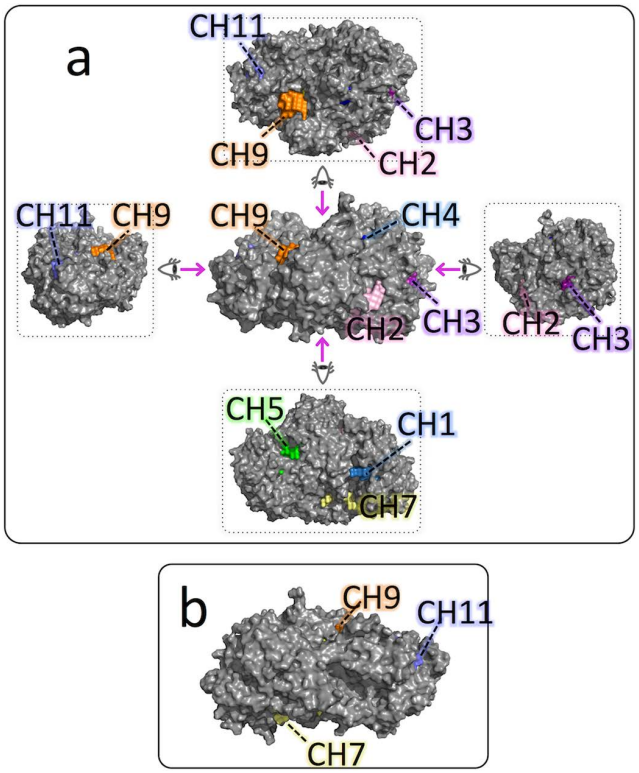

Figure 10 View of the surface of the HSO-DMG complex and eleven water channels determined from the MD simulation. The surface of the HSO is the area in gray. (a) From the same direction of Figure 1(a). The views from the top, bottom, left side, and right side are also included. (b) Seen from the behind.

ously exposed in the surface of the protein. The area of the channels exposed in the surface is not clearly correlated with the frequency of the migration of water molecules.

The combined channel of CH10 and CH11 is not connected to the cavity and have not been found in the previous study. It is notable that the CAVER2 calculations are dependent on the initial starting point of the tunnel search, whereas the method in the present study automatically clarified all the water regions where water molecules resided in the simulation.

Conventional tunnel prediction methods are based on static structures of proteins, and fluctuations and induced fitting have been considered in recent studies using MD simulations [22–29]. The present study also used MD simulation to estimate the distribution of probability densities of rapidly

Table 2 The number of water molecules migrating through the channels from 80 ns to 100 ns

Channel	Corresponding previous tunnel [5]	Outer water region to cavity	Cavity to outer water region	Number of cells at neck[a]
CH1	T1	133	120	4
CH2	T2	43	47	4
CH3	T3	33	37	3
CH4	T4	209	205	1
CH6	–	10	10	1
CH7	–	11	12	1
CH8 and CH9	–	48	41	1

[a] Number of water cells in the narrowest regions of the water channels.

moving water molecules. It was assumed that the fluctuations of the protein were not too large to define rigid channels. Grasping channels as rigid regions is a useful device for understanding the transport process in a protein. The entire structure of the HSO molecule and its local structures were highly conserved in the simulation (Figs. 4 and 6), indicating not only the high precision of the simulation, but also the high rigidity of HSO; accordingly, HSO was a good target for the analysis of solvent channels in the present approach. Furthermore, the permeation of water molecules between the channels through the protein region was not found, except in the combined channels, indicating that the channels detected in this study are well defined.

Molecular transport is a non-equilibrium process; the rate of transport is controlled by the free energy of the transition state, which is different among channels. It is thus natural that a specific channel should be selected for the transport of a particular small molecule. Based on the structural and kinetic studies [5], selective transport was proposed, whereby the iminium intermediate (5-oxazolidinone) exits through T3, *i.e.*, CH3. Among the channels from the cavity, CH4, CH6, CH7, and the set of CH8 and CH9 appear to be too narrow for the transport of molecules, except very small molecules like water (Table 2). Of the remaining channels (CH1, CH2, and CH3), CH1 and CH2 are hydrophilic, and thus the hydrophobic CH3 is the most probable pathway for the intermediate.

Conclusion

A precise100-ns MD simulation was performed for the HSO-DMG complex. A cluster analysis on the small rectangular cells with high water probabilities resulted in the detection of eleven water channels. CH1, CH2, CH3, and CH4 corresponded to the tunnels from the large cavity in the previous study. CH6, CH7, and the combined CH8 and CH9 were narrow channels from the large cavity. CH5 and the combined CH10 and CH11 were not connected with the cavity. No permeation of water molecules between the channels was found except in the combined channels. The results of the present analysis are consistent with the selective transport hypothesis.

Acknowledgement

This study was supported by the MEXT-Supported Program for Strategic Research Foundation at Private Universities. The computations were performed using Research Center for Computational Science, Okazaki, Japan.

Conflict of Interest

G.W., D.N., A.H., H.S., and S.Y. declare that they have no conflict of interest.

Author Contributions

G.W., H.S., and S.Y. directed the entire project and cowrote the manuscript. D.N. and A.H. carried out the molecular dynamics simulation.

References

[1] Suzuki, H. Sarcosine oxidase: structure, function, and the application to creatinine determination. *Amino Acids* **7**, 27–43 (1994).
[2] Jorns, M. S. Properties and catalytic function of the two nonequivalent flavins in sarconsine oxidase. *Biochemistry* **24**, 3189–3194 (1985).
[3] Ida, K., Moriguchi, T. & Suzuki, H. Crystal structure of heterotetrameric sarcosine oxidase from *Corynebacterium* sp. U-96. *Biochem. Biophys. Res. Commun.* **333**, 359–366 (2005).
[4] Chen, Z. W., Hassan-Abdulah, A., Zhao, G., Jorns, M. S. & Mathews, F. S. Heterotetrameric sarcosine oxidase: structure of a diflavin metalloenzyme at 1.85 Å resolution. *J. Mol. Biol.* **360**, 1000–1018 (2006).
[5] Moriguchi, T., Ida, K., Hikima, T., Ueno, G., Yamamoto, M. & Suzuki, H. Channeling and conformational changes in the heterotetrameric sarcosine oxidase from *Corynebacterium* sp. U-96. *J. Biochem.* **148**, 491–505 (2010).
[6] Damborský, J., Petřek, M., Banáš, P. & Otyepka, M. Identification of tunnels in proteins, nucleic acids, inorganic materials and molecular ensembles. *Biotechnol. J.* **2**, 62–67 (2007).
[7] Labute, P. Protonate3D: assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins* **75**, 187–205 (2009).
[8] Molecular Operating Environment (MOE) 2013.08. (Chemical Computing Group Inc., Motreal, 2013).
[9] Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).
[10] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
[11] Bayly, C. I., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269–10280 (1993).
[12] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. A. Jr., *et al.* Gaussian 03 Revision C. 02. (Gaussian Inc., Wallingford CT, 2004).
[13] Calimet, N. & Simonson, T. Cys_xHis_y-Zn^{2+} interactions: Possibilities and limitations of a simple pairwise force field. *J. Mol. Graph. Model.* **24**, 404–411 (2006).
[14] Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).
[15] Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
[16] Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
[17] Hess, B., Bekker, H., Berendsen, H. J. C. & Fraajie, J. G. E. M. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
[18] Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an N-log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

[19] R Core Team (2015). R: A Language and Environment for Statistical Computing. (R Foundation for Statistical Computing, Vienna, 2015) available at http://www.r-project.org/.
[20] Suzuki, M. Purification and some properties of sarcosine oxidase from *Corynebacterium* sp. U-96. *J. Biochem.* **89**, 599–607 (1981).
[21] Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).
[22] Fishelovitch, D., Shaik, S., Wolfson, H. J. & Nussinov, R. Theoretical characterization of substrate access/exit channels in the human cytochrome P450 3A4 enzyme: involvement of phenylalanine residues in the gating mechanism. *J. Phys. Chem. B* **113**, 13018–13025 (2009).
[23] Raunest, M. & Kandt, C. dxTuber: detecting protein cavities, tunnels and clefts based on protein and solvent dynamics. *J. Mol. Graph. Model.* **29**, 895–905 (2011).
[24] Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., *et al.* CAVER 3.0: a Tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput. Biol.* **8**, e1002708 (2012).
[25] Bucci, A. & Abrams, C. F. Oxygen pathways and allostery in monomeric sarcosine oxidase via single-sweep free-energy reconstruction. *J. Chem. Theory Comput.* **10**, 2668–2676 (2014).
[26] Sehnal, D., Vařeková, R. S., Berka, K., Pravda, L., Navrátilová, V., Banáš, P., *et al.* MOLE 2.0: advanced approach for analysis of biomacromolecular channels. *J. Cheminform.* **5**, 39–51 (2013).
[27] Craig, I. R., Pfleger, C., Gohlke, H., Essex, J. W. & Spiegel, K. Pocket-space maps to identify novel binding-site conformations in proteins. *J. Chem. Inf. Model.* **51**, 2666–2679 (2011).
[28] Paramo, T., East, A., Garźon, D., Ulmschneider, M. B. & Bond, P. J. Efficient characterization of protein cavities within molecular simulation trajectories: *trj_cavity*. *J. Chem. Theory Comput.* **10**, 2151–2164 (2014).
[29] Kingsley, L. J. & Lill, M. A. Including ligand-induced protein flexibility into protein tunnel prediction. *J. Comput. Chem.* **35**, 1748–1756 (2014).