

## Review Article

Kirill Zinovjev\*, J. Javier Ruiz-Pernía\* and Iñaki Tuñón\*

# The quantum revolution in enzymatic chemistry: combining quantum and classical mechanics to understand biochemical processes

<https://doi.org/10.1515/pac-2025-0500>

Received April 29, 2025; accepted June 20, 2025

**Abstract:** This paper reflects the authors' personal journey in applying quantum chemistry methods to understand one of the most important processes in nature: enzymatic catalysis. The integration of quantum mechanics with biomolecular simulations represents one of the most significant advances in computational enzymology over the past few decades. This approach has revolutionized our understanding of enzyme function and catalytic mechanisms, and has provided powerful tools for enzyme design and optimization. Combined quantum mechanics/molecular mechanics (QM/MM) methods have transformed theoretical studies of enzymatic reactions from qualitative descriptions to quantitative predictions, capable of guiding experimental work with unprecedented accuracy.

**Keywords:** Enzymes; machine learning potentials; QM/MM methods; quantum science and technology; transition state theory.

## The foundation of modern computational enzymology

The 2013 Nobel Prize in Chemistry marked a milestone for computational biochemistry, signalling the maturity achieved by methods developed during previous decades and their application to the study of enzymatic catalysis. The Royal Swedish Academy of Sciences awarded the prize to three theoretical chemists – Martin Karplus, Michael Levitt and Arieh Warshel – “*for the development of multiscale models for complex chemical systems*”. This recognition highlighted the transformative impact of their work, particularly in combining quantum mechanics (QM) and molecular mechanics (MM) to simulate biomolecular processes, addressing the fundamental challenge of balancing computational accuracy with efficiency when modelling large systems.<sup>1</sup>

At its core, the QM/MM approach divides the enzymatic system into at least two regions: a quantum mechanical (QM) region encompassing the active site where chemical transformations occur, and a molecular mechanical (MM) region comprising the remainder of the protein and solvent environment (see Fig. 1). The QM region is treated using methods that can describe electronic structure, allowing for the modelling of bond breaking and formation, charge transfer, and polarization effects. Meanwhile, the MM region is represented using classical force fields that efficiently capture structural and environmental effects through simplified potential energy functions. The idea of applying different computational methods to different parts of the same

**Article note:** A collection of invited papers to celebrate the UN's proclamation of 2025 as the International Year of Quantum Science and Technology.

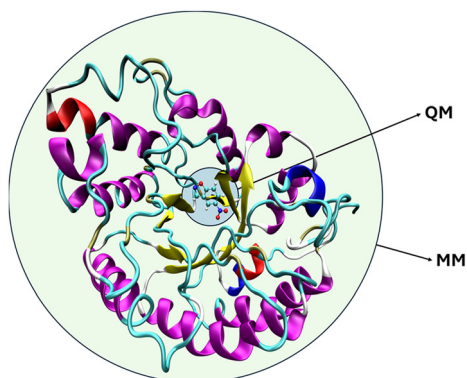
**\*Corresponding authors:** Kirill Zinovjev, J. Javier Ruiz-Pernía and Iñaki Tuñón, Departamento de Química Física, Universidad de Valencia, 46100 Burjassot, Spain, e-mail: kirill.zinovjev@uv.es (K. Zinovjev), j.javier.ruiz@uv.es (J. Javier Ruiz-Pernía), Ignacio.tunon@uv.es (I. Tuñón). <https://orcid.org/0000-0002-6995-1838> (I. Tuñón)

molecular system was first put into practice with the calculation of the spectra of organic molecules with  $\pi$ -electrons.<sup>2</sup> The first application of QM/MM methods to the study of enzymatic reactivity was the study of carbonium ion formation in the active site of lysozyme.<sup>3</sup> This seminal contribution already introduced a key idea for understanding enzymatic catalysis that has remained basically unchanged to this day: *“It is found that electrostatic stabilization is an important factor in increasing the rate of the reaction step that leads to the formation of the carbonium ion intermediate.”*

The fundamental insight that made QM/MM revolutionary was its ability to overcome the restrictions inherent to quantum mechanical calculations when applied to very large systems. QM/MM methods enabled researchers to use quantum mechanical calculations to study reactions taking place in enzymatic active sites, representing the surrounding protein environment with computationally efficient yet realistic molecular mechanics descriptions. In the decades following the pioneering work of Karplus, Levitt, and Warshel, QM/MM methods have evolved considerably, both regarding methodology and applications.<sup>4–7</sup> The original concepts have been refined with improved boundary treatments, more sophisticated quantum mechanical methods, and enhanced sampling techniques. Computational power advancements have enabled increasingly accurate calculations on larger systems, expanding the scope and impact of QM/MM studies in enzymology and drug design.<sup>8,9</sup> QM/MM approaches have become essential tools for understanding how enzymes achieve their remarkable catalytic efficiencies, providing insights that would be inaccessible through experimental methods alone. This theoretical framework has transformed our conceptual understanding of enzyme function from the simplified lock-and-key model to a detailed electronic description of enzymatic catalysis. The foundational principles established by Karplus, Levitt, and Warshel created a methodological bridge between the quantum world and biological macromolecules, enabling researchers to explore the fundamental physical processes underlying enzymatic catalysis.

The 2013 Nobel Prize in Chemistry celebrated the groundwork for computational enzymology by enabling atomic-level simulations of enzymatic reactions, while the 2024 Nobel Prize in Chemistry represents its logical culmination, a conceptual leap from understanding to design. The 2024 laureates – David Baker, Demis Hassabis, and John Jumper – developed computational tools to design sequences and structures of enzymes.<sup>10,11</sup> This progression closes the loop between computational analysis and synthetic biology, with QM/MM acting as a critical connection. By elucidating reaction mechanisms, QM/MM simulations generated fundamental principles, such as transition-state stabilization and electrostatic preorganization, nowadays employed to engineer functional active sites. Together, these tools have transformed enzymatic catalysis from a natural phenomenon to be dissected into a technology to be engineered. The 2013 and 2024 Nobels thus bookend a revolution: QM/MM illuminated *how* enzymes work, while modern design tools now answer *how to build them*.

This contribution reflects our personal experience in the use of quantum chemistry methods to analyze, replicate and predict enzyme catalysis. We briefly discuss the main methodological options, the theoretical framework needed to apply these methods to enzymatic catalysis and the main lessons learned about enzymatic catalysis during the last decades. We also discuss some of the possible directions for the field in the near future. The celebration of the International Year of Quantum Science and Technology is an excellent opportunity to recall



**Fig. 1:** Schematic representation of the division of enzyme-substrate complex into a QM and a MM subsystems.

what quantum chemistry has contributed and will contribute not only to understand nature, but also to improve the life of human beings, assisting in the development of new weapons to fight diseases or of new catalysts for the chemical industry.

## Quantum methods in enzymatic reactivity and their connection to rate theory

The computational analysis of chemical reactivity is necessarily based on the use of quantum chemistry tools. A correct description of the changes in the electronic structure of the reacting molecules is required to quantify the energy changes taking place during the evolution of the nuclei from a rearrangement corresponding to reactants to one corresponding to products. This can be routinely performed for small systems if the surroundings can be ignored, as in gas phase reactions, or if it can be simplified to a passive unreactive environment, as in continuum model treatments. However, the situation is completely different in the case of enzymes. In this case the electronic changes associated with the chemical reaction take place in a complex and flexible environment that influences the energetics of the process and, that in turn, responds to changes in the chemical system, becoming an active agent of the process. Treating the full system, reacting region plus environment, on an equal footing quantum mechanical basis for a full exploration of the configurational space is out of reach even today, except for limited applications in relatively small systems.<sup>12</sup> As discussed below, this limitation could be alleviated in the near future with the aid of Artificial Intelligence.<sup>13,14</sup>

The participation of the protein environment in the chemical reaction can be modelled using different strategies. Even if we have already introduced QM/MM methods, because of historical reasons, we will follow here an order based in the complexity of the models. A first approach involves simplifying the full protein–ligand system into a minimal model that includes only the atoms directly involved in bond breaking and formation, along with the nearest neighbours that significantly influence the energetics of the process. The so-called quantum chemical cluster approach has been successfully used to unravel complex enzymatic reaction mechanisms, allowing the use of high-level quantum chemistry methods that would not be affordable in other models.<sup>15</sup> Cluster models must be necessarily complemented with configurational restrictions to avoid that the active site configuration falls apart in the absence of the rest of the system (protein, solvent, counterions,...). For this purpose, the coordinates of some atoms, typically C $\alpha$  atoms, are frozen at their x-ray values. Obviously, the drastic simplifications inherent in these models, including the neglect of long-range electrostatic effects and a limited protein flexibility, inevitably come at a cost. In some cases, cluster models are embedded into a continuum model, with low dielectric constants mimicking protein hydrophobic cores, in an attempt to capture long-range electrostatic effects.<sup>16</sup> However, cluster models do not generally allow capturing fine effects such as the structural consequences of mutations on the protein backbone or changes in the conformational diversity of the protein.

The cluster model is usually employed in combination with the harmonic version of Transition State Theory (TST) to calculate rate constants or the impact of changes and modifications on the rate constant.<sup>17</sup> Stationary structures corresponding to reactants, products, transition states (TSs) and possible intermediates are localized on the potential energy surface (PES) and characterized through the calculation of the hessian matrix. Then their partition functions are evaluated using the standard harmonic-rigid rotor approximation. Potential energy differences and partition functions, or alternatively thermal and entropic corrections to the potential energy, can be then used to estimate rate constants. This well-established procedure can introduce some artifacts in the calculation of vibrational frequencies due to the fact that some atoms remain fixed at their initial positions.<sup>18</sup> This approach also overlooks the rugged nature of the PES in complex systems, which can make predictions based on a single stationary structure insufficient to accurately represent the ensemble averages corresponding to each state.<sup>19</sup>

A better representation of the interplay between the chemical system and the flexible protein environment, while remaining computationally feasible, can be achieved through hybrid QM/MM approaches. There is a wide variety of QM/MM methods available in the literature and the reader is referred to some reviews to learn about

their technical details.<sup>6,9,20</sup> Typically, the interaction between the two subsystems is described using a van der Waals energy term, an electrostatic term and, if needed, terms to account for bonding contributions between both subsystems. A defining feature of these methods is the way in which the QM subsystem is embedded into the MM environment. In the mechanical embedding approach, the MM environment does not polarize the QM electron density. The electrostatic interaction between both subsystems is described by means of partial point charges placed on the positions of the QM atoms. These charges can be fixed (sometimes termed as molecular mechanics embedding) or fitted to the electrostatic potential created by the QM subsystem structure in vacuo. In the electrostatic embedding approach, the charges of the MM subsystem enter in the self-consistent-field procedure, allowing for the polarization of the QM subsystem under the influence of the environment.<sup>21</sup> This last scheme is highly recommended to describe enzymatic reactions because active sites often contain charged groups that can significantly polarize the electronic density of the reacting subsystem.<sup>6,22</sup> The next step in the methodological ladder is the polarization embedding, where the polarizability of MM atoms is also included. In this case not only the QM subsystem, but also the MM atoms can respond to the influence of the QM subsystem and other MM atoms.<sup>23,24</sup> Applications of this last scheme are mostly limited to the analysis of spectroscopic electronic transitions, where the consideration of the electronic response of the solvent to changes in the electronic density of a solute can be important to quantify the transition energies.

QM/MM schemes provide a more realistic and complete representation of the environment, where specific and long-range interactions of the reactive system with the environment can be included. One of the advantages associated with the use of QM/MM schemes is that these hybrid schemes facilitate the use of simulation methods, in particular Molecular Dynamics (MD). QM/MM MD simulations can be used to generate equilibrium and non-equilibrium trajectories of the system under study, providing valuable information about enzymatic reactions.<sup>25</sup> MD simulations can be used to explore many different configurations of the system that could remain inaccessible during the exploration of the PES. Protein configurational changes can take place in the Michaelis complex after substrate binding or during the reaction progress to facilitate the stabilization of the TS.<sup>26,27</sup> The configurational search can be promoted using enhanced sampling techniques, such as metadynamics,<sup>28</sup> umbrella sampling (US)<sup>29</sup> or accelerated Molecular Dynamics (aMD),<sup>30</sup> to name just a few of the available techniques.<sup>31</sup> This conformational sampling makes possible the calculation of ensemble properties, such as the free energy, that can be connected to experimental thermodynamic and kinetic observables.<sup>32</sup> Analysis of QM/MM MD trajectories can also be used to decipher the role of protein structure in catalysis, including subtle effects such as allostery, where a change occurring far from the active site is transmitted through a cascade of conformational changes, ultimately reaching the active site and significantly altering its catalytic efficiency.<sup>33</sup>

Regarding the study of enzymatic reactions, the most obvious application of QM/MM MD simulations is the calculation of the associated free energy profiles. After the selection of an adequate reaction coordinate ( $\xi$ ), an issue we discuss below, enhanced sampling techniques can be used to force the system to visit the full range of values of this coordinate, from reactant to products. The probability distribution of the reaction coordinate obtained from these biased simulations can be then used to recover the underlying free energy profile ( $G_\xi$ ), using for example the weighted histogram analysis method (WHAM) or other reweighting schemes.<sup>34,35</sup> The calculation of the reaction free energy profile provides a different perspective in the use of TST. First, the transition state (TS) can be variationally optimized. TST overestimates the rate constant because it neglects the possibility of recrossing trajectories once the TS has been surpassed.<sup>36</sup> This error can be minimized placing the TS in the position that minimizes the reaction flux, that is at the top of the reaction free energy profile. Second, the estimated free energy profile is now obtained considering the conformational diversity of the system and without assuming a harmonic behaviour for the vibrational modes. Thus, the rate constant can be directly obtained from the difference between the free energy value at the TS and at the reactants ( $\Delta G_\xi^\ddagger$ ) as:<sup>36</sup>

$$k = \kappa_\xi \cdot \frac{\omega_R}{2\pi} e^{-\frac{\Delta G_\xi^\ddagger}{RT}} \quad (1)$$

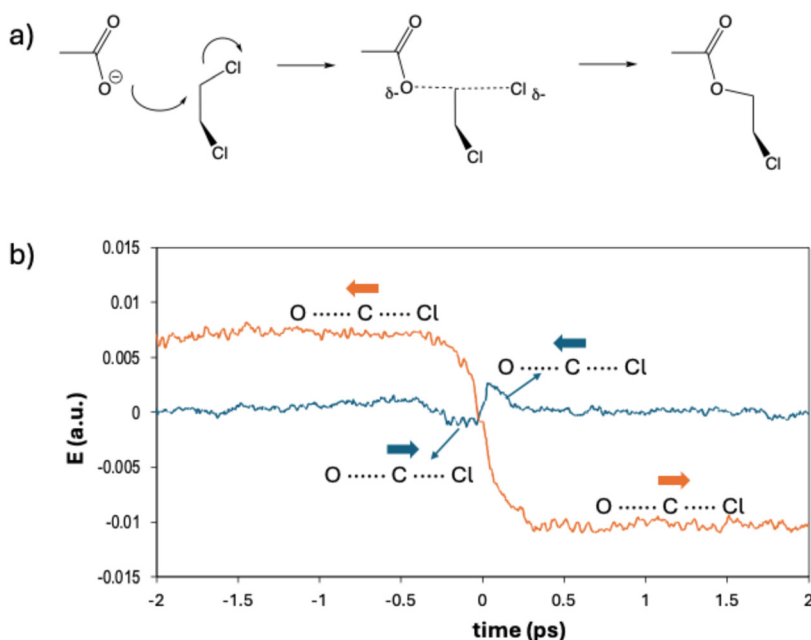
where  $\omega_R$  is the reaction coordinate frequency at the reactants well and  $\kappa_\xi$  is the transmission coefficient that corrects for the fraction of trajectories that recross the dividing surface defined by  $\xi = \xi^\ddagger$ . The transmission

coefficient is usually taken as unity in most of the applications. Deviations from unity can be due to the choice of an incomplete reaction coordinate, which would impose a practical limitation that can be overcome by optimizing the reaction coordinate, and to the complexity of the potential energy of the system, which would impose a fundamental limit to the validity of this theory. As commented below, exhaustive analysis of the limits of TST in enzymatic processes has shown that this theory can be safely used to understand these reactions.<sup>37,38</sup>

## Lessons learned about enzymatic catalysis

Enzymes catalyse chemical reactions by means of electrostatic interactions. The electrostatic environment created by the enzymatic active site has been optimized by evolution to accelerate chemical reactions, selecting the correct residues and placing them at the correct distances and orientations with respect to the reactive fragments. These electrostatic interactions lower the activation free energy with respect to the uncatalyzed process, stabilizing TSs relative to the reactants more than the solvent does. Key elements to explain the electrostatic origin of catalysis are the concepts of preorganization and reorganization, introduced by Warshel.<sup>39</sup> The active site charges and dipoles that interact with the TS are kept together by the protein three-dimensional structure, resulting in an electrostatic environment that is largely independent of the changes taking place in the chemical system. The active site is preorganized for catalysis thanks to the protein structure. Instead, in aqueous solution, water molecules' dipoles must be reoriented to follow the electronic changes taking place during the chemical reaction, which has an associated free energy cost termed as reorganization. In enzymes, this reorganization is reduced because the electrostatic environment provided by the active site is already configured to optimally accommodate the transition state of the reaction. To achieve this, protein folding must place together some charges and dipoles in a reduced space. Enzymatic active sites are thus not optimized for stability but for catalytic efficiency, resulting in a well-known trade-off between these two properties.<sup>40</sup>

Due to the long-range nature of electrostatic interactions QM/MM hybrid models are preferred over cluster models to explain the catalytic efficiency of enzymes. While reaction mechanism can often be explored using cluster models, the impact of distal mutations or other subtle structural changes can only be captured with a more realistic modelling of the protein. Continuum models, often used in combination with cluster models, can even result in electrostatic effects contrary to those produced by the protein in a more faithful representation. This is clearly illustrated with the example of  $S_N2$  reactions, as that represented in Fig. 2a. The nucleophilic attack by an aspartate residue on 1,2-dichloroethane is catalysed by the enzyme Haloalkane Dehalogenase (DhlA). For this type of  $S_N2$  reactions it is well-known that an increase in solvent polarity results in a diminution of the reaction rate. This is due to the fact that the solvent reaction field stabilizes more the reactants and products states than the TS, where the charge is more distributed. Figure 2b shows the average electric field created by the solvent on the carbon atom that suffers a nucleophilic attack obtained for reactive trajectories crossing the TS at  $t = 0$  (with  $t > 0$  corresponding to the evolution towards products and  $t < 0$  to reactants). This electric field was computed from QM/MM simulations in aqueous solution and in the active site of an enzyme.<sup>25</sup> The reaction field created in aqueous solution along the nucleophile-leaving group vector acts as a force against the evolution towards the TS, destabilizing this state with respect to reactants and products. Instead, the electric field created by the protein is not a reaction field. QM/MM simulations on DhlA enzyme show how the enzymatic active site is able to provide an electrostatic stabilization of the TS even if this state is less polar than the reactants.<sup>25,41</sup> The electric field component along the nucleophile-leaving group vector is now much smaller than in water and does not act as a force against the evolution towards the TS. In fact, when the trajectories are close to the TS ( $-200 \text{ fs} > t > 200 \text{ fs}$ ) the electric field created by the enzyme on the carbon atom favours the TS configuration, providing effective stabilization with respect to the reactants. Enzymes must provide this stabilization of the TS without destabilizing the reactants in the active site, otherwise the binding of the substrate could be compromised. In the present example the nucleophile (an aspartic residue) is part of the enzyme structure. In other cases, a cofactor can be recruited to favour substrate binding.<sup>42</sup>

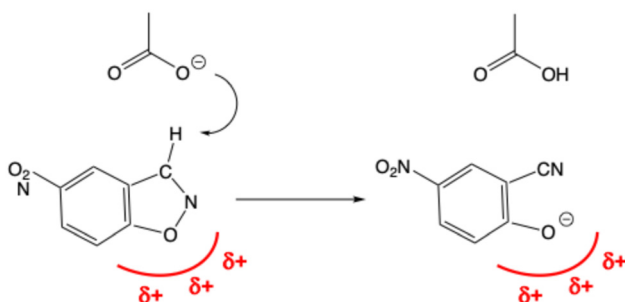


**Fig. 2:** Electrostatic effects in enzyme catalysis. a) Schematic representation of the  $S_N2$  reaction between aspartate and 1,2-dichloroethane. b) Time evolution of the electric field along the nucleophile-leaving group vector averaged for reactive trajectories in aqueous solution (blue) and in the active site of DhIA (orange) going from reactants ( $t < 0$ ) to products ( $t > 0$ ) after crossing the TS at  $t = 0$ . Adapted with permission from ref. 25.

## Lessons learned about enzyme design

QM/MM calculations have played a fundamental role in understanding enzymatic reactions and guiding the rational design of novel biocatalysts, particularly in the context of *de novo* enzyme engineering. A prominent example is the case of Kemp Eliminases (KEs), a family of artificial enzymes designed to catalyze the conversion of benzisoxazoles into salicylonitriles, a reaction not catalyzed by natural enzymes (see Fig. 3). The success of KE engineering relied on the use of theoretical enzymes (theozymes) to model an optimal arrangement of active site residues capable of stabilizing the transition state, primarily by electrostatic interactions with the developing negative charge on the oxygen atom of the substrate. This theozyme-based configuration was subsequently used as a blueprint to design surrounding protein scaffolds. The resulting designs were ranked on the basis of the catalytic geometry and the computed binding free energy for the modelled reaction TS. Designed proteins were further optimized through directed evolution, increasing catalytic efficiency by  $> 200$ -fold.<sup>43</sup>

Later developments based on X-ray crystallography and MD simulations, such as the engineered HG-3 enzyme, introduced structural modifications to enhance the positioning of the catalytic base to favor substrate binding and to exclude water molecules from the active site.<sup>44</sup> Directed evolution then further refined the catalytic efficiency, culminating in the highly active HG-3.17 variant designed by Hilvert and co-workers, which achieved rate constants approaching those of natural enzymes.<sup>45</sup> QM/MM approaches demonstrated that the enhanced activity of evolved KE variants was not merely due to improved substrate positioning but also to a more favorable electrostatic environment that lowered activation barriers. This conclusion was supported by linear



**Fig. 3:** Schematic representation of a kemp elimination reaction in a kemp eliminase enzyme.



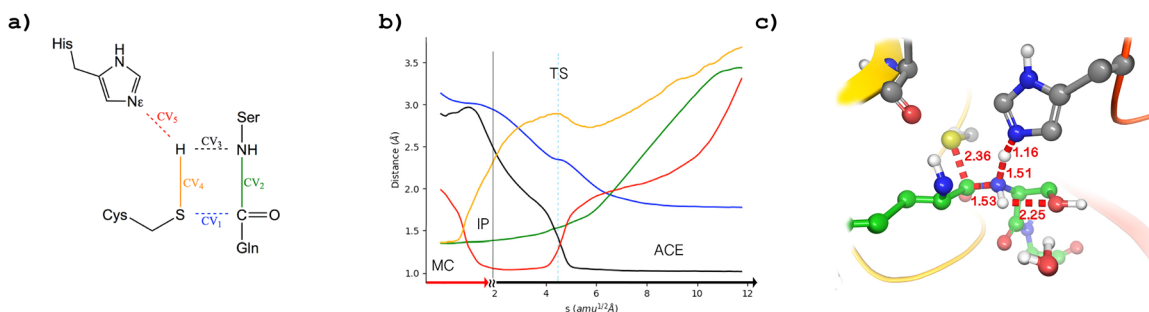
correlations between activation free energies and the electrostatic potential exerted by the enzyme on the oxygen atom that is negatively charged during the reaction (see Fig. 3),<sup>46</sup> reinforcing the notion that electrostatic effects are critical determinants of catalytic efficiency. Beyond understanding natural and designed enzymes, QM/MM calculations have also informed strategies for computational enzyme design. For example, Head-Gordon and co-workers leveraged electric field optimization to predict mutations that enhanced the catalytic performance of a KE variant, demonstrating that computationally guided modifications could substantially improve reaction rates.<sup>47,48</sup> In these studies, the authors enhanced the catalytic efficiency of a *de novo* Kemp Eliminas (KE15) by computationally identifying several mutations that were later validated experimentally. These mutations improved transition state stabilization by optimizing both the electric fields and substrate positioning within the active site. Interestingly, the Asp130Lys mutation was suggested after analysis of long-range electrostatic destabilizing effects. These findings underscore the potential of simulations not only to rationalize how enzymes work but also to guide the design of novel catalysts with tailored functionalities. By elucidating how directed evolution enhances catalysis through electrostatic effects on the transition state, QM/MM simulations can contribute to enzyme engineering, offering a powerful tool for developing efficient biocatalysts for industrial and biomedical applications.

## Lessons learned about QM/MM methods

As previously discussed, the predominant role of electrostatic interactions makes QM/MM the method of choice for computational studies of enzyme catalysis. QM/MM simulations of enzymatic reactions must strike a balance between two key aspects: the accuracy of the energy function used to describe the chemical transformation, and the extent of sampling required to obtain well-converged thermodynamic properties. In practice, computational cost is the limiting factor, necessitating a careful trade-off between the quality of the hamiltonian and the length of the simulations. Only recently has the advancement of computational hardware and software made it feasible to achieve hundreds of picoseconds of sampling in QM/MM simulations with medium-sized QM subsystems (~100 atoms) treated at a reasonable DFT or ab initio level.<sup>49</sup>

Standard applications of QM/MM methods employ DFT or ab initio hamiltonians for the exploration of potential energy surfaces, while QM/MM MD simulations, that require a considerably larger computational effort, are carried out with lower-level energy descriptions for the QM subsystem: semiempirical or tight-binding methods. These methods can introduce severe systematic errors in the evaluation of the relative free energies along the reaction coordinate but they can be sufficient to capture the enzymatic catalytic effect, the difference between the activation free energies obtained for the reaction in aqueous solution and in the enzyme.<sup>50</sup> Several strategies have been developed to attenuate systematic errors associated with the use of these lower-level hamiltonians. One possibility is to add a correction to the energy of the QM subsystem obtained as the difference between single-point calculations carried out at low and high QM levels.<sup>51</sup> Another possibility is the use of specific parameters in low-level hamiltonians developed to describe correctly a particular chemical reaction or type of reactions. This technique has been employed for example to correct the AM1 hamiltonian for the description of the hydride transfer reaction in Dihydrofolate Reductase<sup>52</sup> or to improve the DFTB3 description of phosphate hydrolysis reactions.<sup>53</sup>

On the sampling side, the use of higher-level QM methods implies a significant computational effort. This can potentially lead to missing some configurations relevant for the process under study. One of the major challenges, even with cheaper semiempirical hamiltonians, is the treatment of chemical reactions that need more than a single geometrical coordinate for their description. Some mechanisms can be followed by sampling along a single distance or a combination of distances. For example, an antisymmetric combination of distances is a good reaction coordinate for many transfer reactions, as the  $S_N2$  example presented in Fig. 2. However, some chemical reactions involve more than a single chemical event. The deprotonation of a nucleophile can be coupled to the nucleophilic attack and to the protonation of the leaving group, as found in the proteolysis mechanism of the Main Protease of SARS-CoV-2 (see Fig. 4).<sup>54,55</sup> In such a case, proper sampling of the reaction process may involve the use of more than one distinguished coordinate (up to five bond distances are shown in Fig. 4), which increases



**Fig. 4:** Exploration of complex enzymatic reactions. a) Bond distances involved in the acylation reaction catalysed by the main protease of SARS-CoV-2. B) Evolution of the bond distances along the minimum free energy path. C) Schematic representation of the transition state. Adapted with permission from ref. 54.

exponentially the computational cost. The most popular strategy to explore the free energy dependence along a coordinate is Umbrella Sampling.<sup>29</sup> In this method, the full range of possible values of the coordinate is divided into  $N$  intervals, each of them being explored in a different simulation or window. These windows are then combined to obtain the full probability distribution of the coordinate and then the associated free energy change.<sup>34</sup> The number of simulations required to explore the  $d$  coordinates (dimensions) of a system under study will be approximately  $N^d$ . This is known as the curse of dimensionality. Several strategies have been developed to mitigate this problem. Metadynamics is a successful technique to explore multidimensional free energy landscapes at lower costs, but also exponentially depends on the number of coordinates to be explored.<sup>56,57</sup> One strategy to avoid the curse of dimensionality is the development of methods to trace the minimum free energy path (MFEP). Instead of exploring the whole multidimensional free energy surface, these methods focus on the MFEP region.<sup>58–61</sup> Once the MFEP associated with a particular mechanism is traced, the advance of the system along this path (or path-CV) can be used as an efficient reaction coordinate to collect the sampling needed to trace the free energy profile, reducing the dimensionality of the problem to just one.<sup>62,63</sup> This is illustrated in Fig. 4 for the acylation reaction of the Main Protease of SARS-CoV-2. While several bond distances are involved in the acylation process, they can be integrated into a single path-CV (denoted as  $s$  in panel b).

## Lessons learned about the validity of transition state theory for enzymatic reactions

TST has served as the conceptual framework for interpreting enzymatic reaction rates. The combination of QM/MM simulations for the calculation of the activation free energy along an appropriate reaction coordinate and equation (1) allows the calculation of rate constants and of the effects of several factors on them. TST has been extensively validated in the case of enzyme catalysis, particularly when corrections for quantum effects and dynamical recrossing are properly incorporated. TST has been successfully used to reproduce or predict various kinetic properties, including rate constants, kinetic isotope effects,<sup>64,65</sup> the impact of point mutations on enzymatic efficiency<sup>66,67</sup> and the temperature dependence of reaction rates.<sup>33,68</sup>

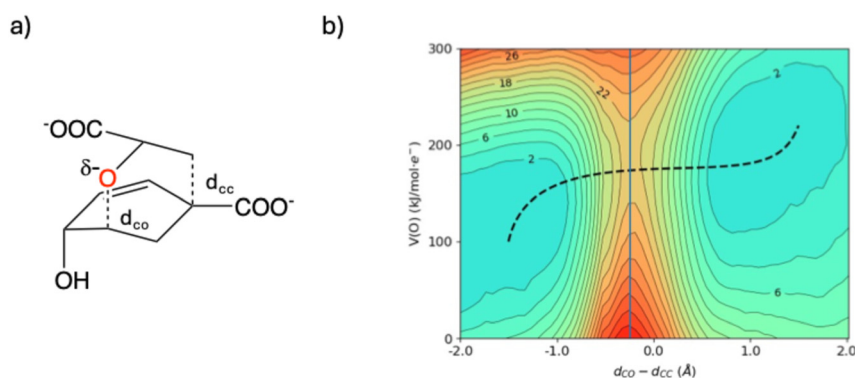
In spite of its apparent success, the validity of TST for explaining enzymatic reactivity has been questioned, with some authors proposing that specific enzyme vibrations, named as promoting vibrations, might actively promote or drive the chemical reaction.<sup>69</sup> According to this hypothesis, the enzyme's dynamics – encoded in its structure – could facilitate the reaction in ways not fully captured by TST. Regarding this point it is important to emphasize that TST is a dynamical theory, in which velocities are thermally averaged under the assumption of an equilibrium distribution. In enzymatic active sites, frequent molecular collisions typically ensure that the system relaxes much faster than the average waiting time between reaction events, thereby supporting one of the central assumptions underlying TST. Obviously, this is not to say that enzymes do not promote chemical reactions but



rather than protein motions can be adequately incorporated in the framework of TST, either as an ingredient of the reaction coordinate (when the protein explicitly participates in the reaction through formation of a covalent intermediate) or assuming that they are at equilibrium. Such a confusion regarding the validity of the TST for enzymatic processes comes from the fact that TST results are sensitive to the choice of the reaction coordinate, and a poor choice may lead to a conclusion that TST is invalid for a system in hand.<sup>70</sup> However, when the reaction coordinate is carefully chosen, the TST was shown to provide excellent descriptions of enzymatic processes even in the cases where the importance of promoting vibrations was hypothesised.<sup>71</sup>

Another feature of enzymatic systems is that any changes in the environment that are required for the reaction to take place are typically slower than the reactive event itself, thus effectively decoupling the slow protein motions from fast (albeit rare) chemical transitions. This can be illustrated with the analysis of the impact of protein electrostatics, a key contribution to catalysis, along the reaction progress. Figure 5 shows the free energy surface for a Claisen reaction, the transformation of chorismate into prephenate catalysed by BsCM enzyme.<sup>50</sup> The free energy was explored using as coordinates an antisymmetric combination of the lengths of the bonds broken and formed during the process and the electrostatic potential that the MM environment creates on the ether oxygen, which develops a negative charge during the process. The MFEP shows that protein fluctuations favouring the stabilization of this charge precede TS crossing, which is essentially governed by internal degrees of freedom of the substrate, while the environment can be considered at equilibrium.<sup>71</sup> The role of protein motions, globally slower than substrate motions leading to the TS, could be relevant in early and late stages of the reaction but not during barrier crossing.<sup>72</sup> In other words, the antisymmetric coordinate is sufficient to define a dividing surface that closely approximates the transition state of the reaction, again supporting the validity of TST for such processes.

In connection with our previous discussion, one of the major challenges in applying TST to enzymatic reactions involves the selection of an appropriate reaction coordinate. Using the variational principle the TS is identified as the maximum along the free energy profile calculated along the selected coordinate, which is critical to ensure the validity of TST assumptions. Enzymatic reactions are highly multidimensional problems, where several chemical events can take place simultaneously (as illustrated in Fig. 4) and where strong intermolecular interactions couple the protein and substrate degrees of freedom of the system. This can make the selection of an adequate reaction coordinate even more difficult. A poorly chosen reaction coordinate can lead to hysteresis and discontinuity problems in the evaluation of the free energy profile using enhanced QM/MM simulations.<sup>73</sup> From a more fundamental perspective, a poorly chosen reaction coordinate can lead to a definition of the TS dividing surface that exhibits a large number of dynamical recrossings, resulting in a transmission coefficient (see Eq. (1)) significantly lower than unity. In that case the transmission coefficient must be evaluated using rare event trajectories<sup>74</sup> or Grote-Hynes theory.<sup>75</sup> In contrast, when the dividing surface is defined using a ‘perfect’ reaction coordinate, the equilibrium committer – that is, the set of configurations in phase space where the probability of reaching the product state before returning to the reactant is exactly 0.5 – TST has been shown to accurately predict enzymatic reaction rates,



**Fig. 5:** Environmental participation in enzymatic reaction coordinates. a) Transition state for the Claisen rearrangement of chorismate to prephenate. b) AM1/OPLS free energy surface obtained for the reaction in the active site of BsCM as a function of an internal coordinate and the electrostatic potential created by the enzyme on the ether oxygen atom. Isoenergetic lines are drawn every 2 kcal mol<sup>-1</sup>. The dividing surface along the antisymmetric RC is shown in blue. The minimum free energy path is shown as a black dashed line.

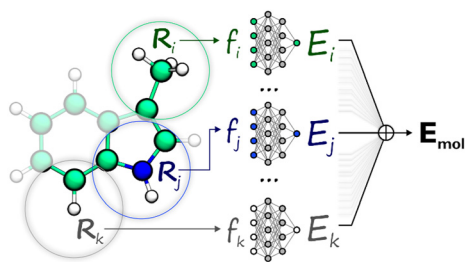
with transmission coefficients exceeding 0.9.<sup>38</sup> The residual recrossings observed even when using the equilibrium committor as the reaction coordinate reflects the intrinsic complexity of the potential energy surface and indicate that the error associated with the core assumptions of TST in enzymatic systems is minimal. Of course, the practical implementation of TST can still introduce significant inaccuracies. Therefore, major efforts should be directed toward improving free energy calculation methodologies, including enhanced sampling strategies, refinement of the underlying Hamiltonians and the selection of appropriate reaction coordinates.

For enzymatic reactions involving the transfer of light particles such as a proton, a hydrogen atom or a hydride, quantum mechanical effects become particularly important. These quantum effects manifest in two primary forms: zero-point energy contributions and tunnelling. In this case, the classical treatment of the motion along the reaction coordinate assumed in TST must be corrected to account for these effects. Several methodologies have been proposed to introduce these effects into QM/MM simulations. Truhlar and coworkers proposed the inclusion of a multiplying prefactor in the TST expression for the rate constant that accounts for tunnelling, while activation free energies are corrected for zero-point energies.<sup>76</sup> Other strategies to include the quantum nature of nuclei motion in QM/MM simulations are obtaining the adiabatic vibrational wave functions solving the time-independent Schrödinger equation<sup>77</sup> or quantizing them via centroid path-integral simulations.<sup>78,79</sup> In connection with the previous discussion, an interesting question is the selection of an adequate reaction coordinate to drive the simulations and that simultaneously allows for a proper quantization of nuclear motions. Recent studies have shown that some coordinates may fail to isolate environmental effects and can bias the description of TSs for H-transfer reactions, affecting the evaluation of some sensitive properties such as kinetic isotope effects.<sup>80</sup>

## The future of quantum methods applied to enzymology. The machine learning era

As discussed in this work, QM/MM simulations have provided a deep understanding of the theoretical principles underlying the extraordinary catalytic efficiency of the enzymes and also of the practical guidelines for rational enzyme design. However, their application has always been hindered by the huge computational cost of the QM potentials needed to describe the chemical transformations during the enzyme lifecycle. To adequately sample the conformational space relevant to an enzymatic process, one would have to perform calculations on the order of  $10^5$ – $10^6$  energy estimates, bringing the total cost easily up to a million CPU hours. More precise approaches, such as coupled cluster,<sup>81</sup> generally considered to provide “chemical precision” of the resulting energy estimates (errors below 1 kcal/mol) are orders of magnitude more expensive and therefore out of reach for QM/MM MD simulations. The high computational cost of QM methods forces researchers to either sacrifice the accuracy, resorting to cheap but approximate semiempirical potentials or to significantly reduce the exploration of the conformational landscape, sometimes bringing the simplification to the extreme: reducing the description of the conformational ensemble to energy estimates of selected structures or employing the so-called cluster models, described above.

A way to overcome the seemingly inevitable compromise between accuracy and computational cost of the QM potentials has recently emerged from the rapidly growing field of Machine Learning (ML). A ML model can be trained on a large dataset of QM calculations to learn the relationship between the molecular structure and the total energy. Once trained, such ML interatomic potential (MLIP) can provide molecular energies and forces significantly faster than the reference QM method without much loss in accuracy, following a scheme such as that represented in Fig. 6. Depending on the reference level of theory, the speed-up could be from  $10^2$  for cheap DFT methods to  $10^5$  and beyond for the most precise wavefunction-based approaches.<sup>82</sup> Although fundamentally being an interpolation, various MLIPs have been introduced that can be applied to chemical compounds that were not present in the training dataset.<sup>83–86</sup> There is also on-going work on developing generic reactive MLIPs, that can provide a reasonable description of transition states for a wide range of chemical reaction types.<sup>87</sup> The need for a large training set is also being addressed: a recently introduced equivariant graph neural network (EqGNN)



**Fig. 6:** Schematic representation of a typical MLIP architecture. The local environment of each atom ( $R_x$ ) is converted to a constant-size atomic feature vector ( $F_x$ ). The features can be either predefined or learned from the data (e.g. using graph neural networks). The energy of the molecular system ( $E_{\text{mol}}$ ) is then calculated as a sum of atomic energies ( $E_x$ ) calculated from the feature vectors by some ML model (usually a deep neural network or a Gaussian process regression model).

architecture<sup>88–90</sup> showed remarkable data efficiency, providing stable potentials with as few as hundreds of reference QM calculations.<sup>91</sup> These advancements suggest that ML potentials are on track to become a reliable alternative to QM methods. In their current state, ML potentials are still inferior to MM in terms of computational cost,<sup>82,92</sup> making it attractive to combine MLIPs with MM forcefields in a hybrid scheme similar to QM/MM.<sup>93–97</sup> Such “ML/MM” potentials would be particularly useful for biomolecular simulations, where the size of the system, a large variety of the atomic environments and the need for nanosecond-scale simulations makes the description of the entire system with an MLIP problematic. Currently, there are only a very few examples of application of ML/MM potentials to enzymatic reactions,<sup>94,98,99</sup> but given the rapid advances in this field we expect ML/MM to become a common approach to computational studies of enzymatic catalysis in the next few years.

The dramatic decrease of the computational cost offered by ML and ML/MM approaches opens intriguing possibilities. The complexity of biological systems offers an excellent opportunity for the development of accurate and efficient ML potentials to describe systems of high interest. One example of biochemical systems that are currently beyond the reach of computational chemistry are the ones that couple chemical reactivity with molecular motion, such as ATP synthase, myosin and topoisomerases to mention a few. Understanding the mechanism of action of these complex molecular machines requires both an accurate description of electronic rearrangements in the active site and a description of large-scale conformational changes, requiring extended sampling. The computational description of photophysical effects in biosystems, fundamentally quantum-mechanical in nature, will also be accelerated: MLIPs can provide both accurate description of the conformational ensemble and the electronic structure of biochromophores. This accuracy is essential to properly describe the interaction between living matter and light, including such phenomena as photosynthesis and vision. First glimpses of such symbiosis of ML and photobiology are already appearing.<sup>97,100</sup>

Finally, the availability of much cheaper reactive molecular potentials will allow routine inclusion of nuclear quantum effects (NQE) in the simulations of enzymatic systems. As discussed above, classical treatment of nuclear motion and reaction rates fails in cases where hydrogen atoms are involved in the reaction, which is the case of many enzymatic reactions. For instance, the classical picture completely ignores kinetic isotope effects (KIEs) – the change in the reaction rate upon isotopic substitution of an atom in the substrate. KIEs arise due to zero-point energy and tunnelling effect experienced by the nuclei. These effects can be described by employing path-integral molecular dynamics (PIMD)-based approaches. PIMD and related methods provide quantum statistical properties from a series of coupled simulations with nuclei described classically.<sup>101</sup> Currently, the high computational cost of QM/MM must be multiplied by the number of coupled simulations, making PIMD prohibitively expensive for routine applications to the study of enzymatic reactivity. The superior efficiency of MLIPs and ML/MM compared to QM-based approaches will pave the way for a full quantum treatment of the biochemical systems, leading to a deeper understanding of the importance of quantum phenomena in living organisms.

## Conclusions

The use of quantum mechanical methods in enzymatic catalysis has changed our understanding of biological reactivity. Since the excellent work developed by Karplus, Levitt, and Warshel, awarded with the 2013 Nobel Prize in Chemistry, QM/MM methods have become essential for studying reaction mechanisms in complex biological

systems. These approaches allow us to model bond-making and bond-breaking events from first principles, while also considering the complex dynamic and electrostatic effects of the surrounding protein environment. This has led to new insights in how enzymes achieve their remarkable catalytic efficiency, often far exceeding the rate of counterpart reactions in solution. Furthermore, QM/MM simulations provide a convenient platform to investigate quantum phenomena such as hydrogen tunnelling, which may play a key role in enhancing enzyme rates, particularly in proton- and hydride-transfer processes. These findings continue to guide the rational design of biomimetic and synthetic catalysts.

QM/MM methods are fundamental for the validation and refinement of enzyme designs, offering detailed insights into how active sites can be optimized and how reaction mechanisms develop within a biological context. These methods allow for a deeper understanding of the catalytic process, enabling the identification of potential inefficiencies and the suggestion of improvements to enhance the catalytic performance of designs. By focusing on these details, QM/MM simulations help prioritize those designs that have higher chances for experimental success, saving both time and resources in the development of new biocatalysts. However, the predictive capabilities of QM/MM remain limited by the complex interplay between enzyme structure, dynamics, and function. To design highly effective biocatalysts, it is essential to integrate quantum mechanical accuracy with a broader understanding of practical factors beyond the chemical step. This includes the stability and expression levels of the protein and its solubility, which are crucial for ensuring that the computationally optimized designs perform effectively at room conditions.

Accurate QM/MM calculations are still computationally demanding, especially when large systems or extensive conformational sampling are involved. One promising solution is the use of machine learning potentials (and their combination with MM models) trained on quantum mechanical data, which can greatly reduce computational cost while keeping chemical accuracy. If these models can be made transferable across related enzymes, they could help explore many more enzyme variants with a significant reduction in the computational effort. Looking ahead, continued progress in QM/MM and their combination with machine learning methods will be essential to better understand and engineer enzymes with desired functions. Enzyme catalysis is not only about the local chemistry at the active site, but also about how the entire protein environment shapes and supports the reaction. Capturing this complexity more accurately will bring us closer to designing enzymes as efficient and versatile as those evolved by nature.

**Acknowledgments:** The authors acknowledge financial support from grant PID2021-123332OB-C22 funded by MCIN/AEI/10.13039/501100011033/and by “ERDF A way of making Europe” and from grant PROMETEO CIPROM/2021/079 of Generalitat Valenciana. KZ acknowledges a Maria Zambrano fellowship from the Ministerio de Universidades (Spain).

**Research ethics:** Not applicable.

**Informed consent:** Not applicable.

**Author contributions:** All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

**Use of Large Language Models, AI and Machine Learning Tools:** NotebookLM was used to fix grammatical errors.

**Conflict of interest:** All other authors state no conflict of interest.

**Research funding:** The authors acknowledge financial support from grant PID2021-123332OB-C22 funded by MCIN/AEI/10.13039/501100011033/and by “ERDF A way of making Europe” and from grant PROMETEO CIPROM/2021/079 of Generalitat Valenciana. KZ acknowledges a Maria Zambrano fellowship from the Ministerio de Universidades (Spain).

**Data availability:** Not applicable.

## References

1. The Nobel Prize in Chemistry 2013 was Awarded Jointly to Martin Karplus, Michael Levitt and Arieh Warshel "For the Development of Multiscale Models for Complex Chemical Systems". <https://www.nobelprize.org/prizes/chemistry/2013/summary/>.
2. Warshel, A.; Karplus, M. Calculation of Ground and Excited State Potential Surfaces of Conjugated Molecules. I. Formulation and Parametrization. *J. Am. Chem. Soc.* **1972**, *94* (16), 5612–5625. <https://doi.org/10.1021/ja00771a014>.
3. Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103* (2), 227–249. [https://doi.org/10.1016/0022-2836\(76\)90311-9](https://doi.org/10.1016/0022-2836(76)90311-9).
4. Gao, J.; Truhlar, D. G. Quantum Mechanical Methods for Enzyme Kinetics. *Annu. Rev. Phys. Chem.* **2002**, *53* (1), 467–505; <https://doi.org/10.1146/annurev.physchem.53.091301.150114>.
5. Friesner, R. A.; Guallar, V. Ab Initio Quantum Chemical And Mixed Quantum Mechanics/Molecular Mechanics (QM/MM) Methods for Studying Enzymatic Catalysis. *Annu. Rev. Phys. Chem.* **2005**, *56* (1), 389–427; <https://doi.org/10.1146/annurev.physchem.55.091602.094410>.
6. Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew Chem Int Ed* **2009**, *48* (7), 1198–1229. <https://doi.org/10.1002/anie.200802019>.
7. Boulanger, E.; Harvey, J. N. QM/MM Methods for Free Energies and Photochemistry. *Curr. Opin. Struct. Biol.* **2018**, *49*, 72–76. <https://doi.org/10.1016/j.sbi.2018.01.003>.
8. Lodola, A.; De Vivo, M. The Increasing Role of QM/MM in Drug Discovery. *Adv. Protein Chem. Struct. Biol.* **2012**, *87*, 337–362. <https://doi.org/10.1016/B978-0-12-398312-1.00011-1>.
9. Van Der Kamp, M. W.; Mulholland, A. J. Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52* (16), 2708–2728. <https://doi.org/10.1021/bi400215w>.
10. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
11. Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; De Haas, R. J.; Bethel, N.; Leung, P. J. Y.; Huddy, T. F.; Pellock, S.; Tischer, D.; Chan, F.; Koepnick, B.; Nguyen, H.; Kang, A.; Sankaran, B.; Bera, A. K.; King, N. P.; Baker, D. Robust Deep Learning-based Protein Sequence Design Using ProteinMPNN. *Science* **2022**, *378* (6615), 49–56. <https://doi.org/10.1126/science.add2187>.
12. Kulik, H. J.; Luehr, N.; Ufimtsev, I. S.; Martinez, T. J. Ab Initio Quantum Chemistry for Protein Structures. *J. Phys. Chem. B* **2012**, *116* (41), 12501–12509. <https://doi.org/10.1021/jp307741u>.
13. Wang, T.; He, X.; Li, M.; Li, Y.; Bi, R.; Wang, Y.; Cheng, C.; Shen, X.; Meng, J.; Zhang, H.; Liu, H.; Wang, Z.; Li, S.; Shao, B.; Liu, T.-Y. Ab Initio Characterization of Protein Molecular Dynamics with AI2BMD. *Nature* **2024**, *635* (8040), 1019–1027. <https://doi.org/10.1038/s41586-024-08127-z>.
14. Wang, T.; He, X.; Li, M.; Li, Y.; Bi, R.; Wang, Y.; Cheng, C.; Shen, X.; Meng, J.; Zhang, H.; Liu, H.; Wang, Z.; Li, S.; Shao, B.; Liu, T.-Y. Author Correction: Ab Initio Characterization of Protein Molecular Dynamics with AI2BMD. *Nature* **2025**, *637* (8047), E25. <https://doi.org/10.1038/s41586-024-08556-w>.
15. Sheng, X.; Himo, F. The Quantum Chemical Cluster Approach in Biocatalysis. *Acc. Chem. Res.* **2023**, *56* (8), 938–947. <https://doi.org/10.1021/acs.accounts.2c00795>.
16. Himo, F.; De Visser, S. P. Status Report on the Quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *Commun. Chem.* **2022**, *5* (1), 29. <https://doi.org/10.1038/s42004-022-00642-2>.
17. Peters, B. Transition State Theory. In *Reaction Rate Theory and Rare Events Simulations*; Elsevier: Cambridge, MA, United States, 2017; pp 227–271.
18. Bowling, P. E.; Dasgupta, S.; Herbert, J. M. Eliminating Imaginary Vibrational Frequencies in Quantum-Chemical Cluster Models of Enzymatic Active Sites. *J. Chem. Inf. Model.* **2024**, *64* (9), 3912–3922. <https://doi.org/10.1021/acs.jcim.4c00221>.
19. Roca, M.; Upfold, C. M.; Williams, I. H. Critical Evaluation of Anharmonicity and Configurational Averaging in QM/MM Modelling of Equilibrium Isotope Effects. *Phys. Chem. Chem. Phys.* **2020**, *22* (28), 16267–16276. <https://doi.org/10.1039/D0CP01744B>.
20. Ahmadi, S.; Barrios Herrera, L.; Chehelamirani, M.; Hostaš, J.; Jalife, S.; Salahub, D. R. Multiscale Modeling of Enzymes: Qm-Cluster, QM/MM, and QM/MM/MD: a Tutorial Review. *Int. J. Quant. Chem.* **2018**, *118* (9), e25558. <https://doi.org/10.1002/qua.25558>.
21. Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11* (6), 700–733. <https://doi.org/10.1002/jcc.540110605>.
22. Clemente, C. M.; Capece, L.; Martí, M. A. Best Practices on QM/MM Simulations of Biological Systems. *J. Chem. Inf. Model.* **2023**, *63* (9), 2609–2627. <https://doi.org/10.1021/acs.jcim.2c01522>.
23. Bondanza, M.; Nottoli, M.; Cupellini, L.; Lipparini, F.; Mennucci, B. Polarizable Embedding QM/MM: the Future Gold Standard for Complex (Bio)Systems? *Phys. Chem. Chem. Phys.* **2020**, *22* (26), 14433–14448. <https://doi.org/10.1039/D0CP02119A>.



24. Nottoli, M.; Bondanza, M.; Mazzeo, P.; Cupellini, L.; Curutchet, C.; Loco, D.; Lagardère, L.; Piquemal, J.; Mennucci, B.; Lipparini, F. QM/AMOEBA Description of Properties and Dynamics of Embedded Molecules. *WIREs Comput. Mol. Sci.* **2023**, *13* (6), e1674. <https://doi.org/10.1002/wcms.1674>.
25. Soriano, A.; Silla, E.; Tuñón, I.; Ruiz-López, M. F. Dynamic and Electrostatic Effects in Enzymatic Processes. an Analysis of the Nucleophilic Substitution Reaction in Haloalkane Dehalogenase. *J. Am. Chem. Soc.* **2005**, *127* (6), 1946–1957. <https://doi.org/10.1021/ja046553h>.
26. Bunzel, H. A.; Kries, H.; Marchetti, L.; Zeymer, C.; Mittl, P. R. E.; Mulholland, A. J.; Hilvert, D. Emergence of a Negative Activation Heat Capacity During Evolution of a Designed Enzyme. *J. Am. Chem. Soc.* **2019**, *141* (30), 11745–11748. <https://doi.org/10.1021/jacs.9b02731>.
27. Åqvist, J. Computer Simulations Reveal an Entirely Entropic Activation Barrier for the Chemical Step in a Designer Enzyme. *ACS Catal.* **2022**, *12* (2), 1452–1460. <https://doi.org/10.1021/acscatal.1c05814>.
28. Ray, D.; Parrinello, M. Kinetics from Metadynamics: Principles, Applications, and Outlook. *J. Chem. Theory Comput.* **2023**, *19* (17), 5649–5670. <https://doi.org/10.1021/acs.jctc.3c00660>.
29. Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23* (2), 187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8).
30. Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: a Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.* **2004**, *120* (24), 11919–11929. <https://doi.org/10.1063/1.1755656>.
31. Bernardi, R. C.; Melo, M. C. R.; Schulten, K. Enhanced Sampling Techniques in Molecular Dynamics Simulations of Biological Systems. *Biochim. Biophys. Acta (BBA) – Gen. Subj.* **2015**, *1850* (5), 872–877. <https://doi.org/10.1016/j.bbagen.2014.10.019>.
32. Masgrau, L.; Truhlar, D. G. The Importance of Ensemble Averaging in Enzyme Kinetics. *Acc. Chem. Res.* **2015**, *48* (2), 431–438. <https://doi.org/10.1021/ar500319e>.
33. Bunzel, H. A.; Anderson, J. L. R.; Hilvert, D.; Arcus, V. L.; Van Der Kamp, M. W.; Mulholland, A. J. Evolution of Dynamical Networks Enhances Catalysis in a Designer Enzyme. *Nat. Chem.* **2021**, *13* (10), 1017–1022. <https://doi.org/10.1038/s41557-021-00763-6>.
34. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13* (8), 1011–1021. <https://doi.org/10.1002/jcc.540130812>.
35. Wu, H.; Mey, A. S. J. S.; Rosta, E.; Noé, F. Statistically Optimal Analysis of State-Discretized Trajectory Data from Multiple Thermodynamic States. *J. Chem. Phys.* **2014**, *141* (21), 214106. <https://doi.org/10.1063/1.4902240>.
36. Peters, B. *Reaction Rate Theory and Rare Events*; Elsevier: Amsterdam, Cambridge, MA, 2017.
37. Zinovjev, K.; Tuñón, I. Transition State Ensemble Optimization for Reactions of Arbitrary Complexity. *J. Chem. Phys.* **2015**, *143* (13), 134111. <https://doi.org/10.1063/1.4931596>.
38. Zinovjev, K.; Tuñón, I. Quantifying the Limits of Transition State Theory in Enzymatic Catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114* (47), 12390–12395. <https://doi.org/10.1073/pnas.1710820114>.
39. Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **2006**, *106* (8), 3210–3235. <https://doi.org/10.1021/cr0503106>.
40. Freiburger, M. I.; Guzovsky, A. B.; Wolynes, P. G.; Parra, R. G.; Ferreira, D. U. Local Frustration Around Enzyme Active Sites. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116* (10), 4037–4043. <https://doi.org/10.1073/pnas.1819859116>.
41. Olsson, M. H. M.; Warshel, A. Solute Solvent Dynamics and Energetics in Enzyme Catalysis: the  $S_N2$  Reaction of Dehalogenase as a General Benchmark. *J. Am. Chem. Soc.* **2004**, *126* (46), 15167–15179. <https://doi.org/10.1021/ja047151c>.
42. Roca, M.; Martí, S.; Andrés, J.; Moliner, V.; Tuñón, I.; Bertrán, J.; Williams, I. H. Theoretical Modeling of Enzyme Catalytic Power: Analysis of “Cratic” and Electrostatic Factors in Catechol *O*-Methyltransferase. *J. Am. Chem. Soc.* **2003**, *125* (25), 7726–7737. <https://doi.org/10.1021/ja0299497>.
43. Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. Kemp Elimination Catalysts by Computational Enzyme Design. *Nature* **2008**, *453* (7192), 190–195. <https://doi.org/10.1038/nature06879>.
44. Privett, H. K.; Kiss, G.; Lee, T. M.; Blomberg, R.; Chica, R. A.; Thomas, L. M.; Hilvert, D.; Houk, K. N.; Mayo, S. L. Iterative Approach to Computational Enzyme Design. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (10), 3790–3795. <https://doi.org/10.1073/pnas.1118082108>.
45. Blomberg, R.; Kries, H.; Pinkas, D. M.; Mittl, P. R. E.; Grütter, 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), 418–421. <https://doi.org/10.1038/nature12623>.
46. Świderek, K.; Tuñón, I.; Moliner, V.; Bertrán, J. Revealing the Origin of the Efficiency of the De Novo Designed Kemp Eliminase HG-3.17 by Comparison with the Former Developed HG-3. *Chem. A European J.* **2017**, *23* (31), 7582–7589. <https://doi.org/10.1002/chem.201700807>.
47. 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. <https://doi.org/10.1021/acscatal.7b03151>.
48. Welborn, V. V.; Ruiz Pestana, L.; Head-Gordon, T. Computational Optimization of Electric Fields for Better Catalysis Design. *Nat. Catal.* **2018**, *1* (9), 649–655. <https://doi.org/10.1038/s41929-018-0109-2>.
49. 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. *J. Chem. Phys.* **2023**, *158* (4), 044801. <https://doi.org/10.1063/5.0130886>.
50. Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J.; Field, M. J. A Hybrid Potential Reaction Path and Free Energy Study of the Chorismate Mutase Reaction. *J. Am. Chem. Soc.* **2001**, *123* (8), 1709–1712. <https://doi.org/10.1021/ja003522n>.
51. Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S.; Moliner, V. Hybrid QM/MM Potentials of Mean Force with Interpolated Corrections. *J. Phys. Chem. B* **2004**, *108* (24), 8427–8433. <https://doi.org/10.1021/jp049633g>.

52. Doron, D.; Major, D. T.; Kohen, A.; Thiel, W.; Wu, X. Hybrid Quantum and Classical Simulations of the Dihydrofolate Reductase Catalyzed Hydride Transfer Reaction on an Accurate Semi-Empirical Potential Energy Surface. *J. Chem. Theory Comput.* **2011**, 7 (10), 3420–3437. <https://doi.org/10.1021/ct2004808>.
53. Gaus, M.; Lu, X.; Elstner, M.; Cui, Q. Parameterization of DFTB3/3OB for Sulfur and Phosphorus for Chemical and Biological Applications. *J. Chem. Theory Comput.* **2014**, 10 (4), 1518–1537. <https://doi.org/10.1021/ct401002w>.
54. Ramos-Guzmán, C. A.; Ruiz-Pernía, J. J.; Tuñón, I. Unraveling the SARS-CoV-2 Main Protease Mechanism Using Multiscale Methods. *ACS Catal.* **2020**, 10 (21), 12544–12554. <https://doi.org/10.1021/acscatal.0c03420>.
55. Świderek, K.; Moliner, V. Revealing the Molecular Mechanisms of Proteolysis of SARS-CoV-2 M<sup>pro</sup> by QM/MM Computational Methods. *Chem. Sci.* **2020**, 11 (39), 10626–10630. <https://doi.org/10.1039/D0SC02823A>.
56. Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, 99 (20), 12562–12566. <https://doi.org/10.1073/pnas.202427399>.
57. Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: a Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, 100 (2), 020603. <https://doi.org/10.1103/PhysRevLett.100.020603>.
58. E, W.; Ren, W.; Vanden-Eijnden, E. String Method for the Study of Rare Events. *Phys. Rev. B* **2002**, 66 (5), 052301. <https://doi.org/10.1103/PhysRevB.66.052301>.
59. Zinovjev, K.; Ruiz-Pernía, J. J.; Tuñón, I. Toward an Automatic Determination of Enzymatic Reaction Mechanisms and their Activation Free Energies. *J. Chem. Theory Comput.* **2013**, 9 (8), 3740–3749. <https://doi.org/10.1021/ct400153r>.
60. Zinovjev, K.; Tuñón, I. Adaptive Finite Temperature String Method in Collective Variables. *J. Phys. Chem. A* **2017**, 121 (51), 9764–9772. <https://doi.org/10.1021/acs.jpca.7b10842>.
61. Giese, T. J.; Ekesan, Ş.; McCarthy, E.; Tao, Y.; York, D. M. Surface-Accelerated String Method for Locating Minimum Free Energy Paths. *J. Chem. Theory Comput.* **2024**, 20 (5), 2058–2073. <https://doi.org/10.1021/acs.jctc.3c01401>.
62. Bonomi, M.; Branduardi, D.; Gervasio, F. L.; Parrinello, M. The Unfolded Ensemble and Folding Mechanism of the C-Terminal GB1  $\beta$ -Hairpin. *J. Am. Chem. Soc.* **2008**, 130 (42), 13938–13944. <https://doi.org/10.1021/ja803652f>.
63. Zinovjev, K.; Tuñón, I. Exploring Chemical Reactivity of Complex Systems with Path-Based Coordinates: Role of the Distance Metric. *J. Comput. Chem.* **2014**, 35 (23), 1672–1681. <https://doi.org/10.1002/jcc.23673>.
64. Ruggiero, G. D.; Williams, I. H.; Roca, M.; Moliner, V.; Tuñón, I. QM/MM Determination of Kinetic Isotope Effects for COMT-Catalyzed Methyl Transfer does Not Support Compression Hypothesis. *J. Am. Chem. Soc.* **2004**, 126 (28), 8634–8635. <https://doi.org/10.1021/ja048055e>.
65. González-Lafont, À.; Lluch, J. M. Kinetic Isotope Effects in Chemical and Biochemical Reactions: Physical Basis and Theoretical Methods of Calculation. *WIREs Comput. Mol. Sci.* **2016**, 6 (5), 584–603. <https://doi.org/10.1002/wcms.1268>.
66. Mulholland, A. J. Chemical Accuracy in QM/MM Calculations on Enzyme-Catalysed Reactions. *Chem. Cent. J.* **2007**, 1 (1), 19. <https://doi.org/10.1186/1752-153X-1-19>.
67. Doron, D.; Stojković, V.; Gakhar, L.; Vardi-Kilshain, A.; Kohen, A.; Major, D. T. Free Energy Simulations of Active-Site Mutants of Dihydrofolate Reductase. *J. Phys. Chem. B* **2015**, 119 (3), 906–916. <https://doi.org/10.1021/jp5059963>.
68. Åqvist, J.; Brandsdal, B. O. Computer Simulations of the Temperature Dependence of Enzyme Reactions. *J. Chem. Theory Comput.* **2025**, 21 (3), 1017–1028. <https://doi.org/10.1021/acs.jctc.4c01733>.
69. Schramm, V. L.; Schwartz, S. D. Promoting Vibrations and the Function of Enzymes. Emerging Theoretical and Experimental Convergence. *Biochemistry* **2018**, 57 (24), 3299–3308. <https://doi.org/10.1021/acs.biochem.8b00201>.
70. Zinovjev, K.; Tuñón, I. Reaction Coordinates and Transition States in Enzymatic Catalysis. *WIREs Comput. Mol. Sci.* **2018**, 8 (1), e1329. <https://doi.org/10.1002/wcms.1329>.
71. Tuñón, I.; Laage, D.; Hynes, J. T. Are There Dynamical Effects in Enzyme Catalysis? Some Thoughts Concerning the Enzymatic Chemical Step. *Arch. Biochem. Biophys.* **2015**, 582, 42–55. <https://doi.org/10.1016/j.abb.2015.06.004>.
72. Peters, B. Transition-State Theory, Dynamics, and Narrow Time Scale Separation in the Rate-Promoting Vibrations Model of Enzyme Catalysis. *J. Chem. Theory Comput.* **2010**, 6 (5), 1447–1454. <https://doi.org/10.1021/ct100051a>.
73. Rosta, E.; Woodcock, H. L.; Brooks, B. R.; Hummer, G. Artificial Reaction Coordinate “Tunneling” in Free-Energy Calculations: the Catalytic Reaction of RNase H. *J. Comput. Chem.* **2009**, 30 (11), 1634–1641. <https://doi.org/10.1002/jcc.21312>.
74. Ruiz-Pernía, J. J.; Tuñón, I.; Moliner, V.; Hynes, J. T.; Roca, M. Dynamic Effects on Reaction Rates in a Michael Addition Catalyzed by Chalcone Isomerase. Beyond the Frozen Environment Approach. *J. Am. Chem. Soc.* **2008**, 130 (23), 7477–7488. <https://doi.org/10.1021/ja801156y>.
75. Grote, R. F.; Hynes, J. T. The Stable States Picture of Chemical Reactions. II. Rate Constants for Condensed and Gas Phase Reaction Models. *J. Chem. Phys.* **1980**, 73 (6), 2715–2732. <https://doi.org/10.1063/1.440485>.
76. Truhlar, D. G.; Gao, J.; Alhambra, C.; Garcia-Viloca, M.; Corchado, J.; Sánchez, M. L.; Villà, J. The Incorporation of Quantum Effects in Enzyme Kinetics Modeling. *Acc. Chem. Res.* **2002**, 35 (6), 341–349. <https://doi.org/10.1021/ar0100226>.
77. Agarwal, P. K.; Billeter, S. R.; Hammes-Schiffer, S. Nuclear Quantum Effects and Enzyme Dynamics in Dihydrofolate Reductase Catalysis. *J. Phys. Chem. B* **2002**, 106 (12), 3283–3293. <https://doi.org/10.1021/jp020190v>.
78. Major, D. T.; Gao, J. An Integrated Path Integral and Free-Energy Perturbation–Umbrella Sampling Method for Computing Kinetic Isotope Effects of Chemical Reactions in Solution and in Enzymes. *J. Chem. Theory Comput.* **2007**, 3 (3), 949–960. <https://doi.org/10.1021/ct600371k>.

79. Boekelheide, N.; Salomón-Ferrer, R.; Miller, T. F. Dynamics and Dissipation in Enzyme Catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (39), 16159–16163. <https://doi.org/10.1073/pnas.1106397108>.
80. García-Meseguer, R.; Duboué-Dijon, E.; Martí, S.; Ruiz-Pernía, J. J.; Laage, D.; Tuñón, I.; Hynes, J. T. Molecular Dynamics Simulations for Enzymatic Hydride-Transfer Reactions: Defining Environmental Reaction Coordinates to Capture Transition State Diversity. *J. Chem. Phys.* **2025**, *162* (12), 124118. <https://doi.org/10.1063/5.0254255>.
81. Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, R. J. Multireference Nature of Chemistry: the Coupled-Cluster View. *Chem. Rev.* **2012**, *112* (1), 182–243. <https://doi.org/10.1021/cr2001417>.
82. Jacobs, R.; Morgan, D.; Attarian, S.; Meng, J.; Shen, C.; Wu, Z.; Xie, C. Y.; Yang, J. H.; Artrith, N.; Blaiszik, B.; Ceder, G.; Choudhary, K.; Csanyi, G.; Cubuk, E. D.; Deng, B.; Drautz, R.; Fu, X.; Godwin, J.; Honavar, V.; Isayev, O.; Johansson, A.; Kozinsky, B.; Martiniani, S.; Ong, S. P.; Poltavsky, I.; Schmidt, K.; Takamoto, S.; Thompson, A. P.; Westermayr, J.; Wood, B. M. A Practical Guide to Machine Learning Interatomic Potentials – Status and Future. *Curr. Opin. Solid State and Mater. Sci.* **2025**, *35*, 101214. <https://doi.org/10.1016/j.cossms.2025.101214>.
83. Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: an Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost. *Chem. Sci.* **2017**, *8* (4), 3192–3203. <https://doi.org/10.1039/C6SC05720A>.
84. Yao, K.; Herr, J. E.; Toth, D. W.; Mckintyre, R.; Parkhill, J. The TensorMol-0.1 Model Chemistry: a Neural Network Augmented with Long-Range Physics. *Chem. Sci.* **2018**, *9* (8), 2261–2269. <https://doi.org/10.1039/C7SC04934J>.
85. Devereux, C.; Smith, J. S.; Huddleston, K. K.; Barros, K.; Zubatyuk, R.; Isayev, O.; Roitberg, A. E. Extending the Applicability of the ANI Deep Learning Molecular Potential to Sulfur and Halogens. *J. Chem. Theory Comput.* **2020**, *16* (7), 4192–4202. <https://doi.org/10.1021/acs.jctc.0c00121>.
86. Anstine, D.; Zubatyuk, R.; Isayev, O. AIMNet2: a Neural Network Potential to Meet your Neutral, Charged, Organic, and Elemental-Organic Needs. *Chemistry* **2024**, *20*. <https://doi.org/10.26434/chemrxiv-2023-296ch-v3>.
87. Zhang, S.; Makoś, M. Z.; Jadrich, R. B.; Kraka, E.; Barros, K.; Nebgen, B. T.; Tretiak, S.; Isayev, O.; Lubbers, N.; Messerly, R. A.; Smith, J. S. Exploring the Frontiers of Condensed-Phase Chemistry with a General Reactive Machine Learning Potential. *Nat. Chem.* **2024**, *16* (5), 727–734. <https://doi.org/10.1038/s41557-023-01427-3>.
88. Schütt, K. T.; Unke, O. T.; Gastegger, M. Equivariant Message Passing for the Prediction of Tensorial Properties and Molecular Spectra. *arXiv* **2021**. <https://doi.org/10.48550/ARXIV.2102.03150>.
89. Haghighatlari, M.; Li, J.; Guan, X.; Zhang, O.; Das, A.; Stein, C. J.; Heidar-Zadeh, F.; Liu, M.; Head-Gordon, M.; Bertels, L.; Hao, H.; Leven, I.; Head-Gordon, T. NewtonNet: a Newtonian Message Passing Network for Deep Learning of Interatomic Potentials and Forces. *Digit. Discov.* **2022**, *1* (3), 333–343. <https://doi.org/10.1039/D2DD00008C>.
90. Batatia, I.; Kovács, D. P.; Simm, G. N. C.; Ortner, C.; Csányi, G. MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields. *arXiv* **2022**. <https://doi.org/10.48550/ARXIV.2206.07697>.
91. Vitartas, V.; Zhang, H.; Juraskova, V.; Johnston-Wood, T.; Duarte, F. Active Learning Meets Metadynamics: Automated Workflow for Reactive Machine Learning Potentials. *Chemistry* **2025**, *21*. <https://doi.org/10.26434/chemrxiv-2024-twmlz-v2>.
92. Zuo, Y.; Chen, C.; Li, X.; Deng, Z.; Chen, Y.; Behler, J.; Csányi, G.; Shapeev, A. V.; Thompson, A. P.; Wood, M. A.; Ong, S. P. Performance and Cost Assessment of Machine Learning Interatomic Potentials. *J. Phys. Chem. A* **2020**, *124* (4), 731–745. <https://doi.org/10.1021/acs.jpca.9b08723>.
93. Zeng, J.; Giese, T. J.; Ekesan, Ş.; York, D. M. Development of Range-Corrected Deep Learning Potentials for Fast, Accurate Quantum Mechanical/Molecular Mechanical Simulations of Chemical Reactions in Solution. *J. Chem. Theory Comput.* **2021**, *17* (11), 6993–7009. <https://doi.org/10.1021/acs.jctc.1c00201>.
94. Pan, X.; Yang, J.; Van, R.; Epifanovsky, E.; Ho, J.; Huang, J.; Pu, J.; Mei, Y.; Nam, K.; Shao, Y. Machine-Learning-Assisted Free Energy Simulation of Solution-Phase and Enzyme Reactions. *J. Chem. Theory Comput.* **2021**, *17* (9), 5745–5758. <https://doi.org/10.1021/acs.jctc.1c00565>.
95. Lier, B.; Poliak, P.; Marquetand, P.; Westermayr, J.; Oostenbrink, C. BuRNN: Buffer Region Neural Network Approach for Polarizable-Embedding Neural Network/Molecular Mechanics Simulations. *J. Phys. Chem. Lett.* **2022**, *13* (17), 3812–3818. <https://doi.org/10.1021/acs.jpclett.2c00654>.
96. Zinovjev, K. Electrostatic Embedding of Machine Learning Potentials. *J. Chem. Theory Comput.* **2023**, *19* (6), 1888–1897. <https://doi.org/10.1021/acs.jctc.2c00914>.
97. Mazzeo, P.; Cignoni, E.; Arcidiacono, A.; Cupellini, L.; Mennucci, B. Electrostatic Embedding Machine Learning for Ground and Excited State Molecular Dynamics of Solvated Molecules. *Digit. Discov.* **2024**, *3* (12), 2560–2571. <https://doi.org/10.1039/D4DD00295D>.
98. Lei, Y.-K.; Yagi, K.; Sugita, Y. Efficient Training of Neural Network Potentials for Chemical and Enzymatic Reactions by Continual Learning. *J. Chem. Theory Comput.* **2025**, *21* (5), 2695–2711. <https://doi.org/10.1021/acs.jctc.4c01393>.
99. Sha, X.; Chen, Z.; Xie, D.; Zhou, Y. Modeling Enzyme Reaction and Mutation by Direct Machine Learning/Molecular Mechanics Simulations. *J. Chem. Theory Comput.* **2025**, *21* (9), 4335–4346. <https://doi.org/10.1021/acs.jctc.5c00149>.
100. Zinovjev, K.; Curutchet, C. Improved Description of Environment and Vibronic Effects with Electrostatically Embedded ML Potentials. *J. Phys. Chem. Lett.* **2025**, *16* (3), 774–781. <https://doi.org/10.1021/acs.jpclett.4c02949>.
101. Suleimanov, Y. V.; Aoiz, F. J.; Guo, H. Chemical Reaction Rate Coefficients from Ring Polymer Molecular Dynamics: Theory and Practical Applications. *J. Phys. Chem. A* **2016**, *120* (43), 8488–8502. <https://doi.org/10.1021/acs.jpca.6b07140>.