A Kinetic View of Enzyme Catalysis from Enhanced Sampling QM/MM Simulations

Dhiman Ray,* Sudip Das,* and Umberto Raucci*

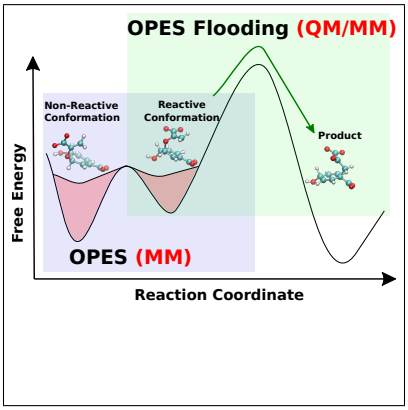
Atomistic Simulations, Italian Institute of Technology, Via Enrico Melen 83, Genova GE 16152, Italy

E-mail: dhiman.ray@iit.it; sudip.das@iit.it; umberto.raucci@iit.it

Abstract

The rate constants of enzyme-catalyzed reactions (k_{cat}) are often approximated from the barrier height of the reactive step. We introduce an enhanced sampling QM/MM approach that directly calculates the kinetics of enzymatic reactions, without introducing the transition state theory assumptions, and takes into account the dynamical equilibrium between the reactive and non-reactive conformations of the enzyme:substrate complex. Our computed k_{cat} values are in order-of-magnitude agreement with the experimental data for two representative enzymatic reactions.

TOC Graphic



Enzymes play a crucial role in the complex interplay of biochemical processes due to their high precision and selectivity in catalysis. Traditionally, enzymatic reactions are described using the Michaelis–Menten model,[1] which introduces the concept of enzyme:substrate complex (ES) as an intermediate state connecting the reactants (enzyme E and substrate S) to the products (P):

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_{\text{cat}}} E + P \tag{1}$$

The equilibrium between the reactants and the enzyme:substrate complex involves non-covalent binding, while the catalytic step entails the formation of the product with a rate constant k_{cat}. The k_{cat} is commonly measured in experiments, and it is often used as a metric to quantify the catalytic efficiency of the enzyme along with the k_{cat}/K_M ratio, where K_M ($K_M = \frac{k_{-1}+k_{\text{cat}}}{k_1}$) is the Michaelis-Menten constant.

However, at a molecular level, not all configurations of the enzyme:substrate complex can lead to product formation. The substrate necessitates a precise arrangement in the active site to initiate a reaction, and thus, only a fraction of conformations within the ES ensemble are catalytically active.[2,3] Therefore, characterizing the enzymatic activity requires not only estimating the kinetics associated with potential reactive pathways but also considering the dynamical equilibrium between reactive (RC) and non-reactive conformations (NRC) of the enzyme:substrate complex.[4] Reaching such a level of molecular resolution is extremely challenging from an experimental perspective. Nonetheless, it becomes more accessible through atomistic simulations, which can aid in differentiating between reactive and non-reactive configurations, predicting k_{cat} and understanding the enzymatic mechanism.[4–7]

So far, most studies computed the free energy landscape of the reactive step and used the effective barrier height between the enzyme:substrate complex and the product as a proxy for the k_{cat}.[7–9] The calculation of the energy barriers has been achieved either by optimizing the structures of reactants and transition states or through enhanced sampling techniques such as umbrella sampling[8,10] and metadynamics.[11,12] These latter methods use collective variables (CVs) specifically designed to encode the slow modes of the system, thereby facilitating the

occurrence of reactive events.

However, these approaches have some limitations. First, the rate constant is usually estimated using the Arrhenius equation which has notable exceptions for enzymatic reactions.[13,14] Furthermore, the applicability of transition state theory for these reactions is rather limited, since it does not take into account the recrossing probability and the diffusion constant in the free energy surface.[15–17] Second, the reactive QM/MM simulations are often initiated from the reaction-ready conformation and not from the free energy minimum corresponding to the most stable conformation of the enzyme:substrate complex.[6,12] Although this approach facilitates the sampling of the reactive events, it leads to an underestimation of the barrier height as the starting point is already near the transition state. Finally, the estimation of the barrier height from the free energy landscape, as derived using CV-based methods, is highly influenced by the choice of collective variables. As demonstrated in several studies,[18–20] projection of the free energy landscape along a suboptimal CVs can lead to an underestimation of the barrier height. Therefore, the assumption of a direct correlation between estimated barrier height and rate constant is not completely justified. Despite these shortcomings, correlating k_{cat} with the barrier height is a common practice due to the simplicity of the overall approach and the availability of enhanced sampling algorithms such as umbrella sampling and metadynamics for computing free energy landscape.

Recently, it has been demonstrated that the kinetics of molecular rare events can be calculated directly using adaptive-bias enhanced sampling algorithms without introducing any assumption of the underlying kinetic model e.g. transition state theory.[21–24] Examples of such methods include infrequent metadynamics,[23] Gaussian accelerated molecular dynamics,[25] variational flooding,[26] Gaussian mixture based enhanced sampling (GAMBES),[27] and On-the-fly probability enhanced sampling (OPES) flooding.[20] It is, therefore, appropriate that such algorithms are used to directly measure the rate constant of the catalytic step instead of the indirect estimation from the free energy landscape.

Furthermore, the importance of distinguishing the lowest free energy minimum of the enzyme:substrate complex from the reactive configuration is being highlighted in literature.[4,8,28–30] Our recent work showed the existence of an equilibrium between the deep free energy minimum of the non-reactive conformation (NRC) and the shallower minimum corresponding to the reactive conformation (RC).[6] This effect is implicitly included in the experimental k_{cat} values and should be taken into account while modeling enzymatic reactivity.

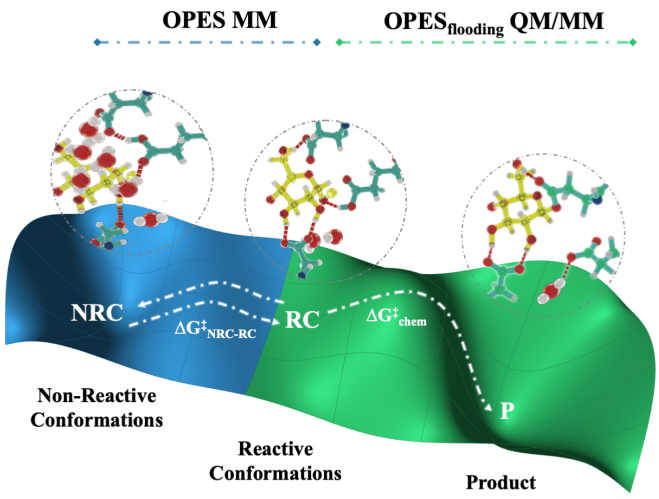


Figure 1: A schematic representation of our proposed enhanced sampling QM/MM protocol for calculation of rate constant of enzymatic reactions. The blue region on the model free energy surface depicts the transition between the reactive and non-reactive conformations of the enzyme:substrate complex. This part is modeled at MM level using OPES simulation. The reactive step (colored in green) is modeled at QM/MM level using unidirectional OPES-flooding simulations.

Here, we propose a general workflow to overcome these limitations by using adaptive bias enhanced sampling methods to explicitly sample the conformational space of the enzyme:substrate complex and to directly estimate the rate constant of the chemical transformation. We argue that the conversion from the enzyme:substrate complex to the product can be better modeled as the following two-step process (Fig. 1):

$$\mathrm{NRC} \underset{k_{\mathrm{RC}\to\mathrm{NRC}}}{\overset{k_{\mathrm{NRC}\to\mathrm{RC}}}{\rightleftharpoons}} \mathrm{RC} \xrightarrow{k_{\mathrm{RC}\to\mathbf{P}}} \mathrm{P} \tag{2}$$

An equilibrium is established between NRC and RC before the unidirectional reactive step. Considering the free energy minimum of the enzyme:substrate complex corresponds to the NRC, the rate constant for the overall reaction (NRC $\rightarrow$ P) k_{cat} is:

$$k_{\text{cat}} = \frac{k_{\text{NRC}\rightarrow\text{RC}}\, k_{\text{RC}\rightarrow\text{P}}}{k_{\text{RC}\rightarrow\text{NRC}}} = \left(\frac{p_{\text{RC}}}{p_{\text{NRC}}}\right) k_{\text{RC}\rightarrow\text{P}} = \frac{\exp(-\beta\Delta G_{\text{NRC}\rightarrow\text{RC}})}{\tau_{\text{RC}\rightarrow\text{P}}} \tag{3}$$

where $p_{\text{RC}}/p_{\text{NRC}}$ is the relative population of RC compared to NRC, and it is given by $p_{\text{RC}}/p_{\text{NRC}} = \exp(-\beta\Delta G_{\text{NRC}\rightarrow\text{RC}})$ where β is the inverse temperature. Considering the reaction from RC to P as first order we can write the reaction time as the inverse of the rate constant: $k_{\text{RC}\rightarrow\text{P}} = \tau^{-1}_{\text{RC}\rightarrow\text{P}}$. We calculate the free energy difference between the reactive and non-reactive conformations ($\Delta G_{\text{NRC}\rightarrow\text{RC}}$) at the molecular mechanics (MM) level of theory as no chemical reaction takes place in this step. As the second step (RC $\rightarrow$ P) involves the chemical transformation from the substrate to the product, we simulate it using the hybrid QM/MM method where the QM region is treated at the DFT level. We use the OPES algorithm[19] for computing free energy landscapes and the OPES flooding algorithm[20] to get a direct estimate of reaction timescale $\tau_{\text{RC}\rightarrow\text{P}}$. A schematic of this protocol is provided in Fig. 1.

The choice of the enhanced sampling algorithm is motivated by the fact that OPES leads to a quicker convergence of the free energy landscape than its predecessor, metadynamics.[11,31] In OPES, the bias potential $V(\mathbf{s})$ is expressed as a function of the reduced dimensional collective variable $\mathbf{s}$ space which encodes the slow modes of the system. $V(\mathbf{s})$ is computed from the instantaneous probability distribution $P(\mathbf{s})$ constructed on-the-fly, and is given by:

$$V(\mathbf{s}) = -\frac{1}{\beta} \ln \frac{p^{\text{tg}}(\mathbf{s})}{P(\mathbf{s})}\,. \tag{4}$$

where $p^{\text{tg}}(\mathbf{s})$ is the target distribution sampled in the presence of bias. As the target distribution one usually uses the well-tempered distribution $p^{\text{tg}}(\mathbf{s}) \propto P(\mathbf{s})^{1/\gamma}$, where $\gamma = \beta\Delta E$ and ΔE is the barrier height that needs to be overcome. The bias potential in the n-th

iteration is expressed as

$$V_n(\mathbf{s}) = (1 - 1/\gamma)\frac{1}{\beta}\ln\left(\frac{P_n(\mathbf{s})}{Z_n} + \epsilon\right), \tag{5}$$

where $P_n(\mathbf{s})$ is the estimated unbiased marginal probability distribution along $\mathbf{s}$ at step n: $P_n(s) = \sum_k^n w_k G_k(s, s_k) / \sum_k^n w_k$. Here $G_k(s, s_k)$ are Gaussian Kernels and w_k is the weight of k-th kernel computed as $w_k = \exp(\beta V_{k-1}(s_k))$. The normalization factor Z and the regularization term ϵ are introduced for numerical stability. The accuracy and efficiency of the OPES algorithm have been established for a variety of molecular rare events.[32–37]

OPES flooding (OPES_f) is a variant of the OPES algorithm specifically introduced for calculating the kinetics of molecular processes.[20] It uses an extra parameter, the excluded region $\chi_{\text{exc}}(s, s_{\text{exc}})$. A careful choice of the s_{exc} parameter ensures that no bias is deposited in the transition state making it possible to recover the unbiased kinetics. The bias is only applied in the region, $s < s_{\text{exc}}$, to reduce the effective barrier height and accelerate transitions.

We tested our approach on two prototypical enzymatic reactions: (i) the chorismate-to-prephenate conversion catalyzed by the *Bacillus subtilis* chorismate mutase enzyme[38] and (ii) the hydrolysis of maltopentaose sugar by human pancreatic α-amylase. Our choice is motivated by the fact that these two reactions have very different mechanisms. The chorismate mutase enzyme exerts its catalytic effect only via electrostatic interactions as the enzyme does not directly participate in the reaction; whereas, amylase directly takes part in the reaction by forming a covalently bound enzyme:substrate intermediate and by participating in a proton transfer event. In addition, the k_{cat} values of these two systems are about one order of magnitude apart which helps us to measure the ability of our method to distinguish their kinetics.

The conversion from chorismate to prephenate is a pericyclic rearrangement reaction responsible for the biosynthesis of several aromatic amino acids[39] (Fig. 2a). Chorismate

mutase catalyzes this reaction by stabilizing its chair-like transition state[40] via the electrostatic effect of an Arginine residue in the active site.[41] Before reaching the transition state the chorismate molecule has to undergo a conformational transition from its thermodynamically stable axial form to a relatively unstable equatorial form that better resembles the TS structure.[38,42] This step is what we refer to as the conformational transition from the non-reactive to reactive state as it is known to have a key impact on the catalytic rate constant.[42] Several studies have established that the electrostatic stabilization of the transition state is the primary determinant of the catalytic efficiency.[41,43] Nevertheless, the role of the conformational dynamics of the enzyme:substrate complex is not negligible. Using OPES simulations along a 2D CV space comprising of the distance (d_1) between the bond forming atoms and a torsion angle (θ), we obtain the free energy landscape of this conformational transition (Fig. 3). We compute the free energy cost of reaching the RC to be 19.9 kJ/mol whereas the timescale of the pericyclic rearrangement reaction, estimated from OPES_f simulations is 124 μs (Fig. 3). The predicted k_{cat}, from Eq. 3, is 2.66 s^{-1} which is in order magnitude agreement with the experimental value of 46$\pm$3 s^{-1}[38] (Table 1). It is noteworthy that the significantly better estimate of k_{cat} compared to only considering $k_{\text{RC}\to\text{P}}$ indicates the important role of substrate conformational dynamics in determining the reaction kinetics.

The other enzyme, human pancreatic α-amylase, cleaves the α(1–4) glycosidic bonds in the amylose component of starch, and hence, is a primary partner in glucose production for energy acquisition.[28,44] Furthermore, α-amylase is also considered a major target for the development of new inhibitors to treat type II diabetes.[45,46] In the catalytic decomposition of starch by α-amylase,[8,28,29,44,47] the rate-limiting first step of the catalytic reaction occurs through the nucleophilic attack by the carboxylate group of Asp197 on the anomeric C_1 carbon and the cleavage of the sugar glycosidic bond C_1–O_{gly} (Fig. 2b), leading to the formation of the covalently bound enzyme:substrate complex. This substitution reaction is coupled to a proton transfer (PT) event between the acidic Glu233 and the glycosidic oxygen O_{gly}. Previous studies[6,8,28,29,47] have found that the transient RCs of the enzyme:substrate

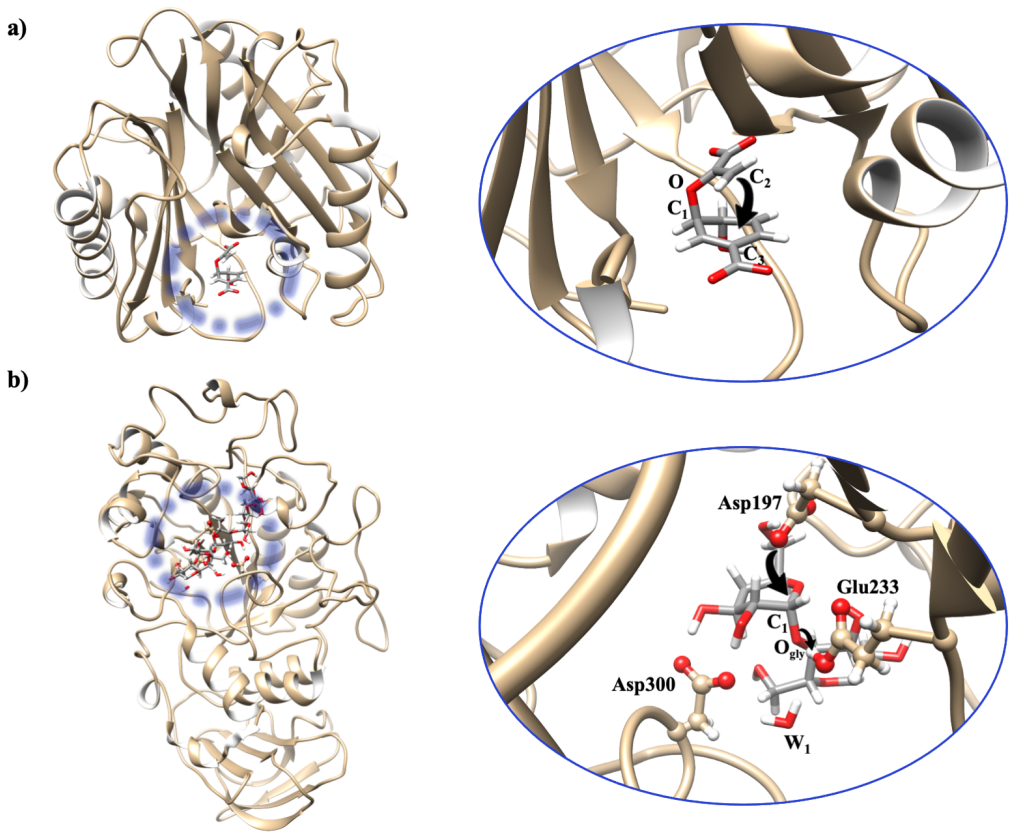


Figure 2: The two prototypical enzymatic reactions studied in this work. (a) Chorismate to prephenate conversion by the *B. subtilis* chorismate mutase and (b) the hydrolysis of maltopentaose sugar by human pancreatic α-amylase. The atoms and residues taking part in the reaction have been highlighted.

complex satisfy the reactive contacts between enzyme active site residues and the substrate. These contacts include (Fig. 2b): (1) proximity of the nucleophilic residue Asp197 to the electrophilic anomeric carbon of the substrate for a facile nucleophilic attack; (2) strong hydrogen bonds among the glycosidic oxygen (O_{gly}), acidic Glu233 residues, and a coordinated water molecule (W_1) to facilitate the proton transfer event; and finally, (3) short hydrogen bonds between Asp300 and the substrate which helps to place the electrophilic anomeric carbon in a proper position for the nucleophilic attack. On the other hand, the stable NRCs of the enzyme:substrate complex lack these important contacts.[6] In addition, our recent study[6] has found the active site water environment to play a crucial role in distinguishing these two states of the Michaelis complex. In the same study, we employed OPES

simulation with deep-targeted discriminant analysis (Deep-TDA) CVs[48] to compute the free energy landscape of the transition between RC and NRC. The free energy cost of reaching the RC has been found to be 6.5 kJ/mol (Table 1).

The rate-limiting chemical transformation of these RCs to the covalently bound enzyme:substrate complex follows two different mechanisms (as identified by our previous work[6]), depending on whether the proton transfer from the enzyme to the substrate takes place via a coordinating water molecule (W_1) (Fig. 2b). Using OPES$_{\mathrm{f}}$ simulations, we computed the reaction time corresponding to both the direct and water-mediated proton transfer. The direct PT reaction is faster compared to the water-mediated one although the difference is less than a factor of two indicating that the likelihood of these two mechanisms is almost identical (Table 1). The estimated k_{cat}, from Eq. 3, including both the mechanisms is 40.2 s^{-1} which is in order of magnitude agreement with the experimental value of 408$\pm$17 s^{-1}.[44]

We observe that only considering the reactive part of the process (i.e. approximating $k_{\mathrm{cat}} = k_{\mathrm{RC}\to\mathrm{P}}$) leads to a significant over-estimation of k_{cat} as the free energy cost to reach the reactive conformation is not taken into account. It also leads to an incorrect ordering of the k_{cat} values of the two reactions studied in this work. For example, the OPES$_{\mathrm{f}}$ simulations, initiated at the reactive conformation, produce a higher reaction timescale ($\tau_{\mathrm{RC}\to\mathrm{P}}$) for the amylase system compared to the chorismate mutase enzyme. This observation is consistent with QM/MM umbrella sampling calculations which showed the activation barrier of reactions catalyzed by chorismate mutase and α-amylase to be 11.8 kcal/mol[49] and 13.9 kcal/mol[8] respectively, using the same DFT functional used in the current work. It is, however, inconsistent with the experimental results which show a higher k_{cat} i.e. faster reaction for the amylase compared to chorismate mutase. Only when the free energy difference between the reactive and non-reactive conformations is taken into consideration, we can recover the correct ordering of the estimated k_{cat} values. This demonstrates the necessity of including the conformational dynamics of the enzyme:substrate complex in the descrip-

tion of the enzymatic reaction mechanism. For both systems, the computed k_{cat} values are one order of magnitude different from the corresponding experimental values. Considering the approximate nature of both the classical force field model and the DFT functional, this discrepancy is within the acceptable range for complex biological systems.[24]

Table 1: Summary of the thermodynamic and kinetic results for both the enzymatic reactions

System	$\Delta G_{NRC\rightarrow RC}$ (kJ/mol)	$\tau_{\text{RC}\rightarrow\text{P}}$ (μs)	$k_{\text{RC}\rightarrow\text{P}}$ (s^{-1})	k_{cat} (s^{-1})	$k_{\text{cat}}^{\text{exp}}$ (s^{-1})
Chorismate mutase	19.9	124	8064	2.66	46±3 [a]
α-Amylase	6.5	1847	541	40.2	408±17 [b]

[a] Ref. 38 [b] Ref. 44

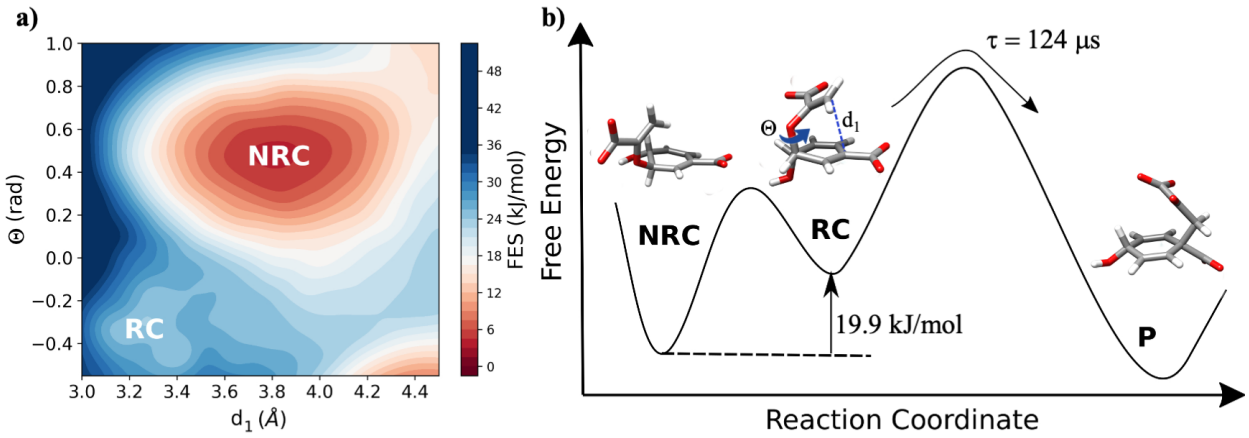


Figure 3: (a) The free energy landscape of the conformational transition between the reactive and the non-reactive states of chorismate in the chorismate mutase enzyme:substrate complex. (b) Depiction of the conformational transition and the reactive step in a single qualitative free energy landscape. Representative structures of the different states are shown along with the two CVs (Θ and d_1) biased for calculating the free energy landscape in subfigure (a). The free energy surface for the α-amylase system has been reported in Ref. 6.

In conclusion, we demonstrated the application of adaptive bias enhanced sampling methods to directly estimate the rate constants of enzyme-catalyzed reactions. Enhanced sampling methods like infrequent metadynamics, variational flooding, and OPES flooding have so far been successfully applied to a variety of chemical, material, and biological systems for calculating kinetics.[24] However, to the best of our knowledge, this work first demonstrates a

successful application of such a method to compute the timescales of an enzymatic reaction at the QM/MM level. Our approach takes into consideration the equilibrium populations of the reactive and non-reactive conformations of the enzyme:substrate complex using enhanced sampling simulations based on molecular mechanics force fields, leading to a significant increase in computational efficiency. We also show that the relative free energy of the reactive conformation can be included via a simple kinetic model, leading to an improved estimate of the k_{cat} values.

One key advantage of our protocol is that we are not required to adhere to the transition state theory approximation to estimate kinetics. So we can directly compare computational k_{cat} values with experimental ones, without calculating the barrier height, a quantity that can be severely impacted by the choice of collective variables. Furthermore, the OPES flooding algorithm only needs to sample one-way transitions from reactant to product. As a result, our approach is computationally more efficient than traditional enhanced sampling methods that try to calculate the free energy landscape of the reactive step, which ideally requires one to sample multiple back-and-forth transitions. Therefore, when integrated with either massively parallel[50,51] or GPU-accelerated[52,53] electronic structure codes, it can be highly efficient in studying systems requiring large QM regions, such as ion channels.[54,55] We could demonstrate the successful application of this protocol for two enzymatic reactions with completely different reaction mechanisms, which is an indication of the transferability of our approach to different enzymes, substrates, and reaction types. Our study also paves the way to modifying enzyme efficiency not only by tuning the transition state stabilization but also by perturbing the conformational dynamics of the enzyme:substrate complex.

Acknowledgement

The authors thank Rui Neves for sharing the input files for the α-mylase system. The authors declare no competing financial interest.

Data Availability Statement

All calculations were performed using open source software packages e.g GROMACS 2022,[56] CP2K v9.1,[57] and PLUMED v2.9.[58] The input files for all simulations reported in this work are available in the following Github repository: `https://github.com/dhimanray/Enzyme_Kinetics_OPES_flooding.git`

Supporting Information Available

Computational details and the statistical analysis of the kinetic data using Kolmogorov-Smirnov test are provided in the supporting information.

References

(1) Michaelis, L.; Menten, M. L., et al. Die kinetik der invertinwirkung. *Biochem. z* **1913**, *49*, 352.

(2) Agarwal, P. K. A biophysical perspective on enzyme catalysis. *Biochemistry* **2018**, *58*, 438–449.

(3) Ramanathan, A.; Savol, A.; Burger, V.; Chennubhotla, C. S.; Agarwal, P. K. Protein conformational populations and functionally relevant substates. *Accounts of chemical research* **2014**, *47*, 149–156.

(4) Hur, S.; Bruice, T. C. The near attack conformation approach to the study of the chorismate to prephenate reaction. *Proceedings of the National Academy of Sciences* **2003**, *100*, 12015–12020.

(5) Agarwal, P. K. Role of protein dynamics in reaction rate enhancement by enzymes. *Journal of the American Chemical Society* **2005**, *127*, 15248–15256.

(6) Das, S.; Raucci, U.; Neves, R. P.; Ramos, M. J.; Parrinello, M. How and When Does an Enzyme React? Unraveling α-Amylase Catalytic Activity with Enhanced Sampling Techniques. *ACS Catalysis* **2023**, *13*, 8092–8098.

(7) Ojeda-May, P.; Mushtaq, A. U.; Rogne, P.; Verma, A.; Ovchinnikov, V.; Grundstrom, C.; Dulko-Smith, B.; Sauer, U. H.; Wolf-Watz, M.; Nam, K. Dynamic connection between enzymatic catalysis and collective protein motions. *Biochemistry* **2021**, *60*, 2246–2258.

(8) Neves, R. P.; Fernandes, P. A.; Ramos, M. J. Role of Enzyme and Active Site Conformational Dynamics in the Catalysis by α-Amylase Explored with QM/MM Molecular Dynamics. *Journal of Chemical Information and Modeling* **2022**, *62*, 3638–3650.

(9) Villa, J.; Warshel, A. Energetics and dynamics of enzymatic reactions. 2001.

(10) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics* **1977**, *23*, 187–199.

(11) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proceedings of the national academy of sciences* **2002**, *99*, 12562–12566.

(12) Raich, L.; Nin-Hill, A.; Ardèvol, A.; Rovira, C. Enzymatic cleavage of glycosidic bonds: strategies on how to set up and control a QM/MM metadynamics simulation. *Methods in enzymology* **2016**, *577*, 159–183.

(13) Truhlar, D. G.; Kohen, A. Convex Arrhenius plots and their interpretation. *Proceedings of the National Academy of Sciences* **2001**, *98*, 848–851.

(14) Roy, S.; Schopf, P.; Warshel, A. Origin of the non-Arrhenius behavior of the rates of enzymatic reactions. *The Journal of Physical Chemistry B* **2017**, *121*, 6520–6526.

(15) Truhlar, D. G. Transition state theory for enzyme kinetics. *Archives of biochemistry and biophysics* **2015**, *582*, 10–17.

(16) Doron, D.; Kohen, A.; Nam, K.; Major, D. T. How accurate are transition states from simulations of enzymatic reactions? *Journal of Chemical Theory and Computation* **2014**, *10*, 1863–1871.

(17) Hynes, J. T.; Laage, D.; Tuñón, I.; Moliner, V. *Simulating Enzyme Reactivity: Computational Methods in Enzyme Catalysis*; The Royal Society of Chemistry, 2016; pp 54–88.

(18) Bal, K. M.; Fukuhara, S.; Shibuta, Y.; Neyts, E. C. Free energy barriers from biased molecular dynamics simulations. *The Journal of Chemical Physics* **2020**, *153*.

(19) Invernizzi, M.; Parrinello, M. Rethinking metadynamics: From bias potentials to probability distributions. *The journal of physical chemistry letters* **2020**, *11*, 2731–2736.

(20) Ray, D.; Ansari, N.; Rizzi, V.; Invernizzi, M.; Parrinello, M. Rare Event Kinetics from Adaptive Bias Enhanced Sampling. *Journal of Chemical Theory and Computation* **2022**, *18*, 6500–6509.

(21) Grubmüller, H. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Physical Review E* **1995**, *52*, 2893.

(22) Voter, A. F. Hyperdynamics: Accelerated molecular dynamics of infrequent events. *Physical Review Letters* **1997**, *78*, 3908.

(23) Tiwary, P.; Parrinello, M. From metadynamics to dynamics. *Physical review letters* **2013**, *111*, 230602.

(24) Ray, D.; Parrinello, M. Kinetics from Metadynamics: Principles, Applications, and Outlook. *Journal of Chemical Theory and Computation* **2023**, *19*, 5649–5670.

(25) Miao, Y.; Feher, V. A.; McCammon, J. A. Gaussian accelerated molecular dynamics: Unconstrained enhanced sampling and free energy calculation. *Journal of chemical theory and computation* **2015**, *11*, 3584–3595.

(26) McCarty, J.; Valsson, O.; Tiwary, P.; Parrinello, M. Variationally optimized free-energy flooding for rate calculation. *Physical review letters* **2015**, *115*, 070601.

(27) Debnath, J.; Parrinello, M. Gaussian mixture-based enhanced sampling for statics and dynamics. *The Journal of Physical Chemistry Letters* **2020**, *11*, 5076–5080.

(28) Pinto, G. P.; Bras, N. F.; Perez, M. A.; Fernandes, P. A.; Russo, N.; Ramos, M. J.; Toscano, M. Establishing the catalytic mechanism of human pancreatic α-amylase with QM/MM methods. *Journal of chemical theory and computation* **2015**, *11*, 2508–2516.

(29) Neves, R. P.; Cunha, A. V.; Fernandes, P. A.; Ramos, M. J. Towards the Accurate Thermodynamic Characterization of Enzyme Reaction Mechanisms. *ChemPhysChem* **2022**, *23*, e202200159.

(30) Maria-Solano, M. A.; Serrano-Hervás, E.; Romero-Rivera, A.; Iglesias-Fernández, J.; Osuna, S. Role of conformational dynamics in the evolution of novel enzyme function. *Chemical Communications* **2018**, *54*, 6622–6634.

(31) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Physical review letters* **2008**, *100*, 020603.

(32) Ansari, N.; Rizzi, V.; Parrinello, M. Water regulates the residence time of Benzamidine in Trypsin. *Nature Communications 2022 13:1* **2022**, *13*, 1–9.

(33) Rizzi, V.; Bonati, L.; Ansari, N.; Parrinello, M. The role of water in host-guest interaction. *Nature Communications* **2021**, *12*, 1–7.

(34) Raucci, U.; Rizzi, V.; Parrinello, M. Discover, Sample, and Refine: Exploring Chemistry

with Enhanced Sampling Techniques. *The Journal of Physical Chemistry Letters* **2022**, *13*, 1424–1430.

(35) Karmakar, T.; Invernizzi, M.; Rizzi, V.; Parrinello, M. Collective variables for the study of crystallisation. *Molecular Physics* **2021**, *119*, e1893848.

(36) Mambretti, F.; Raucci, U.; Yang, M.; Parrinello, M. How Does Structural Disorder Impact Heterogeneous Catalysts? The Case of Ammonia Decomposition on Non-stoichiometric Lithium Imide. *ACS Catalysis* **2023**, *14*, 1252–1256.

(37) Yang, M.; Raucci, U.; Parrinello, M. Reactant-induced dynamics of lithium imide surfaces during the ammonia decomposition process. *Nature Catalysis* **2023**, *6*, 829–836.

(38) Kast, P.; Asif-Ullah, M.; Hilvert, D. Is chorismate mutase a prototypic entropy trap?-Activation parameters for the Bacillus subtilis enzyme. *Tetrahedron letters* **1996**, *37*, 2691–2694.

(39) Haslam, E. Shikimic acid: metabolism and metabolites. *John Wiley and Sons, New York* **1993**,

(40) Sogo, S. G.; Widlanski, T. S.; Hoare, J. H.; Grimshaw, C. E.; Berchtold, G. A.; Knowles, J. R. Stereochemistry of the rearrangement of chorismate to prephenate: chorismate mutase involves a chair transition state. *Journal of the American Chemical Society* **1984**, *106*, 2701–2703.

(41) Claeyssens, F.; Ranaghan, K. E.; Lawan, N.; Macrae, S. J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. Analysis of chorismate mutase catalysis by QM/MM modelling of enzyme-catalysed and uncatalysed reactions. *Organic & biomolecular chemistry* **2011**, *9*, 1578–1590.

(42) Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M. Substrate conformational transitions in

the active site of chorismate mutase: Their role in the catalytic mechanism. *Proceedings of the National Academy of Sciences* **2001**, *98*, 9032–9037.

(43) Burschowsky, D.; van Eerde, A.; Ökvist, M.; Kienhöfer, A.; Kast, P.; Hilvert, D.; Krengel, U. Electrostatic transition state stabilization rather than reactant destabilization provides the chemical basis for efficient chorismate mutase catalysis. *Proceedings of the National Academy of Sciences* **2014**, *111*, 17516–17521.

(44) Brayer, G. D.; Sidhu, G.; Maurus, R.; Rydberg, E. H.; Braun, C.; Wang, Y.; Nguyen, N. T.; Overall, C. M.; Withers, S. G. Subsite mapping of the human pancreatic α-amylase active site through structural, kinetic, and mutagenesis techniques. *Biochemistry* **2000**, *39*, 4778–4791.

(45) Jayaraj, S.; Suresh, S.; Kadeppagari, R.-K. Amylase inhibitors and their biomedical applications. *Starch-Stärke* **2013**, *65*, 535–542.

(46) Oliveira, H.; Fernandes, A.; F. Brás, N.; Mateus, N.; de Freitas, V.; Fernandes, I. Anthocyanins as antidiabetic agents—in vitro and in silico approaches of preventive and therapeutic effects. *Molecules* **2020**, *25*, 3813.

(47) Santos-Martins, D.; Calixto, A. R.; Fernandes, P. A.; Ramos, M. J. A buried water molecule influences reactivity in α-amylase on a subnanosecond time scale. *ACS Catalysis* **2018**, *8*, 4055–4063.

(48) Trizio, E.; Parrinello, M. From enhanced sampling to reaction profiles. *The Journal of Physical Chemistry Letters* **2021**, *12*, 8621–8626.

(49) Ranaghan, K. E.; Shchepanovska, D.; Bennie, S. J.; Lawan, N.; Macrae, S. J.; Zurek, J.; Manby, F. R.; Mulholland, A. J. Projector-based embedding eliminates density functional dependence for QM/MM calculations of reactions in enzymes and solution. *Journal of Chemical Information and Modeling* **2019**, *59*, 2063–2078.

(50) Bolnykh, V.; Olsen, J. M. H.; Meloni, S.; Bircher, M. P.; Ippoliti, E.; Carloni, P.; Rothlisberger, U. Extreme scalability of DFT-based QM/MM MD simulations using MiMiC. *Journal of chemical theory and computation* **2019**, *15*, 5601–5613.

(51) Olsen, J. M. H.; Bolnykh, V.; Meloni, S.; Ippoliti, E.; Bircher, M. P.; Carloni, P.; Rothlisberger, U. MiMiC: a novel framework for multiscale modeling in computational chemistry. *Journal of chemical theory and computation* **2019**, *15*, 3810–3823.

(52) Cruzeiro, V. W. D.; Wang, Y.; Pieri, E.; Hohenstein, E. G.; Martínez, T. J. TeraChem protocol buffers (TCPB): Accelerating QM and QM/MM simulations with a client–server model. *The Journal of Chemical Physics* **2023**, *158*.

(53) Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Isborn, C. M.; Kokkila-Schumacher, S. I.; Li, X.; Liu, F.; Luehr, N.; Snyder Jr, J. W. et al. TeraChem: A graphical processing unit-accelerated electronic structure package for large-scale ab initio molecular dynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2021**, *11*, e1494.

(54) Chiariello, M. G.; Alfonso-Prieto, M.; Ippoliti, E.; Fahlke, C.; Carloni, P. Mechanisms Underlying Proton Release in CLC-type F–/H+ Antiporters. *The journal of physical chemistry letters* **2021**, *12*, 4415–4420.

(55) Chiariello, M. G.; Bolnykh, V.; Ippoliti, E.; Meloni, S.; Olsen, J. M. H.; Beck, T.; Rothlisberger, U.; Fahlke, C.; Carloni, P. Molecular basis of CLC antiporter inhibition by fluoride. *Journal of the American Chemical Society* **2020**, *142*, 7254–7258.

(56) 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*, 19–25.

(57) Kühne, T. D.; Iannuzzi, M.; Del Ben, M.; Rybkin, V. V.; Seewald, P.; Stein, F.; Laino, T.; Khaliullin, R. Z.; Schütt, O.; Schiffmann, F. et al. CP2K: An electronic

structure and molecular dynamics software package-Quickstep: Efficient and accurate electronic structure calculations. *The Journal of Chemical Physics* **2020**, *152*.

(58) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Computer physics communications* **2014**, *185*, 604–613.