

ADVANCED REVIEW **OPEN ACCESS**

Advances in the Simulations of Enzyme Reactivity in the Dawn of the Artificial Intelligence Age

 Katarzyna Świderek¹  | Joan Bertran²  | Kirill Zinovjev³  | Iñaki Tuñón³  | Vicent Moliner¹ 

¹Biocomp Group, Institute of Advanced Materials (INAM), Universitat Jaume I, Castellón, Spain | ²Departament de Química, Universitat Autònoma de Barcelona, Bellaterra, Spain | ³Departament de Química Física, Universitat de València, Burjassot, Spain

Correspondence: Iñaki Tuñón (ignacio.tunon@uv.es) | Vicent Moliner (moliner@uji.es)

Received: 31 July 2024 | **Revised:** 9 December 2024 | **Accepted:** 10 December 2024

Associate Editor: Modesto Orozco | **Editor-in-Chief:** Peter R. Schreiner

Funding: Authors thank financial support from grants PGC2018-094852-B-C21 and PID2021-123332OB-C22 funded by MCIN/AEI/[10.13039/501100011033](https://doi.org/10.13039/501100011033)/ and by “ERDF A way of making Europe” and from grant PROMETEO CIPROM/2021/079 of Generalitat Valenciana. K.Ś. thanks the Spanish Ministerio de Ciencia, Innovación y Universidades (Grant PID2019-107098RJ-I00). K.Z. acknowledges a Maria Zambrano fellowship by Ministerio de Universidades (Spain).

Keywords: allosteric effects | conformational dynamics | machine learning techniques | QM/MM | reaction dynamics

ABSTRACT

The study of natural enzyme catalytic processes at a molecular level can provide essential information for a rational design of new enzymes, to be applied in more efficient and environmentally friendly industrial processes. The use of computational tools, combined with experimental techniques, is providing outstanding milestones in the last decades. However, apart from the complexity associated with the nature of these large and flexible biomolecular machines, the full enzyme catalyzed process involves different physical and chemical steps. Consequently, from the computational point of view, a deep understanding of every single step requires the selection of a proper computational technique to get reliable, robust and useful results. In this article, we summarize the different computational techniques and their use in the study of every single step of the catalytic process, including conformational diversity, allostery and those to study the chemical steps, as well as in the design of new enzymes. Because of the impact of artificial intelligence in all aspects of science during the last years, special attention has been applied to methods based on these techniques, their foundations and some selected recent applications.

1 | Introduction

Living organisms require the use of enzymes to accelerate their chemical processes, otherwise most of these processes would not take place at rates compatible with life. These biological catalysts are large, complex, and flexible molecular machines, usually proteins, that dramatically accelerate chemical reactions by comparison with their counterpart reactions in solution. The exceptional features of enzymes are not limited to their rate enhancements, which in many cases reach values of up to 10^{23} relative to the uncatalyzed reaction, but to the fact that they usually work in aqueous solution under mild conditions of pH, pressure, or temperature,

most of them are specific in substrate recognition and formation of products and they show limited or non-unwanted side products. These qualities of enzymes represent an advantage in comparison with homogeneous catalysts, thus explaining the interest in the use of biocatalysts in industrial applications. In addition, the use of biotechnological strategies opens the door to reducing the required reaction steps by using a one-pot procedure, allowing the decrease of energy consuming steps such as separation and purification of intermediates, with the obvious consequent economic savings. Consequently, the natural biochemical reactions in living organisms are a source of inspiration for designing new enzyme-based processes. In this design, wild-type enzymes must

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). *WIREs Computational Molecular Science* published by Wiley Periodicals LLC.

be modified for industrial applications, because new non-natural substrates or products are usually required, as well as specific technical conditions different from the standard physiological media in cells. Thus, a deep knowledge of the enzyme catalytic processes at a molecular level is essential for a rational design of new efficient enzymes. In this regard, computational methods are highly valuable tools to unveil the enzyme mechanism at molecular level. The results of theoretical studies, merged with the information derived from experimental techniques, have demonstrated to be a good strategy for the (re)design of new enzymes for future industrial applications. However, apart from the complexity of studying the chemical reaction itself, the full enzyme catalyzed process involves different physical and chemical steps. As shown in Figure 1, starting from the solvent-separated species, substrate (S) and enzyme (E), the full catalytic process requires the formation of a non-covalent reactant or Michaelis complex (E:S) that precedes the chemical step(s) where the substrate is transformed into products complex (E:P). Finally, the products release takes place to obtain the desired product in solution and the regeneration of the enzyme. Because of the flexibility of the enzyme, a myriad of different conformations appears in every state along the process. Consequently, from the computational point of view, a deep understanding of every single step requires the selection of a proper computational technique to get reliable and robust results.

The present review is organized as follows: In the next section a brief summary of the different methods employed to study enzyme reactivity is presented, containing a subsection devoted

to the possible potentials with special emphasis on the recent machine learning (ML) potentials, followed by a description of those methods designed to explore the different steps of the full enzyme catalyzed process; those methods to study conformational diversity and allostery and those to study the chemical steps. Then some selected applications are described in the following section, and a section of future perspectives is closing the review. It is important to point out that because of the flexibility of enzymes, together with the different nature of the steps along the enzyme catalyzed process, there is no a universal method to study all phenomena involved in enzyme catalysis. In addition, the complexity of every of these steps makes that all proposed computational techniques have inherent limitations. Major pros and cons of the different methods and techniques discussed in the present review will be highlighted along the text.

2 | Methods in Enzyme Reactivity

2.1 | Molecular Mechanics, Quantum Mechanics and Hybrid Potentials

The first step in performing computer simulations on enzymatic systems is selecting the potential to be used, a function returning the energy and gradients of the system for a given set of nuclear coordinates. Traditionally, this function can be based in first principles, as in the case of quantum mechanics (QM) potentials, fitted to some simple but physically-grounded

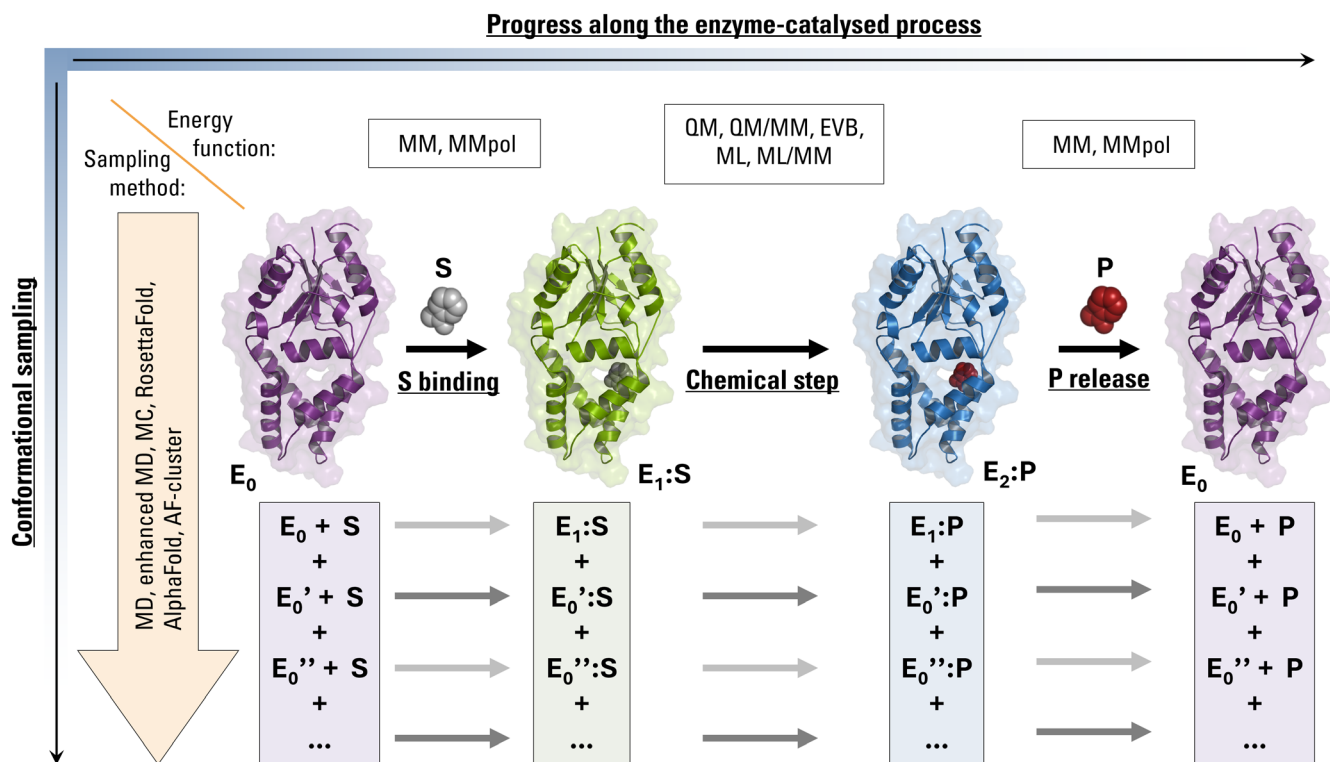


FIGURE 1 | Schematic representation of the different theoretical methods applied to study the different steps of an enzyme catalyzed process. The Y axis represents those methods required to carry out the sampling of the conformational space in any of the states appearing along the catalyzed process, molecular dynamics (MD) or Monte Carlo (MC). The x axis represents the progress along the catalyzed process, from the binding of the substrate (S), the chemical steps (E:S to E:P), and the product (P) release. Binding of the substrate and product release can be studied with molecular mechanics, MM, force fields, while quantum or hybrid quantum/classical potentials (QM, QM/MM) are required to explore the breaking and forming bonds during the chemical steps.

expressions, as for molecular mechanics (MM) force fields (FF), or a combination of both, hybrid (QM/MM) potentials, in which a small subset of atoms is described at a QM level, whereas the rest of the system is treated by a simpler (and faster) MM FF. Obviously, those steps of the full catalyzed process that involve forming or breaking bonds (the chemical steps) will require the use of QM potentials, while less accurate and computational demanding MM potentials can be employed to study the conformational diversity and allostery involving the substrate (or other ligand) binding steps, or the products release. As discussed in the next section, the boundaries between these approaches are being blurred with the use of ML potentials, where the fitting process achieves a new level of complexity, potentially matching first-principles calculations for any set of coordinates.

Simulations using *ab initio*, density functional theory (DFT) or semiempirical Hamiltonians provide accurate results of the chemical step(s) but they are computationally demanding for large systems. An initial option is reducing or truncating the full system up to some hundreds of atoms, fixing the position of some atoms on the surface of the resulting model to the original coordinates of the crystallographic data and applying a continuum model to mimic the rest of the missing environment. However, despite valuable information on reaction mechanisms of the chemical steps of enzyme catalyzed processes achieved by means of these so-called quantum chemical cluster approaches [1], these methods are limited to the exploration of potential energy surfaces for a reduced subset of atoms. An interesting alternative is the use of last generation semiempirical descriptions, such as the GFN2-xTB, which are designed for the fast calculation of the energy and its derivatives [2]. This method offers a reasonable alternative for the sampling of low-energy conformations in systems with roughly 1000 atoms. Automated conformational sampling of these medium-sized systems can be carried out with codes such as CREST (abbreviation for Conformer–Rotamer Ensemble Sampling Tool) [3]. In any case, any method based on the exploration of the potential energy surface can face limitations when trying to describe, for example, the effect of a mutation that causes extended conformational changes, as the exploration can be trapped around a particular subset of structures.

In general, running MD simulations with QM potentials, required to carry out adequate statistical simulations, is still extremely expensive for full enzymatic systems. Thus, it is generally accepted that the use of QM/MM potentials is the most adequate option in order to incorporate the effect of the full environment (protein and solvent molecules) in the study of the chemical reaction taking place in the active site of enzymes. These multiscale methods, where both subsystems are coupled through van der Waals and electrostatic interactions, can be basically classified between those that solve the electronic Hamiltonian under the field created by the environment, and methods based on valence bond descriptions, empirical valence bond methods [4], where the QM subsystem is described by means of two or more MM valence states, and mixed through an off-diagonal Hamiltonian parameterized term. Despite their general applicability QM/MM methods suffer also of some drawbacks derived from the arbitrary introduction of a boundary in the system that separates the QM and MM subsystems. Effects such as the exchange of molecules between either subsystems or charge transfer are

not easy to incorporate in traditional schemes, although there are some strategies that allow flexible definitions of the subsystems [5]. In addition, because of the amount of variables to take into account in the QM/MM schemes, the comparison between calculations carried out by different laboratories, which is one of the foundations of Science, is not an easy task.

While most of the simulations of biological systems employ traditional MM FF based on parameters (partial charges and van der Waals) derived from *ab initio* calculations on model compounds, recent advances have been oriented to improve their quality, or/and the speed of the simulations based on these kinds of FF, in the context of multiscale QM/MM simulations. Regarding the former target, efforts have been oriented to the development of polarizable FF, where the fixed partial charges assigned to the atoms can be modified by the effect of the changes in the QM wave function [6]. In this regard, it is worth mentioning the AMOEBA polarizable FF, which has been developed and successfully applied for the study of reactions in solution [7] and in enzymes [8]. Despite being much more computationally expensive than the traditional fixed charges FFs, advanced polarizable FFs have shown to be especially successful on protein-substrate binding processes studied within QM/MM calculations in highly charged active sites [6–9]. Efforts to develop faster FF have been based on the application of computationally efficient ML techniques. This is described in the following section because of the relevance and impact irrupted in the last decade in the scientific community.

2.2 | Machine Learning Potentials

In recent years, advances in ML techniques have led to their application to numerous aspects of protein science, from protein engineering and experiment design to atomistic computer simulations. Former applications have been reviewed recently and we invite interested readers to consult the existing literature [10]. Here we focus on the recent advances in the development of ML potentials (MLPs) to get universal and faster FF, capable of describing enzymatic reactivity.

MLPs are a class of supervised ML models that provide energy and forces for a given state of a molecular system. Modern MLPs are constructed in a way that the resulting function respects all the symmetries of the molecular potentials, that is they are invariant to translation, rotation, and inversion of the system as well as permutation of atoms of the same kind (hydrogens, carbons etc.). This is achieved by expressing the total energy of the system as a sum of atomic contributions with different “atomic energy” models being used for each chemical species. Each atomic contribution is calculated based on the immediate environment of each atom, expressed as a set of features. In contrast to conventional representations, such as Cartesian coordinates or Z-matrix, the feature vectors used in MLPs gave the same length independently of the composition of the environment. This idea, originally proposed by Behler and Parrinello [11], has important advantages. Apart from ensuring the above-mentioned invariances, it allows to develop models applicable to systems with different numbers of atoms and chemical composition, paving the way to the development of generic MLPs. An example of such generic models is the family of ANI potentials [12–15] that provide energies for ground

state geometries across a wide chemical space. The main culprit of MLPs is the reliance on extensive reference data set to train the model on, which can sometimes contain tens of thousands of single point QM calculations. As a potential solution to this issue, a combination of equivariant neural networks [16] and message passing graph neural networks [17] was recently shown to provide exceptional data efficiency when training MLPs. The MACE architecture proposed by Batatia et al. [18] when combined with active learning protocol [19] was shown to provide a stable MLP with as little as 45 reference structures for a simple S_N2 reaction and 565 reference structures for a glycosylation reaction.

While being dramatically faster than QM methods, MLPs are still orders of magnitude slower than the MM FFs. Moreover, for condensed phase simulations (most common for enzymatic processes) the reference QM data is expensive to obtain due to the high number of atoms in the simulations system. Therefore, as with QM/MM, significant performance gains are expected by treating most of the system with MM, while applying a more expensive MLP only in the region of the molecular system where MM does not provide an adequate description (such as the active site of an enzyme). Such ML/MM hybrid methods promise to provide simulations with precision close to DFT/MM MD while performance-wise being comparable to classical MM MD and, consequently, larger sampling can be carried out based on ML/MM MD simulations at a reasonable computer cost. ML/MM also benefits from a significantly smaller size of the ML region compared to pure ML simulations and therefore dramatically reduces the computational cost associated to dataset generation.

In contrast to QM/MM potentials, where the polarization of the QM part by the MM region is treated in the QM method of choice by introducing the corresponding electrostatic interaction term into the Hamiltonian, the description of the interaction between the ML and MM regions is not trivial. This conceptual drawback of ML/MM has been a topic of active research in the recent years. A straightforward approach to the development of ML/MM potentials is to train an ML model that directly reproduces the reference QM/MM energies based on both QM and MM configurations. One of the first studies that aimed to train an ML/MM potential based on QM/MM energies was presented by Bösel et al. [20] In this work, a feed forward neural network (FNN) potential was expanded to include features representing the MM environment. The authors demonstrated that directly learning the QM/MM energy requires obtaining an extensive dataset and that instead a Δ -learning correction (learning the difference between high and low levels of theory) provides significantly better results with much smaller training sets and smaller model architecture. Hofstetter et al. [21] applied the same approach to graph convolutional NNs (CNN) that reduces the number of parameters as well as computational cost compared with FNNs.

Another approach that aims to correct QM/MM potentials to a higher level of theory is deep potential—range correction (DPRc) proposed by Zeng et al. [22] DPRc focuses on improving the short-range interactions between QM and MM regions, while keeping the long-range ones treated at the low level of theory. Notably, the model was trained using an active learning approach that improved both the quality and transferability of the resulting model [13]. The increased computational efficiency provided by DPRc allowed to perform path integral MD simulations to describe kinetic

isotope effects in RNA cleavage reaction at a level of theory comparable to DFT/MM [23]. DPRc was recently added to Amber simulation package through integration with DeepMD Kit [24, 25].

Some studies focus on correcting only a specific aspect of cheap semiempirical methods when used in the context of QM/MM. Snyder et al. developed an ML-based correction based on Sparse Gaussian Process Regression for AM1/MM potential to reach the accuracy of DFT/MM in the resulting free energy profiles [26]. Remarkably, their correction depends exclusively on QM geometry, thereby correcting the average discrepancy between low and high QM/MM levels, instead of the potential energy itself. Kim et al. presented another simple correction scheme that focuses uniquely on improving the treatment of polarization of the QM region in AM1/MM by introducing chaperone atomic polarizabilities that account for underestimated polarization of the QM region provided by AM1 [27].

A drawback of ML/MM models that directly predict the QM/MM energy is that the model must be retrained whenever one decides to switch to another QM method or MM forcefield. This is different from QM/MM schemes that are agnostic to the specific choice of QM and MM potentials. A step toward reusable ML/MM models are approaches that treat the MM environment as an external field that polarizes the QM region. Gastegger et al. presented FieldSchNet, a model able to describe solvent effects on molecular spectra and reactions [28]. Instead of explicitly providing the MM atoms as the model input, the MM region is represented by electrostatic fields generated by MM point charges. A similar philosophy, but with a radically different model architecture was applied by Pan et al. [29] There, both electrostatic potential and electric fields of the MM environment were used as the model inputs to predict the energies and forces of the ML/MM system.

Recently, Lei et al. generalized these approaches by training a message-passing equivariant NN based on higher order expansion of MM potential over the QM atoms [30]. The quality of the resulting model is showcased by applying it to two systems: an $SN2$ reaction between methylfluoride and chlorine ion in solution and an enzymatic proton transfer. In both cases the model was able to recover both the geometry and energetics of the reference DFT/MM simulations. It should be noted that training of the model required a significant number of reference energy estimates (120,000 and 42,000 for the S_N2 reaction and the enzymatic proton transfer, respectively).

A fundamentally different approach is to separate the internal energy of the QM subsystem (obtained in the absence of an MM environment) from the interaction between the QM and MM subsystems. This allows training a separate model for the interaction term, resulting in an ML/MM scheme agnostic to the choice of the gas phase ML potential. In its simplest form, the interaction term can be calculated at the MM level of theory, that is, by assigning MM point charges to the ML atoms and obtaining the interaction energy as a sum of pairwise Coulomb and Lennard–Jones terms. This approach was employed by Galvelis et al. [31], allowing ML/MM MD simulation at up to 127 ns/day, which is the range typical for pure MM MD, but with the internal energy of the ML region provided by ANI-1x, a generic gas phase ML potential [13]. The disadvantage of such an approach is that the polarization of the ML region by MM point charges or conformational dependence of

the atomic charges of the ML region is neglected, preventing then its use for the study of chemical processes where the electronic distribution suffers important changes.

A more rigorous version of such “MM embedding” was presented by Song et al. [32] While their approach also neglects the polarization effects, the gas phase ML model provides not only the energy and forces but also the partial atomic charges. This allows to include the conformational dependence on the electronic structure of the ML region on the ML/MM interaction energy.

Finally, one of us proposed an electrostatic machine learning embedding (EMLE) approach [33, 34] that aims at recovering the full electrostatic embedding energy, including polarization of the ML region. EMLE scheme allowed to obtain a generic ML/MM model using only a relatively small (7000) training set of diverse compounds. This is achieved by combining charge equilibration [35] and minimal basis iterative stockholder (MBIS) partitioning scheme [36] to build a model of the electronic density and Thole polarizability model [37] to describe the response of the ML region to the external field generated by the MM subsystem. Also, it is worth mentioning that, while, in principle, EMLE is agnostic to the choice of the charge partitioning scheme, Grassano et al. recently showed that MBIS provides the best results in predicting the embedding energies [38].

All the above-mentioned techniques aim to recover the quality of state-of-the-art QM/MM potentials with high-level QM theory, usually with electrostatic embedding. Providing such “standard” QM/MM energies and forces at a reduced cost is definitely an approach worth pursuing. However, state-of-the-art QM/MM also has known drawbacks. For instance, it tends to overpolarize the QM region at the boundary [39], resulting in overestimated interaction energies between QM and MM subsystems. A way to overcome this problem while at the same time reducing the computational cost was proposed by Lier et al. [40] Their buffer region neural network (BuRNN) method introduces a buffer region that is subjected to full electronic polarization by the inner QM region. This buffer region is treated both quantum mechanically and classically, with interactions between the QM and buffer regions described by deep NNs. These NNs are trained to predict the difference between the QM calculations of the buffer region with and without the inner QM region. In contrast to the above-mentioned techniques that aim to recover the precision of state-of-the-art QM/MM, this approach minimizes overpolarization and other artifacts thus, in principle, providing a better description of the system overall. This is an example of how ML techniques can not only accelerate the multiscale potentials but also improve their quality.

2.3 | Methods to Study Conformational Diversity and Allostery

2.3.1 | Structure

The starting point for study chemical reactions catalyzed by enzymes is the structure of the protein scaffold. Traditionally, these are downloaded from the protein data bank (PDB), which contains ca. 250 K protein structures obtained mostly from x-ray

diffraction studies, but also from other techniques such as nuclear magnetic resonance (NMR) or Cryo-Electron Microscopy. However, this number is not large enough for the recent applications of protein design based on ML techniques. In addition, there are many enzymatic systems whose 3D structures have not been solved yet, despite their relevance for subsequent computational studies (i.e., computer design of enzymes capable of degrading synthetic polymers isolated from genome mining). Nowadays, the problem of determining the structure of a protein from its sequence is considered to be almost solved thanks to computational approaches such as AlphaFold [41, 42] or RoseTTAFold [43]. Using multiple-sequence alignment (MSA) and extensive structural data these methods are able to predict the most probable structure associated to a given sequence, reaching confidence levels close to experimental techniques. However, and despite the dramatically improved accuracy predicted by these ML based method (i.e., see the results of the Critical Assessment of Protein Structure Prediction competitions), AI models been trained on x-ray structures and then they tend to provide as solution the most stable configuration under crystallization conditions, which is not necessarily the active conformation of the protein [44]. The simplified view of proteins as a single active conformation, represented by the x-ray structure, responsible for a given function in its activity cycle has been already surpassed by a more complex representation, where protein dynamics play an essential role (see Figure 1). Proteins can open/close their active site, suffer conformational changes to accommodate the substrate and/or respond to changes in the experiments, and some of these conformational changes are key ingredients for the evolution of the enzyme along its catalytic cycle. In this regard, AF-Cluster is a novel method that clusters multiple-sequence alignments (MSAs) by sequence similarity to enable AlphaFold2 to predict multiple conformational states of proteins [45]. Unlike default AlphaFold predictions that are often limited to a single structure, AF-cluster successfully predicts alternative conformations for metamorphic proteins, such as KaiB, experimentally validated by nuclear magnetic resonance. Additionally, AF-Cluster can be used to predict mutations that alter the conformational state and to identify alternative states in other protein families. This approach broadens the scope of structural prediction by elucidating protein energy landscapes and alternative functional states, which is crucial for understanding biological processes.

2.3.2 | Free Energy Landscape

Proteins conformational changes can be seen as trajectories between basins that appear in a highly multidimensional free energy landscape (FEL) that results from the interplay between intra and intermolecular forces, as schematically shown in Figure 2. The depth of the wells on this FEL correlates with the probability of finding the protein in the different conformations, while the height separating these wells is linked to the time scale of the transitions among them. The complexity of this FEL was first described for myoglobin, where a hierarchical structure of states and substates that can perform different functions or the same function but with different rates was proposed [46]. The biological function of a protein depends on multiple substates [47], that appear along the catalytic cycle of the enzyme, which must bind the substrate, catalyze one or more chemical steps, and

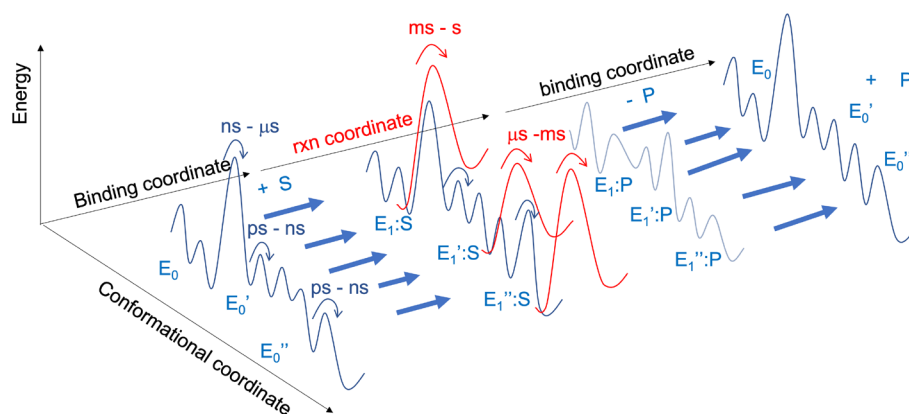


FIGURE 2 | Free energy landscape of the full catalytic process involving conformational changes (conformational coordinate) together with binding and chemical transformations (binding and reaction coordinate). Each of the states appearing along the process is populated by myriads of different conformations separated by local conformational changes happening in many cases in shorter timescales (ps–ns), but there are slower conformational changes that require (ns–μs). Changes between different chemical states can proceed in a larger timescale (from μs to s) than the conformational ones.

finally release the products. To perform all these actions, the enzyme must explore different conformations, which can be favored or not depending on the interactions established between the protein and the substrate in its different forms (reactants, intermediates, products) and also by other factors such as the presence of additional ligands (including allosteric ligands), pH, temperature, etc. The identification of those factors controlling the population of the most reactive enzymatic conformations is a very active research field, as briefly discussed here.

The substrate binding step involves a diffusion process of the substrate from the solvent to the enzyme active site (see Figure 1). However, because of the flexible nature of proteins, this process can be preceded by or coupled with significant conformational changes of the protein ranging from local (involving mostly sidechain changes) to large conformational changes (at the protein domain scale). In fact, debates on how this particular step of the catalytic process takes place are associated with the proposals of the induced-fit mechanism or the conformational selection mechanism, also known as the conformational shift or population-shift model (as recently summarized in a review by Nam et al. [48]). In other words, whether the enzyme undergoes a conformational change upon binding of its substrate, rearranging its active site to accommodate the substrate within a new conformation (induced fit), or the apoenzyme can adopt multiple pre-existing conformations, one of them being the most favorable one to bind the substrate and then to proceed to the following chemical steps (conformational shift). Further discussions on the induced-fit mechanism or the conformational shift mechanism are presented in section 3.1. Deciphering which of the two alternative mechanisms is employed by enzymes requires the exploration of the FEL, which must be obtained from the use of appropriate computational methods combined with information derived from experimental techniques. In addition to active site effects, the substrate or any other small molecule can bind at a distal site from the active site (the allosteric site), shifting the conformational ensemble of the protein and as a consequence modulating the activity of the active site (the orthosteric site). This is known as allosteric regulation [49], a term introduced in the sixties by Monod and Jacob [50]. Changes in

the activation energies or the diffusion coefficients in the prefactor of the catalyzed reaction rate have been observed to occur upon this coupling or long-range cooperativity [51], but dynamic allostery has also been identified as responsible for the adaptation of enzymes to environmental niches [52]. Moreover, from a broader perspective, allosteric effects can also be induced by distal single-point mutations or covalent attachments, which are being used by strategies for re-designing enzymes (see Section 3.2). Finally, the relevance of conformational dynamics can also involve the dynamics of protein loops (opening or closure, see section 3.1), which can be required for the binding of the substrate, the product release, modulating the activity of the chemical reaction and turnover [53].

Original attempts to study the E:S binding step were based on the microscopic linear response approximation (LRA) and related approaches [54]. These include the linear interaction energy method (LIE) [55], where the nonelectrostatic term is estimated by scaling the average van der Waals interactions, or the less accurate semi-macroscopic methods such as the protein dipoles Langevin dipoles (PDL/D/S) in its LRA form (PDL/D/S-LRA), the molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA) or the molecular mechanics/generalized Born surface area (MM/GBSA) [56]. Despite being popular methods, the accuracy of these predictions is not high, apart from lacking the possibility of a detailed description of the conformational space of the apoenzyme and the Michaelis complex. Alternative approaches such as the alchemical “double decoupling methods” [57] or the “pathway methods”, appeared as more promising. The former, also called the double annihilation method [58], is based on thermodynamic cycles in which the interaction of the ligand with its surroundings is progressively switched off in both the active site and the bulk solution [59]. This is done through a series of nonphysical (alchemical) overlapping states using thermodynamic integration (TI) or free energy perturbation (FEP) procedures. Nowadays, these methods provide relatively accurate predictions of relative binding energies of similar substrates to a common enzyme active site. In the “pathway methods” the ligand is physically pulled away from the binding site

and then the binding free energy can be computed as a potential of mean force (PMF). There are some studies where these methods are reviewed and compared [60]. In practice, a critical issue with either alchemical or pathway methods is to achieve sufficient sampling to improve the statistical convergence in calculations of absolute binding free energy of the substrate to the enzyme. This can be achieved by means of different approaches such as MD or MC simulations. However, the exploration of the FEL for the full conformational space of the protein and/or protein: substrate complex, which provides crucial information on the thermodynamics and the kinetics for the protein conformational transitions [61], may require the study of rare events, where free energy barriers higher than the thermal energy can prevent the adequate exploration of the full molecular configurational space by unbiased MD simulations. Thus, energy barriers separating different minima will determine the time scale of the conformational transformation (Figure 2): fast conformational changes of the protein occur in the picosecond to nano-second time scales, whereas transitions between states separated by higher energy barriers (i.e., reactants state to products states, or some larger scale conformational changes such as loop motions and domain rearrangements, as discussed below) take place in microseconds to seconds. Consequently, the time scales of these processes are often beyond the reach of straightforward MD simulations. To overcome this problem, a variety of enhanced sampling algorithms have been developed, which can be classified as either unbiased methods or biased methods. In the former, an external bias potential is applied along a chosen set of meaningful coordinates, commonly referred to as collective variables (CVs), along which a biasing potential is applied to facilitate barrier-crossing events. The underlying FEL can then be recovered by reweighting the sampled configurations by keeping track of the introduced energy. The free-energy difference between two states is then obtained from a probability histogram along a distinguished reaction coordinate. Examples of biased methods are metadynamics (MTD) or any of its refined variants [62, 63], steered MD [64], gaussian accelerated MD [65], replica exchange molecular dynamics (REMD) [66], parallel tempering [67], or umbrella sampling (US) [68]. In contrast, in unbiased MD algorithms, such as transition path sampling (TPS) [69], an ensemble of short reactive trajectories connecting a predefined initial and final state is collected without modifying the energy function and without prior knowledge of the transition state (TS) region. In its simplest form, a path sampling simulation samples transitions from one state A to another state B by performing MC importance sampling in path space. Initially, a TPS simulation requires defining the stable states and providing an initial trajectory connecting these stable states. When two states are well-defined, TPS provides mechanistic insights into the transition, which can be used in biased methods to compute free-energy profiles [70]. Examples of other unbiased approaches include transition interface sampling (TIS) [71], or the memoryless transition networks often referred to as Markov state model (MSM) [72].

Computational methods to locate allosteric sites and to design allosteric modulators have been recently reviewed in this journal [73, 74]. The transmission of the allosteric signal to the orthosteric site has been usually analyzed in terms of preexisting

allosteric pathways, encoded in the protein structure and dynamics. There are several computational strategies to find these allosteric pathways based on the analysis of correlated motions among different residues during an MD simulation of the enzyme. The Eigenvector centrality (EC) method uses the adjacency matrix, defined from the correlation coefficients between two nodes of the protein as obtained from MD simulations [75]. An example of methods developed to redesign enzymes based on these ideas is the Shortest Path Map (SPM) tool of Osuna and co-workers, which has been successfully applied for enhancing the, for instance, the stand-alone activity of tryptophan synthases [76], and the esterase activity of hydroxynitrile lyases [77]. The method relies on the construction of a graph based on the computed mean distances and correlation values obtained from MD simulations, combined with ancestral sequence reconstruction to reduce the number of possible positions as well as to provide specific amino-acid substitutions.

Amaro and co-workers presented a dynamical network analysis tool, weighted implementation of suboptimal paths (WISP), that based on aligned molecular dynamics trajectory and the generation of a dynamic interdependence correlation matrix, is able to calculate optimal and suboptimal communication pathways between residues [78]. The method, which is of interest for studying the molecular mechanisms of allostery, has been successfully applied to describe the allostery of an amidotransferase. More recently, they have developed an open-source scientific software package, NetSci, that leverages GPU acceleration and a k-nearest-neighbor algorithm that, by means of improved correlation metrics, and it allows to estimate of the mutual information (MI) between data in a set of several orders of magnitude faster than CPU-based implementations [79]. NetSci is shown to be a useful program for the estimation of MI and the generalized correlation (GC) within general datasets, the analysis of intraprotein communication and for understanding the allosteric mechanisms.

Kamerlin and co-workers have developed computational tools based on non-covalent interaction networks, such as the key interactions finder (KIF) [80] and key interaction networks (KIN), that are able to extract insight from static structures and dynamic conformational ensembles of protein, have provided useful information of enzyme mechanistic regulation that can be applied to protein engineering. These tools, KIF and KIN, generate a set of prioritized residues and interactions as targets for experimental protein engineering, based on the analysis of structures derived from experiments (i.e., PDB structures) or from MD simulation for evolutionary groups of proteins, yielding conserved networks of static or dynamic interactions.

Dokholyan and co-workers develop a computationally efficient network-based method, Ohm, to identify and characterize allosteric communication networks within proteins, relying solely on the structure of the protein of interest [81]. Ohm automatically determines allosteric network architecture and identifies critical coupled residues correlating well with nuclear magnetic resonance CHEMical Shift Covariance Analysis (NMR CHESCA) result.

Dynamical nonequilibrium MD simulations (D-NEMD) can be also used to propagate perturbations in the active site and

locate allosteric pathways and sites through inspection of atomic fluctuations [82]. It is important to stress that the existence of allosteric pathways does not imply a mechanism for the transmission of some strain energy between distant sites, such as an energy stroke, something unlikely to be possible between distant sites in a solvated structure where energy dissipation can happen in short time scales [83]. Instead, these pathways unveil the coupling between degrees of freedom of neighbor residues that can extend in large networks due to the complexity of the potential energy function governing protein dynamics. These pathways then show a mechanistic basis for the alteration of the protein FEL that can be reflected in local changes in sites apparently distant from those where the perturbation was acting.

The potential applications of methods that identify allosteric pathways can be boosted by the combination with ML methods that predict protein structure. These methods, such as AlphaFold [41, 42] or RoseTTAFold [43], may be used to generate multiple conformations, which can be then used as initial structures to explore allosteric pathways. These ML methods can be seen as a shortcut for the exploration of the protein FEL, allowing better identification of those states or substates participating in protein function and the possible communication pathways existing at each conformational state [45]. Osuna et al. used a combination of AlphaFold 2 and the SPM method to identify conformationally relevant positions more efficiently than just using enhanced MD simulations [84]. At this point, it is worth mentioning the existence of bioinformatic tools, such as CAVER [85] or AQUA-DUCT [86], that can complement the studies based on MD simulations to carry out protein analysis on tunnels, pores, channels, pockets and cavities. The results of these analysis can facilitate protein engineering (or drug design) by providing information on the characteristics of transportation pathways and internal binding cavities, which can be relevant on the identification of positions that can be subjected to mutagenesis.

ML methods can also be trained with the huge amount of data extracted from simulations of conformational changes and used to reduce the complexity of the problem, extracting the principal CVs governing such processes. The application of ML techniques to describe conformational changes in biomolecules has shown substantial advances in recent times [87–90]. These techniques excel at identifying essential degrees of freedom and reaction coordinates from high-dimensional molecular simulation data, enabling more accurate and efficient characterizations of conformational dynamics. The advantage of ML reaction coordinates is their ability to capture highly nonlinear relationships and handle noisy data. Additionally, the integration of explainable artificial intelligence methods [87, 91], enhances the interpretability of these models by identifying and ranking critical features driving conformational transitions. Nevertheless, while these methods have achieved notable success in simpler systems, extending their applicability to more complex biomolecular systems remains an ongoing challenge.

2.4 | Methods to Study Reactivity

The generation of the FEL of the full enzyme catalyzed process requires the exploration of the evolution of the relevant

chemical coordinates of the protein and substrate (catalytic coordinate) as well as the conformational space of the system along every single step (conformational coordinate), as schematically summarized in Figure 2. In this regard, it is important to point out that the energy barrier that separates the different conformational states in the apo and the substrate: enzyme forms (either in the binding, the reactant complex, the product complex and the product release) are usually lower than those of the chemical steps. This means that rare events with free energy barriers much higher than the thermal energy are more frequent in the chemical steps than in the previously discussed binding step. Thus, the examples of the biased and unbiased methods mentioned in the previous section are also used in the exploration of chemical reactivity, probably being the US, MTD, and TPS the most popular ones. Regarding the biased methods, most of the selected CVs selected to apply the bias are in this case related to the breaking and forming bonds, that is, distances or coordination numbers. Then, if the reaction coordinate can be properly described with a limited number of coordinates the method renders a realistic free energy surface. However, the degrees of freedom associated with the protein can be equally relevant during the reaction and must be included in the set of CVs used to drive the chemical step to get reliable results, making the problem of generating the associated FEL quite challenging. In such cases, the reaction can be explored by means of free energy perturbation (FEP) methods, which implies the sampling of the environment (usually the MM region in QM/MM simulations) along a previously traced potential energy path (i.e., IRC) generated from a TS located at the QM/MM level. The method is not only cheaper (which means that higher level Hamiltonians can be used to describe the QM sub-set of atoms) but, keeping in mind that the sampling is performed along the IRC, the free energy profile is obtained along a more realistic reaction coordinate. However, since there is no sampling on the chemical system, the result could be biased by the fact that just one TS structure is used [92].

Finally, as an alternative for sampling the chemical reaction steps over a large number of degrees of freedom, chain-of-states methods, such as nudged elastic band [93], or the string method [94], can be used to limit the sampling on the minimum energy or free energy paths (MEP or MFEP), thus dramatically reducing the computational cost of the simulations. As recently stated by York and co-workers [95], many variations of the string method have been developed during the past decades to tackle different particular problems related, not only to the reactivity but to the substrate binding and conformational changes of the protein.

Among variants of the string method, the adaptive string method (ASM) [96] was particularly successful in the determination of enzymatic reaction mechanisms [97–100]. The assumption behind ASM (as well as the majority of chain-of-states methods) is that a complex process can be described by monitoring changes in only a handful of local degrees of freedom, such as interatomic distances or dihedral angles. For instance, in the case of enzymatic reactivity, the reaction coordinate is typically defined in terms of the lengths of the chemical bonds broken or formed during the reaction. The success of ASM indicates that this reductionist assumption is valid for the processes mentioned above.

3 | Advances in Applications

3.1 | Conformational Changes and Reactivity: From Allostery to Biological Motors

As discussed above, the FEL determines the statistical outcome of protein dynamics, but the FEL itself can be also seen as a dynamic entity that can be modified by external stimulus: thermodynamic forces that alter the protein ensemble. Changes in the temperature, the addition of an active site substrate or of an allosteric ligand, and alterations of the pH or salt concentration, among others, can induce variations in the FEL, and then in the depths of conformational minima and the heights of the barriers separating them. In principle, as mentioned in Section 2.3, the changes in the FEL could be classified according to the consequences on the topology of the FEL. Thus, these can be associated with the relative populations of conformational minima already existing on the original FEL (conformational shift), the emergence of new minima (induced fit) or just a modification of the free energy curvature along conformational coordinates (landscape reshaping), as shown in Figure 3. We will now discuss recent applications to the study of conformational changes in enzymes using this simplified classification.

As discussed in the next section, distal or allosteric mutations may have an impact on the conformational ensemble of the active site and then on the enzymatic catalytic properties, an effect that can be rationalized using Figure 3. Mulholland and coworkers used one of the aforementioned allosteric pathway methods, SPM, to study the changes in the allosteric networks at the TS ensemble of different variants of Kemp eliminases (KE) generated by directed evolution [101]. These enzymes catalyze the conversion of benzisoxazoles into salicylonitriles, see Figure 4. The improved enzyme design, which included up to six mutations, showed an increased network of correlated motions centered at the active site and also reduced energetic and structural fluctuations at the TS, suggesting a tighter interaction and then a better stabilization of the TS in the evolved variant. This KE enzyme has also been used to discuss the effects of temperature on the enzymatic FEL. Eyring fits of the enzymatic second-order rate constant display curvature, which has been attributed to a negative activation heat capacity [102]. According to the authors, the origin of the reduced heat capacity at the TS ensemble with respect to the ground state would be a reduction of the energy fluctuations at the TS. In other words, the FEL along conformational coordinates orthogonal to the reaction coordinate (see Figure 2) would be narrower at the TS than at the ground state, because enzyme-substrate interactions are

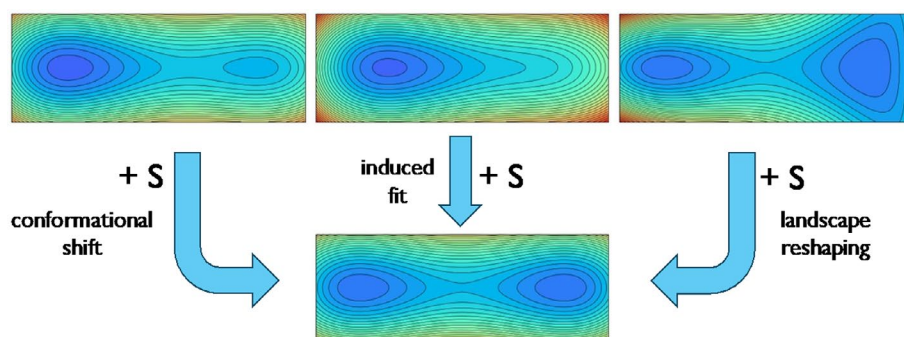


FIGURE 3 | Different types of changes on the Free Energy Landscape (FEL) along conformational coordinates of the enzyme due to ligand binding (here denoted with + S) or any other physical or chemical transformation. Changes along substrate conformational coordinates are ignored here. In a conformational shift the relative depth of two free energy wells is altered, modulating then the population of the two conformations, one of them ready for the binding of the substrate. In an induced fit a new well, corresponding to a new conformation, appears on the (FEL). In the case of landscape reshaping the depth of the two wells is not necessarily altered, just the width of the well along some coordinates, affecting thus the fluctuations of the system in that state. A combination of two or more of these effects can appear in the FEL of an enzyme when the system is altered.

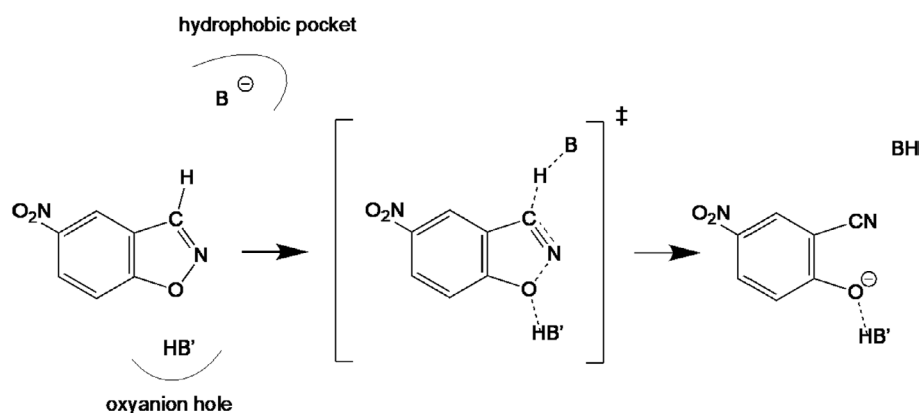


FIGURE 4 | Kemp elimination reaction.

tightened when the substrate reaches the TS charge distribution. These improved interactions would then result in a landscape reshaping, as shown in Figure 3. However, another explanation for the observed curvature has been proposed. Åqvist and co-workers suggested an alternative kinetic model to explain the observation of curvature in the Eyring plot, suggesting that instead of a negative activation heat capacity, the key ingredient is an equilibrium between ground state conformations [103]. This equilibrium would be highly temperature dependent, reflecting a conformational shift between ground state conformations (the relaxed and the reactive complexes, being the latter uphill in free energy). This different interpretation, which can be recast in terms of the changes appearing on the enzymatic FEL, has ignited an interesting controversy [104, 105]. More recently, Mulholland and coworkers have shown that the two interpretations can be complementary and that the emergence of new enzyme-substrate stable conformations advanced along the reaction coordinate (Transition Like Conformations) can be accompanied by changes in the heat capacity [106], demonstrating the cooperativity between conformational changes and TS stabilization: conformational changes that tighten enzyme-TS interactions are favored as the system approaches this point of the reaction path. In other words, changes in the FEL along the direction of the chemical change must also be reflected in other directions, narrowing the FEL and thus decreasing the heat capacity, as illustrated in Figure 3. More recently, Åqvist et al. have argued that while heat capacity changes can be expected during the binding process, changes during the chemical step would not be large enough to cause a noticeable variation in the system's heat capacity [107]. The reason would be that while during the binding process, some molecules appear/disappear in the active site, during the chemical reaction only the electronic density is redistributed, while the same set of atoms is kept.

Obviously, not only allosteric effects can change the protein FEL. The binding of a substrate into the active site can certainly alter the FEL of the enzyme. A well-known example is provided by flexible loops that frequently decorate active sites and that can suffer dramatic conformational changes upon binding. Loop closure/opening motions are one of the most recurrent conformational changes observed in the catalytic cycles of enzymes. The binding of the substrate modifies the FEL of these loops inducing a conformational shift from open to closed conformations (see Figure 2). Loop motions may favor binding, modulate substrate selectivity, improve active site preorganization for the chemical step and may also be needed for product release [53, 108–110]. L-Lactate Dehydrogenase (LDH) is one of the enzymatic systems employed to discuss the role of the opening and closure of loops. It was originally assumed that the full catalytic process starts with the binding of the cofactor and the substrate followed by the enclosure of the active site by a mobile loop of the protein before the reaction takes place. Using the x-ray structures of LDH with open and loop conformation [111], we explored the chemical step within the open and closed loop conformations. The reactivity measured for the different protein conformations, after quantizing the vibrations and adding the contributions of recrossing and tunneling effects, was in agreement with the range of rate constants measured in single-molecule spectrometry studies [112]. The presence of different reactive conformations agreed with some of the pioneering QM/MM studies by one of us [113]. Our calculations suggested that the closure of the

active site is required for the inverse process, the oxidation of lactate to pyruvate.

The critical role of loops dynamics in enzyme activity has increased its interest as possible objectives in enzymatic design [114]. In this sense, the methods mentioned in the previous section to identify allosteric pathways can be useful to identify residues that regulate loop dynamics and then to identify possible mutational hotspots for loop engineering. This strategy, together with enhanced sampling simulations, has been used in the case of two Protein Tyrosine Phosphatases, the human PTP1B and the *Yernisia* PTP YopH, a paradigmatic example where the closure of the WPD-loop is a requirement to place the reacting fragments in close contact with the substrate as it contains one of the catalytic residues [110]. Kamerlin, Hengge and coworkers prepared several chimeric enzymes with point mutations along the WPD-loop, demonstrating that mutations along the loop can alter the equilibrium between the open and closed states affecting the enzymatic activity of the variants [115, 116]. Recent mechanistic studies on the WPD-loop motion have shown that the conformational change between the open and closed states is clearly activated, this is there are two different conformational minima separated by a significant free energy barrier that explains that the loop motion takes place in the millisecond/s timescale [117, 118]. These studies have stressed that the conformational coordinate that determines the barrier between conformations is a combination of backbone torsions and not the mere displacement of the loop. The FEL of the loop along displacement coordinates is mostly diffusive, with small barriers associated with the breaking of intermolecular interactions, while most of the activation free energy barrier that determines the experimental rate constants for the loop conformational change is associated with backbone torsions. For both PTP1B and YoPH, the torsional change is located in four central residues of the WPD-loop, while the rest of the backbone is strongly coupled, influencing the dynamics of conformational transitions [118]. These mechanistic studies pave the way for the rational modification of the loop's free energy landscape, identifying those factors that determine the equilibrium between open and closed forms and the rate of the transitions. As a recent example with implications in enzymes design, and in connection with other studies mentioned above, it has been recently shown that altering the active site loop dynamics can enhance standalone activity of the Tryptophan Synthase Alpha Subunit (TrpA). The SPM method was here used to suggest four mutations that stabilize the catalytically active closed conformation, improving the catalytic activity of TrpA [119].

Another example of FEL modification coupled to the chemical step is provided by protein motors. Protein motors are essential constituents for cellular functions like intracellular transport or the replication of genetic material. These proteins are enzymes that use the chemical energy, mostly coming from the exergonic ATP hydrolysis, to perform directed motion [120]. As in the case of allosteric changes, the functioning of a protein motor should not be viewed as a direct transmission of some excess energy produced during the chemical reaction to some mobile parts of the protein. Two main mechanisms have been proposed to explain the behavior of biological motors: power strokes and Brownian ratchets. In the power stroke proposal, the energy of the excess energy of the chemical reaction drives a conformational change in the protein,

while in a Brownian ratchet mechanism, the conformational change occurs in conditions of thermal equilibrium, driven by thermal forces. The conditions in which protein motors work, surrounded by a dense and randomized environment, seem more compatible with the Brownian mechanism [66]. Recently, some of us have shown that a conformational shift mechanism coupled to ATP hydrolysis could be the basis for the translocation of Zika helicase along a nucleic acid strand [121]. This enzyme contains two active sites, one to bind the nucleic acid strand and another to bind ATP and a magnesium ion, both sites connected through flexible motifs. The FEL associated with one of these motifs, motif V, changes substantially when ATP is transformed into ADP and an inorganic phosphate is released from the active site. The loss of a hydrogen bond established between this motif and the γ -phosphate group of ATP changes the conformational preference of the motif, favoring a conformation that is substantially higher in free energy when ATP is present. As in the case of the opening and closing of active site loops, the conformational change can be described in terms of the backbone torsional angles. The presence of a γ -phosphate sensor, a protein residue able to form hydrogen bond interactions with this phosphate group when present in the active site, could be a general mechanism to change the conformational preference of a flexible region of a protein, which in turn can modulate the interactions between the protein and the nucleic acid, facilitating their relative movement [122].

3.2 | Advances in Enzymatic Design

As mentioned, natural biochemical reactions in living organisms are a source of inspiration for designing new enzyme-based processes. The different computer-assisted engineering techniques that have been proposed are intrinsically related to the interpretations of how wild-type enzymes work and what is the origin of their catalytic activity. Thus, electrostatic interactions have been traditionally recognized as the most important contribution to catalysis [123]. The complex tridimensional structure of the enzyme generates a highly preorganized environment perfectly arranged for efficiently stabilizing the TS of the reaction by electrostatic interactions more efficiently than water does when the reaction takes place in an aqueous solution [123–125]. Computational studies from different groups support this proposal [126]. However, because of the inherent flexibility and dimensions of enzymes, a myriad of different conformations can be adopted by the protein, not all of them adequate for substrate binding and/or product release [61]. This connects with the other main interpretation of the enzyme activity and, in turn, to the other group of approaches to design new enzymes, suggesting that they are not only stabilizing the TS by electrostatic interactions but enzymes also show a perfect capability to modulate the possible conformational ensembles by their inherent conformationally dynamic nature [127, 128].

3.2.1 | Electrostatic Effects

As recently stated in a review by some of us [124], the relevance of electrostatic effects in enzyme catalysis, demonstrated by

many groups, including ourselves, and more specifically in the relative stabilization of the TS with respect to the reactants state (reflected in a reduction of the activation free energy), justify the efforts made to obtain improved enzymatic designs based on this premise.

Based on the results of a series of experimental and computational studies, the analysis of the electrostatic potential and the electric field generated by the protein allowed us to illustrate how these concepts work in different enzymes and how they can be used to guide protein design in cases where the electronic density at the TS differs significantly from that of the reactants [124]. However, some of the difficulties associated with the use of electrostatics to guide the selection of mutations on a protein scaffold to improve or generate certain catalytic activity are the long-range nature of these interactions and their dependence on protein dynamics. This indirect electrostatic effect can be more difficult to reproduce by current simulation techniques because it requires the correct sampling of the conformational spaces available to the enzyme variants. Despite these difficulties, the usefulness of electrostatic concepts must serve as an incentive for the development of new methodological approaches for their incorporation in the protein design process.

The KE reaction mentioned in the previous section, see Figure 4, is one of the most popular benchmarks for engineering strategies based on this kind of foundation, basically because the reaction implies a proton transfer from a carbon atom to a heteroatom thus involving a dramatic electron density reorganization. Several remarkable achievements have been published since the first *de novo* design by Houk, Tawfik, Baker, and co-workers [129], including HG3, HG-3 and especially HG-3.17, designed by Hilvert and coworkers, that showed an efficiency close to natural enzymes [130]. A deep analysis of the KE reaction catalyzed by these three different variants was carried out in our laboratory based on MD simulations with hybrid QM/MM potentials [131]. Our studies showed that the high reactivity of HG-3.17 was related to a proper electrostatic preorganization of the environment that creates a favorable electrostatic potential for the reaction to proceed, especially on the oxygen leaving group that is negatively charged as the reaction proceeds. Similar to the conclusions derived from other studies on different enzymes [59, 92, 124, 132], a linear correlation between the activation free energies and the electrostatic potential generated by the protein in the leaving group was obtained [131].

Head-Gordon and co-workers arrived at similar conclusions on the relevance of electrostatic preorganization in the efficient designs of previous KE enzymes and suggested that efficient computationally designed enzymes could be achieved with minimal experimental intervention using electric field optimization as guidance [133]. In this case, authors predicted (and later tested experimentally) four mutations that enhanced the electric fields and chemical positioning of the substrate to stabilize the TS.

Another example of the use of the electrostatic effects to guide the design of enzymes has been applied by some of us to transform an esterase into an amidase [134]. We proposed a strategy

based on the combination of the best electrostatic properties of different promiscuous enzymes with activity on a common reaction. The first step of the protocol was to study the reaction mechanism of two promiscuous esterases; *Candida antarctica* lipase B (CALB) and the esterase from *Bacillus subtilis* (Bs2), by means of US methods with QM/MM MD simulations. Once the amidase mechanisms were established, an analysis of the electrostatic properties of the chemical system along the reaction showed that a positive charge was developed in the active site histidine residue in the rate-limiting TS. Consequently, this TS could be stabilized by generating a negative electrostatic potential on this residue through site-directed mutagenesis. For this purpose, first, the results of evolutive analysis of the protein geometries using the CNN deep learning approach allowed the conclusion that Bs2 would be a more robust protein scaffold to perform mutations in order to improve the, in this case, amidase activity without dramatically perturbing the structure of the protein [135]. The alignment of the structures of both systems allowed for identifying key residues with a different impact on the amidase reaction in the two enzymes. Residues of Bs2 with an unfavorable effect on catalysis were substituted by those at equivalent spatial positions in CALB that exert a favorable effect. Our *in silico* method predicted an, a priori, increase of 4 orders of magnitude on the measurable rate constant after a single distant mutation. However, deeper computational insights revealed a significant shift in the pKa of the inserted aspartate, resulting in a more modest catalytic effect and showing the relevance of proper assignment of the protonation state of titratable residues in setting up the models. While a variant with much better activity was not obtained, there was excellent agreement between the predicted activation-free energies, the electrostatic potential generated by the enzyme in the identified site, and the experimentally determined rate constants. Supported by the good agreement between theoretical and experimental results, the presented computer-guided rational design can become a valuable component in the toolkit of enzyme redesign and engineering, upon optimization [134].

3.2.2 | Conformational Selection

The challenge of designing new enzymes based on conformational selection criteria lies in identifying mutations that enhance the population of specific conformations, especially if located in distal sites from the active site cavity [53, 76, 127]. As mentioned, the differences in protein conformations can involve not only relatively small changes of residues side-chains but more dramatic differences encompassing allosteric networks that regulate loop motion. The development of computational techniques to monitor these dynamics and identifying possible sites for putative mutations that regulate the enzyme activity has suffered significant progress in the last decade, despite the number of successful productions is still scarce.

Directed evolution methods have shown that the catalytic activity of an enzyme can often be enhanced by distal mutations [127, 136]. Some mutations taking place far from the active site can provoke an allostery effect on the enzymatic FEL, inducing a conformational change in the active site that favors the population of those conformational states that are able to catalyze more efficiently a target chemical reaction. This conformational shift

(see Figure 3) caused by mutations was shown to be operative in the case of the retro-aldolases, enzymes that catalyze a complex chemical process with several steps [137]. The low activity of the initial computational designs was improved by several rounds of directed evolution, with the most evolved variant presenting catalytic activity close to natural aldolases [138, 139]. Long MD simulations showed that the mutations introduced by directed evolution were able to stabilize the catalytically active conformations, thus modifying the enzymatic FEL with a direct impact on its catalytic activity [76]. The identification of allosteric pathways can be also used to locate mutational hotspots with the potential to alter the conformation properties of the active site. In this way, methods that identify allosteric pathways can be useful to guide the rational redesign of enzymes or to rationalize the effects of directed evolution. MD simulations and the SPM method were applied to analyze the effect of the mutations introduced by directed evolution to convert the inefficient when isolated Tryptophan Synthase B subunit (TrpB) into an efficient standalone enzyme [140]. SPM was able to assign most of the mutations that appeared during directed evolution to positions that are relevant to change the population of conformational states that can be stabilized by the allosteric partner (subunit A) or through mutations in the standalone version [141]. This strategy was then used to design new stand-alone TrpB variants that were subsequently validated experimentally [76].

Among other recent significant contributions, the relevance of the conformational changes on the enzyme activity profile has been experimentally supported by several studies. Recently, Thompson, Chica, and co-workers described a computational design method for increasing the catalytic efficiency of de novo KE [129, 130] relying using crystallographically enhanced conformational sampling [142]. Thus, using structural ensembles generated from dynamics-based refinement against X-ray diffraction data collected from crystals of KE enzymes to identify productive substrates for efficient catalysis, authors designed new enzymes displaying k_{cat}/K_M values improved by 100–250-fold, without relying on directed evolution and high-throughput screening.

Finally, protein engineering strategies have been also oriented to shift the optimum temperature of enzymes, which can be a technical advantage in industrial applications. As an example, Åqvist and co-workers recently predicted a computational redesign of the psychrophilic α -amylase from an Antarctic bacterium by selecting a set of mutations that stabilize the enzyme-substrate interaction based on MD/EVB simulations of the catalytic reaction at different temperatures [143]. Analysis of the computational results, verified by experiments, revealed that the critical surface loop controlling the temperature dependence approaches the target conformation observed in a mesophilic ortholog.

4 | Future Perspectives

During the last decades, molecular simulations of enzymatic systems have been successfully used to complement structural and kinetic studies, providing a framework to understand the origin of their amazing catalytic properties. Nowadays, this can be rationalized based on the combination of two a priori contradictory hypotheses: the electrostatic preorganization of the

active site and the dynamic conformational selection of the protein scaffold of the enzyme. The electrostatic stabilization of the TS of the chemical step(s) is achieved after or simultaneously with the optimisation of the conformational ensemble along the full catalytic process that favor the population of the most reactive conformations [144, 145].

The maturity reached by computational simulations is demonstrated by the fact that these tools can be used not only to understand the origin of the catalytic features of enzymes and to rationalize the consequences of mutations on different enzymatic properties (such as their stability, kinetics, selectivity, etc) but also to predict which could be the effect of a particular mutation. The success in this area will represent the systematic use of computational tools to design new improved enzymes for a wide range of applications.

Despite the undeniable achievements to date, there is still one dimension of enzymatic chemistry in which computational methods have yet to progress: complexity. Enzymes are complex chemical entities, containing many atoms and combining simultaneously structure and flexibility. This complexity is reflected in the diversity of the chemistry that can happen in their active sites, which is in turn modulated by the large variety of interactions and conformations that can come into play during the catalytic cycle. ML techniques offer now a way to assist with this complexity in computational simulations.

ML has already revolutionized the analysis of complex biological data. By training models on existing large datasets of protein structures and dynamics, one cannot only predict the predominant structure of biomolecules but also some conformational changes, identify active sites, and propose novel mutations for enhanced functionality. ML's ability to learn from massive datasets can uncover patterns and relationships beyond traditional computational methods. In this sense, ML methods are used to analyze the outcome of molecular simulations and to uncover which coordinates are adequate to represent protein conformational spaces and to drive changes between conformations. In the future, ML algorithmic developments can be expected to improve the selection of collective variables to be used as order parameters or as reaction coordinates in complex conformational changes. These methods should be combined with detailed first principles analysis that determine which conformational changes are activated or controlled by diffusion, and which types of coordinates contribute to the transitions. Obviously, proper computational methods must be selected depending on the nature of the particular step under study. Moreover, the separation between the physical and chemically controlled steps is sometimes not clear as the formation of the non-covalent reactive enzyme: substrate complex can involve not diffusion of the substrate to the active site and protein conformational changes but a significant advance in the chemical reaction coordinate toward the TS of the chemical step. Thus, our computational tools must be capable of exploring events taking place along multiple temporal and spatial scales, thus requiring computational methods capable of integrating different levels of theory ranging from classical to quantum potentials.

Nowadays, we are just starting to see the possibilities of ML methods to generate potentials able to describe enzymatic

chemical reactions at a moderate computational cost but with the accuracy of high-level quantum methods. The diversity of approaches to the development of ML/MM potentials and the relatively limited number of practical applications to enzymatic reactions underscore the early stage of this research area. The need for model retraining, the low transferability of MLPs to other systems, and the lack of availability in standard MD codes currently impede the adoption of ML/MM methods by the wider computational enzymology community. Nevertheless, thanks to rapid ongoing advances along these research lines, one can expect ML/MM to become the method of choice for simulations of enzymatic catalysis in the next few years. High precision and significantly lower computational cost of MLPs compared with DFT and ab initio methods will lead to a dramatic increase in the amount of sampling feasible with modern hardware. This will open exciting opportunities, such as studies of phenomena at longer timescales (e.g., coupling between chemical and conformational changes), precise virtual screening of protein-ligand complexes for drug design applications, and systematic prediction of mutations to improve enzyme activity and stability to mention a few.

Finally, modern computational tools, including AI-based techniques and fundamental principles of enzyme design, are having a large impact not only from a scientific perspective but are also becoming increasingly significant in industrial applications, such as the production of animal feed, beverages, detergents, food products, pharmaceuticals, and textiles. Application of computer-based techniques can dramatically reduce the experimental cost and provide new innovative and efficient enzyme catalyzed pathways to get high value chemicals that otherwise would require traditional non-environmentally friendly synthetic methods. Moreover, ML techniques can assist in data pre-processing, classification and prediction, cluster analysis and mining multimedia, for scientific programming in industrial applications [146, 147]. Obviously, a broad understanding of the advantages and limitations of the different techniques is required for the correct choice of the specific computational technique.

Author Contributions

Katarzyna Świderek: writing – original draft (equal), writing – review and editing (equal). **Joan Bertran:** writing – original draft (equal), writing – review and editing (equal). **Kirill Zinovjev:** writing – original draft (equal), writing – review and editing (equal). **Iñaki Tuñón:** conceptualization (equal), supervision (equal), writing – original draft (equal), writing – review and editing (equal). **Vicent Moliner:** conceptualization (equal), project administration (equal), supervision (equal), writing – original draft (equal), writing – review and editing (equal).

Acknowledgments

Authors thank financial support from grants PGC2018-094852-B-C21 and 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. K.Ś. thanks the Spanish Ministerio de Ciencia, Innovación y Universidades (Grant PID2019-107098RJ-I00). K.Z. acknowledges a Maria Zambrano fellowship by Ministerio de Universidades (Spain).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Related WIREs Articles

[The challenge of predicting distal active site mutations in computational enzyme design](#)

[Rational design of allosteric modulators: Challenges and successes](#)

[Enhanced sampling strategies for molecular simulation of DNA](#)

References

1. X. Sheng and F. Himo, "The Quantum Chemical Cluster Approach in Biocatalysis," *Accounts of Chemical Research* 56 (2023): 938–947.
2. C. Bannwarth, S. Ehlert, and S. Grimme, "GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method With Multipole Electrostatics and Density-Dependent Dispersion Contributions," *Journal of Chemical Theory and Computation* 15 (2019): 1652–1671.
3. P. Pracht, S. Grimme, C. Bannwarth, et al., "CREST—A Program for the Exploration of Low-Energy Molecular Chemical Space," *Journal of Chemical Physics* 160 (2024): 114110.
4. (a) A. Warshel and M. Levitt, "Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme," *Journal of Molecular Biology* 103 (1976): 227–249. (b) J. Åqvist and A. Warshel, "Simulation of Enzyme Reactions Using Valence Bond Force Fields and Other Hybrid Quantum/Classical Approaches," *Chemical Reviews* 93 (1993): 2523–2544.
5. (a) H. C. Watanabe and Q. Cui, "Quantitative Analysis of QM/MM Boundary Artifacts and Correction in Adaptive QM/MM Simulations," *Journal of Chemical Theory and Computation* 15 (2019): 3917–3928. (b) B. Kirchhoff, E. O. Jónsson, A. O. Dohn, T. Jacob, and H. Jónsson, "Elastic Collision Based Dynamic Partitioning Scheme for Hybrid Simulations," *Journal of Chemical Theory and Computation* 17 (2021): 5863–5875.
6. J. Nochebuena, S. Naseem-Khan, and G. A. Cisneros, "Development and Application of Quantum Mechanics/Molecular Mechanics Methods With Advanced Polarizable Potential," *WIREs Computational Molecular Science* 11 (2020): e1515.
7. S. Chatterjee, J. Nochebuena, and G. A. Cisneros, "Impact of an Ionic Liquid Solution on Horseradish Peroxidase Activity," *Journal of the American Chemical Society* 146 (2024): 13247–13257.
8. D. Loco, L. Lagardere, G. A. Cisneros, et al., "Towards Large Scale Hybrid QM/MM Dynamics of Complex Systems With Advanced Point Dipole Polarizable Embeddings," *Chemical Science* 10 (2019): 7200–7211.
9. J. W. Ponder, C. Wu, P. Ren, et al., "Current Status of the AMOEBA Polarizable Force Field," *Journal of Physical Chemistry B* 114 (2010): 2549–2564.
10. J. Yang, F. Z. Li, and F. H. Arnold, "Opportunities and Challenges for Machine Learning-Assisted Enzyme Engineering," *ACS Central Science* 10 (2024): 226–241.
11. J. Behler and M. Parrinello, "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces," *Physical Review Letters* 98 (2007): 146401.
12. J. S. Smith, O. Isayev, and A. E. Roitber, "ANI-1: An Extensible Neural Network Potential With DFT Accuracy at Force Field Computational Cost," *Chemical Science* 8 (2017): 3192–3203.
13. J. S. Smith, B. Nebgen, N. Lubbers, O. Isayev, and A. E. Roitber, "Less Is More: Sampling Chemical Space With Active Learning," *Journal of Chemical Physics* 148 (2018): 241733.
14. J. S. Smith, B. T. Nebgen, R. Zubatyuk, et al., "Approaching Coupled Cluster Accuracy With a General-Purpose Neural Network Potential Through Transfer Learning," *Nature Communications* 10 (2019): 2903.
15. C. Devereux, J. S. Smith, K. K. Huddleston, et al., "Extending the Applicability of the ANI Deep Learning Molecular Potential to Sulfur and Halogens," *Journal of Chemical Theory and Computation* 16 (2020): 4192–4202.
16. M. Geiger and T. Smidt, "e3nn: Euclidean Neural Networks." 10.48550/arXiv.2207.09453.
17. F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The Graph Neural Network Model," *IEEE Transactions on Neural Networks* 20 (2009): 61–80.
18. I. Batatia, D. P. Kovács, G. N. C. Simm, C. Ortner, and G. Csányi, "MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields," arXiv:2206.07697, 10.48550/arXiv.2206.07697.
19. V. Vitartas, H. Zhang, V. Juraskova, T. Johnston-Wood, and F. Duarte, "Active Learning Meets Metadynamics: Automated Workflow for Reactive Machine Learning Potentials," <https://doi.org/10.26434/chemrxiv-2024-twmlz>.
20. L. Bösel, M. Thürlmann, and S. Riniker, "Machine Learning in QM/MM Molecular Dynamics Simulations of Condensed-Phase Systems," *Journal of Chemical Theory and Computation* 17 (2020): 2641–2658.
21. A. Hofseter, L. Bösel, and S. Riniker, "Graph-Convolutional Neural Networks for (QM)ML/MM Molecular Dynamics Simulations," *Physical Chemistry Chemical Physics* 24 (2022): 22497–22512.
22. J. Zeng, T. J. Giese, S. Ekesan, and D. M. York, "Development of Range-Corrected Deep Learning Potentials for Fast, Accurate Quantum Mechanical/Molecular Mechanical Simulations of Chemical Reactions in Solution," *Journal of Chemical Theory and Computation* 17 (2021): 6993–7009.
23. T. J. Giese, J. Zeng, S. Ekesan, and D. M. York, "Combined QM/MM, Machine Learning Path Integral Approach to Compute Free Energy Profiles and Kinetic Isotope Effects in RNA Cleavage Reactions," *Journal of Chemical Theory and Computation* 18 (2022): 4304–4317.
24. T. J. Giese, J. Zeng, L. Lerew, et al., "Software Infrastructure for Next-Generation QM/MM-ΔMFP Force Fields," *Journal of Physical Chemistry B* 128 (2024): 6257–6271.
25. Y. Tao, T. J. Giese, S. Ekesan, et al., "Amber Free Energy Tools: Interoperable Software for Free Energy Simulations Using Generalized Quantum Mechanical/Molecular Mechanical and Machine Learning Potentials," *Journal of Chemical Physics* 160 (2024): 224104.
26. R. Snyder, B. Kim, X. Pan, Y. Shao, and J. Pu, "Bridging Semiempirical and Ab Initio QM/MM Potentials by Gaussian Process Regression and Its Sparse Variants for Free Energy Simulation," *Journal of Chemical Physics* 159 (2021): 54107.
27. B. Kim, Y. Shao, and J. Pu, "Doubly Polarized QM/MM With Machine Learning Chaperone Polarizability," *Journal of Chemical Theory and Computation* 17 (2023): 7682–7695.
28. M. Gastegger, K. T. Schütt, and K.-R. Müller, "Machine Learning of Solvent Effects on Molecular Spectra and Reactions," *Chemical Science* 12 (2021): 11473–11483.
29. X. Pan, J. Yang, R. Van, et al., "Machine-Learning-Assisted Free Energy Simulation of Solution-Phase and Enzyme Reactions," *Journal of Chemical Theory and Computation* 17 (2021): 75745–75758.
30. Y.-K. Lei, K. Yagi, and Y. Sugita, "Learning QM/MM Potential Using Equivariant Multiscale Model," *Journal of Chemical Physics* 160 (2024): 214109.
31. R. Galvelis, A. Valera-Rial, S. Doerr, et al., "NNP/MM: Accelerating Molecular Dynamics Simulations With Machine Learning Potentials

- and Molecular Mechanics,” *Journal of Chemical Information and Modeling* 63 (2023): 5701–5708.
32. K. Song, S. Käser, K. Töpfer, L. I. Vazquez-Salazar, and M. Meuwly, “PhysNet Meets CHARMM: A Framework for Routine Machine Learning/Molecular Mechanics Simulations,” *Journal of Chemical Physics* 159 (2023): 24125.
33. K. Zinovjev, “Electrostatic Embedding of Machine Learning Potentials,” *Journal of Chemical Theory and Computation* 19 (2023): 1888–1897.
34. K. Zinovjev, L. Hedges, R. Montagud Andreu, C. Woods, I. Tuñón, and M. van der Kamp, “Emle-Engine: A Flexible Electrostatic Machine Learning Embedding Package for Multiscale Molecular Dynamics Simulations,” *Journal of Chemical Theory and Computation* 20 (2024): 4514–4522.
35. A. K. Rappé, “Goddard III WA Charge Equilibration for Molecular Dynamics Simulations,” *Journal of Physical Chemistry* 95 (1991): 3358–3363.
36. T. Verstraelen, S. Vandenbrande, F. Heidar-Zadeh, et al., “Minimal Basis Iterative Stockholder: Atoms in Molecules for Force-Field Development,” *Journal of Chemical Theory and Computation* 12 (2016): 3894–3912.
37. B. T. Thole, “Molecular Polarizabilities Calculated With a Modified Dipole Interaction,” *Chemical Physics* 59 (1981): 341–350.
38. J. S. Grassano, I. Pickering, A. E. Roitberg, M. C. González Lebrero, D. A. Estrin, and J. A. Semelak, “Assessment of Embedding Schemes in a Hybrid Machine Learning/Classical Potentials (ML/MM) Approach,” *Journal of Chemical Information and Modeling* 64 (2024): 4047–4058.
39. H. M. Senn and W. Thiel, “QM/MM Methods for Biomolecular Systems,” *Angewandte Chemie International Edition* 48 (2009): 1198–1229.
40. B. Lier, P. Poliak, P. Marquetand, J. Westermayr, and C. Oostenbrink, “BuRNN: Buffer Region Neural Network Approach for Polarizable-Embedding Neural Network/Molecular Mechanics Simulations,” *Journal of Physical Chemistry Letters* 13 (2022): 3812–3818.
41. J. Jumper, R. Evans, A. Pritzel, et al., “Highly Accurate Protein Structure Prediction With AlphaFold,” *Nature* 596 (2021): 583–590.
42. J. Abramson, J. Adler, J. Dunger, et al., “Accurate Structure Prediction of Biomolecular Interactions With AlphaFold 3,” *Nature* 630 (2024): 493–500.
43. M. Baek, F. DiMaio, I. Anishchenko, et al., “Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network,” *Science* 373 (2021): 871–876.
44. S. Reardon, “Five Protein-Design Questions That Still Challenge AI,” *Nature* 635 (2024): 246–248.
45. H. K. Wayment-Steele, A. Ojoawo, R. Otten, et al., “Predicting Multiple Conformations via Sequence Clustering and AlphaFold2,” *Nature* 625 (2024): 832–839.
46. A. Ansari, J. Berendzen, S. F. Bownde, et al., “Protein States and Proteinquakes,” *Proceedings of the National Academy of Sciences* 82 (1985): 5000–5004.
47. K. Henzler-Wildman and D. Kern, “Dynamic Personalities of Proteins,” *Nature* 450 (2007): 964–972.
48. K. Nam, Y. Shao, D. T. Major, and M. Wolf-Watz, “Perspectives on Computational Enzyme Modeling: From Mechanisms to Design and Drug Development,” *ACS Omega* 9 (2024): 7393–7412.
49. (a) Q. Cui and M. Karplus, “Allostery and Cooperativity Revisited,” *Protein Science* 17 (2008): 1295–1307. (b) J. Guo and H. X. Zhou, “Protein Allostery and Conformational Dynamics,” *Chemical Reviews* 116 (2016): 6503–6515.
50. J. Monod and F. Jacob, “General Conclusions: Telenomic Mechanisms in Cellular Metabolism, Growth, and Differentiation,” *Cold Spring Harbor Symposia on Quantitative Biology* 26 (1961): 389–401.
51. X. Q. Yao and D. Hamelberg, “Dissecting the Allosteric Fine-Tuning of Enzyme Catalysis,” *JACS Au* 4 (2024): 837–846.
52. H. G. Saavedra, J. O. Wrab, J. A. Anderson, J. Li, and V. J. Hilser, “Dynamic Allostery Can Drive Cold Adaptation in Enzymes,” *Nature* 558 (2018): 324–328.
53. M. Corbella, G. P. Pinto, and S. C. L. Kamerlin, “Loop Dynamics and the Evolution of Enzyme Activity,” *Nature Reviews Chemistry* 7 (2023): 536–547.
54. (a) F. S. Lee, Z. T. Chu, and A. Warshel, “Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by the POLARIS and ENZYMIx Programs,” *Journal of Computational Chemistry* 14 (1993): 161–185. (b) Y. Y. Sham, Z. T. Chu, H. Tao, and A. Warshel, “Examining Methods for Calculations of Binding Free Energies: LRA, LIE, PDL-D-LRA, and PDL-D/S-LRA Calculations of Ligands Binding to an HIV Protease,” *Proteins: Structure, Function, and Bioinformatics* 39 (2000): 393–407.
55. (a) T. Hansson, J. Marelus, and J. Åqvist, “Ligand Binding Affinity Prediction by Linear Interaction Energy Methods,” *Journal of Computer-Aided Molecular Design* 12 (1998): 27–35. (b) J. Åqvist and T. Hansson, “Estimation of Binding Free Energy for HIV Proteinase Inhibitors by Molecular Dynamics Simulations,” *Protein Engineering* 8 (1995): 1137–1144.
56. (a) J. Srinivasan, T. E. Cheatham, P. Cieplak, P. A. Kollman, and D. A. Case, “Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate–DNA Helices,” *Journal of the American Chemical Society* 120 (1998): 9401–9409. (b) P. A. Kollman, I. Massova, C. Reyes, et al., “Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models,” *Accounts of Chemical Research* 33 (2000): 889–897.
57. J. Gao, K. Kuczera, B. Tidor, and M. Karplus, “Hidden Thermodynamics of Mutant Proteins: A Molecular Dynamics Analysis,” *Science* 244 (1989): 1069–1072.
58. (a) W. L. Jorgensen, J. K. Buckner, S. Boudon, and J. Tirado-Rives, “Efficient Computation of Absolute Free Energies of Binding by Computer Simulations. Application to the Methane Dimer in Water,” *Journal of Chemical Physics* 89 (1988): 3742–3746. (b) M. K. Gilson, J. A. Given, B. L. Bush, and J. A. McCammon, “The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review,” *Biophysical Journal* 72 (1997): 1047–1069.
59. K. Świderek, S. Martí, and V. Moliner, “Theoretical Studies of HIV-1 Reverse Transcriptase Inhibition,” *Physical Chemistry Chemical Physics* 14 (2012): 12614–12624.
60. D. Trzesniak, A. Pitschna, E. Kunz, and W. F. van Gunsteren, “A Comparison of Methods to Compute the Potential of Mean Force,” *ChemPhysChem* 8 (2007): 162–169.
61. G. Casadevall, C. Duran, and S. Osuna, “AlphaFold2 and Deep Learning for Elucidating Enzyme Conformational Flexibility and Its Application for Design,” *JACS Au* 3 (2023): 1554–1562.
62. (a) A. Laio and M. Parrinello, “Escaping Free-Energy Minima,” *Proceedings. National Academy of Sciences United States of America* 99 (2002): 12562–12566. (b) A. Barducci, G. Bussi, and M. Parrinello, “Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method,” *Physical Review Letters* 100 (2008): 20603.
63. D. Ray and M. Parrinello, “Kinetics From Metadynamics: Principles, Applications, and Outlook,” *Journal of Chemical Theory and Computation* 19 (2023): 5649–5670.
64. S. Izrailev, S. Stepaniants, B. Isralewitz, et al., “Computational Molecular Dynamics: Challenges, Methods, Ideas,” 1997 *Proceedings*

of the 2nd International Symposium on Algorithms for Macromolecular Modelling, Berlin, 39–65.

65. D. Hamelberg, J. Mongan, and J. A. McCammon, “Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules,” *Journal of Chemical Physics* 120 (2004): 11919–11929.

66. R. H. Swendsen and J. Wang, “Replica Monte Carlo Simulation of Spin-Glasses,” *Physical Review Letters* 57 (1986): 2607–2609.

67. U. H. Hansmann, “Parallel Tempering Algorithm for Conformational Studies of Biological Molecules,” *Chemical Physics Letters* 281 (1997): 140–150.

68. G. Torrie and J. Valleau, “Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling,” *Journal of Computational Physics* 23 (1977): 187–199.

69. (a) P. G. Bolhuis, C. Dellago, and D. Chandler, “Sampling Ensembles of Deterministic Transition Pathways,” *Faraday Discussions* 110 (1998): 421–436. (b) P. Bolhuis, D. Chandler, C. Dellago, and P. Geissler, “Transition Path Sampling: Throwing Ropes Over Mountain Passes, in the Dark,” *Annual Review of Physical Chemistry* 53 (2002): 291–318.

70. B. Mohr, T. van Heesch, P. de Alba, A. Ortiz, and J. Vreede, “Enhanced Sampling Strategies for Molecular Simulation of DNA,” *WIREs Computational Molecular Science* 14 (2024): e1712.

71. O. Valsson, P. Tiwary, and M. Parrinello, “Enhancing Important Fluctuations: Rare Events and Metadynamics From a Conceptual Viewpoint,” *Annual Review of Physical Chemistry* 67 (2016): 159–184.

72. (a) W. Wang, S. Cao, L. Zhu, and X. Huang, “Constructing Markov State Models to Elucidate the Functional Conformational Changes of Complex Biomolecules,” *Wiley Interdisciplinary Reviews: Computational Molecular Science* 8 (2018): e1343. (b) B. E. Husic and V. S. Pande, “Markov State Models: From an Art to a Science,” *Journal of the American Chemical Society* 140 (2018): 2386–2396.

73. C. Chatzigeorgoulas and Z. Cournia, “Rational Design of Allosteric Modulators: Challenges and Successes,” *WIREs Computational Molecular Science* 11 (2021): e1529.

74. D. Duan Ni, Z. Chai, Y. Wang, et al., “Along the Allostery Stream: Recent Advances in Computational Methods for Allosteric Drug Discovery,” *WIREs Computational Molecular Science* 12 (2022): e1585.

75. C. F. A. Negre, U. N. Morzan, H. P. Hendrickson, and B. S. Batista, “Eigenvector Centrality for Characterization of Protein Allosteric Pathways,” *Proceedings of the National Academy of Sciences* 115 (2018): E12201–E122208.

76. M. A. Maria-Solano, T. Kinateter, J. Iglesias-Fernández, R. Sterner, and S. Osuna, “In Silico Identification and Experimental Validation of Distal Activity-Enhancing Mutations in Tryptophan Synthase,” *ACS Catalysis* 11 (2021): 13733–13743.

77. G. Casadevall, C. Pierce, B. Guan, et al., “Designing Efficient Enzymes: Eight Predicted Mutations Convert a Hydroxynitrile Lyase Into an Efficient Esterase,” *bioRxiv* (2023). bioRxiv [Preprint]. 2023.08.23.554512, <https://doi.org/10.1101/2023.08.23.554512>.

78. A. T. van Wart, J. Durrant, L. Votapka, and R. E. Amaro, “Weighted Implementation of Suboptimal Paths (WISP): An Optimized Algorithm and Tool for Dynamical Network Analysis,” *Journal of Chemical Theory and Computation* 10 (2014): 511–517.

79. A. M. Stokely, L. W. Votapka, M. T. Hock, et al., “NetSci: A Library for High Performance Biomolecular Simulation Network Analysis Computation,” *Journal of Chemical Information and Modeling* 64, no. 20 (2024): 7966–7976, <https://doi.org/10.26434/chemrxiv-2024-fjrpp>.

80. (a) R. M. Crean, J. S. G. Slusky, P. M. Kasson, and S. C. L. Kamerlin, “KIF—Key Interactions Finder: A Program to Identify the Key Molecular Interactions That Regulate Protein Conformational Changes,” *Journal of Chemical Physics* 158 (2023): 144114. (b) D. Yehorova, R. Crean, P. M. Kasson, and S. C. L. Kamerlin, “Key Interaction Networks: Identifying

Evolutionarily Conserved Non-Covalent Interaction Networks Across Protein Families,” *Protein Science* 33 (2024): e4911.

81. J. Wang, A. Jain, L. R. McDonald, C. Gambogi, A. L. Lee, and N. V. Dokholyan, “Mapping Allosteric Communications Within Individual Proteins,” *Nature Communications* 11, no. 3 (2020): 862.

82. H. T. H. Chan, A. S. F. Oliveira, C. J. Schofield, A. J. Mulholland, and F. Duarte, “Dynamical Nonequilibrium Molecular Dynamics Simulations Identify Allosteric Sites and Positions Associated With Drug Resistance in the SARS-CoV-2 Main Protease,” *JACS Au* 3 (2023): 1767–1774.

83. R. D. Astumian, “Irrelevance of the Power Stroke for the Directionality, Stopping Force, and Optimal Efficiency of Chemically Driven Molecular Machines,” *Biophysical Journal* 108 (2015): 291–303.

84. C. Duran, G. Casadevall, and S. Osuna, “Harnessing Conformational Dynamics in Enzyme Catalysis to Achieve Nature-Like Catalytic Efficiencies: The Shortest Path Map Tool for Computational Enzyme Redesign,” *Faraday Discussions* 252 (2024): 306–322.

85. E. Chovancova, A. Pavelka, P. Benes, et al., “CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures,” *PLoS Computational Biology* 8 (2012): e1002708.

86. T. Magdziarz, K. Mitusinska, M. Bzówka, et al., “AQUA-DUCT 1.0: Structural and Functional Analysis of Macromolecules From an Intramolecular Voids Perspective,” *Bioinformatics* 36 (2020): 2599–2601.

87. N. Naleem, C. R. A. Abreu, K. Warmuz, M. Tong, S. Kirmizialtin, and M. E. Tuckerman, “An Exploration of Machine Learning Models for the Determination of Reaction Coordinates Associated With Conformational Transitions,” *Journal of Chemical Physics* 159 (2023): 34102.

88. H. Fu, H. Bian, X. Shao, and W. Cai, “Collective Variable-Based Enhanced Sampling: From Human Learning to Machine Learning,” *Journal of Physical Chemistry Letters* 15 (2024): 1774–1783.

89. H. Chen and C. Chipot, “Chasing Collective Variables Using Temporal Data-Driven Strategies,” *QRB Discovery* 4 (2023): e2.

90. S. Bhakat, “Collective Variable Discovery in the Age of Machine Learning: Reality, Hype and Everything in Between,” *RSC Advances* 12 (2022): 25010–25024.

91. T. Kikutsuji, Y. Mori, K.-i. Okazaki, T. Mori, K. Kim, and N. Matubayasi, “Explaining Reaction Coordinates of Alanine Dipeptide Isomerization Obtained From Deep Neural Networks Using Explainable Artificial Intelligence (XAI),” *Journal of Chemical Physics* 156 (2022): 154108.

92. K. Świderek, S. Marti, I. Tuñón, V. Moliner, and J. Bertran, “Peptide Bond Formation Mechanism Catalyzed by Ribosome,” *Journal of the American Chemical Society* 137 (2015): 12024–12034.

93. H. Jónsson, G. Mills, and K. W. Jacobsen, “Chapter Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions,” in *Classical and Quantum Dynamics in Condensed Phase Simulations* (Singapore: World Scientific, 1998), 385–404.

94. (a) E. Weinan, W. Ren, and E. Vanden-Eijnden, “String Method for the Study of Rare Events,” *Physical Review B: Condensed Matter and Materials Physics* 66 (2002): 52301. (b) E. Weinan, W. Ren, and E. Vanden-Eijnden, “Simplified and Improved String Method for Computing the Minimum Energy Paths in Barrier-Crossing Events,” *Journal of Chemical Physics* 126 (2007): 164103.

95. T. J. Giese, S. Ekesan, E. McCarthy, Y. Tao, and D. M. York, “Surface-Accelerated String Method for Locating Minimum Free Energy Paths,” *Journal of Chemical Theory and Computation* 20 (2024): 2058–2073.

96. K. Zinovjev and I. Tuñón, “Adaptive Finite Temperature String Method in Collective Variables,” *Journal of Physical Chemistry. A* 121 (2017): 9764–9772.

97. R. García-Meseguer, E. Ortí, I. Tuñón, J. J. Ruiz-Pernía, and J. Aragó, “Insights Into the Enhancement of the Poly(Ethylene Terephthalate)

- Degradation by FAST-PETase From Computational Modeling,” *Journal of the American Chemical Society* 145 (2023): 19243–19255.
98. C. L. Tooke, P. Hinchliffe, M. Beer, et al., “Tautomer-Specific Deacylation and Ω -Loop Flexibility Explain the Carbapenem-Hydrolyzing Broad-Spectrum Activity of the KPC-2 β -Lactamase,” *Journal of the American Chemical Society* 145 (2023): 7166–7180.
99. C. A. Ramos-Guzmán, J. J. Ruiz-Pernía, K. Zinovjev, and I. Tuñón, “Unveiling the Mechanistic Singularities of Caspases: A Computational Analysis of the Reaction Mechanism in Human Caspase-1,” *ACS Catalysis* 13 (2023): 7166–7172.
100. H. Jarbeen, M. Beer, J. Spencer, M. W. van der Kamp, H. A. Bunzel, and A. J. Mulholland, “Electric Fields Are a Key Determinant of Carbapenemase Activity in Class A β -Lactamases,” *ACS Catalysis* 14 (2024): 4348–4361.
101. H. A. Bunzel, J. L. R. Anderson, D. Hilvert, V. L. Arcus, M. W. van der Kamp, and A. J. Mulholland, “Evolution of Dynamical Networks Enhances Catalysis in a Designer Enzyme,” *Nature Chemistry* 13 (2021): 1017–1022.
102. H. A. Bunzel, H. Kries, L. Marcheti, et al., “Emergence of a Negative Activation Heat Capacity During Evolution of a Designed Enzyme,” *Journal of the American Chemical Society* 141 (2019): 11745–11748.
103. J. Åqvist, “Computer Simulations Reveal an Entirely Entropic Activation Barrier for the Chemical Step in a Designer Enzyme,” *ACS Catalysis* 12 (2022): 1452–1460.
104. A. Lear, J. L. Ross Anderson, D. Hilvert, V. L. Arcus, M. W. van der Kamp, and H. A. Bunzel, “Mulholland AJ Commento n: “Computer Simulations Reveal an Entirely Entropic Activation Barrier for the Chemical Step in a Designer Enzyme,”” *ACS Catalysis* 13 (2023): 10527–10530.
105. J. Åqvist, “Reply to Comment on: “Computer Simulations Reveal an Entirely Entropic Activation Barrier for the Chemical Step in a Designer Enzyme,”” *ACS Catalysis* 13 (2023): 10007–10009.
106. E. J. Walker, C. J. Hamill, R. Crean, et al., “Cooperative Conformational Transitions Underpin the Activation Heat Capacity in the Temperature Dependence of Enzyme Catalysis,” *ACS Catalysis* 14 (2024): 4379–4394.
107. L. Koenekoop and J. Åqvist, “Computational Analysis of Heat Capacity Effects in Protein–Ligand Binding,” *Journal of Chemical Theory and Computation* 20 (2024): 5708–5716.
108. J. P. Richard, T. L. Amyes, B. Goryanova, and X. Zhai, “Enzyme Architecture: On the Importance of Being in a Protein Cage,” *Current Opinion in Chemical Biology* 21 (2014): 1–10.
109. J. P. Richard, “Protein Flexibility and Stiffness Enable Efficient Enzymatic Catalysis,” *Journal of the American Chemical Society* 141 (2019): 3320–3331.
110. (a) R. M. Crean, M. Biler, M. W. van der Kamp, A. C. Hengge, and S. C. L. Kamerlin, “Loop Dynamics and Enzyme Catalysis in Protein Tyrosine Phosphatases,” *Journal of the American Chemical Society* 143 (2021): 3830–3845. (b) R. M. Crean, M. Biler, M. Corbella, et al., “Correction to “Loop Dynamics and Enzyme Catalysis in Protein Tyrosine Phosphatases,”” *Journal of the American Chemical Society* 144 (2022): 10091–10093.
111. K. Świderek, A. Panczakiewicz, A. Bujacz, G. Bujacz, and P. Paneth, “Modeling of Isotope Effects on Binding Oxamate to Lactic Dehydrogenase,” *Journal of Physical Chemistry B* 113 (2009): 12782–12789.
112. K. Świderek, I. Tuñón, S. Martí, and V. Moliner, “Protein Conformational Landscapes and Catalysis. Influence of Active Site Conformations in the Reaction Catalyzed by L-Lactate Dehydrogenase,” *ACS Catalysis* 5 (2015): 1172–1185.
113. (a) V. Moliner, A. J. Turner, and I. H. Williams, “Transition-State Structural Refinement With GRACE and CHARMM: Realistic Modelling of Lactate Dehydrogenase Using a Combined Quantum/Classical Method,” *Chemical Communications* (1997): 1271–1272. (b) A. Turner, V. Moliner, and I. H. Williams, “Transition-State Structural Refinement With GRACE and CHARMM: Flexible QM/MM Modelling for Lactate Dehydrogenase,” *Physical Chemistry Chemical Physics* 1 (1999): 1323–1331.
114. A. M. Damry and C. J. Jackson, “The Evolution and Engineering of Enzyme Activity Through Tuning Conformational Landscapes,” *Protein Engineering, Design & Selection* 34 (2021): 1–6.
115. R. Shen, R. M. Crean, S. J. Johnson, S. C. L. Kamerlin, and A. C. Hengge, “Single Residue on the WPD-Loop Affects the pH Dependency of Catalysis in Protein Tyrosine Phosphatases,” *JACS Au* 1 (2021): 646–659.
116. R. Shen, R. M. Crean, K. J. Olsen, et al., “Insights Into the Importance of WPD-Loop Sequence for Activity and Structure in Protein Tyrosine Phosphatases,” *Chemical Science* 13 (2022): 13524–13540.
117. J. B. Greisman, L. Willmore, C. Y. Yeh, et al., “Discovery and Validation of the Binding Poses of Allosteric Fragment Hits to Protein Tyrosine Phosphatase 1b: From Molecular Dynamics Simulations to X-Ray Crystallography,” *Journal of Chemical Information and Modeling* 63 (2023): 2644–2650.
118. K. Zinovjev, P. Guénou, C. A. Ramos-Guzmán, J. J. Ruiz-Pernía, D. Laage, and I. Tuñón, “Activation and Friction in Enzymatic Loop Opening and Closing Dynamics,” *Nature Communications* 15 (2024): 2490.
119. C. Duran, T. Kinader, C. Hiefinger, R. Sterner, and S. Osuna, “Altering Active-Site Loop Dynamics Enhances Standalone Activity of the Tryptophan Synthase Alpha Subunit,” *ACS Catalysis* 14 (2024): 16986–16995.
120. M. Schliwa and G. Woehlke, “Molecular Motors,” *Nature* 422 (2003): 759–765.
121. A. García-Martínez, K. Zinovjev, J. J. Ruiz-Pernía, and I. Tuñón, “Conformational Changes and ATP Hydrolysis in Zika Helicase: The Molecular Basis of a Biomolecular Motor Unveiled by Multiscale Simulations,” *Journal of the American Chemical Society* 145 (2023): 24809–24819.
122. R. D. Vale, “Milligan RA the Way Things Move: Looking Under the Hood of Molecular Motor Proteins,” *Science* 288 (2000): 88–95.
123. (a) A. Warshel, “Electrostatic Origin of the Catalytic Power of Enzymes and the Role of Preorganized Active Sites,” *Journal of Biological Chemistry* 273 (1998): 27035–27038. (b) A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. Liu, and M. H. M. Olsson, “Electrostatic Basis for Enzyme Catalysis,” *Chemical Reviews* 106 (2006): 3210–3235.
124. J. J. Ruiz-Pernía, K. Świderek, J. Bertran, V. Moliner, and I. Tuñón, “Electrostatics as a Guiding Principle in Understanding and Designing Enzymes,” *Journal of Chemical Theory and Computation* 20 (2024): 1783–1795.
125. S. Martí, M. Roca, J. Andres, et al., “Theoretical Insights in Enzyme Catalysis,” *Chemical Society Reviews* 33 (2004): 98–107.
126. S. S. Chaturvedi, S. Vargas, P. Ajmera, and A. N. Alexandrova, “Directed Evolution of Protoglobin Optimizes the Enzyme Electric Field,” *Journal of the American Chemical Society* 146 (2024): 16670–16680.
127. S. Osuna, “The Challenge of Predicting Distal Active Site Mutations in Computational Enzyme Design,” *Wiley Interdisciplinary Reviews: Computational Molecular Science* 11 (2021): e1502.
128. (a) S. L. Lovelock, R. Crawshaw, S. Basler, et al., “The Road to Fully Programmable Protein Catalysis,” *Nature* 606 (2022): 49–58. (b) N. Tokuriki and D. S. Tawfik, “Protein Dynamism and Evolvability,” *Science* 324 (2009): 203–207. (c) E. C. Campbell, G. J. Correy, P. D. Mabbitt, A. M. Buckle, N. Tokuriki, and C. J. Jackson, “Laboratory Evolution of Protein Conformational Dynamics,” *Current Opinion in Structural Biology* 50 (2018): 49–57. (d) R. M. Crean, J. M. Gardner, and

- S. C. L. Kamerlin, "Harnessing Conformational Plasticity to Generate Designer Enzymes," *Journal of the American Chemical Society* 142 (2020): 11324–11342.
129. D. Rothlisberger, O. Khersonsky, A. M. Wollacott, et al., "Kemp Elimination Catalysts by Computational Enzyme Design," *Nature* 453 (2008): 190–195.
130. (a) H. K. Privett, G. Kiss, T. M. Lee, et al., "Iterative Approach to computational enzyme design," *Proceedings of National Academy of Sciences. United States of America* 109 (2012): 3790–3795. (b) R. Blomberg, H. Kries, D. M. Pinkas, et al., "Precision Is Essential for Efficient Catalysis in an Evolved Kemp Eliminate," *Nature* 503 (2013): 418–421.
131. (a) K. Świderek, I. Tuñón, V. Moliner, and J. Bertran, "Protein Flexibility and Preorganization in the Design of Enzymes. The Kemp Elimination Catalyzed by HG3.17," *ACS Catalysis* 5 (2015): 2587–2595. (b) K. Świderek, I. Tuñón, V. Moliner, and J. Bertran, "Revealing the Origin of the Efficiency of the De Novo Designed Kemp Eliminate HG-3.17 by Comparison With the Former Developed HG-3," *Chemistry* 23 (2017): 7582–7589.
132. K. Świderek, I. Tuñón, I. H. Williams, and V. Moliner, "Insights on the Origin of Catalysis on Glycine N-Methyltransferase From Computational Modeling," *Journal of the American Chemical Society* 140 (2018): 4327–4334.
133. (a) A. Bhowmick, S. C. Sharma, and T. Head-Gordon, "The Importance of the Scaffold for de Novo Enzymes: A Case Study With Kemp Eliminate," *Journal of the American Chemical Society* 139 (2017): 5793–5800. (b) V. Vaissier, S. C. Sharma, K. Schaettle, T. Zhang, and T. Head-Gordon, "Computational Optimization of Electric Fields for Improving Catalysis of a Designed Kemp Eliminate," *ACS Catalysis* 8 (2018): 219–227. (c) V. V. Welborn, L. Ruiz Pestana, and T. Head-Gordon, "Computational Optimization of Electric Fields for Better Catalysis Design," *Nature Catalysis* 1 (2018): 649–655.
134. M. A. Galmes, A. R. Nodling, K. He, L. Y. P. Luk, K. Świderek, and V. Moliner, "Computational Design of an Amidase by Combining the Best Electrostatic Features of Two Promiscuous Hydrolases," *Chemical Science* 13 (2022): 4779–4787.
135. M. A. Galmes, A. R. Nodling, L. Y. P. Luk, K. Świderek, and V. Moliner, "Combined Theoretical and Experimental Study to Unravel the Differences in Promiscuous Amidase Activity of Two Nonhomologous Enzymes," *ACS Catalysis* 11 (2021): 8635–8644.
136. K. Świderek, I. Tuñón, V. Moliner, and J. Bertran, "Computational Strategies for the Design of New Enzymatic Functions," *Archives of Biochemistry and Biophysics* 582 (2015): 68–79.
137. D. De Raffele, S. Martí, and V. Moliner, "QM/MM Theoretical Studies of a de Novo Retro-Aldolase Design," *ACS Catalysis* 9 (2019): 2482–2492.
138. L. Giger, S. Caner, R. Obexer, et al., "Evolution of a Designed Retro-Aldolase Leads to Complete Active Site Remodeling," *Nature Chemical Biology* 9 (2013): 494–498.
139. R. Obexer, A. Godina, X. Garrabou, et al., "Emergence of a Catalytic Tetrad During Evolution of a Highly Active Artificial Aldolase," *Nature Chemistry* 9 (2017): 50–56.
140. J. Murciano-Calles, D. K. Romney, S. Brinkmann-Chen, A. R. Buller, and F. H. Arnold, "A Panel of TrpB Biocatalysts Derived From Tryptophan Synthase Through the Transfer of Mutations That Mimic Allosteric Activation," *Angewandte Chemie, International Edition* 55 (2016): 11577–11581.
141. M. A. Maria-Solano, J. Iglesias-Fernández, and S. Osuna, "Deciphering the Allosterically Driven Conformational Ensemble in Tryptophan Synthase Evolution," *Journal of the American Chemical Society* 141 (2019): 13049–13056.
142. R. V. Rakotoharisoa, B. Seifinoferest, N. Zarifi, et al., "Design of Efficient Artificial Enzymes Using Crystallographically Enhanced Conformational Sampling," *Journal of the American Chemical Society* 146 (2024): 10001–10013.
143. F. van der Ent, S. Skagseth, B. A. Lund, et al., "Computational Design of the Temperature Optimum of an Enzyme Reaction," *Science Advances* 9 (2023): eadi0963.
144. S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, and J. Bertran, "Preorganization and Reorganization as Related Factors in Enzyme Catalysis: The Chorismate Mutase Case," *Chemistry—A European Journal* 9 (2003): 984–991.
145. S. D. Schwartz, "Protein Dynamics and Enzymatic Catalysis," *Journal of Physical Chemistry B* 127 (2023): 2649–2660.
146. (a) A. Sharma, G. Gupta, T. Ahma, S. Mansoor, and B. Kaur, "Enzyme Engineering: Current Trends and Future Perspectives," *Food Reviews International* 37, no. 2 (2019): 121–154, <https://doi.org/10.1080/87559129.2019.1695835>. (b) P. Jha, S. Mahato, P. K. Jana, S. Maurya, and I. Chihi, eds., *Artificial Intelligence Based Solutions for Industrial Applications* (Boca Raton, FL: CRC Press, 2024).
147. C. K. K. Reddy, P. R. Anisha, S. Khan, M. M. Hanafiah, L. Pamulaparty, and R. M. Mohana, eds., *Sustainability in Industry 5.0: Theory and Applications* (Boca Raton, FL: CRC Press, 2024).