Conformational changes allow processing of bulky substrates by a haloalkane dehalogenase with a small and buried active site

Received for publication, October 7, 2017, and in revised form, May 28, 2018. Published, Papers in Press, June 1, 2018, DOI 10.1074/jbc.RA117.000328

Piia Kokkonen1,‡, David Bednar1,‡, Veronika Dockalova1,‡, Zbynek Prokop4,‡, and Jiri Damborsky1,‡

From the ‡ Loschmidt Laboratories, Department of Experimental Biology, Research Centre for Toxic Compounds in the Environment (RECETOX), Masaryk University, Kamenné 5/A 13, 625 00 Brno, Czech Republic and ‡§ International Centre for Clinical Research, St. Anne’s University Hospital, Pekarska 53, 656 91 Brno, Czech Republic

Edited by John M. Denu

Haloalkane dehalogenases catalyze the hydrolysis of halogen–carbon bonds in organic halogenated compounds and as such are of great utility as biocatalysts. The crystal structures of the haloalkane dehalogenase DhlA from the bacterium from Xanthobacter autotrophicus GJ10, specifically adapted for the conversion of the small 1,2-dichloroethane (DCE) molecule, display the smallest catalytic site (110 Å3) within this enzyme family. However, during a substrate-specificity screening, we noted that DhlA can catalyze the conversion of far bulkier substrates, such as the 4-(bromomethyl)-6,7-dimethoxy-coumarin (220 Å3). This large substrate cannot bind to DhlA without conformational alterations. These conformational changes have been previously inferred from kinetic analysis, but their structural basis has not been understood. Using molecular dynamic simulations, we demonstrate here the intrinsic flexibility of part of the cap domain that allows DhlA to accommodate bulky substrates. The simulations displayed two routes for transport of substrates to the active site, one of which requires the conformational change and is likely the route for bulky substrates. These results provide insights into the structure–dynamics function relationships in enzymes with deeply buried active sites. Moreover, understanding the structural basis for the molecular adaptation of DhlA to 1,2-dichloroethane introduced into the biosphere during the industrial revolution provides a valuable lesson in enzyme design by nature.

The haloalkane dehalogenase (HLD, EC 3.8.1.5)DhlA from Xanthobacter autotrophicus GJ10 has the smallest catalytic site among the solved crystal structures of HLDs. Despite this hindrance, it catalyzes the hydrolysis of carbon–halogen bonds in a wide variety of substrates and displays notable catalytic efficiency for the generally poorly processed small substrate 1,2-dichloroethane (DCE) (1, 2). DhlA has been the target of numerous protein engineering and computational studies which have revealed that the sequence and structure of the cap domain control the substrate specificity and the catalytic activity of this enzyme (3–10). Kinetic experiments provided the experimental evidences (Fig. 1) that DhlA can adopt two distinct conformational states in the reaction mixture: The open and the closed states (11, 12). The switching between these states affects mostly the release of the halide ion which is the rate-limiting step of the catalytic cycle. The complex pipe-shaped dependence of the observed kinetic rate against the halide concentration indicates slow conformational transitions in two parallel kinetic routes. These conformational changes involve both conformational selection and induced fit mechanisms. In the first route, a slow enzyme isomerization step is followed by rapid halide binding, which was predominant at low halide concentrations. The second route involves a rapid binding into an initial collision complex followed by an induced slow conformational step and prevailed at higher halide concentrations. The overall rate of the halide release is slow and limited by these conformational changes. Yet, the solved crystal structures of DhlA are almost identical with a root mean square deviation (RMSD) of Cα atoms, RMSD < 0.5 Å, and display only the closed state (3, 4, 11, 13–16). The closed state in crystal structures displays a tightly packed enzyme with a catalytic site of a volume barely able to accommodate DCE (Fig. 2). Surprisingly, during a substrate-specificity screening with DhlA, we noted that this enzyme can catalyze the debromination of 4-(bromomethyl)-6,7-dimethoxy-coumarin (COU).4 COU is more than three times larger than the natural substrate DCE and two times larger than the active site (Fig. 2). This observation and the lack of molecular level data inspired us to conduct molecular dynamics (MD) simulations to elucidate the structural basis for the catalytically important conformational changes in DhlA.

Results

Steady-state kinetics of DhlA with COU confirms the conversion of a bulky substrate

The fluorescence of COU increases after the conversion to the product (4-hydroxymethyl-6,7-dimethoxycoumarin), which enables the determination of steady-state kinetics. The

This work was supported by Grant Agency of the Czech Republic Grants GA16-08096S and GA16-07965S; Czech Ministry of Education Grants LQ1605, LO1214, LM2015051, LM2015047, and LM2015055; and European Union Grant 676559. J. D. and Z. P. are cofounders of the biotechnology spin-off Enantis Ltd.

This article contains Figs. S1–S13.

1To whom the correspondence may be addressed. E-mail: zbynek@chemi.muni.cz.
2To whom the correspondence may be addressed. E-mail: jiri@chemi.muni.cz.
3The abbreviations used are: HLD, haloalkane dehalogenase; aMD, accelerated molecular dynamics; COU, 4-(bromomethyl)-6,7-dimethoxy-coumarin; DCE, 1,2-dichloroethane; HTMD, high throughput molecular dynamics; MD, molecular dynamics; RMSD, root mean square deviation.
4V. Dockalova et al., manuscript in preparation.

© 2018 by The American Society for Biochemistry and Molecular Biology, Inc. Published in the U.S.A.

This is an Open Access article under the CC BY license.
DhlA changes conformation to accommodate bulky substrates

The rate at which DhlA catalyzes COU dehalogenation is slow, yet significantly different from the abiologic control (Table 1 and Fig. S1). Previous studies focusing on DhlA substrate specificity show that it catalyzes a variety of compounds and it prefers smaller and nonycyclic substrates, which is consistent with its constricted active site (4, 17). DhlA must employ some conformational changes to adapt to bulkier substrates. Earlier attempts to capture the conformational dynamics of HLDs DhlA and DhaA using NMR spectroscopy by the Dick B. Janssen group (University of Groningen, The Netherlands) and by our group were unsuccessful. Thus, we opted for MD simulations to study the conformational dynamics of DhlA at the molecular level.

MD simulations with COU suggest two binding pathways

The initial analysis of the DhlA and COU interaction was performed with the accelerated molecular dynamics (aMD) simulation using Amber 14 and five COU molecules (concentration 22 mM) introduced into the solvent around the enzyme (18, 19). A high concentration of COU was used to increase the statistical possibility of binding and acceleration was used to enhance the binding of COU as it was expected to be slow based on the low turnover rates (Table 1). In these simulations, we observed binding of COU to the entrance of both the main tunnel p1 as well as to two different locations of the p3 tunnel (Fig. 3). The simulations revealed that the cap domain of DhlA displays conformational changes, thus opening the p3 tunnel to such an extent that it can accommodate up to two COU molecules in the locations p3a and p3b. The cap domain has been determined as the most flexible part of DhlA in the previous studies employing MD simulations, and it was also suspected to contribute to the open and closed states in the previous kinetic studies (10, 20, 21). The opening was measured as the distance between Lys-176C and Pro-223C, which is 5.7 Å in the crystal structure PDB ID 2YXP and ~8.5 Å in the open state (Figs. S2 and S3).

To ensure that the opening of the p3 tunnel is not caused by the force field or the underlying software code, eight classical 200-ns MD simulations of ligand-free DhlA were conducted using CHARMM36 force field and AceMD (22, 23). A similar cap domain opening was observed in five of these simulations, thus diminishing the possibility that we were observing a software- or force field-related artifact and supporting the biological feasibility of the conformational change (Fig. S4). The opening of the cap domain does not enlarge the tunnel p1 as dramatically as the tunnel p3 (Fig. S4). The ligand-free simulations showed that the conformational change occurs without the inductive effect of the ligand. Furthermore, the open conformation has good stereochemistry (<0.5% of residues in disallowed regions in the Ramachandran plot), and it is reversible as some simulations lead back to the closed conformation (Fig. S4) (24).

The conformational dynamics of Glu-56

The initial accelerated simulations did not generate the productive complex where the nucleophilic attack would be possible at the active site even though COU bound to the access tunnels p1, p3a, and p3b. The binding of COU should occur in a longer timescale based on the experimental results, making it difficult to observe in random simulations. Furthermore, because of its large size, COU cannot be directly docked to the active site of DhlA for the analysis of the unbinding process and introducing a biasing force might lead to unnatural binding routes or modes. Therefore, DCE was used to start simulations from the productive enzyme–substrate complex to detect whether similar binding routes to COU are also employed by this smaller substrate.

The latest high-resolution crystal structure of DhlA (PDB ID 2YXP) was used in the calculations and simulations (13). This structure, as all the other published crystal structures of DhlA (PDB IDs 1B6G, 1EDE, 2HAD, 1E6B, 2DHE, 2EDA, 2DH, and 2EDC), has a hydrogen bond between the side chain of buried Glu-56 and the backbone carbonyl oxygen of Val-219 (Fig. 4). During the initial aMD simulations with deprotonated Glu-56, this hydrogen bond was not formed and Glu-56 rotated from the hydrophobic pocket to face the active site within the first few femtoseconds of the equilibration simulation. Glu-56 rotates because the hydrophobic pocket does not provide any interaction partners for a charged glutamic acid residue.

The H++ web server calculated Glu-56 a pK_a value around 7, depending on the crystal structure used (6.9 for 2YXP, 7.1 for 1EDE, 6.7 for 1B6G, and 6.6 for 2DH), resulting in the initial simulations having been conducted with the deprotonated form at pH 7.5 (25). To study the effect of the hydrogen bond between Glu-56 side chain and Val-219, the DhlA–DCE complex was simulated with both protonated and deprotonated Glu-56. The run time of these simulations was 7150 ns and 8400 ns for the deprotonated and protonated Glu-56, respectively. The RMSD plots of the Ca atoms of the backbone, the Ca atoms of the moving cap domain helix (residues 166–186), and the Ca atoms excluding the moving cap domain helix (Figs. S5–S11) indicate that the moving cap domain helix contributes most to the RMSD fluctuations and that the rest of the protein remains stable in these simulations.

Effect of the protonation state of Glu-56 on the conformational behavior of DhlA

The opening of the cap domain is possible with both protonation states of Glu-56 (Table 2). When Glu-56 is protonated, it adopts both the conformation observed in the crystal structure as well as the conformation facing the active site (Fig. 4). In the
DhlA changes conformation to accommodate bulky substrates

Figure 2. The volume of the catalytic site of DhlA (PDB ID 2YXP) compared with the van der Waals volume of the two substrates. Residues 156–196 of the cap domain are colored cyan in the DhlA structure. The volume of the DhlA catalytic site was calculated using Castp web server (43) and the van der Waals volume of the ligands using the methodology by Zhao et al. (60).

Table 1
Kinetic constants for the reaction of DhlA with COU determined at 30 °C and pH 8.0

The Michael constant $K_{m}$, turnover number $k_{cat}$, and equilibrium dissociation constant for enzyme-product complex $K_i$ were calculated by global fitting of the kinetic data. The S.E. was calculated from the covariance matrix during nonlinear regression, and the lower and upper limits for each parameter were derived from the confidence contour analysis.

<table>
<thead>
<tr>
<th></th>
<th>$K_m$</th>
<th>$k_{cat}$</th>
<th>$K_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value ± S.E.</td>
<td>$49 \pm 2$</td>
<td>$1.24 \pm 0.04$</td>
<td>$26 \pm 1$</td>
</tr>
<tr>
<td>Lower, upper bound</td>
<td>44, 56</td>
<td>1.18, 1.33</td>
<td>24, 28</td>
</tr>
</tbody>
</table>

Figure 3. The two substrate binding routes explored by COU molecules in the accelerated MD simulations. On the left, a single molecule is bound to the entrance of the main tunnel p1 of HLDs and on the right two molecules are bound to p3a and p3b locations in the tunnel p3 which has been opened by a conformational change by the cap domain residues shown in cyan. The other three to four molecules are interacting with different parts of the protein surface or stacking together. The halide stabilizing Trp-125 and Trp-175 and the catalytic Asp-124 are shown as sticks to illustrate the active site.

DhlA changes conformation to accommodate bulky substrates

Deprotonated form Glu-56 adopts only the conformation facing the active site, similarly to the aMD simulations. Based on the predicted $pK_a$ value of Glu-56, we estimate it to display both protonation states in the reaction mixture at pH 8, with a strong excess of the deprotonated form. Most of the crystal structures of DhlA have been determined in pH 5–6, which might explain the abundance of the protonated form of Glu-56 observed in them.

The protonation state of Glu-56 influences the frequency of the opening of the cap domain and the packing of the closed conformation. When Glu-56 is protonated, the hydrogen bond between it and Val-219 stabilizes the closed structure and makes the opening less frequent (Table 2). Deprotonated Glu-56 adopts only the conformation where it faces the active site, thus making the packing of the protein core more efficient and the distance Lys-176C₆H₁₂−Pro-223C₆H₁₂ in the closed state slightly shorter (~0.5 Å, Table 2).

The effect of the mutation F172W on the conformational behavior of DhlA

To study the effect of mutations to the conformational changes, we used the crystal structure of DhlA with mutation F172W (PDB ID 1HDE). This mutation was described by Schanstra et al. in 1996 to display a faster enzyme isomerization step than the WT (3). The results from the deprotonated Glu-56 simulations of this mutant confirm that the conformational change is faster (Table 2). The open conformation is more common than the closed one. The closed conformation also displays a larger distance between the two helices than what was observed in the WT simulations.

DCE binds through two routes, one of which requires a conformational change

During the simulations with the DCE–DhlA complex, the substrate left and rebound to the enzyme through two pathways: (i) through the main crystallographically observable tunnel p1 and (ii) through the newly opened p3 tunnel (Fig. 5). This corresponds with the pathways probed by COU in the aMD simulations (Fig. 3). Both of the previously mentioned simulations complement previous kinetic studies, where the binding and the release of the halide ion proceeds via two distinct routes, one of which requires a conformational change (Fig. 1) (3, 4, 9, 11, 16). Similarly to the experiments, the closed state (represented in the crystal structures) is more common than the open one (20). The binding through the open conformation must be slower because the enzyme needs to return to the closed conformation to form the reactive complex. The formation of the reactive complex in the open conformation is not possible because one of the halide-stabilizing residues (Trp-175) is located on the gating cap domain helix. The opening causes this residue to move too far from the other halide-stabilizing residue (Trp-125) to stabilize the position of the ion during the reaction.

Table 1

<table>
<thead>
<tr>
<th></th>
<th>$K_m$</th>
<th>$k_{cat}$</th>
<th>$K_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value ± S.E.</td>
<td>$49 \pm 2$</td>
<td>$1.24 \pm 0.04$</td>
<td>$26 \pm 1$</td>
</tr>
<tr>
<td>Lower, upper bound</td>
<td>44, 56</td>
<td>1.18, 1.33</td>
<td>24, 28</td>
</tr>
</tbody>
</table>
Discussion

The conformational changes in DhlA take place at the residues 155–187 of the cap domain (20, 21). The conformational change enables the substrate molecule to enter the active site through the p3 tunnel in addition to the p1 tunnel. The conformational change also increases the volume of the active site. The p1 tunnel has been termed the main tunnel, as it is observable in the crystal structures and corresponds to the main transport route in other HLDs (10, 26, 27). Analogous opening of the cap domain and the p3 tunnel has not been observed in simulations carried out with other HLDs.

The unique conformational flexibility of the cap domain of DhlA explains why it is the only natural dehalogenase with high activity toward DCE and 1,2-dichloropropane (17, 28, 29). Evolutionarily, the cap domain of DhlA has been formed by a repetition of its sequence (residues Val-156–Ala-160 and Val-165–Ala-169) which causes an extension in the loop preceding the helices of the cap domain (Figs. 2 and 3) (9). The cap domain of DhlA also lacks an α-helix connecting the two helices lining the p1 tunnel (loop near the p3b route in Figs. 3 and 5). These two flexible loops allow DhlA to open the p3 tunnel and to hydrate the active site with a large number of water molecules. This is beneficial for the release of a tightly bound chloride ion product. The halide release is known to be the slowest step for the conversion of DCE by DhlA (11). DCE is clearly one of the most recalcitrant substrates of HLDs when many of the family members cannot catalyze its dehalogenation. We believe that the conformational change is a result of a molecular adaptation of DhlA to DCE, whereas conversion of bulkier substrates like COU, is a side effect of this adaptation (17).

The protonation state of Glu-56 affects the structure and the function of DhlA (Table 2). A protonatable residue and a buried charge in this position are not common among HLDs. In other HLDs, the position analogous to Glu-56 is occupied by an Asn or a Gln residue which functions as the second halide-stabilizing residue, directly involved in the catalytic machinery (30). In DhlA, the second halide stabilizing residue Trp-175 is located on the flexible cap domain helix. Independent of its protonation state, Glu-56 can turn to face the active site. The turning of Glu-56 hinders the reactive binding of the halogenated sub-

**Table 2**

<table>
<thead>
<tr>
<th>States</th>
<th>DhlA Glu-56 protonated</th>
<th>DhlA Glu-56 deprotonated</th>
<th>DhlA F172W Glu-56 deprotonated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Open</td>
<td>Closed</td>
<td>Open</td>
</tr>
<tr>
<td>Lys-176_Cα-Pro-223_Cα (Å)</td>
<td>11.1 ± 0.3 (10.5–1.6)</td>
<td>6.4 ± 0.1 (6.2–6.5)</td>
<td>12.1 ± 0.4 (11.4–13.0)</td>
</tr>
<tr>
<td>ΔG (kcal/mol)</td>
<td>1.1 ± 0.3 (0.2–1.6)</td>
<td>0.1 ± 0.3 (0.5–0.6)</td>
<td>0.1 ± 0.3 (0.5–0.6)</td>
</tr>
<tr>
<td>Equilibrium probability (%)</td>
<td>27 ± 6</td>
<td>73 ± 6</td>
<td>47 ± 11</td>
</tr>
</tbody>
</table>
strates, as it interferes with the interaction network between the substrate and the halide-stabilizing residues Trp-125 and Trp-175. The partial negative charge of Glu-56 side chain also repulses partially negatively charged DCE molecule. The protonation state of Glu-56 can change because of the contact with the water molecules at the active site. The conformational flexibility of Glu-56 enables it to disturb the interactions of the halide ion and the halide-stabilizing residues, further benefiting the release of the charged halide ion product.

Transitions between open and closed states help to control the entry of substrates, the release of products, and the access of water molecules to and from the active site in many enzymes (31–35). The structural basis for these transitions is generally the movement of a domain or a part of it, such as the cap domain helix in the case of DhlA. The open state generally allows the entry of molecules to the active site, whereas the closed state ensures the correct organization for the chemical step. Even though the transitions are common in enzymes, our understanding of their importance for enzymatic catalysis is poorly understood (36, 37). Kinetic experiments can reveal the presence of transitions between the open and closed states, but NMR studies or molecular dynamics simulations are needed to study the states at the molecular level. Here, we used the adaptive sampling scheme to understand the molecular mechanisms of transitions without inducing a strong biasing force to the system (38). It is a very powerful method in finding various enzyme conformations in a relatively short simulation time because new simulations are started from the less-sampled states and the system does not get stuck in the local minima. We recommend the adaptive sampling scheme utilizing molecular dynamics simulations when deciphering the changes that cause the open and closed states of enzymes. Understanding the molecular mechanisms of the different conformations can lead to the design of more efficient enzymes and potent inhibitors.

In summary, the cap domain of DhlA can open a second access tunnel and make room in the active site to accommodate bulky substrates. These substrates can have a volume of up to twice the volume of the (closed) active site in the crystal structure. The flexibility of enzyme domains is not necessarily evident from the crystal structures, and thus one should be careful not to exclude substrates from specificity screenings based on the comparison of their size with the active site volume. Conformational dynamics of HLD DhlA is a result of its molecular adaptation to the conversion of the highly recalcitrant molecule DCE, released to the environment very recently, during the industrial revolution (5). Understanding the structural basis of conformational dynamics in DhlA provides useful instructions for the computational design of catalytically efficient biocatalysts. We have recently demonstrated that de novo tunnels can be introduced to the enzymes with buried active sites, like HLDs (39). In addition to engineering the access tunnels, the introduction of dynamic elements providing the transitions between the open and closed states offers unexplored possibilities in the computational design of enzymes.

### Experimental procedures

#### Steady-state kinetic analysis

The steady-state kinetic experiments were performed in the microtiter plates at 30 °C using spectrofluorometer FLUOstar Optima. The reaction mixture contained 176 μl of 50 mM phosphate buffer pH 8.0, 20 μl of substrate dissolved in DMSO, and 4 μl of DhlA. The reaction was initiated by the addition of the substrate through an automatic syringe pump exactly 10 min after the substrate dissolved and was monitored for 91 min. The fluorescent intensity was measured from the top using excitation short-pass filter 360 nm and emission band-pass filter 460–10 nm. Before each measurement of fluorescent intensity, the microtiter plate was shaken for 2 s. The experiment was conducted for five substrate concentrations 120 μM, 80 μM, 40 μM, 20 μM, and 10 μM; the concentration of DhlA was 3.5 μM. The reaction took place in eight wells accompanied by another eight wells of abiotic control (phosphate buffer). For all substrate concentrations, the measurement was independently repeated with three different substrate aliquots.

The measured fluorescent intensity of product and substrate (i.e. abiotic control) was converted to product concentration according to quadratic calibration curves calculated from fluorescent intensity after total conversion. The substrate concentration 120 μM was not included in the calculation because the fluorescent intensity of the product no longer increases with this concentration and higher. This aberration is most likely caused by aggregation quenching. The fluorescent intensity of spontaneous hydrolysis of the substrate was subtracted from the converted data according to Equation 1,

\[
f(\text{c}_{\text{prod}}) = FI - f(\text{c}_{\text{subs}})
\]  

(Eq. 1)

where \(f(\text{c}_{\text{prod}})\) is formula of calibration curve for product, \(FI\) is observed fluorescent intensity, and \(f(\text{c}_{\text{subs}})\) is formula of calibration curve for substrate.

The kinetic data were fitted globally by using dynamic kinetic simulation program KinTek Global Kinetic Explorer version 5.2 (KinTek, Snow Shoe, PA). Data fitting used numerical integration of rate equations derived from the input model in Equation 2,

\[
E + S \leftrightarrow ES \rightarrow EP \leftrightarrow E + P
\]  

(Eq. 2)

where \(E\) is an enzyme; \(S\) and \(P\) are substrate and product, respectively; \(ES\) and \(EP\) are enzyme–substrate and enzyme–product complex, respectively; \(K_{\text{m}}\) is Michaelis constant; \(k_{\text{cat}}\) is turnover number; and \(K_i\) is equilibrium dissociation constant for enzyme–product complex. The best fit was reached by searching a set of kinetic parameters that produce a minimum \(\chi^2\) value using nonlinear regression based on the Levenberg-Marquardt method. Residuals were normalized by \(\sigma\) value for each data point. The S.E. was calculated from the covariance matrix during nonlinear regression in globally fitting all fluorescence traces (40). In addition to the S.E. values, more rigorous analysis of the variation of the kinetic parameters was accomplished with confidence contour analysis by FitSpace Explorer (KinTek). In this analysis, the lower and upper limits
DhIA changes conformation to accommodate bulky substrates

for each parameter were derived from the confidence contours for the $\chi^2$ threshold at the boundary of 0.95 (41).

Computational studies

Structure preparation—The crystal structure of DhIA (PDB ID 2YXP) was downloaded from RSCB Protein Data Bank (13, 42). The hydrogen atoms were added to the structure with H++ web server at pH 7.5 (25). The protonation state of Glu-56 was manually changed in the protonated structure. The water molecules from the crystal structure, which did not overlap with the protonated structure, were retained.

The partial atomic charges for the ligands COU and DCE were calculated using the PyRED server (44). Input geometries were optimized by Gaussian 2009 D01 program interfaced with this server, and a multiorientation RESP fit with RESP + A1A charge model was performed. The partial charges were implemented into Amber force field using the Antechamber module of Amber Tools 15 (45, 46).

Accelerated MD simulations—Energy minimization and accelerated MD simulations were performed using the PMEMD module of Amber 14 with the ff14SB force field (19, 47). Initially, the investigated systems were minimized by 500 steps of the steepest descent followed by 500 steps of the conjugate gradient over five rounds with decreasing harmonic restraints. The restraints were applied as follows: 500 kcal/mol × 1 Å$^{-2}$ on all the heavy atoms of the protein, and then 500, 125, 25, and 0 kcal/mol × 1 Å$^{-2}$ on backbone atoms only. The subsequent MD simulations employed periodic boundary conditions, using the particle mesh Ewald method for the treatment of interactions beyond 10 Å cut-off, and a 2-fs time step with the SHAKE algorithm to fix all bonds containing hydrogens (48, 49). Equilibration simulations consisted of two steps: (i) 20 ps of gradual heating from 0 to 310 K at a constant volume, using a Langevin thermostat with a collision frequency of 1.0 ps$^{-1}$, and with harmonic restraints of 5.0 kcal/mol × 1 Å$^{-2}$ on the position of all protein atoms, and (ii) 2000 ps of unrestrained MD at 310 K using the Langevin thermostat, and constant pressure of 1.0 bar using pressure coupling constant of 1.0 ps. The acceleration parameters from a single 20 ns unbiased MD simulation were $E_{\text{dih}} = 5124$; $\sigma_{\text{dih}} = 216$; $E_{\text{rot}} = -100116$; $\alpha_{\text{rot}} = 7218$; $\lambda_{\text{rot}} = 7\%$; and $\lambda_{\text{dih}} = 27\%$. Other simulation settings were as in the second step of the equilibration MD and the coordinates were saved with 0.1-ns interval. The trajectories were analyzed using Cpptraj and visualized in VMD 1.9.1 and PyMol 1.7.1.

MD simulations with CHARMM force field—The system for AceMD and CHARMM36 force field as prepared with a cubic TIP3P water box reaching at least 10 Å from the protein atoms and a 0.1 M Cl$^{-}$ and Na$^+$ ion concentration using the High Throughput Molecular Dynamics (HTMD) CHARMM builder module (52). The simulations used CHARMM36 parameters for the ions and the protein atoms. Initially, the investigated systems were minimized by 500 steps of conjugate gradient. Then the system was heated and minimized as follows: (i) 25,000 steps (100 ps) of NVT thermalization with Berendsen barostat to 310 K with constraints on all heavy atoms of the protein, (ii) 250,000 steps (1 ns) of NPT equilibration with Langevin thermostat with 0.1 kcal/mol × 1 Å$^{-2}$ constraints on heavy atoms and 1.0 kcal/mol × 1 Å$^{-2}$ on the protein atoms, and (iii) 250,000 steps (1 ns) of NPT equilibration with Langevin thermostat without constraints. During the equilibration simulations, holonomic constraints were on all hydrogen-heavy atom bond terms and the mass of hydrogen atoms was scaled with factor 4, enabling the 4-fs time step (22, 53–55). The simulations employed the default settings of AceMD from the HTMD protocol production v6, including periodic boundary conditions, using the particle mesh Ewald method for the treatment of the interactions beyond 9 Å cut-off, electrostatic interactions suppressed > 4 bond terms away from each other and the smoothing and switching van der Waals and electrostatic interaction were employed from 7.5 Å to the cut-off value (55). The production MD simulations were run for each system using the same settings as the last step of the equilibration for 200 ns. The coordinates were saved with 0.1 ns interval, and the resulting DCD-trajectories were analyzed using Cpptraj and visualized in VMD 1.9.1 and PyMol 1.7.1.

Adaptive sampling MD simulations with docked DCE—DCE was docked to the catalytic site with Autodock Vina (56). The docking grid was calculated 18 Å from the coordinates corresponding to the catalytic oxygen of Asp-124. Water molecules clashing with the docked pose of the ligand were removed. The system was solvated using the tLeap module of Amber Tools 15 in a cubic water box of TIP3P water molecules (19, 57). Cl$^{-}$ and Na$^+$ ions were added to neutralize the charge of the protein and to get a final concentration of 0.1 M.

The systems were equilibrated using the Equilibration v1 module of HTMD (22, 52, 55). The system was first minimized using conjugate-gradient method for 500 steps. Then the system was heated and minimized as follows: (i) 500 steps (2 ps) of NVT thermalization with Berendsen barostat to 310 K with 0.1 kcal/mol × 1 Å$^{-2}$ constraints on heavy atoms and 1.0 kcal/mol × 1 Å$^{-2}$ on the protein atoms, (ii) 1250 steps (5 ps) of NPT equilibration with Langevin thermostat with 0.1 kcal/mol × 1 Å$^{-2}$ constraints on heavy atoms and 1.0 kcal/mol × 1 Å$^{-2}$ on the protein atoms, and (iii) 1250 steps (5 ps) of NPT equilibration with Langevin thermostat without constraints (22, 53–55).

HTMD was used to perform adaptive sampling of the substrate leaving the binding site from the docked position with the same simulation settings as described for the CHARMM system production simulations (38). The 50-ns production runs were started with the files resulting from the equilibration, and they employed the same settings as the last step of the equilibration. Adaptive sampling was run using the distance between the DCE molecule and the nucleophile Asp-124 as the reaction coordinate (58). The reaction coordinate was changed after ~3 μs of simulation time to the distance of Lys-176 and Pro-223 Cα atoms to enhance the sampling of the opening.

The simulations were made into a simulation list using HTMD, and water was filtered out and crashed simulations with the length <50 ns were omitted (52). The total simulation time used in calculations was $142 \times 50$ ns for the deprotonated

---

state and 168 × 50 ns for the protonated state. Multiple Markov state models were created to study the effect of clustering, the usage of different lag times and the amount of Markov states. In the final model, the dynamics of the cap domain were mapped with the distances of Cα atoms of residues Lys-176 and Pro-223 (Fig. S2). The statistics between the open and closed states were estimated using a bootstrap method by 1000 times resampling random 50% of the original data. The bootstrapped data were clustered using MiniBatchKMeans algorithm to 100 clusters after which a two-state Markov model was built with a lag time of 20 ns (59). The implied time scales plot and the Chapman-Kolmogorov test of the obtained Markov state models are shown in Figs. S11 and S12.

Adaptive sampling simulations of DhlA mutant F172W—
The structure of the F172W mutant (PDB ID 1HDE) with the docked DCE was prepared and equilibrated similarly as the DhlA structure. The adaptive sampling simulations used the distance of Lys-176 and Pro-223 as the reaction coordinate, similar settings as the simulations with DhlA and DCE, and the total simulation time was 104 × 50 ns. The Markov state models and bootstrapping were built similarly as described with DhlA and DCE. The implied time scales plot and the Chapman-Kolmogorov test of the obtained Markov state model are shown in Fig. S13.

Author contributions—P. K., Z. P., and J. D. resources; P. K., D. B., and V. D. data curation; P. K. and D. B. software; P. K., D. B., V. D., and Z. P. formal analysis; P. K., D. B., V. D., and Z. P. writing-original draft; D. B. and V. D. validation; Z. P. and J. D. conceptualization; Z. P. and J. D. funding acquisitions; J. D. supervision; J. D. project administration; J. D. writing-review and editing.

Acknowledgments—Computational resources were supplied by the Ministry of Education, Youth and Sports of the Czech Republic under the Projects CESNET (Project No. LM2015042) and CERIT-Scientific Cloud (Project No. LM2015085).

References

DhA changes conformation to accommodate bulky substrates


51. Schrödinger, LLC: The PyMOL Molecular Graphics System, Version 1.7.4


