



ACADÉMIE  
DES SCIENCES  
INSTITUT DE FRANCE

# *Comptes Rendus*

---

## *Chimie*


Esther Pruna Cortada and Sílvia Osuna

**Shortest Path Map correlation-based tool for capturing functionally relevant allosteric networks and its application in enzyme design**

Volume 28 (2025), p. 383-395

Online since: 8 April 2025

<https://doi.org/10.5802/crchim.389>

 This article is licensed under the  
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.  
<http://creativecommons.org/licenses/by/4.0/>



*The Comptes Rendus. Chimie are a member of the  
Mersenne Center for open scientific publishing*  
[www.centre-mersenne.org](http://www.centre-mersenne.org) — e-ISSN : 1878-1543



Review article

# Shortest Path Map correlation-based tool for capturing functionally relevant allosteric networks and its application in enzyme design

Esther Pruna Cortada<sup>®,<sup>a</sup></sup> and Sílvia Osuna<sup>®,\*,<sup>a</sup>,<sup>b</sup></sup>

<sup>a</sup> Institut de Química Computacional i Catàlisi and Departament de Química, Universitat de Girona, Parc R+i Univ. Girona, Ed. Monturiol, c/ Emili Grahit 91, 17003 Girona, Spain

<sup>b</sup> ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain

E-mail: [silvia.osuna@udg.edu](mailto:silvia.osuna@udg.edu) (S. Osuna)

**Abstract.** The impact of distal mutations on enzyme design is analogous to the allosteric regulation effect observed in effector binding or substrate transport within heterocomplexes. Building on this analogy, we employed molecular dynamics simulations to estimate the conformational landscape of enzymes and developed the Shortest Path Map (SPM) tool. This tool identifies key conformationally relevant positions that regulate enzymatic function. In this review, we highlight various applications of the SPM method in its use as a tool for investigating the effects of distal mutations on the conformational landscape, characterizing allosteric regulation, and rationally designing enzymes. A brief description of the method is provided together with several examples reported over the years (from us and other labs) in which the SPM method has been successfully applied.

**Keywords.** Molecular dynamics simulations, Correlation-based tools, Enzyme design, Shortest Path Map, Allostery.

*Manuscript received 13 November 2024, revised 14 February 2025, accepted 26 February 2025.*

## 1. Introduction

Enzyme design seeks to develop biocatalysts with improved “à la carte” characteristics by altering the amino acid sequences of natural enzymes or generating new sequences and tertiary structures. The interest in enzyme design and engineering stems from the unique potential benefits of these catalysts, such as their ability to work efficiently under mild biological conditions while achieving exceptional efficiency, selectivity, and specificity. Enzyme design is intellectually stimulating, serving as a stringent test of our understanding of enzyme stability, folding, evolution, and catalysis.

The process of enzyme design, whether starting from a natural or computationally generated

scaffold, involves selecting specific residues for mutation, creating new variants, and applying screening methods to evaluate the desired properties [1]. There are two primary approaches that can be effectively combined: rational design [2–4] and directed evolution (DE) [5–7]. Rational design focuses on predetermined key residues, identified through the analysis of multiple sequence alignments (MSAs), structural evaluation of active sites, substrate-binding tunnels, reaction mechanisms, and key intermediates formed during the reaction, using techniques like quantum mechanics, molecular mechanics, molecular dynamics (MD), and Monte Carlo simulations [3,8]. Rational methods typically target as mutation hotspots active site residues [9,10] or those involved in the bottleneck regions of tunnels, using tools like CAVER, AQUADUCT, and HotSpot Wizard [8,11,12]. Inspired by

\*Corresponding author

the deep-learning (DL) AF2 approach [13–15], recognized with the 2024 Noble Prize in Chemistry, many DL techniques have been recently developed for generating new protein sequences and scaffolds, which then are further refined using sequence design methods such as ProteinMPNN [16–22]. All these different designs generated by means of traditional methods or DL can also be further improved with DE. Recognized with the 2018 Nobel Prize in Chemistry, DE traditionally utilized random mutagenesis but has advanced to integrate bioinformatics [23–28], rational methods [29,30], sequence analysis [31,32], NMR data [33], and high-throughput screening [34], often guided by machine-learning models [35,36]. The DE approach is particularly powerful for improving low-activity computational enzyme designs [37–39] and enhancing side activities [40,41]. Numerous lab-engineered enzymes have been developed over the years, contributing to drug production, biotherapeutics, and fragrance production [42].

One key advantage of DE is its ability to introduce mutations across the entire protein unlike traditional rational design strategies, which often focus on the active site or substrate tunnels for helping substrate binding/product release and/or altering the water content [8,43]. The DE studies have shown significant activity increases through mutations distant from the active site, which are difficult to predict computationally [6,44–46]. This has been observed in various enzyme families, where in many cases mutations occur around 15 Å from the active site [44]. Intriguingly, the impact of mutations on the enzyme turnover ( $k_{\text{cat}}$ ) is not directly correlated to their proximity to the active site, which contrasts with the more deterministic role of active site mutations in specificity [6]. Such regulation of the enzyme catalytic activity by means of distal mutations suggests that allostery (i.e., regulation of catalytic activity by effector and/or protein binding) plays a key role in many proteins [47]. Molecular dynamics simulations have explained how such mutations influence enzyme conformational dynamics and enhance catalytic activity by altering the network of non-covalent interactions [46,48], thus regulating the flexibility of dynamical elements such as loops or lids gating the active site access [40,49,50]. Although computational models can rationalize these changes, predicting which distal mutations will affect activity remains challenging [46,48,51]. The insights gained

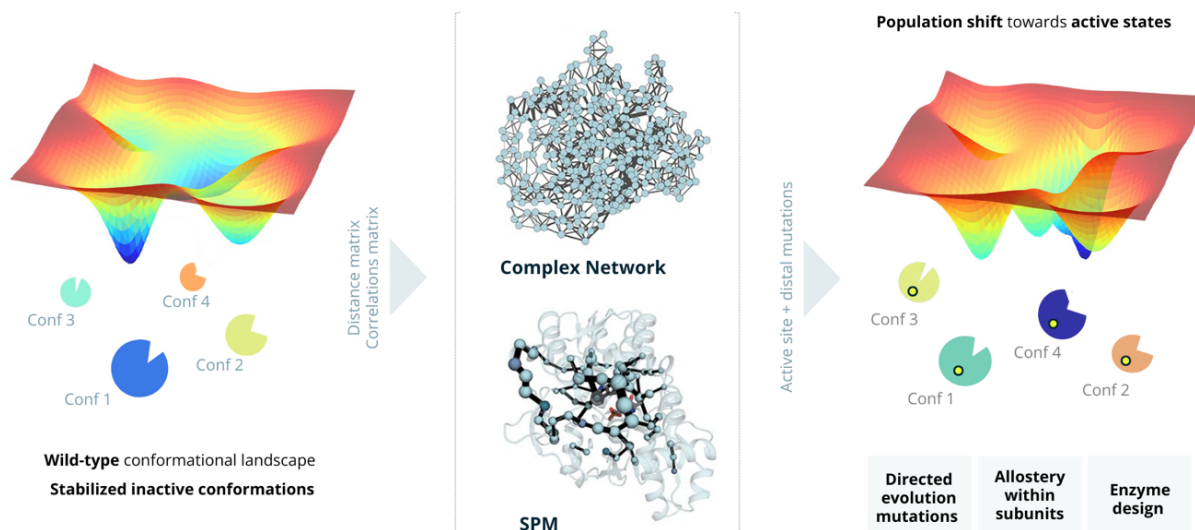
from DE about the role of distal mutations regulating enzymatic function could help develop computational approaches for designing efficient enzymes with natural-like functionality [44].

The influence of distal mutations in enzyme design parallels the allosteric regulation seen in effector binding or substrate transport within heterocomplexes. These mutations can shift the conformational landscape of the enzyme, favoring a subset of conformations that enhance catalysis. Recognizing this similarity, we explored the application of MD simulations for estimating the enzymes' conformational landscape and extracting key conformationally relevant positions for their use in enzyme design [2,48]. We developed the Shortest Path Map (SPM) tool [52], which constructs a graph based on mean distances and correlation values between residues as determined from MD simulations. Unlike previous allosteric studies that identified regions of importance [53], the SPM method identifies specific residue pairs that significantly influence enzyme dynamics (Figure 1). This precise identification of key residues is valuable in enzyme design, enabling the creation of targeted libraries for mutagenesis.

## 2. SPM tool

The SPM builds a graph based on mean distances and correlation values computed from MD simulations, following a strategy similar to the protocol used by Sethi et al. [53]. For each protein residue, a node is created centered on the carbon alpha atom. The next step is to establish edges between node pairs. An edge is formed between pairs of nodes whose  $C_{\alpha}$  atoms remain within 6 Å of each other throughout the MD simulation. The edge distance is then derived from the calculated correlation values, which represent the information transfer along that edge. Residue pairs with higher correlation values (closer to 1 or -1) have shorter edge distances while less correlated pairs (values closer to 0) have longer edge distances. This approach results in a graph with nodes and edges reflecting residue proximity and correlation, which is then simplified as illustrated in Figure 1.

Unlike previous allosteric studies that focus on identifying communities within the first complex graph generated, we compute the shortest path lengths. The SPM identifies which edges of the graph



**Figure 1.** Computational enzyme design based on inducing a population shift on the enzyme conformational landscape towards active states by using the SPM method. Left: The wild-type enzyme conformational landscape features multiple stable states, which are predominantly catalytically inactive (labeled Conf 1, Conf 2, Conf 3, and Conf 4). This represents the enzyme conformational heterogeneity, but the enzyme presents poor catalytic efficiency as active states are poorly populated. Center: By computing the distance and the correlation matrices from the multiple replica MD simulations, the first complex graph network is generated, which is further simplified to obtain the SPM network. SPM identifies potential mutational hotspots that can influence the global conformational landscape and induce a shift in the population distribution towards catalytically active states. Right: By introducing targeted mutations at both active and distal sites, the conformational landscape is reshaped to favor catalytically active states. This approach has a wide range of applications, including directed evolution mutation identification (Section 2.1), elucidation and optimization of allosteric pathways (Section 2.2), and finally, rational enzyme design (Section 2.3).

are the shortest (i.e., most correlated) and most frequently used when passing through all the protein residues. After normalizing all edges, only those with the highest contribution are depicted and visualized in the 3D structure. The key advantage of the SPM is that it directly pinpoints critical residues rather than broad regions, making it especially valuable for enzyme design by enabling the construction of small libraries of hotspot positions.

The SPM has proven effective in identifying mutation spots, either within the active site or at distal locations targeted by DE. It has also been applied for elucidating the allosteric regulation existing in multimeric protein structures and more recently for enzyme design. The following sections discuss several cases where SPM has been applied: (1) identification of DE mutation hotspots, (2) rationalization of

allosterically relevant residues, and (3) for enzyme design.

### 2.1. SPM captures DE mutations

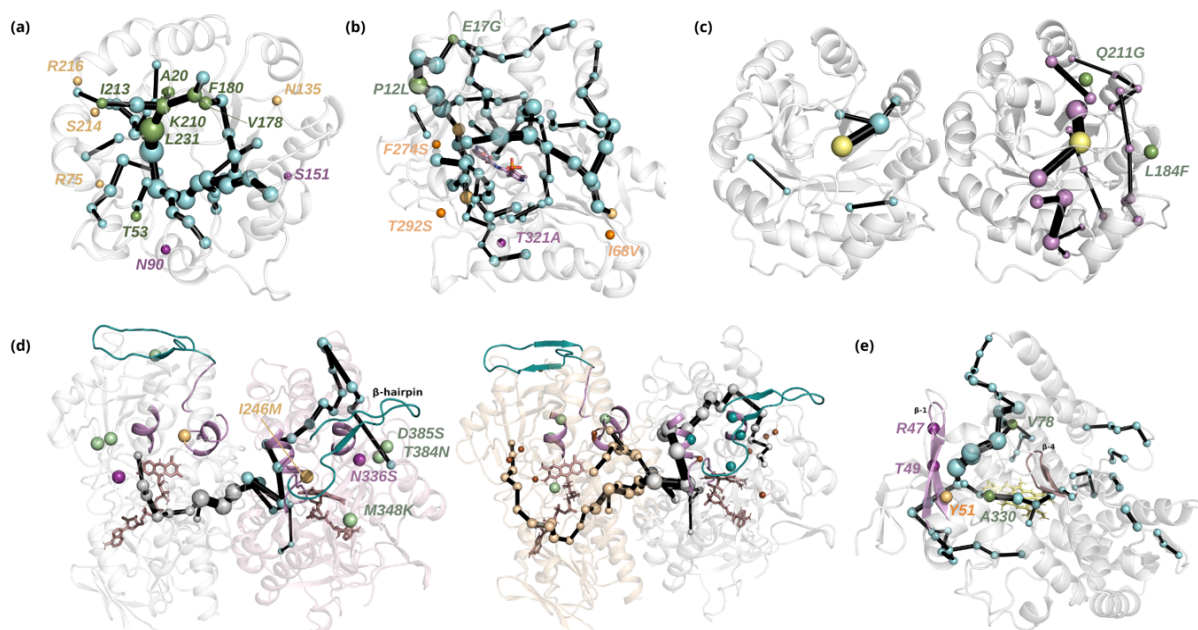
The DE approach has become an invaluable strategy in enzyme design engineering, enabling the discovery of enzymes with improved or novel functionalities through iterative rounds of mutation and selection [5–7]. The key advantage of DE is its ability to identify beneficial mutations across the entire protein, which are often difficult to predict computationally, but it lacks mechanistic insights into how the identified mutations affect enzyme dynamics and conformational changes. Computational models can rationalize DE mutations, but the challenge persists in the prediction of these key mutations.

The SPM's predictive capabilities have been validated in several key case studies, where it successfully identified mutations that mirrored those obtained through DE. For instance, in the case of mechanistically complex **retro-aldolase (RA) enzymes**, DE introduced multiple mutations scattered throughout the enzyme, which drastically increased its catalytic activity towards the abiological aldol substrate methodol [23,45,48]. The SPM tool could predict most of the mutational hotspots observed in the experimentally evolved enzymes. Specifically, the SPM identified seven of the thirteen mutations introduced by DE while four others were located at adjacent positions and only two were more than six positions away in the sequence (Figure 2a). Of note is that designed RAs are  $(\beta\alpha)_8$ -barrel enzymes, which is a fold shared by many enzymes suggesting that the application of the SPM tool might be quite broad. A similar result was observed in the complex homodimeric **monoamine oxidase from *Aspergillus niger* (MAO-N)** [49,54,55], where DE was employed to increase catalytic activity towards chiral amines and broaden the enzyme's substrate scope [56]. Five positions were introduced by DE for obtaining the MAO-N D5 evolved variant [57]. The key DE mutations include one active site mutation, one mutation at the entrance tunnel, and three additional distant mutations up to 18–19 Å away from the active site, all of them, despite being computationally challenging to predict, identified by the SPM tool or located at adjacent positions (Figure 2d). The computational evaluation of the conformational changes induced by these mutations by means of Markov state models indicated major changes in beta-hairpin conformation between wild-type and D5 variants. Additionally, the SPM applied to D5 contains the DE positions for generating D9 and D11 variants [57]. This agreement between computational predictions and experimental outcomes underscores the SPM's effectiveness in identifying functionally relevant distal positions and highlights its ability to predict key residues regulating enzyme activity by analyzing the conformational dynamics of the starting enzyme [44].

The SPM tool was also employed by the Mulholland Laboratory to analyze shifts in dynamical networks within the transition-state (TS) ensemble along the DE of a computationally designed **Kemp eliminase**, which catalyzes a proton abstraction from carbon by a base [58]. An SPM analysis of the

closed states uncovered a dynamical network that connected the active site to more distal regions of the protein, which indicates fine-tuning of its dynamics by remote mutations (Figure 2c). Mutations such as Q211G and L184F, introduced during DE, were also identified by SPM as key contributors to this network, facilitating the global response of the enzyme scaffold to the TS. Mutation Q211G increased flexibility, fine-tuning the dynamic response while L184F enhanced packing by tightening the interactions between neighboring solvent-exposed loops, further stabilizing the TS. These mutations were essential in reducing the structural fluctuations in the TS ensemble, thereby improving catalytic preorganization.

Another striking example of SPM's utility to identify the mutations introduced by DE comes from work on the heterocomplex **tryptophan synthase from *Pyrococcus furiosus* (*Pf*TrpS)**. The TrpB complex, a subunit of the tryptophan synthase complex (TrpS), was subjected to DE to create an efficient stand-alone enzyme variant independent of the alpha subunit (TrpA) with improved catalytic activity [59,60]. The evolved variant, *Pf*TrpB-0B2, contained six distal mutations that restored the enzyme's ability to access essential conformational states in the absence of its TrpA protein binding partner [61,62] (Figure 2b). Two of the distal mutations, P12L and E17G, introduced by DE are located near the TrpA–TrpB interface. Another mutation, F247S, is found in a known tunnel while T292S is situated in one of the loops close to the essential COMM domain that covers the active site where the pyridoxal phosphate cofactor is located. Additionally, two more mutations, I68V and T321A, are present in more remote regions of the enzyme. Of the six mutations in *Pf*TrpB-0B2, the first two were directly predicted by SPM and three were making persistent non-covalent interactions with SPM-identified positions. Mutation T292S, which significantly enhanced enzyme activity, modulated COMM domain closure via interactions with D300, supporting SPM predictions. The only residue not captured by the SPM is T321, which is considered to play a minor role in COMM domain conformational dynamics. The comparison of the free energy landscapes between the wild type and the evolved *Pf*TrpB-0B2 indicated that these distal mutations recovered the conformational heterogeneity observed for the wild type when forming complexes with TrpA [61]. This study highlights the



**Figure 2.** Some examples of the application of the SPM method for capturing DE mutations. (a) SPM is applied to retro-aldolase enzyme. Of the thirteen mutations introduced by DE, seven were directly identified by SPM (green) with an additional four located at adjacent positions (yellow). Only two mutations (purple) were positioned more than six residues away in sequence. (b) Six distal mutations in *Pf*TrpB were introduced by DE to generate a stand-alone *Pf*TrpB-0B2 variant. Of these, five mutations were identified by SPM: two directly captured by SPM (green) and three adjacent to SPM-identified positions (orange). The only distal mutation not captured by SPM is shown in purple. (c) The SPM in 1A53-2 (left) and 1A53-2.5 (right) depicts the interaction network centered on the transition state and extending to the closed state of the protein. The two mutations found by DE are directly involved in the SPM. (d) Right: The five mutations in MAON-WT identified by DE to generate MAON-D5 evolved variant are shown as spheres. The key DE mutations include one active site mutation (purple), one mutation at the entrance tunnel (yellow), and three additional distant mutations (green). Left: The DE mutations on MAON-D5 towards MAON-D9 variant are also captured by SPM and are shown in dark green while previous mutations are shown in brown. The beta-hairpin structure is highlighted in teal, and the entrance loop is represented in violet in both images. The FAD cofactors are depicted as sticks in dark violet. (e) The SPM pathway of the CYP450 enzyme included the previously identified DE mutations V78 and A330 (green). SPM also depicts the long-distance connection between beta-sheet 1 (purple) and beta-sheet 4 (light violet). The key mutation Y51I, identified by SPM as an adjacent position, is shown in yellow.

potential of the SPM in identifying key residues for improving enzyme function, offering a powerful tool for targeting allosterically regulated systems.

The SPM was also applied to study the epistatic effects and conformational dynamics in the stepwise evolution of a **cytochrome P450-BM3 monooxygenase** engineered for regioselective and stereoselective hydroxylation of a steroid [63]. The previously identified DE mutations V78 and A330 were

successfully included in the SPM pathway of the parent enzyme (Figure 2e). Additionally, combinatorial saturation mutagenesis revealed that the mutant R47I/T49I/Y51I/F87A exhibited improved activity compared to the wild type, achieving 94%  $2\beta$ -selectivity and 67% substrate conversion. Computational analysis showed that while mutations R47I and T49I alone did not affect selectivity, the Y51I mutation shifted substrate orientation, increasing

2 $\beta$ -selectivity. The SPM successfully identified the Y51 position providing the key residue responsible for both activity and selectivity. Furthermore, the SPM pathway detected a long-distance communication network between  $\beta$ 4 and  $\beta$ 1, involving positions R47, T49, and Y51.

These selected examples show that the SPM method is a valuable computational tool for capturing the conformational effects of mutations introduced by DE. By pinpointing key residues and pathways that regulate conformational transitions, the SPM can be, in principle, used to enhance the efficiency of enzyme engineering efforts by predicting conformationally relevant positions not strictly located at the active site. The successful application of the SPM in enzymes like RA, monoamine oxidase, tryptophan synthase, Kemp eliminase, and cytochrome P450 illustrates its potential to complement and rationalize DE approaches, accelerating the development of enzymes with optimized functionalities through a deeper understanding of their conformational landscapes.

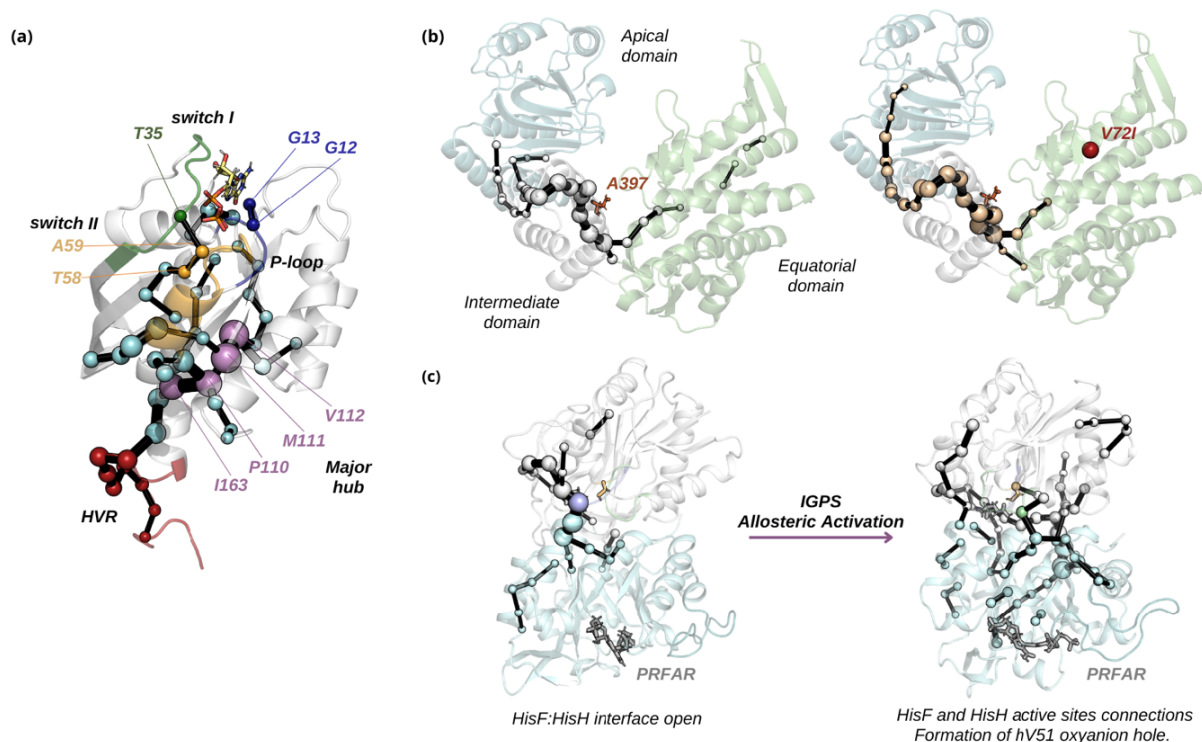
## 2.2. SPM identifies allosteric pathways

Allosteric regulation is a fundamental process in which distant residues within a protein communicate to regulate its activity. This long-range communication allows proteins to respond to various signals, adjusting their function accordingly. Understanding the specific pathways through which allosteric signals travel is crucial to enzyme design and for understanding how mutations alter protein function. The SPM has proven effective in mapping allosteric pathways in key proteins, offering insights into how mutations can alter their regulation and function. For instance, the SPM was applied to 5  $\mu$ s atomistic MD simulations of the **K-Ras4B protein** embedded in a phospholipid membrane to investigate its allosteric regulation [64]. Allosteric regulation in K-Ras, particularly the oncogenic K-Ras4B, is key to its function in cellular signaling. K-Ras switches between an active GTP-bound state and an inactive GDP-bound state with allosteric changes in key regions of the G domain (switches I and II), enabling effector recruitment and signal transduction. Mutations, especially in Gly12 and Gly13, can disrupt this regulation, locking K-Ras in an active state and driving cancer progression [64]. The SPM

identified a prominent allosteric pathway starting at the hypervariable region (HVR), a near-membrane region that might be receiving allosteric messages and transferring them to the major hub, located on helix  $\alpha$ 5, loop  $\alpha$ 3– $\beta$ 5, and the N-terminus of sheet  $\beta$ 5 (Figure 3a). Branching out to key areas in the G domain, the allosteric route extends to regions crucial to K-Ras function, including the P-loop, Thr58, and Ala59 on switch II and Thr35 on switch I. The analysis of the allosteric communication existing in K-Ras4B was also assessed by means of other measures: distance fluctuation analysis [65,66], anisotropic thermal diffusion [67], and dynamical non-equilibrium simulations [68], which provided complementary information on how allostery is transmitted when hydrolyzing GTP [64]. The SPM therefore provided useful insights into K-Ras allostery that help in identifying potential allosteric sites for therapeutic targeting, advancing drug design efforts for this protein.

In a recent study, the allosteric regulation mechanism of the human mitochondrial heat shock protein (**Hsp60**), which is responsible for controlling proteostasis, is elucidated using the SPM tool [69]. This study examines the effect of V72I point mutation, linked to hereditary spastic paraplegia SPG13, which is characterized by progressive weakness and spasticity of the lower limbs. An SPM analysis was performed on MD simulations of both the wild-type (M WT) and mutant (M V72I) Hsp60 monomers (Figure 3b). The results indicate that the extensive interdomain allosteric communication present in the wild-type monomer is entirely disrupted in the V72I mutant, leading to decreased efficiency in communication across the intermediate domain, which contains the critical catalytic residue Asp 397. By mapping the allosteric network in the Hsp60 monomer, this work offers valuable insights into the pathogenic effects of the V72I mutation, demonstrating how specific distal mutations can alter protein function and contribute to disease and highlighting the potential applicability of SPM in rationalizing allosterically regulated mechanisms.

Imidazole glycerol phosphate synthase (IGPS) enzyme has been a model for studying allosteric regulation [70–72]. The enzyme IGPS consists of two subunits, HisH and HisF, with HisH catalyzing glutamine hydrolysis and HisF delivering ammonia through an internal tunnel. Bound 30 Å from the



**Figure 3.** Some examples of the application of the SPM method for identifying allosteric networks. (a) K-Ras4B: SPM reveals the allosteric network connecting the membrane-embedded HVR region (dark red) to the central hub (violet), which extends to switch II (orange) and further connects to switch I (green). Notably, the critical mutations Gly12 and Gly13 in the P-loop (blue) are also captured by SPM. (b) Hsp60: The allosteric network links the equatorial domain (wheat), intermediate domain (white), and apical domain (dark red). Introducing the V72I mutation disrupts this network, disrupting the communication between the equatorial domain and the active site, represented by residue A397 in dark orange. (c) IGPS: SPM analysis shows the progression of IGPS from the open HisH–HisF interface to an allosterically active state, where multiple pathways connect HisF and HisH active sites, facilitating the formation of the hV51 oxyanion hole.

HisH active site, PRFAR acts as an allosteric effector, enhancing glutaminase activity by 4500-fold [73]. In this case, a time-evolution Shortest Path Map (te-SPM) analysis is applied to MD simulations, specifically by examining accelerated MD simulations in 600 ns time intervals, uncovering the progression of allosteric activation [74] (Figure 3c). Initially, correlated motions are limited to the HisH subunit and the HisF–HisH interface. However, as the te-SPM analysis progresses, it reveals concerted motions that are activated upon productive closure of the HisF–HisH interface, extending throughout the entire HisF subunit, the interdomain region, and the HisH active site, leading to the formation of the hV51

oxyanion hole. This dynamic process aligns with  $\mu$ s–ms motions observed by NMR and suggests that IGPS allosteric activation follows a “violin model” of dynamics-based allostery, similar to what is found in some protein kinases [75]. The truncated SPM approach for the time-evolution analysis can be also used to decipher the molecular basis of allosteric mechanisms in related allosterically regulated enzymes.

All the provided examples support the idea that the SPM tool is a useful strategy for uncovering the intricate mechanisms of allosteric regulation in proteins by identifying critical pathways of residue communication. So far, the SPM has offered deep insights

into how proteins modulate their functions and how mutations can disrupt this regulation. Therefore, the SPM serves as a promising tool for rational design of targeted therapeutics and enzyme engineering.

### 2.3. Conformationally driven enzyme design with SPM

The application of the SPM in different unrelated enzymes demonstrates that analyzing an enzyme conformational ensemble can pinpoint key positions that significantly impact its conformational dynamics. Notably, the SPM identifies positions throughout the enzyme, not just in the active site, allowing for the prediction of distal mutations. This suggests a strategy for conformationally driven enzyme design involving the reconstruction of the enzyme conformational landscape, detection of critical positions using the SPM, and identification of the specific amino acid substitution for each mutation through experimental or computational methods.

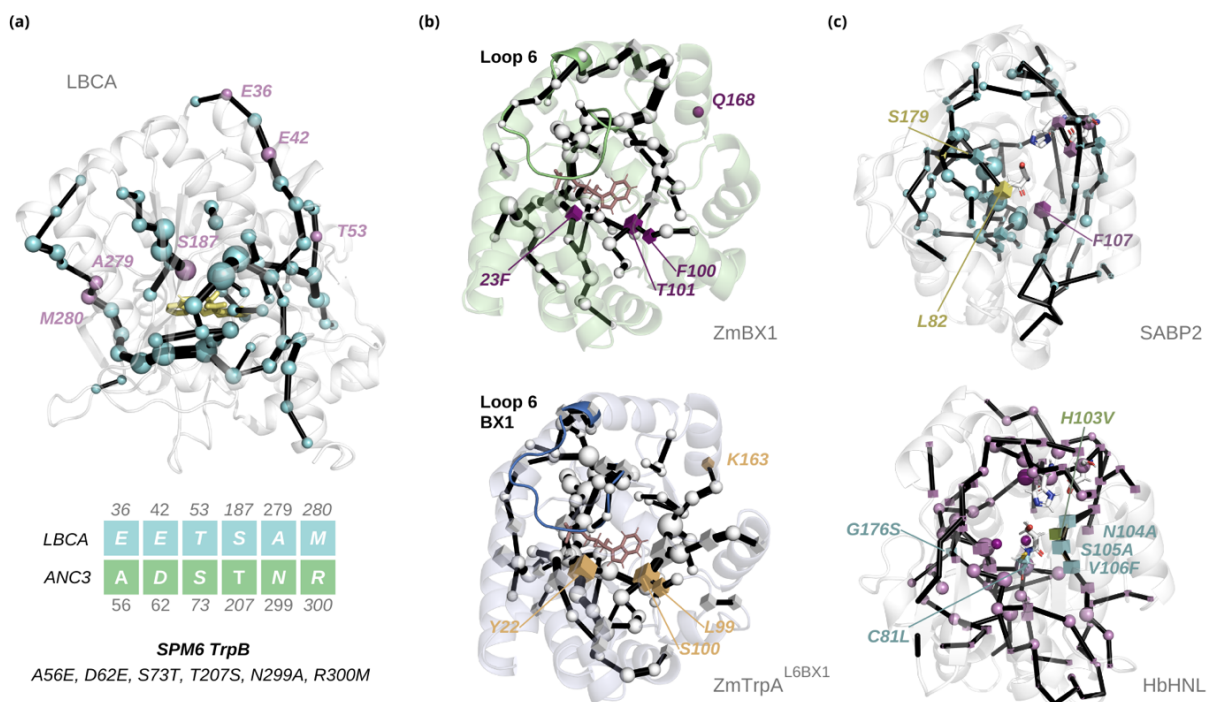
We engineered a **stand-alone functional TrpB** by combining the SPM with ancestral sequence reconstruction (ASR) [76]. Interestingly, ASR revealed that the last bacterial common ancestor (LBCA) is allosterically inhibited by TrpA, thus presenting high stand-alone activity [77,78]. Over the course of evolution, this inhibition shifted towards allosteric activation; specifically ANC3 TrpS is the first ancestral enzyme displaying allosteric activation by TrpA. By coupling SPM and sequence conservation analysis between LBCA and ANC3 TrpB, six key positions were rationally identified, leading to the SPM6 variant, which achieved a sevenfold increase in catalytic activity [76] (Figure 4a). This improvement of activity obtained by the experimental testing of a few (<5) variants was similar to that obtained after multiple rounds of DE in *PfTrpB* that required the generation and screening of more than 3000 variants [59]. Two additional variants were developed that further validate the approach: SPM3, based on ANC2 TrpB, involved three mutations; SPM8, derived from an additional SPM analysis of ANC3 TrpS instead of TrpB, added two allosterically relevant positions [76]. While SPM8 achieved similar catalytic improvement to SPM6, SPM3 showed a smaller enhancement, which was expected from the reduced number of mutations. An additional comparison of SPM with MSA revealed the complementarity of the

two methods, as they identified distinct mutation sites with only one shared residue mutation between them [76].

In another recent study also focused on a tryptophan synthase, the SPM tool was employed to enhance the **stand-alone activity of *Zea mays* TrpA (*ZmTrpA*)** by identifying key conformationally relevant positions [79] (Figure 4b). Unlike TrpB, where a shift in allosteric regulation was observed along its phylogenetic tree, LBCA TrpA was already allosterically activated by TrpB [77,78]. This made more difficult the identification of positions for stand-alone activity using traditional methods like MSA. However, the structurally similar stand-alone enzyme *ZmBX1* [80] from the secondary metabolism of maize served as a model for designing an enhanced TrpA variant [79].

The initial approach involved engineering the *ZmTrpA* to the *ZmTrpA*-L6Bx1 variant by switching the L6 loop with the L6 found in *ZmBX1* [80], as L6 closed conformation observed in *ZmBX1* was reported to be crucial to TrpA stand-alone activity [79]. This resulted in a significant increase in catalytic turnover ( $k_{cat}$ ) but at the expense of increasing  $K_M$ , indicating the need for additional improvements to achieve better catalytic efficiency [80]. To further enhance the stand-alone activity of *ZmTrpA*, the SPM method was applied to identify key conformationally relevant positions by comparing the intramolecular pathways of *ZmTrpA*-L6Bx1 and *ZmBX1* at the catalytically activated state. The motivation for specifically targeting the IGP-bound catalytically activated state was the experimental observation that the conformational transition to achieve the catalytically active state is rate-determining [81]. By comparing the generated SPM for the starting *ZmTrpA*-L6Bx1 scaffold and *ZmBX1*, a subset of non-conserved SPM-identified positions was selected for mutation, leading to the *ZmTrpA*-SPM4-L6Bx1 variant, which exhibits a 178-fold improvement in catalytic efficiency towards IGP cleavage [79]. With this study, the power of SPM in identifying both active site and distal mutations to boost stand-alone activity is highlighted and paves the way for enzyme design in those enzymatic systems where conformational change is rate-limiting.

The SPM tool was also successfully applied to predict mutations that converted **hydroxynitrile lyase (HbHNL)** from the rubber tree into an



**Figure 4.** Some examples of the application of the SPM method for rational enzyme design. (a) Designing stand-alone TrpB variants based on the ancestral ANC3-TrpB scaffold: SPM identifies a set of conformationally relevant positions in LBCA-TrpB (violet), which after sequence analysis conservation with ANC3-TrpB residues led to the SPM6 TrpB design exhibiting a sevenfold increase in catalytic activity with respect to ANC3 WT. (b) Designing stand-alone TrpA enzymes: the comparison of the SPM graphs computed considering IGP-bound catalytically activated states of the stand-alone ZmBX1 (top) and ZmTrpA-L6BX1 (bottom) allowed the generation of the ZmTrpA-SPM4-L6BX1 variant. The selected positions mutated are shown in purple for ZmBX1 and in yellow for ZmTrpA-L6BX1. ZmTrpA-SPM4-L6BX1 presents a 178-fold higher stand-alone catalytic efficiency compared to ZmTrpA. (c) Designing efficient esterases from hydroxynitrile lyases: the introduction of the three obvious active site mutations for providing an oxyanion hole pocket to *HbHNL* and one additional stabilization mutation yielded HNL3V, which shows poor esterase activity. The SPM analysis applied to the reference SABP2 esterase (top) and HNL3V (bottom) is depicted, with spheres indicating residues shared between both proteins and cubes marking differing residues. Four/five SPM positions were selected for mutagenesis, yielding HNL7 and HNL8 that both exhibit a 300-fold enhancement in esterase catalytic activity as compared to *HbHNL*. The incorporation of evolutionary information into SPM provided HNL7T, which displays a 1450-fold increase in catalytic efficiency, thus surpassing that of the SABP2 esterase taken as reference.

efficient esterase by comparison with tobacco **esterase (SABP2)** [82] (Figure 4c). Both enzymes are present in the  $\alpha$ , $\beta$ -hydrolase fold and despite the fact that they share the same catalytic Ser-His-Asp triad, they exhibit very different forms of reactivity. Initial efforts to engineer *HbHNL* into an esterase produced the HNL3V variant, which focused on the

three obvious active site mutations (T11G, E79H, K236M) to introduce the oxyanion hole residues and to replace the polar site by a hydrophobic site as observed in SABP2. However, these three obvious mutations only showed limited esterase activity. To further enhance the esterase activity, SPM analysis of HNL3V, a stabilized version of HNL3 for experimental

purposes, and SABP2 revealed five additional positions outside the active site, which were connected to the catalytic residues and could potentially affect the oxyanion hole and the catalytic Asp preorganization. By replacing these positions with those from SABP2, the HNL6V, HNL7V, and HNL8V variants were created, enhancing esterase activity by 300-fold. The SPM-guided mutations targeted correlated movements essential for catalysis by removing unwanted movements in HNL3V and adding the missing ones from SABP2. To this end, the residues in HNL3V were replaced with the corresponding ones from SABP2. Notably, most of the engineered mutations were outside the active site, demonstrating the ability of the SPM to identify distal residues contributing to catalytic function.

The combination of SPM analysis with an MSA of esterases resulted in the identification of the most effective variant, HNL7TV. Among the mutations tested, the N104A substitution produced the largest increase in turnover rate ( $k_{\text{cat}}$ ). However, sequence alignment revealed that most esterases naturally have a threonine at this position rather than an alanine. Introducing the N104T substitution led to the best overall variant with approximately twice the catalytic efficiency ( $k_{\text{cat}}/K_{\text{M}}$ ) of SABP2, thus increasing the esterase catalytic efficiency of the starting *HbHNL* scaffold by 1450-fold. Despite this impressive improvement, the turnover rate ( $k_{\text{cat}}$ ) of HNL7TV remained 13-fold lower than SABP2, indicating a need for further refinement. Nevertheless, the dramatic increase in catalytic efficiency demonstrated the ability of the SPM to predict which residues outside the active site contribute to catalytic activity [82].

### 3. Conclusions

Recognizing the high similarity between allosteric processes and the effect exerted by distal active site mutations on the enzyme catalytic activity, we hypothesized that the computational identification of allosteric networks could bring interesting insights into enzymatic function and design [44]. To that end, we developed the SPM method [44,52], which identifies a set of conformationally relevant positions based on the application of graph theory to the distance and correlation matrices computed from MD data. Over the years, we (and others) have applied the SPM to identify conformationally relevant

positions and compared those with residues mutated in DE experiments [2,44,48,58,63]. Interestingly, in many different enzyme classes and families, the application of SPM has revealed that some of the DE mutations are introduced in conformationally relevant sites, thus suggesting its potential application for enzyme design.

The SPM has also been used for elucidating the intricate mechanisms of allosteric regulation in different enzymatic and protein systems [64,69,74]. The SPM allosteric networks identify critical pathways of residues, some of which are experimentally known to have a large effect on enzyme/protein regulation. We have also tested the applicability of SPM, especially if combined with ASR and/or MSA, to rationally design new improved enzymes with enhancements in catalytic efficiency that range from 7- to 1450-fold [76,79,82]. Altogether these examples show the potential of the SPM method for understanding the effect of distal mutations into the enzyme conformational dynamics and how they affect enzymatic function. More importantly, integrating SPM with other computational techniques such as MSA and ASR corresponds to a successful approach for rational enzyme design.

### Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

### References

- [1] E. L. Bell, W. Finnigan, S. P. France, et al., "Biocatalysis", *Nat. Rev. Meth. Primers* **1** (2021), article no. 46.
- [2] M. A. Maria-Solano, E. Serrano-Hervás, A. Romero-Rivera, J. Iglesias-Fernández and S. Osuna, "Role of conformational dynamics in the evolution of novel enzyme function", *Chem. Commun.* **54** (2018), pp. 6622–6634.
- [3] A. Romero-Rivera, M. Garcia-Borràs and S. Osuna, "Computational tools for the evaluation of laboratory-engineered biocatalysts", *Chem. Commun.* **53** (2017), pp. 284–297.
- [4] J. Damborsky and J. Brezovsky, "Computational tools for designing and engineering enzymes", *Curr. Opin. Chem. Biol.* **19** (2014), pp. 8–16.
- [5] F. H. Arnold, "The nature of chemical innovation: new enzymes by evolution", *Quart. Rev. Biophys.* **48** (2015), pp. 404–410.

- [6] A. Currin, N. Swainston, P. J. Day and D. B. Kell, "Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently", *Chem. Soc. Rev.* **44** (2015), pp. 1172–1239.
- [7] Y. Wang, P. Xue, M. Cao, T. Yu, S. T. Lane and H. Zhao, "Directed evolution: methodologies and applications", *Chem. Rev.* **121** (2021), pp. 12384–12444.
- [8] C. E. Sequeiros-Borja, B. Surpeta and J. Brezovsky, "Recent advances in user-friendly computational tools to engineer protein function", *Brief. Bioinform.* **22** (2020), article no. baa150.
- [9] J. J. Ruiz-Pernía, K. Świderek, J. Bertran, V. Moliner and I. Tuñón, "Electrostatics as a guiding principle in understanding and designing enzymes", *J. Chem. Theory Comput.* **20** (2024), pp. 1783–1795.
- [10] C. Jerves, R. P. P. Neves, S. L. da Silva, M. J. Ramos and P. A. Fernandes, "Rate-enhancing PETase mutations determined through DFT/MM molecular dynamics simulations", *New J. Chem.* **48** (2024), pp. 45–54.
- [11] J. Stourac, O. Vavra, P. Kokkonen, J. Filipovic, G. Pinto, J. Brezovsky, J. Damborsky and D. Bednar, "Caver Web 1.0: identification of tunnels and channels in proteins and analysis of ligand transport", *Nucleic Acids Res.* **47** (2019), W414–W422.
- [12] L. Sumbalova, J. Stourac, T. Martinek, D. Bednar and J. Damborsky, "HotSpot Wizard 3.0: web server for automated design of mutations and smart libraries based on sequence input information", *Nucleic Acids Res.* **46** (2018), W356–W362.
- [13] A. W. Senior, R. Evans, J. Jumper, et al., "Improved protein structure prediction using potentials from deep learning", *Nature* **577** (2020), pp. 706–710.
- [14] A. Ourmazd, K. Moffat and E. E. Lattman, "Structural biology is solved — now what?", *Nat. Meth.* **19** (2022), pp. 24–26.
- [15] J. Jumper, R. Evans, A. Pritzel, et al., "Highly accurate protein structure prediction with AlphaFold", *Nature* **596** (2021), pp. 583–589.
- [16] S. Ovchinnikov and P.-S. Huang, "Structure-based protein design with deep learning", *Curr. Opin. Chem. Biol.* **65** (2021), pp. 136–144.
- [17] W. Gao, S. P. Mahajan, J. Sulam and J. J. Gray, "Deep learning in protein structural modeling and design", *Patterns* **1** (2020), article no. 100142.
- [18] Y. Ming, W. Wang, R. Yin, M. Zeng, L. Tang, S. Tang and M. Li, "A review of enzyme design in catalytic stability by artificial intelligence", *Brief. Bioinform.* **24** (2023), pp. 1–19.
- [19] A. Lauko, S. J. Pellock, I. Anischanka, et al., "Computational design of serine hydrolases", *Science* **0** (2025), article no. eadu2454.
- [20] P. Notin, N. Rollins, Y. Gal, C. Sander and D. Marks, "Machine learning for functional protein design", *Nat. Biotechnol.* **42** (2024), pp. 216–228.
- [21] J. Dauparas, I. Anishchenko, N. Bennett, et al., "Robust deep learning-based protein sequence design using ProteinMPNN", *Science* **378** (2022), pp. 49–56.
- [22] B. I. M. Wicky, L. F. Milles, A. Courbet, et al., "Hallucinating symmetric protein assemblies", *Science* **378** (2022), pp. 56–61.
- [23] L. Jiang, E. A. Althoff, F. R. Clemente, et al., "De novo computational design of retro-aldol enzymes", *Science* **319** (2008), pp. 1387–1391.
- [24] D. Rothlisberger, O. Khersonsky, A. M. Wollacott, et al., "Kemp elimination catalysts by computational enzyme design", *Nature* **453** (2008), pp. 190–195.
- [25] J. B. Siegel, A. Zanghellini, H. M. Lovick, et al., "Computational design of an enzyme catalyst for a stereoselective bimolecular Diels–Alder reaction", *Science* **329** (2010), pp. 309–313.
- [26] R. Kourist, H. Jochens, S. Bartsch, et al., "The  $\alpha/\beta$ -hydrolase fold 3DM database (ABHDB) as a tool for protein engineering", *ChemBioChem* **11** (2010), pp. 1635–1643.
- [27] R. K. Kuipers, H.-J. Joosten, W. J. H. van Berkel, et al., "3DM: systematic analysis of heterogeneous superfamily data to discover protein functionalities", *Proteins: Struct. Funct. Bioinform.* **78** (2010), pp. 2101–2113.
- [28] R. K. P. Kuipers, H.-J. Joosten, E. Verwiel, et al., "Correlated mutation analyses on super-family alignments reveal functionally important residues", *Proteins: Struct. Funct. Bioinform.* **76** (2009), pp. 608–616.
- [29] R. Blomberg, H. Kries, D. M. Pinkas, P. R. E. Mittl, M. G. Grütter, H. K. Privett, S. L. Mayo and D. Hilvert, "Precision is essential for efficient catalysis in an evolved Kemp eliminase", *Nature* **503** (2013), pp. 418–421.
- [30] M. Pavlova, M. Klvana, Z. Prokop, et al., "Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate", *Nat. Chem. Biol.* **5** (2009), pp. 727–733.
- [31] T. A. Addington, R. W. Mertz, J. B. Siegel, et al., "Prediction and ranking of mutations required for functional interconversion of enzymes", *J. Mol. Biol.* **425** (2013), pp. 1378–1389.
- [32] A. Pavelka, E. Chovancova and J. Damborsky, "HotSpot Wizard: a web server for identification of hot spots in protein engineering", *Nucleic Acids Res.* **37** (2009), W376–W383.
- [33] S. Bhattacharya, E. G. Margheritis, K. Takahashi, et al., "NMR-guided directed evolution", *Nature* **610** (2022), pp. 389–393.
- [34] H. Xiao, Z. Bao and H. Zhao, "High throughput screening and selection methods for directed enzyme evolution", *Ind. Eng. Chem. Res.* **54** (2015), pp. 4011–4020.
- [35] K. K. Yang, Z. Wu and F. H. Arnold, "Machine-learning-guided directed evolution for protein engineering", *Nat. Methods* **16** (2019), pp. 687–694.
- [36] S. Mazurenko, Z. Prokop and J. Damborsky, "Machine learning in enzyme engineering", *ACS Catal.* **10** (2020), pp. 1210–1223.
- [37] P. A. Romero and F. H. Arnold, "Exploring protein fitness landscapes by directed evolution", *Nat. Rev. Mol. Cell Biol.* **10** (2009), pp. 866–876.
- [38] C. Jaeckel, P. Kast and D. Hilvert, "Protein design by directed evolution", *Annu. Rev. Biophys.* **37** (2008), pp. 153–173.
- [39] H. Renata, Z. J. Wang and F. H. Arnold, "Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution", *Angew. Chem. Int. Ed.* **54** (2015), pp. 3351–3367.

- [40] E. Campbell, M. Kaltenbach, G. J. Correy, et al., “The role of protein dynamics in the evolution of new enzyme function”, *Nat. Chem. Biol.* **12** (2016), pp. 944–950.
- [41] R. B. Leveson-Gower, C. Mayer and G. Roelfes, “The importance of catalytic promiscuity for enzyme design and evolution”, *Nat. Rev. Chem.* **3** (2019), pp. 687–705.
- [42] R. Buller, S. Lutz, R. J. Kazlauskas, R. Snajdrova, J. C. Moore and U. T. Bornscheuer, “From nature to industry: harnessing enzymes for biocatalysis”, *Science* **382** (2023), article no. eadh8615.
- [43] A. Gora, J. Brezovsky and J. Damborsky, “Gates of enzymes”, *Chem. Rev.* **113** (2013), pp. 5871–5923.
- [44] S. Osuna, “The challenge of predicting distal active site mutations in computational enzyme design”, *WIREs Comput. Mol. Sci.* **11** (2021), article no. e1502.
- [45] R. Obexer, A. Godina, X. Garrabou, P. R. E. Mittl, D. Baker, A. D. Griffiths and D. Hilvert, “Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase”, *Nat. Chem.* **9** (2017), pp. 50–56.
- [46] G. Jiménez-Osés, S. Osuna, X. Gao, et al., “The role of distant mutations and allosteric regulation on LovD active site dynamics”, *Nat. Chem. Biol.* **10** (2014), pp. 431–436.
- [47] K. Gunasekaran, B. Ma and R. Nussinov, “Is allostery an intrinsic property of all dynamic proteins?”, *Proteins: Struct. Funct. Bioinform.* **57** (2004), pp. 433–443.
- [48] A. Romero-Rivera, M. Garcia-Borràs and S. Osuna, “Role of conformational dynamics in the evolution of retroaldolase activity”, *ACS Catal.* **7** (2017), pp. 8524–8532.
- [49] C. Curado-Carballada, F. Feixas, J. Iglesias-Fernández and S. Osuna, “Hidden conformations in *Aspergillus niger* monoamine oxidase are key for catalytic efficiency”, *Angew. Chem. Int. Ed.* **58** (2019), pp. 3097–3101.
- [50] D. Petrović, V. A. Risso, S. C. L. Kamerlin and J. M. Sanchez-Ruiz, “Conformational dynamics and enzyme evolution”, *J. R. Soc. Interface* **15** (2018), article no. 20180330.
- [51] P. Campitelli, T. Modi, S. Kumar and S. B. Ozkan, “The role of conformational dynamics and allostery in modulating protein evolution”, *Annu. Rev. Biophys.* **49** (2020), pp. 267–288.
- [52] G. Casadevall, J. Casadevall, C. Duran and S. Osuna, “The shortest path method (SPM) webserver for computational enzyme design”, *Protein Eng. Des. Select.* **37** (2024), article no. gzae005.
- [53] A. Sethi, J. Eargle, A. A. Black and Z. Luthey-Schulten, “Dynamical networks in tRNA:protein complexes”, *Proc. Natl. Acad. Sci. USA* **106** (2009), pp. 6620–6625.
- [54] V. F. Batista, J. L. Galman, D. C. G. A. Pinto, A. M. S. Silva and N. J. Turner, “Monoamine oxidase: tunable activity for amine resolution and functionalization”, *ACS Catal.* **8** (2018), pp. 11889–11907.
- [55] F. Zhao, D. Masci, S. Ferla, et al., “Monoamine oxidase (MAO-N) biocatalyzed synthesis of indoles from indolines prepared via photocatalytic cyclization/arylation dearomatization”, *ACS Catal.* **10** (2020), pp. 6414–6421.
- [56] R. Carr, M. Alexeeva, A. Enright, T. S. C. Eve, M. J. Dawson and N. J. Turner, “Directed evolution of an amine oxidase possessing both broad substrate specificity and high enantioselectivity”, *Angew. Chem. Int. Ed.* **42** (2003), pp. 4807–4810.
- [57] D. Ghislieri, A. P. Green, M. Pontini, S. C. Willies, I. Rowles, A. Frank, G. Grogan and N. J. Turner, “Engineering an enantioselective amine oxidase for the synthesis of pharmaceutical building blocks and alkaloid natural products”, *J. Am. Chem. Soc.* **135** (2013), pp. 10863–10869.
- [58] 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”, *Nat. Chem.* **13** (2021), pp. 1017–1022.
- [59] A. R. Buller, S. Brinkmann-Chen, D. K. Romney, M. Herger, J. Murciano-Calles and F. H. Arnold, “Directed evolution of the tryptophan synthase  $\beta$ -subunit for stand-alone function recapitulates allosteric activation”, *Proc. Natl. Acad. Sci. USA* **112** (2015), pp. 14599–14604.
- [60] E. Watkins-Dulaney, S. Straathof and F. Arnold, “Tryptophan synthase: biocatalyst extraordinaire”, *ChemBioChem* **22** (2021), pp. 5–16.
- [61] M. A. Maria-Solano, J. Iglesias-Fernández and S. Osuna, “Deciphering the allosterically driven conformational ensemble in tryptophan synthase evolution”, *J. Am. Chem. Soc.* **141** (2019), pp. 13049–13056.
- [62] A. R. Buller, P. van Roye, J. K. B. Cahn, R. A. Scheele, M. Herger and F. H. Arnold, “Directed evolution mimics allosteric activation by stepwise tuning of the conformational ensemble”, *J. Am. Chem. Soc.* **140** (2018), pp. 7256–7266.
- [63] C. G. Acevedo-Rocha, A. Li, L. D’Amore, et al., “Pervasive cooperative mutational effects on multiple catalytic enzyme traits emerge via long-range conformational dynamics”, *Nat. Commun.* **12** (2021), article no. 1621.
- [64] M. Castelli, F. Marchetti, S. Osuna, A. S. F. Oliveira, A. J. Mulholland, S. A. Serapian and G. Colombo, “Decrypting allostery in membrane-bound K-Ras4B using complementary in silico approaches based on unbiased molecular dynamics simulations”, *J. Am. Chem. Soc.* **146** (2024), pp. 901–919.
- [65] A. Triveri, C. Sanchez-Martin, L. Torielli, et al., “Protein allostery and ligand design: computational design meets experiments to discover novel chemical probes”, *J. Mol. Biol.* **434** (2022), article no. 167468.
- [66] M. Castelli, K. Bhattacharya, E. Abboud, S. A. Serapian, D. Picard and G. Colombo, “Phosphorylation of the Hsp90 Co-Chaperone Hop changes its conformational dynamics and biological function”, *J. Mol. Biol.* **435** (2023), article no. 167931.
- [67] N. Ota and D. A. Agard, “Intramolecular signaling pathways revealed by modeling anisotropic thermal diffusion”, *J. Mol. Biol.* **351** (2005), pp. 345–354.
- [68] A. S. F. Oliveira, G. Ciccotti, S. Haider and A. J. Mulholland, “Dynamical nonequilibrium molecular dynamics reveals the structural basis for allostery and signal propagation in biomolecular systems”, *Eur. Phys. J. B* **94** (2021), article no. 144.
- [69] L. Torielli, F. Guarra, H. Shao, J. E. Gestwicki, S. A. Serapian and G. Colombo, “How a pathogenic mutation impairs Hsp60 functional dynamics from monomeric to fully assembled states”, preprint, bioRxiv, 2024, 2024.2009.2009.611948.

- [70] S. Beismann-Driemeyer and R. Sterner, "Imidazole glycerol phosphate synthase from *thermotoga maritima*: quaternary structure, steady-state kinetics, and reaction mechanism of the bienzyme complex", *J. Biol. Chem.* **276** (2001), pp. 20387–20396.
- [71] I. Rivalta, G. P. Lisi, N.-S. Snoeberger, G. Manley, J. P. Loria and V. S. Batista, "Allosteric communication disrupted by a small molecule binding to the imidazole glycerol phosphate synthase protein–protein interface", *Biochemistry* **55** (2016), pp. 6484–6494.
- [72] B. Reisinger, J. Sperl, A. Holinski, et al., "Evidence for the existence of elaborate enzyme complexes in the Paleoproterozoic era", *J. Am. Chem. Soc.* **136** (2014), pp. 122–129.
- [73] G. P. Lisi, K. W. East, V. S. Batista and J. P. Loria, "Altering the allosteric pathway in IGPS suppresses millisecond motions and catalytic activity", *Proc. Natl. Acad. Sci. USA* **114** (2017), E3414–E3423.
- [74] C. Calvó-Tusell, M. A. Maria-Solano, S. Osuna and F. Feixas, "Time evolution of the millisecond allosteric activation of imidazole glycerol phosphate synthase", *J. Am. Chem. Soc.* **144** (2022), pp. 7146–7159.
- [75] L. G. Ahuja, S. S. Taylor and A. P. Kornev, "Tuning the "violin" of protein kinases: the role of dynamics-based allostery", *IUBMB Life* **71** (2019), pp. 685–696.
- [76] M. A. Maria-Solano, T. Kinatader, J. Iglesias-Fernández, R. Sterner and S. Osuna, "In silico identification and experimental validation of distal activity-enhancing mutations in tryptophan synthase", *ACS Catal.* **11** (2021), pp. 13733–13743.
- [77] M. Schupfner, K. Straub, F. Busch, R. Merkl and R. Sterner, "Analysis of allosteric communication in a multienzyme complex by ancestral sequence reconstruction", *Proc. Natl. Acad. Sci. USA* **117** (2020), pp. 346–354.
- [78] F. Busch, C. Rajendran, K. Heyn, S. Schlee, R. Merkl and R. Sterner, "Ancestral tryptophan synthase reveals functional sophistication of primordial enzyme complexes", *Cell Chem. Biol.* **23** (2016), pp. 709–715.
- [79] C. Duran, T. Kinatader, C. Hiefinger, R. Sterner and S. Osuna, "Altering active site loop dynamics enhances stand-alone activity of tryptophan synthase alpha-subunit", *ACS Catal.* **14** (2024), pp. 16986–16995.
- [80] M. Schupfner, F. Busch, V. H. Wysocki and R. Sterner, "Generation of a stand-alone tryptophan synthase  $\alpha$ -subunit by mimicking an evolutionary blueprint", *ChemBioChem* **20** (2019), pp. 2747–2751.
- [81] V. Kulik, E. Hartmann, M. Weyand, M. Frey, A. Gierl, D. Niks, M. F. Dunn and I. Schlichting, "On the structural basis of the catalytic mechanism and the regulation of the alpha subunit of tryptophan synthase from *Salmonella typhimurium* and BX1 from maize, Two evolutionarily related enzymes", *J. Mol. Biol.* **352** (2005), pp. 608–620.
- [82] G. Casadevall, C. Pierce, B. Guan, et al., "Designing efficient enzymes: eight predicted mutations convert a hydroxynitrile lyase into an efficient esterase", preprint, bioRxiv, 2023, 2023.2008.2023.554512.