

# MD Simulations Reveal Complex Water Paths in Squalene–Hopene Cyclase: Tunnel-Obstructing Mutations Increase the Flow of Water in the Active Site

Camilla Gustafsson,<sup>†</sup> Serguei Vassiliev,<sup>‡</sup> Charlotte Kürten,<sup>§</sup> Per-Olof Syrén,<sup>†,§</sup> and Tore Brinck<sup>\*,†</sup>

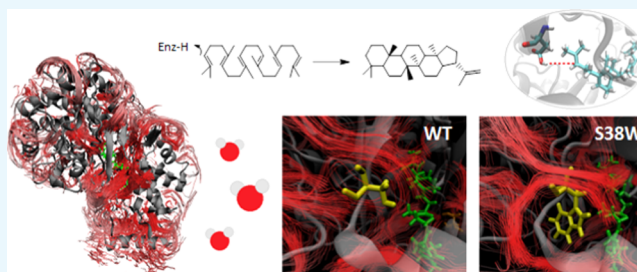
<sup>†</sup>Applied Physical Chemistry, School of Chemical Science and Engineering, KTH Royal Institute of Technology, Teknikringen 36, 100 44 Stockholm, Sweden

<sup>‡</sup>Department of Biological Sciences, Brock University, Mackenzie Chown F 234, 1812 Sir Isaac Brock Way, St. Catharines, Ontario L2S 3A1, Canada

<sup>§</sup>Science for Life Laboratory, Stockholm—School of Biotechnology, Division of Proteomics and Nanobiotechnology, KTH Royal Institute of Technology, Tomtebodavägen 23a, 171 65 Solna, Sweden

## Supporting Information

**ABSTRACT:** Squalene–hopene cyclase catalyzes the cyclization of squalene to hopanoids. A previous study has identified a network of tunnels in the protein, where water molecules have been indicated to move. Blocking these tunnels by site-directed mutagenesis was found to change the activation entropy of the catalytic reaction from positive to negative with a concomitant lowering of the activation enthalpy. As a consequence, some variants are faster and others are slower than the wild type (wt) in vitro under optimal reaction conditions for the wt. In this study, molecular dynamics (MD) simulations have been performed for the wt and the variants to investigate how the mutations affect the protein structure and the water flow in the enzyme, hypothetically influencing the activation parameters. Interestingly, the tunnel-obstructing variants are associated with an increased flow of water in the active site, particularly close to the catalytic residue Asp376. MD simulations with the substrate present in the active site indicate that the distance for the rate-determining proton transfer between Asp376 and the substrate is longer in the tunnel-obstructing protein variants than in the wt. On the basis of the previous experimental results and the current MD results, we propose that the tunnel-obstructing variants, at least partly, could operate by a different catalytic mechanism, where the proton transfer may have contributions from a Grotthuss-like mechanism.



## INTRODUCTION

Terpenes comprise the most diversified class of natural products of high importance for life.<sup>1</sup> The key to the myriad of observed multicyclic products is a terpene cyclase-catalyzed carbocationic polycyclization cascade of a prefolded linear polyisoprene substrate that is chaperoned in essentially one reactive conformation. The stereospecific terpene cyclization in the hydrophobic active site is initiated by either phosphate release (class I mechanism) or protonation of the terminal isoprene or oxirane (class II mechanism), which yields a carbocation that propagates in the molecule upon ring closure through carbon–carbon bond formation.<sup>2</sup> The final carbocation is quenched by either deprotonation or water addition depending on the availability of a suitable base and the spatial organization of water in the active site.<sup>3</sup>

The cyclization step and subsequent product modification by tailoring enzymes determine the bioactivities of terpenes that range from mediating cell membrane rigidity, defense, to signaling.<sup>4,5</sup> Triterpenes constitute one important class of terpenes with applications as medicines and fine chemical synthons.<sup>6</sup> The potent antibacterial, anticancer, and antiviral

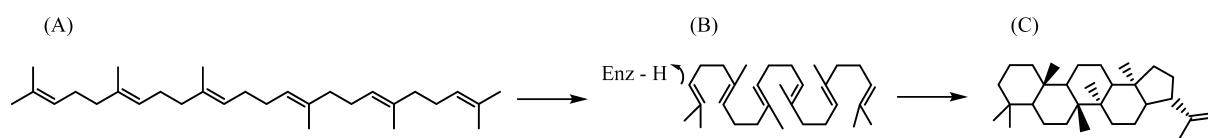
bioactivities displayed by some cyclic terpenes,<sup>6–8</sup> as well as their high synthetic potential as new biofuels,<sup>9,10</sup> have spurred increasing interest from the biotechnological and chemical industries.

Stereospecific carbon–carbon bond formation is of high importance in synthetic organic chemistry and biocatalysis.<sup>11,12</sup> Terpene cyclases are of high synthetic interest<sup>13–15</sup> as they enable stereospecific polycyclizations that are challenging to achieve by synthetic chemistry.<sup>6</sup> Therefore, an enhanced understanding of the reaction mechanism displayed by these fascinating enzymes would be beneficial for chemical and biotechnological industries. Squalene–hopene cyclase (SHC) has served as a model system to enhance our understanding of triterpene biosynthetic machineries. SHC converts the linear molecule squalene to the pentacyclic triterpenes hopene (Figure 1) and hopanol.

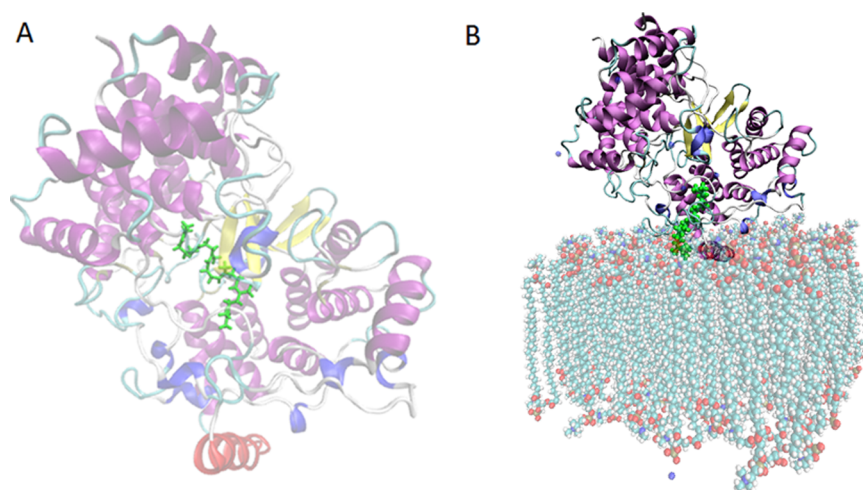
Received: July 28, 2017

Accepted: November 20, 2017

Published: November 30, 2017



**Figure 1.** Conversion from the linear squalene (A) to a prefolded geometry (B) in the active site of SHC in the polycyclization reaction to hopene (C).



**Figure 2.** Structure and membrane association of squalene-hopene cyclase from *Alicyclobacillus acidocaldarius* (PDB entry 1UMP<sup>17</sup>). (A) The prefolded squalene (green, stick) is shown in the active site. The membrane-binding helix is shown in red. (B) The protein inserted into a membrane patch, the substrate is located in the hydrophobic channel (green, stick).

SHC is a monotopic enzyme<sup>16</sup> with a hydrophobic channel toward the membrane helix that allows for acquisition of the hydrophobic substrate. The membrane-binding region of the enzyme is highly lipophilic, and it is submerged in the nonpolar region of the membrane (Figure 2).

The substrate is guided into the prefolded conformation for ring closure when entering the active site, and the shape of the active site stabilizes the substrate in a prefolded arrangement for the polycyclization to take place. The cyclization cascade is then initiated by protonation of a terminal isoprene double bond from the catalytic residue Asp376 (Figure 1), which yields a carbocation that is stabilized from premature quenching by hydrophobic side chains in the active site.<sup>14</sup> The  $pK_a$  of Asp376 is modulated by a hydrogen bond to the protonated His451 and by the close vicinity to the negatively charged Asp374:Asp377 pair.<sup>18</sup> The latter has also been postulated to stabilize the protonated substrate.<sup>19</sup>

In a study performed by Schwab et al.,<sup>20</sup> a higher mobility of the membrane-binding region was identified when a ligand was bound to the active site. However, there was no significant change of the narrowest part of the hydrophobic channel, either in fluctuations or the absolute structure.

The polycyclization reaction is highly exergonic and the formation of hopene theoretically releases about 48 kcal/mol.<sup>20</sup> It has been suggested that this energy release affects the structure of the protein in such a way that the hydrophobic channel is expanded enough to accommodate the bulky product.<sup>17</sup> Others have advocated that energy is released and dissipated much faster than the timescale of the necessary conformational changes.<sup>20,21</sup>

A previous study by Syrén et al. detected a network of tunnels leading from the active site to the surface of the protein through an *in silico* CAVER<sup>22</sup> analysis.<sup>23</sup> It was suggested that expulsion of water molecules from the active site to the bulk

could compensate for the entropic cost of prefolding the substrate in one reactive conformation.<sup>23</sup> The identified highest ranked tunnels (from CAVER) were obstructed with rational *in silico* protein design by introducing bulky amino acids at selected positions. *In vitro* experimental kinetic characterization showed that the variants display significantly altered temperature dependencies of catalysis and entropies and enthalpies of activation compared to those of the wild type (wt), suggesting the importance of unobstructed water flow for the performed catalytic reaction (Table 1). All variants, except for the control

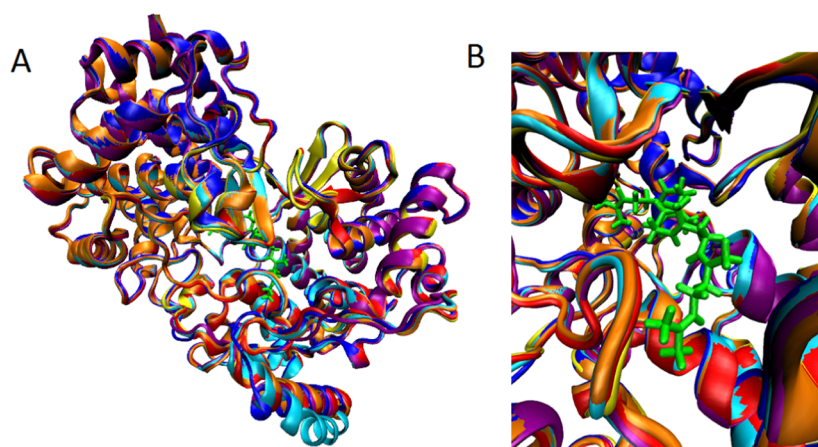
**Table 1.** Kinetic Experimental Data from Syrén et al.<sup>23</sup> Displaying Differences in Entropic ( $T\Delta S^\ddagger$ ) and Enthalpic ( $\Delta H^\ddagger$ ) Contributions to the Activation Free Energy ( $\Delta G^\ddagger$ ) between the Different Protein Variants<sup>a</sup>

protein variant	wt	S38W	S168F	S168W	F605W
$T\Delta S^\ddagger$ (kcal/mol)	16	-41	-16	-13	-7.1
$\Delta H^\ddagger$ (kcal/mol)	31	-23	-0.75	4.4	7.1
$\Delta G^\ddagger$ (kcal/mol)	15	18	15	17	14
$(k_{cat}/K_M)_H/(k_{cat}/K_M)_D$	2.0	1.9	1.6	1.0	2.1

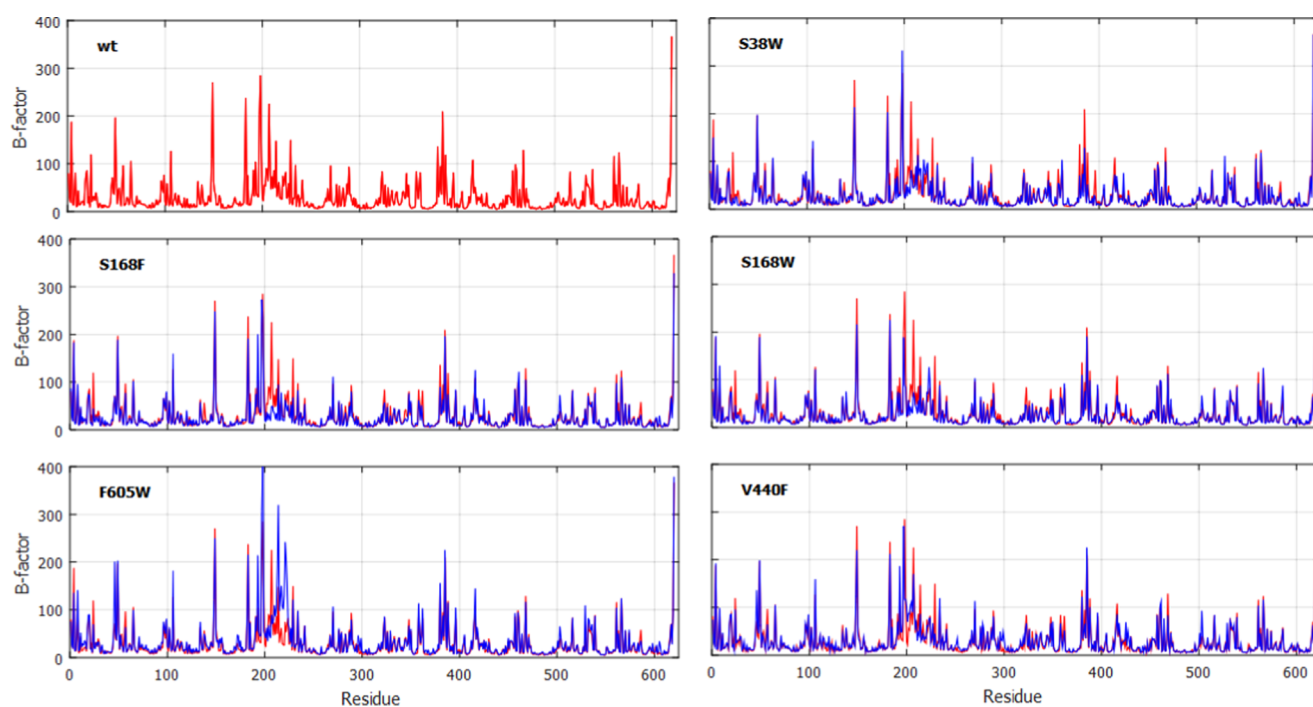
<sup>a</sup>Data is given for apparent  $k_{cat}/K_M$  at 328 K and pH 6. Deuterium kinetic isotope effects (DKIEs)  $[(k_{cat}/K_M)_H/(k_{cat}/K_M)_D]$  are taken from the Supporting Information of ref 23.

mutation, showed decreased activation enthalpy ( $\Delta H^\ddagger$ ) and activation entropy ( $\Delta S^\ddagger$ ) compared to those of the wt. However, the exact mechanism by which these mutations affect the cyclization reaction remained unknown.

The impact of water within enzyme structures has been shown to have important functions in the catalytic activities of enzymes.<sup>24,25</sup> It has been found that water molecules can improve the steric fit of ligand binding, mediate hydrogen bonds between enzyme and active site, and affect interaction



**Figure 3.** (A) Wt (blue) superposed onto the mutated protein variants S38W (yellow), S168F (orange), S168W (red), F605W (cyan), and V440F (purple). The structures used for generation of the figure are averaged structures from the trajectories. Squalene is indicated in green. The membrane patch is not visualized in the following images for clarity, but was included in all simulations. (B) A close-up of the protein backbone structure for the different protein variants in the active site. Squalene is indicated in green.



**Figure 4.** Investigated variants show small variations in B-factors. B-factors from MD simulations are given for the wt (red) and protein variants (blue). For reference, wt data is plotted in each subplot. One highly mobile region of the protein is constituted by the residues around the membrane-binding region (residues 216–234). The analysis shown was performed on the last 100 ns of the 150 ns trajectory.

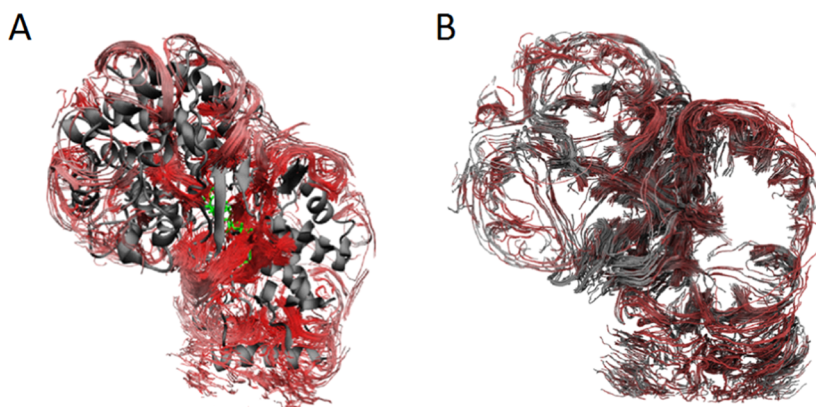
energies.<sup>24</sup> Product release can be facilitated by refilling the active site by a flow of water molecules from the bulk solution upon product release.<sup>26</sup> It has previously been suggested that a network of water molecules within tunnels that connect the active site to the bulk water on the outside of the protein might be a general feature of all enzymes to enable protonation, as well as rehydrating and dehydrating the active site during product release and substrate binding.<sup>27</sup> In a study by Schiebel et al.,<sup>27</sup> mutations were introduced in water channels in saFAB1 (*Staphylococcus aureus* enoyl-acyl carrier protein reductase) and were found to affect the catalytic efficiency of the enzyme. These mutations were suggested to hinder the flow of water out from the active site. It was additionally argued that because the mutations are at a distance from the active site, it is more likely

to be global processes that are disturbed than local processes in the active site.

The aim of the current study was to further investigate the previously suggested water tunnels in SHC<sup>23</sup> to study the movements of water molecules between the active site and the surface of SHC in more detail, and to analyze how the introduction of bulky amino acids, predicted to block the water tunnels, affect the flow of water molecules in the active site. Therefore, we investigated the dynamics of the tunnels and water flow in the active site during sub-microsecond MD simulations of the wt as well as different protein variants.

The tunnel-blocking variants were found to have an increased flow of water around the catalytic residue Asp376 in the active site compared to that in the wt. The higher water





**Figure 5.** Streamline analysis showing the flow of water (red lines) between the bulk solution and the active site. (A) Streamlines through the whole protein with the substrate in the active site. (B) Streamlines visualized through a clipping plane through the empty protein cut in half (protein not included for visualization).

flow is associated with a larger distance for the proton transfer between Asp376 and the substrate for the rate-determining proton transfer step of the catalytic reaction. However, there is no correlation between the protonation distance and the activation free energy, and we propose that the designed protein variants function through a different mechanism to that of the wt; one that at least partly could be dependent on Grothuss-like proton transfer.

## RESULTS AND DISCUSSION

The stability of the protein structure without a membrane has previously been established in a simulation by Schwab et al.<sup>20</sup> However, a recent study of membrane protein has indicated that both membrane and protein need to be considered explicitly to reproduce the hydration dynamics of the system.<sup>28</sup> Therefore, the membrane was included in our model to account for potential contributions from water associated to the membrane. The protein structures (both 1UMP and 2SQC) were found to be stable for the full-length simulation at 330 K in water (see SI for root-mean square deviation (RMSD) plots) with and without the membrane patch.

Visual inspection of the protein structures superposed on top of the wt shows that the protein structures stay intact, and do not change significantly over the course of the simulations, for any of the protein variants. An analysis of the RMSDs was performed on averaged structures from the trajectories. The RMSD values for the backbone of the mutated protein variants compared to that of the wt were S38W: 0.510 Å; S168F: 0.385 Å; S168W: 1.270 Å; F605W: 0.405 Å; and V440F: 0.509 Å. The mutations do not introduce any major visible changes in the protein three-dimensional structure, as is depicted in Figure 5.

**Protein Flexibility.** An analysis of the B-factors for each amino acid residue in the protein detected only small variations in the flexibility between the different variants. However, some parts of the enzyme are more flexible than others. Long distance effects of introduced mutations on the mobility of protein loops have previously been discussed.<sup>29</sup>

The membrane-binding region (residues 216–234) was found to display high B-factors and thus seems to be flexible in comparison to other parts of the enzyme (Figure 4). This is in agreement with a study by Wendt et al.,<sup>30</sup> where the membrane-binding helix was described as the most mobile region, based on the variations of the B-factors from the crystal structure. There were some differences in the membrane-

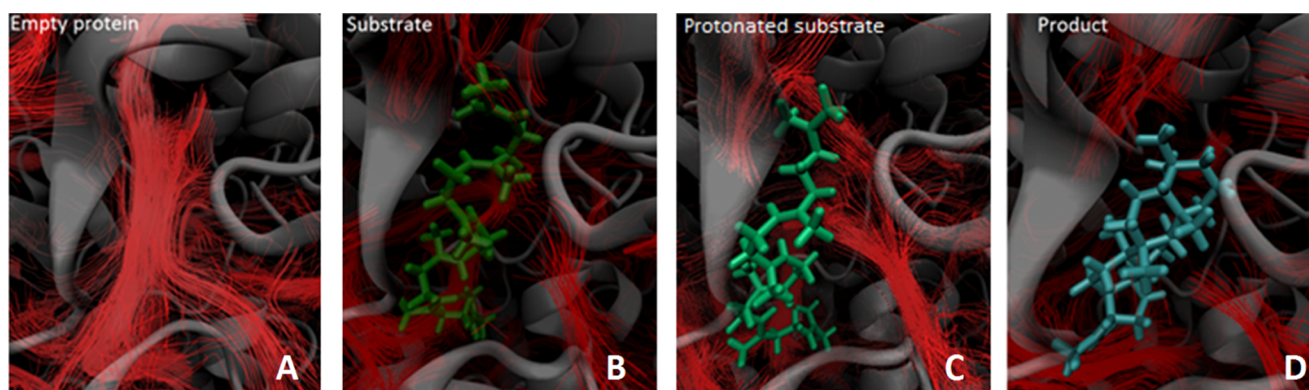
binding helix between the different protein variants (Figures 3 and 4). For one of the variants, F605W, the helix is slightly twisted compared to those of the other variants (Figure 3). Interestingly, this variant was the fastest at all studied temperatures among the variants analyzed by Syrén et al. (Table 1). The rate increase of F605W is probably connected to other factors, but, at least, the conformational change does not seem to have a negative impact on the catalytic conversion. In all variants, the N- and C-terminal regions show high flexibility, as could be expected for terminal regions of the protein chain.

Previous studies have shown that there is a relationship between the activity of an enzyme and the rigidity of the residues in the active site. Some flexibility of the residues in the active site is necessary for the enzyme to perform its catalytic function throughout the catalytic cycle.<sup>31</sup> An increased stiffness of residues has been associated with a decrease of activity for the barnase enzyme.<sup>32</sup> Focusing on the flexibility of the amino acids in the DXDD motif (Asp374, Asp376, and Asp377), which is involved in the initial protonation step,<sup>33</sup> some minor differences in B-factors can be noted (Table S2 in SI), but the changes do not follow the trends of the kinetic behavior of the experimental results. Thus, our analysis indicates that the variations in B-factors of the DXDD motif cannot explain the measured experimental differences between the mutated variants for this enzyme.

**Substrate Entrance and Product Release.** The time-scales of substrate entrance and product release were investigated over the MD trajectories. Simulations with the substrate in the hydrophobic channel (Figure 2) show that the substrate can move through the channel during this time frame. In simulations with the substrate located in the hydrophobic channel at the start of the simulation, the substrate leaves the channel and moves into the membrane.

Wendt et al. suggested that the product leaves through the same channel as that through which the substrate enters.<sup>19</sup> There are two mobile loops attached to the membrane, and the required widening of the channel could occur by membrane fluctuations.<sup>30</sup> However, in this study, the product does not leave the active site during the simulation (even when extended to 180 ns). It is likely that the hydrophobic channel needs to expand to accommodate the bulky pentacyclic hopene for the transport from the active site to the membrane. To simulate the heat that is released during the reaction, MD was also run at a higher temperature (350 K) to allow for a possible energy-





**Figure 6.** Streamlines in (A) empty protein, (B) protein with substrate (green) in the active site, (C) protonated substrate (green), and (D) protein with product (blue) in the active site.

dependent widening of the channel. However, the product does not leave the active site during the course of the simulation at elevated temperature. It is difficult to assign the significance of this result, because the higher temperature may not reproduce the effect of the energy dissipation, and any associated conformational change, that results from the energy release from the reaction. Furthermore, the timescale of the simulation may not be sufficient for the event to take place.

**Analysis of Water Tunnels.** Visual inspection of the trajectories shows that no stationary water molecules exist in the active site, or in close proximity to the active site on the timescale of 0.1  $\mu$ s, independently of whether the protein is empty, has the substrate in the hydrophobic channel or active site, or the product is in the active site. There is also a spontaneous exchange of water molecules between the protein interior and the bulk water.

The tunnels identified in the CAVER analysis (data not shown) clearly display a dynamic behavior, and the tunnels open and close over the time frame of these simulations. Because CAVER analysis does not reveal if there are water molecules moving or residing in these tunnels, an additional streamline analysis was performed.

**Streamline Analysis.** A streamline analysis performed on the SHC wt shows that water moves continuously through certain regions of the protein (Figure 5) when simulated in a water box. The most intense water flow takes place in the region in and around the active site. The streamline is a curve that is at any point tangent to the velocity (or flow) field. The streamline describes the path of a massless particle in a flow field. The flow field in our work is constructed from the direction and magnitude of the fastest water diffusion component in each volume cell. The diffusion parameters are obtained by analyzing and averaging the motions of all water molecules throughout the time frame of a simulation, as described in the SI. Thus, each streamline represents the most probable path of a water molecule in a simulation. For clarity, we show only streamlines originating in the cells with a magnitude and anisotropy of diffusion higher than the predefined threshold.

Interestingly, the streamline analyses of the (i) empty protein, (ii) protein with prefolded substrate, (iii) protonated substrate, and (vi) product show that water molecules are moving in and out of the active site during the entire modeled reaction sequence independent of active site occupation. Figure 6 shows the flow of water in the active site of the empty protein as well when it contains substrate, protonated substrate, and

product. The flow of water is highest when the active site is empty, and the lowest flow is observed in the presence of the bulky product.

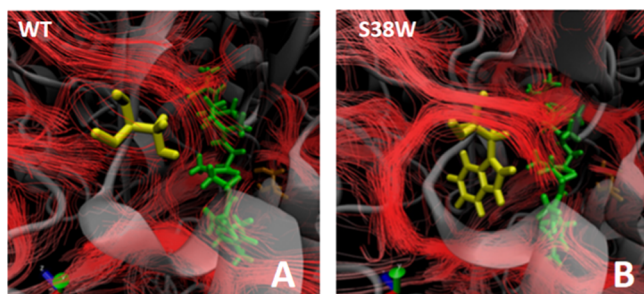
The rate-determining step of this reaction is the initial protonation of the substrate, as indicated by a significant deuterium kinetic isotope effect.<sup>23</sup> Interestingly, the streamline analysis shows a difference in the flow of water between the systems with the protonated and unprotonated substrate in the active site (Figure 6).

Syrén et al.,<sup>23</sup> on the basis of their CAVER analysis, introduced mutations to block the tunnels and thereby hypothetically hinder the flow of water to and from the active site. The temperature-dependent kinetic properties of these protein variants were determined experimentally and analyzed by means of Eyring transition state theory (Table 1). It was found that all tunnel-blocking mutations resulted in a lowering of the activation enthalpy ( $\Delta H^\ddagger$ ) and a lowering of the activation entropy ( $\Delta S^\ddagger$ ) compared to those of the wt, whereas some variants were faster and some slower than the wt at the normal reaction temperature (328 K).

To investigate the impact of the introduced mutations on the global water dynamics, a streamline analysis was initially performed with the S38W variant. Introducing a tryptophan at this tunnel lining position obstructs one of the (by CAVER) highest ranked tunnels in the protein,<sup>23</sup> which is associated with profound experimental effects on the temperature dependence, and with that the activation parameters of catalysis (Table 1).<sup>23</sup> Changing S38 to a tryptophan in silico forces water to move around the side chain of the bulkier tryptophan, which further impacts the motion of water molecules in the solvated active site (Figure 7).

The streamline analysis was subsequently performed on all of the protein variants; Figure 8 displays the flow of water in the active site for the different protein variants. As the activation parameters ( $\Delta H^\ddagger$ ,  $\Delta S^\ddagger$ , and  $\Delta G^\ddagger$ ) of the different protein variants were found to differ in the experimental study,<sup>23</sup> we focused the analysis on the region in the active site where the rate-determining step occurs, i.e., around the catalytic residue Asp376 and the substrate, as shown in Figure 8.

The streamline analysis of the wt and variants (Figure 8) shows that there are changes in the flow of water in the active site. In particular, there is an increased flow of water between the substrate and the catalytic residue Asp376 in S38W, S168F/S168W, and F605W compared to that in the wt and V440F. This is surprising, considering that the mutations were designed to hinder the flow of water in the tunnels. The streamline

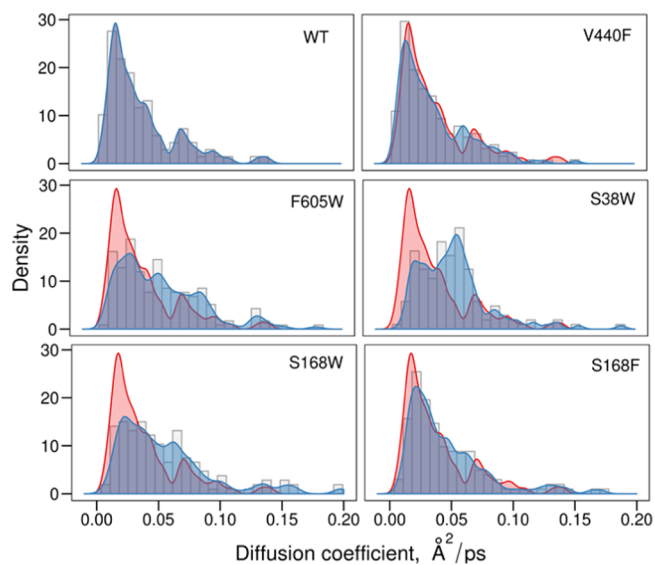


**Figure 7.** Streamline analysis on (A) the wt and (B) S38W, indicating the flow of water in the active site. The substrate (squalene) is indicated in green, and S38 (A) and S38W (B) in yellow. The flow of water is altered near the introduced mutation.

analysis indeed shows that the water paths in the vicinities of the mutated residues are altered, leading to changes in active site solvation.

**Analysis of Diffusion Constants in the Active Site.** The streamline analysis images (Figure 8) confirm that there are more water molecules moving in the region between the catalytic residue and the substrate in the mutated protein variants (except for control mutation V440F). An investigation of the distribution of the diffusion constants in this region around the catalytic residue shows further differences between the wt and protein variants. Figure 9 shows how the distributions of diffusion coefficients change for the wt and each protein variant in a 5.5 Å region around Asp376.

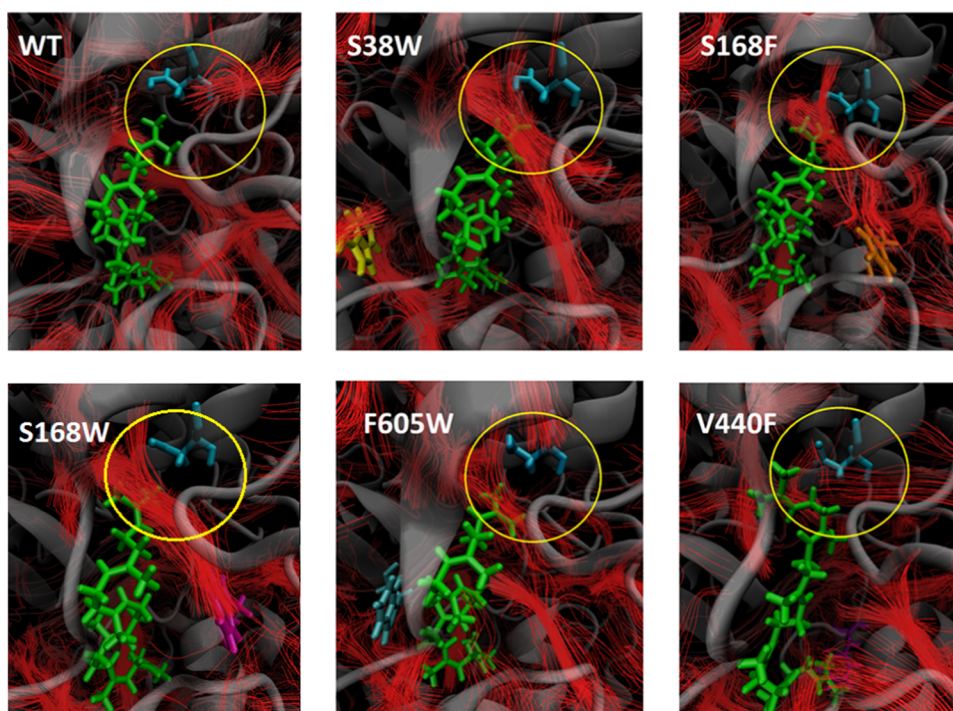
From Figure 9, it can be concluded that the wt and control mutation (V440F) have generally lower diffusion in the investigated region compared to that of the mutated protein variants F605W, S168F/W, and S38W. Although the Gibbs free energy of activation varies among the mutated variants, they all have in common that they have a much lower activation



**Figure 9.** Distribution of diffusion coefficients in a 5.5 Å region around the catalytic residue for the wt and protein variants (blue). The distribution of the wt is also inserted (red) for facilitated comparison with each protein variant.

entropy than that of the wt (Table 1). The S38W variant, which has its peak for the distribution most shifted toward the higher diffusion coefficient, has the most negative activation entropy. This is also the protein variant with the highest Gibbs free energy barrier at 328 K.

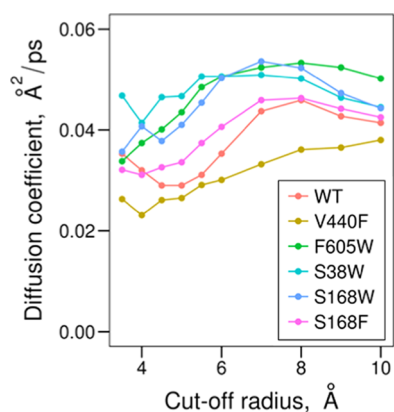
The introduced mutations do not only alter the location of the water streams in the active site, but also the diffusion in the catalytic region. Looking at the median values for the diffusion coefficients in Figure 12, the trend is sustained independently of the chosen cut-off region around Asp376 for the different



**Figure 8.** Streamlines around the substrate and the catalytic residue Asp376 (blue) in the wt, S38W, S168F, S168W, F605W, and V440F. The substrate (green) stays in the active site during the simulation. Each mutated residue is also indicated in each subplot. Gray: protein backbone.



protein variants. There is generally a lower flow in the active site of the wt and control mutation (V440F) compared to that of the other protein variants (Figure 10). This is an interesting observation that shows how altered water access pathways by enzyme design induce a reconfigured water flow.



**Figure 10.** Median values for the diffusion coefficients at different cut-off radius values around the catalytic residue for the different protein variants.

**Amount of Water around the Catalytic Residue and the Protonation Distance.** An analysis of the number of water molecules within 5 Å of the catalytic residue Asp376 over the course of the simulation confirms that the wt and control mutation (V440F) on average have fewer water molecules surrounding the substrate than the designed variants (Figure 11). The average number of water molecules within 5 Å of Asp376 is also higher in the variants as shown in Figure 11.

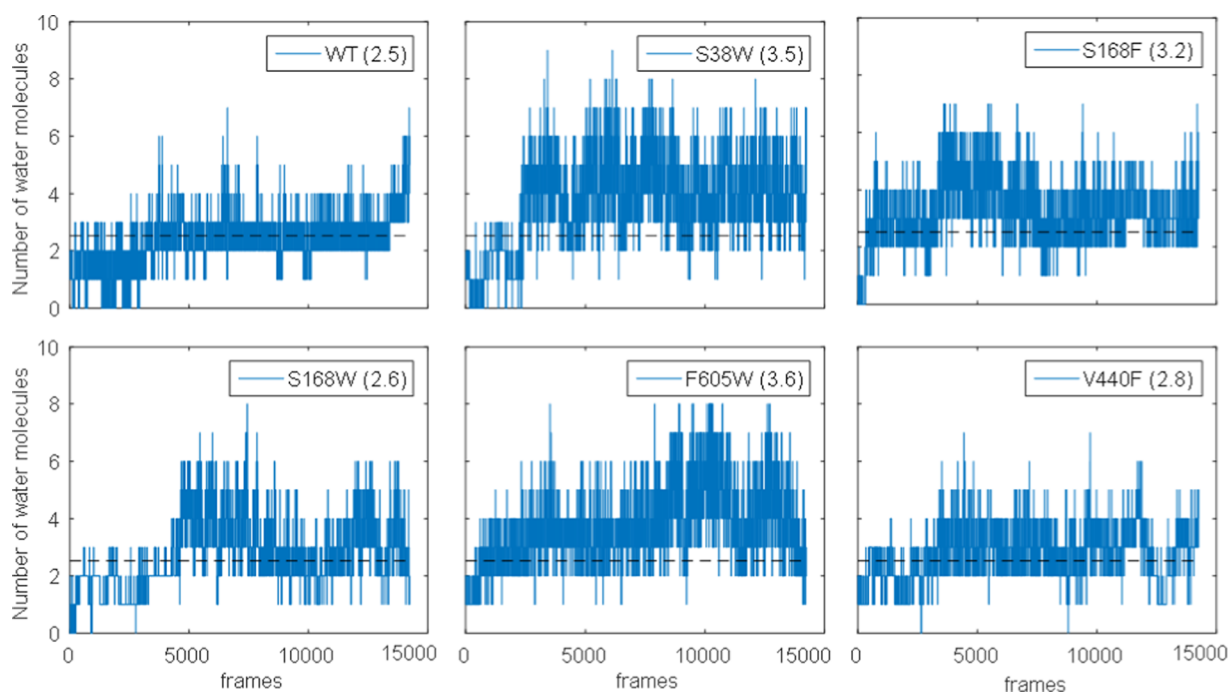
The water molecules in the active site were hypothesized to affect the rate-determining proton transfer; if there are water molecules between the catalytic residue and the terminal

isoprene group, the direct proton transfer from Asp376 to the substrate will be hindered. To quantify this effect, the distance between the acidic proton of Asp376 and the carbon on the substrate that receives the proton was compared for the different variants (Figure S2 in SI). We will refer to this distance as the “protonation distance”, although the actual distance of proton transfer is shorter by ca. 1.1 Å (the typical distance of a C–H bond).

The effect of the distance between the catalytic residue and the terminal isoprene group of the substrate on the catalytic efficiency of SHC has previously been investigated on the basis of shorter MD simulations.<sup>34,35</sup> Hammer et al. identified a correlation between the fraction of structures with a protonation distance of 2.9 Å or shorter and the catalytic activity for different unnatural substrates.<sup>35</sup> However, a similar correlation for the natural substrate and the different variants could not be found herein. The variations in the protonation distance along MD trajectories for the different protein variants are displayed in Figure 12.

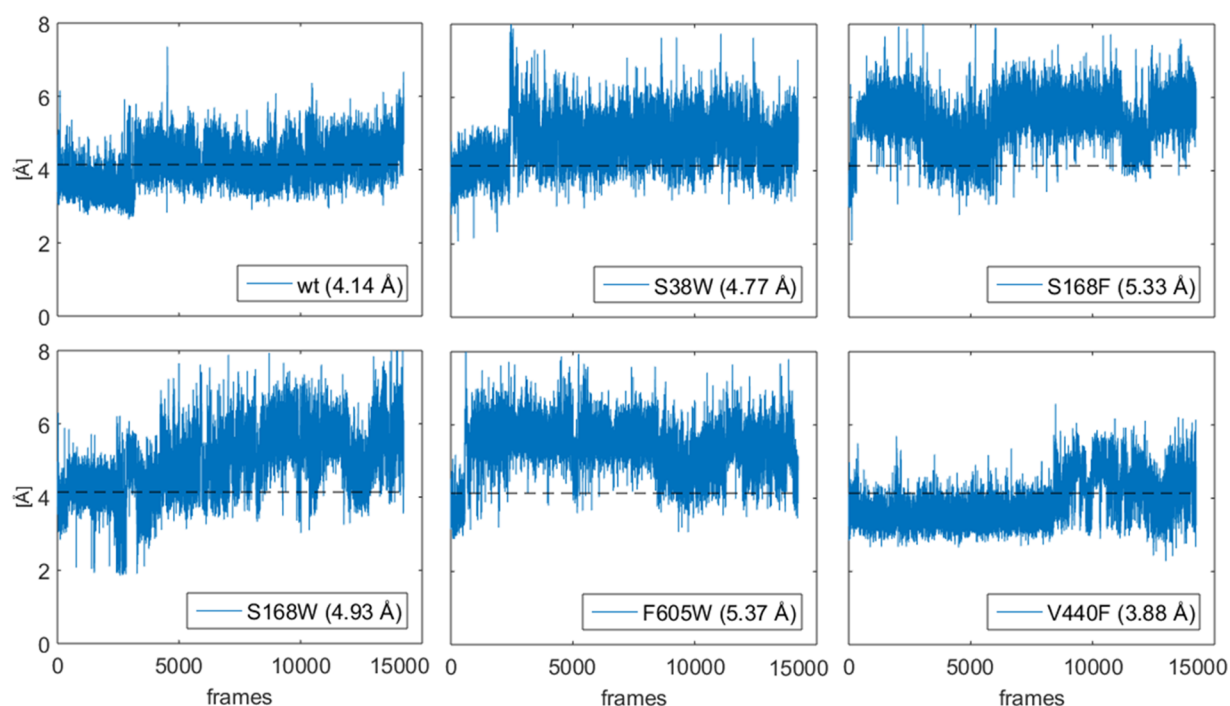
The different plots of Figure 12 show that the protonation distance on average is shorter in the wt and the control mutation V440F than in the other variants. Furthermore, the figure clearly demonstrates the importance of a longer simulation time for this type of analysis. The first 20–30 ns are not representative of the state that the system occupies for the remaining simulation time in the majority of these trajectories (V440 changes its behavior after a longer time, ca. 80 ns). Thus, in the subsequent analysis we removed the first 50 ns and used the last 100 ns for statistics.

Table 2 indicates a qualitative correlation between the average number of water molecules around Asp376 and the average protonation distance for the different variants. Thus, it seems that on average, the larger protonation distance in the designed variants is due to intercalation of water between Asp376 and the substrate.



**Figure 11.** Number of water molecules within 5 Å of the catalytic residue Asp376 over the trajectory. The dotted black lines in all subplots are the average value for the wt. The averages for each variant are stated in the legend of each subplot.





**Figure 12.** Protonation distance for the wt and protein variants. The dotted black lines in all figures are the average value for the wt. The averages for each variant are stated in the legend of each subplot.

**Table 2.** Average Value of Number of Water Molecules and Average Protonation Distance between the Catalytic Residue Asp376 and the Substrate for the wt and Protein Variants, the Fraction of Structures Identified as Near Attack Conformers (NACs), and the Corresponding  $\Delta G_{\text{NAC}}$ <sup>a</sup>

	average number of water molecules around Asp376	average protonation distance (Å)	fraction of structures ( $p_{\text{NAC}}$ in %) with a protonation distance shorter than 3.0 (3.5) Å	$\Delta G_{\text{NAC}}$ (kcal/mol) for a protonation distance shorter than 3.0 (3.5) Å
wt	2.9	4.29	0.04 (4.03)	5.2 (2.2)
S38W	4.0	4.88	0.01 (0.63)	6.2 (3.4)
S168F	3.2	5.45	0.00 (0.24)	>6.2 (4.0)
S168W	3.1	5.29	0.00 (0.42)	>6.2 (3.7)
F605W	4.0	5.37	0.00 (0.05)	>6.2 (5.1)
V440F	2.6	3.67 <sup>b</sup>	2.21 <sup>b</sup> (41.88) <sup>b</sup>	2.6 <sup>b</sup> (0.6) <sup>b</sup>

<sup>a</sup>The average values were calculated from the last 100 ns of the simulation trajectories. <sup>b</sup>The V440F protonation distance becomes on average longer after around 80 ns, and therefore the average protonation distance is likely to be overestimated and  $\Delta G_{\text{NAC}}$  underestimated for V440F in the table (see Figure 12).

On the basis of the concept of near attack conformers (NAC),<sup>36,37</sup> the free energy of activation for the catalytic reaction has a contribution from the free energy of NAC-formation ( $\Delta G_{\text{NAC}}$ ), i.e., the system has to pass a NAC before it can reach the transition state, and the free energy cost of reaching the NAC is a fraction of the total activation free energy.  $\Delta G_{\text{NAC}}$  is computed from the fraction of structures labeled as NACs ( $p_{\text{NAC}}$ ) by  $\Delta G_{\text{NAC}} = -RT \ln(p_{\text{NAC}})$ . We have previously shown<sup>38,39</sup> that  $\Delta G_{\text{NAC}}$  can be used in enzyme design to rank the catalytic efficiencies of different protein variants for a particular reaction. Herein, we define the NAC as a structure where the distance between the acidic hydrogen of Asp376 and the accepting carbon of the substrate is within the sum of the van der Waals radii of the interacting atoms.

Hammer et al. used a similar definition in their study of unnatural substrates, but they used the term transition state structure instead of NAC to denote the relevant structures.<sup>35</sup>

Studies of interactions in solid crystals have found that the mean non-bonding H–C distance of organic compounds is close to 3.0 Å, and we used this distance as an estimate of the H–C van der Waals distance and the NAC criterion in our analysis (the sum of the Bondi van der Waals radii of H and C gives 2.9 Å).<sup>40</sup> However, as this criterion gives a very low NAC frequency for the protein variants that results in low precision, we have, for comparison, also used a slightly longer H–C distance (3.5 Å) as the NAC criterion (Table 2).

The relatively short average protonation distance and the relatively high frequency of NACs reported in Table 2 for the wt and the control mutant are consistent with a direct proton transfer between Asp376 and the substrate. On the other hand, the tunnel-blocking variants consistently have a larger average protonation distance and a much lower, in most cases negligible, NAC frequency; observations that seem to preclude a direct proton transfer mechanism in these variants. There is no overall correlation between  $\Delta G_{\text{NAC}}$  and the free energy of activation for the catalytic reaction because some of the designed variants are faster than the wt and others are slower at 328 K. On the other hand, a low NAC frequency was found to be associated with a negative activation entropy; the designed variants in all cases have a negative activation entropy whereas it is positive for the wt. One plausible explanation for these observations is that the tunnel-obstructing variants operate to some extent by a different mechanism than that of the wt. It can be noted that using the longer H–C distance as the NAC criterion, as expected, decreases  $\Delta G_{\text{NAC}}$ , but the relative difference in  $\Delta G_{\text{NAC}}$  ( $\Delta \Delta G_{\text{NAC}}$ ) between variants shows similar variations as with the shorter H–C distance. Thus, the qualitative interpretation of the results is not sensitive to the NAC criterion.

We can speculate that the variants, in addition to direct proton transfer, can operate via a Grotthuss-like mechanism where the transfer of a proton from Asp376 to the substrate is mediated via one or more adjacent water molecules. Such a mechanism will lead to a lower activation entropy as it will require the positioning of a water molecule during the proton transfer, but it could reduce the activation enthalpy, and in principle, also the activation free energy. With the exception of S168W (where the deuterium kinetic isotope effect (DKIE) was 1.0), the observed DKIE of the wt and the variants lies in the range 1.6–2.1 (Table 1); a DKIE of this size is consistent with both a direct proton transfer or a Grotthuss-like proton transfer, and thus the DKIE does not provide a method to distinguish between the two mechanisms.<sup>41,42</sup> In this context, it should be mentioned that a previous study reported an altered reaction mechanism by introducing mutations in a tunnel in haloalkane dehalogenase LinB.<sup>43</sup>

The introduced mutations could also affect the  $pK_a$  of Asp376, and thereby indirectly affect the reaction rate. The most plausible reasons behind a  $pK_a$  change are a change in the solvation of Asp376 or in the positions of the surrounding residues; in particular, the positions of His451, Asp374, and Asp377 and their protonation states are expected to be important for the  $pK_a$  of Asp376. However, it should be noted that there is not likely to be a simple relationship between the  $pK_a$  of Asp376 and the reaction rate. The barrier for proton transfer from Asp376 will decrease with decreasing  $pK_a$ , but if the  $pK_a$  is much lower than pH, only a tiny fraction of Asp376 will be protonated and the reaction rate will decrease significantly (the reaction rate is expected to reach a maximum at a pH close to the  $pK_a$ ).<sup>44</sup> Furthermore, the barrier is not only dependent on the  $pK_a$  but also on the distance that the proton has to be transferred, and mutations that change the  $pK_a$  may simultaneously change this distance. Still, it would be very interesting to experimentally investigate how the activation parameters and the DKIE of the wt and the variants vary with the pH. Unfortunately, computations of  $pK_a$  values of systems of this complexity are still affected with too great uncertainties to be directly applicable.

Water molecules have previously been recognized to play several different important roles in cell biology.<sup>45</sup> They impact the binding of ligands to receptors and can determine ligand selectivity.<sup>46</sup> A key role of water in the reprotonation of catalytically active amino acids has been acknowledged, and Wendt et al. proposed that the protonation of the catalytic acid in squalene–hopene cyclase from *A. acidocaldarius* likely occurs through a tunnel, identified in the crystal structure,<sup>30</sup> linking the catalytic aspartate with the bulk solvent via a chain of six hydrogen-bonded water molecules. Mutations of similar proton channels that mediate solvent-assisted “Grotthuss-like” proton transfers were shown to affect the catalytic activity of other enzymes, such as the enoyl-acyl carrier protein reductase from *S. aureus*, where the catalytic efficacy was decreased 10000-fold in the water channel variants compared to that in the wt.<sup>27</sup>

Moreover, water molecules are prone to intercepting highly-reactive intermediates to tune product distribution, which is of high relevance for terpene cyclases that form several carbocations in their catalyzed polycyclization reactions. A structural role of water molecules in terpene cyclase active sites has been discussed in light of a sesquiterpene cyclase from *Aspergillus terreus*. Water molecules lining the hydrophobic active site were suggested to direct product distribution by contributing to prefolding of the substrate.<sup>47</sup>

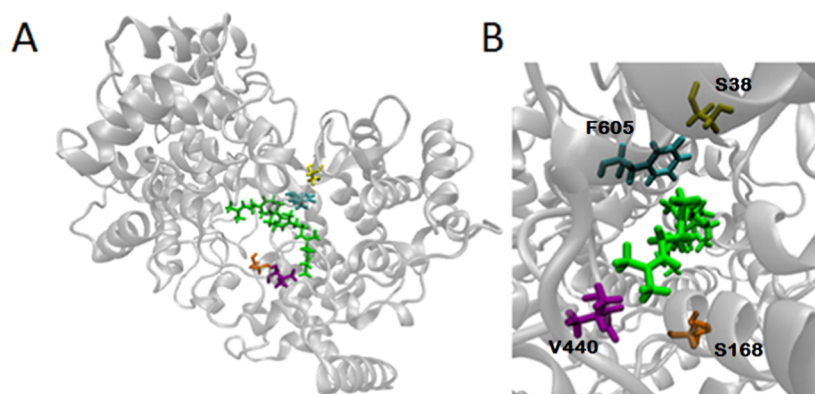
This study has provided an enhanced understanding of active site solvation throughout the reaction sequence of SHC. It has also shown how the introduced mutations, which were designed to obstruct the water paths leading to the active site, change the flow of water in the active site. The importance of water dynamics has also been studied recently by David et al.;<sup>48</sup> they introduced mutations in water channels that altered the dynamics of water access to the active site.

Surprisingly, the enzyme variants with obstructed water tunnels enable an increased flow of water in the active site by forcing water molecules to change path. This leads to an increase in the distance the proton has to travel in the rate-determining proton transfer step of the catalytic reaction. These variants, in contrast to the wt, all have a negative activation entropy, whereas their activation free energy varies, and for some variants, it is lower than that for the wt at 328 K.

We suggest that these observations can be explained by changes in the protonation step of the reaction mechanism in the variants, which may partially operate by a Grotthuss-like proton transfer. The SHC system seems to be very complex, and the previously identified changes in activation parameters are likely to arise from a combination of actions in the enzyme, and probably not solely from these identified changes in water flow with the associated altered protonation distance. It is possible that the introduced mutations also affect other global processes not investigated in this study. Further investigations will be necessary to fully understand the complex catalytic mechanism of SHC. In particular, we foresee that an investigation on how changes in pH affect the activation parameters and the deuterium kinetic isotope effect of the wt and variants can help elucidate variations in the mechanism.

## ■ CONCLUSIONS

Water molecules were found to move continuously between the bulk solution and the active site through specific paths in the protein; there are water molecules in the active site before the entering of the substrate and throughout the reaction. Protein variants previously designed to obstruct water paths leading to the active site by the introduction of bulky side chains were found to change the water path locally at the substitution site. Interestingly, these variants were found to be associated with an *increased* flow of water in the active site, particularly in the vicinity of the catalytic amino acid Asp376. Simulations with the substrate present in the active site indicate that the distance for the rate-determining proton transfer between Asp376 and the substrate is longer in the tunnel-obstructing protein variants than that in the wt and the control mutation V440F. However, the NAC analysis shows no correlation between the protonation distance and the activation free energy of the catalytic reaction at the simulation temperature. The common denominator for the tunnel-obstructing variants is that they have a negative activation entropy, whereas it is positive for the wt. On the basis of the observations of this study, we speculate that these variants could operate by a Grotthuss-like proton transfer in the rate-determining step, which is consistent with both the negative activation entropy and a larger protonation distance. Further experimental and theoretical studies will be needed to increase the mechanistic understanding of this complex enzyme.



**Figure 13.** (A) Location of introduced mutations S38 (yellow), S168 (orange), F605 (cyan), and control mutation V440 (purple). The substrate, squalene, is indicated in green. S38, S168, and F605 are considered to be in the second sphere of the active site, V440 is outside the active site. (B) A close-up of the active site and the location of the native residues and the substrate.

## METHODS

To investigate patterns of water movement in the protein, MD simulations were conducted with AMBER<sup>49</sup> on the bacterial SHC from *A. acidocaldarius* (PDB entries: 2SQC<sup>30</sup> and 1UMP<sup>17</sup>). The monomeric form (chain A) of the protein was selected for in silico analysis. The SHC structure has been reported as a dimer<sup>30</sup> in its native form, but it has been suggested that the dimerization might be of transient nature.<sup>20</sup>

Both crystal structures 2SQC and 1UMP contain an analogue of the substrate crystallized with the protein. The analogue is either localized in the channel in (2SQC) or in the active site (1UMP). The analogue molecules were replaced by squalene and hopene molecules (Figure 1) in preparation for the MD simulations.

The monotopic membrane protein SHC was attached to an 80 Å × 80 Å palmitoyl oleoyl phosphatidylcholine (POPC) membrane, constructed in visual molecular dynamics<sup>50</sup> (with charmm36 parameters). The binding residues (residues 216–234) were manually inserted into the membrane (Figure 2). All molecules (both POPC and associated water molecules) from the membrane that were within 3 Å of the protein were removed to avoid steric clashes.

The water molecules from the crystal structure were retained, and additional TIP3P water molecules were added by Leap<sup>51</sup> in a box extending 10 Å around the edge of the protein–membrane complex. In total, about forty thousand water molecules were included in the simulation box.

**Introduction of Mutations.** Mutations were introduced manually by replacing the affected residue in each PDB-file, and processing each file with Leap, to investigate whether introduction of bulky amino acids in place of the native residue with a small side chain changes the flow of water through the previously identified tunnels<sup>23</sup> in the protein.

Protein variants included mutations of S38, S168, and F605 to tryptophan and phenylalanine (S38W, S168F, S168W, and F605W) that block three of the previously identified tunnels<sup>23</sup> in the enzyme. In addition, we introduced a control mutation V440F (Figure 13). V440 is considered to be neither in the active site nor in a tunnel, and the kinetic behavior of the V440F mutant is similar to that of the wt.

**MD Simulations.** The simulations were performed with the AMBER software using force field ff14SB<sup>52</sup> for the proteins, lipid14<sup>53</sup> for the membrane, and generalized Amber force field (GAFF)<sup>54</sup> for the ligands. Parameters for the GAFF force field for the substrate and product were developed with ante-

chamber,<sup>51</sup> see SI for details. The simulations were around 150 ns long.

Some simulations were extended for up to 300 ns to study the processes during a longer time frame. Five parallel simulations of 20 ns length each were performed for further investigation of the sample space for the wt and the S38W mutant.

An additional simulation at higher temperature (350 K) of the protein with the product in the active site was performed to analyze whether the conformation of the hydrophobic channel would be more inclined to change to a wider diameter in order for the product to be able to be released from the active site to the membrane. A previous study<sup>20</sup> suggested that due to the high flexibility of the channel, it is likely that it also serves as an exit channel for release of the bulky product.

Simulations were performed with (i) the empty enzyme, (ii) the substrate (squalene) in different locations of the substrate entrance channel, (iii) squalene in the active site, for each of the protein variants, (iv) protonated squalene in the active site, and (v) product in the active site. More information about which PDB structure was used for each case can be found in the SI, Table S1. Post processing of the trajectories was performed using the CPPTRAJ program<sup>55</sup> in AmberTools.<sup>56</sup>

The stability of the protein outside its native environment of a bilayer lipid membrane was also investigated (under the same conditions as when the membrane was included, see SI for a more detailed description).

**Water Analysis.** The network of water tunnels in SHC was further investigated in silico with the use of different computational tools. CAVER<sup>22</sup> was developed to identify tunnels and pores in proteins. It was originally used for investigating these features in crystal structures or other static images of macromolecules. As CAVER is not measuring water dynamics, we complemented it with a method developed by Vassiliev et al.<sup>57,58</sup> to track the flow of water through a protein using streamline visualization of diffusion tensor fields. Previous water diffusion pathway studies have focused on proteins that have well-defined channels and a more simple geometry than SHC, such as aquaporins.<sup>59–62</sup> A recent study also investigated water dynamics in a de novo tunnel.<sup>63</sup> The data analysis technique applied in this study locates areas of highly anisotropic motions of water molecules from the simulation. It finds the most likely pathways of water diffusion in the protein based on the water motions detected in the MD simulations.



The streamline analysis was conducted on the 150 ns trajectories to investigate how the water molecules move in the enzyme. The water movements in the empty enzyme were compared to water movements in the enzyme with substrate, protonated substrate, and product in the active site. Water movements were also compared between the wt and the mutated protein variants. Details of streamline analysis are described in the SI.

## ■ ASSOCIATED CONTENT

### ● Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.7b01084.

Preparation of PDB files; detailed description of simulation procedure; RMSD plots for simulations; extended B-factor analysis for DXDD motif; details about streamline analysis method (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [tore@kth.se](mailto:tore@kth.se).

### ORCID

Serguei Vassiliev: 0000-0001-7351-3251

Tore Brinck: 0000-0003-2673-075X

### Author Contributions

C.G. performed simulations and analysis. The manuscript was written through contributions from all authors. All authors have given approval to the final version of the manuscript.

### Funding

C.G. and T.B. received funding from Swedish Research Council (VR), grant: 621-2013-04165. S.V. received funding from Discovery grant from the National Science and Engineering Research Council of Canada. C.K. and P.-O.S. received funding from Swedish Research Council (VR), grant: 621-2013-5138.

### Notes

The authors have nothing to disclose.

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The Swedish Research Council (VR) is acknowledged for financial support (T.B.), and Barbora Kozlikova and the CAVER team are greatly acknowledged for the time and effort to upgrade their program for these simulations.

## ■ REFERENCES

- (1) Oldfield, E.; Lin, F. Y. Terpene Biosynthesis: Modularity Rules. *Angew. Chem., Int. Ed.* **2012**, *51*, 1124–1137.
- (2) Wendt, K. U.; Schulz, G. E.; Corey, E. J.; Liu, D. R. Enzyme Mechanisms for Polycyclic Triterpene Formation. *Angew. Chem., Int. Ed.* **2000**, *39*, 2812–2833.
- (3) Potter, K. C.; Zi, J.; Hong, Y. J.; Schulte, S.; Malchow, B.; Tantillo, D. J.; Peters, R. J. Blocking Deprotonation with Retention of Aromaticity in a Plant Ent-Copalyl Diphosphate Synthase Leads to Product Rearrangement. *Angew. Chem., Int. Ed.* **2016**, *55*, 634–638.
- (4) Chen, M.; Harris, G. G.; Pemberton, T. A.; Christianson, D. W. Multi-Domain Terpenoid Cyclase Architecture and Prospects for Proximity in Bifunctional Catalysis. *Curr. Opin. Struct. Biol.* **2016**, *41*, 27–37.
- (5) Sáenz, J. P.; Grosser, D.; Bradley, A. S.; Lagny, T. J.; Lavrynenko, O.; Broda, M.; Simons, K. Hopanoids as Functional Analogues of Cholesterol in Bacterial Membranes. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 11971–11976.

- (6) Thimmappa, R.; Geisler, K.; Louveau, T.; O'Maille, P.; Osbourn, A. Triterpene Biosynthesis in Plants. *Annu. Rev. Plant Biol.* **2014**, *65*, 225–257.

- (7) Ajikumar, P. K.; Xiao, W.-H.; Tyo, K. E. J.; Wang, Y.; Simeon, F.; Leonard, E.; Mucha, O.; Phon, T. H.; Pfeifer, B.; Stephanopoulos, G. Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia coli*. *Science* **2010**, *330*, 70–74.

- (8) Ro, D.-K.; Paradise, E. M.; Ouellet, M.; Fisher, K. J.; Newman, K. L.; Ndungu, J. M.; Ho, K. A.; Eachus, R. A.; Ham, T. S.; Kirby, J.; et al. Production of the Antimalarial Drug Precursor Artemisinic Acid in Engineered Yeast. *Nature* **2006**, *440*, 940–943.

- (9) Jiang, Z.; Kempinski, C.; Bush, C. J.; Nybo, S. E.; Chappell, J. Engineering Triterpene and Methylated Triterpene Production in Plants Provides Biochemical and Physiological Insights into Terpene Metabolism. *Plant Physiol.* **2016**, *170*, 702–716.

- (10) Peralta-Yahya, P. P.; Zhang, F.; del Cardayre, S. B.; Keasling, J. D. Microbial Engineering for the Production of Advanced Biofuels. *Nature* **2012**, *488*, 320–328.

- (11) Nestl, B. M.; Hammer, S. C.; Nebel, B. A.; Hauer, B. New Generation of Biocatalysts for Organic Synthesis. *Angew. Chem., Int. Ed.* **2014**, *53*, 3070–3095.

- (12) Sukumaran, J.; Hanefeld, U. Enantioselective C-C Bond Synthesis Catalysed by Enzymes. *Chem. Soc. Rev.* **2005**, *34*, 530–542.

- (13) Seitz, M.; Syrén, P. O.; Steiner, L.; Klebensberger, J.; Nestl, B. M.; Hauer, B. Synthesis of Heterocyclic Terpenoids by Promiscuous Squalene-Hopene Cyclases. *ChemBioChem* **2013**, *14*, 436–439.

- (14) Hammer, S. C.; Syrén, P.; Seitz, M.; Nestl, B. M.; Hauer, B. Squalene Hopene Cyclases: Highly Promiscuous and Evolvable Catalysts for Stereoselective C-C and C-X Bond Formation. *Curr. Opin. Chem. Biol.* **2013**, *17*, 293–300.

- (15) Yoosuf-Aly, Z.; Faraldos, J. A.; Miller, D. J.; Allemann, R. K. Chemoenzymatic Synthesis of the Alarm Pheromone (+)-Verbenone from Geranyl Diphosphate. *Chem. Commun.* **2012**, *48*, 7040.

- (16) Seckler, B.; Poralla, K. Characterization and Partial Purification of Squalene - Hopene. *Biochim. Biophys. Acta* **1986**, *881*, 356–363.

- (17) Reinert, D. J.; Balliano, G.; Schulz, G. E. Conversion of Squalene to the Pentacarbocyclic Hopene. *Chem. Biol.* **2004**, *11*, 121–126.

- (18) Dang, T.; Prestwich, G. D. Site-Directed Mutagenesis of Squalene-Hopene Cyclase: Altered Substrate Specificity and Product Distribution. *Chem. Biol.* **2000**, *7*, 643–649.

- (19) Wendt, K. U.; Poralla, K.; Schulz, G. E. Structure and Function of a Squalene Cyclase. *Science* **1997**, *277*, 1811–1815.

- (20) Schwab, F.; van Gunsteren, W. F.; Zagrovic, B. Computational Study of the Mechanism and the Relative Free Energies of Binding of Anticholesteremic Inhibitors to Squalene-Hopene Cyclase. *Biochemistry* **2008**, *47*, 2945–2951.

- (21) Rajamani, R.; Gao, J. Balancing Kinetic and Thermodynamic Control: The Mechanism of Carbocation Cyclization by Squalene Cyclase. *J. Am. Chem. Soc.* **2003**, *127*, 12768–12781.

- (22) Chovancova, E.; Pavelka, A.; Benes, P.; Strnad, O.; Brezovsky, J.; Kozlikova, B.; Gora, A.; Sustr, V.; Klvana, M.; Medek, P.; et al. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* **2012**, *8*, No. e1002708.

- (23) Syrén, P.-O.; Hammer, S. C.; Claasen, B.; Hauer, B. Entropy Is Key to the Formation of Pentacyclic Terpenoids by Enzyme-Catalyzed Polycyclization. *Angew. Chem., Int. Ed.* **2014**, *53*, 4845–4849.

- (24) Helms, V.; Wade, R. C. Thermodynamics of Water Mediating Protein-Ligand Interactions in Cytochrome P450cam: A Molecular Dynamics Study. *Biophys. J.* **1995**, *69*, 810–824.

- (25) Pavlova, M.; Klvana, M.; Prokop, Z.; Chaloupkova, R.; Banas, P.; Otyepka, M.; Wade, R. C.; Tsuda, M.; Nagata, Y.; Damborsky, J. Redesigning Dehalogenase Access Tunnels as a Strategy for Degrading an Anthropogenic Substrate. *Nat. Chem. Biol.* **2009**, *5*, 727–733.

- (26) Meyer, E. Internal Water Molecules and H-Bonding in Biological Macromolecules: A Review of Structural Features with Functional Implications. *Protein Sci.* **1992**, *1*, 1543–1562.

- (27) Schiebel, J.; Chang, A.; Merget, B.; Bommineni, G. R.; Yu, W.; Spagnuolo, L. A.; Baxter, M. V.; Tareilus, M.; Tonge, P. J.; Kisker, C.;

et al. An Ordered Water Channel in *Staphylococcus aureus* FabI: Unraveling the Mechanism of Substrate Recognition and Reduction. *Biochemistry* **2015**, *54*, 1943–1955.

(28) Fiset, O.; Páslack, C.; Barnes, R.; Isas, J. M.; Langen, R.; Heyden, M.; Han, S.; Schäfer, L. V. Hydration Dynamics of a Peripheral Membrane Protein. *J. Am. Chem. Soc.* **2016**, *138*, 11526–11535.

(29) Benkovic, S. J.; Hammes-Schiffer, S. A Perspective on Enzyme Catalysis. *Science* **2003**, *301*, 1196–1202.

(30) Wendt, K. U.; Lenhart, A.; Schulz, G. E. The Structure of the Membrane Protein Squalene-Hopene Cyclase at 2.0 Å Resolution. *J. Mol. Biol.* **1999**, *286*, 175–187.

(31) Kokkinidis, M.; Glykos, N. M.; Fadoulglou, V. E. Protein Flexibility and Enzymatic Catalysis. *Adv. Protein Chem. Struct. Biol.* **2012**, *87*, 181.

(32) Meiering, E. M.; Serrano, L.; Fersht, A. R. Effect of Active Site Residues in Barnase on Activity and Stability. *J. Mol. Biol.* **1992**, *225*, 585–589.

(33) Sato, T.; Hoshino, T. Functional Analysis of the DXDDTA Motif in Squalene-Hopene Cyclase by Site-Directed Mutagenesis Experiments: Initiation Site of the Polycyclization Reaction and Stabilization Site of the Carbocation Intermediate of the Initially Cyclized A-Ring. *Biosci. Biotechnol. Biochem.* **1999**, *63*, 2189–2198.

(34) Hammer, S. C.; Marjanovic, A.; Dominicus, J. M.; Nestl, B. M.; Hauer, B. Squalene Hopene Cyclases Are Protonases for Stereoselective Brønsted Acid Catalysis. *Nat. Chem. Biol.* **2015**, *11*, 121–126.

(35) Hammer, S. C.; Syrn, P.-O.; Hauer, B. Substrate Pre-Folding and Water Molecule Organization Matters for Terpene Cyclase Catalyzed Conversion of Unnatural Substrates. *ChemistrySelect* **2016**, *1*, 3589–3593.

(36) Bruice, T. C.; Lightstone, F. C. Ground State and Transition State Contributions to the Rate of Intramolecular and Enzymic Reactions. *Acc. Chem. Res.* **1999**, *32*, 127–136.

(37) Hur, S.; Bruice, T. C. The near Attack Conformation Approach to the Study of the Chorismate to Prephenate Reaction. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12015–12020.

(38) Linder, M.; Hermansson, A.; Liebeschuetz, J.; Brinck, T. Computational Design of a Lipase for Catalysis of the Diels-Alder Reaction. *J. Mol. Model.* **2011**, *17*, 833–849.

(39) Linder, M.; Johansson, A. J.; Olsson, T. S. G.; Liebeschuetz, J.; Brinck, T. Computational Design of a Diels-Alderase from a Thermophilic Esterase: The Importance of Dynamics. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 1079–1095.

(40) Rowland, R. S.; Taylor, R. Intermolecular Nonbonded Contact Distances in Organic Crystal Structures: Comparison with Distances Expected from van Der Waals Radii. *J. Phys. Chem.* **1996**, *100*, 7384–7391.

(41) Hwang, J. K.; Warshel, A. How Important Are Quantum Mechanical Nuclear Motions in Enzyme Catalysis? *J. Am. Chem. Soc.* **1996**, *118*, 11745–11751.

(42) Marx, D. Proton Transfer 200 Years after Von Grothuss: Insights from Ab Initio Simulations. *ChemPhysChem* **2006**, *7*, 1848–1870.

(43) Biedermannová, L.; Prokop, Z.; Gora, A.; Chovancová, E.; Kovács, M.; Damborský, J.; Wade, R. C. A Single Mutation in a Tunnel to the Active Site Changes the Mechanism and Kinetics of Product Release in Haloalkane Dehalogenase LinB. *J. Biol. Chem.* **2012**, *287*, 29062–29074.

(44) Hyland, L. J.; Tomaszek, T. A.; Meek, T. D. Human Immunodeficiency Virus-1 Protease. 2. Use of pH Rate Studies and Solvent Kinetic Isotope Effects To Elucidate Details of Chemical Mechanism. *Biochemistry* **1991**, *30*, 8454–8463.

(45) Ball, P. More than a Bystander. *Nature* **2011**, *478*, 467–468.

(46) Snyder, P. W.; Mecinovic, J.; Moustakas, D. T.; Thomas, S. W.; Harder, M.; Mack, E. T.; Lockett, M. R.; Héroux, A.; Sherman, W.; Whitesides, G. M. Mechanism of the Hydrophobic Effect in the Biomolecular Recognition of Arylsulfonamides by Carbonic Anhydrase. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 17889–17894.

(47) Chen, M.; Chou, W. K. W.; Al-Lami, N.; Faraldos, J. A.; Allemann, R. K.; Cane, D. E.; Christianson, D. W. Probing the Role of Active Site Water in the Sesquiterpene Cyclization Reaction Catalyzed by Aristolochene Synthase. *Biochemistry* **2016**, *55*, 2864–2874.

(48) David, B.; Irague, R.; Jouanneau, D.; Daligault, F.; Czjzek, M.; Sanejouand, Y. H.; Tellier, C. Internal Water Dynamics Control the Transglycosylation/Hydrolysis Balance in the Agarase (AgaD) of *Zobellia galactanivorans*. *ACS Catal.* **2017**, *7*, 3357–3367.

(49) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolosváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Kollman, P. A.; et al. *AMBER 14*; University of California: San Francisco, 2014.

(50) Humphrey, W.; Dalke, A.; Schulten, K. VMD—Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.

(51) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247–260.

(52) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

(53) Dickson, C. J.; Madej, B. D.; Skjerve, Å. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **2014**, *10*, 865–879.

(54) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(55) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *3084*–3095.

(56) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolosváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Wu, X.; Kollman, P. A.; et al. *AMBER 14 Reference Manual*; University of California: San Francisco, 2014.

(57) Vassiliev, S.; Zaraiskaya, T.; Bruce, D. Exploring the Energetics of Water Permeation in Photosystem II by Multiple Steered Molecular Dynamics Simulations. *Biochim. Biophys. Acta, Bioenerg.* **2012**, *1817*, 1671–1678.

(58) Vassiliev, S.; Comte, P.; Mahboob, A.; Bruce, D. Tracking the Flow of Water through Photosystem II Using Molecular Dynamics and Streamline Tracing. *Biochemistry* **2010**, *49*, 1873–1881.

(59) de Groot, B. L.; Grubmüller, H. Water Permeation across Biological Membranes: Mechanism and Dynamics of Aquaporin-1 and GlpF. *Science* **2001**, *294*, 2353–2357.

(60) De Groot, B. L.; Frigato, T.; Helms, V.; Grubmüller, H. The Mechanism of Proton Exclusion in the Aquaporin-1 Water Channel. *J. Mol. Biol.* **2003**, *333*, 279–293.

(61) Zhu, F.; Tajkhorshid, E.; Schulten, K. Theory and Simulation of Water Permeation in Aquaporin-1. *Biophys. J.* **2004**, *86*, 50–57.

(62) Chakrabarti, N.; Roux, B.; Pomès, R. Structural Determinants of Proton Blockage in Aquaporins. *J. Mol. Biol.* **2004**, *343*, 493–510.

(63) Brezovsky, J.; Babkova, P.; Degtjarik, O.; Fortova, A.; Gora, A.; Iermak, I.; Rezacova, P.; Dvorak, P.; Smananova, I. K.; Prokop, Z.; et al. Engineering a de Novo Transport Tunnel. *ACS Catal.* **2016**, *6*, 7597–7610.