Investigating the Non-Electrostatic Component of Substrate Positioning Dynamics

Yaoyukun Jiang,[1,#] Ning Ding,[1,#] Qianzhen Shao,[1,#] Sebastian L. Stull,[1] Zihao Cheng,[1] and Zhongyue J. Yang[1-5,*]

[1]*Department of Chemistry, Vanderbilt University, Nashville, Tennessee 37235, United States*

[2]*Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37235, United States*

[3]*Vanderbilt Institute of Chemical Biology, Vanderbilt University, Nashville, Tennessee 37235, United States* [4]*Data Science Institute, Vanderbilt University, Nashville, Tennessee 37235, United States* [5]*Department of Chemical and Biomolecular Engineering, Vanderbilt University, Nashville, Tennessee 37235, United States*

ABSTRACT: Substrate positioning dynamics (SPD), which orients the substrate to a reactive conformation in the active site, is critical in mediating enzyme catalysis. However, given that conformational changes often accompany variations in the enzyme interior electrostatics, it remains unknown whether SPD contains a non-electrostatic component that independently mediates catalysis, or originates primarily from perturbation of enzyme interior electrostatics. This study integrated computational and experimental approaches to investigate the non-electrostatic component of SPD using Kemp eliminase (KE) as a model enzyme. A molecular dynamics-derived descriptor, substrate positioning index (SPI), was used to quantify the impact of protein dynamics on substrate positioning. Using high throughput enzyme modeling, we selected 7 KE variants for kinetic assessment – these variants involved significantly different SPD but similar interior enzyme electrostatics. We observed a valley-shaped, two-segment piecewise linear correlation between the experimentally characterized activation free energies and SPI values. The trend is further validated using previously reported kinetic data. An optimal SPI value, corresponding to the lowest activation free energy, was observed for R154W, a surface mutation located distantly from the active site. Compared to the wild type, R154W involves favorable SPD that increases the proportion of reactive conformations for substrate deprotonation. These results indicate the presence of the non-electrostatic component of SPD, a concrete factor that mediates catalysis by tuning the population of reactive conformation.

Keywords: Substrate Positioning Dynamics, Enzyme Kinetics, Mutation Effect, Electrostatic Effect, Kemp Eliminase

[#]Y. Jiang, N. Ding, and Q. Shao contributed equally.

1. Introduction

Elucidating the catalytic origin of enzymes enhances the fundamental comprehension of chemistry, which in turn guides the development of engineering strategies to create enzyme variants for chemical synthesis,[1-3] waste degradation,[4-7] fuel production,[8-11] disease diagnosis, and treatment.[12-14] Protein dynamics have been widely reported as a factor to mediate catalysis.[15-20] Their complexity is attributed to the broad range of time scales over which correlated protein motions can occur.[15, 19, 21-28] For example, residue vibrations and collision have been proposed to facilitate transition state (TS) barrier crossing in the sub-picosecond time scale (e.g., lactate dehydrogenase, alcohol dehydrogenase, and purine nucleoside phosphorylase).[15, 21, 22] Residue and loop motions have been proposed to facilitate the positioning of substrates to form reactive conformation (or near-attack conformation[29]) in the pico- to nanosecond time scale (e.g., dihydrofolate reductase, chitinase, β-lactamase, retro-aldolase, Kemp eliminase, glycoside hydrolase, Cytochrome P450, and soybean lipoxygenase).[25-28, 30-37] Conformational change of loops and domains have been demonstrated to enable substrate binding, solvent shielding, or product releasing in the nanosecond to millisecond time scale (e.g., triosephosphate isomerase and adenylate kinase).[19, 23, 38, 39]

Substrate positioning dynamics (SPD) is perhaps the most extensively studied form of protein dynamics in enzyme catalysis, as it directly orients the substrates for an energetically favorable barrier crossing and desired selectivity.[33, 36, 40-52] Experimentally, the impact of SPD on catalysis has been investigated using mutagenesis, where a significant change of kinetics was observed upon mutating a key residue hypothesized to govern conformational dynamics (e.g.,

Gly15 in DHFR). Nonetheless, a major pitfall is that SPD is sensitively coupled to the variation of enzyme interior electric field,[16, 53] which is a well-established physical factor underlying the high catalytic efficiency of enzymes based on theoretical,[54] computational,[34, 55-57] spectroscopy,[58-60] kinetic and mutagenesis studies[53]. Upon mutation, any change in SPD likely affects the projection of the enzyme electric field along the reacting bond. As evidence, Wu *et al*. showed experimentally that the SPD mediates catalysis through tuning electrostatics in ketosteroid isomerase.[58] In this case, the correlation between the change of SPD and that of enzyme catalytic efficiency appears to be confounded by enzyme interior electrostatics. On the other hand, rate-enhancing mutants have been created through strategies that apparently optimize SPD.[49, 61] For one, Broom *et al*. observed 700-fold rate acceleration in Kemp eliminase HG4 after multiple rounds of mutagenesis that turned out to rigidify the dynamic motion of active site residues.[49]

Unless the impact of electrostatics is factored out in mutagenesis,[18, 62-64] it will remain elusive whether SPD contains a non-electrostatic component that independently mediates catalysis (*bottom*, Figure 1) or originates primarily from perturbation of enzyme interior electrostatics (*top*, Figure 1). The answer to the question will not only deepen our fundamental understanding of the catalytic origin of enzymes but also inform us whether dynamics-related descriptors should be considered as a general and independent factor for the computational engineering of biocatalysts.

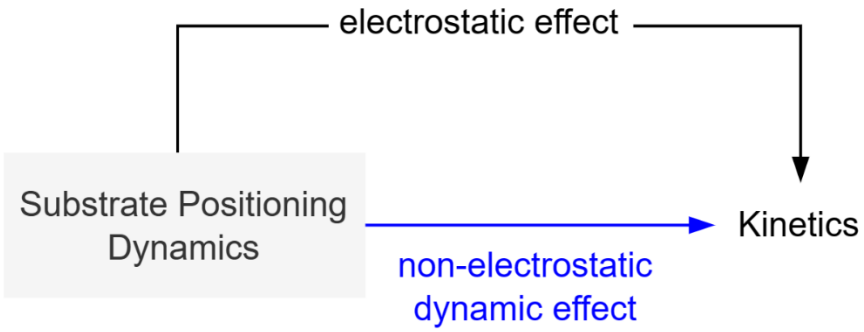


Figure 1. The correlation between substrate positioning dynamics and enzyme kinetics. (Top) The correlation is confounded by enzyme interior electrostatics. (Bottom) Substrate positioning dynamics contains a non-electrostatic component that directly mediates enzyme catalysis.

In this work, we investigated the non-electrostatic component of substrate positioning dynamics using Kemp eliminase (KE07-R7-2) as the model enzyme.[35, 65] We adopted a computer-guided mutagenesis approach that integrates high-throughput enzyme modeling with experimental kinetic assays. Through *in silico* screening enabled by EnzyHTP,[66] we identified single-point KE variants involving significantly different SPD but highly similar interior electrostatics. The turnover rate and Michaelis constant of these selected KE mutants were then characterized experimentally using kinetic assays. Based on these data, we examined the correlation between the activation free energies and substrate positioning index to investigate the non-electrostatic component of SPD.

2. Computational and Experimental Methods

Computational Methods. We employed EnzyHTP, a software developed by our lab, to perform high-throughput computational screening of Kemp eliminase (KE) mutants.[66] A job script was prepared that leveraged EnzyHTP functions to automate the process of enzyme structure preparation, random mutation generation,[67] folding stability assessment,[68, 69] molecular dynamics simulation using AMBER18,[67] quantum mechanics calculation using Gaussian16[70] and TeraChem,[71, 72] and post-analysis of results using PyMol.[73] The workflow starts from KE07-R7-2,[39] the "wild-type" structure used in this study and then creates and simulates 98 random KE variants with single amino acid substitution. The configurations of the EnzyHTP functions are detailed in Supporting Information, Text S1.

Experimental Methods and Characterization. The enzymes were expressed in *Escherichia coli* BL21(DE3) using a pET-29b(+) vector (Novagen) and purified using Ni-NTA resin (Invitrogen). Kinetic parameters were determined using 5-nitro-1,2-benzoxazole as the substrate, with concentrations ranging from 5 to 1500 μM. The reactions were initiated by adding 50 μL of the enzyme (8 μM final concentration) to 150 μL of the substrate in a 96-well plate (Corning-Costa) at 25 °C in 25 mM HEPES (pH 7.25), 100 mM NaCl, 5% glycerol and 1.5% (v/v) acetonitrile.[35] The formation of the product was monitored at 380 nm using a SpectraMax iD3 microplate reader (Molecular Device). V_{max} and K_m were calculated by nonlinear regression with the Michaelis-Menten model using GraphPad Prism software (Version 8).[74] Three biologically independent replicates were used to calculate means and standard deviations. More details can be found in the Supporting Information, Text S1.

3. Results and Discussion

3.1 The Model Enzyme for the Study: Kemp Eliminase

We used Kemp eliminase (KE), the first known *de novo*-designed enzyme, as the model enzyme in this study. KE accelerates the conversion of benzisoxazole to cyanophenol through C–H deprotonation followed by ring opening (*top*, Figure 2).[34, 35, 39, 50, 56, 65, 75-78] The substrate 5-nitro-1,2-benzoxazole undergoes deprotonation catalyzed by the carboxylate of E101, generating 2-hydroxy-5-nitrobenzonitrile through a single transition state (*top*, Figure 2). KEs involve a general acid-base mechanism, which is enabled by active site residues, including A9, I11, S48, W50, E101, Y128, H201, R202, and K222 (*bottom*, Figure 2). Specifically, Glu101 serves as the general base that deprotonates the substrate, Lys222 acts as the hydrogen bond donor to stabilize the phenoxide intermediate, and Trp50 acts as the π-stacking residue to stabilize the substrate binding and charge-

separated transition state. Four polar residues（Ser48, Tyr128, His201, and Arg202) likely stabilize the substrate binding or transition state through electrostatic or polar interactions. The nonpolar residues Ala9 and Ile11 likely favor substrate binding via dispersion interactions.

In this study, we employed KE07-R7-2 as the "wild-type" scaffold.[35, 65] KE07-R7-2 was obtained from seven rounds of directed evolution based on the parent scaffold KE07.[65] We chose this model enzyme for the following reasons. First, KE is a single-substrate enzyme whose reaction mechanism and strategies of kinetic characterization have been well-established. Second, kinetic parameters have been experimentally characterized by Bhowmick *et al*.[35] for KE07-R7-2 mutants that were identified computationally based on the correlated motion of residues. These parameters can provide references for comparison in this study. Third, we have performed benchmarks on KE07-R7-2 and its variants and determined the optimal computational strategies to ensure accuracy and reproducibility.[79]

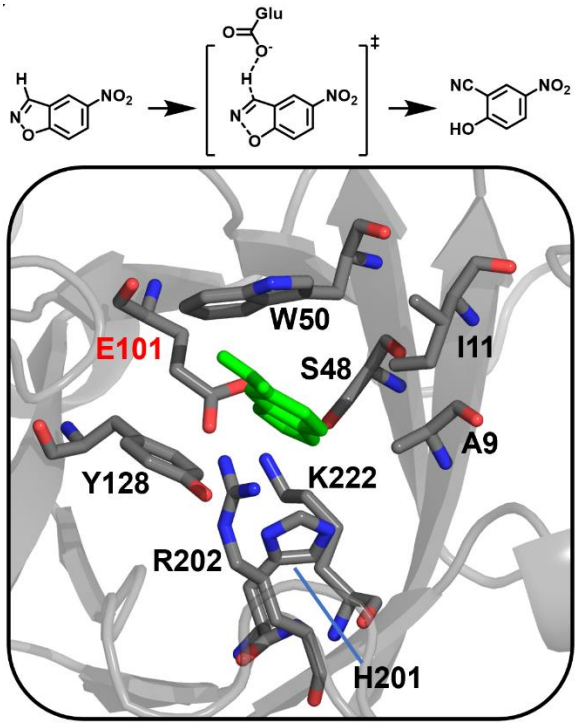


Figure 2. The model enzyme, Kemp eliminase, used in the study. (Top) A putative reaction mechanism. The carboxylic group of Glu101 deprotonates the C–H bond on the substrate, 5-nitro-1,2-benzoxazole. (Bottom) The active site structure of Kemp eliminase, KE07-R7-2. The active site residues and the substrate are shown in stick. The substrate is shown in green, and the carbon, oxygen, and nitrogen of the residues are shown in gray, red, and blue, respectively. The catalytic base is labeled in red, and others are labeled in black. The structure is derived from the crystal structure with a PDB ID of 5D38.[39]

3.2 A Quantitative Metric for Substrate Positioning Dynamics

To quantitatively describe the impact of protein dynamics on substrate positioning, we introduced a computational descriptor derived from molecular dynamics (MD) simulations. Existing descriptors for substrate position dynamics (SPD), including mechanism-based bond parameters (e.g. length of a certain H-bond) and root-mean-square deviation (RMSD) of the active site, do not directly inform the dynamic response of substrate to the conformational fluctuation of the active site residues.[44, 49] To address this, we defined a substrate positioning index (SPI), based on the ratio of solvent-accessible surface area (SASA) between the substrate and the active site residues that constitute the enzyme binding pocket (i.e., $SASA_{sub}/SASA_{pkt}$, Figure 3). For the same substrate binding to various enzyme mutants, a greater SPI value means that protein dynamics positions the substrate more tightly in the active site. The descriptor was first introduced in our prior study of mutation effects in lactonase *Sso*Pox,[46] where a piecewise linear correlation was observed between the activation free energies and SPI values for various lactonase mutants-substrate pairs. An optimal range of SPI was identified that enables the non-native substrate to react as efficiently as the native substrate.

Technically, SPI was first computed for an individual MD snapshot, with $SASA_{sub}$ and $SASA_{pkt}$ evaluated separately, and then averaged over all snapshots taken along the MD trajectories (Supporting Information, Text S1). The residues incorporated in the calculation of $SASA_{pkt}$ are illustrated in Figure 2 – the selection was guided by our previous benchmark study.[79] We also conducted a test using solvent-exclusive surface area (SESA) in lieu of SASA to compute the SPI (Figure S4). The results indicate SASA to be a better descriptor for substrate positioning effects.

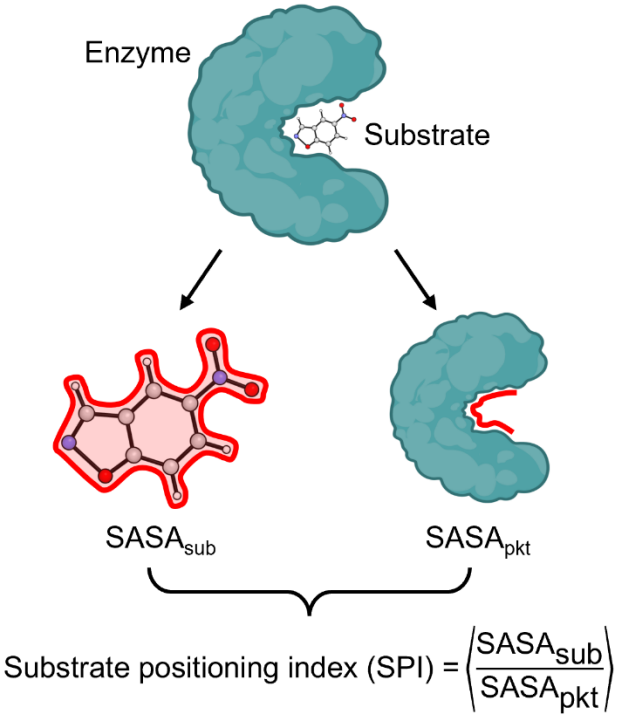


Figure 3. The definition of substrate positioning index (SPI). The solvent-accessible surface area (SASA) of the substrate ($SASA_{sub}$) and enzyme pocket ($SASA_{pkt}$) are evaluated separately by isolating these two parts from the complex. The SPI value is computed as an average over all snapshots sampled along the MD trajectories.

3.3 Identify KE Mutants Mediated by Substrate Positioning Dynamics

We designed a high-throughput computational workflow to identify single amino acid mutations with significant variation in substrate positioning dynamics but minimal change in interior enzyme electrostatics (Figure 4). Notably, substrate positioning dynamics is represented by SPI as described in Section 3.2, while interior enzyme electrostatics is represented by the interior electric field change of the breaking C–H bond, ΔEF_{C-H} (Supporting Information, Text S1). Enabled by EnzyHTP,[66] we constructed the high-throughput computational workflow (Figure 4a) that first built structural models for 98 KE variants, then conducted a folding stability test to identify mutants with a thermally stable fold at room temperature (Table S1), and eventually performed a functional test to select mutations that perturb SPI significantly but electric field minimally (Table S2).

Specifically, the 98 KE variants used in this study consist of a "wild-type" enzyme (KE07-R7-2[39, 65]), 17 mutants reported by Bhowmick *et al*.[35] and were involved in our previous benchmark study,[79] and 80 randomly-generated mutants using EnzyHTP.[66] The 80 randomly-generated mutants were then sent for the folding stability test, which was introduced to minimize unexpressed and misfolded mutants for experimental characterization. Through the stability test, 61 mutants were retained with their folding free energies of less than 10 Rosetta Energy Units (calculated using *cartesian_ddg*[68, 69]). These mutants were used for the following MD simulations, in which the average SPI and ΔEF_{C-H} were calculated using snapshots sampled from MD production trajectories. To factor out the impact of electrostatics, we selected the mutants whose averaged electric field strength was determined to be within ±2.88 MV/cm compared to the wild-type enzyme (WT, KE07-R7-2). This range corresponds to the fluctuation of electrostatic stabilization energy (ΔG_{ele}) of ±0.1 kcal/mol, in which ΔG_{ele} was estimated by the projection of electric field on the reacting dipole of C–H bond (Text S1). Within these mutants, we further selected 5 KE

variants for a kinetic assessment, including N247W, K4M, R154W, K37Q, and WT. These variants are evenly distributed across an SPI range from 1.30 to 1.70 (orange dots in Figure 4b).

After the first round of kinetic measurements, R154W was found to exhibit a 1.4-fold increase in k_{cat}/K_M compared to the WT (Table S2). As such, we further selected D14F and E185A for the second round of kinetic measurement (white dots in Figure 4b) because they have a similar SPI value to R154W. Eventually, we obtained 7 variants to investigate the impact of substrate positioning dynamics on enzyme kinetics. These mutation spots involve large distances to the substrate, from 17.4 Å for K4M to 23.0 Å for N247W. Because they involve similar electric field strength along the reacting bond, we expect that the influence of interior enzyme electrostatics is minimized.

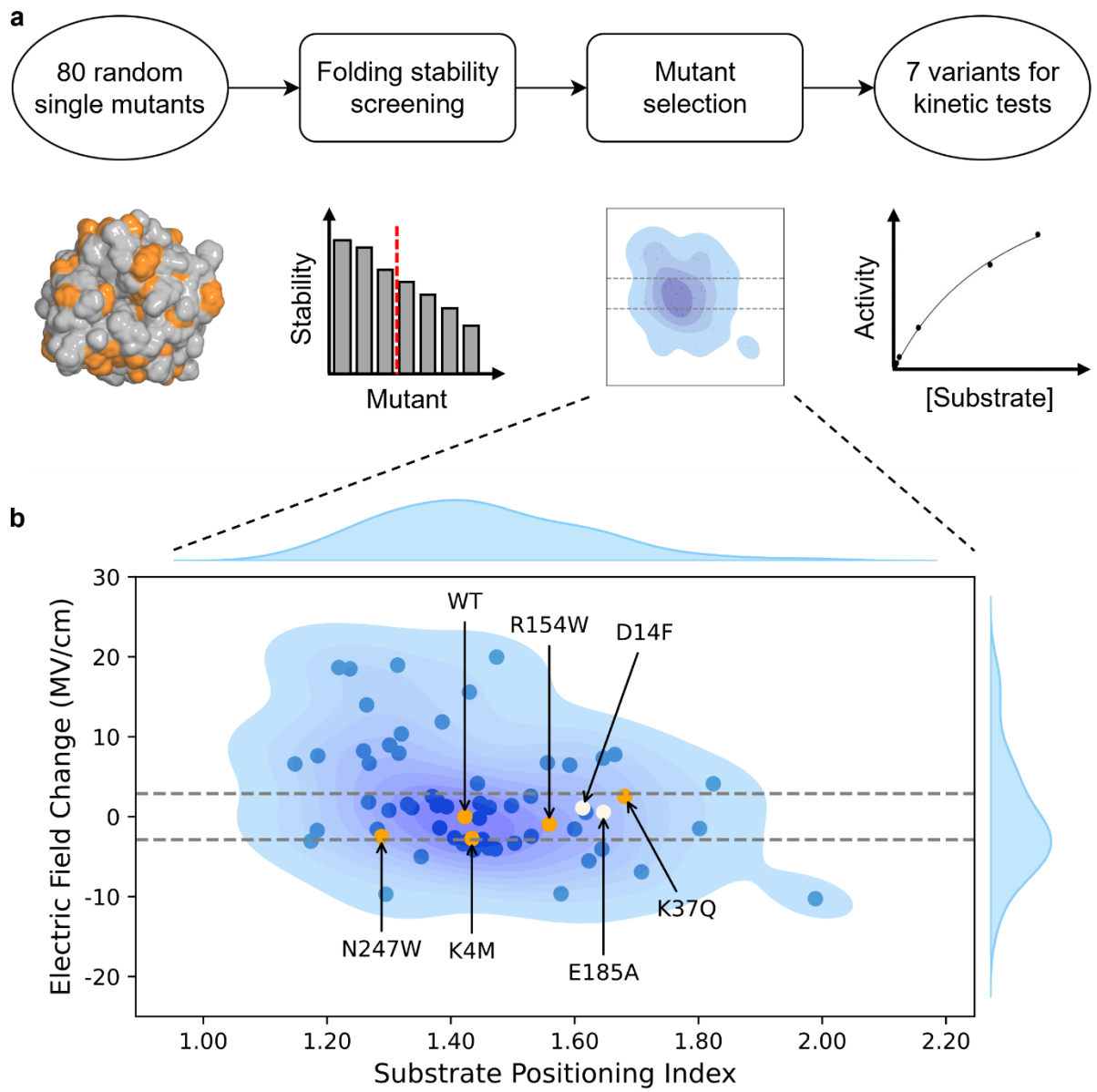


Figure 4. The computational protocol for the selection of KE variants for kinetic assessment. (a) Computational workflow to screen for mutants with a single amino acid substitution that significantly affects the substrate positioning dynamics but minimally the electric field strength on the breaking bond. The protocol contains a stability screening step and a mutational selection step. (b) Distribution of the electric field change, ΔEF_{C-H} versus the substrate positioning index, SPI. The ΔEF_{C-H} of WT is set to be the reference (0 MV/cm). The dashed lines show the EF_{C-H} range

cutoff from which mutants with a small electrostatic effect were selected. The dots include the data points for 63 KE variants, including 61 randomly-generated variants that pass the folding stability test, E185A reported by Bhowmick *et al*.[35], and the wild-type enzyme. Orange and white dots are the mutants selected for the first and second rounds of kinetic characterization, respectively. Other mutants are shown in blue. The blue shade under the dots shows the data distribution in the 2D space and the curves with blue fill on the sides show the data distribution of EF_{C-H} or SPI.

3.4 Non-electrostatic Component of Substrate Positioning Dynamics

To evaluate the non-electrostatic component of substrate-positioning dynamics, we investigated the correlation between free activation energy barrier changes, $\Delta\Delta G_{\text{eff}}^{\ddagger}$, versus the SPI (Figure 5). $\Delta\Delta G_{\text{eff}}^{\ddagger}$ is calculated using the equation $\Delta\Delta G_{\text{eff}}^{\ddagger} = -RT \ln \frac{k_{\text{cat}}^{\text{mutant}}/K_{\text{M}}^{\text{mutant}}}{k_{\text{cat}}^{\text{WT}}/K_{\text{M}}^{\text{WT}}}$, where k_{cat} and K_{M} are the experimentally measured turnover number and Michaelis constant, respectively, and the superscripts denote the variant to be the wild-type or its mutants. The temperature T is set at 298 K and R represents the gas constant. $\Delta\Delta G_{\text{eff}}^{\ddagger}$ represents the mutation effect on kinetics and directly informs whether a mutation accelerates (positive sign) or slows down (negative sign) the overall reaction.

Figure 5a shows a valley-shaped, two-segment piecewise linear correlation between the $\Delta\Delta G_{\text{eff}}^{\ddagger}$ and SPI values for the seven selected KE variants whose kinetics were experimentally characterized in this work. The first linear segment involves a gradual drop of $\Delta\Delta G_{\text{eff}}^{\ddagger}$ from 0.11 kcal/mol (N247W) to -0.21 kcal/mol (R154W), which accompanies the increase of SPI value from 1.29 (N247W) to 1.56 (R154W). The K4M and WT, the two variants in between, are similar in both SPI (~1.42) and $\Delta\Delta G_{\text{eff}}^{\ddagger}$ (~0.00 kcal/mol). The second linear segment involves a gradual

elevation of $\Delta\Delta G_{\text{eff}}^{\ddagger}$ from -0.21 kcal/mol (R154W) to 0.31 kcal/mol (K37Q), which accompanies the increase of SPI value from 1.56 (R154W) to 1.68 (K37Q). The SPI values are 1.61 in D14F and 1.65 in E185A, corresponding to a $\Delta\Delta G_{\text{eff}}^{\ddagger}$ of -0.08 and 0.05 kcal/mol, respectively. The Pearson correlation coefficients for the two linear segments are -0.98 and 0.96, respectively. R154W exhibits the most favorable $\Delta\Delta G_{\text{eff}}^{\ddagger}$ value with an SPI of 1.56. To validate the observed trend, we tested the correlation by incorporating data points reported by Bhowmick *et al.*[35], including H201A, M62A, N25S, K162A, K132M, H84Y, and L170A, which fall into the same selection cutoff of ΔEF_{C-H} (±2.88 MV/cm) to ensure small electrostatic deviation (orange dots, Figure 5b). As a result, the two-segment piecewise linear correlation still holds with a Pearson coefficient of -0.83 and 0.83, respectively. A similar trend has also been observed in our previous work for lactonase *Sso*Pox.[46] Noticeably, a recent study by Bååth *et al.*[80] showed that the turnover of poly(ethylene terephthalate) hydrolases is initially enhanced and then diminished as the binding affinity between enzyme and the substrate becomes weaker (i.e., Sabatier principle[81, 82]). The similarity in trend is likely resulted from the fact that both SPI and binding affinity are related to the influence of active site cavity on the substrate – an optimal cavity should favor both the recruitment of substrate into the active site cavity (i.e., binding affinity) and the positioning of the substrate for chemical activation (i.e., SPI).

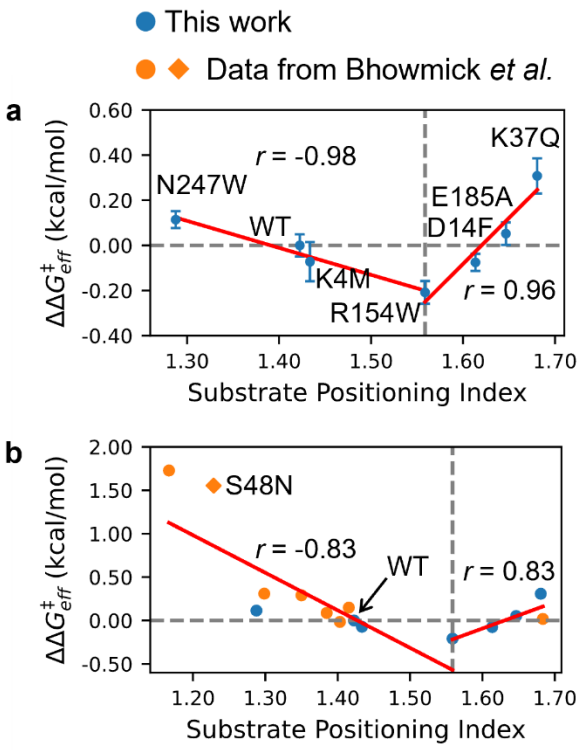


Figure 5. The correlation between change of activation free energy ($\Delta\Delta G^{\ddagger}_{\text{eff}}$) versus the substrate positioning index (SPI) for (a) KE variants that are experimentally tested in this work (blue) and (b) additional variants reported by Bhowmick *et al.*[35] (orange). For each data point shown in (a), the mean and standard error are derived from three independently repeated kinetic measurements. In each plot, the horizontal dashed line indicates the position of $\Delta\Delta G^{\ddagger}_{\text{eff}} = 0$. The vertical dashed line indicates the position of SPI = 1.56 where the beneficial mutant R154W is located. The vertical dashed line is also the boundary of the two-segment piecewise linear fitting. The fitting lines are shown in red and labeled with the corresponding Pearson correlation coefficient (*r*). The data point of R154W is included in both fitting lines. S48N is not included in the fitting because its electric field strength is outside the selection window of ±2.88 MV/cm.

The results shown in Figure 5 indicate a concrete correlation between the change of activation free energy versus the variation of SPD. The increase of SPI from 1.30 to 1.70 reflects

the process in which mutations reshape protein dynamics, gradually positioning the substrate toward a more compact active-site conformation. The presence of an optimal SPI value (i.e., ~1.56 for R154W) suggests the formation of a reactive conformation in which the substrate is positioned for favorable barrier crossing (will be investigated further in Section 3.6). This observation is consistent with previous hypotheses that SPD promotes the formation of reactive conformation.[33, 40-43] Beyond the optimal range, further increasing SPI leads to the compressed active-site cavity, which likely jeopardizes the population of reactive conformation. Since the selected mutations have been constrained within a small range of electrostatic deviation, the observed trend is largely mediated by the non-electrostatic component of SPD.

Despite being an independent factor that mediates enzyme catalysis, the change of magnitude in $\Delta\Delta G^{\ddagger}_{\text{eff}}$ caused by SPD appears to be relatively small (i.e., -0.2 to 0.4 kcal/mol). In contrast, $\Delta\Delta G^{\ddagger}_{\text{eff}}$ has been observed to change over 3-7 kcal/mol when mediated by the change of interior electric field through mutagenesis.[59] To experimentally test the severity of unfavorable substrate positioning effects, we characterized the kinetic parameters for S48N (Figure 5b). This mutant involves a favorable electrostatic environment (a ΔEF_{C-H} of 5.68 MV/cm), but a small SPI value (1.23) that substantially deviates from the predicted optimal range (1.56). The $\Delta\Delta G^{\ddagger}_{\text{eff}}$ value of S48N was measured to be +1.56 kcal/mol, which is >10-fold slower than the wild-type enzyme at room temperature. This negative impact of substrate-positioning dynamics is projected to be even worse if the electrostatic contribution is entirely factored out in S48N.

Notably, Xie and Warshel[83] recently reported that the statistical energy (E_{MaxEnt}) displays a strong anti-correlation (correlation values are -0.88 and -0.89 for $\log(k_{\text{cat}}/K_{\text{M}})$ and $\log k_{\text{cat}}$, respectively) to the corresponding activity ($\log(k_{\text{cat}}/K_{\text{M}})$ or $\log k_{\text{cat}}$) in the KE variants reported by

Bhowmick *et al.*[35] The statistical energy E_{MaxEnt} quantifies the fitness of a specific sequence in evolution, which is different from SPI that is derived from MD simulations. Although not within the scope of the current study, further investigations into the relationship between substrate positioning dynamics and the evolutionary profile of sequences may inform the synergy between the free energy landscape and fitness landscape that mediates enzyme catalysis in evolution.[61]

3.5 Substrate Positioning Dynamics Mediates the Sampling of Reactive Conformations

We hypothesized that the non-electrostatic component of SPD mediates enzyme kinetics by perturbing the population of reactive conformation. To validate this hypothesis, we calculated the mass-weighted RMSD relative to the active site of the idealized transition state (iTS),[65] i.e., $\mathrm{RMSD}_{\mathrm{iTS}}$, for each KE variant. This iTS was constructed by maximizing both the stability of the active site conformation and the affinity to the transition state while maintaining protein stability. The structural model of iTS is shown in Figure 6a. Using this structure as the reference, we calculated the mass-weighted RMSD of each MD snapshot as $\mathrm{RMSD}_{\mathrm{iTS}} = \sqrt{\frac{\sum_{i=1}^{\mathrm{N}} m_i (X_i - X_{\mathrm{iTS}})^2}{M}}$, where X represents the coordinate of an atom, i denotes the i^{th} atom in this snapshot, and iTS denotes the corresponding atom in the reference iTS structure. m_i is the mass of the i^{th} atom. N is the total number of heavy atoms and M is the total mass. We have confirmed that $\mathrm{RMSD}_{\mathrm{iTS}}$ is an effective descriptor for with a decent linear correlation (a Pearson coefficient of 0.82, Figure S5).

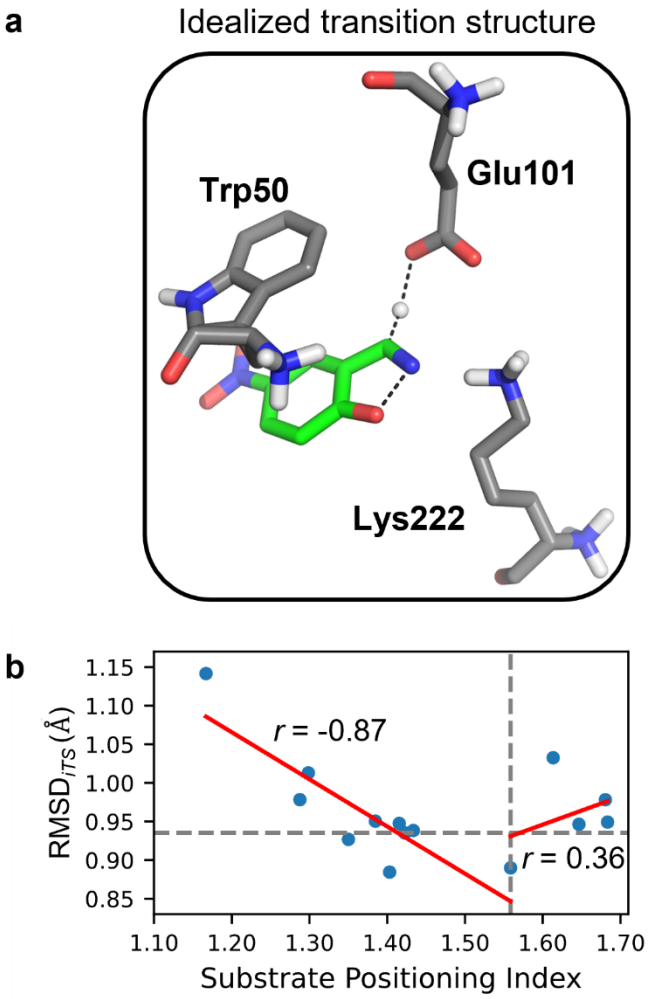


Figure 6. The impact of substrate positioning dynamics on the reaction conformation. (a) Structure of the idealized transition state optimized from QM calculations.[65] The carbon of the amino acid residues and the substrate are shown in gray and green, respectively. Nitrogen and oxygen are shown in blue and red, respectively. The polar hydrogen is shown in white. The dashed lines indicate the breaking or forming bonds. (b) Scatter plots for the root-mean-square deviation from the idealized transition state, i.e., $RMSD_{iTS}$ versus the substrate positioning index of the KE variants selected within the electric field range of ±2.88 MV/cm. The horizontal dashed line indicates the value of WT $RMSD_{iTS}$ (0.94 Å). The vertical dashed line indicates the position of SPI = 1.56 where the most beneficial mutant, i.e., R154W locates. The vertical dashed line is also

the boundary of the two-segment piecewise linear fitting. The fitted lines are shown in red and labeled with the corresponding Pearson correlation coefficient.

As shown in Figure 6b, the correlation between the average $RMSD_{iTS}$ and SPI values also follows a valley-shaped, two-segment piecewise linear trend. The $RMSD_{iTS}$ decreases from 1.14 Å (H201A) to 0.88 Å (K132M) as SPI increases from 1.17 (H201A) to 1.56 (R154W) whereas $RMSD_{iTS}$ starts to increase beyond the SPI value of 1.56 and reaches a local maximum of 1.03 Å in D14F. The Pearson coefficients for the two fitting lines are -0.87 and 0.36. This result informs more physical details behind the valley-shaped correlation pattern. During the first linear segment, the increase of SPI leads to the reduction of active-site pocket space, which enhances the sampling of reactive conformations that resemble the active-site geometry of an idealized TS. However, when the pocket further shrinks and surpasses the optimal SPI range, the active site tends to populate in a non-reactive conformation that deviates significantly from the iTS. As such, the non-electrostatic component of SPD promotes enzyme kinetics by shifting the conformation ensemble towards TS-like geometries. This may help lower the conformational entropy cost during the transition from reactant to transition state. We hypothesize that this two-segment piecewise linear correlation trend, which is similar to the "volcano plot" broadly observed in catalysis,[84] may exist universally in enzymes when the electrostatic contributions are factored out. The specific SPI value for optimal enzyme kinetics, however, is likely to be case-dependent.

3.6 The Molecular Mechanism Underlying the Impact of R154W on Substrate Positioning

To understand the molecular mechanism of how mutations mediate substrate positioning dynamics, we conducted conformational analyses on R154W (SPI: 1.56) and compared it against the results of WT (SPI: 1.42). As shown in Figure 7a, R154W is a remote mutation which locates

on the surface of the enzyme and is spatially distant from the active site (i.e., ~19.4 Å away from the active site). Compared to WT, the $SASA_{pkt}$ of R154W decreases by 18.30 $Å^2$. To identify which residue contributes the most to the change of active-site pocket, we decomposed the $SASA_{pkt}$ into contributions of individual residues (Table S5). The decomposition shows that Trp50 contributes over 84% of the overall decrease. As shown in Figure 7b, the large reduction in $SASA_{pkt}$ is driven by the shortening of spatial proximity between Trp50 and Ser144. This is supported by the downshift of distance distribution between Ser144 O_γ and Trp50 H_ε ($d_{O_\gamma - H_\varepsilon}$) upon mutation (average $d_{O_\gamma - H_\varepsilon}$ values: 4.45 Å in R154W; 6.87 Å in WT, Figure 7c). Furthermore, the formation of hydrogen bond between Ser144 O_γ and Trp50 H_ε is observed in R154W (around 2.72 Å) but is absent in WT.

Compared to WT, the close contact between Ser144 and Trp50 eliminates the accessible space of the substrate, forcing it to adopt a conformation that is parallel to the sidechain of Trp50 (Figure 7b). This conformation directs the breaking C–H bond towards the carboxylic group of Glu101. The $RMSD_{iTS}$ distribution of R154W shifts towards smaller values, generating more conformations that resemble the iTS (Figure 7c). This likely reduces the activation entropy cost, which ultimately reduces the activation barrier. Notably, a similar phenomenon has been observed in HG3, another member of the Kemp eliminase family.[61] Otten *et al.* showed that the evolved HG3 variants have more ordered side-chain orientations, leading to optimal positioning of the residues crucial to the chemical transformation and constraining of the ligand in the reactive pose.

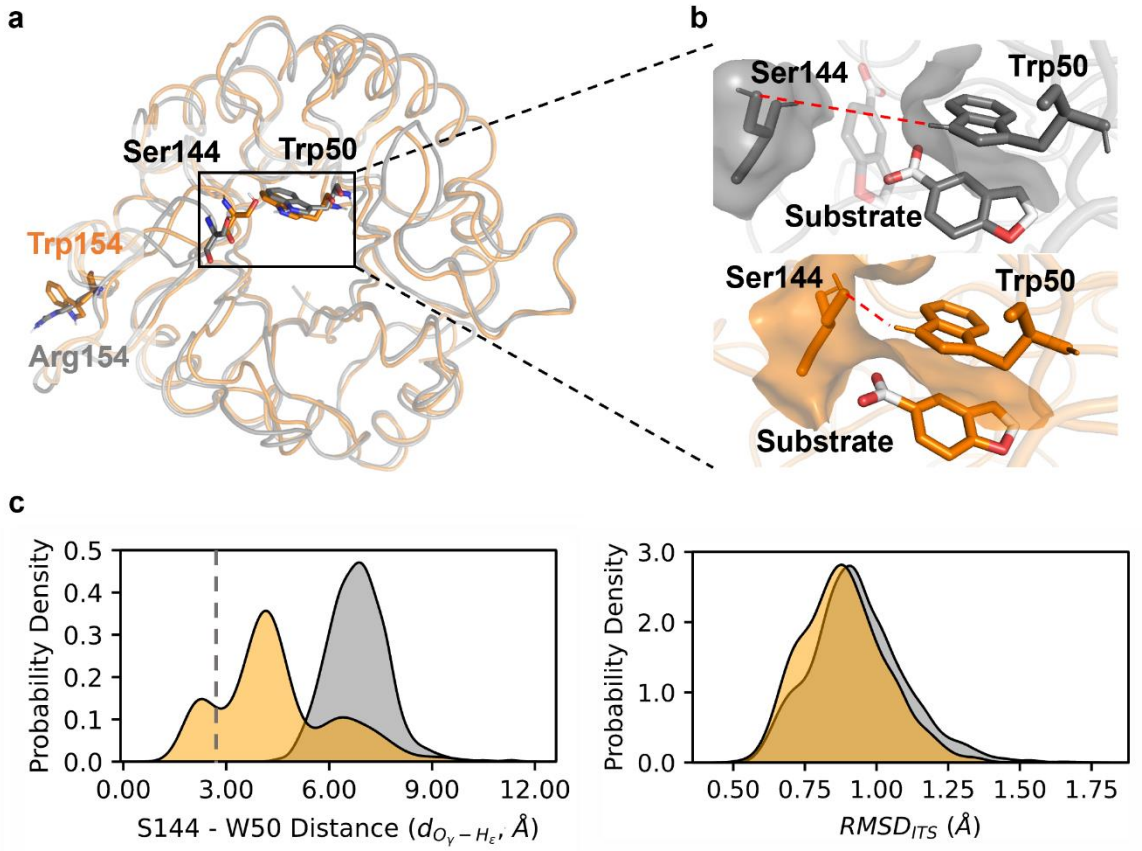


Figure 7. The molecular mechanism underlying the impact of R154W mutation on substrate positioning dynamics. (a) Structure overlay of WT (gray) and R154W (orange) with the residues at the sites of 50, 144, and 154 shown in stick. (b) Surfaces of Trp50 and Ser144 in WT and R154W. The distance between Ser144 O_γ and Trp50 H_ε, $d_{O_\gamma - H_\varepsilon}$, is shown as the red dashed line. The opaque substrate indicates its favorable position in the active site, while the transparent substrate illustrates a potentially unfavorable position in the WT. (c) Distribution of the distance between Ser144 O_γ and Trp50 H_ε, $d_{O_\gamma - H_\varepsilon}$ in the WT and R154W (*left*) and RMSD to the idealized transition station (*right*). In the $d_{O_\gamma - H_\varepsilon}$ distribution, the gray vertical dashed line represents the sum of van der Waals radii for oxygen and hydrogen (2.72 Å).

4. Conclusions

In this work, we combined computational and experimental approaches to investigate the non-electrostatic component of substrate positioning dynamics (SPD) in mediating enzyme kinetics using Kemp eliminase (KE) as a model system. To quantitatively describe SPD, we introduced a molecular dynamic-derived descriptor, substrate positioning index (SPI), which is defined using the ratio of solvent-accessible surface area between the substrate and the enzyme active site residues. We designed a high-throughput computational workflow to identify stable KE variants that involve similar interior enzyme electrostatics but distinct SPI values.

The resulting KE variants were characterized using kinetic assays. The correlation between activation free energies and SPI values demonstrates a valley-shaped, two-segment piecewise linear relationship. The trend was validated using additional KE data reported by Bhowmick *et al*. The presence of an optimal SPI value was observed in R154W, which corresponds to the lowest activation free energy among the selected mutants. We further investigated the relationship between SPI and the root-mean-square deviation of each conformational ensemble from an idealized active-site transition state model. The results show that the non-electrostatic component of SPD promotes enzyme kinetics by shifting the conformation ensemble towards TS-like geometries. To understand the molecular details behind how mutation reshapes the SPD, we performed conformational analyses on R154W and compared the results against the WT. We found that this distal mutation has a significant impact on the conformational distribution at the active site, where the mutation enables a hydrogen bonding between Ser144 and Trp50, limiting the accessible space of the substrate and positioning the substrate towards chemical activation.

These results indicate the presence of a non-electrostatic component of SPD in mediating enzyme catalysis. To promote catalysis, SPD has to position the substrate in an optimal active-site cavity to favor barrier crossing. The study implies that SPD should be considered as an independent

factor in developing strategies for pinpointing rate-enhancing mutants for biocatalysis. The study also highlights SPI as a descriptor that informs the impact of the mutation on substrate positioning dynamics. SPI can be easily calculated from molecular mechanics modeling and implemented in high-throughput computational workflows for computational enzyme engineering.

ASSOCIATED CONTENT

Supporting Information. Detailed computational and experimental methods; folding free energy change upon mutation for 80 randomly generated single mutants of KE07-R7-2; restraints applied in the MD simulation; values of molecular dynamics-derived descriptors and computed kinetic parameters for the 15 variants; primer sequences used in this study; Sanger sequencing chromatograms depicting site-directed mutagenesis in KE07-R7-2 variants; SDS-PAGE analysis of the purified KE07-R7-2 variants; scatter plots for the efficiency-enhancing free energy barrier changes upon mutation versus the substrate positioning index calculated using solvent-exclusive surface area; experimentally characterized kinetic parameters of the purified KE07-R7-2 variants; scatter plots for the correlation between the change of activation free energy versus the root-mean-square deviation from the idealized transition state; solvent-accessible surface area decomposition of KE07-R7-2 and its R154W mutant.

AUTHOR INFORMATION

Corresponding Author

*Email: zhongyue.yang@vanderbilt.edu Phone: 615-343-9849

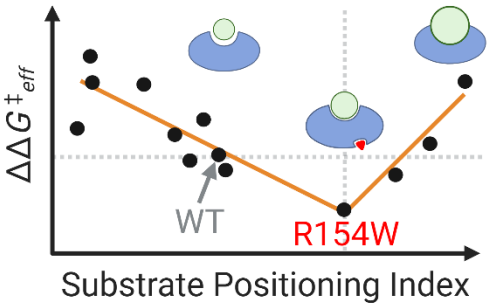
Notes

The authors declare no competing financial interest.

Acknowledgments

This research was supported by the startup grant from Vanderbilt University. Z. J. Yang, Y. Jiang, N. Ding, and Q. Shao are supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R35GM146982. Z. J. Yang thanks the sponsorship from Rosetta Commons Seed Grant Award and the Dean's Faculty Fellowship in the College of Arts and Science at Vanderbilt. S. L. Stull acknowledges financial support from the Vanderbilt Undergraduate Summer Research Program and the Department of Computer Science. This work used SDSC Dell Cluster with AMD Rome HDR IB at Expanse from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by National Science Foundation grants BIO200057.[85]

For Table of Contents Only



References

(1) Koeller, K. M.; Wong, C.-H. Enzymes for Chemical Synthesis. *Nature* **2001**, *409* (6817), 232-240. DOI: 10.1038/35051706.
(2) Strohmeier, G. A.; Pichler, H.; May, O.; Gruber-Khadjawi, M. Application of Designed Enzymes in Organic Synthesis. *Chem. Rev.* **2011**, *111* (7), 4141-4164. DOI: 10.1021/cr100386u From NLM Medline.
(3) Petchey, M. R.; Grogan, G. Enzyme-Catalysed Synthesis of Secondary and Tertiary Amides. *Adv. Synth. Catal.* **2019**, *361* (17), 3895-3914. DOI: 10.1002/adsc.201900694.
(4) Austin, H. P.; Allen, M. D.; Donohoe, B. S.; Rorrer, N. A.; Kearns, F. L.; Silveira, R. L.; Pollard, B. C.; Dominick, G.; Duman, R.; El Omari, K.; et al. Characterization and Engineering

of a Plastic-Degrading Aromatic Polyesterase. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (19), E4350-E4357. DOI: 10.1073/pnas.1718804115.
(5) Knott, B. C.; Erickson, E.; Allen, M. D.; Gado, J. E.; Graham, R.; Kearns, F. L.; Pardo, I.; Topuzlu, E.; Anderson, J. J.; Austin, H. P.; et al. Characterization and Engineering of a Two-Enzyme System for Plastics Depolymerization. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (41), 25476-25485. DOI: 10.1073/pnas.2006753117 From NLM Medline.
(6) Tiso, T.; Narancic, T.; Wei, R.; Pollet, E.; Beagan, N.; Schroder, K.; Honak, A.; Jiang, M.; Kenny, S. T.; Wierckx, N.; et al. Towards Bio-Upcycling of Polyethylene Terephthalate. *Metab. Eng.* **2021**, *66*, 167-178. DOI: 10.1016/j.ymben.2021.03.011 From NLM Medline.
(7) Ellis, L. D.; Rorrer, N. A.; Sullivan, K. P.; Otto, M.; McGeehan, J. E.; Román-Leshkov, Y.; Wierckx, N.; Beckham, G. T. Chemical and Biological Catalysis for Plastics Recycling and Upcycling. *Nat. Catal.* **2021**, *4* (7), 539-556. DOI: 10.1038/s41929-021-00648-4.
(8) Chundawat, S. P.; Beckham, G. T.; Himmel, M. E.; Dale, B. E. Deconstruction of Lignocellulosic Biomass to Fuels and Chemicals. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2*, 121-145. DOI: 10.1146/annurev-chembioeng-061010-114205.
(9) Yang, B.; Dai, Z.; Ding, S.-Y.; Wyman, C. E. Enzymatic Hydrolysis of Cellulosic Biomass. *Biofuels* **2011**, *2* (4), 421-449. DOI: 10.4155/bfs.11.116.
(10) Sweeney, M. D.; Xu, F. Biomass Converting Enzymes as Industrial Biocatalysts for Fuels and Chemicals: Recent Developments. *Catalysts* **2012**, *2* (2), 244-263.
(11) Horn, S. J.; Vaaje-Kolstad, G.; Westereng, B.; Eijsink, V. G. Novel Enzymes for the Degradation of Cellulose. *Biotechnol. Biofuels* **2012**, *5* (1), 45. DOI: 10.1186/1754-6834-5-45 From NLM PubMed-not-MEDLINE.
(12) Gordon, S. R.; Stanley, E. J.; Wolf, S.; Toland, A.; Wu, S. J.; Hadidi, D.; Mills, J. H.; Baker, D.; Pultz, I. S.; Siegel, J. B. Computational Design of an Alpha-Gliadin Peptidase. *J. Am. Chem. Soc.* **2012**, *134* (50), 20513-20520. DOI: 10.1021/ja3094795.
(13) Sun, S.; Jiang, D.; Fan, M.; Li, H.; Jin, C.; Liu, W. Selection of a Versatile Lactobacillus Plantarum for Wine Production and Identification and Preliminary Characterisation of a Novel Histamine-Degrading Enzyme. *Int. J. Food Sci. Technol.* **2020**, *55* (6), 2608-2618. DOI: 10.1111/ijfs.14514.
(14) Samadi, N.; Heiden, D.; Klems, M.; Salzmann, M.; Rohrhofer, J.; Weidmann, E.; Koidl, L.; Jensen-Jarolim, E.; Untersmayr, E. Gastric Enzyme Supplementation Inhibits Food Allergy in a BALB/c Mouse Model. *Nutrients* **2021**, *13* (3). DOI: 10.3390/nu13030738 From NLM Medline.
(15) Schwartz, S. D. Protein Dynamics and Enzymatic Catalysis. *J. Phys. Chem. B* **2023**. DOI: 10.1021/acs.jpcb.3c00477.
(16) Welborn, V. V. Structural Dynamics and Computational Design of Synthetic Enzymes. *Chem. Catalysis.* **2022**, *2* (1), 19-28. DOI: 10.1016/j.checat.2021.10.009.
(17) Petrovic, D.; Kamerlin, S. C. L. Molecular Modeling of Conformational Dynamics and Its Role in Enzyme Evolution. *Curr. Opin. Struct. Biol.* **2018**, *52*, 50-57. DOI: 10.1016/j.sbi.2018.08.004.
(18) Kamerlin, S. C. L.; Warshel, A. At the Dawn of the 21st Century: Is Dynamics the Missing Link for Understanding Enzyme Catalysis? *Proteins* **2010**, *78* (6), 1339-1375. DOI: 10.1002/prot.22654.
(19) 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-U927. DOI: 10.1038/nature06407.

(20) Radkiewicz, J. L.; Brooks, C. L. Protein Dynamics in Enzymatic Catalysis: Exploration of Dihydrofolate Reductase. *J. Am. Chem. Soc.* **2000**, *122* (2), 225-231. DOI: 10.1021/ja9913838.
(21) Saen-oon, S.; Quaytman-Machleder, S.; Schramm, V. L.; Schwartz, S. D. Atomic Detail of Chemical Transformation at the Transition State of an Enzymatic Reaction. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (43), 16543-16548. DOI: 10.1073/pnas.0808413105.
(22) Schwartz, S. D.; Schramm, V. L. Enzymatic Transition States and Dynamic Motion in Barrier Crossing. *Nat. Chem. Biol.* **2009**, *5* (8), 552-559. DOI: 10.1038/nchembio.202.
(23) Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; et al. Intrinsic Motions Along an Enzymatic Reaction Trajectory. *Nature* **2007**, *450* (7171), 838-U813. DOI: 10.1038/nature06410.
(24) Hanson, J. A.; Duderstadt, K.; Watkins, L. P.; Bhattacharyya, S.; Brokaw, J.; Chu, J. W.; Yang, H. Illuminating the Mechanistic Roles of Enzyme Conformational Dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (46), 18055-18060. DOI: 10.1073/pnas.0708600104.
(25) Bhabha, G.; Lee, J.; Ekiert, D. C.; Gam, J.; Wilson, I. A.; Dyson, H. J.; Benkovic, S. J.; Wright, P. E. A Dynamic Knockout Reveals That Conformational Fluctuations Influence the Chemical Step of Enzyme Catalysis. *Science* **2011**, *332* (6026), 234-238. DOI: 10.1126/science.1198542 (acccessed 2023/04/26).
(26) Agarwal, P. K.; Billeter, S. R.; Rajagopalan, P. T. R.; Benkovic, S. J.; Hammes-Schiffer, S. Network of Coupled Promoting Motions in Enzyme Catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (5), 2794-2799. DOI: 10.1073/pnas.052005999.
(27) Cannon, W. R.; Singleton, S. F.; Benkovic, S. J. A Perspective on Biological Catalysis. *Nat. Struct. Biol.* **1996**, *3* (10), 821-833. DOI: 10.1038/nsb1096-821.
(28) Epstein, D. M.; Benkovic, S. J.; Wright, P. E. Dynamics of the Dihydrofolate-Reductase Folate Complex - Catalytic Sites and Regions Known to Undergo Conformational Change Exhibit Diverse Dynamical Features. *Biochemistry* **1995**, *34* (35), 11037-11048. DOI: 10.1021/bi00035a009.
(29) Hur, S.; Bruice, T. C. The near Attack Conformation Approach to the Study of the Chorismate to Prephenate Reaction. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (21), 12015-12020. DOI: 10.1073/pnas.1534873100.
(30) Watney, J. B.; Agarwal, P. K.; Hammes-Schiffer, S. Effect of Mutation on Enzyme Motion in Dihydrofolate Reductase. *J. Am. Chem. Soc.* **2003**, *125* (13), 3745-3750. DOI: 10.1021/ja028487u.
(31) Hammes-Schiffer, S. Impact of Enzyme Motion on Activity. *Biochemistry* **2002**, *41* (45), 13335-13343. DOI: 10.1021/bi0267137.
(32) Ramanathan, A.; Agarwal, P. K. Evolutionarily Conserved Linkage between Enzyme Fold, Flexibility, and Catalysis. *Plos. Biol.* **2011**, *9* (11). DOI: 10.1371/journal.pbio.1001193.
(33) Norberg, A. L.; Dybvik, A. I.; Zakariassen, H.; Mormann, M.; Peter-Katalinić, J.; Eijsink, V. G. H.; Sørlie, M. Substrate Positioning in Chitinase a, a Processive Chito-Biohydrolase from Serratia Marcescens. *FEBS Lett.* **2011**, *585* (14), 2339-2344. DOI: 10.1016/j.febslet.2011.06.002.
(34) Bhowmick, A.; Sharma, S. C.; Head-Gordon, T. The Importance of the Scaffold for de Novo Enzymes: A Case Study with Kemp Eliminase. *J. Am. Chem. Soc.* **2017**, *139* (16), 5793-5800. DOI: 10.1021/jacs.6b12265.
(35) Bhowmick, A.; Sharma, S. C.; Honma, H.; Head-Gordon, T. The Role of Side Chain Entropy and Mutual Information for Improving the de Novo Design of Kemp Eliminases KE07 and KE70. *Phys. Chem. Chem. Phys.* **2016**, *18* (28), 19386-19396. DOI: 10.1039/c6cp03622h From NLM Medline.

(36) Ruscio, J. Z.; Kohn, J. E.; Ball, K. A.; Head-Gordon, T. The Influence of Protein Dynamics on the Success of Computational Enzyme Design. *J. Am. Chem. Soc.* **2009**, *131* (39), 14111-14115. DOI: 10.1021/ja905396s.
(37) Thielges, M. C.; Chung, J. K.; Fayer, M. D. Protein Dynamics in Cytochrome P450 Molecular Recognition and Substrate Specificity Using 2D IR Vibrational Echo Spectroscopy. *J. Am. Chem. Soc.* **2011**, *133* (11), 3995-4004. DOI: 10.1021/ja109168h.
(38) Liao, Q. H.; Kulkarni, Y.; Sengupta, U.; Petrovic, D.; Mulholland, A. J.; van der Kamp, M. W.; Strodel, B.; Kamerlin, S. C. L. Loop Motion in Triosephosphate Isomerase Is Not a Simple Open and Shut Case. *J. Am. Chem. Soc.* **2018**, *140* (46), 15889-15903. DOI: 10.1021/jacs.8b09378.
(39) Hong, N. S.; Petrovic, D.; Lee, R.; Gryn'ova, G.; Purg, M.; Saunders, J.; Bauer, P.; Carr, P. D.; Lin, C. Y.; Mabbitt, P. D.; et al. The Evolution of Multiple Active Site Configurations in a Designed Enzyme. *Nat. Commun.* **2018**, *9* (1), 3900. DOI: 10.1038/s41467-018-06305-y From NLM Medline.
(40) Hamre, A. G.; Jana, S.; Reppert, N. K.; Payne, C. M.; Sorlie, M. Processivity, Substrate Positioning, and Binding: The Role of Polar Residues in a Family 18 Glycoside Hydrolase. *Biochemistry* **2015**, *54* (49), 7292-7306. DOI: 10.1021/acs.biochem.5b00830.
(41) Patra, N.; Ioannidis, E. I.; Kulik, H. J. Computational Investigation of the Interplay of Substrate Positioning and Reactivity in Catechol O-Methyltransferase. *Plos One* **2016**, *11* (8). DOI: 10.1371/journal.pone.0161868.
(42) Hu, S. S.; Offenbacher, A. R.; Thompson, E. M.; Gee, C. L.; Wilcoxen, J.; Carr, C. A. M.; Prigozhin, D. M.; Yang, V.; Alber, T.; Britt, R. D.; et al. Biophysical Characterization of a Disabled Double Mutant of Soybean Lipoxygenase: The "Undoing" of Precise Substrate Positioning Relative to Metal Cofactor and an Identified Dynamical Network. *J. Am. Chem. Soc.* **2019**, *141* (4), 1555-1567. DOI: 10.1021/jacs.8b10992.
(43) Mehmood, R.; Qi, H. W.; Steeves, A. H.; Kulik, H. J. The Protein's Role in Substrate Positioning and Reactivity for Biosynthetic Enzyme Complexes: The Case of SyrB2/SyrB1. *ACS Catal.* **2019**, *9* (6), 4930-4943. DOI: 10.1021/acscatal.9b00865.
(44) Yabukarski, F.; Biel, J. T.; Pinney, M. M.; Doukov, T.; Powers, A. S.; Fraser, J. S.; Herschlag, D. Assessment of Enzyme Active Site Positioning and Tests of Catalytic Mechanisms through X-Ray-Derived Conformational Ensembles. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (52), 33204-33215. DOI: 10.1073/pnas.2011350117 From NLM Medline.
(45) Mehmood, R.; Vennelakanti, V.; Kulik, H. J. Spectroscopically Guided Simulations Reveal Distinct Strategies for Positioning Substrates to Achieve Selectivity in Nonheme Fe(II)/Alpha-Ketoglutarate-Dependent Halogenases. *ACS Catal.* **2021**, *11* (19), 12394-12408. DOI: 10.1021/acscatal.1c03169.
(46) Jiang, Y.; Yan, B.; Chen, Y.; Juarez, R. J.; Yang, Z. J. Molecular Dynamics-Derived Descriptor Informs the Impact of Mutation on the Catalytic Turnover Number in Lactonase across Substrates. *J. Phys. Chem. B* **2022**, *126* (13), 2486-2495. DOI: 10.1021/acs.jpcb.2c00142 From NLM Medline.
(47) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; Clair, J. L. S.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; et al. Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science* **2010**, *329* (5989), 309-313. DOI: 10.1126/science.1190239.

(48) Blomberg, R.; Kries, H.; Pinkas, D. M.; Mittl, P. R. E.; Grutter, M. G.; Privett, H. K.; Mayo, S. L.; Hilvert, D. Precision Is Essential for Efficient Catalysis in an Evolved Kemp Eliminase. *Nature* **2013**, *503* (7476), 418-+. DOI: 10.1038/nature12623.
(49) Broom, A.; Rakotoharisoa, R. V.; Thompson, M. C.; Zarifi, N.; Nguyen, E.; Mukhametzhanov, N.; Liu, L.; Fraser, J. S.; Chica, R. A. Ensemble-Based Enzyme Design Can Recapitulate the Effects of Laboratory Directed Evolution in Silico. *Nat. Commun.* **2020**, *11* (1). DOI: 10.1038/s41467-020-18619-x.
(50) Khersonsky, O.; Kiss, G.; Rothlisberger, D.; Dym, O.; Albeck, S.; Houk, K. N.; Baker, D.; Tawfik, D. S. Bridging the Gaps in Design Methodologies by Evolutionary Optimization of the Stability and Proficiency of Designed Kemp Eliminase KE59. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (26), 10358-10363. DOI: 10.1073/pnas.1121063109.
(51) Haataja, T.; Gado, J. E.; Nutt, A.; Anderson, N. T.; Nilsson, M.; Momeni, M. H.; Isaksson, R.; Valjamae, P.; Johansson, G.; Payne, C. M.; et al. Enzyme Kinetics by GH7 Cellobiohydrolases on Chromogenic Substrates Is Dictated by Non-Productive Binding: Insights from Crystal Structures and MD Simulation. *FEBS J.* **2023**, *290* (2), 379-399. DOI: 10.1111/febs.16602.
(52) Offenbacher, A. R.; Sharma, A.; Doan, P. E.; Klinman, J. P.; Hoffman, B. M. The Soybean Lipoxygenase-Substrate Complex: Correlation between the Properties of Tunneling-Ready States and Endor-Detected Structures of Ground States. *Biochemistry* **2020**, *59* (7), 901-910. DOI: 10.1021/acs.biochem.9b00861.
(53) Hanoian, P.; Liu, C. T.; Hammes-Schiffer, S.; Benkovic, S. Perspectives on Electrostatics and Conformational Motions in Enzyme Catalysis. *Acc. Chem. Res.* **2015**, *48* (2), 482-489. DOI: 10.1021/ar500390e.
(54) Warshel, A. Electrostatic Origin of the Catalytic Power of Enzymes and the Role of Preorganized Active Sites. *J. Biol. Chem.* **1998**, *273* (42), 27035-27038. DOI: 10.1074/jbc.273.42.27035.
(55) Welborn, V. V.; Head-Gordon, T. Computational Design of Synthetic Enzymes. *Chem. Rev.* **2019**, *119* (11), 6613-6630. DOI: 10.1021/acs.chemrev.8b00399.
(56) Vaissier, V.; Sharma, S. C.; Schaettle, K.; Zhang, T.; Head-Gordon, T. Computational Optimization of Electric Fields for Improving Catalysis of a Designed Kemp Eliminase. *ACS Catal.* **2018**, *8* (1), 219-227. DOI: 10.1021/acscatal.7b03151.
(57) Yang, Z. Y.; Liu, F.; Steeves, A. H.; Kulik, H. J. Quantum Mechanical Description of Electrostatics Provides a Unified Picture of Catalytic Action across Methyltransferases. *J. Phys. Chem. Lett.* **2019**, *10* (13), 3779-3787. DOI: 10.1021/acs.jpclett.9b01555.
(58) Wu, Y. F.; Fried, S. D.; Boxer, S. G. A Preorganized Electric Field Leads to Minimal Geometrical Reorientation in the Catalytic Reaction of Ketosteroid Isomerase. *J. Am. Chem. Soc.* **2020**, *142* (22), 9993-9998. DOI: 10.1021/jacs.0c00383.
(59) Fried, S. D.; Boxer, S. G. Electric Fields and Enzyme Catalysis. *Annu. Rev. Biochem.* **2017**, *86*, 387-415. DOI: 10.1146/annurev-biochem-061516-044432.
(60) Fried, S. D.; Boxer, S. G. Measuring Electric Fields and Noncovalent Interactions Using the Vibrational Stark Effect. *Acc. Chem. Res.* **2015**, *48* (4), 998-1006. DOI: 10.1021/ar500464j.
(61) Otten, R.; Padua, R. A. P.; Bunzel, H. A.; Nguyen, V.; Pitsawong, W.; Patterson, M.; Sui, S.; Perry, S. L.; Cohen, A. E.; Hilvert, D.; et al. How Directed Evolution Reshapes the Energy Landscape in an Enzyme to Boost Catalysis. *Science* **2020**, *370* (6523), 1442-1446. DOI: 10.1126/science.abd3623.

(62) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H. B.; Olsson, M. H. M. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **2006**, *106* (8), 3210-3235. DOI: 10.1021/cr0503106.
(63) Lameira, J.; Bora, R. P.; Chu, Z. T.; Warshel, A. Methyltransferases Do Not Work by Compression, Cratic, or Desolvation Effects, but by Electrostatic Preorganization. *Proteins* **2015**, *83* (2), 318-330. DOI: 10.1002/prot.24717.
(64) Jindal, G.; Warshel, A. Misunderstanding the Preorganization Concept Can Lead to Confusions About the Origin of Enzyme Catalysis. *Proteins* **2017**, *85* (12), 2157-2161. DOI: 10.1002/prot.25381.
(65) Rothlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; et al. Kemp Elimination Catalysts by Computational Enzyme Design. *Nature* **2008**, *453* (7192), 190-195. DOI: 10.1038/nature06879.
(66) Shao, Q.; Jiang, Y.; Yang, Z. J. EnzyHTP: A High-Throughput Computational Platform for Enzyme Modeling. *J. Chem. Inf. Model.* **2022**, *62* (3), 647-655. DOI: 10.1021/acs.jcim.1c01424.
(67) *Amber 2018*; University of California, San Francisco, 2018.
(68) Park, H.; Bradley, P.; Greisen, P.; Liu, Y.; Mulligan, V. K.; Kim, D. E.; Baker, D.; DiMaio, F. Simultaneous Optimization of Biomolecular Energy Functions on Features from Small Molecules and Macromolecules. *J. Chem. Theory Comput.* **2016**, *12* (12), 6201-6212. DOI: 10.1021/acs.jctc.6b00819.
(69) Frenz, B.; Lewis, S. M.; King, I.; DiMaio, F.; Park, H.; Song, Y. F. Prediction of Protein Mutational Free Energy: Benchmark and Sampling Improvements Increase Classification Accuracy. *Front. Bioeng. Biotech.* **2020**, *8*. DOI: 10.3389/fbioe.2020.558247.
(70) *Gaussian 16 Rev. C.01*; Wallingford, CT, 2016.
(71) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009**, *5* (10), 2619-2628. DOI: 10.1021/ct9003004 From NLM PubMed-not-MEDLINE.
(72) Titov, A. V.; Ufimtsev, I. S.; Luehr, N.; Martinez, T. J. Generating Efficient Quantum Chemistry Codes for Novel Architectures. *J. Chem. Theory Comput.* **2013**, *9* (1), 213-221. DOI: 10.1021/ct300321a From NLM PubMed-not-MEDLINE.
(73) *The Pymol Molecular Graphics System, Version 2.4*; 2015.
(74) *Graphpad Prism Version 8.3.1 for Macos*; 2019.
(75) Alexandrova, A. N.; Rothlisberger, D.; Baker, D.; Jorgensen, W. L. Catalytic Mechanism and Performance of Computationally Designed Enzymes for Kemp Elimination. *J. Am. Chem. Soc.* **2008**, *130* (47), 15907-15915, Article. DOI: 10.1021/ja804040s.
(76) Khersonsky, O.; Rothlisberger, D.; Dym, O.; Albeck, S.; Jackson, C. J.; Baker, D.; Tawfik, D. S. Evolutionary Optimization of Computationally Designed Enzymes: Kemp Eliminases of the KE07 Series. *J. Mol. Biol.* **2010**, *396* (4), 1025-1042. DOI: 10.1016/j.jmb.2009.12.031.
(77) Caselle, E. A.; Yoon, J. H.; Bhattacharya, S.; Rempillo, J. J. L.; Lengyel, Z.; D'Souza, A.; Moroz, Y. S.; Tolbert, P. L.; Volkov, A. N.; Forconi, M.; et al. Kemp Eliminases of the Alleycat Family Possess High Substrate Promiscuity. *ChemCatChem* **2019**, *11* (5), 1425-1430. DOI: 10.1002/cctc.201801994 From NLM PubMed-not-MEDLINE.
(78) Wang, P. Y.; Zhang, J.; Zhang, S. Y.; Lu, D. N.; Zhu, Y. S. Using High-Throughput Molecular Dynamics Simulation to Enhance the Computational Design of Kemp Elimination Enzymes. *J. Chem. Inf. Model.* **2023**, *63* (4), 1323-1337. DOI: 10.1021/acs.jcim.3c00002.

(79) Jiang, Y. Y. K.; Stull, S. L.; Shao, Q. Z.; Yang, Z. J. Convergence in Determining Enzyme Functional Descriptors across Kemp Eliminase Variants. *Electron. Struct.* **2022**, *4* (4). DOI: 10.1088/2516-1075/acad51.
(80) Baath, J. A.; Jensen, K.; Borch, K.; Westh, P.; Kari, J. Sabatier Principle for Rationalizing Enzymatic Hydrolysis of a Synthetic Polyester. *JACS Au* **2022**, *2* (5), 1223-1231. DOI: 10.1021/jacsau.2c00204.
(81) Schaller, K. S.; Molina, G. A.; Kari, J.; Schiano-di-Cola, C.; Sørensen, T. H.; Borch, K.; Peters, G. H. J.; Westh, P. Virtual Bioprospecting of Interfacial Enzymes: Relating Sequence and Kinetics. *ACS Catalysis* **2022**, *12* (12), 7427-7435. DOI: 10.1021/acscatal.2c02305.
(82) Kari, J.; Schaller, K.; Molina, G. A.; Borch, K.; Westh, P. The Sabatier Principle as a Tool for Discovery and Engineering of Industrial Enzymes. *Current Opinion in Biotechnology* **2022**, *78*, 102843. DOI: https://doi.org/10.1016/j.copbio.2022.102843.
(83) Xie, W. J.; Warshel, A. Natural Evolution Provides Strong Hints About Laboratory Evolution of Designer Enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119* (31). DOI: 10.1073/pnas.2207904119.
(84) Wodrich, M. D.; Sawatlon, B.; Busch, M.; Corminboeuf, C. The Genesis of Molecular Volcano Plots. *Acc. Chem. Res.* **2021**, *54* (5), 1107-1117. DOI: 10.1021/acs.accounts.0c00857.
(85) Towns, J.; Cockerill, T.; Dahan, M.; Foster, I.; Gaither, K.; Grimshaw, A.; Hazlewood, V.; Lathrop, S.; Lifka, D.; Peterson, G. D.; et al. XSEDE: Accelerating Scientific Discovery. *Comput. Sci. Eng.* **2014**, *16* (5), 62-74. DOI: 10.1109/MCSE.2014.80.