How and When Does an Enzyme React? Unraveling α-Amylase Catalytic Activity with Enhanced Sampling Techniques

Sudip Das,[1] Umberto Raucci,[1] Rui P. P. Neves,[2] Maria J. Ramos,[2,*] Michele Parrinello[1,*]

[1] *Italian Institute of Technology, Genova GE, Italy*
[2] *LAQV@REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Porto, Portugal*

Abstract

Enzymatic catalysis is a complex process that can involve multiple conformations of the enzyme:substrate complex and several competitive reaction pathways, resulting in a multi-dimensional free energy landscape. The study of enzymatic activity often requires deep knowledge of the system to establish the catalytic mechanism and identify the possible reactive conformations of the complex. Here, we present an enhanced sampling and machine learning-based approach to explore the catalytic reaction space and characterize the transformation from reactive to non-reactive conformations with minimal *a priori* knowledge of the system. We applied this approach to study the rate-determining step of the glycolysis reaction of maltopentose catalyzed by human pancreatic α-amylase, an important enzyme in glucose production as well as a major drug target for the treatment of type-II diabetes. We unravel the complexity of the enzymatic reaction, reveal three binding modes of the substrate in the catalytic pocket, and highlight the role of water in the catalytic process and in the stepwise conversion of reaction-ready to non-reactive conformations. Overall, these insights offer atomistic details on the catalytic mechanism and dynamics of the active site, allowing to shed light on two fundamental questions in enzymatic catalysis, that is how and when does an enzyme react?

Introduction

Enzymes are highly efficient molecular systems that catalyze complex chemical processes in living cells. Enzymatic catalysis relies on the concerted action of amino acid residues, substrate, and sometimes water molecules arranged in or close to the active site. They operate cooperatively, promoting the binding of substrates and their chemical transformation, ultimately resulting in the release of the products. Since several species and many degrees of freedom are involved in this process, the conformational space of the active site can be extremely complex[1] (**Fig. 1a**).

In most studies, attention has been focused on the search for the transition state (TS) and its activation energy ($\Delta G^{\ddagger}_{\text{chem}}$). This is of major importance because the TS is the eye of the needle through which the system has to pass for the reaction to occur. However, the catalytic activity is often the result of a complex process that requires first the enzyme:substrate adduct to be arranged in a reactive conformation. These reactive conformations are referred to as reaction-ready or near-attack conformation (NAC), and the associated energy will be called ($\Delta G^{\ddagger}_{\text{NAC}}$)[2, 3] (**Fig. 1a**). In many cases, the probability of the system visiting such conformations is very low. For instance, in the isomerization of chorismate to prephenate, this probability can be as low as 10^{-4} %[2]. Additionally, the arrangement of water molecules in the active site can influence the reactivity. In the case of human pancreatic α-amylase, which we shall report below, the activation energy can change by as much as 20 kcal/mol depending on the orientation of a single water molecule[4]. Additional complications arise because the substrate can have several binding modes and multiple reaction pathways[5-9].

Evaluating $\Delta G^{\ddagger}_{\text{chem}}$, $\Delta G^{\ddagger}_{\text{NAC}}$ and finding the possible reactive pathways could be extremely challenging, and molecular dynamics simulations could be a valuable tool in the study of these complex enzymatic processes. Nevertheless, these reactions occur on a timescale that far exceeds present day computational capabilities. State-of-the-art enhanced sampling techniques allow overcoming these limitations, and designing a workflow to study enzymatic processes in a statistical mechanics compliant way[10-19].

Two main tools will allow us to i) blindly explore reaction mechanisms occurring in the catalytic pocket, and ii) compute free energy barriers ($\Delta G^{\ddagger}_{\text{NAC}}$) between reactive and non-reactive conformations. The first is the on-the-fly probability enhanced sampling (OPES) method[20, 21]. OPES, like several enhanced sampling approaches, is based on the use of collective variables

(CVs) which are meant to describe the slow modes of the system. The determination of collective variables appropriate for use in the present context is the second set of tools. To promote enzymatic reactions in a blind way, we used a CV derived from spectral graph theory[22, 23]. In particular, we represent a molecule as a graph whose vertices and edges are its atoms and chemical bonds, respectively, and the CV is the maximum eigenvalue (λ^{max}) of the symmetric adjacency matrix associated to the graph[22, 23]. When used in the OPES context[24], this CV allows the discovery of new chemical pathways[25]. To promote the sampling of NAC and compute the associated energetics, we used machine learning-based CVs also trained to account for the water arrangement inside the enzymatic cavity.

We exploited these tools to perform several reaction discovery simulations at the QM/MM level, starting from different initial conformations (ICs) of the enzyme:substrate complex (**Fig. 1b**). The MD engine samples different reaction pathways, and each IC can be classified as reactive or non-reactive depending on the bias required to promote the reaction. The ICs are characterized using two classes of descriptors: the reactive contacts of the substrate with the enzyme and the functions that describe the water environment inside the enzymatic cavity. These descriptors are then fed to a Neural Network and the resulting CVs are used to sample the conformational landscape of the enzyme:substrate complex and to calculate $\Delta G^{\ddagger}_{\mathrm{NAC}}$.

As a proof of concept, we investigated the rate-determining step of the glycolysis reaction of a maltopentose sugar catalyzed by human pancreatic α-amylase (HPA) (**Fig. 2a**), an important enzyme in the glucose production for energy acquisition in humans, as well as a major target for the drug treatment of type-II diabetes[26, 27]. With minimal *a priori* knowledge, this pipeline has helped us to unravel the complexity of this enzymatic reaction, reveal three binding modes, and highlight the role of water in the conversion from non-reactive to reactive conformations.

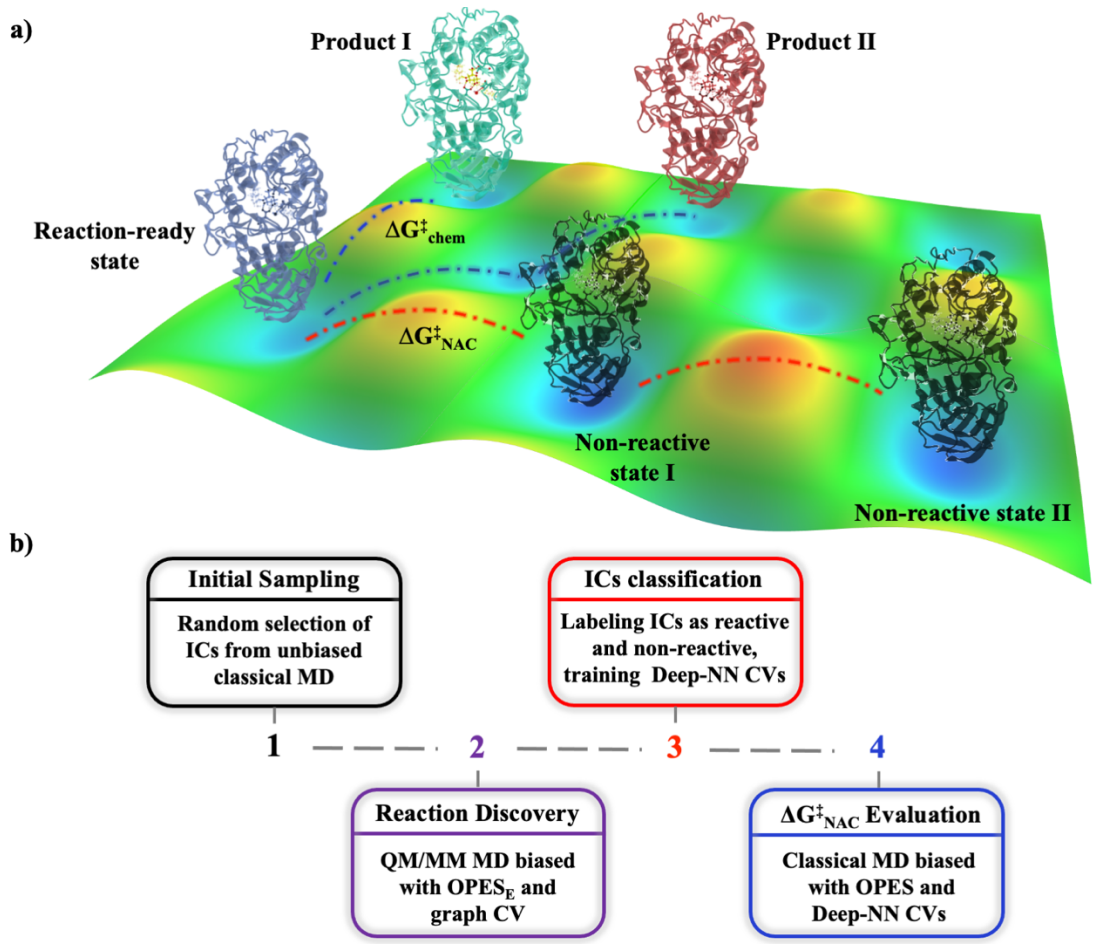


Figure 1. Enhanced Sampling to explore the reaction space of the enzyme:substrate complex. (a) Schematic representation of the complexity of the free energy landscape: the substrate can be bound in multiple ways to the catalytic active site, and only one can have the optimal atomic setup necessary to react. Furthermore, several competitive reaction pathways can be accessible. $\Delta G^{\ddagger}_{chem}$ (free energy required for the chemical transformation) and $\Delta G^{\ddagger}_{NAC}$ (free energy required to reach the reaction-ready state) are also highlighted. (b) Schematic representation of our workflow. The first step involves the generation of initial configurations (ICs) via classical unbiased molecular dynamics, followed by blind exploration of the reaction space via biased QM/MM simulations (step 2). Steps 3 and 4 focus on the machine learning-based (Deep Neural Network or Deep-NN) classification of the ICs as reactive and non-reactive, and estimation of the $\Delta G^{\ddagger}_{NAC}$.

Results and Discussion

HPA consists of three structural domains with the active site located in the largest one made of eight-stranded parallel β-barrel[28] (**Fig. 2a**). The maltopentose polysaccharide (G5) is bound in the catalytic pocket, forming short hydrogen bonds between the third glucose unit (henceforth referred as the reactive sugar ring) and the enzyme residues Asp197, Glu233 and Asp300[28, 29] (**Fig. 2b**). In addition, a crystallographic water molecule (W_1) hydrogen bonded to Glu233 is also present in the active site.

We started by performing a 1μs of classical MD on a solvated enzyme:substrate system, built from an available crystallographic structure of the human pancreatic α-amylase in complex with the acarbose inhibitor (PDB ID: 1CPU) (computational details are reported in the Supplementary Information (SI)), from which ten initial configurations of the enzyme:substrate complex were randomly selected (a snapshot of each IC is shown in **Fig. S1**). The discovery phase was initiated from these ICs, and consisted in performing ten sets of QM/MM molecular dynamics simulations for an aggregate time of ~2 ns using barrier cutoff values between 12 and 60 kcal/mol in the explore version of OPES ($OPES_E$) (the reaction discovery workflow is explained in details in the SI).

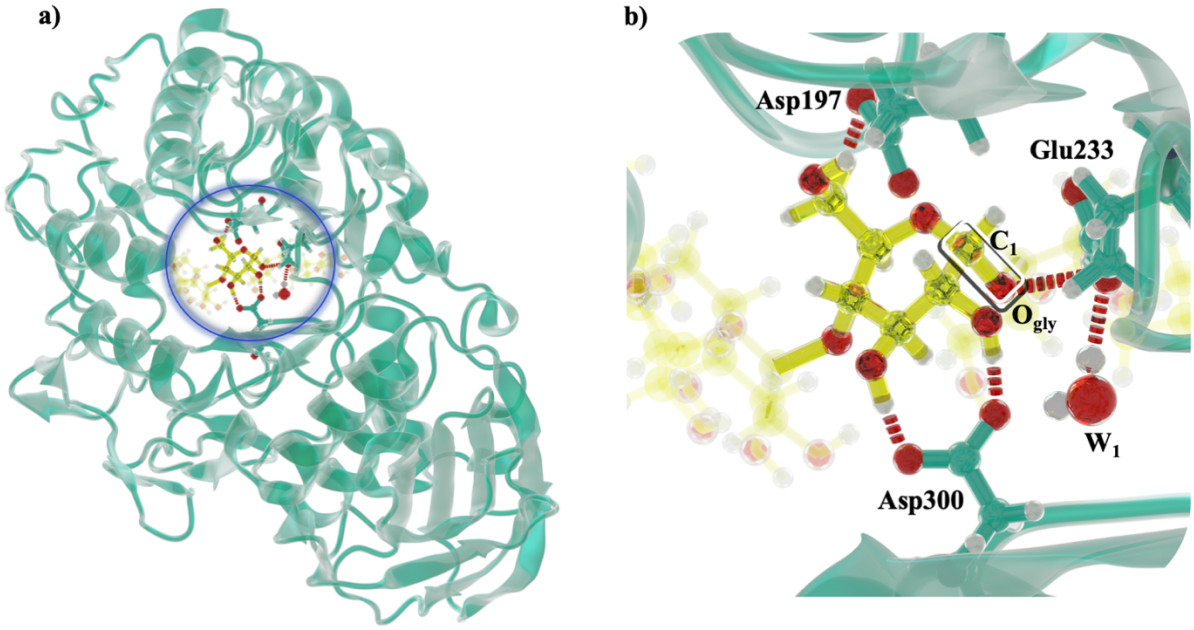


Figure 2. (a) Human Pancreatic α-Amylase, with its active site highlighted within a blue circle (PDB code: 1CPU). (b) Zoomed view of the catalytic pocket showing the maltopentose substrate hydrogen-bonded to the catalytic residues Asp197, Glu233 and Asp300 through its third glucose unit. As discovered from the biased QM/MM simulations, this binding mode promotes cleavage of the glycosidic bond between the third and fourth glucose rings (C_1-O_{gly} bond highlighted within a rectangle), leading to maltotriose and maltose. Furthermore, in the crystallographic structure of ligand-bound HPA[28], a water molecule (W_1) is hydrogen-bonded to the carboxylic side chain of Glu233. The hydrogen bonds are represented by red dotted lines. For clarity, the sugar units attached to the reactive sugar ring is transparent.

The outcomes of the discovery phase are shown in **Fig. 3** and **Table S1** and **S2**. Discovery simulations starting from a small barrier value (ε ~12 kcal/mol) were already able to explore the enzymatic reaction space. Indeed, four of the ten ICs successfully underwent chemical reactions within 5-17 ps, whereas for the other six ICs, no chemical reaction was observed in the 100 ps long QM/MM MD simulations. The observed reaction mechanism consisted of a nucleophilic

attack by the carboxylate group of Asp197 on the C_1 carbon and a cleavage of the sugar glycosidic bond C_1-O_{gly} (**Fig. 3**), 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}. In these four reactive simulations, we observed two proton transfer mechanisms: i) a direct PT between Glu233 and O_{gly} (observed in one simulation, **Fig. 3a**); ii) a water mediated PT where the crystallographic water molecule hydrogen bonded to Glu233 (W_1 in **Fig. 2b**) acts as a bridge between the Glu233 and O_{gly} (observed in three simulations, **Fig. 3b**).

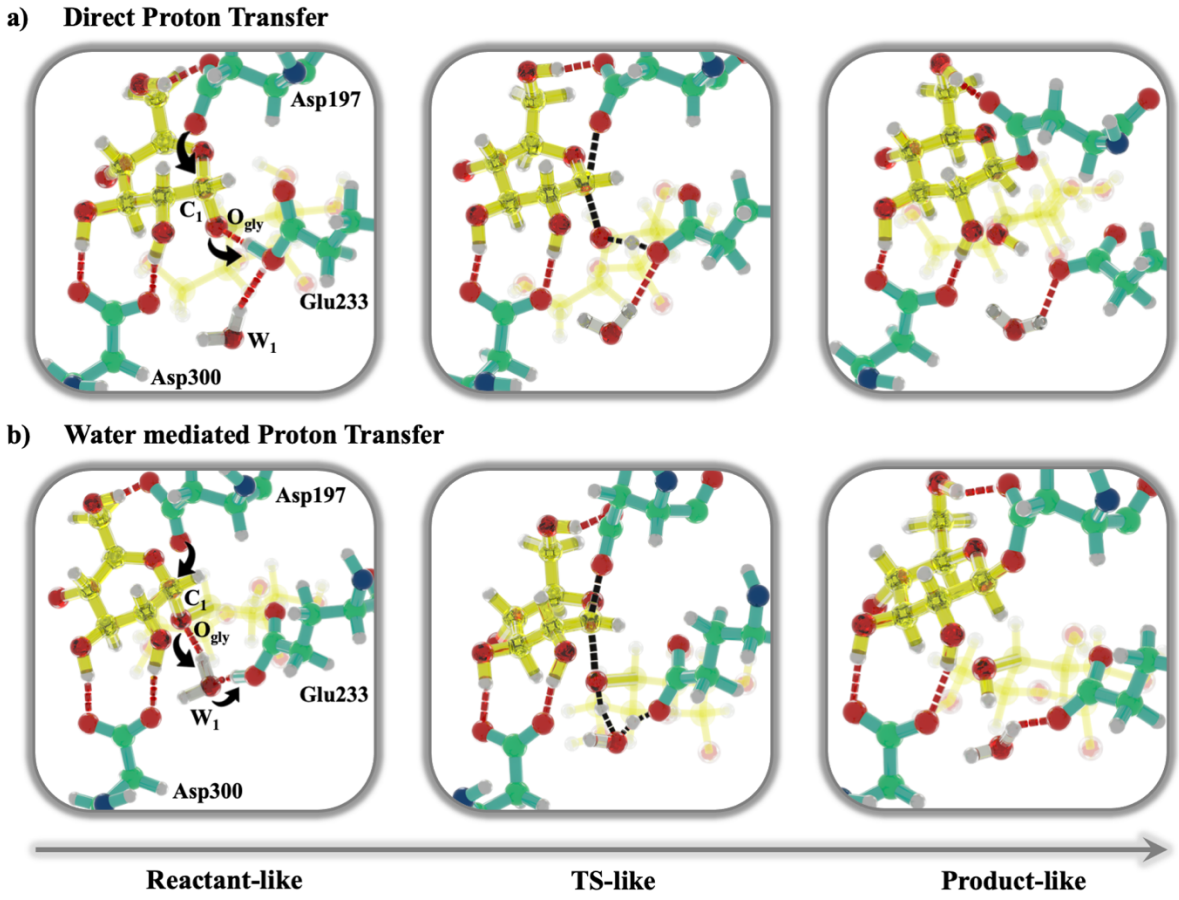


Figure 3. Pathways discovered from QM/MM biased molecular dynamics. (a) Direct proton transfer mechanism comprising nucleophilic attack of one of the carboxylic oxygens of Asp197 on the C_1 center and direct transfer of the acidic proton from Glu233 to O_{gly}. (b) Water mediated mechanism comprising nucleophilic attack of one of the carboxylic oxygens of Asp197 on the C_1 center and proton transfer mediated by the crystallographic water W_1. The black arrows in the left panels indicate the flow of electrons during the reaction. The black dotted lines in the central panels represent covalent bonds formed or broken during the reaction. The hydrogen bonds are represented by red dotted lines. For clarity, the sugar unit attached to the reactive sugar ring is transparent.

We further analyzed the reaction mechanism by examining the time evolution of some structural parameters, and inspected the centers of the maximally localized Wannier functions during the reaction. This analysis suggested that the glycosylic process follows a S_N2-like nucleophilic substitution reaction, where the electrophilic center C_1 undergoes an sp^2 hybridization (**Fig. S2** and **S3**). This mechanism is in overall agreement with that proposed in the literature based on chemical-intuition driven approaches[29-31].

For the six ICs that did not react with an $OPES_E$ ε value of 12 kcal/mol, we performed new simulations increasing ε to 36 kcal/mol; however, reactive events were still not observed (**Table S2**). Only when ε was increased to 48 kcal/mol we could observe two ICs that followed the water-mediated PT mechanism within 35 ps. We further increased the ε value to 60 kcal/mol for the remaining four unreactive ICs, observing chemical reactions via direct PT for two initial configurations while the other two did not react in 100 ps (**Table S2**).

Based on the results of the discovery simulations, four ICs reacted with a low ε value (12 kcal/mol), four reacted only with very high ε values (48 or 60 kcal/mol), and two ICs did not react at all within 100 ps regardless of the ε. The different reactivity of the ICs and the need for increasing external biases to promote reactive events, suggested that these ICs span the enzymatic reactant space to a good extent, consisting of both reaction-ready and non-reactive basins of the enzyme-substrate complex. We assumed that the four ICs that led to a reactive event with a small applied bias potential, were in a proper molecular arrangement for the reaction to occur; hence, these configurations were labeled as reaction-ready or reactive conformations. The remaining ICs were not in a reaction-ready state, therefore, they required extra bias to reach a reactive conformation and then react. These configurations were considered non-reactive.

To identify the structural motifs that distinguish reactive and non-reactive conformations, we analyzed all ICs and discovery dynamics in terms of reacting contacts and their solvation. From this analysis, we identified four criteria to discriminate the two ensembles of conformations: 1) availability of the nucleophile, 2) availability of proton donor, 3) structural hydrogen bonds, and 4) hydration environment (**Fig. 4a**).

For a successful nucleophilic attack, one of the carboxylate oxygens of the nucleophilic residue (O_2@Asp197) should be hydrogen bonded to the -CH_2OH of the reactive sugar ring, while the other carboxylic oxygen (O_1@Asp197) should be close to the point of attack (within ~5 Å of C_1

of the reactive sugar ring, criterion 1 in **Fig. 4a**)[4, 29-31]. Furthermore, the nucleophilic oxygen O_1@Asp197 should be available for attack, namely it should not be involved in a hydrogen bond with Glu233. In addition, the acidic proton of Glu233 should be within 5 Å of the glycosidic oxygen O_{gly} (criterion 2 in **Fig. 4a**) for successful proton transfer via a direct or water-mediated mechanism. The reactive conformations should also have structural hydrogen bonds between the carboxylate group of Asp300 and the hydroxyl groups of the reactive sugar ring (C2-OH and C3-OH) along with a value of approximately 120° for the N-C_α-C_β-C_γ dihedral angle of Asp300, which helps to hold the sugar ring in a proper position for nucleophilic attack (criterion 3 in **Fig. 4a**). Finally, the reactive conformations should have no more than two water molecules (one is W_1 required for catalysis, while the other is hydrogen bonded to W_1) within 5 Å of the reactive sugar ring, in such a way that the hydrogen bond network of the reactive contacts is not perturbed by the presence of the solvent.

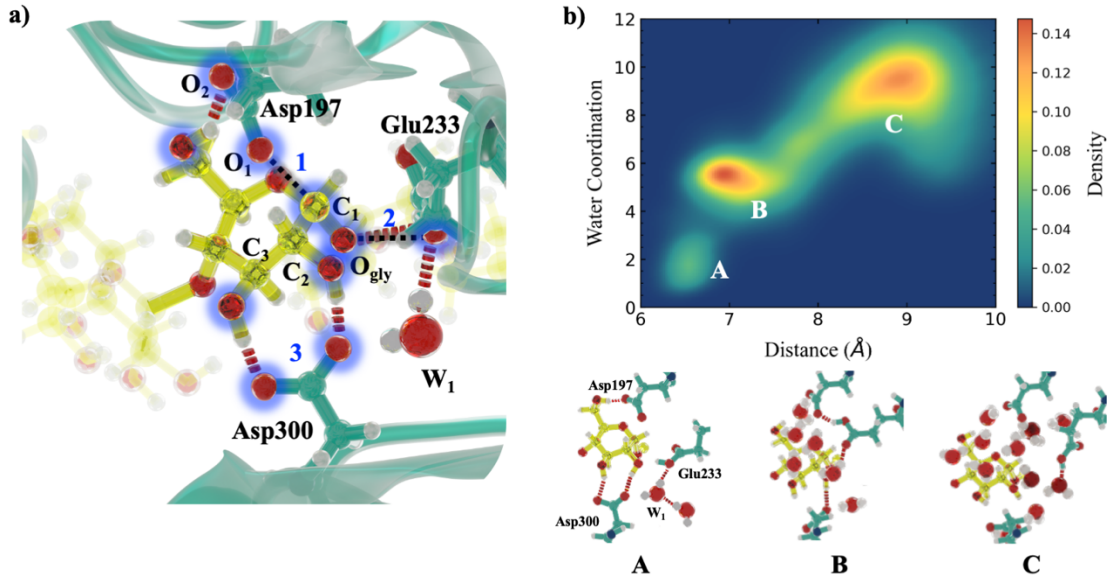


Figure 4. (a) Maltopentose bound to the active site of HPA in a reaction-ready conformation. The criteria to discriminate reactive and non-reactive configurations are reported: 1 availability of the nucleophile, 2 availability of proton donor, 3 structural hydrogen bonds. Atoms involved in reactive contacts are highlighted by blue circles. Criteria 1 and 2 are visualized as black dotted lines representing the distances between O_1@Asp197 and C_1, and between O_{gly} and O@Glu233, respectively. Hydrogen bonds are represented by red dotted lines. (b) Top: Two-dimensional density map projected into the Distance-Water Coordination space, calculated from ten unbiased MD trajectories. Here, Distance represents the distance of the center of mass of the reactive sugar ring from the C_α atom of the nucleophilic residue Asp197, while Water Coordination represents the number of water molecules within 5 Å of the reactive sugar ring. Bottom: Molecular representation of the three substrate-binding modes A, B, and C, as identified from the density map.

We performed ten unbiased MD trajectories starting from the equilibrated crystallographic structure (see computational details section of SI), and used these criteria to identify reactive and non-reactive conformations. **Figure 4b** shows the 2D distribution of the distance of the reactive sugar ring from the C_α of the nucleophilic residue Asp197, and the water coordination around 5 Å of the reactive sugar ring in the unbiased MD simulations.

This analysis clearly showed three possible substrate-binding modes to the enzymatic active site. The reaction-ready basin (state A, **Fig. 4b**) is characterized by having the sugar ring very close to the nucleophilic residue Asp197 (the distance between the center of mass of the reactive ring and C_α of Asp197 is ~ 6.5 Å), and satisfies all four criteria. Furthermore, state A contains two water molecules located near the acidic proton of Glu233, which can assist the proton transfer reaction as discussed earlier. Another substrate-binding mode (state B, **Fig. 4b**) has the substrate slightly away from the nucleophilic center (~ 7 Å) and it is characterized by six water molecules around the reactive sugar ring that break some of the reactive contacts, making the nucleophile and acidic proton unavailable for catalysis, since they form hydrogen bonds with each other. Hence, state B is a non-reactive state. Finally, in state C the reactive ring is far from the nucleophile (~ 9 Å) and solvated by around ten water molecules, which also does not satisfy any of the criteria required for a reactive conformation. Overall, during the transition from state A to C, solvent molecules enter into the active site pocket pushing the maltopentose substrate away from the nucleophile, thus affecting its ability to react.

Although we have identified these three substrate-binding modes, the full characterization of the enzymatic activity requires the estimation of the free energy landscape of the reactive-non-reactive transformation, and the corresponding $\Delta G^{\ddagger}_{\mathrm{NAC}}$. Recovering such a complex free energy landscape is extremely challenging because several collective degrees of freedom are involved (i.e., reactive contacts, and hydration environment), resulting in a multidimensional free energy hypersurface. To reduce the dimensionality of the problem, we used machine learning-based tools, particularly Deep Targeted Discriminant Analysis[32] (Deep-TDA), which allows the construction of highly efficient non-linear collective variables that can distinguish among multiple states.

To drive the transition between the three binding poses we defined two Deep-CVs (an extended discussion on these CVs is reported in the SI). The first collective variable (Contact CV) was trained using the nucleophile-substrate distance, and all the reactive contacts as descriptors for the

neural network. The second CV (Water CV) describes the hydration environment inside the catalytic active pocket, by identifying the location of hydration spots around the reactive contacts where water molecules reside for 0.5 ns or beyond. These hydration spots are used as centers around which we calculated the water coordination that, in turn, was fed into the neural network (extended details are provided in refs. [33, 34]).

To verify the capability of these Deep-CVs to discriminate among the three states, we projected the ten unbiased MD trajectories into the two-dimensional Deep CVs space observing a path connecting states A, B, and C (**Fig. S9**). We used this path to define a path collective variable[35] in the Deep-CV space, which has been subsequently biased with the OPES method to obtain the free energy landscape of the A, B, C transformation.

The converged free energy profile is reported in **Fig. 5a.** The first local minimum along the path is the crystallographic state[28] (state X-ray in **Fig. 5a**) which satisfies all the criteria for reactive conformations, and shows only one water molecule (W_1) in the region surrounded by O_{gly}, Glu233 and Asp300. Close to the X-ray state, there is another reactive state (A) which contains one more water molecule that solvates W_1. This molecule enters from the bulk (**Fig. 5b**) and can take part in the second step of glycolysis reaction, hydrolysing the covalently bound enzyme:substrate complex to the final products[29]. The rearrangement of the reaction-ready state A begins when these two water molecules move on top of the reactive sugar ring, coming very close to the reactive contacts (**Fig. 5a** and **5b**). This is followed by a small number of additional solvent molecules entering the area that the water molecules have vacated, eventually moving into the active site and transforming state A into state A_1. This is the rate-limiting step in the transition from the reactive state to state C, as it is associated with the highest free energy barrier ($\Delta G^{\ddagger}_{NAC} \sim 5$ kcal/mol).

The presence of these water molecules in the pocket gradually breaks the structural hydrogen bond between the reactive sugar ring and Asp300 (from state A_1 to state B and then state B_1), weakening the reactive contacts. Finally, the loss of all direct electrostatic contacts of the reactive sugar ring with the active site residues in state B_1 pulls the sugar away from the nucleophilic residue Asp197. As a result, the bulk solvent floods the active site, solvating the reactive sugar ring (transition from state B_1 to C), and another non-reactive substrate-binding mode, state C, is reached. Additionally, the free energy profile suggests that the reaction-ready states (X-ray and A) reside in metastable basins with a free energy ~ 2 kcal/mol higher when compared to the non-reactive states B and C.

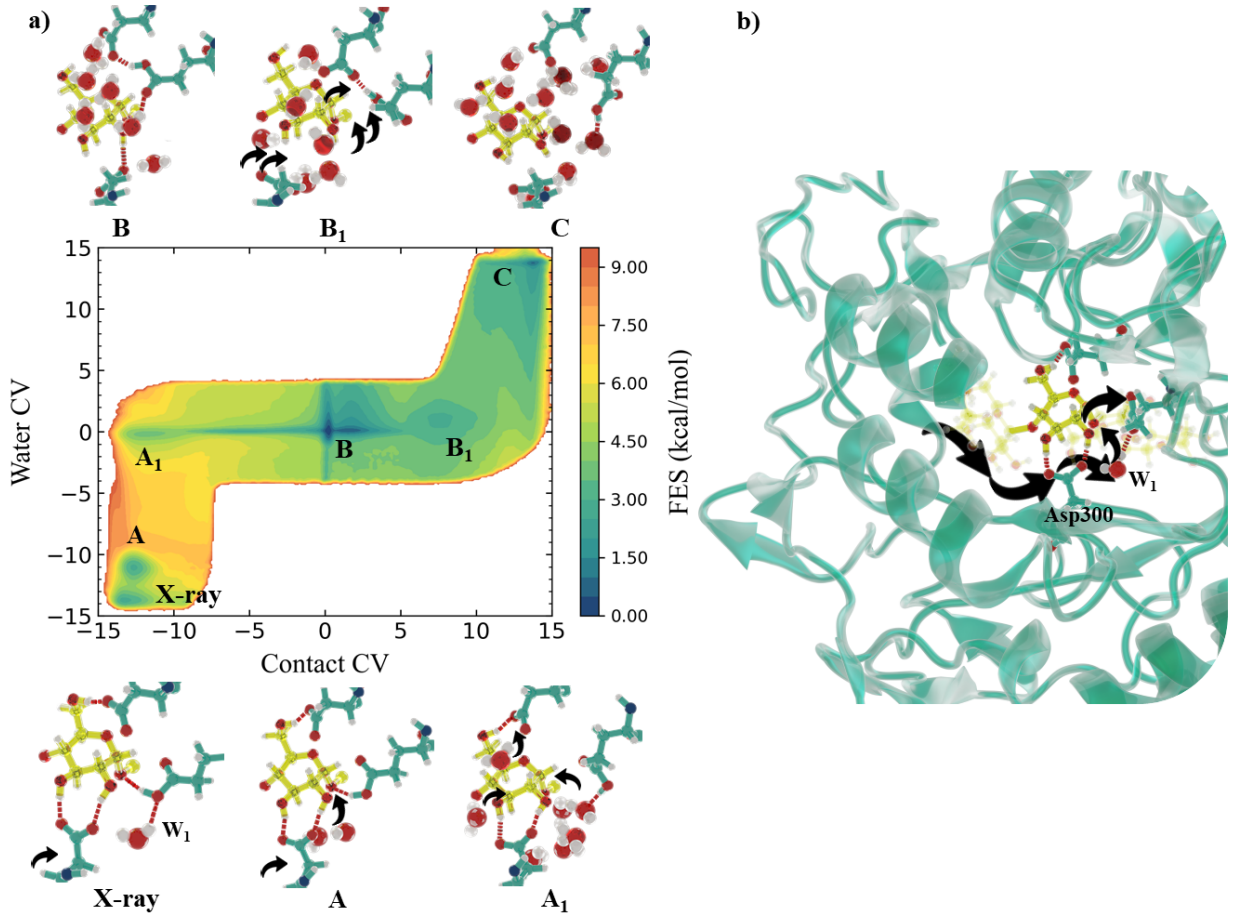


Figure 5. (a) Two-dimensional free energy surface (FES) in the Contact-Water CVs' space (see SI for details). Moving along the Contact CV axis (from -15 to 15), the reactive enzyme:substrate contacts are gradually lost as explained in the text. Along the Water CV axis (from -15 to 15), water molecules gradually enter the active pocket. In the X-ray state (Water CV = -14) only W_1 is present, while in state A a water from the bulk is solvating W_1. States A_1, B and B_1 (Water CV=0) are characterized by approximately six water molecules around the reactive sugar ring, whereas state C (Water CV = 14) has about ten solvent molecules. The FES represents the free energy of the transition between the reaction-ready (state X-ray and A) and inactive basins (states B and C), and provides an estimate of $\Delta G^{\ddagger}_{NAC}$. The molecular structures of the corresponding basins found in the FES are also shown. The black arrows represent the movement of water molecules from one state to another. (b) Path followed by water molecules to enter the HPA active site from the bulk solvent. First, the water molecules reach the substrate-binding residue Asp300, coming close to W_1 (state X-ray to A) and then on top of the reactive sugar ring (state A to A_1), finally moving further into the enzyme pocket (state B_1 to C).

Conclusions

In conclusion, our approach allowed understanding the complexity of the glycosylation reaction catalyzed by human pancreatic α-amylase. Using state-of-the-art enhanced sampling techniques, we revealed the details of its catalytic mechanism at the atomistic level, discovering three different binding modes in the active pocket, and unraveling the role of water molecules in the reactive process, and in the stepwise conversion of reactive to non-reactive conformations. These atomistic

insights on the ligand binding/unbinding process to the active site can aid in the rational design of novel inhibitors able to obstruct the catalytic pocket of human pancreatic α-amylase more strongly. Overall, our workflow represents a promising tool for exploring the complex conformational landscape of the enzymatic active site, unraveling the underlying catalytic mechanism, and thus shedding light on two fundamental questions in enzymatic catalysis, that is how and when does an enzyme react? Furthermore, its potential can be increased when combined with massively parallel[36, 37] or GPU-accelerated[38, 39] QM/MM calculations which allow the treatment of large QM regions[40-42].

ASSOCIATED CONTENT

Supporting Information

Computational details (system preparation, classical and QM/MM MD simulations); Enhanced sampling simulations (reaction discovery, Deep-TDA and Path CVs, ΔG_{NAC} estimation and FES convergence); Supplementary figures and tables.

AUTHOR INFORMATION

Corresponding Author

michele.parrinello@iit.it; mjramos@fc.up.pt

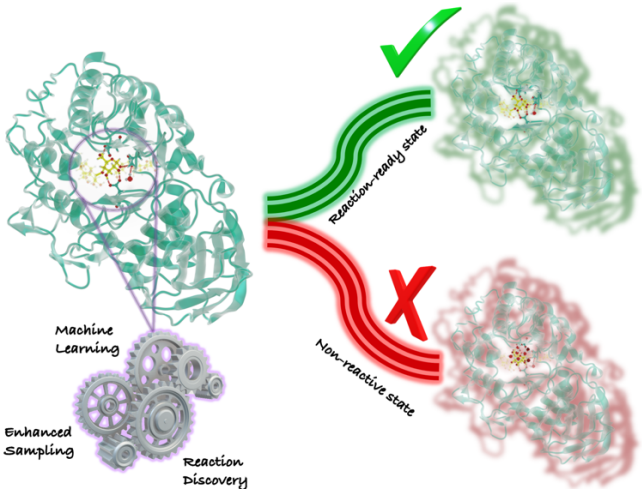
Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgments

The authors would like to thank Enrico Trizio, Narjes Ansari, Valerio Rizzi and Dhiman Ray for helpful discussions. We acknowledge the financial support of the Associate Laboratory for Green Chemistry Unit – LAQV, which is financed by national funds through FCT/MCTES – Portuguese Foundation for Science and Technology within the scope of the project UIDB/50006/2020. RPPN thanks FCT for funding through the Individual Call to Scientific Employment Stimulus (Ref. 2021.00391.CEECIND/CP1662/CT0003).

Table of Contents



References

(1) Benkovic, S. J.; Hammes, G. G.; Hammes-Schiffer, S. Free-Energy Landscape of Enzyme Catalysis. *Biochemistry* **2008**, *47* (11), 3317-3321. DOI: 10.1021/bi800049z.
(2) 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* (21), 12015-12020. DOI: 10.1073/pnas.1534873100 (acccessed 2023/02/23).
(3) Brás, N. F.; Fernandes, P. A.; Ramos, M. J. QM/MM Studies on the β-Galactosidase Catalytic Mechanism: Hydrolysis and Transglycosylation Reactions. *Journal of Chemical Theory and Computation* **2010**, *6* (2), 421-433. DOI: 10.1021/ct900530f.
(4) 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* (5), 4055-4063. DOI: 10.1021/acscatal.7b04400.
(5) Pomeranz Krummel, D. A.; Altman, S. Multiple binding modes of substrate to the catalytic RNA subunit of RNase P from Escherichia coli. *Rna* **1999**, *5* (8), 1021-1033. DOI: 10.1017/s1355838299990416 From NLM.
(6) Pfaff, L.; Gao, J.; Li, Z.; Jäckering, A.; Weber, G.; Mican, J.; Chen, Y.; Dong, W.; Han, X.; Feiler, C. G.; et al. Multiple Substrate Binding Mode-Guided Engineering of a Thermophilic PET Hydrolase. *ACS Catalysis* **2022**, *12* (15), 9790-9800. DOI: 10.1021/acscatal.2c02275.
(7) Wong, C. P.; Awakawa, T.; Nakashima, Y.; Mori, T.; Zhu, Q.; Liu, X.; Abe, I. Two Distinct Substrate Binding Modes for the Normal and Reverse Prenylation of Hapalindoles by the Prenyltransferase AmbP3. *Angewandte Chemie International Edition* **2018**, *57* (2), 560-563. DOI: https://doi.org/10.1002/anie.201710682.
(8) Bak-Jensen, K. S.; André, G.; Gottschalk, T. E.; Paës, G.; Tran, V.; Svensson, B. Tyrosine 105 and Threonine 212 at Outermost Substrate Binding Subsites –6 and +4 Control Substrate Specificity, Oligosaccharide Cleavage Patterns, and Multiple Binding Modes of Barley α-Amylase 1*. *Journal of Biological Chemistry* **2004**, *279* (11), 10093-10102. DOI: https://doi.org/10.1074/jbc.M312825200.
(9) Wong, N. R.; Sundar, R.; Kazanis, S.; Biswas, J.; Pochapsky, T. C. Conformational heterogeneity suggests multiple substrate binding modes in CYP106A2. *J Inorg Biochem* **2023**, *241*, 112129. DOI: 10.1016/j.jinorgbio.2023.112129 PubMed.
(10) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proceedings of the National Academy of Sciences* **2002**, *99* (20), 12562-12566. DOI: doi:10.1073/pnas.202427399.
(11) Valsson, O.; Tiwary, P.; Parrinello, M. Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Annual review of physical chemistry* **2016**, *67*, 159-184.
(12) Bussi, G.; Laio, A. Using metadynamics to explore complex free-energy landscapes. *Nature Reviews Physics* **2020**, *2* (4), 200-212. DOI: 10.1038/s42254-020-0153-0.
(13) Gao, J.; Ma, S.; Major, D. T.; Nam, K.; Pu, J.; Truhlar, D. G. Mechanisms and Free Energies of Enzymatic Reactions. *Chemical Reviews* **2006**, *106* (8), 3188-3209. DOI: 10.1021/cr050293k.
(14) Kollman, P. Free energy calculations: Applications to chemical and biochemical phenomena. *Chemical Reviews* **1993**, *93* (7), 2395-2417. DOI: 10.1021/cr00023a004.
(15) Adcock, S. A.; McCammon, J. A. Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chemical Reviews* **2006**, *106* (5), 1589-1615. DOI: 10.1021/cr040426m.

(16) Antoniou, D.; Basner, J.; Núñez, S.; Schwartz, S. D. Computational and Theoretical Methods to Explore the Relation between Enzyme Dynamics and Catalysis. *Chemical Reviews* **2006**, *106* (8), 3170-3187. DOI: 10.1021/cr0503052.
(17) Zuckerman, D. M.; Chong, L. T. Weighted Ensemble Simulation: Review of Methodology, Applications, and Software. *Annual Review of Biophysics* **2017**, *46* (1), 43-57. DOI: 10.1146/annurev-biophys-070816-033834.
(18) Wang, L.; Friesner, R. A.; Berne, B. J. Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). *The Journal of Physical Chemistry B* **2011**, *115* (30), 9431-9438. DOI: 10.1021/jp204407d.
(19) Fratev, F.; Sirimulla, S. An Improved Free Energy Perturbation FEP+ Sampling Protocol for Flexible Ligand-Binding Domains. *Scientific Reports* **2019**, *9* (1), 16829. DOI: 10.1038/s41598-019-53133-1.
(20) Invernizzi, M.; Parrinello, M. Rethinking Metadynamics: From Bias Potentials to Probability Distributions. *The Journal of Physical Chemistry Letters* **2020**, *11* (7), 2731-2736. DOI: 10.1021/acs.jpclett.0c00497.
(21) Invernizzi, M.; Piaggi, P. M.; Parrinello, M. Unified Approach to Enhanced Sampling. *Physical Review X* **2020**, *10* (4), 041034. DOI: 10.1103/PhysRevX.10.041034.
(22) Pietrucci, F.; Andreoni, W. Graph Theory Meets Ab Initio Molecular Dynamics: Atomic Structures and Transformations at the Nanoscale. *Physical Review Letters* **2011**, *107* (8), 085504. DOI: 10.1103/PhysRevLett.107.085504.
(23) Raucci, U.; Rizzi, V.; Parrinello, M. Discover, Sample, and Refine: Exploring Chemistry with Enhanced Sampling Techniques. *The Journal of Physical Chemistry Letters* **2022**, *13* (6), 1424-1430. DOI: 10.1021/acs.jpclett.1c03993.
(24) Invernizzi, M.; Parrinello, M. Exploration vs Convergence Speed in Adaptive-Bias Enhanced Sampling. *Journal of Chemical Theory and Computation* **2022**, *18* (6), 3988-3996. DOI: 10.1021/acs.jctc.2c00152.
(25) Raucci, U.; Sanchez, D. M.; Martínez, T. J.; Parrinello, M. Enhanced Sampling Aided Design of Molecular Photoswitches. *Journal of the American Chemical Society* **2022**, *144* (42), 19265-19271. DOI: 10.1021/jacs.2c04419.
(26) Jayaraj, S.; Suresh, S.; Kadeppagari, R. K. Amylase inhibitors and their biomedical applications. *Starch-Stärke* **2013**, *65* (7-8), 535-542.
(27) 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* (17), 3813.
(28) 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* (16), 4778-4791. DOI: 10.1021/bi9921182.
(29) Pinto, G. P.; Brás, N. F.; Perez, M. A. S.; 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* (6), 2508-2516. DOI: 10.1021/acs.jctc.5b00222.
(30) Neves, R. P. 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* (15), 3638-3650. DOI: 10.1021/acs.jcim.2c00691.
(31) Neves, R. P.; Cunha, A. V.; Fernandes, P. A.; Ramos, M. J. Towards the Accurate Thermodynamic Characterization of Enzyme Reaction Mechanisms. *ChemPhysChem* **2022**, *23* (13), e202200159.
(32) Trizio, E.; Parrinello, M. From Enhanced Sampling to Reaction Profiles. *The Journal of Physical Chemistry Letters* **2021**, *12* (35), 8621-8626. DOI: 10.1021/acs.jpclett.1c02317.
(33) Ansari, N.; Rizzi, V.; Parrinello, M. Water regulates the residence time of Benzamidine in Trypsin. *Nature Communications* **2022**, *13* (1), 5438. DOI: 10.1038/s41467-022-33104-3.
(34) Hydration spot. https://github.com/narjesansari/Hydration_spot.git (accessed.
(35) Branduardi, D.; Gervasio, F. L.; Parrinello, M. From A to B in free energy space. *The Journal of chemical physics* **2007**, *126* (5), 054103.
(36) 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* (6), 3810-3823. DOI: 10.1021/acs.jctc.9b00093.
(37) 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* (10), 5601-5613. DOI: 10.1021/acs.jctc.9b00424.
(38) 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* (4), 044801. DOI: 10.1063/5.0130886.
(39) Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Isborn, C. M.; Kokkila-Schumacher, S. I. L.; 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. *WIREs Computational Molecular Science* **2021**, *11* (2), e1494. DOI: https://doi.org/10.1002/wcms.1494.
(40) 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* (18), 4415-4420. DOI: 10.1021/acs.jpclett.1c00361.
(41) 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* (16), 7254-7258. DOI: 10.1021/jacs.9b13588.
(42) Walker, A. R.; Wu, B.; Meisner, J.; Fayer, M. D.; Martínez, T. J. Proton Transfer from a Photoacid to a Water Wire: First Principles Simulations and Fast Fluorescence Spectroscopy. *The Journal of Physical Chemistry B* **2021**, *125* (45), 12539-12551. DOI: 10.1021/acs.jpcb.1c07254.