Beyond the Chemical Step

The Role of Substrate Access in Acyltransferase from Mycobacterium smegmatis

Carvalho, Henrique F.; Mestrom, Luuk; Hanefeld, Ulf; Pleiss, Jürgen

DOI
10.1021/acscatal.4c00812

Publication date
2024

Document Version
Final published version

Published in
ACS Catalysis

ACS Catalysis

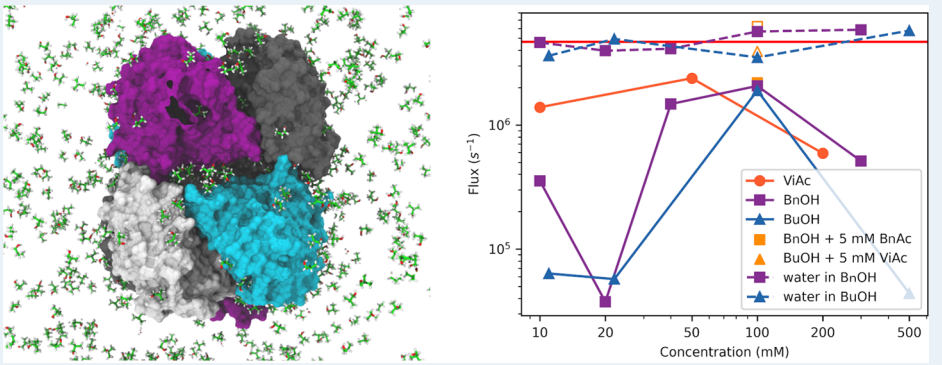
pubs.acs.org/acscatalysis Research Article

Beyond the Chemical Step: The Role of Substrate Access in Acyltransferase from *Mycobacterium smegmatis*

Henrique F. Carvalho, Luuk Mestrom, Ulf Hanefeld, and Jürgen Pleiss*

Cite This: *ACS Catal.* 2024, 14, 10077−10088 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Acyltransferase from *Mycobacterium smegmatis* is a versatile enzyme, which catalyzes the transesterification of esters in aqueous media due to a kinetic preference of the synthesis reaction over the thermodynamically favored hydrolysis reaction. In the active octamer, the active site is deeply buried and connected to the protein surface by long and hydrophobic substrate access channels. The role of the access channel in controlling catalytic activity and substrate specificity was investigated by molecular dynamics simulations and Markov-state models, and the thermodynamics and kinetics of binding of acyl donors, acceptors, and water were compared. Despite the hydrophobic nature of the substrate access channel, water is present in the channel and competes with the acyl acceptors for access to the active site. The binding free energy profiles in the access channel and the flux of butyl and benzyl alcohol and vinyl acetate were analyzed in the concentration range between 10 and 500 mM and compared to water. The flux showed a maximum at an alcohol concentration of 50−100 mM, in agreement with experimental observations. At the maximum, the flux of alcohol approaches 50% of the flux of water, which explains the high transesterification rate as compared to hydrolysis. The molecular origin of this effect is due to the accumulation of alcohol molecules along the access channel. Extensive molecular dynamics simulations and analysis of trajectories by a Markov-state model provided insights into the role of the access channel in activity and specificity by controlling access and binding of competing substrates.

KEYWORDS: *acylation, transesterification, binding free energy, kinetics, Markov-state model, reaction mechanism, molecular dynamics simulation*

INTRODUCTION

Evolved enzymes catalyze a chemical reaction through a series of partial reaction steps. In a hydrolysis reaction catalyzed by a serine hydrolase, the ester substrate binds to the protein surface and then diffuses to the active site. In the first partial reaction, an acyl enzyme is formed, and the alcohol product leaves the active site. Subsequently, a water molecule enters the active site as the substrate of the second partial reaction, releasing the acid product. For an enzyme such as butyrylcholine esterase, where substrate access is fast, the rate-limiting step is the acylation or deacylation reaction, depending on the substrate.[1] However, in other enzymes, the rate-limiting and specificity-determining step is not the chemical step, but the rate of substrate access through a tunnel,[2] product exit,[3] or conformational changes within the enzyme during the reaction.[4] Substrate access might govern not only the overall rate of a reaction but also other biochemical properties of an enzyme such as substrate specificity and regio- and stereoselectivity. Substrate access has been demonstrated to be controlled by gatekeeping residues[5] at the entrance or by the properties of the substrate access tunnel. As a consequence, engineering substrate tunnel has been a promising target for enzyme engineering for increasing reaction rates, shifting

Received: February 5, 2024
Revised: May 23, 2024
Accepted: June 12, 2024

https://doi.org/10.1021/acscatal.4c00812
ACS Catal. 2024, 14, 10077−10088

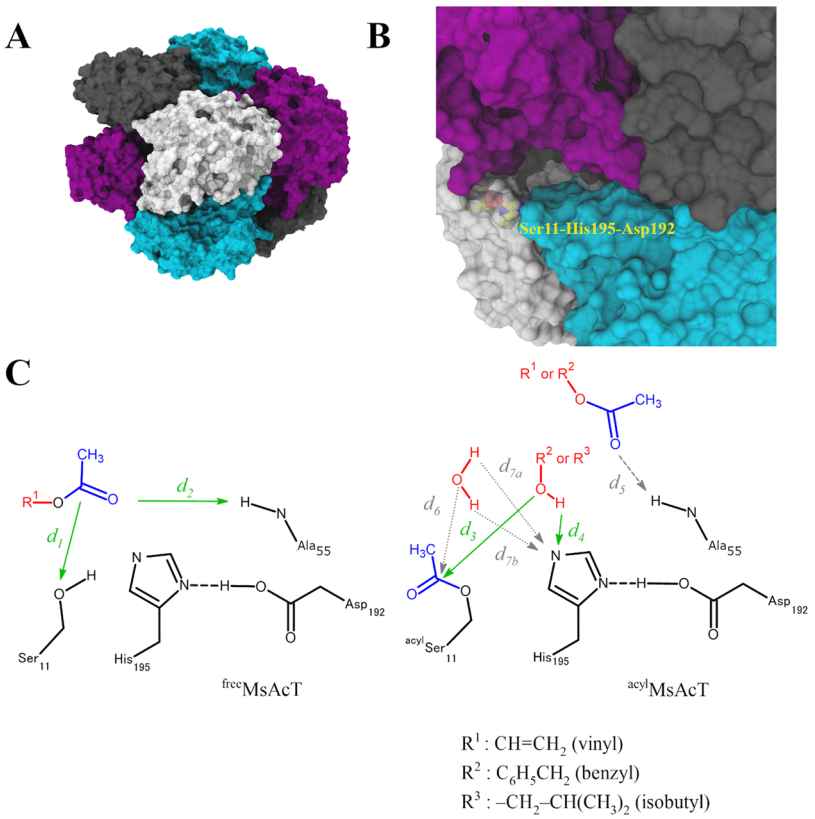


Figure 1. (A) *Ms*AcT octameric surface and (B) active site access tunnel in detail. Individual monomers in different surface colors representation. Catalytic His195 and Ser11 located within the tunnel in yellow surface representation. (C) Monitored substrate–protein interactions. Interactions of acyl donors (d_1–d_2) with the free-enzyme form (left) and interactions of acyl acceptors (d_3–d_4) and water (d_6–d_7) with the acyl enzyme (right).

substrate specificity, and optimizing regio- and stereoselectivity.[6,7]

Serine hydrolases such as esterases and lipases are widely used enzymes, which in aqueous media catalyze the stereoselective and regioselective hydrolysis of ester bonds.[8] Because most serine hydrolases are stable and active in organic solvents, they also catalyze the reverse reactions, esterification, and transesterification.[9] All serine hydrolases have the same catalytic mechanisms based on a catalytic triad of serine, histidine, and aspartic or glutamic acid.[10] The transition state is stabilized by residues of the oxyanion hole. One subclass of hydrolases, the acyltransferases, catalyze the transesterification of activated acids such as thioesters in aqueous media, although hydrolysis is thermodynamically preferred in the presence of water.[11] The observed high yield of ester product is caused by their kinetic preference of synthesis to hydrolysis.[12−15] The acyltransferase from *Mycobacterium smegmatis* (*Ms*AcT) is a versatile enzyme because it also accepts nonactivated esters such as ethyl acetate and shows an exceptionally wide substrate spectrum.[16,17] *Ms*AcT has been successfully applied to synthesize a broad range of esters,[12,18,19] alcohols,[20] peracids,[21] and amides.[22−25] Previously, the potential of *Ms*AcT as a catalyst for the enantioselective synthesis of cyanohydrins and secondary propargyl alcohols in aqueous media was assessed.[26]

In order to understand the exceptionally high transesterification activity of *Ms*AcT in aqueous media, the properties of the active site were investigated by computational studies[27,28] Because the intrinsic energy barriers of the transesterification and hydrolysis reaction were similar, it was proposed that a more efficient substrate binding was the basis for favoring transesterification over hydrolysis.

In addition to the shape and physicochemical properties of the active site, substrate access channels have been demonstrated to play a pivotal role in mediating selectivity and specificity in many hydrolases, such as lipases,[7,29] haloalkane dehalogenases,[2] and acylases.[30] *Ms*AcT forms an octamer with deeply buried active sites, which are connected to the protein surface by long and hydrophobic substrate access channels.[31] The aim of this study was to investigate the role of these channels in enzyme kinetics, selectivity, and inhibition.[2,32−35] Therefore, we modeled the binding of substrate and water and the accessibility to the active site of a *Ms*AcT octamer by molecular dynamics (MD) simulations, analyzed the trajectories by Markov-state models (MSMs) to evaluate kinetic constants, and performed a mutual information (MI) analysis to characterize the ligand–ligand and ligand–protein interactions.[36,37] MSMs have been increasingly applied to interpret enzyme kinetics[29,38−40] and to explore the rate-limiting molecular events beyond the chemical conversion step,[41,42] such as binding of substrate molecules to the protein surface and substrate access to the active site. By combining molecular modeling with experimental measurements provided novel insights into the mechanisms driving transesterification reactions of *Ms*AcT in aqueous media.

METHODS

Experimental Section. The transesterification reaction using *Ms*AcT for benzyl alcohol (BnOH) was experimentally characterized previously and the enzymatic constants determined, with $K_{M,BnOH}$ = 13.6 mM and $k_{cat,BnOH}$ = 2586 ± 2.4 s^{-1},

https://doi.org/10.1021/acscatal.4c00812
ACS Catal. 2024, 14, 10077−10088

and for vinyl acetate (ViAc), $K_{M,ViAc}$ = 0.0123 mM and $k_{cat,ViAc}$ = 182 ± 2.5 s^{-1}.[12]

*Ms*AcT was expressed, purified, and quantified as described earlier.[12,26] The enzymatic transesterification of ViAc by isobutanol (*i*-BuOH) is based on earlier work describing the same reaction for ViAc and BnOH and reporting the k_{cat} and K_m for both ViAc and BnOH above. The assay is based on the spectroscopic measurements of the conversion of NADH in the *Sc*ADH catalyzed reduction of acetaldehyde released during ester production.

The assay was performed in polyacrylate 1 cm cuvettes by monitoring the conversion of 0.25 mM NADH at 340 nm and 20 °C using an extinction coefficient of 6.221 mM^{-1} cm^{-1}. Before the addition of *Ms*AcT, the reaction was monitored until no chemical background hydrolysis of ViAc occurred. Enzymatic reactions were started by the addition of *Ms*AcT with a final concentration of isobutanol (0.0−200 mM), vinyl acetate (ViAc, 5.4 mM), NADH (0.25 mM), KPi (200 mM, pH 7.5), *Ms*AcT (24 ng mL^{-1}), and *Sc*ADH (50 U mL^{-1}, Sigma-Aldrich A7011). All measurements were performed in triplicates. The addition of additional *Sc*ADH did not result in a higher response for indirect acetaldehyde detection.

MD Simulations. The structure of *Ms*AcT was simulated in its homooctameric state (Figure 1), with all of the 8 subunits (chain A−H) modeled in unbound form (^{free}*Ms*AcT) or as acyl enzyme (^{acyl}*Ms*AcT), the latter corresponding to the native sequence containing the catalytic Ser11 covalently bound to an acetyl moiety. For construction of both forms, the crystal structure of *Ms*AcT (PDB ID: 2Q0S, 1.5 Å resolution)[31] was used. The adduct atoms were removed, and serine (Ser11) for ^{free}*Ms*AcT and acetyl-serine (^{acyl}Ser11) for ^{acyl}*Ms*AcT was constructed. In order to obtain force field parameters for ^{acyl}Ser11, its coordinates were extracted, and the structure was modified by acetylation and amidation of the termini. The OPLS[43] all-atom (OPLS-AA) force field parameters of the capped ^{acyl}Ser11 were obtained from the LigParGen server using the 1.14*CM1A-LBCC charge model.[44,45] The partial charges of the backbone atoms were adjusted to the values of the standard OPLS-AA force field, and the partial charges of the 10 side-chain atoms were modified by −0.005 each, in order to result in a neutral ^{acyl}Ser11 moiety. The nonbonded terms of atoms C_β, $H_{\beta1}$, and $H_{\beta2}$ as well as bond length, bond angle, and dihedral parameters of the C_β and Oγ bond were taken from OPLS-AA. The OPLS-AA force field for the protein was obtained by the tool *pdb2gmx* of GROMACS 2018.[46] A pH of 7.0 was assumed and the PROPKA 3 tool was used to calculate the protonation state of residues, with the side chains of Asp/Glu and Lys/Arg negatively and positively charged, respectively, terminal residues charged, and all histidine residues including the catalytic His195 were neutral, with the hydrogen on the $N_{\delta1}$ atom.[47] As a result, the *Ms*AcT octamer had a net charge of −72. OPLS-AA parameters for isobutyl alcohol, benzyl alcohol, benzyl acetate, vinyl acetate, and benzyl acetate were retrieved from the LigParGen server, using as input the corresponding SMILES codes.

MD simulations were performed using GROMACS version 2018[46] with an integration time of 2 fs, a constant pressure of 1 bar, and a constant temperature of 294 K (NPT ensemble) under periodic boundary conditions. The v-rescale algorithm was used for temperature coupling,[48] and the Berendsen algorithm was used for pressure coupling with τ_p = 10 fs.[49] Electrostatics were calculated by a smooth particle mesh Ewald summation.[50] A single octameric protein structure was placed in the center of a cubic simulation box, and variable number of substrate molecules were inserted at random locations and orientations. The box size (from 15 to 25 nm) and number of ligand molecules (from 50 to 1350) were chosen to obtain ligand concentrations ranging from 5 to 500 mM (Table S1). In total, 14 protein−ligand mixtures at different concentrations of ligand were prepared: 3 of ^{free}*Ms*AcT/ViAc, 4 of ^{acyl}*Ms*AcT/*i*-BuOH, 5 of ^{acyl}*Ms*AcT/BnOH, 1 ternary mixture of ^{acyl}*Ms*AcT/*i*-BuOH/ViAc, and 1 of ^{acyl}*Ms*AcT/BnOH/BnAc. For each ligand concentration, 10 independent replicates were prepared by the use of random seeds for ligand location and orientation (total 140 systems). The boxes containing the protein and ligand molecules were solvated with SPC/E water molecules[51] in variable number (from 4.2 × 10^5 to 1.6 × 10^6, Table S1). Net charge neutralization was made by replacement of 72 water molecules with Na^+ ions. Replicates at each concentration varied according to the position and orientation or the ligands and the total number of atoms (from 4.6 × 10^5 to 1.6 × 10^6, Table S1). Each system was first minimized for a maximum of 10,000 steps by a steepest descent algorithm (force threshold of 1000 $kJ{\cdot}mol^{-1}{\cdot}nm^{-1}$), followed by 1000 steps of the conjugate gradient algorithm (force threshold of 400 $kJ{\cdot}mol^{-1}{\cdot}nm^{-1}$). Subsequently, a short solvent equilibration simulation was made for 5 ns to allow for water molecules to equilibrate in the box by applying position restraints on ligand atoms and on protein heavy atoms (force constant of 1000 $kJ{\cdot}mol^{-1}{\cdot}nm^{-1}$). Finally, each system was simulated for 200 ns, and trajectory frames were saved every 5 ps. As control, ^{acyl}*Ms*AcT with only water was simulated. In total, 30 μs of simulation data were generated, analyzed, and visualized by VMD.[52]

Featurization. *Ms*AcT-catalyzed transesterification in aqueous conditions has been described by two half-reactions: the acylation of the free enzyme by an acyl donor and the transfer of the acyl group from the acyl enzyme to an acyl acceptor.[27] The acylation of the free enzyme proceeds when an ester substrate binds to the active site, followed by a nucleophilic attack of the catalytic serine, thus forming the acyl-enzyme intermediate and releasing the first alcohol product. The transfer of the acyl group proceeds when either a water or an alcohol acts as the second substrate by performing a nucleophilic attack to the acyl-enzyme intermediate, which leads to the release of the final product. The two half-reactions were modeled separately. The acylation was modeled as a protein/ester mixture with ^{free}*Ms*AcT and vinyl acetate (ViAc) as acyl donors. The acyl transfer was modeled as a protein/alcohol mixture with ^{acyl}*Ms*AcT and either BnOH or *i*-BuOH as acyl acceptors. The competing hydrolysis reaction was modeled using SPC/E water molecules acting as the solvent in protein/alcohol mixtures. Additionally, two ternary protein/ester/alcohol mixtures were prepared: the first with ^{acyl}*Ms*AcT, the substrate ViAc, and the acyl acceptor *i*-BuOH; the second with ^{acyl}*Ms*AcT, the product benzyl acetate (BnAc), and the acyl acceptor *i*-BnOH. For each reactant, the pathway of access to each of the 8 active sites of the *Ms*AcT octamer was investigated. Trajectories were featurized using the near attack conformation (NAC)[53] concept in the context of serine hydrolases,[29] by using the distance toward NAC formation (d_{NAC}) as a continuous one-dimensional latent variable

https://doi.org/10.1021/acscatal.4c00812
ACS Catal. 2024, 14, 10077−10088

$$d_{NAC} = \sqrt{\frac{\sum_{i=1}^{D} d_i^2}{D}}$$

where D is an arbitrary number of distances d used to describe a ligand pose relative to the protein. For the modeled *Ms*AcT/ligand mixtures, we defined distances between the functional atoms of the ligands and the corresponding active site atoms of *Ms*AcT (Figure 1). For the acylation reaction, $d_{NAC,donor}$ was calculated from the distance between the carbonyl carbon of the acyl donor and the hydroxyl oxygen of Ser11 (d_1) and the distance between the carbonyl oxygen of the acyl donor and the backbone amide hydrogen of Ala55 (d_2). d_1 represents the nucleophilic attack by the catalytic Ser11 residue and d_2 the stabilization of the substrate by the oxyanion hole residues. For the acyl transfer reaction, $d_{NAC,acceptor}$ was calculated from the distance between the hydroxyl oxygen of the acyl acceptors and the carbonyl carbon of ^{acyl}Ser11 (d_3) and the distance between the hydroxyl hydrogen of the acyl acceptors and the *tele* nitrogen of His195 (d_4). d_3 and d_4 represent the nucleophilic attack by the substrate on the acyl-serine intermediate. For the ternary systems, the distance between the carbonyl oxygen of the ester and the backbone amide hydrogen of Ala55 (d_5) of the oxyanion hole was monitored to represent the approach of the ester molecules toward the active site. For the hydrolysis reaction, the distance between the oxygen of water and the carbonyl carbon of ^{acyl}Ser11 (d_6) and the distance between the hydrogen and the nitrogen of His195 ($d_7 = \min(d_{7a}, d_{7b})$) was monitored to represent the nucleophilic attack of the water molecule toward the acyl-serine intermediate. Distances d_1−d_7 were calculated independently for each replicate ($n_{replicates}$ = 10), between all active sites ($n_{references}$ = 8), and all ligand molecules ($n_{molecules}$) for each frame of the 200 ns production runs (n_{frames} = 40,000), using the MDAnalysis Python library.[54,55] The corresponding d_{NAC} values were calculated and encoded as a three-dimensional matrix ($n_{frames} \times n_{molecules} \times n_{references}$), which was analyzed by *in-house* built Python scripts.[56] The range of d_{NAC} values is specific to each system: the lower limit is due to van der Waals repulsive terms of the OPLS-AA force field, and the upper limit is the size of the simulation box. The three-dimensional (3D) matrix describing a d_{NAC}-featurized trajectory varied along the $n_{molecules}$ dimension as the number of ligand molecules varied for each system, while the number of frames and subunits was constant.

Discretization. *Binding Free Energy Profile.* The d_{NAC}-featurized trajectories were used to generate ligand binding free energy profiles.[29,30] For each ligand concentration, trajectories were discretized based on the distribution of d_{NAC} values in a set of spherical shells with origin at d_{NAC} = 0 Å and with 0.25 Å thickness. For each shell i, the average ($N(i)_{enzyme}$) and standard deviation of d_{NAC} counts of all references, all molecules, and all replicates were calculated over the total number of sampled frames. Shell distributions were converted to the corresponding free energy values by the Boltzmann equation

$$\frac{\Delta G(i)}{k_B T} = -\ln \frac{N(i)_{enzyme}}{N(i)_{reference}}$$

where k_B is the Boltzmann constant and T is the temperature in Kelvin. The reference distribution $N(i)_{reference}$ of d_{NAC} values corresponds to the number of counts for each shell i assuming an isotropic distribution of molecules at an arbitrary concentration in the absence of *Ms*AcT. $N(i)_{reference}$ is dependent on the effective bulk concentration of the ligand after equilibration. The effective bulk concentration in the simulation boxes was determined by fitting $N(i)_{reference}$ to $N(i)_{enzyme}$ in the subset of shells far from the protein surface. The resulting binding free energy profiles were used to identify regions in d_{NAC} latent space with a repulsive or attractive potential toward NAC formation across different ligand concentrations. Convergence of the resulting profiles was evaluated based on the analysis of associated errors obtained by different combinations of time intervals and replicates (data not shown).

Collective Ligand Encoding. The simultaneous sampling of regions in d_{NAC} latent space by different ligand molecules was evaluated using a combinatorial discretization scheme.[29] The d_{NAC} latent space was partitioned based on common features identified in binding free energy profiles across the ligand and concentration space. The number of resulting shells as well as their limits and thickness were set fixed for all *Ms*AcT−ligand mixtures, and five regions were defined: active site (**A**), tunnel (**T**), entrance to the tunnel (**E**), surface of the protein (**S**), and bulk (**B**). For every 5 ps trajectory frame, each region was checked to be *sampled* if there was at least one d_{NAC} value among all protein−ligand pairs to be contained within its boundaries, or *not sampled* otherwise. By considering the sampling of the five regions, 2^5 states were defined, with each state defined according to the set of *sampled* regions, e.g., *AB* if only regions **A** and **B** were *sampled*, or *AESB* if only the region **T** was *not sampled*. This combinatorial discretization scheme is independent of the number of ligand molecules and acts as a dimensionality reduction technique by generating a two-dimensional state-discretized trajectory of size $n_{frames} \times n_{references}$. This was especially useful for analyzing water ($10^5 < n_{molecules} < 10^6$), where handling of a single d_{NAC}-featurized trajectory without dimensionality reduction would have excessively large memory requirements.

Ligand−Water Encoding. The state-discretized trajectories of the ligands and water were combined and analyzed further. Only the effective set of sampled states were considered: 10 ligands states (*ATESB*, *ATES*, *AESB*, *TESB*, *ATSB*, *ESB*, *ASB*, *TSB*, *SB*, and *B*) and 2 water states (*ATESB* and *TESB*), yielding 10 × 2 ligand−water states that were labeled using the superscript *Water* when the water *ATESB* state was sampled.

Mutual Information. The state-discretized trajectories were used for mutual information measurements as well as correlation coefficient determination using scikit-learn.[57] For the binary systems, the trajectories were compared along the $n_{references}$ dimension to evaluate cross-talk between different active site subunits. For ternary systems, the state-discretized trajectories of each ligand were used to evaluate the cross-talk between different ligand molecules. For water, the state-discretized trajectories of water were used to evaluate the cross-talk between the state-discretized trajectories of acyl acceptor molecules.

Markov-State Models. The state-discretized trajectories of size $n_{frames} \times n_{references}$ were used to evaluate the kinetics of transitions of the ligands between different regions along the binding pathway. Therefore, the state-discretized trajectories were concatenated to a single trajectory of length $n_{frames \times references}$, which was used for generating a Bayesian MSM using the PyEMMA software package.[58] The lag times τ selected for each system were validated by the convergence of implied time scales and by the Chapman−Kolmogorov test.[59−61] The mean first passage times (MFPT) between

https://doi.org/10.1021/acscatal.4c00812
ACS Catal. 2024, 14, 10077−10088

states as well as their corresponding stationary distributions were calculated for all concentrations and for all protein−ligand systems. Transition Path Theory[62−64] was employed to evaluate the flux of the binding pathway from the Bayesian MSMs. The states used as source or sink varied for each protein−ligand mixture but were kept fixed for different concentrations. The resulting forward/backward committor probabilities, pathways, and net fluxes were combined to evaluate the effects of the ligand concentration on the kinetics of active site access along the binding pathway.

RESULTS

Kinetic Characterization of Acyltransferase from *M. smegmatis* (*Ms*AcT). In an extension of the previous coupled spectrophotometric assay, the kinetics of *Ms*AcT upon transesterification of vinyl acetate by *i*-BuOH to isobutyl acetate was investigated (Figures 2 and S1−S4).[12] A K_M value

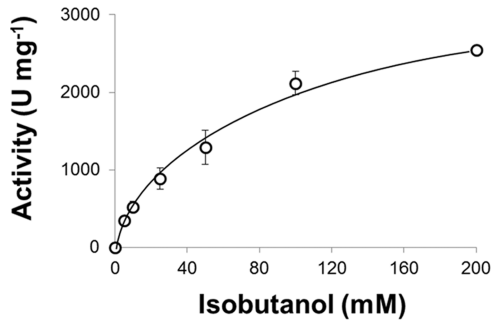


Figure 2. Enzymatic synthesis of isobutyl acetate using *i*-BuOH and ViAc. Reaction conditions: *i*-BuOH (0−200 mM), ViAc (5.4 mM), NADH (0.25 mM), KPi (200 mM, pH 7.5), *Ms*AcT (24 ng mL^{-1}), and ScADH (50 U mL^{-1}). The error bars show the standard deviation of triplicates.

of 72 ± 12 mM and v_{max} value of 3480 ± 260 U mg^{-1} were determined. Thus, the K_M toward isobutyl acetate was higher than the previously determined K_M (13.6 mM) toward BnOH, most likely due to a lack of aromatic interactions with the aromatic active site residues within *Ms*AcT. The hydrolysis of isobutyl acetate was not investigated.

Molecular Dynamics Results. *Transesterification Reaction.* The interactions between *Ms*AcT and substrates of the transesterification reaction were monitored by featurization of MD simulation trajectories with d_{NAC} as the latent variable, which encodes the location of substrate molecules relative to the active sites of the *Ms*AcT octamer. Trajectory frames were visually inspected to map the latent variable representation to the location of substrate molecules along the substrate access pathway from the protein surface to the active site (Figure 3A). Binding free energy profiles of ligands were calculated based on the distribution of discretized d_{NAC} values, using as reference their distribution at the same bulk concentration in the absence of the protein. The resulting binding free energy profiles (Figure 3B) encode the regions of the d_{NAC} space with attractive or repulsive Gibbs free energy value ΔG (in units of k_BT). Significant differences were found between the acyl donor vinyl acetate (ViAc), and the two acyl acceptors BnOH and *i*-BuOH. For all substrates and all concentrations, there was either a local or global ΔG minimum at $d_{NAC} \approx 4$ Å, which corresponds to a productive binding pose of substrate molecules within the active site pocket. At lower values of d_{NAC}, there was a steep increase in ΔG values, caused by steric repulsion between the substrate and the side chains of active site residues. We focused on the details of the access pathway toward the productive binding pose at $d_{NAC} \approx 4$ Å in the following sections.

The acyl donor ViAc was modeled at concentrations of 10, 50, and 200 mM. The binding free energy profile revealed a mostly attractive potential along the access pathway. There was an inflection point at $d_{NAC,donor} \approx 6$ Å, corresponding to molecules located within the tunnel that connects the catalytic dimer interface and the active site pocket. For the lowest concentration of 10 mM, the binding potential was more attractive at the active site while the local maximum, located at the tunnel entrance, was at its highest value. For concentrations of 50 and 200 mM, the attractive potential shifted toward the tunnel entrance. Above 8 Å, the ΔG values were similar at all concentrations, indicating an overall high affinity of ViAc toward the ^{free}*Ms*AcT, including the catalytic dimer interface.

The acyl acceptor BnOH was modeled at concentrations of 10, 20, 40, 100, and 300 mM. The binding free energy profile changed significantly at different concentrations. The presence of the global ΔG minimum at $d_{NAC,acceptor} \approx 4$ Å was found only at concentrations higher than 40 mM. A local ΔG minimum occurred within the tunnel at $d_{NAC,acceptor} \approx 7$ Å, which was more pronounced at concentrations below 40 mM. At these concentrations, the BnOH molecules accumulate within the tunnel and sample the active site pocket only rarely, resulting in high uncertainties of ΔG values at $d_{NAC,acceptor} < 5$ Å. At higher concentrations, and for ΔG values within the $4 < d_{NAC,acceptor} < 10$ Å region, the uncertainty was largely reduced, indicating accumulation of molecules in the tunnel and the active site. For all BnOH concentrations, a second local minimum was observed at $d_{NAC,acceptor} \approx 15$ Å, corresponding to the tunnel entrance located at the catalytic dimer interface, similar to the local minimum found for ViAc. The experimentally determined dissociation constant of $K_{M,BnOH}$ = 13.6 mM is close to the concentrations (10 and 20 mM) where accumulation of BnOH molecules occurred in the tunnel.[12] Above these concentrations, the BnOH molecules sampled the active site more extensively. Interestingly, no saturation was observed for ViAc even at the highest concentration of 200 mM, despite $K_{M,ViAc}$ being three orders of magnitude lower than that of $K_{M,BnOH}$. This points to different binding mechanisms of esters and alcohols, which will be discussed below.

The second acyl acceptor, *i*-BuOH, was modeled at concentrations of 11, 22, 100, and 500 mM. The binding free energy profiles exhibited a distinct concentration dependency from those of BnOH. For $d_{NAC,acceptor} < 7$ Å and for low *i*-BuOH concentrations (11 and 22 mM), ΔG was positive and had a high uncertainty, indicating a low sampling of the tunnel and active site. At higher concentrations (100 and 500 mM), above the experimentally determined $K_{M,i\text{-}BuOH}$, ΔG was still repulsive although the uncertainty decreased. In contrast to the other substrates, a saturation of the tunnel entrance was observed at 500 mM, resulting in a global ΔG minimum at $d_{NAC,acceptor} \approx 15$ Å. *i*-BuOH shares a similar tunnel signature compared to that of ViAc, given the presence of a local maximum at $d_{NAC,acceptor} = 8$ Å for all concentrations.

For all substrates, the latent variable d_{NAC} yielded consistent results, both at the level of energy variations ($-2.5 < \Delta G < 7.5$ k_BT) and range under the influence of the protein ($2 < d_{NAC} <$

https://doi.org/10.1021/acscatal.4c00812
ACS Catal. 2024, 14, 10077−10088

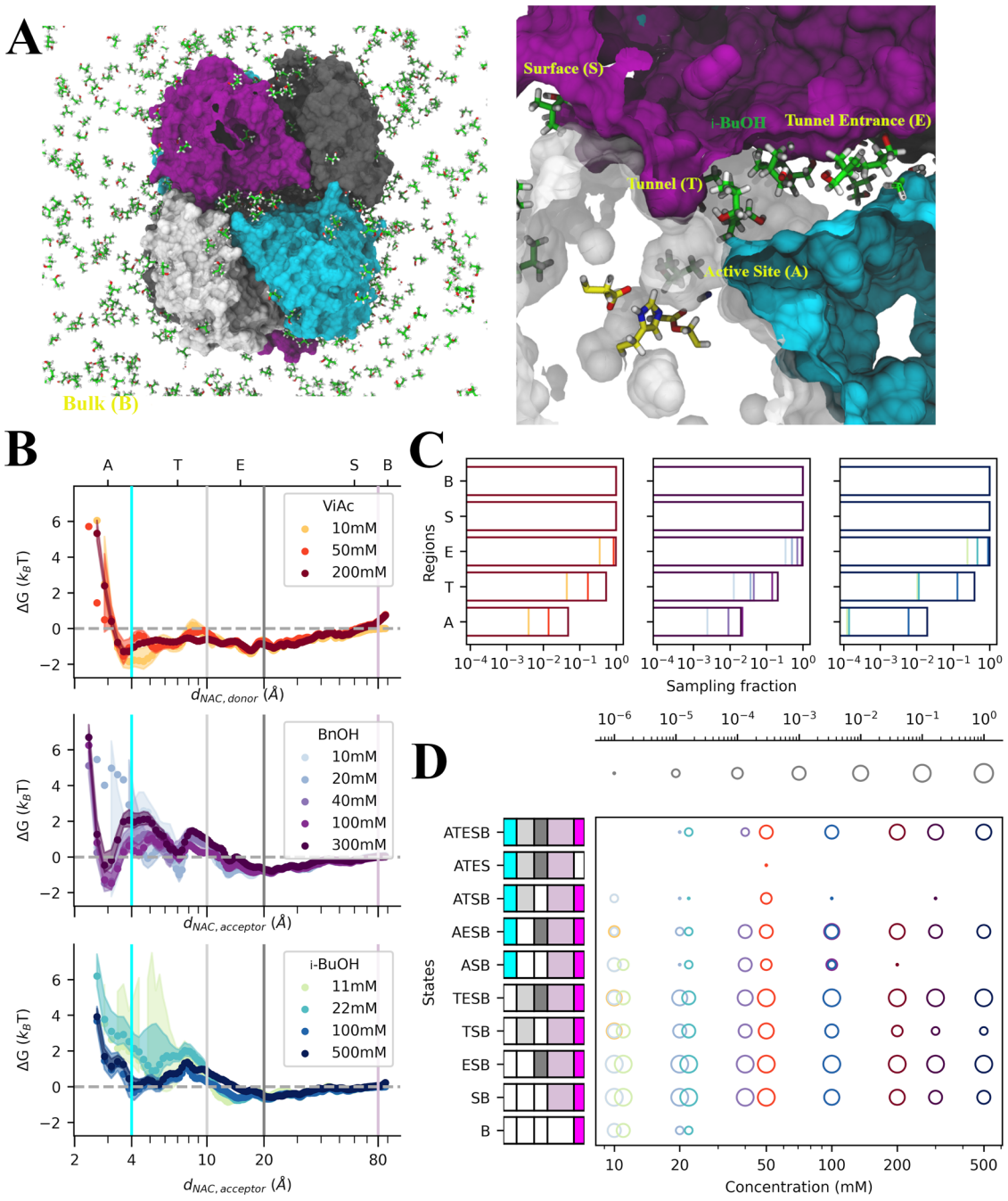


Figure 3. Binding mechanism of transesterification substrates. (A) Snapshot of substrate *i*-BuOH distribution in the *Ms*AcT surface and within the active site tunnel. *i*-BuOH molecules are shown in green licorice representation. Corresponding locations along d_{NAC} regions have yellow labels. (B) Binding free energy profiles of d_{NAC} as a function of substrate concentration. Circles correspond to the mean ΔG values (in k_BT units) per d_{NAC} shell over all calculated protein−substrate pairs, subunits, and replicates. Shaded regions correspond to the associated standard deviations. A, T, E, and S d_{NAC} shell maximal radii in colored solid lines: A—cyan, T—light gray, E—dark gray, S—gray, and E—pink. Values of d_{NAC} are represented in the logarithmic scale for clarity. (C) Sampled regions as a function of substrate concentration. (D) Sampled states as a function of concentration. Circles correspond to the sampling fraction in logarithmic scale for clarity.

80 Å). The major differences found between substrates and across concentrations were within the $4 < d_{NAC} < 20$ Å region, indicating that the substrates had distinct interactions with the protein along the binding pathway. The ester substrate ViAc displayed more favorable interactions with the protein and less saturation than the two alcohols. *i*-BuOH had a more repulsive potential toward the active site, which is consistent with its 5-fold higher experimentally determined dissociation constant compared to that of BnOH.

To further investigate the binding kinetics of the three substrates, we assigned five regions **A**, **T**, **E**, **S**, and **B** along the substrate access pathway (Figure 3C), based on the ΔG profiles: the active site (**A**, $d_{NAC} \leq 4$ Å), the tunnel (**T**, $4 < d_{NAC} \leq 10$ Å), the entrance to the tunnel, located at the dimer catalytic interface (**E**, $10 < d_{NAC} \leq 20$ Å), the remaining octameric protein surface (**S**, $20 < d_{NAC} \leq 80$ Å), and the bulk (**B**, $d_{NAC} > 80$ Å). Sampling of regions **A**, **T**, and **E** depended sensitively on the substrate concentration, whereas at all

https://doi.org/10.1021/acscatal.4c00812
ACS Catal. 2024, 14, 10077−10088

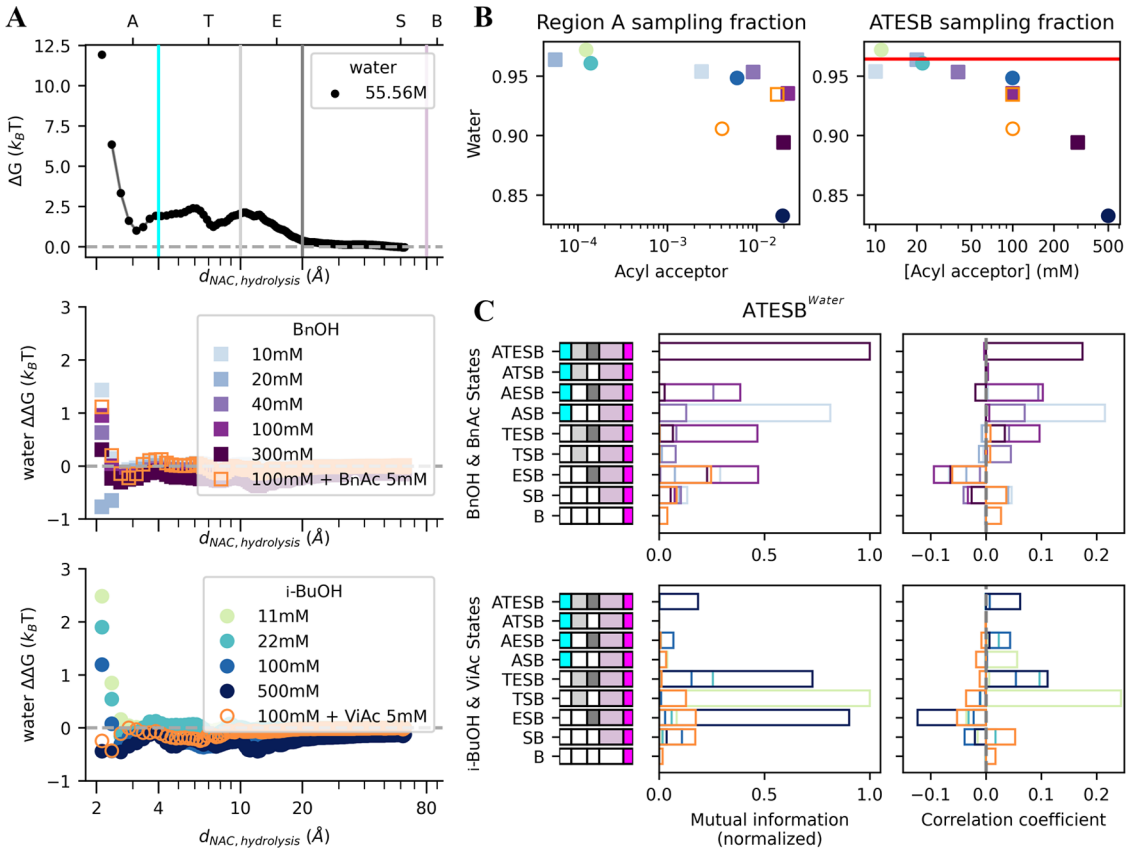


Figure 4. Binding mechanism of water. (A) Binding free energy profiles of d_{NAC} for pure water (top) and under the presence of BnOH (middle) and *i*-BuOH (bottom). Circles correspond to the mean ΔG values (in k_BT units) per dNAC shell over all calculated protein−substrate pairs, subunits, and replicates. Shaded regions correspond to the associated standard deviations. A, T, E, and S d_{NAC} shell maximal radii in solid lines: A—cyan, T—light gray, E—dark gray, S—gray, and E—pink. Values of d_{NAC} are represented in the logarithmic scale for clarity. (B) Sampled region **A** by water as a function of acyl acceptor sampling (left) and sampled *ATESB* by water as a function of acyl acceptor concentration (right). (C) Mutual information (left) and correlation coefficients (right) between water *ATESB* states and BnOH, *i*-BuOH states.

concentrations, substrate molecules were found in the regions **S** and **B**. Sampling of region **A** varied according to the corresponding K_M values, i.e., it was highest for ViAc, followed by BnOH and lowest for *i*-BuOH. Sampling of **T** was also highest for ViAc, followed by *i*-BuOH and BnOH. This inversion between BnOH and *i*-BuOH is a consequence of the saturation of BnOH in the tunnel region. The region **E** became saturated in all substrates at increasing concentration, which was most pronounced for *i*-BuOH, followed by BnOH and ViAc. At the level of the catalytic interface, we note that regions **A** and **T** were sampled independently between the constituting subunits for all ligands (Supporting Information and Figure S5A).

Further elucidation of the binding pathways of the three substrates was obtained by employing a combinatorial discretization scheme that allows to identify which regions (**A**, **T**, **E**, **S**, and **B**) are simultaneously sampled at a given time step of the trajectory (Figure 3D, details in the Methods section). All of the ligands sampled a similar set of states. In the *B* state, the substrate molecules are only present in bulk, and it was only sampled at the lowest concentrations. In contrast, the *ATESB* state, which corresponds to all regions being sampled, was increasingly sampled at higher substrate concentrations. This was the case for the other partially bound states such as the states *AESB*, *TESB*, *ESB*, and *SB*. The states *ATSB*, *ASB*, and *TSB* were sampled only at the lowest concentrations, corresponding to access to the active site without saturation of the full binding pathway. At the level of the catalytic interface, we note that the states that include the entrance and the tunnel (*TESB*, *TSB*, *ESB*) were mutually sampled by the two neighboring subunits, meaning that binding of ligand molecules in these regions in one subunit favored sampling of the corresponding regions in the other subunit (Supporting Information and Figure S5B).

The results of ternary systems show, in a manner similar to those of binary systems, that the binding pathway is prone to saturation within the tunnel and catalytic dimer interface when two different types of ligands are present (Supporting Information and Figure S6), which makes the system prone to substrate and product inhibition.

Hydrolysis Reaction. The interactions between *Ms*AcT and water were monitored through featurization of simulation trajectories with the $d_{NAC,hyd}$ latent variable, which allowed for individual tracking of all solvent water molecules relative to the eight catalytic subunits (Figure 4A). For the ^{acyl}*Ms*AcT in pure water, the binding free energy profiles show that the potential is repulsive along the full binding pathway with barriers similar to those for the ester and the two alcohol substrates. The repulsive potential is caused by saturation of the active site and tunnel with water molecules, which are filled with water molecules despite their hydrophobicity. Binding free energy profiles of water were also calculated for the acyl-enzyme systems with and without alcohols. The difference in $\Delta\Delta G$ was small along the binding pathway, except for the active site

https://doi.org/10.1021/acscatal.4c00812
ACS Catal. 2024, 14, 10077−10088

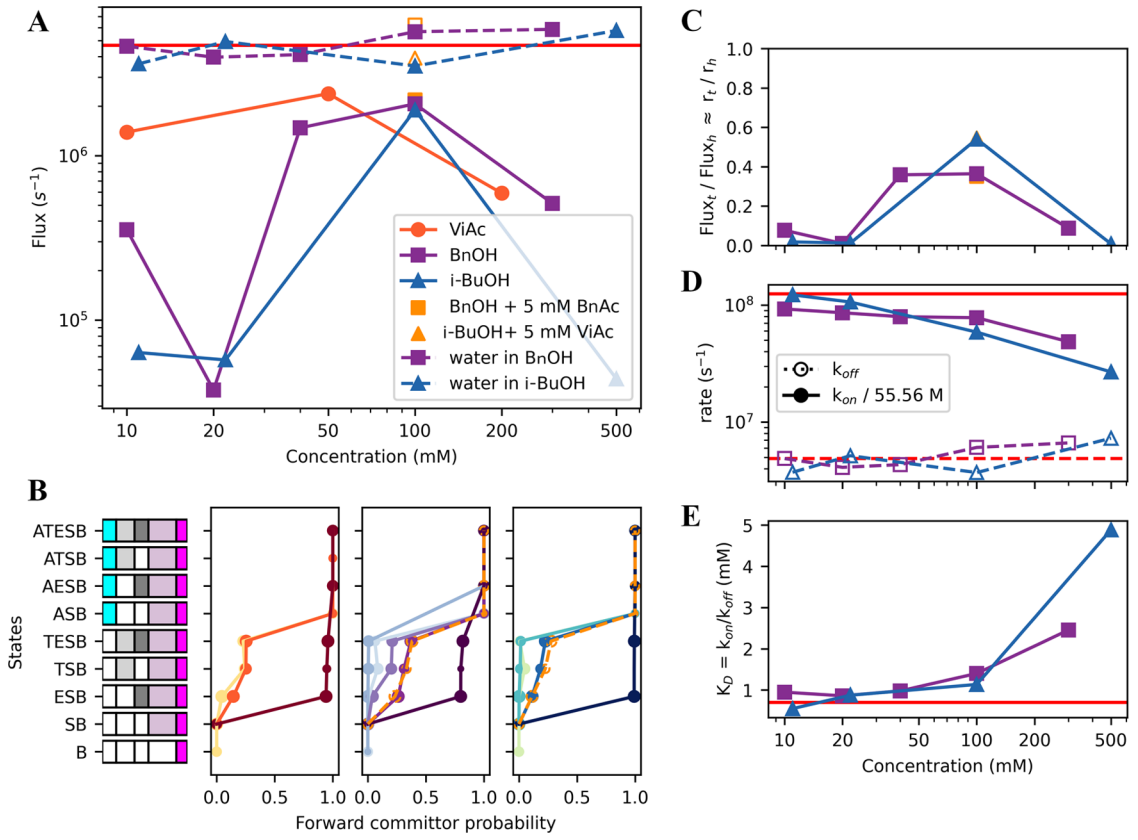


Figure 5. Kinetics of the binding of transesterification substrates and water. (A) Net flux (in s^{-1}) of reactive trajectories as a function of substrate concentration. Source states: *B*, *SB*; sink states: *ASB*, *AESB*, *ATSB*, and *ATESB*. Red line corresponds to net flux of pure water. (B) Forward committor probabilities of individual states for each substrate and concentration (ViAc left, BnOH center, and *i*-BuOH right). States connected at each concentration identified by solid colored lines (source and sink states have forward committor probabilities of 0 and 1 by definition). (C) Ratios of acyl acceptor-to-water net fluxes as a function of acyl acceptor concentration. (D) Binding (k_{on}, solid circles) and unbinding (k_{off}, open circles) rates of water as a function of acyl acceptor concentration. Red lines correspond to k_{on} (solid) and k_{off} (dashed) of pure water. (E) Dissociation constant for water as a function of acyl acceptor concentration. Solid lines correspond to the dissociation constant in pure water.

pocket, where $\Delta\Delta G$ reached up to 1.5 k_BT in the case of BnOH and up to 2.5 k_BT in the case of *i*-BuOH (Figure 4A). This indicates that the presence of acyl acceptors leads to a displacement of water molecules from the active site. Indeed, the sampling of region **A** by water (Figure 4B) also varied considerably for simulations containing the acyl acceptors, with a decrease of 12% in the sampling fraction when the concentration of acyl acceptors was equal or higher than 100 mM. Because water was always present in regions **T**, **E**, **S**, and **B**, only two water states were sampled, $TESB^{Water}$ and $ATESB^{Water}$ (Figure 4B). Sampling of the $ATESB^{Water}$ state decreased as the concentration of the acyl acceptors increased, pointing to a competition between water and the two acyl acceptors for active site access. We note that for the ternary systems, there was no further decrease in the $ATESB^{Water}$ sampling fraction compared to that of the binary system containing only BnOH, since BnAc does not sample the active site. However, for the ternary system containing *i*-BuOH and ViAc, there was a significant drop in the level of sampling of the active site by water because ViAc also samples the active site.

Further details of binding effects within the active site were obtained by mutual information and correlation coefficient measurements. Mutual information provides information about the simultaneous occurrence of ligand and water states, while the correlation coefficient provides insights if the mutual occurrence is linear. In the case of BnOH, the mutual information was highest between the states $ATESB^{Water}$ and $ATESB^{BnOH}$ at 300 mM BnOH (Figure 4C). The high mutual information and a positive correlation coefficient between the two states that include region **A** indicate that both BnOH and water simultaneously sample the active site. A different scenario was obtained in the case of *i*-BuOH, where the highest mutual information was obtained for the states $ATESB^{Water}$ and $TSB^{i\text{-}BuOH}$ at 11 mM, with a positive correlation coefficient, followed by the set of states $ATESB^{Water}$ and $ESB^{i\text{-}BuOH}$ or $TESB^{i\text{-}BuOH}$ at 500 mM, with varying correlation coefficients. Therefore, sampling of the active site by water did not occur simultaneously with sampling by *i*-BuOH. Considering this difference between the two acyl acceptors and the higher differences in ΔG of water in the *i*-BuOH system, the results suggest that BnOH binding does not require desolvation of the active site, in contrast to *i*-BuOH binding.

Binding Kinetics. The binding kinetics of the active site access was modeled by MSM models, using as input the state-discretized trajectories of the acyl donor, the two acyl acceptors, and water. The lag times for each ligand and concentration were chosen based on the convergence of implied time scales and the Chapman−Kolmogorov test. The flux of molecules diffusing from bulk to the active site was calculated using TPT by considering [*B*, *SB*] as the source

https://doi.org/10.1021/acscatal.4c00812
ACS Catal. 2024, 14, 10077−10088

states and [*ASB*, *AESB*, *ATSB*, *ATESB*] as the sink states (Figure 5A). For water, the states *TESB* and *ATESB* were considered as the source and sink states, respectively.

The flux for water was highest and remained nearly constant in both binary and ternary systems of BnOH or *i*-BuOH. In contrast, the fluxes of the three substrates were concentration-dependent, with a maximum at 100 mM for *i*-BuOH and BnOH and 50 mM for ViAc. For all ligands, the peak flux was approximately 50% of the water flux. The decrease of flux at the highest concentrations is consistent with the gradual saturation of the binding pathway at higher concentrations.

The analysis of forward committor probabilities provided further insights into the kinetics of active site access for the acyl acceptors (Figure 5B). At low concentrations, the forward committor probabilities of the intermediate states *ESB*, *TSB*, and *TESB* were low, indicating a high probability of molecules going back to the bulk after entering the tunnel and its entrance rather than proceeding toward the active site. At high concentrations, the forward committor probabilities of these intermediate states was close to unity, indicating that molecules sampling the tunnel and its entrance were more likely to move toward the active site than the back to the bulk. From the analysis of fluxes and forward committor probabilities, the access of the two acyl acceptors to the active site is slow and unrestrained at low concentrations, fast and restrained at optimal concentrations, and slow and restrained at high concentrations.

The ratio of 0.5 between the flux of the acyl acceptors and the flux of water (Figure 5C) was similar to the experimentally determined ratio between the transesterification rate and hydrolysis rate (r_t/r_h) of about 0.5. Considering the differences in concentrations between acyl acceptors and water, the transesterification reaction is largely favored given the fast kinetics of acyl acceptor binding.

Finally, we analyzed whether the presence of the two acyl acceptors had an effect on the binding/unbinding kinetics of water to the active site (Figure 5D). The effect was more pronounced on the binding (k_{on}) rate of water than on the unbinding rate (k_{off}), indicating that the presence of acyl acceptors leads to a slowdown of the access of water to the active site. The binding/unbinding rates were used to calculate the dissociation constant of water, K_D (Figure 5E), which increased at the highest concentrations by 2-fold in the case of BnOH and almost 5-fold in the case of *i*-BuOH. Together, these results show that the transesterification reaction benefits from the accumulation of acyl acceptor molecules along the binding pathway to the point where the kinetics of active site access approach those of water, despite the large differences in concentration.

DISCUSSION

It is been questioned why enzymes have such a complex structure and large size,[65] whereas the active site consists of only a few residues. The task of an enzyme is attracting desired substrates, preorienting them and guiding them toward the active site, placing them in the active site according to the desired selectivity, and catalyzing a series of chemical reactions by stabilizing the respective transition states. At the same time, the role of the enzyme is to protect the active site from unwanted compounds, which might block the active site and compete with the desired substrates. Attraction of uncharged substrates is achieved by binding to the protein surface,[30,66] followed by a two-dimensional diffusion on the protein surface. Substrate specificity is mainly achieved within the active site access channel by a 1000:1 enrichment of the desired substrate (alcohol) compared to the undesired substrate (water). Thus, the substrate specificity of an enzyme such as *Ms*AcT is obtained by a combination of positive and negative design (attracting the desired and occluding the undesired substrate), and the substrate access channel plays a crucial role in explaining the observed biochemical properties and in engineering of improved enzymes. Despite the wide range of promising applications endowed by *Ms*AcT as a transesterification catalyst in aqueous solvents,[16] so far only few studies have addressed the molecular origins of the kinetically favored transesterification reaction. Two major driving forces for transesterification have been proposed: a preferred binding of the alcohol molecule in the hydrophobic active site[28] or a relative shift of activation energies toward the acyl transfer reaction.[27,67] We propose that the kinetics of transesterification and hydrolysis is dominated by the thermodynamics and kinetics of the binding pathways preceding the chemical step: the access path taken by substrate molecules from the bulk toward the eight active sites of the *Ms*AcT octamer.

The analysis of extensive molecular dynamics simulations by Markov-state models demonstrated that thermodynamics and kinetics of the binding pathways differ, with free energies of binding being more favorable for acyl acceptors than for water due to the hydrophobic nature of the active site pocket and the tunnel. Nonetheless, these favorable interactions with alcohol molecules do not prevent water from accessing the active site since water is found both in the tunnel and in the active site pocket even when the alcohol molecules are also present. Therefore, the origin of the high transesterification-to-hydrolysis ratio can be attributed to the kinetics of binding. Despite the concentrations of alcohols and water differing by more than two orders of magnitude, the rates of alcohol binding are close to water. This kinetic effect results from the accumulation of alcohol molecules along the tunnel and in the active site pocket. For benzyl alcohol as acyl acceptor, our results are in agreement with the experimentally determined activity profiles of *Ms*AcT and its engineered variants, where a jump in r_t/r_h occurs for concentrations above the experimental K_M value (13.6 mM) and its maximum value is also obtained at 100 mM.[68] This observed accumulation of alcohol in the tunnel is similar to that reported for the case of methanol with CALB.[29,69] Decreased *Ms*AcT activity at high substrate concentrations has also been reported recently for the synthesis of 2-phenethyl acetate using different acyl donors.[19] Indeed, the hydrophobicity of active site pockets and tunnels has been shown to be a good predictor of acyltransferase activity.[70,71] These effects add up to the favorable thermodynamics and kinetics of the transesterification reaction when compared to the hydrolysis reaction.[27] While the formation of the tertiary structure is essential for active site formation, in *Ms*AcT, as in other systems,[31,72−75] it is the formation of the quaternary structure through oligomerization of catalytic subunits that allows for the formation of the tunnels so that favorable interactions with substrates can take place. The results indicate that to engineer *Ms*AcT toward better acyltransferase activity, mutations with hydrophobic amino acids should be included to decrease the affinity for water (negative design) while also aiming at increasing protein−substrate interactions along the entrance to the access channel (positive design).

https://doi.org/10.1021/acscatal.4c00812
ACS Catal. 2024, 14, 10077−10088

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.4c00812.

Molecular dynamics system composition, kinetic data, and additional MD simulation results (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jürgen Pleiss − *Institute of Biochemistry and Technical Biochemistry, University of Stuttgart, 70569 Stuttgart, Germany*; orcid.org/0000-0003-1045-8202; Email: Juergen.Pleiss@itb.uni-stuttgart.de

Authors

Henrique F. Carvalho − *Institute of Biochemistry and Technical Biochemistry, University of Stuttgart, 70569 Stuttgart, Germany*

Luuk Mestrom − *Department of Biotechnology, TU Delft, 2629 HZ Delft, The Netherlands*

Ulf Hanefeld − *Department of Biotechnology, TU Delft, 2629 HZ Delft, The Netherlands*; orcid.org/0000-0002-4102-6165

Complete contact information is available at: https://pubs.acs.org/10.1021/acscatal.4c00812

Funding

The authors acknowledge financial support by the German Research Foundation (EXC 2075) and Grant ERA-IB-15-110 of the ERANET on Industrial Biotechnology.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the High Performance Computing Center Stuttgart (HLRS) for supplying computing resources.

■ ABBREVIATIONS

i-BuOH, isobutanol; BnOH, benzyl alcohol; ViAc, vinyl acetate; BuAc, isobutyl acetate; BnAc, benzyl acetate

■ REFERENCES

(1) Smith, A. J. T.; Müller, R.; Toscano, M. D.; Kast, P.; Hellinga, H. W.; Hilvert, D.; Houk, K. N. Structural Reorganization and Preorganization in Enzyme Active Sites: Comparisons of Experimental and Theoretically Ideal Active Site Geometries in the Multistep Serine Esterase Reaction Cycle. *J. Am. Chem. Soc.* **2008**, *130* (46), 15361−15373.

(2) Chaloupková, R.; Sýkorová, J.; Prokop, Z.; Jesenská, A.; Monincová, M.; Pavlová, M.; Tsuda, M.; Nagata, Y.; Damborský, J. Modification of Activity and Specificity of Haloalkane Dehalogenase from Sphingomonas Paucimobilis UT26 by Engineering of Its Entrance Tunnel*. *J. Biol. Chem.* **2003**, *278* (52), 52622−52628.

(3) Fierke, C. A.; Johnson, K. A.; Benkovic, S. J. Construction and Evaluation of the Kinetic Scheme Associated with Dihydrofolate Reductase from *Escherichia Coli*. *Biochemistry* **1987**, *26* (13), 4085−4092.

(4) Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, S. J.; Karplus, M.; Kern, D. A Hierarchy of Timescales in Protein Dynamics Is Linked to Enzyme Catalysis. *Nature* **2007**, *450* (7171), 913−916.

(5) Bhattacharya, M.; Toth, M.; Smith, C. A.; Vakulenko, S. B. Bulky "Gatekeeper" Residue Changes the Cosubstrate Specificity of Aminoglycoside 2-Phosphotransferase IIa. *Antimicrob. Agents Chemother.* **2013**, *57* (8), 3763−3766.

(6) Kokkonen, P.; Bednar, D.; Pinto, G.; Prokop, Z.; Damborsky, J. Engineering Enzyme Access Tunnels. *Biotechnol. Adv.* **2019**, *37* (6), No. 107386.

(7) Lafaquière, V.; Barbe, S.; Puech-Guenot, S.; Guieysse, D.; Cortés, J.; Monsan, P.; Siméon, T.; André, I.; Remaud-Siméon, M. Control of Lipase Enantioselectivity by Engineering the Substrate Binding Site and Access Channel. *ChemBioChem* **2009**, *10* (17), 2760−2771.

(8) Houde, A.; Kademi, A.; Leblanc, D. Lipases and Their Industrial Applications. *Appl. Biochem. Biotechnol.* **2004**, *118* (1), 155−170.

(9) Vilas Bôas, R. N.; de Castro, H. F. A Review of Synthesis of Esters with Aromatic, Emulsifying, and Lubricant Properties by Biotransformation Using Lipases. *Biotechnol. Bioeng.* **2022**, *119* (3), 725−742.

(10) Rauwerdink, A.; Kazlauskas, R. J. How the Same Core Catalytic Machinery Catalyzes 17 Different Reactions: The Serine-Histidine-Aspartate Catalytic Triad of α/β-Hydrolase Fold Enzymes. *ACS Catal.* **2015**, *5* (10), 6153−6176.

(11) Contente, M. L.; Roura Padrosa, D.; Molinari, F.; Paradisi, F. A Strategic Ser/Cys Exchange in the Catalytic Triad Unlocks an Acyltransferase-Mediated Synthesis of Thioesters and Tertiary Amides. *Nat. Catal.* **2020**, *3* (12), 1020−1026.

(12) Mestrom, L.; Claessen, J. G. R.; Hanefeld, U. Enzyme-Catalyzed Synthesis of Esters in Water. *ChemCatChem* **2019**, *11* (7), 2004−2010.

(13) Marsden, S. R.; Mestrom, L.; McMillan, D. G. G.; Hanefeld, U. Thermodynamically and Kinetically Controlled Reactions in Biocatalysis–from Concepts to Perspectives. *ChemCatChem* **2020**, *12* (2), 426−437.

(14) Jan, A.-H.; Dubreucq, E.; Drone, J.; Subileau, M. A Glimpse into the Specialization History of the Lipases/Acyltransferases Family of CpLIP2. *Biochim. Biophys. Acta, Proteins Proteomics* **2017**, *1865* (9), 1105−1113.

(15) Jan, A.-H.; Subileau, M.; Deyrieux, C.; Perrier, V.; Dubreucq, É. Elucidation of a Key Position for Acyltransfer Activity in Candida Parapsilosis Lipase/Acyltransferase (CpLIP2) and in Pseudozyma Antarctica Lipase A (CAL-A) by Rational Design. *Biochim. Biophys. Acta, Proteins Proteomics* **2016**, *1864* (2), 187−194.

(16) Cannazza, P.; Donzella, S.; Pellis, A.; Contente, M. L. Mycobacterium Smegmatis Acyltransferase: The Big New Player in Biocatalysis. *Biotechnol. Adv.* **2022**, *59*, No. 107985.

(17) Domínguez de María, P. Across the Board: Pablo Domínguez de María on the Biocatalytic Synthesis of Esters and Amides in Aqueous Media. *ChemSusChem* **2020**, *13* (21), 5611−5613.

(18) Szymańska, K.; Odrozek, K.; Zniszczoł, A.; Torrelo, G.; Resch, V.; Hanefeld, U.; Jarzębski, A. B. MsAcT in Siliceous Monolithic Microreactors Enables Quantitative Ester Synthesis in Water. *Catal. Sci. Technol.* **2016**, *6* (13), 4882−4888.

(19) Li, H.; Qin, F.; Huang, L.; Jia, W.; Zhang, M.; Li, X.; Shu, Z. Enzymatic Synthesis of 2-Phenethyl Acetate in Water Catalyzed by an Immobilized Acyltransferase From. *RSC Adv.* **2022**, *12* (4), 2310−2318.

(20) Zdun, B.; Kopińska, I.; Dranka, M.; Reiter, T.; Kroutil, W.; Borowiecki, P. Chemoenzymatic Synthesis of Optically Active Alcohols Possessing 1, 2, 3, 4-Tetrahydroquinoline Moiety Employing Lipases or Variants of the Acyltransferase from Mycobacterium Smegmatis. *Catalysts* **2022**, *12* (12), No. 1610, DOI: 10.3390/catal12121610.

(21) Wiermans, L.; Hofzumahaus, S.; Schotten, C.; Weigand, L.; Schallmey, M.; Schallmey, A.; De María, P. D. Transesterifications and Peracid-Assisted Oxidations in Aqueous Media Catalyzed by Mycobacterium Smegmatis Acyl Transferase. *ChemCatChem* **2013**, *5* (12), 3719−3724.

(22) Padrosa, D. R.; Contente, M. L. Multi-Gram Preparation of Cinnamoyl Tryptamines as Skin Whitening Agents through a Chemo-Enzymatic Flow Process. *Tetrahedron Lett.* **2021**, *86*, No. 153453.

(23) Contente, M. L.; Pinto, A.; Molinari, F.; Paradisi, F. Biocatalytic N -Acylation of Amines in Water Using an Acyltransferase from Mycobacterium Smegmatis. *Adv. Synth. Catal.* **2018**, *360* (24), 4814−4819.
(24) Christodoulou, M. S.; Contente, M. L.; Dallavalle, S.; Pinto, A. Enzymatic Amide Bond Formation: Synthesis of Aminooxo-Acids through a Mycobacterium Smegmatis Acyltransferase. *Green Chem.* **2022**, *24* (11), 4432−4436.
(25) Pinna, C.; Martino, P. A.; Meroni, G.; Sora, V. M.; Tamborini, L.; Dallavalle, S.; Contente, M. L.; Pinto, A. Biocatalyzed Synthesis of Vanillamides and Evaluation of Their Antimicrobial Activity. *J. Agric. Food Chem.* **2022**, *70* (1), 223−228.
(26) de Leeuw, N.; Torrelo, G.; Bisterfeld, C.; Resch, V.; Mestrom, L.; Straulino, E.; van der Weel, L.; Hanefeld, U. Ester Synthesis in Water: Mycobacterium Smegmatis Acyl Transferase for Kinetic Resolutions. *Adv. Synth. Catal.* **2018**, *360* (2), 242−249.
(27) Kazemi, M.; Sheng, X.; Kroutil, W.; Himo, F. Computational Study of Mycobacterium Smegmatis Acyl Transferase Reaction Mechanism and Specificity. *ACS Catal.* **2018**, *8* (11), 10698−10706.
(28) Kazemi, M.; Sheng, X.; Himo, F. Origins of Enantiopreference of Mycobacterium Smegmatis Acyl Transferase: A Computational Analysis. *Chem. - Eur. J.* **2019**, *25* (51), 11945−11954.
(29) Carvalho, H. F.; Ferrario, V.; Pleiss, J. Molecular Mechanism of Methanol Inhibition in CALB-Catalyzed Alcoholysis: Analyzing Molecular Dynamics Simulations by a Markov State Model. *J. Chem. Theory Comput.* **2021**, *17*, 6570 DOI: 10.1021/acs.jctc.1c00559.
(30) Ferrario, V.; Fischer, M.; Zhu, Y.; Pleiss, J. Modelling of Substrate Access and Substrate Binding to Cephalosporin Acylases. *Sci. Rep.* **2019**, *9* (1), No. 12402, DOI: 10.1038/s41598-019-48849-z.
(31) Mathews, I.; Soltis, M.; Saldajeno, M.; Ganshaw, G.; Sala, R.; Weyler, W.; Cervin, M. A.; Whited, G.; Bott, R. Structure of a Novel Enzyme That Catalyzes Acyl Transfer to Alcohols in Aqueous Conditions. *Biochemistry* **2007**, *46* (31), 8969−8979.
(32) Oh, C.; Ryu, B. H.; An, D. R.; Nguyen, D. D.; Yoo, W.; Kim, T.; Ngo, T. D.; Kim, H. S.; Kim, K. K.; Kim, T. D. Structural and Biochemical Characterization of an Octameric Carbohydrate Acetylesterase from Sinorhizobium Meliloti. *FEBS Lett.* **2016**, *590* (8), 1242−1252.
(33) Banno, M.; Komiyama, Y.; Cao, W.; Oku, Y.; Ueki, K.; Sumikoshi, K.; Nakamura, S.; Terada, T.; Shimizu, K. Development of a Sugar-Binding Residue Prediction System from Protein Sequences Using Support Vector Machine. *Comput. Biol. Chem.* **2017**, *66*, 36−43.
(34) Müller, H.; Godehard, S. P.; Palm, G. J.; Berndt, L.; Badenhorst, C. P. S.; Becker, A.-K.; Lammers, M.; Bornscheuer, U. T. Discovery and Design of Family VIII Carboxylesterases as Highly Efficient Acyltransferases. *Angew. Chem., Int. Ed.* **2021**, *60* (4), 2013−2017.
(35) Godehard, S. P.; Müller, H.; Badenhorst, C. P. S.; Stanetty, C.; Suster, C.; Mihovilovic, M. D.; Bornscheuer, U. T. Efficient Acylation of Sugars and Oligosaccharides in Aqueous Environment Using Engineered Acyltransferases. *ACS Catal.* **2021**, *11* (5), 2831−2836.
(36) Perišić, O.; Wriggers, W. Mechanism for the Unfolding of the TOP7 Protein in Steered Molecular Dynamics Simulations as Revealed by Mutual Information Analysis. *Front. Mol. Biosci.* **2021**, *8*, No. 696609, DOI: 10.3389/fmolb.2021.696609.
(37) Wang, K.; Zhou, R.; Tang, J.; Li, M. GraphscoreDTA: Optimized Graph Neural Network for Protein−Ligand Binding Affinity Prediction. *Bioinformatics* **2023**, *39* (6), No. btad340, DOI: 10.1093/bioinformatics/btad340.
(38) Bringas, M.; Lombardi, L. E.; Luque, F. J.; Estrin, D. A.; Capece, L. Ligand Binding Rate Constants in Heme Proteins Using Markov State Models and Molecular Dynamics Simulations. *ChemPhysChem* **2019**, *20* (19), 2451−2460.
(39) Silva, D.-A.; Bowman, G. R.; Sosa-Peinado, A.; Huang, X. A Role for Both Conformational Selection and Induced Fit in Ligand Binding by the LAO Protein. *PLOS Comput. Biol.* **2011**, *7* (5), 1−11.
(40) Gordon, S. E.; Weber, D. K.; Downton, M. T.; Wagner, J.; Perugini, M. A. Dynamic Modelling Reveals 'Hotspots' on the Pathway to Enzyme-Substrate Complex Formation. *PLoS Comput. Biol.* **2016**, *12* (3), 1−17.
(41) Lodola, A.; Sirirak, J.; Fey, N.; Rivara, S.; Mor, M.; Mulholland, A. J. Structural Fluctuations in Enzyme-Catalyzed Reactions: Determinants of Reactivity in Fatty Acid Amide Hydrolase from Multivariate Statistical Analysis of Quantum Mechanics/Molecular Mechanics Paths. *J. Chem. Theory Comput.* **2010**, *6* (9), 2948−2960.
(42) Zhang, Y.; Kua, J.; McCammon, J. A. Influence of Structural Fluctuation on Enzyme Reaction Energy Barriers in Combined Quantum Mechanical/Molecular Mechanical Studies. *J. Phys. Chem. B* **2003**, *107* (18), 4459−4463.
(43) Jorgensen, W. L.; Tirado-rives, J. The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(44) Dodda, L. S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen Web Server: An Automatic OPLS-AA Parameter Generator for Organic Ligands. *Nucleic Acids Res.* **2017**, *45* (W1), W331−W336.
(45) Dodda, L. S.; Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L. 1.14*CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations. *J. Phys. Chem. B* **2017**, *121* (15), 3864−3870.
(46) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.
(47) Olsson, M. H. M.; SØndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical p K a Predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525−537.
(48) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), No. 014101, DOI: 10.1063/1.2408420.
(49) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684.
(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.
(51) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem. A* **1987**, *91*, 6269−6271.
(52) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14* (1), 33−38.
(53) Bruice, T. C.; Lightstone, F. C. Ground State and Transition State Contributions to the Rates of Intramolecular and Enzymatic Reactions. *Acc. Chem. Res.* **1999**, *32* (2), 127−136.
(54) Gowers, R.; Linke, M.; Barnoud, J.; Reddy, T.; Melo, M.; Seyler, S.; Domański, J.; Dotson, D.; Buchoux, S.; Kenney, I.; Beckstein, O. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* **2016**, *32*, 98−105.
(55) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* **2011**, *32* (10), 2319−2327.
(56) McKinney, W.et al. In: *Data Structures for Statistical Computing in Python*. Proceedings of the 9th Python in Science Conference, 2010; pp 51−56.
(57) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-Learn: Machine Learning in {P}ython. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.
(58) Scherer, M. K.; Trendelkamp-Schroer, B.; Paul, F.; Pérez-Hernández, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J. H.; Noé, F. PyEMMA 2: A Software Package for Estimation,

Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.* **2015**, *11* (11), 5525−5542.
(59) Husic, B. E.; Pande, V. S. Markov State Models: From an Art to a Science. *J. Am. Chem. Soc.* **2018**, *140* (7), 2386−2396.
(60) Prinz, J. H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J. D.; Schtte, C.; Noé, F. Markov Models of Molecular Kinetics: Generation and Validation. *J. Chem. Phys.* **2011**, *134* (17), No. 174105, DOI: 10.1063/1.3565032.
(61) Swope, W. C.; Pitera, J. W.; Suits, F. Describing Protein Folding Kinetics by Molecular Dynamics Simulations. 1. Theory. *J. Phys. Chem. B* **2004**, *108* (21), 6571−6581.
(62) E, W.; Vanden-Eijnden, E. Towards a Theory of Transition Paths. *J. Stat. Phys.* **2006**, *123* (3), 503−523.
(63) Metzner, P.; Schütte, C.; Vanden-Eijnden, E. Transition Path Theory for Markov Jump Processes. *Multiscale Model. Simul.* **2009**, *7* (3), 1192−1219.
(64) Noé, F.; Schütte, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. Constructing the Equilibrium Ensemble of Folding Pathways from Short Off-Equilibrium Simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (45), 19011−19016.
(65) Britt, B. M. For Enzymes, Bigger Is Better. *Biophys. Chem.* **1997**, *69* (1), 63−70.
(66) Ferrario, V.; Pleiss, J. Molecular Simulations of Enzymes under Non-Natural Conditions. *Eur. Phys. J.: Spec. Top.* **2019**, *227* (14). 1631.
(67) Himo, F.; de Visser, S. P. Status Report on the Quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *Commun. Chem.* **2022**, *5* (1), No. 29, DOI: 10.1038/s42004-022-00642-2.
(68) Godehard, S. P.; Badenhorst, C. P. S.; Müller, H.; Bornscheuer, U. T. Protein Engineering for Enhanced Acyltransferase Activity, Substrate Scope, and Selectivity of the Mycobacterium Smegmatis Acyltransferase MsAcT. *ACS Catal.* **2020**, *10* (14), 7552−7562.
(69) Mangiagalli, M.; Carvalho, H.; Natalello, A.; Ferrario, V.; Pennati, M. L.; Barbiroli, A.; Lotti, M.; Pleiss, J.; Brocca, S. Diverse Effects of Aqueous Polar Co-Solvents on Candida Antarctica Lipase B. *Int. J. Biol. Macromol.* **2020**, *150*, 930−940.
(70) Müller, H.; Becker, A. K.; Palm, G. J.; Berndt, L.; Badenhorst, C. P. S.; Godehard, S. P.; Reisky, L.; Lammers, M.; Bornscheuer, U. T. Sequence-Based Prediction of Promiscuous Acyltransferase Activity in Hydrolases. *Angew. Chem., Int. Ed.* **2020**, *59* (28), 11607−11612.
(71) Jost, E.; Kazemi, M.; Mrkonjić, V.; Himo, F.; Winkler, C. K.; Kroutil, W. Variants of the Acyltransferase from Mycobacterium Smegmatis Enable Enantioselective Acyl Transfer in Water. *ACS Catal.* **2020**, *10* (18), 10500−10507.
(72) Decker, D.; Meng, M.; Gornicka, A.; Hofer, A.; Wilczynska, M.; Kleczkowski, L. A. Substrate Kinetics and Substrate Effects on the Quaternary Structure of Barley UDP-Glucose Pyrophosphorylase. *Phytochemistry* **2012**, *79*, 39−45.
(73) Führing, J.; Damerow, S.; Fedorov, R.; Schneider, J.; Münster-Kühnel, A.-K.; Gerardy-Schahn, R. Octamerization Is Essential for Enzymatic Function of Human UDP-Glucose Pyrophosphorylase. *Glycobiology* **2013**, *23* (4), 426−437.
(74) Lai, X.; Wu, J.; Chen, S.; Zhang, X.; Wang, H. Expression, Purification, and Characterization of a Functionally Active Mycobacterium Tuberculosis UDP-Glucose Pyrophosphorylase. *Protein Expression Purif.* **2008**, *61* (1), 50−56.
(75) Bosco, M. B.; Machtey, M.; Iglesias, A. A.; Aleanzi, M. UDPglucose Pyrophosphorylase from Xanthomonas Spp. Characterization of the Enzyme Kinetics, Structure and Inactivation Related to Oligomeric Dissociation. *Biochimie* **2009**, *91* (2), 204−213.

https://doi.org/10.1021/acscatal.4c00812
ACS Catal. 2024, 14, 10077−10088