Water will find a way: transport through narrow tunnels in hydrolases

Carlos Sequeiros-Borja[1,2], Aravind Selvaram Thirunavukarasu[1,2], Cedrix J. Dongmo Foumthuim[3], Jan Brezovsky[1,2]*

1 International Institute of Molecular and Cell Biology, Warsaw, Poland

2 Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland

3 National Institute of Nuclear Physics (INFN), Sezione di Roma Tor Vergata, Rome, Italy

Corresponding author: Jan Brezovsky, E-mail: janbre@amu.edu.pl, jbrezovsky@iimcb.gov.pl

ABSTRACT

An aqueous environment is necessary for life as we know it, and water is required for almost all biochemical processes at a molecular level. Proteins employ water molecules in many ways. Hence, proteins need to transport water molecules across their internal network of tunnels to reach the desired action sites, either inside them or acting as molecular pipes to control cellular osmotic pressure. Even though water is an essential player in enzymatic activity and stability, its transport has been mostly neglected, with water transport studies mainly focused on the transport across membrane proteins. The transport of molecules through a protein's tunnel network is hard to study experimentally, rendering molecular dynamics simulations the most popular approach to study such events. In this study, we focused on the transport of water molecules across three different α/β-hydrolases: haloalkane dehalogenase, epoxide hydrolase, and lipase. Employing 5 μs adaptive simulation per system, we observed that only a few tunnels were responsible for the majority of water transport in dehalogenase, contrasting with a higher diversity of tunnels in other enzymes. Interestingly, water molecules could traverse narrow tunnels with even sub-angstrom bottlenecks, which is surprising given the commonly accepted water molecule radii of 1.4 Å. Our analysis of the transport events in such narrow tunnels showed a markedly increased number of hydrogen bonds formed between the water molecules and protein, likely compensating for the steric penalty of the process. Overall, these commonly disregarded narrow tunnels accounted for ~20% of total water transport observed, highlighting the need to move past the standard geometrical limits on the functional tunnels to account for relevant transport processes properly. Finally, we showed how the obtained insights could be applied to explain the differences in a mutant of the human soluble epoxide hydrolase associated with a higher incidence of ischemic stroke.

KEYWORDS: water transport, protein tunnels, molecular dynamics simulations, hydrolases,

Introduction

An aqueous environment is necessary for life as we know it, and water is required for almost all biochemical processes at a molecular level. Water can affect the folding of proteins,[1–3] their dynamics,[4] and actively participate in chemical transformations.[5] Proteins can employ water molecules in different ways: as proton donors or receivers,[6] molecular stabilizers,[7] molecular lubricants enhancing protein dynamics,[8] and as significant contributors behind the enthalpy-entropy compensation in protein-ligand binding.[9] In this sense, proteins need to transport water molecules across their internal network of tunnels to reach the desired action sites, either inside them or acting as molecular pipes to control cellular osmotic pressure.

The transport of molecules through a protein's tunnel network has proven challenging to study experimentally, with a few exceptions.[10] Therefore, the most suitable approach to study such events is molecular dynamics (MD) simulations, a methodology with its own limitations. Nevertheless, the principal advantage of MD is that it provides a detailed atomic picture of the transport process itself, the tunnel pathways employed by the molecules, and even approximated transport rates.[11–13] Even though water is an essential player in enzymatic activity and stability, its transport has been somehow neglected, with water transport studies mainly focused on the transport across membrane proteins like aquaporins.[14–17]

Water transport at a molecular level has been extensively studied in the nanomaterials field, where researchers have been testing the particularities and limits of water transport through different molecular barriers.[18–21] On the other hand, in biological sciences, water transport in proteins is mainly focused on the water transport across membranes by specialized trans-membrane proteins, like aquaporins. Within these proteins, the so-called single-file transport of water can be considered the limit in space required for transport.[13,22,23] Here, water molecules are

arranged in a single lane and move across tunnels of at least one water molecule wide, a commonly accepted radius of 1.4 Å.[24–26]

Recent research has shown that the methods to evaluate a protein's tunnel network and the migration of small molecules do not always yield equal results. There, some water molecules could not be traced to a defined tunnel, showcasing either the method's limitations in detecting tunnels or that water molecules could traverse across the protein through inconspicuous empty spaces.[27] In this paper, we hypothesize that water transport in proteins can occur across considerably narrow tunnels, with their bottleneck radius below 1.4 Å, and that this contribution is not negligible. To test our hypothesis, we focus on the tunnel network across three members of the α/β-hydrolase superfamily, emphasizing the tunnels used for water transport. Following this, we characterize these narrow tunnels and their interactions with water molecules. We show that water molecules can overcome Van der Waals repulsion by an increase in the number of H-bonds. Next, to showcase the utility of the insights obtained, we analyze two variants of the human epoxide hydrolase, a wild type and a mutant product of a single nucleotide polymorphism (E470G), associated with an increased risk of ischemic stroke.[28] We show that these narrow tunnels can pinpoint structural differences in protein variants that are not easily discernible with traditional analyses. Finally, we suggest revisiting the current considerations on the parameters of functional tunnels.

Materials and Methods

Preparation of structures

The crystallographic structures for the systems were obtained from the Protein Data Bank[29] (PDB): haloalkane dehalogenase (Hal) from *Rhodococcus rhodochrous*[30] (PDB: 4E46), epoxide hydrolase I (Epx) from *Solanum tuberosum*[31] (PDB: 2CJP), open[32] and closed[33] states of lipase

(Lip) from *Diutina rugosa* (previously *Candida rugosa*) (PDB: 1CRL and 1TRH, respectively), and human epoxide hydrolase[34] (hEpx) (PDB: 4X6X). In the case of Epx, two chains corresponding to two distinct enzyme structures were present in the asymmetric unit of the respective PDB record. To provide more conformational diversity, both structures were taken through the follow-up modeling after synchronizing their composition by removing residue K2 present only in chain A. To obtain the E470G variant of hEpx, residue E470 was renamed to Glycine, and the side-chain atoms were removed from hEpx. All the structures were processed by removing all non-protein molecules except the crystallographic waters.

The initial protonation states were determined with the H++ 3.0 webserver[35] at pH 8.5[36] for Hal, 6.8[37] for Epx, 8.0[38] for Lip, and 7.4[39] for hEpx and E470G. For the Lip system, residue E208 was differentially protonated for the open and closed states, therefore, in the view of building a uniform topology, this residue was manually de-protonated in both conformations. For all proteins, water molecules were initially placed around the solute using a tandem approach based on the 3D Reference Interaction Site Model theory[40] (3D-RISM) and the Placevent[41] algorithm. All predicted 3D-RISM waters were subsequently combined with crystallographic waters, keeping only the water molecules at least 2 Å away from the protein using EDIAscorer,[42] a method to compute electron density for individual atoms in a crystal structure. The resulting system was then processed with the tleap module of the Amber18[43] package. Each protein was placed at the center of a periodic truncated octahedral box, with a distance of 10 Å away from the edges, and solvated with the 3-charge, 4-point rigid OPC water model.[44] Na+ and Cl- ions were initially added to neutralize the charge of the system, and subsequently to reach a salt concentration of 0.1 M. The MD simulations were performed with the pmemd.cuda module of Amber18,[45,46] employing the

ff14SB[47] force field. Finally, the Hydrogen Mass Repartitioning method[48] was applied to the topology to enable a simulation timestep of 4 fs.

MD protocol

The systems were subjected to five minimization cycles with a stepwise release of positional restraints of the protein atoms. The minimization cycles consisted of 100 steps of the steepest descent minimization algorithm, followed by 400 steps of the conjugated gradient. In the first minimization cycle, harmonic positional restraints were imposed on all heavy atoms of the protein with a force constant of 500 $kcal \cdot mol^{-1} \cdot Å^{-2}$. In subsequent cycles, restrictions were applied only to the backbone atoms with force constants of 500, 125, 25, and 0.0001 $kcal \cdot mol^{-1} \cdot Å^{-2}$ sequentially. After the minimization workflow, a short heating round in the canonical NVT ensemble was applied while keeping the protein's heavy atoms restrained with a force constant of 5 $kcal \cdot mol^{-1} \cdot Å^{-2}$. The heating cycle consisted of 20 ps of MD from 0 to 200 K using the Langevin thermostat with a collision frequency of 2 ps^{-1}, a coupling constant of 1 ps, and a simulation timestep of 4 fs. The long-range electrostatic interactions were computed using the Particle Mesh Ewald summation[49,50] and all bonds involving hydrogen atoms were constrained using the SHAKE[51] algorithm. Following, three cycles of equilibration were performed. First, in the NVT ensemble, the temperature was raised to the target value of 310 K in 100 ps, employing the same parameters as previously described. The temperature was kept constant for 900 ps and during the procedure, harmonic positional restraints were applied to the protein's heavy atoms with a force constant of 5 $kcal \cdot mol^{-1} \cdot Å^{-2}$. Next, 1 ns of MD was run in the NPT ensemble, pressure was controlled with the weak-coupling Berendsen barostat with a coupling constant of 1 ps and positional restraints were applied only to the backbone atoms using the same force constant as in the previous stage. Lastly, another 201 ns of MD in the NPT ensemble with the same settings as previous was performed,

however without any restrictions. The last snapshots from simulations of all variants of each system were employed as their initial seeding structures for the adaptive MD simulations.

The production stage was performed with the High-Throughput Molecular Dynamics[52] package (HTMD v1.13.10) and Amber18 as the MD engine. For HTMD settings, five parallel MD runs per epoch were set with ten epochs in total. The HTMD package allows the building of Markov State Models (MSM) after each epoch, then giving preference to the less sampled conformational states in the following rounds, thereby enhancing the sampling of the system.[52–54] To build the MSM, the dihedral angles of a selection of residues were employed, and such residues were selected in the following manner. Initially, the geometric center of the active site residues for each system was calculated. Residues N41, D106, W107, E130, and H272 for Hal; D105, Y154, Y234, D265, and H300 for Epx; G123, G124, S209, A210, E341, and H449 for Lip; and D335, Y383, Y466, and H524 for hEpx and E470G were used as centers of selection spheres (numbering according to their PDB id). The dihedral angles of all residues within a sphere of 12 Å radius for Hal, Epx, hEpx, and E470G and 18 Å radius for Lip were used as metric for MSM building, with a TICA[55] dimension of 3 and a lag time of 2. The equilibration phase in HTMD consisted of 2 short MD rounds of 250 ps in NVT and NPT ensembles, during which the systems were heated from 0 to 310 K with a Langevin thermostat and harmonic positional restraints to the backbone atoms with a force constant of 5 $kcal \cdot mol^{-1} \cdot Å^{-2}$. Finally, a 100 ns unrestrained MD was performed in the canonical NVT ensemble, with Berendsen barostat and a saving frequency of 10 ps.

Tunnel network and water transport analyses

The trajectories were analyzed with the cpptraj[56] module of Amber18, CAVER[57] v3.02, AQUA-DUCT[58] v1.0.11, TransportTools[59] v0.9.0, and Python3 *in-house* scripts. For the tunnel network calculations with CAVER, the divide-and-conquer approach[60] was employed, the probe radius was

set to 0.7 Å, shell radius to 3.0 Å for Hal, Epx, hEpx, and E470G, and 5.0 Å for Lip. As CAVER starting atoms for tunnel search: D106-CG, W107-CD2, F168-N and L246-N for Hal; W106-CD2, L181-N, Y235-CB and D265-CA for Epx; G123-O, E126-CG, G342-CA and T416-HA for Lip; Y466-CD1 and H524-CA for hEpx and E470G were used, and the remaining parameters were left as default. These atoms were selected for two reasons: first, their center of geometry presented one of the lowest root mean square fluctuation (RMSF), and second, this center of geometry was inside the deepest part of the cavities described in the literature for each protein.[61–70] Furthermore, to achieve consistent tunnel entrances, HDBSCAN v0.8.27[71] was employed at the tunnel endpoints for clustering with the following parameters: *cluster_selectioin_epsilon* of 1.5, *allow_single_cluster* to True, *min_samples* and *min_cluster_size* to 5. Then, the formed clusters were taken as independent tunnels, and the tunnels identified as noise were discarded. For water tracking analyses, the same atoms described earlier were used for the object definition in AQUADUCT, employing a sphere of 6.0 Å for Hal and 4.0 Å for Epx, Lip, hEpx, and E470G. The remaining parameters were left as default. In TransportTools, the average-link hierarchical clustering method with a threshold of 1.0 Å was employed, with the exact matching analysis for the assignment of transport events. All remaining settings were set as default.

For each water transport event assigned by TransportTools to a defined tunnel for at least 70 % of its duration, a post-processing step was performed to obtain more detailed information about the nature of the transport. Here, a tunnel is considered a collection of *tunnel spheres* along a central line, and in each frame of a transport event, the water molecule involved is assigned to the closest *tunnel sphere* (**Figure 1**). Following this idea, a transport event contains a collection of *event spheres* to which a water molecule was assigned, restricting a single *event sphere* per frame. Here, we have focused our following analyses on the frames where the water molecule was

assigned to the smallest sphere along the event; we will refer to this sphere as the *minimal event sphere* (**Figure 1**). For these selected frames, the radius of the *minimal event sphere* and the hydrogen bonds between the water molecule and the protein were considered. The H-bonds were calculated with cpptraj.

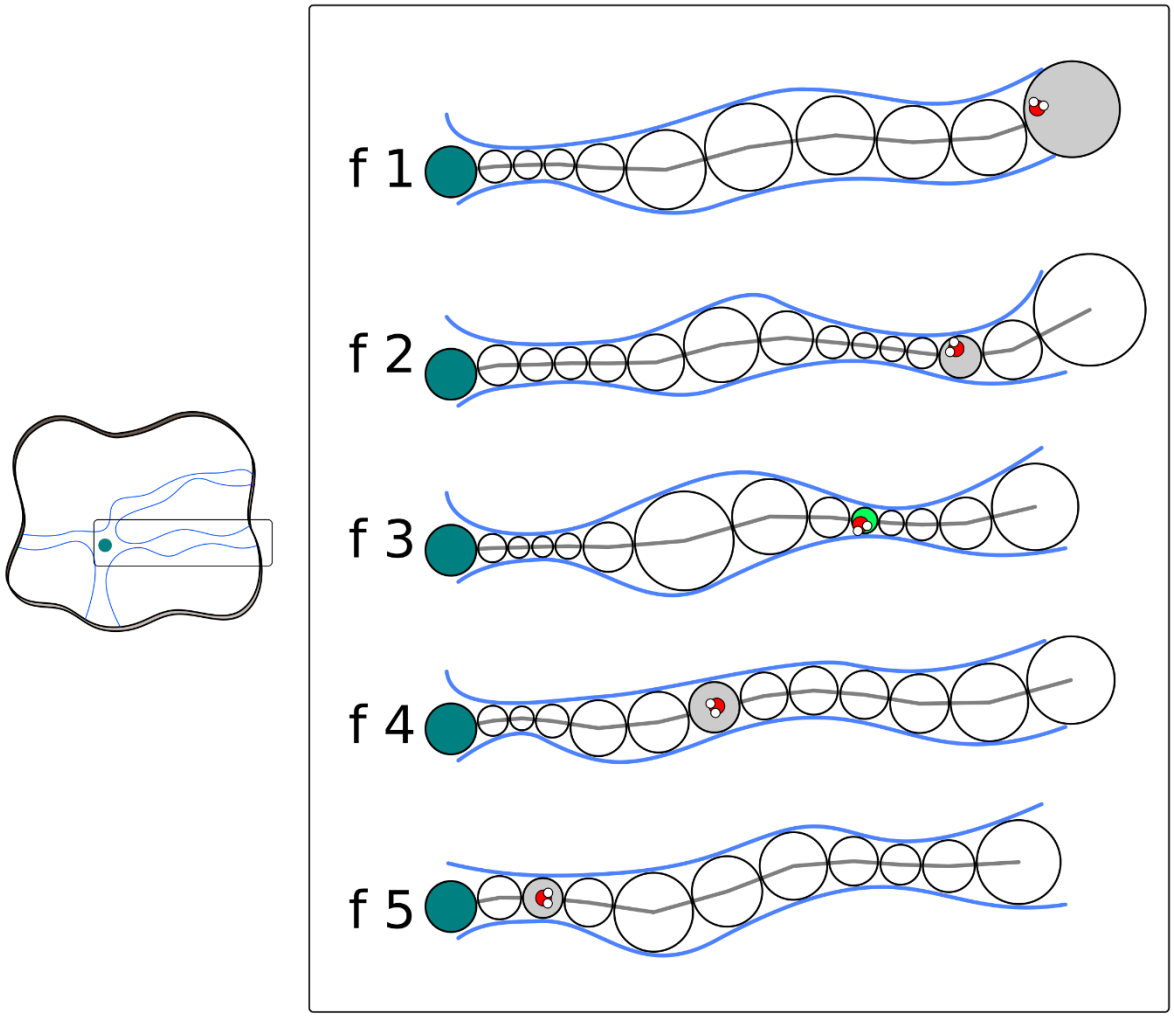

Figure 1. Water transport event and *minimal event sphere* definition. A protein can contain several tunnels and transport events (left), and in one tunnel, a transport event is composed of several frames (enclosed right). In each frame of a transport event, a tunnel is present as a set of *tunnel spheres* (empty circles) along a centerline (gray line) that joins the active site (turquoise circles) with the exterior. In each frame, the water molecules are assigned to a single *event sphere* of the tunnel (gray circles), and among all of them, the smallest corresponds to the *minimal event sphere* (green circle).

Results and Discussion

Tunnel networks identified in Hal, Epx, and Lip from adaptive MD simulations. The MD simulations for all the proteins showed a stable behavior during the 5 μs of simulation, with root mean square deviation (RMSD) averages of 1, 2, and 2.5 Å for Hal, Epx, and Lip, respectively (**Figures S1-S3**). Although these differences in RMSD might raise some concerns about the stability of the systems, the RMSF analyses showed that this is not the case. Not only because the sizes of the proteins are different, but in Hal, the protein is tightly packed and fluctuations are only occurring at the N- and C-terminal regions (**Figure S4**). Similarly, in Epx and Lip, some regions have considerably higher flexibility than the rest of the system (**Figures S5 and S6**). However, they correspond to regions with large b-factors in the crystals, attesting to high flexibility.

The results obtained with TransportTools for Hal show the presence of all tunnels described in the literature[61], namely *p1*, *p2a*, *p2b*, *p2c*, and *p3*[62] (**Figure S7a,b and Table S1**), with all of them used for water transport at some extent. The results also show new tunnels used for water transport: a tunnel near *p1* (called the *top* tunnel here), located closer to helix α5' (**Figure S7c**), and another tunnel perpendicular to the *p1* tunnel and opposite to the *p2* tunnel (called the *side* tunnel here), between the α/β and cap domains (**Figure S7d**). For Epx, the three main tunnels described by Mitusińska *et al.* are present,[63] however tunnel *TM1* is formed by two TransportTools superclusters (**Figure S8a,b and Table S1**). The results also show numerous new tunnels that transport water, albeit in considerably smaller quantities, which may partly correspond to the outliers observed by Mitusińska *et al.* (**Figure S8c,d and S9**). The Lip results showed the predominant presence of water transport of the *main* tunnel described in previous studies[64–67] (**Figure S10a,b and Table S1**). Additionally, several important tunnels for water transport were identified, albeit most are in the same region of the main tunnel (**Figure S10c,d**) and fewer in different regions (**Figure S11**).

The results also revealed the presence of the hypothesized *ester* exit tunnel,[68] however, only two water molecules were transported via this tunnel (**Figure S10a,b and Table S1**). The tunnel network analysis for the three systems showed the presence of all tunnels described in the literature. Furthermore, several tunnels not described before were found, mainly for Epx and Lip.

Water transport within the tunnel networks of Hal, Epx, and Lip. Overall, we have collected 2222, 17786, and 11086 water transport events that could be unambiguously assigned to tunnels found in Hal, Epx, and Lip, respectively (**Table S2**), constituting 23-84% of all transport events analyzed. Limited amount of events (<3%) could not be assigned even to the overall superclusters generated by TransportTools across all 50 simulations of each hydrolase. However, the predominant reason for the misalignment between the water events and tunnels originated from the attempts to exactly match these events to particular tunnels found by CAVER in the respective part of the trajectory in which the transport event occurred, which is in line with the recent observations for eight soluble epoxide hydrolases.[27] In general, the source of such mismatches can be traced to the approximation of asymmetric internal voids by spherical tunnels produced by CAVER,[72] which is further promoted by the fact that for tractability of the tunnel clustering, only the tunnel with the best throughput is maintained in each analyzed frame,[73] restricting the description of the entire tunnel geometry.

For the assigned events, the distribution of *minimal event sphere* radii for Hal showed a distribution with two peaks, with the major peak at 1.6-1.7 Å and a minor at 1.1-1.2 Å (**Figure 2**). A similar trend was observed for Epx, with a minor peak at 1.2-1.3 Å and a major one at 2.2-2.3 Å (**Figure 2**). However, this trend changed in Lip, where the curve showed a clear single peak at 1.5-1.6 Å (**Figure 2**). Observing these distributions by the tunnels involved, it was visible that the bimodal nature of Epx transport was mainly due to two tunnels (**Figure S12**). A somewhat similar

phenomenon was observed for Hal, in which two tunnels form a smaller peak around 1.0-1.2 Å (**Figure S12**). This bimodal behavior was not the case in Lip since all tunnels featured a peak at 1.5-1.6 Å (**Figure S12**). The cumulative distribution results displayed a sigmoidal behavior for all systems, although Epx deviated slightly due to its stronger bimodal nature (**Figure 2**). Considering 1.4 Å as the most widely accepted radius for a molecule of water,[24–26] the number of transport events below this threshold were 22.7%, 17.7%, and 25.0% for Hal, Epx, and Lip respectively (**Figure 2**), not a negligible quantity at all. For proteins specialized in water transport like aquaporins, the accepted minimum radius of a functional state is the size of a water molecule, resulting in the so-called single-file water transport.[22] If we consider the results obtained for the three proteins studied here, it is visible that water can be transported below the limits of specialized proteins. Hence, some explanations for these phenomena could be: although water could be transported through tunnels of radius below 1.4 Å, this process is not efficient, nor fast enough to fulfill the water requirements for the cell as aquaporins do; aquaporins could, in fact, transport water through narrower tunnel's radii too; finally, as aquaporins are specialized to control water transport across membranes, the requirements are different than inside a protein's tunnel network, hence the limitations may also be different. Considering these ideas and intrigued by these results, we focused on unraveling the characteristics and limits of water transport through tunnels of radius below 1.4 Å.

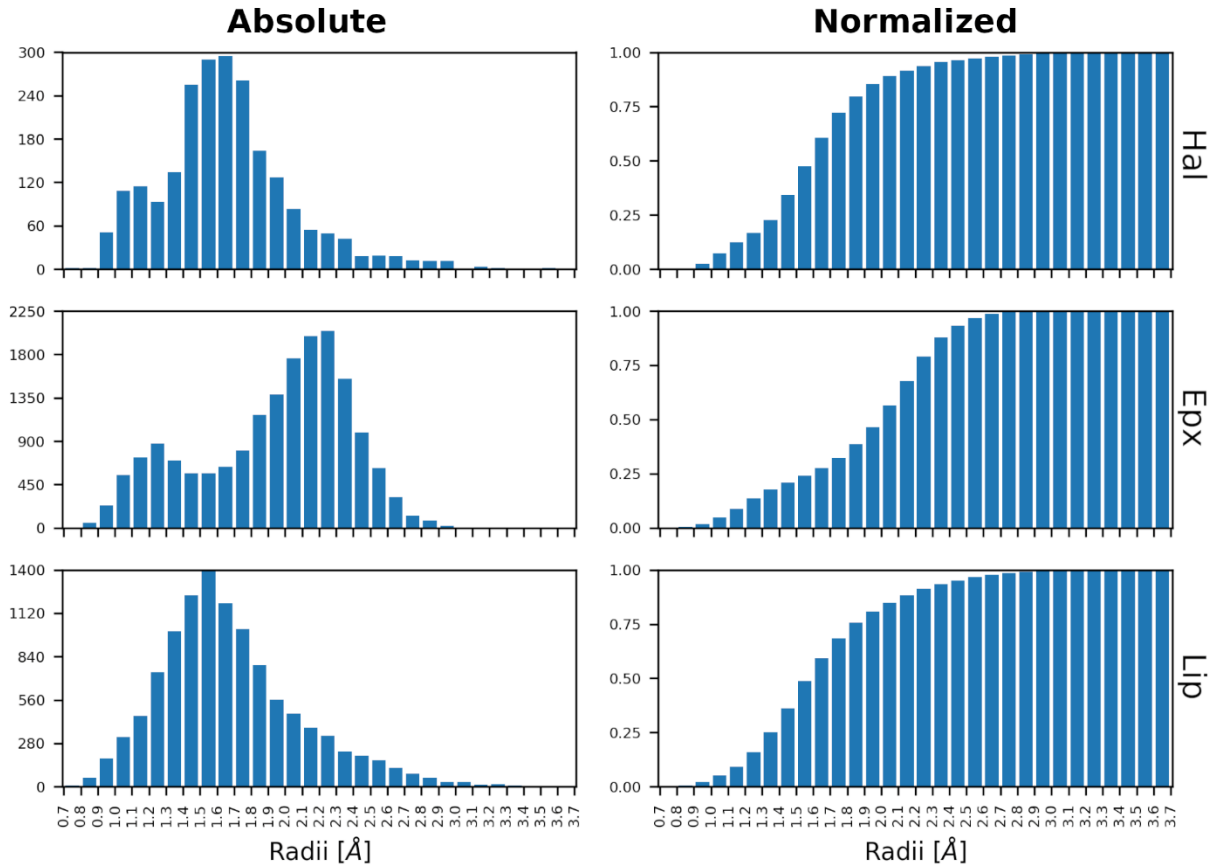

Figure 2. Distribution of *minimal event sphere radii* for water transport in Hal, Epx, and Lip. The absolute number of transport events and the corresponding normalized cumulative distribution, respectively.

Water transport through narrow tunnels. We analyzed H-bonds between the protein and water for each frame at which the water molecule was inside the *minimal event sphere*. Notably, water tends to make several H-bonds with the protein at a lower *minimal event sphere* radius, a common trend among the three systems studied here (**Figure 3**). When considering all radii for the analysis, this trend was maintained and further reinforced by the fact that the radius appears to have an inverse correlation with the number of H-bonds (**Figure S13**).

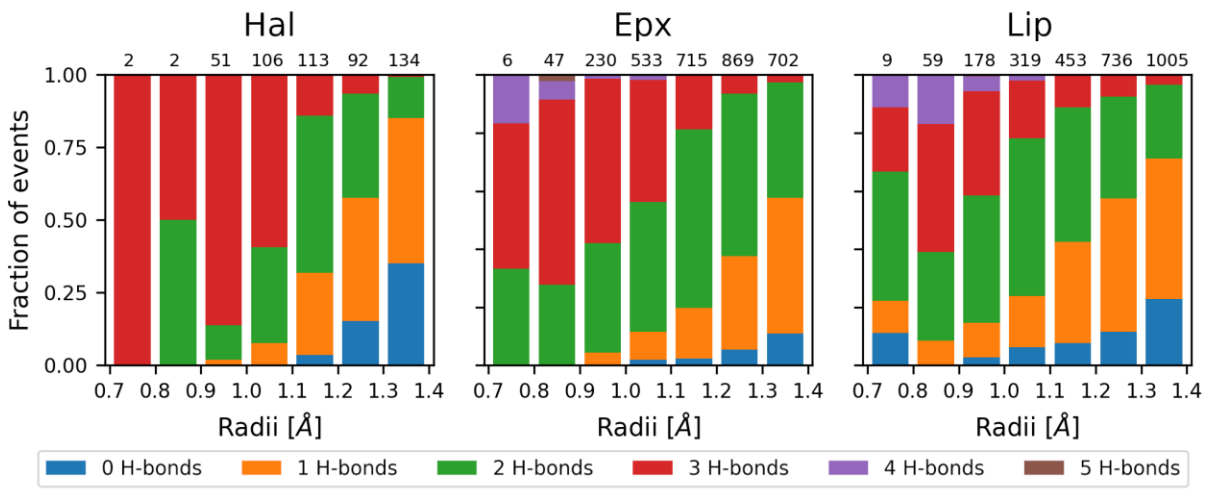

Figure 3. **Hydrogen bond distribution of water events at narrow tunnels.** H-bonds formed between water and protein at narrow tunnels below 1.4 Å bottleneck. Normalized distributions for Hal, Epx, and Lip. The total amount of events in each bin is shown at the top.

Hence, we hypothesized that the water molecules need to make several H-bonds to stabilize their transport through narrow tunnels to overcome the repulsive Van der Waals forces between water and protein atoms (**Figure 4**). Although not for transport but wetting, a similar behavior was observed in the mutational study of the trans-membrane protein 175 from *Chamaesiphon minutus*.[26] By mutating the hydrophobic gate residues Ile and Leu to the polar Asn, the researchers observed an increased wetting behavior at the gate, although the radius of the tunnel remained practically the same. On the other hand, it is striking to observe that the H-bond analysis performed here gives the opposite result that the one stipulated with the single-file water transport, where the presence of H-bond donors/acceptors along the whole tunnel length increases the energy barrier for water transport.[22,23] In nanomaterials, to our knowledge, a pore of radius 2.85 Å in a graphdiyne membrane was considered as the ultimate size limit for effective water transport.[20] Even though such a radius is twice that of a water molecule, in these artificial structures, their entirely hydrophobic nature prevents water from passing through pores physically open.[18]

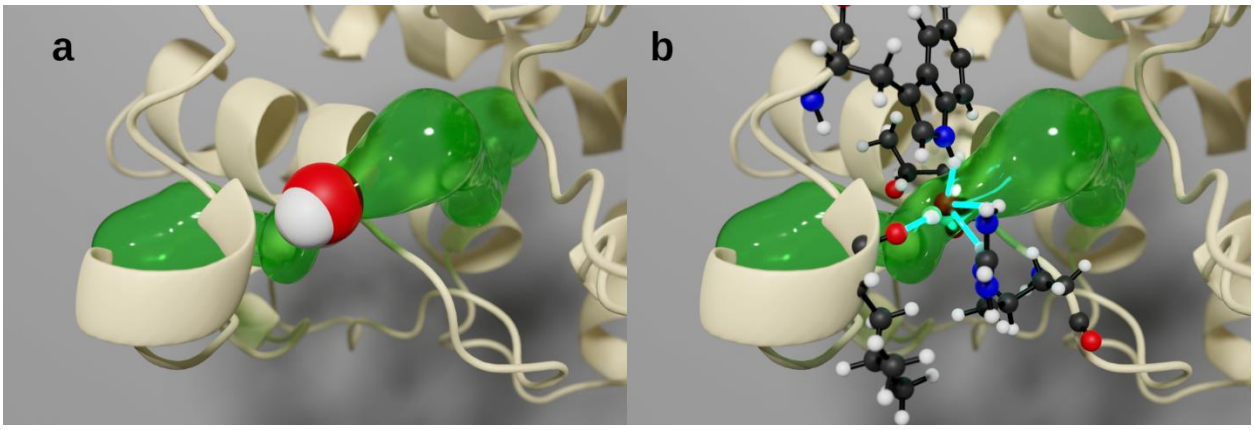

Figure 4. A w**ater molecule making hydrogen bonds while traversing a narrow tunnel.** a) Water molecule (Van der Waals volume) inside Hal (beige cartoon), traversing a narrow tunnel (green surface). The volume of the water molecule is larger than the available space in the tunnel. b) Same frame as a, but different representation. The water molecule (ball and sticks) inside Hal (beige cartoon), making several H-bonds (cyan sticks) with protein sidechain and backbone atoms (ball and sticks) while traversing a narrow tunnel (green surface).

Furthermore, the analysis of which atoms on the residues were responsible for the H-bonding network showed a predominance of the backbone atoms (**Figure 5**). On the complete distribution of events, no common trend was to be found, however, the backbone atoms were the most prevalent for *minimal event sphere* radii below 1.4 Å (**Figure S14**). Considering that for each H-bond analyzed, more than one atom could make contacts, the distributions changed visibly for Hal and Epx, not only in quantity but also in the position of the peaks (**Figure S14**). In the Lip case, only the quantity changed, keeping the peak almost at the same radius (**Figure S14**). Even though there is no clear trend on the type of atoms that make H-bonds on water transport, it is safe to conclude that backbone atoms play an important role, especially at low radii. With this in mind, we can hypothesize that water transport through considerably narrow tunnels is fold-specific and could be preserved inside protein families. The recent study on epoxide hydrolases supports this observation showing that tunnels are conserved structural features of proteins.[74] Although mutational studies on tunnels have shown changes in specificity[75] and catalytic rate[76], the targets were always the major tunnels, and the effects observed were related to substrates/products intake/release.

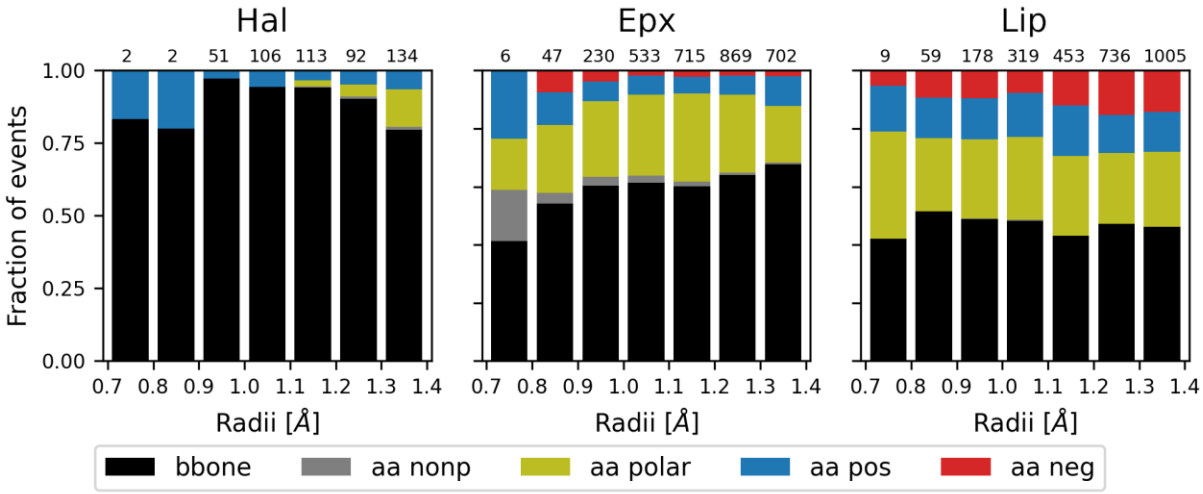

Figure 5. Distribution of atoms responsible for hydrogen bonding in water transport. Protein atoms involved in H-bonding with water in transport events for Hal, Epx, and Lip. Sidechain atoms were considered non-polar (Ala, Gly, Ile, Leu, Met, Phe, Pro, Trp, Val), polar (Asn, Cys, Gln, Ser, Thr, Tyr), positive (Arg, His, Lys) or negative (Asp, Glu) charged. The total amount of events for each bin is shown at the top.

Case study: human epoxide hydrolase and risk of ischemic stroke. To showcase the utility and importance of usually disregarded narrow tunnels, a comparative analysis between the human soluble epoxide hydrolase wild type (hEpx) and the variant E470G was performed. Since hEpx employs water as a co-substrate to hydrolyze epoxides to their corresponding diols, water is a vital component for hEpx's correct functioning.[77,78] hEpx is involved in the hydrolysis of epoxyeicosatrienoic acids, which have vasodilator actions and anti-inflammatory properties and are also analgesic, antifibrotic, and antihypertensive.[79] Polymorphisms of the gene that encodes hEpx have been linked to different enzyme activity rates,[39] and consequently also associated with incident ischemic stroke susceptibility or protection.[28] One of the variants that caught our attention (E470G) is correlated to an increased risk of ischemic stroke in the African-American population, but the molecular mechanism behind this correlation is not yet understood.[28]

Results from the MD simulations of 5 μs for each variant (hEpx and E470G) showed a minimal difference in RMSD on average (**Figure 6a**) and across all the simulations (**Figures S15 and S16**). Similarly, the average RMSF values of both variants did not considerably differ, except for region

S407-E446 (**Figure 6b**), which corresponds to a loop region on the cap domain of the protein (**Figure 6c**). The same region was considerably more flexible in E470G simulations when compared to hEpx (**Figures S17 and S18**), indicating increased flexibility produced by the mutation, albeit without the mutated position inside this region (**Figure 6b,c**). The distribution of *minimal event sphere* radii for hEpx and E470G shows a bell-shaped distribution, with a peak at ~1.6 Å in both cases (**Figure S19**). On the cumulative distribution of events, the proportion of events below 1.4 Å drops to 12.4 and 13.8 % in hEpx and E470G respectively (**Figure S19**), when comparing with Hal, Epx, and Lip. This reduction in transport events may be due to the considerable increase of total transport events between proteins (30 times increase compared with Hal and hEpx). The distribution of events by tunnels shows that some tunnels preferred narrow radii, shifting the distribution away from the major tunnel (**Figure S20**), although it was not as marked as for Epx or Hal (**Figure S12**). The H-bonding analyses present similar behavior and trends observed in Hal, Epx, and Lip, with a predominance of H-bonds at lower *minimal event sphere* radii (**Figure S21**) and the backbone atoms as major contributors to such contacts (**Figure S22**).

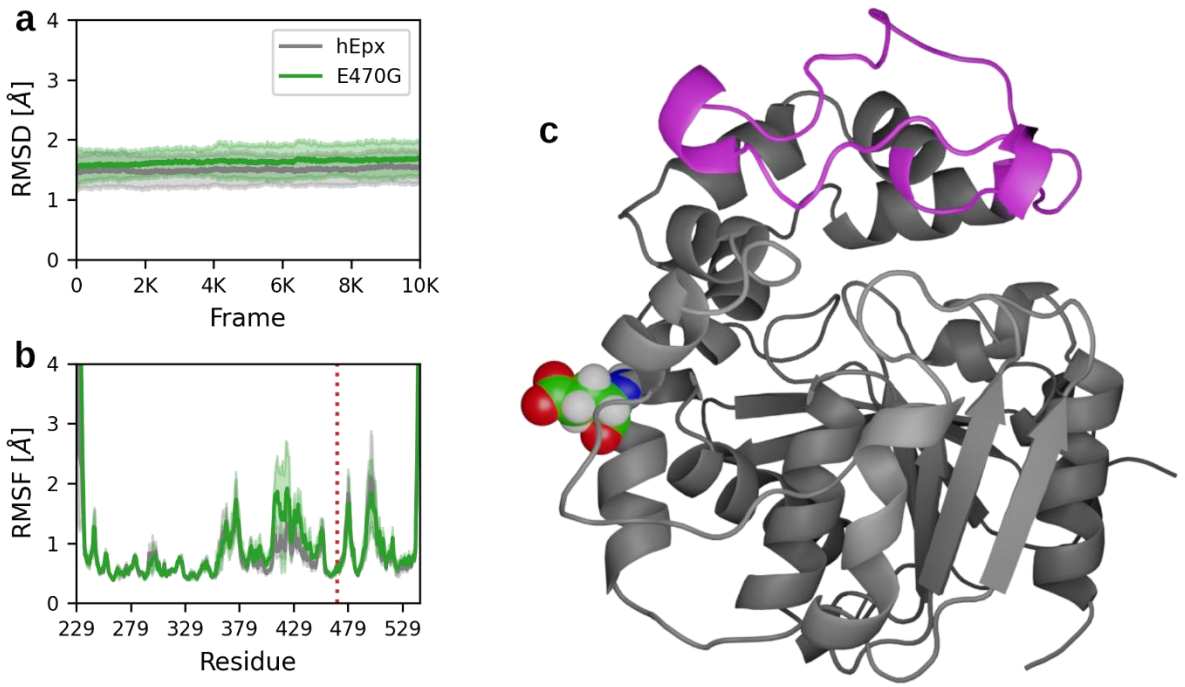

Figure 6. Stability and flexibility of human epoxide hydrolase wild type and E470G mutant. Average RMSD (a) and RMSF (b) are shown as solid lines with standard deviation as shaded regions for 5 μs adaptive MD simulations. The position of the mutation is marked with a vertical dotted red line in b. c) Cartoon representation of hEpx with the most flexible region in magenta and the mutated residue as spheres.

The MD results present all the tunnels in hEpx described in the literature,[63,69,70] including the four newly identified ones (**Table 1 and Figures S23-S25**). When we focused only on the tunnels that were used for water transport, the *Tc/m* tunnel was the major contributor in both variants (**Figure 7a and Table S3**), although its relevance was decreased in E470G compared to hEpx, by ~16%. Interestingly, in a recent study of epoxide hydrolases,[70] *Tm1* tunnel was found as the major contributor of water transport in the hEpx, however, this discrepancy can be the result of differences in simulation time and water model employed (TIP3P versus OPC).[80,81]

A view of the average bottleneck radius of the tunnels shows that most of them have a radius below the broadly accepted 1.4 Å radius for a water molecule (**Figure 7b and Table S3**). Although it is not surprising that the *Tm1* and *Tc/m* tunnels with large bottleneck radii transport approximately 78 and 64% of the total water in hEpx and E470G, respectively (**Table 1 and Table**

S3), the contribution of narrower tunnels should not be disregarded, especially in E470G, where the mutation seems to have modified the tunnel usage considerably (**Table 1 and Table S3**). In almost all tunnels, the volume of transported waters was increased in E470G compared to hEpx (**Table S3**), likely due to the reduction of transport via the *Tc/m* tunnel (**Figure 7a and Table 1**). The most interesting case induced by mutation is for a newly identified *Tnew_A* tunnel, whose bottleneck radius increased by 0.128 Å while its importance for water transport increased 20-fold (**Table 1 and Table S3**). The mutation also opened the *Tcap1* tunnel and increased the flow in the *Tside* tunnel in E470G (**Figure 7a and Table 1**). Similarly, a minimal reduction in the bottleneck radius of the *Tm5* tunnel by 0.085 Å in the E470G mutant resulted in the decreased transport by almost 25% of the wild-type hEpx (**Figure 7a and Table 1**).

Table 1. Utilization of tunnels found in hEpx and E470G mutant

Tunnel	**Tunnel IDs**	**Water transport**		
		WT	**E470G**	**Difference**
Tm1	2, 4, 8	22.37%	24.48%	-2.12%
Tm2	16	0.83%	2.45%	-1.62%
Tm3	3, 20, 37	4.66%	5.90%	-1.25%
Tm5 ■	6	5.14%	1.66%	**3.49%**
Tc/m	1, 29, 34	55.86%	39.63%	16.22%
Tg	12	0.76%	1.40%	-0.64%
Tside ●	5, 7	9.13%	17.49%	**-8.37%**
Tcap1 ⅄	14, 24	0.00%	2.27%	**-2.27%**
Tcap2	15	0.01%	0.10%	-0.08%
Tcap4	9, 39	0.09%	0.04%	0.05%
Tnew_A ★	10	0.21%	4.20%	**-3.98%**
Tnew_B	11	0.03%	0.02%	0.01%
Tnew_C	22	0.28%	0.16%	0.12%
Tnew_D	23	0.62%	0.19%	0.43%

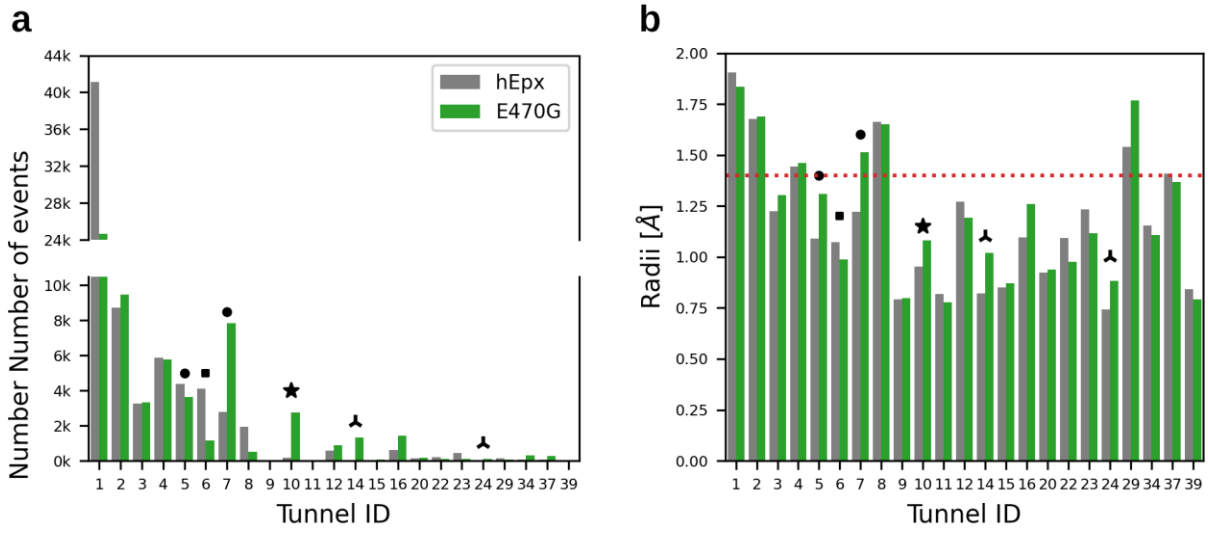

Figure 7. Tunnels employed by water in Human epoxide hydrolase wild type and E470G mutant. a) the Total number of transport events by the tunnel for Hepx and E470G. b) Average bottleneck radius for tunnels employed by water in Hepx and E470G, the accepted radius for a water molecule marked with a horizontal red dotted line. Narrow tunnels with altered water transport are marked with black shapes corresponding to the naming in **Table 1**.

We hypothesize that the altered water transport and tunnel usage by hEpx and E470G can lead to phenotypic changes observed in those variants. The reduced water flow in E470G could render a more proper environment for catalysis, resulting in increased activity of this variant observed in experimental tests.[39] Although a deep and comprehensive structural explanation for the differences observed in these two variants is out of the scope of this paper, we hypothesize that the differences in water flow observed can point forward to a proper explanation of their phenotypes. Additionally, the entry of a few water molecules into the active site can dramatically improve the overall efficiency of LinB86 dehalogenase,[82] where the new inflow of waters via the engineered narrow tunnel markedly promotes the rate-limiting product release. We want to clarify that the differences in water transport between hEpx and E470G cannot be considered a cause-effect to an increased risk of ischemic stroke. However, these differences are more insightful than a simple conformational analysis showing an increase in flexibility of some regions. Furthermore, the differences observed in tunnel usage were found in tunnels whose bottleneck radius could be easily

overlooked if the focus is only on wider tunnels. These results also increase the possible targets for protein engineering, where the opening or closing of tunnels has been successfully employed to alter relevant functional properties of enzymes.[82]

Conclusions

Through this paper, we have demonstrated that water molecules can traverse extremely narrow tunnels even of subangstrom radii. The amount of these events is roughly 20% in the proteins evaluated here, reconciling the discrepancy between geometry-based and ligand-tracking methods for tunnel identification.[27] We also showed that to overcome repulsion during the transport, water molecules create several H-bonds mainly with the backbone. Moreover, we found that the amount of H-bonds created is inversely proportional to the *minimal event sphere* radius. Although at this point, we lack sufficient data to support the extrapolation of these observations outside the α/β-hydrolase fold, our observation indicates the need to rethink the classical limits on the functional tunnels in enzymes, moving past the pure dimensions of their bottleneck. Finally, with the insights obtained, we propose a structural and functional hypothesis for the differences observed in the Human epoxide hydrolase and a single point mutant variant that could be traced to the minor alteration in the tunnel dimension resulting in major rewiring of the water transport flows through the active site.

ASSOCIATED CONTENT

Supporting Information. The following files are available free of charge. RMSD and RMSF plots for all Hal, Epx, Lip, hEpx an E470G adaptive simulations; Visualization of tunnels identified in adaptive simulations of Hal, Epx, Lip hEpx an E470G; Distribution of *minimal event sphere* radii in Hal, Epx, Lip, hEpx an E470G by tunnels, H-bonds, and interacting atoms; Distribution of

minimal event sphere radii in hEpx and E470G; Literature-based naming of the identified tunnels in Hal, Epx, and Lip; Water usage and average bottlenecks of tunnels in hEpx and E470G (PDF).

Data and Software Availability. The data underlying this study are available in the published article, the Supporting Information, and in Zenodo repositories at https://doi.org/10.5281/zenodo.7966082 (Hal), https://doi.org/10.5281/zenodo.7966059 (Epx), https://doi.org/10.5281/zenodo.7966092 (Lip), https://doi.org/10.5281/zenodo.7965661 (hEpx), and https://doi.org/10.5281/zenodo.7965881 (E470G). The data were deposited as Python3 scripts, binary files, plain text as well as PDB-formatted and AMBER-formatted structural data, which can be processed via numerous freely available SW packages. No tools with restricted access are needed.

AUTHOR INFORMATION

Corresponding Author

* janbre@amu.edu.pl or jbrezovsky@iimcb.gov.pl.

Author Contributions

C.S-B. set up, performed, and analyzed simulations of hEpx and E470G, analyzed water transport in all five systems, interpreted data, and prepared the figures and manuscript draft. A.S.T. set up, performed, and analyzed simulations of Hal. C.J.D.F. set up, performed, and analyzed simulations of Lip. J.B. conceived and coordinated the project, designed the calculation, set up, performed, and analyzed simulations of Epx, and interpreted data. The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

This work was supported by the National Science Centre, Poland (grant no. 2017/26/E/NZ1/00548). The computations were performed at the Poznan Supercomputing and Networking Center. C.S-B. and A.S.T. are recipients of a scholarship associated with the POWER project (grant nos. POWR.03.02.00-00-I022/16 and POWR.03.02.00-00-I006/17, respectively).

ABBREVIATIONS

Hal, haloalkane dehalogenase from *Rhodococcus rhodochrous*; Epx, potato epoxide hydrolase 1; Lip, lipase from *Diutina rugosa* (previously *Candida rugosa*); MD, molecular dynamics; hEpx, human epoxide hydrolase; E470G, E470G mutant of human epoxide hydrolase; MSM, Markov State Models; RMSD, root mean square deviation; RMSF, root mean square fluctuation.

REFERENCES

(1) Levy, Y.; Onuchic, J. N. Water Mediaton in Protein Folding and Molecular Recognition. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35* (1), 389–415. https://doi.org/10.1146/annurev.biophys.35.040405.102134.

(2) Rhee, Y. M.; Sorin, E. J.; Jayachandran, G.; Lindahl, E.; Pande, V. S. Simulations of the Role of Water in the Protein-Folding Mechanism. *Proc. Natl. Acad. Sci.* **2004**, *101* (17), 6456–6461. https://doi.org/10.1073/pnas.0307898101.

(3) Dongmo Foumthuim, C. J.; Giacometti, A. Solvent Quality and Solvent Polarity in Polypeptides. *Phys. Chem. Chem. Phys.* **2023**, *25* (6), 4839–4853. https://doi.org/10.1039/d2cp05214h.

(4) Bellissent-Funel, M. C.; Hassanali, A.; Havenith, M.; Henchman, R.; Pohl, P.; Sterpone, F.; Van Der Spoel, D.; Xu, Y.; Garcia, A. E. Water Determines the Structure and Dynamics of Proteins. *Chem. Rev.* **2016**, *116* (13), 7673–7697. https://doi.org/10.1021/acs.chemrev.5b00664.

(5) Pocker, Y. Water in Enzyme Reactions: Biophysical Aspects of Hydration-Dehydration Processes. *Cell. Mol. Life Sci.* **2000**, *57* (7), 1008–1017. https://doi.org/10.1007/pl00000741.

(6) Hassanali, A.; Giberti, F.; Cuny, J.; Kühne, T. D.; Parrinello, M. Proton Transfer through the Water Gossamer. *Proc. Natl. Acad. Sci.* **2013**, *110* (34), 13723–13728. https://doi.org/10.1073/pnas.1306642110.

(7) Li, G.; Wang, B.; Resasco, D. E. Water-Mediated Heterogeneously Catalyzed Reactions.

ACS Catal. **2020**, *10* (2), 1294–1309. https://doi.org/10.1021/acscatal.9b04637.
(8) Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2008**, *108* (1), 74–108. https://doi.org/10.1021/cr068037a.
(9) Breiten, B.; Lockett, M. R.; Sherman, W.; Fujita, S.; Al-Sayah, M.; Lange, H.; Bowers, C. M.; Heroux, A.; Krilov, G.; Whitesides, G. M. Water Networks Contribute to Enthalpy/Entropy Compensation in Protein–Ligand Binding. *J. Am. Chem. Soc.* **2013**, *135* (41), 15579–15584. https://doi.org/10.1021/ja4075776.
(10) Kokkonen, P.; Sykora, J.; Prokop, Z.; Ghose, A.; Bednar, D.; Amaro, M.; Beerens, K.; Bidmanova, S.; Slanska, M.; Brezovsky, J.; Damborsky, J.; Hof, M. Molecular Gating of an Engineered Enzyme Captured in Real Time. *J. Am. Chem. Soc.* **2018**, *140* (51), 17999–18008. https://doi.org/10.1021/jacs.8b09848.
(11) Zhu, F.; Tajkhorshid, E.; Schulten, K. Pressure-Induced Water Transport in Membrane Channels Studied by Molecular Dynamics. *Biophys. J.* **2002**, *83* (1), 154–160. https://doi.org/10.1016/S0006-3495(02)75157-6.
(12) Pluhackova, K.; Schittny, V.; Bürkner, P.; Siligan, C.; Horner, A. Multiple Pore Lining Residues Modulate Water Permeability of GlpF. *Protein Sci.* **2022**, *31* (10), e4431. https://doi.org/10.1002/pro.4431.
(13) Jensen, M. Ø.; Mouritsen, O. G. Single-Channel Water Permeabilities of Escherichia Coli Aquaporins AqpZ and GlpF. *Biophys. J.* **2006**, *90* (7), 2270–2284. https://doi.org/10.1529/biophysj.105.073965.
(14) de Groot, B. L.; Grubmüller, H. Water Permeation Across Biological Membranes: Mechanism and Dynamics of Aquaporin-1 and GlpF. *Science (80-.).* **2001**, *294* (5550), 2353–2357. https://doi.org/10.1126/science.1066115.
(15) Murata, K.; Mitsuoka, K.; Hirai, T.; Walz, T.; Agre, P.; Heymann, J. B.; Engel, A.; Fujiyoshi, Y. Structural Determinants of Water Permeation through Aquaporin-1. *Nature* **2000**, *407* (6804), 599–605. https://doi.org/10.1038/35036519.
(16) Saparov, S. M.; Pohl, P. Beyond the Diffusion Limit: Water Flow through the Empty Bacterial Potassium Channel. *Proc. Natl. Acad. Sci.* **2004**, *101* (14), 4805–4809. https://doi.org/10.1073/pnas.0308309101.
(17) Hadidi, H.; Kamali, R. Molecular Dynamics Study of Water Transport through AQP5-R188C Mutant Causing Palmoplantar Keratoderma (PPK) Using the Gating Mechanism Concept. *Biophys. Chem.* **2021**, *277*, 106655. https://doi.org/10.1016/j.bpc.2021.106655.
(18) Beckstein, O.; Biggin, P. C.; Sansom, M. S. P. A Hydrophobic Gating Mechanism for Nanopores. *J. Phys. Chem. B* **2001**, *105* (51), 12902–12905. https://doi.org/10.1021/jp012233y.
(19) Secchi, E.; Marbach, S.; Niguès, A.; Stein, D.; Siria, A.; Bocquet, L. Massive Radius-Dependent Flow Slippage in Carbon Nanotubes. *Nature* **2016**, *537* (7619), 210–213. https://doi.org/10.1038/nature19315.
(20) Xu, J.; Zhu, C.; Wang, Y.; Li, H.; Huang, Y.; Shen, Y.; Francisco, J. S.; Zeng, X. C.; Meng, S. Water Transport through Subnanopores in the Ultimate Size Limit: Mechanism from Molecular Dynamics. *Nano Res.* **2019**, *12* (3), 587–592. https://doi.org/10.1007/s12274-018-2258-7.
(21) Coalson, R. D. Driven Water/Ion Transport through Narrow Nanopores: A Molecular Dynamics Perspective. *Faraday Discuss.* **2018**, *209* (0), 249–257. https://doi.org/10.1039/c8fd00073e.
(22) Horner, A.; Pohl, P. Single-File Transport of Water through Membrane Channels. *Faraday*

Discuss. **2018**, *209* (0), 9–33. https://doi.org/10.1039/c8fd00122g.
(23) Pfeffermann, J.; Goessweiner-Mohr, N.; Pohl, P. The Energetic Barrier to Single-File Water Flow through Narrow Channels. *Biophysical Reviews*. Springer Science and Business Media Deutschland GmbH December 23, 2021, pp 913–923. https://doi.org/10.1007/s12551-021-00875-w.
(24) Li, A.-J.; Nussinov, R. A Set of van Der Waals and Coulombic Radii of Protein Atoms for Molecular and Solvent-Accessible Surface Calculation, Packing Evaluation, and Docking. *Proteins Struct. Funct. Genet.* **1998**, *32* (1), 111–127. https://doi.org/10.1002/(SICI)1097-0134(19980701)32:1<111::AID-PROT12>3.0.CO;2-H.
(25) Gerstein, M.; Chothia, C. Packing at the Protein-Water Interface. *Proc. Natl. Acad. Sci.* **1996**, *93* (19), 10167–10172. https://doi.org/10.1073/pnas.93.19.10167.
(26) Lynch, C. I.; Klesse, G.; Rao, S.; Tucker, S. J.; Sansom, M. S. P. Water Nanoconfined in a Hydrophobic Pore: Molecular Dynamics Simulations of Transmembrane Protein 175 and the Influence of Water Models. *ACS Nano* **2021**, *15* (12), 19098–19108. https://doi.org/10.1021/acsnano.1c06443.
(27) Mitusińska, K.; Bzówka, M.; Magdziarz, T.; Góra, A. Geometry-Based versus Small-Molecule Tracking Method for Tunnel Identification: Benefits and Pitfalls. *J. Chem. Inf. Model.* **2022**, *62* (24), 6803–6811. https://doi.org/10.1021/acs.jcim.2c00985.
(28) Fornage, M.; Lee, C. R.; Doris, P. A.; Bray, M. S.; Heiss, G.; Zeldin, D. C.; Boerwinkle, E. The Soluble Epoxide Hydrolase Gene Harbors Sequence Variation Associated with Susceptibility to and Protection from Incident Ischemic Stroke. *Hum. Mol. Genet.* **2005**, *14* (19), 2829–2837. https://doi.org/10.1093/hmg/ddi315.
(29) Burley, S. K.; Berman, H. M.; Bhikadiya, C.; Bi, C.; Chen, L.; Di Costanzo, L.; Christie, C.; Duarte, J. M.; Dutta, S.; Feng, Z.; Ghosh, S.; Goodsell, D. S.; Green, R. K.; Guranovic, V.; Guzenko, D.; Hudson, B. P.; Liang, Y.; Lowe, R.; Peisach, E.; Periskova, I.; Randle, C.; Rose, A.; Sekharan, M.; Shao, C.; Tao, Y. P.; Valasatava, Y.; Voigt, M.; Westbrook, J.; Young, J.; Zardecki, C.; Zhuravleva, M.; Kurisu, G.; Nakamura, H.; Kengaku, Y.; Cho, H.; Sato, J.; Kim, J. Y.; Ikegawa, Y.; Nakagawa, A.; Yamashita, R.; Kudou, T.; Bekker, G. J.; Suzuki, H.; Iwata, T.; Yokochi, M.; Kobayashi, N.; Fujiwara, T.; Velankar, S.; Kleywegt, G. J.; Anyango, S.; Armstrong, D. R.; Berrisford, J. M.; Conroy, M. J.; Dana, J. M.; Deshpande, M.; Gane, P.; Gáborová, R.; Gupta, D.; Gutmanas, A.; Koča, J.; Mak, L.; Mir, S.; Mukhopadhyay, A.; Nadzirin, N.; Nair, S.; Patwardhan, A.; Paysan-Lafosse, T.; Pravda, L.; Salih, O.; Sehnal, D.; Varadi, M.; Vařeková, R.; Markley, J. L.; Hoch, J. C.; Romero, P. R.; Baskaran, K.; Maziuk, D.; Ulrich, E. L.; Wedell, J. R.; Yao, H.; Livny, M.; Ioannidis, Y. E. Protein Data Bank: The Single Global Archive for 3D Macromolecular Structure Data. *Nucleic Acids Res.* **2019**, *47* (D1), D520–D528. https://doi.org/10.1093/nar/gky949.
(30) Stepankova, V.; Khabiri, M.; Brezovsky, J.; Pavelka, A.; Sykora, J.; Amaro, M.; Minofar, B.; Prokop, Z.; Hof, M.; Ettrich, R.; Chaloupkova, R.; Damborsky, J. Expansion of Access Tunnels and Active-Site Cavities Influence Activity of Haloalkane Dehalogenases in Organic Cosolvents. *ChemBioChem* **2013**, *14* (7), 890–897. https://doi.org/10.1002/cbic.201200733.
(31) Mowbray, S. L.; Elfström, L. T.; Ahlgren, K. M.; Andersson, C. E.; Widersten, M. X-Ray Structure of Potato Epoxide Hydrolase Sheds Light on Substrate Specificity in Plant Enzymes. *Protein Sci.* **2006**, *15* (7), 1628–1637. https://doi.org/10.1110/ps.051792106.
(32) Grochulski, P.; Li, Y.; Schrag, J. D.; Bouthillier, F.; Smith, P.; Harrison, D.; Rubin, B.; Cygler, M. Insights into Interfacial Activation from an Open Structure of Candida Rugosa

Lipase. *J. Biol. Chem.* **1993**, *268* (17), 12843–12847. https://doi.org/10.1016/s0021-9258(18)31464-9.
(33) Grochulski, P.; Li, Y.; Schrag, J. D.; Cygler, M. Two Conformational States of Candida Rugosa Lipase. *Protein Sci.* **1994**, *3* (1), 82–91. https://doi.org/10.1002/pro.5560030111.
(34) Takai, K.; Chiyo, N.; Nakajima, T.; Nariai, T.; Ishikawa, C.; Nakatani, S.; Ikeno, A.; Yamamoto, S.; Sone, T. Three-Dimensional Rational Approach to the Discovery of Potent Substituted Cyclopropyl Urea Soluble Epoxide Hydrolase Inhibitors. *Bioorg. Med. Chem. Lett.* **2015**, *25* (8), 1705–1708. https://doi.org/10.1016/j.bmcl.2015.02.076.
(35) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: Automating PK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Res.* **2012**, *40* (W1), W537–W541. https://doi.org/10.1093/nar/gks375.
(36) Koudelakova, T.; Bidmanova, S.; Dvorak, P.; Pavelka, A.; Chaloupkova, R.; Prokop, Z.; Damborsky, J. Haloalkane Dehalogenases: Biotechnological Applications. *Biotechnol. J.* **2013**, *8* (1), 32–45. https://doi.org/10.1002/biot.201100486.
(37) Elfström, L. T.; Widersten, M. Catalysis of Potato Epoxide Hydrolase, StEH1. *Biochem. J.* **2005**, *390* (2), 633–640. https://doi.org/10.1042/bj20050526.
(38) Miranda, M.; Urioste, D.; Andrade Souza, L. T.; Mendes, A. A.; De Castro, H. F. Assessment of the Morphological, Biochemical, and Kinetic Properties for Candida Rugosa Lipase Immobilized on Hydrous Niobium Oxide to Be Used in the Biodiesel Synthesis. *Enzyme Res.* **2011**, *2011* (1). https://doi.org/10.4061/2011/216435.
(39) Przybyla-Zawislak, B. D.; Srivastava, P. K.; Vázquez-Matías, J.; Mohrenweiser, H. W.; Maxwell, J. E.; Hammock, B. D.; Bradbury, J. A.; Enayetallah, A. E.; Zeldin, D. C.; Grant, D. F. Polymorphisms in Human Soluble Epoxide Hydrolase. *Mol. Pharmacol.* **2003**, *64* (2), 482–490. https://doi.org/10.1124/mol.64.2.482.
(40) Luchko, T.; Gusarov, S.; Roe, D. R.; Simmerling, C.; Case, D. A.; Tuszynski, J.; Kovalenko, A. Three-Dimensional Molecular Theory of Solvation Coupled with Molecular Dynamics in Amber. *J. Chem. Theory Comput.* **2010**, *6* (3), 607–624. https://doi.org/10.1021/ct900460m.
(41) Sindhikara, D. J.; Yoshida, N.; Hirata, F. Placevent: An Algorithm for Prediction of Explicit Solvent Atom Distribution-Application to HIV-1 Protease and F-ATP Synthase. *J. Comput. Chem.* **2012**, *33* (18), 1536–1543. https://doi.org/10.1002/jcc.22984.
(42) Meyder, A.; Nittinger, E.; Lange, G.; Klein, R.; Rarey, M. Estimating Electron Density Support for Individual Atoms and Molecular Fragments in X-Ray Structures. *J. Chem. Inf. Model.* **2017**, *57* (10), 2437–2447. https://doi.org/10.1021/acs.jcim.7b00391.
(43) Case, D. A.; Ben-Shalom, I. Y.; Brozel, S. R.; Cerutti, D. S.; Cheatham, T. E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 18. University of California: San Francisco 2018.
(44) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5* (21), 3863–3871.

https://doi.org/10.1021/jz501780a.
(45) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9* (9), 3878–3888. https://doi.org/10.1021/ct400314y.
(46) Le Grand, S.; Götz, A. W.; Walker, R. C. SPFP: Speed without Compromise - A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2013**, *184* (2), 374–380. https://doi.org/10.1016/j.cpc.2012.09.022.
(47) 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* (8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.
(48) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11* (4), 1864–1874. https://doi.org/10.1021/ct5010406.
(49) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092. https://doi.org/10.1063/1.464397.
(50) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593. https://doi.org/10.1063/1.470117.
(51) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. . Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327–341. https://doi.org/10.1016/0021-9991(77)90098-5.
(52) Doerr, S.; Harvey, M. J.; Noé, F.; De Fabritiis, G. HTMD: High-Throughput Molecular Dynamics for Molecular Discovery. *J. Chem. Theory Comput.* **2016**, *12* (4), 1845–1852. https://doi.org/10.1021/acs.jctc.6b00049.
(53) Hruska, E.; Abella, J. R.; Nüske, F.; Kavraki, L. E.; Clementi, C. Quantitative Comparison of Adaptive Sampling Methods for Protein Dynamics. *J. Chem. Phys.* **2018**, *149* (24), 244119. https://doi.org/10.1063/1.5053582.
(54) Betz, R. M.; Dror, R. O. How Effectively Can Adaptive Sampling Methods Capture Spontaneous Ligand Binding? *J. Chem. Theory Comput.* **2019**, *15* (3), 2053–2063. https://doi.org/10.1021/acs.jctc.8b00913.
(55) Pérez-Hernández, G.; Paul, F.; Giorgino, T.; De Fabritiis, G.; Noé, F. Identification of Slow Molecular Order Parameters for Markov Model Construction. *J. Chem. Phys.* **2013**, *139* (1), 015102. https://doi.org/10.1063/1.4811489.
(56) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084–3095. https://doi.org/10.1021/ct400341p.
(57) Chovancova, E.; Pavelka, A.; Benes, P.; Strnad, O.; Brezovsky, J.; Kozlikova, B.; Gora, A.; Sustr, V.; Klvana, M.; Medek, P.; Biedermannova, L.; Sochor, J.; Damborsky, J. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* **2012**, *8* (10), e1002708. https://doi.org/10.1371/journal.pcbi.1002708.
(58) Magdziarz, T.; Mitusińska, K.; Bzówka, M.; Raczyńska, A.; Stańczak, A.; Banas, M.; Bagrowska, W.; Góra, A. AQUA-DUCT 1.0: Structural and Functional Analysis of Macromolecules from an Intramolecular Voids Perspective. *Bioinformatics* **2020**, *36* (8),

2599–2601. https://doi.org/10.1093/bioinformatics/btz946.
(59) Brezovsky, J.; Thirunavukarasu, A. S.; Surpeta, B.; Sequeiros-Borja, C. E.; Mandal, N.; Sarkar, D. K.; Dongmo Foumthuim, C. J.; Agrawal, N. TransportTools: A Library for High-Throughput Analyses of Internal Voids in Biomolecules and Ligand Transport through Them. *Bioinformatics* **2022**, *38* (6), 1752–1753. https://doi.org/10.1093/bioinformatics/btab872.
(60) Sequeiros-Borja, C.; Surpeta, B.; Marchlewski, I.; Brezovsky, J. Divide-and-Conquer Approach to Study Protein Tunnels in Long Molecular Dynamics Simulations. *MethodsX* **2023**, *10*, 101968. https://doi.org/10.1016/j.mex.2022.101968.
(61) Petřek, M.; Otyepka, M.; Banáš, P.; Košinová, P.; Koča, J.; Damborský, J. CAVER: A New Tool to Explore Routes from Protein Clefts, Pockets and Cavities. *BMC Bioinformatics* **2006**, *7* (1), 316. https://doi.org/10.1186/1471-2105-7-316.
(62) Klvana, M.; Pavlova, M.; Koudelakova, T.; Chaloupkova, R.; Dvorak, P.; Prokop, Z.; Stsiapanava, A.; Kuty, M.; Kuta-Smatanova, I.; Dohnalek, J.; Kulhanek, P.; Wade, R. C.; Damborsky, J. Pathways and Mechanisms for Product Release in the Engineered Haloalkane Dehalogenases Explored Using Classical and Random Acceleration Molecular Dynamics Simulations. *J. Mol. Biol.* **2009**, *392* (5), 1339–1356. https://doi.org/10.1016/j.jmb.2009.06.076.
(63) Mitusińska, K.; Magdziarz, T.; Bzówka, M.; Stańczak, A.; Gora, A. Exploring Solanum Tuberosum Epoxide Hydrolase Internal Architecture by Water Molecules Tracking. *Biomolecules* **2018**, *8* (4), 143. https://doi.org/10.3390/biom8040143.
(64) Rodríguez-Salarichs, J.; García De Lacoba, M.; Prieto, A.; Martínez, M. J.; Barriuso, J. Versatile Lipases from the Candida Rugosa-like Family: A Mechanistic Insight Using Computational Approaches. *J. Chem. Inf. Model.* **2021**, *61* (2), 913–920. https://doi.org/10.1021/acs.jcim.0c01151.
(65) Mancheño, J. M.; Pernas, M. A.; Martínez, M. J.; Ochoa, B.; Rúa, M. L.; Hermoso, J. A. Structural Insights into the Lipase/Esterase Behavior in the Candida Rugosa Lipases Family: Crystal Structure of the Lipase 2 Isoenzyme at 1.97Å Resolution. *J. Mol. Biol.* **2003**, *332* (5), 1059–1069. https://doi.org/10.1016/j.jmb.2003.08.005.
(66) Domínguez de María, P.; Sánchez-Montero, J. M.; Sinisterra, J. V.; Alcántara, A. R. Understanding Candida Rugosa Lipases: An Overview. *Biotechnol. Adv.* **2006**, *24* (2), 180–196. https://doi.org/10.1016/j.biotechadv.2005.09.003.
(67) de Melo, J. J. C.; Gonçalves, J. R.; Brandão, L. M. de S.; Souza, R. L.; Pereira, M. M.; Lima, Á. S.; Soares, C. M. F. Evaluation of Lipase Access Tunnels and Analysis of Substance Transport in Comparison with Experimental Data. *Bioprocess Biosyst. Eng.* **2022**, *45* (7), 1149–1162. https://doi.org/10.1007/s00449-022-02731-x.
(68) Foresti, M. L.; Ferreira, M. L. Computational Approach to Solvent-Free Synthesis of Ethyl Oleate Using Candida r Ugosa and Candida a Ntarctica B Lipases. I. Interfacial Activation and Substrate (Ethanol, Oleic Acid) Adsorption. *Biomacromolecules* **2004**, *5* (6), 2366–2375. https://doi.org/10.1021/bm049688o.
(69) Bzówka, M.; Mitusińska, K.; Hopko, K.; Góra, A. Computational Insights into the Known Inhibitors of Human Soluble Epoxide Hydrolase. *Drug Discov. Today* **2021**, *26* (8), 1914–1921. https://doi.org/10.1016/j.drudis.2021.05.017.
(70) Mitusińska, K.; Wojsa, P.; Bzówka, M.; Raczyńska, A.; Bagrowska, W.; Samol, A.; Kapica, P.; Góra, A. Structure-Function Relationship between Soluble Epoxide Hydrolases Structure and Their Tunnel Network. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 193–205.

https://doi.org/10.1016/j.csbj.2021.10.042.
(71) McInnes, L.; Healy, J.; Astels, S. Hdbscan: Hierarchical Density Based Clustering. *J. Open Source Softw.* **2017**, *2* (11), 205. https://doi.org/10.21105/joss.00205.
(72) Brezovsky, J.; Chovancova, E.; Gora, A.; Pavelka, A.; Biedermannova, L.; Damborsky, J. Software Tools for Identification, Visualization and Analysis of Protein Tunnels and Channels. *Biotechnol. Adv.* **2013**, *31* (1), 38–49. https://doi.org/10.1016/j.biotechadv.2012.02.002.
(73) Pavelka, A.; Sebestova, E.; Kozlikova, B.; Brezovsky, J.; Sochor, J.; Damborsky, J. CAVER: Algorithms for Analyzing Dynamics of Tunnels in Macromolecules. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* **2016**, *13* (3), 505–517. https://doi.org/10.1109/tcbb.2015.2459680.
(74) Bzówka, M.; Mitusińska, K.; Raczyńska, A.; Skalski, T.; Samol, A.; Bagrowska, W.; Magdziarz, T.; Góra, A. Evolution of Tunnels in α/β-Hydrolase Fold Proteins—What Can We Learn from Studying Epoxide Hydrolases? *PLOS Comput. Biol.* **2022**, *18* (5), e1010119. https://doi.org/10.1371/journal.pcbi.1010119.
(75) Schmitt, J.; Brocca, S.; Schmid, R. D.; Pleiss, J. Blocking the Tunnel: Engineering of Candida Rugosa Lipase Mutants with Short Chain Length Specificity. *Protein Eng. Des. Sel.* **2002**, *15* (7), 595–601. https://doi.org/10.1093/protein/15.7.595.
(76) Wang, L.; Althoff, E. A.; Bolduc, J.; Jiang, L.; Moody, J.; Lassila, J. K.; Giger, L.; Hilvert, D.; Stoddard, B.; Baker, D. Structural Analyses of Covalent Enzyme-Substrate Analog Complexes Reveal Strengths and Limitations of de Novo Enzyme Design. *J. Mol. Biol.* **2012**, *415* (3), 615–625. https://doi.org/10.1016/j.jmb.2011.10.043.
(77) Wagner, K.; Vito, S.; Inceoglu, B.; Hammock, B. D. The Role of Long Chain Fatty Acids and Their Epoxide Metabolites in Nociceptive Signaling. *Prostaglandins Other Lipid Mediat.* **2014**, *113–115*, 2–12. https://doi.org/10.1016/j.prostaglandins.2014.09.001.
(78) Decker, M.; Adamska, M.; Cronin, A.; Di Giallonardo, F.; Burgener, J.; Marowsky, A.; Falck, J. R.; Morisseau, C.; Hammock, B. D.; Gruzdev, A.; Zeldin, D. C.; Arand, M. EH3 (ABHD9): The First Member of a New Epoxide Hydrolase Family with High Activity for Fatty Acid Epoxides. *J. Lipid Res.* **2012**, *53* (10), 2038–2045. https://doi.org/10.1194/jlr.m024448.
(79) Morisseau, C.; Hammock, B. D. Impact of Soluble Epoxide Hydrolase and Epoxyeicosanoids on Human Health. *Annu. Rev. Pharmacol. Toxicol.* **2013**, *53* (1), 37–58. https://doi.org/10.1146/annurev-pharmtox-011112-140244.
(80) Kadaoluwa Pathirannahalage, S. P.; Meftahi, N.; Elbourne, A.; Weiss, A. C. G.; McConville, C. F.; Padua, A.; Winkler, D. A.; Costa Gomes, M.; Greaves, T. L.; Le, T. C.; Besford, Q. A.; Christofferson, A. J. Systematic Comparison of the Structural and Dynamic Properties of Commonly Used Water Models for Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2021**, *61* (9), 4521–4536. https://doi.org/10.1021/acs.jcim.1c00794.
(81) Agrawal, N.; Brezovsky, J. Impact of Water Models on Structure and Dynamics of Ligand-Transport Tunnels in Enzymes Derived from Molecular Dynamics Simulations. *bioRxiv* **2023**, 2023.04.19.537534. https://doi.org/10.1101/2023.04.19.537534.
(82) Brezovsky, J.; Babkova, P.; Degtjarik, O.; Fortova, A.; Gora, A.; Iermak, I.; Rezacova, P.; Dvorak, P.; Smatanova, I. K.; Prokop, Z.; Chaloupkova, R.; Damborsky, J. Engineering a de Novo Transport Tunnel. *ACS Catal.* **2016**, *6* (11), 7597–7610. https://doi.org/10.1021/acscatal.6b02081.