Structural-based Modeling in Protein Engineering, a Must Do

Sergi Roda[1], Ana Robles-Martín[1], Ruite Xiang[1], Masoud Kazemi[1], Victor Guallar[1,2,]*

[1]Barcelona Supercomputing Center (BSC), Barcelona, Spain

[2]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.

ABSTRACT

Biotechnological solutions will be a key aspect in our immediate future society where optimized enzymatic processes through enzyme engineering might be an important solution for waste transformation, clean energy production, biodegradable materials, and green chemistry, for example. Here we advocate the importance of structural-based bioinformatics and molecular modeling tools in such developments. We summarize our recent experiences indicating a great prediction/success ratio, and suggest that an early *in silico* phase should be performed in enzyme engineering studies. Moreover, we demonstrate the potential of a new technique combining Rosetta and PELE which could provide a faster and more automated procedure, an essential aspect for a broader use.

KEYWORDS: enzymology, computational chemistry, enzyme engineering, protein, PELE, enzyme-substrate, molecular modeling, bioprospecting.

INTRODUCTION

Our society is facing extraordinary challenges that urge us to find (bio)technological solutions. While we observe a nascent societal turn toward a more responsible consumption, it is clear that such a change is being implanted at a very slow rate. Moreover, sustainability, and the hypothetical green deal, will still require such technologies capable of, for example, waste transformation, cleaner energy production, biodegradable materials, green chemistry, etc. Part of these developments will take place in the form of industrial enzymatic processes, involving a significant effort in two key aspects: enzyme bioprospecting[1,2] and engineering[3,4]. The first one aims at identifying novel enzymes with optimal or improved properties toward some goal, such as substrate specificity/promiscuity, thermal stability, increased activity, etc, while the second one seeks to produce new variants, after introducing mutations, with a similar objective but from an already characterized enzyme. In both cases, the explosion of data coupled with the extraordinary development of hardware and software tools offers encouraging perspectives.

In the past few years, we have witnessed a significant number of proof of concept studies where computer simulations make a difference in selecting and delivering improved enzymatic variants[5–11]. The potential of running massive parallel computational experiments together with the improvement of the simulation's quality, has provided some of the finest examples in recent enzyme engineering. Furthermore, current developments in sequence annotation and automated biochemical characterization will soon provide enough big data to develop the next wave of enzyme optimization tools based on machine learning. Nonetheless, while significant efforts have been reported in the area, it is still in its early stages[12]. We want to discuss here, however, more traditional structural-based bioinformatics and molecular modeling techniques and, in particular, those developed in our laboratory. We want to convince you that today, an enzyme engineering campaign should start with a thorough

modeling effort, in a similar way that pharmaceutical companies carry on any drug development project. Thus, we aim at demonstrating the maturity of structural-based modeling techniques in enzymatic biotechnology, at the level of both bioinformatics and molecular mechanics.

RESULTS AND DISCUSSION

It is all about the structure. Protein engineering has successfully incorporated structural-based computational selection and design techniques thanks to its fast and low-cost implementation. Accordingly, we find multiple laboratories devoted to methods development and applications; software pieces such as Rosetta[13,14] ORBIT[15,16], OSPREY[17], 3DM[18], and FoldX[19] have become widespread techniques nowadays. In our group, we typically combine biochemical and biophysical molecular modeling techniques into protocols that allow describing the substrate binding events and (if necessary) electronic properties that take place in the catalytic process[20]. While each application might require different protocols, a typical procedure (Figure 1) includes the following steps:

(1) Global enzyme and substrate binding search. We use our PELE (Protein Energy Landscape Exploration) software[21,22], a Monte Carlo (MC) sampling method that includes protein structure prediction techniques, to carry out an unconstrained ligand exploration. This is intended mostly when searching for potential active/binding sites or probing the substrate's possibilities for entering buried ones (study of migration pathways). In this step, we measure different metrics such as potential catalytic distances, the substrate solvent accessible surface area, or the enzyme-substrate interaction energy profile, where the lowest energy minima correspond with the main binding modes. While we might already enter some specific mutations in this initial exploration, this is typically reserved for step 2.

(2) *In silico* mutagenesis and local exploration. Once the active or binding sites of interest have been identified, or different binding modes found within a given active site, we proceed with a higher resolution local search coupled with variant generation. At this point, we follow different metrics describing the protein-substrate interactions: catalytic distances, interaction energies, time (MC steps) of residence, solvent accessibility, etc.

(3) Quantum Mechanics/Molecular Mechanics (QM/MM) evaluation. In some particular cases, as when aiming at increasing oxidation rates, QM/MM calculations are performed on selected snapshots. These might provide additional valuable metrics, such as the estimation of the spin densities on the substrate[23–25].

(4) Mutant selection and experimental validation. Mutants are selected based on the different computational metrics combined with a sequence conservation study, which we typically perform using HotSpot Wizard[26]. Selected variants are proposed for *in vitro* validation, the outcome of which drives typically a second round of *in silico* screening where we aim at combining multiple experimentally proved mutations.

These different steps, along with additional implementations shown below, have been developed in a modular fashion and implemented in a modeling platform. Current efforts, in addition, are focusing on porting these modular packages into the BioExcel building blocks and biocontainers formats.

Next, some specific applications will be revised. The potential of performing a global enzyme-substrate search is illustrated in the rational design of a highly stable manganese peroxidase (MnP6)[27], which we activated (zero initial activity) toward 2,2′-azinobis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) oxidation after the introduction of two nonconserved surface mutations far away from the active site[28]. In this particular case,

simulations allowed the characterization of a substrate binding site that challenged the established one. QM/MM spin density calculations, which included the heme compound I prosthetic group, further indicated the presence of a strong radical character in the substrate, indicative of its oxidation. After experimental validation of the *in silico* proposed double mutant, we obtained a comparable specificity constant to that of active peroxidases.

Obviously, in most cases, a global search is not necessary and we can proceed to a local exploration of enzyme-substrate molecular interactions in the active site, coupled with variant generation. Using such a procedure, for example, we successfully engineered a double mutant high-redox-potential laccase with enhanced aniline oxidation and stability in an acidic medium[29]. One of the mutations was a negatively charged residue which improved the chemical environment of the active site, modifying the redox potential of aniline and stabilizing the oxidized form. Finally, after experimental validation, the double mutant laccase increased its catalytic activity with a 2-fold increase in the turnover number.

A local analysis with PELE was also used to engineer a *Marasmius rotula* unspecific peroxidase (*Mro*UPO) for substrate modulation. In this particular case, substrate (active site) entrance simulations were performed to generate variants along the entrance pathway that significantly affected the substrate specificity profile[30].

The potential of quickly and accurately probing variants was exploited in implementing the first *in silico* atomistic directed evolution protocol. A fungal laccase was engineered for activity increase, obtaining a redox potential boost from 740 to 790 mV, with a concomitant improvement in thermal and acidic pH stability[31]. Interestingly, parallels *in silico* and *in vitro* directed evolution studies were attempted, achieving analogous results leading to the first single mutant. Molecular modeling, however, allowed a second round of saturated

mutagenesis on selected positions, leading to the final double mutant. Importantly, a round of *in silico* directed evolution, mapping all possible single mutants on ~40-50 amino acids surrounding the active site, was accomplished in only 2 days of modest supercomputing resources (~64-128 computing cores). These examples (see also Table 1), along with many others performed recently for enzyme characterization[32], indicate the maturity of atomistic molecular modeling.

Structural-based active site analysis has also been recently used in our laboratory for bioprospecting, a key aspect when building an enzyme-driven biotechnological process. When aiming for substrate promiscuity, for example, enzyme-substrate active site diffusion simulations can accurately determine and quantify experimental substrates[33]. These simulations, however, are still too demanding when aiming at screening large (sequence) data. Thus, we developed a structural bioinformatics descriptor, the effective volume, capable of predicting esterases with a promiscuity substrate profile. This property normalizes the volume of the active site by the solvent accessible surface area of the catalytic triad: while increasing the size might increase promiscuity, a too exposed active site will quickly reduce the number of esters hydrolyzed. The full calculation requires about ~20 min per protein in a single computing core, thus being able to quickly analyze on the order of thousands of sequences. To address genome or metagenome data, however, predictions should be performed much faster, probably at the sequence space level; here machine learning techniques could come in handy. For this purpose, an ensemble classifier was developed by combining 3 classical machine learning algorithms: KNN (k-nearest neighbors), SVM (support vector machines), and a linear model, specifically the RidgeClassifier implementation in scikit-learn. This ensemble approach can identify promiscuous esterases, those with activity toward more than 20 substrates (from a 96 substrates set), with an MCC

(Matthew correlation coefficient) score of 0.76 in the test set compared to the mean score of 0.67 ± 0,09 if we take into account individually the models[34]. Thus, better and more robust predictions can be produced by combining multiple learners.

Having studied the rules for esterase substrate promiscuity, we further converted a low-promiscuous serine ester hydrolase into a high-promiscuous one, while maintaining high catalytic efficiency[35]. Our goal, and that of most modeling campaigns, was to provide a rather small set of mutants for experimental validation with a high success rate. Using our PELE pipeline, we provided 11 mutants, 4 of which increased the substrate promiscuity, two of them with lowered catalytic efficiencies, and the other two acquired prominent promiscuity and high activity levels (k_{cat} up to ca. 152.124 min^{-1} in the WT compared to k_{cat} up to ca. 216.103 min^{-1} and 119.348 min^{-1} in the successful mutants). In the end, we were able to computationally convert a low-promiscuous esterase (16 hydrolyzed esters out of 96 tested) into a prominent one (63 out of 96) while maintaining high efficiency.

How long does it take to perform such designs? The average time for a first round of computational design takes approximately 2-4 weeks, where the first 1-2 days are employed in learning about the system and in the molecular model preparation, followed by a comprehensive mutational analysis using all or part of the protocols above described. From this first round, we usually select the 5-10 mutants to be experimentally tested where we typically see an approximated success rate of 30-40%. Depending on the experimental validation results, we might perform an additional round of design refinement, which usually lasts another 2 weeks.

How accurate is this modeling? Are we talking about only success cases? This is a recurrent question we always get when presenting these results. The straight answer is that it is over

80% of success, based on the 9 successful engineering projects from the last ~5 years, summarized in Table 1, and the two only projects whose mutants did not add any significant improvement and were not published. Clearly, we need a good structure to work with, but other than that our modeling techniques provide a fast and reliable prediction in almost all systems we attempted.

Additional structural modeling methods and studies by other groups. Clearly, structural-based engineering through modeling has been studied by many other laboratories. While we do not aim here for an exhaustive review, we want to underline some state-of-the-art studies, highlighting the differences with our methods. One of the best known examples, and a pioneer in many aspects, is the work by professor David Baker. In Baker's lab, they have improved the properties of many enzymes[36–41] through engineering and, in particular, have centered on designing *de novo* enzymes to catalyze non-natural reactions[7,8,42] by developing and applying the Rosetta[13,14,43] software. In fact, catalytically active natural enzymes have been found after their computational designs[44,45]. Thus, they have clearly shown the potential of structural-based modeling to design artificial active sites, enabling the smart exploration of the sequence space in proteins. The Rosetta software uses a complex combination of MC and molecular dynamics (MD) techniques, based on intercalation of structural know protein segments, such as trimers and ninemers. The overall procedure, however, has not been optimized for substrate placement. Furthermore, in Fleishman's lab, they have expanded the potential of Rosetta's structural calculations with evolution-guided design[46–51], based on the phylogenetic information encoded in the protein family (which they named FuncLib[46]). An example includes reshaping the substrate and cofactor specificity of two natural enzymes (acetyl-CoA synthetase, propionyl-CoA reductase) to enhance the

bypass of CO_2 fixation with glycolate (which does not exist in nature), avoiding carbon-releasing photorespiration[49].

Another laboratory that leverages the advantages of structure-based computational modeling is the Kamerlin's group, which has a special focus on the study of atomistic protein dynamics, using MD, to enhance enzymes or create new ones[52–58]. In a recent study, they showed the importance of conformational flexibility by studying a loop of a tyrosine phosphatase, which contains a residue that acts as an acid/base catalyst. Mutation of a noncatalytic residue involved in the dynamics of this loop changed the pH-rate profile of the enzyme[57]. In Janssen & Wijma's lab, they also use structural-based MD modeling to improve the thermal stability of enzymes (and proteins) with their own computational workflow[39,59–63] (FRESCO[59]). They aim to reduce the library of variants to be experimentally tested, but assuring an enhanced stability of the protein. To present an example, they increased the thermostability of the haloalkane dehalogenase LinB by 23ºC (increase in apparent melting temperature) based on energy calculations, designing disulfide bonds, MD simulations, and rational inspection[61]. Also, the enzyme had increased solvent tolerance, showing their hypothesis that the improvement of enzyme stability will lead to the improvement of other properties of the catalyst. While these examples show the potential of a dynamical conformational sampling, modeling hundreds of variants through MD is a significantly expensive effort, particularly when aiming at coupling the dynamics with a robust enzyme-substrate exploration. It is here where PELE offers a competitive advantage, providing an atomistic and flexible exploration that quickly maps the enzyme-substrate energy landscape[32].

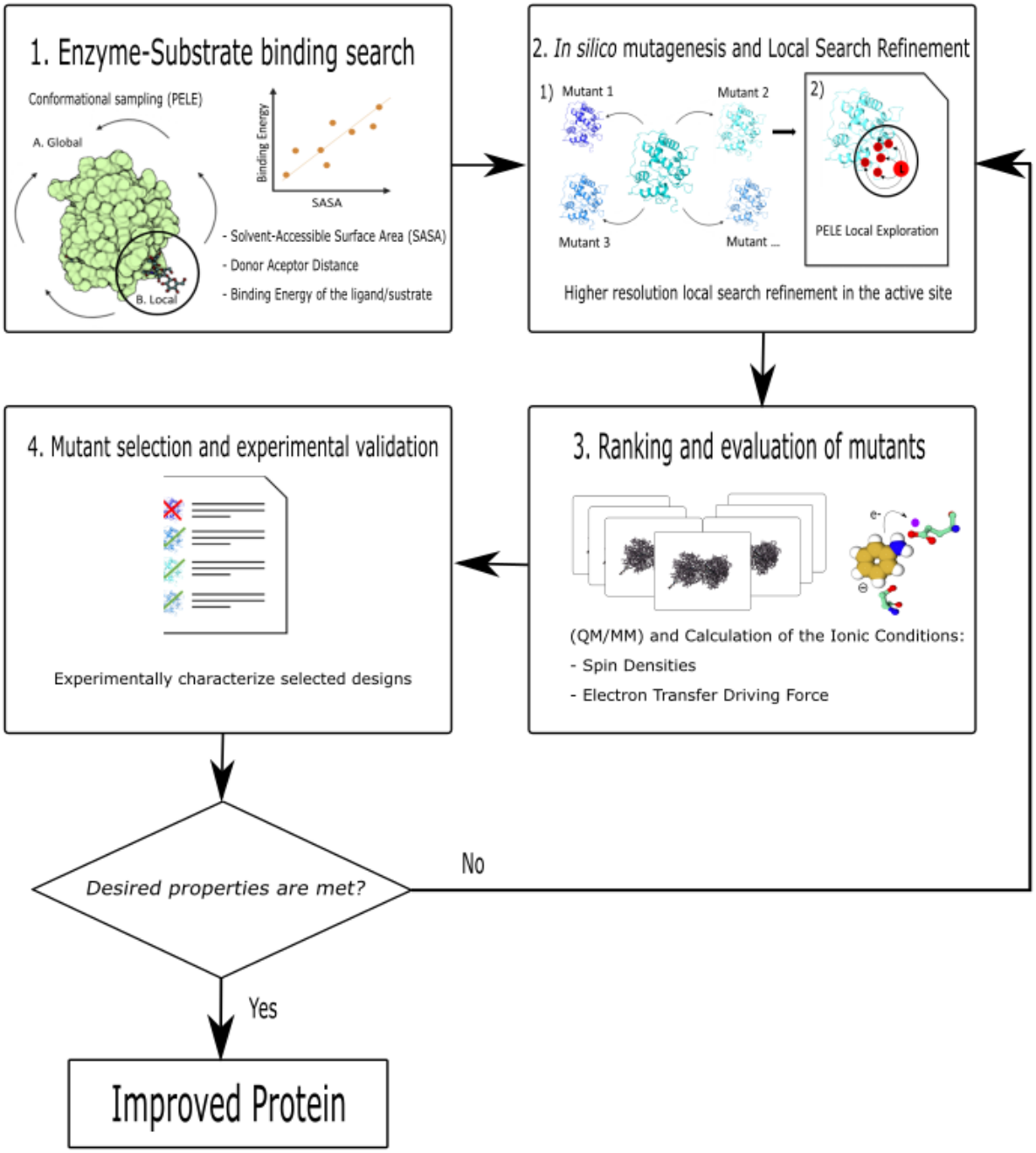


Figure 1. General approach for structural-based rational design of proteins. The structure used as a visual example corresponds with PDB code 5JD4.

Table 1. Collection of all our recent cases of enzyme engineering based on structural-based modeling in our laboratory.

Enzyme/Year	PDB code	Mutations	Approach	Improved property
Lake Arreo Esterase 5 (LAE5) 2021	5JD3	I16V/I92A/W96G	SiteMap & Glide & Local explorations with PELE & MD	Increased substrate promiscuity (4-fold increase)[35]
Marasmius rotula unspecific peroxygenase (*Mro*UPO) 2020	5FUJ	I153F/S156F	MD & Glide & Local explorations with PELE	Increased activity, regioselectivity, and enantioselectivity toward α-linolenic acid (polyunsaturated fatty acids)[30,64]
Pleurotus eryngii aryl alcohol oxidase (*Per*AAO) 2019	3FIM	F501W/I500M	Local explorations with PELE	Enhanced oxidation of chiral benzyl alcohols with high enantioselectivity (160-fold increase in activity)[65,66]
Lake Arreo Esterase 6 (LAE6) 2020/2018	5JD4	L214H/E25D/(R23G)	Global and local explorations with PELE & MD	Added an extra hydrolase site[5,67]
Fungal high-redox-potential laccase (HRPL) 2019	2HRG	A162V/A458L	Glide & Local explorations with PELE & MD	Improved redox potential, acid, and thermal stability (7.5-fold increase in catalytic efficiency, 18-fold increase in tolerance at pH=2, ~2.5-fold increase in thermostability)[31]
Basidiomycete PM1 HRPL 2016	5ANH	N207S/N263D	Local explorations with PELE & QM/MM spin density calculations	Improved aniline oxidation to enhance polyaniline production (2-fold k_{cat} increase)[29]
Highly stable manganese peroxidase (MnP6) 2016	4CZN	G139H/N218H	Local explorations with PELE & QM/MM spin density calculations	Enabled ABTS oxidation with comparable specificity compared to active peroxidases[28]

Toluene 4-monooxygenase (T4MO) 2014	3DHG	F269W	Migration pathway study with PELE	Changes in regiospecificity and activity due to increased radius of one of the channels (2.1-fold increase in activity)[68]
Pleurotus eryngii aryl alcohol oxidase (*Per*AAO) 2012	3FIM	F501A	Local explorations with PELE & QM/MM calculations	Higher stereoselectivity toward secondary aromatic alcohols (~3-fold increase in S/R ratio)[69]

Plurizymes, a new way of thinking catalysis. Let us now turn our review into one of our most significant enzyme engineering successes, the development of plurizymes.

The design of active sites into noncatalytic protein scaffolds has been a compelling idea[6,7,11,42,54,70–78]. However, designing *de novo* active sites has proven to be a difficult task, providing, in most cases, residual activity levels, representing a current challenge for biochemistry[6,42,74,79,80]. This includes the more recent efforts in adding artificial additional active sites, a nascent field of study. The fact that these enzymes might enable one-pot cascade reactions makes them an eco-friendly potential alternative for inorganic catalysts (as well as complex enzyme mixtures) in different industrial sectors.

In our lab, we computationally designed an esterase variant with both native and artificial hydrolase active sites by site-directed mutagenesis of the enzyme's WT sequence[5,67], what we call a plurizyme (using the Latin root pluri- and -zyme from the word enzyme). The design was based on a three-step procedure. It started with an initial search for (noncatalytic) substrate (esters) binding sites using a global exploration of the protein surface with PELE[5,32,67], similar to step 1 defined above. This first step is key: the absence of alternative

substrate binding sites would cease the design of a plurizyme, or require the engineering of a new binding site, involving more mutations and complicating the protocol. Once identified, we introduced different combinations of possible catalytic triads (serine/histidine/aspartate or glutamate) and evaluated them with a local exploration of the substrate around the new putative active site with PELE. At this point we aim for the following: (1) that the substrate remains at the site so the local minimum is preserved, (2) that catalytic poses can be achieved (ester carbon within 3-4 Å from the serine nucleophilic oxygen and catalytic triad H-bonds within reasonable distances and angles), and (3) oxyanion hole stabilization. The tracking of these 3 factors gives us a sense of whether the substrate will bind or not and if the first step of the hydrolysis reaction could occur. Finally, the best ranked catalytic triads are further examined through extensive (~0.5-1 μs) MD simulations to check the structural integrity of the site and the main catalytic distances involved (Figure 2). These MD simulations aid in the ranking of the different variants designed with PELE, prioritizing those for in vitro validation.

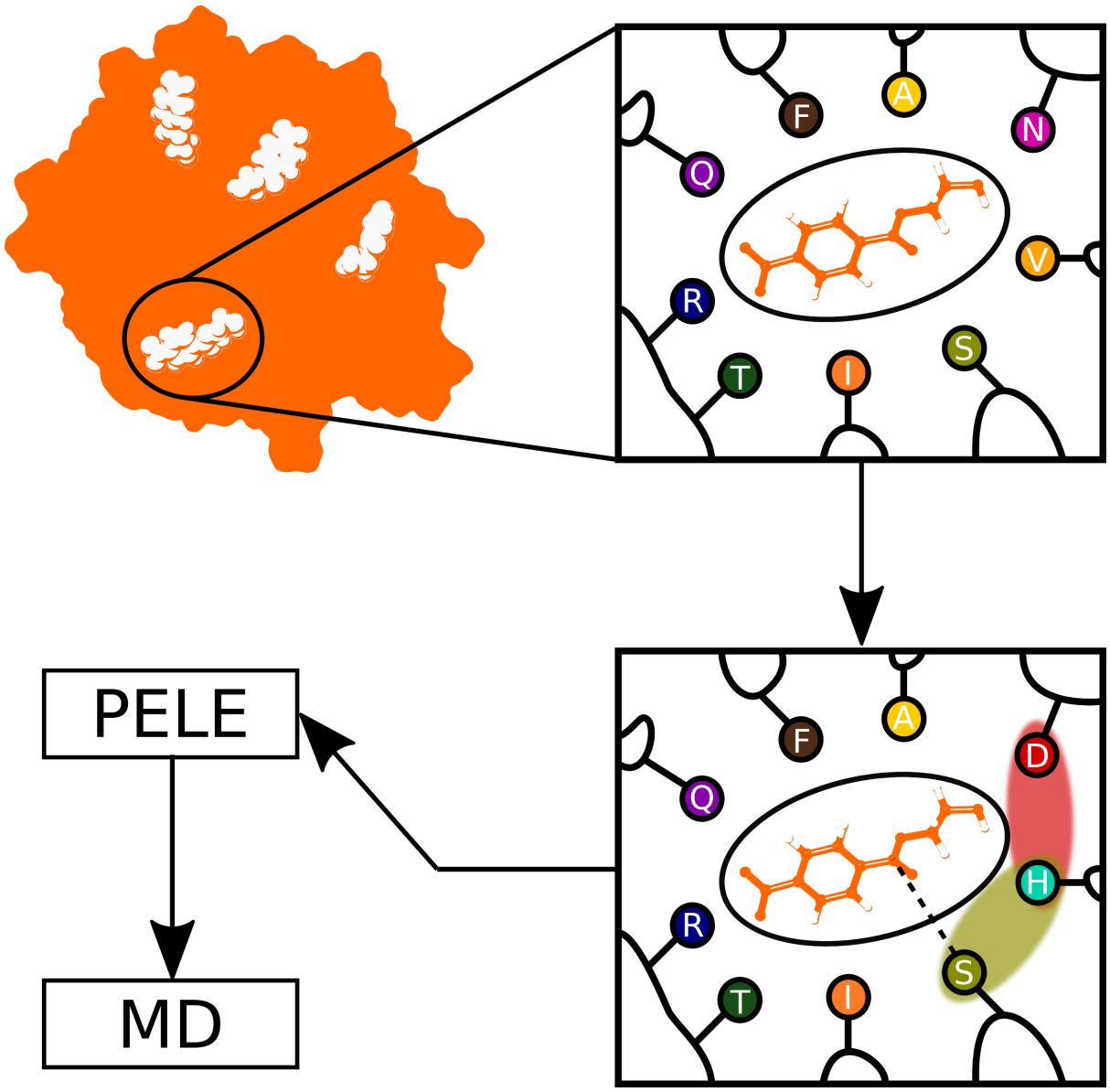


Figure 2. Protocol used in our laboratory to design plurizymes. First, a global exploration of the protein's surface with PELE[5,32,67] to find (noncatalytic) binding sites is performed. Then, (selected) found binding sites are functionalized with one possible combination of a serine/histidine/aspartate (or glutamate) catalytic triad. Finally, the computationally designed variant is tested with PELE and MD simulations. The PDB code of the silhouette of the protein surface is 6SBN. The substrate displayed in the figure is 2-hydroxyethyl terephthalate.

Although there have been other publications related to plurizyme-like systems[73,81–90], the approaches and the final obtained catalyst were different, resulting in several strategies to

tackle a desired cascade process. One of them consisted in the self-assembly of histidine-tyrosine peptides to mimic catalytic microenvironments[73]. Another strategy made use of a noncatalytic protein scaffold (lactococcal multidrug resistance regulator) to add two abiological catalytic sites that could act synergistically[77,87]. Another group was looking for creating plurizymes by coupling small metal nanoparticles with the enzyme, gaining the catalytic properties of the metal-assisted catalyst[88,91–93]. Lastly, a recent publication used an innovative approach to change the native cofactor of an enzyme for a new host that incorporated both a mimic of NADH and a flavin analog, enabling the direct proton and electron transfer between cofactors without being consumed either affected by diffusion[90].

Therefore, several ways to create plurizymes have been performed with ideas and methodology from different fields. Still, the vast majority of the explored strategies are not based on computational predictions to come up with the design. In our lab, we use the power of molecular modeling to design plurizymes[32] and we are currently expanding our approach on other families of enzymes, besides esterases, and even on noncatalytic proteins[94]. In fact, we have one plurizyme being patented, two systems with an added active site experimentally validated waiting to be published, and four more being currently in experimental characterization[95].

Ongoing efforts in plurizyme design automation. As stated, we aim at implementing modular and faster tools, capable of automatic and high throughput enzyme engineering. One of our current working lines is based on combining Rosetta with PELE, which allows the placement of an active site in a putative pocket based on a set of geometrical constraints. In this method, the design process is performed in three steps: design of catalytic residues, design of noncatalytic residues, and a final refinement/analysis. In the first step, all residues in the active site are mutated into Ala and an MC simulation is performed to identify the most

suitable positions for catalytic residues (Ser, His, and Asp in the case of ester hydrolase). In each MC step, one position is mutated into a catalytic residue and the energy of the system, including geometrical constraints, is evaluated using a Metropolis criterion. In addition, to improve the sampling, MC simulations are performed using the adaptive reinforcement learning algorithm designed in our laboratory[96]. In this procedure, several rounds of MC simulations are performed. After each round, the results of MC explorers (from all previous rounds) are clustered based on some metric (e.g. constraint energies) and the next round of MC simulations are performed from the top-ranked clusters. This technique also allows the inclusion of user-defined bias in a simpler manner.

This initial procedure yields a set of possible solutions for the location of catalytic residues which are then used for the placement of the other noncatalytic residues in the second step. Here again, an adaptive MC sampling is performed to identify the noncatalytic residues that are most compatible with the proposed catalytic ones; during this process, the suggested catalytic residues are kept frozen. The design proposals are then optimized by additional side-chain MC simulation to find the most stable packing without imposing any constraints. The overall procedure is performed using the pyrosetta library[97] which allows for fast and efficient implementation of protein engineering algorithms. Finally, the best ranked plurizymes are analyzed with PELE and MD to identify those having the highest affinity for the substrate and exhibiting poses with proper catalytic contacts.

To check the potential of this new computational approach, we performed a retrospective study using the first plurizyme published in our laboratory[5,67], developed from the LAE6 alpha/beta hydrolase from the metagenome of Lake Arreo (Spain). 11 residues around the putative artificial active site, a binding site found in a global PELE exploration, were taken to be mutated, which includes the 3 residues experimentally validated. The catalytic residues

were set to be serine, histidine, and aspartate in order to generate the catalytic triad. 47 active site designs were created for each system. Afterward, the top 10 designs according to their full energies were used for noncatalytic residues design, i.e. the 8 remaining residues of the potential active site. Finally, we performed 4 replicas of 500 ns of apo MD simulations (see Figure 3 caption for details).

In Figure 3, we show the box plot of the main catalytic distances distributions for all 18 variants analyzed (after expanding the top 10 catalytic triad designs with the noncatalytic residues), where we highlight with a red frame the two different mutants validated in our previous studies[5,67]. As it can be seen from Figure 3, the new computational approach successfully recovers the experimentally validated mutants. In addition, we find new mutants that exhibit similar catalytic distances. Importantly, the procedure is fully automatic and easy to implement in a general manner. On the basis of these encouraging results, we aim to develop further the current method to both improve the accuracy and include the substrate in the design procedure.

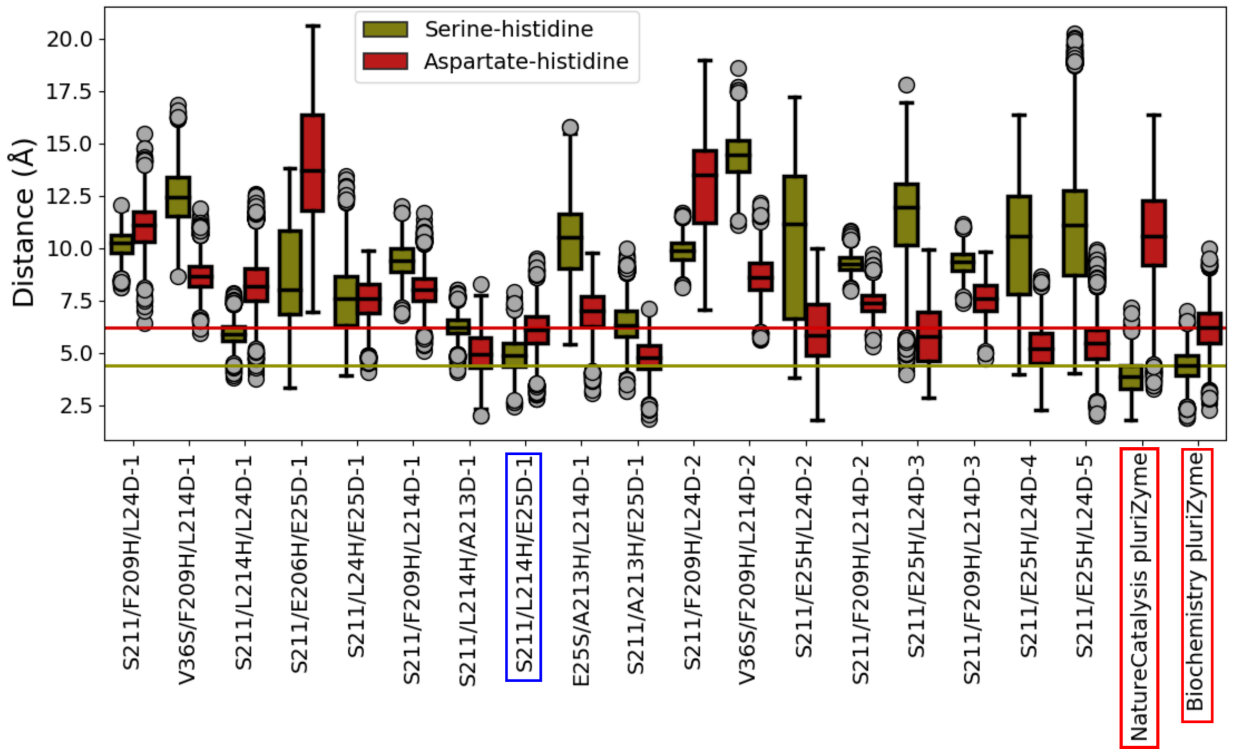


Figure 3. Box plot representing the serine-histidine distance ($d_{H_{\gamma}^{Ser} - N_{\delta/\varepsilon}^{His}}$) and aspartate-histidine distance ($d_{O_{\delta}^{Asp} - H_{\delta/\varepsilon N}^{His}}$) along the 500 ns of the 4 MD replicas performed for all the mutants obtained out of the noncatalytic design of the top catalytic designs. The red frames indicate the experimentally validated mutants published in Nature Catalysis and Biochemistry. The blue frame indicates the catalytic design that recovered the residues used in the experimentally validated ones, but with different noncatalytic residues. The figure was created with the Matplotlib library[98]. MD simulations were performed with OPENMM[99], a TIP3P[100] water box of 8 Å, the AMBER99SB force field[101], Andersen thermostat[102], and MC barostat[103,104]. Production run used an NPT ensemble with a prior NVT equilibration for ~400 ps and a constraint of 10 kcal/(mol·A^2) followed by a 1ns NPT equilibration using a milder constraint of 5 kcal/(mol·A^2).

CONCLUSIONS

With this mini-review article, we aim at demonstrating the potential of our structural-based bioinformatics and molecular modeling techniques in the field of enzyme engineering. The recent development of improved algorithms along with the vast computational resources available nowadays provides an excellent prediction/success ratio. We believe that this performance should motivate an early *in silico* phase in most enzyme engineering studies, similar to what is widely accepted in, for example, drug design. Moreover, we demonstrate the potential of some new techniques combining Rosetta and PELE which could provide a faster and more automated procedure, an essential aspect for broader use.

AUTHOR INFORMATION

Corresponding Author

*E-mail V.G.: victor.guallar@bsc.es.

ORCID

Sergi Roda: 0000-0002-0174-7435

Ana Robles-Martín: 0000-0002-6377-6338

Masoud Kazemi: 0000-0002-0750-8865

Victor Guallar: 0000-0003-3274-2482

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work has also been supported by predoctoral fellowships FPU19/00608 and PRE2020-091825, and the PID2019-106370RB-I00/AEI/10.13039/501100011033 grant from the Spanish Ministry of Science and Innovation.

ABBREVIATIONS

PELE, protein energy landscape exploration; SASA, solvent-accessible surface area; MD, molecular dynamics; WT, wild type; RMSD, root mean square deviation

REFERENCES

(1) Kamble, A.; Srinivasan, S.; Singh, H. In-Silico Bioprospecting: Finding Better Enzymes. *Mol. Biotechnol.* **2019**, *61* (1), 53–59.
(2) Ferrer, M.; Martínez-Martínez, M.; Bargiela, R.; Streit, W. R.; Golyshina, O. V.; Golyshin, P. N. Estimating the Success of Enzyme Bioprospecting through Metagenomics: Current Status and Future Trends. *Microb. Biotechnol.* **2016**, *9* (1), 22–34.
(3) Chen, R. Enzyme Engineering: Rational Redesign versus Directed Evolution. *Trends Biotechnol.* **2001**, *19* (1), 13–14.
(4) Jemli, S.; Ayadi-Zouari, D.; Hlima, H. B.; Bejar, S. Biocatalysts: Application and Engineering for Industrial Purposes. *Crit. Rev. Biotechnol.* **2016**, *36* (2), 246–258.
(5) Alonso, S.; Santiago, G.; Cea-Rama, I.; Fernandez-Lopez, L.; Coscolín, C.; Modregger, J.; Ressmann, A. K.; Martínez-Martínez, M.; Marrero, H.; Bargiela, R.; et al. Genetically Engineered Proteins with Two Active Sites for Enhanced Biocatalysis and Synergistic Chemo- and Biocatalysis. *Nature Catalysis* **2020**, *3* (3), 319–328.
(6) Richter, F.; Blomberg, R.; Khare, S. D.; Kiss, G.; Kuzin, A. P.; Smith, A. J. T.; Gallaher, J.; Pianowski, Z.; Helgeson, R. C.; Grjasnow, A.; et al. Computational Design of Catalytic Dyads and Oxyanion Holes for Ester Hydrolysis. *J. Am. Chem. Soc.* **2012**, *134* (39), 16197–16206.
(7) Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; et al. Kemp Elimination Catalysts by Computational Enzyme Design. *Nature* **2008**, *453* (7192), 190–195.
(8) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; et al. Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science* **2010**, *329* (5989), 309–313.
(9) Nanda, V.; Koder, R. L. Designing Artificial Enzymes by Intuition and Computation. *Nat. Chem.* **2010**, *2* (1), 15–24.
(10) Barrozo, A.; Borstnar, R.; Marloie, G.; Kamerlin, S. C. L. Computational Protein Engineering: Bridging the Gap between Rational Design and Laboratory Evolution. *Int. J. Mol. Sci.* **2012**, *13* (10), 12428–12460.

(11) Faiella, M.; Andreozzi, C.; de Rosales, R. T. M.; Pavone, V.; Maglio, O.; Nastri, F.; DeGrado, W. F.; Lombardi, A. An Artificial Di-Iron Oxo-Protein with Phenol Oxidase Activity. *Nat. Chem. Biol.* **2009**, *5* (12), 882–884.
(12) Mazurenko, S.; Prokop, Z.; Damborsky, J. Machine Learning in Enzyme Engineering. *ACS Catal.* **2020**, *10* (2), 1210–1223.
(13) Rohl, C. A.; Strauss, C. E. M.; Misura, K. M. S.; Baker, D. Protein Structure Prediction Using Rosetta. In *Methods in Enzymology*; Academic Press, 2004; Vol. 383, pp 66–93.
(14) Raman, S.; Vernon, R.; Thompson, J.; Tyka, M.; Sadreyev, R.; Pei, J.; Kim, D.; Kellogg, E.; DiMaio, F.; Lange, O.; et al. Structure Prediction for CASP8 with All-Atom Refinement Using Rosetta. *Proteins* **2009**, *77 Suppl 9*, 89–99.
(15) Desmet, J.; De Maeyer, M.; Hazes, B.; Lasters, I. The Dead-End Elimination Theorem and Its Use in Protein Side-Chain Positioning. *Nature* **1992**, *356* (6369), 539–542.
(16) Ross, S. A.; Sarisky, C. A.; Su, A.; Mayo, S. L. Designed Protein G Core Variants Fold to Native-like Structures: Sequence Selection by ORBIT Tolerates Variation in Backbone Specification. *Protein Sci.* **2001**, *10* (2), 450–454.
(17) Hallen, M. A.; Martin, J. W.; Ojewole, A.; Jou, J. D.; Lowegard, A. U.; Frenkel, M. S.; Gainza, P.; Nisonoff, H. M.; Mukund, A.; Wang, S.; et al. OSPREY 3.0: Open-Source Protein Redesign for You, with Powerful New Features. *J. Comput. Chem.* **2018**, *39* (30), 2494–2507.
(18) Kuipers, R. K.; Joosten, H.-J.; van Berkel, W. J. H.; Leferink, N. G. H.; Rooijen, E.; Ittmann, E.; van Zimmeren, F.; Jochens, H.; Bornscheuer, U.; Vriend, G.; et al. 3DM: Systematic Analysis of Heterogeneous Superfamily Data to Discover Protein Functionalities. *Proteins* **2010**, *78* (9), 2101–2113.
(19) Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX Web Server: An Online Force Field. *Nucleic Acids Res.* **2005**, *33* (Web Server issue), W382–W388.
(20) Kotev, M.; Lecina, D.; Tarragó, T.; Giralt, E.; Guallar, V. Unveiling Prolyl Oligopeptidase Ligand Migration by Comprehensive Computational Techniques. *Biophys. J.* **2015**, *108* (1), 116–125.
(21) Borrelli, K. W.; Vitalis, A.; Alcantara, R.; Guallar, V. PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique. *J. Chem. Theory Comput.* **2005**, *1* (6), 1304–1311.
(22) Municoy, M.; Roda, S.; Soler, D.; Soutullo, A.; Guallar, V. aquaPELE: A Monte Carlo-Based Algorithm to Sample the Effects of Buried Water Molecules in Proteins. *J. Chem. Theory Comput.* **2020**, *16* (12), 7655–7670.
(23) van der Kamp, M. W.; Mulholland, A. J. Combined Quantum Mechanics/molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52* (16), 2708–2728.
(24) Wallrapp, F. H.; Voityuk, A. A.; Guallar, V. In-Silico Assessment of Protein-Protein Electron Transfer. A Case Study: Cytochrome c Peroxidase – Cytochrome c. *PLoS Comput. Biol.* **2013**, *9* (3), e1002990.
(25) Prytkova, T. R.; Kurnikov, I. V.; Beratan, D. N. Coupling Coherence Distinguishes Structure Sensitivity in Protein Electron Transfer. *Science* **2007**, *315* (5812), 622–625.
(26) Sumbalova, L.; Stourac, J.; Martinek, T.; Bednar, D.; Damborsky, J. HotSpot Wizard 3.0: Web Server for Automated Design of Mutations and Smart Libraries Based on Sequence Input Information. *Nucleic Acids Res.* **2018**, *46* (W1), W356–W362.
(27) Fernández-Fueyo, E.; Ruiz-Dueñas, F. J.; Martínez, M. J.; Romero, A.; Hammel, K. E.; Medrano, F. J.; Martínez, A. T. Ligninolytic Peroxidase Genes in the Oyster Mushroom

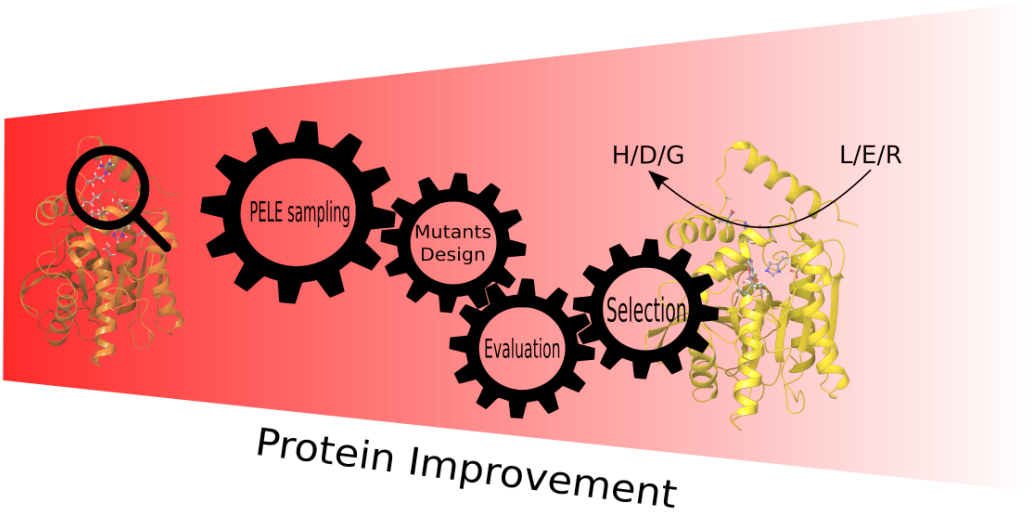
Genome: Heterologous Expression, Molecular Structure, Catalytic and Stability Properties, and Lignin-Degrading Ability. *Biotechnol. Biofuels* **2014**, *7* (1), 2.
(28) Acebes, S.; Fernandez-Fueyo, E.; Monza, E.; Lucas, M. F.; Almendral, D.; Ruiz-Dueñas, F. J.; Lund, H.; Martinez, A. T.; Guallar, V. Rational Enzyme Engineering Through Biophysical and Biochemical Modeling. *ACS Catal.* **2016**, *6* (3), 1624–1629.
(29) Santiago, G.; de Salas, F.; Lucas, M. F.; Monza, E.; Acebes, S.; Martinez, Á. T.; Camarero, S.; Guallar, V. Computer-Aided Laccase Engineering: Toward Biological Oxidation of Arylamines. *ACS Catal.* **2016**, *6* (8), 5415–5423.
(30) Carro, J.; González-Benjumea, A.; Fernández-Fueyo, E.; Aranda, C.; Guallar, V.; Gutiérrez, A.; Martínez, A. T. Modulating Fatty Acid Epoxidation vs Hydroxylation in a Fungal Peroxygenase. *ACS Catal.* **2019**, *9* (7), 6234–6242.
(31) Mateljak, I.; Monza, E.; Lucas, M. F.; Guallar, V.; Aleksejeva, O.; Ludwig, R.; Leech, D.; Shleev, S.; Alcalde, M. Increasing Redox Potential, Redox Mediator Activity, and Stability in a Fungal Laccase by Computer-Guided Mutagenesis and Directed Evolution. *ACS Catal.* **2019**, *9* (5), 4561–4572.
(32) Roda, S.; Santiago, G.; Guallar, V. Mapping Enzyme-Substrate Interactions: Its Potential to Study the Mechanism of Enzymes. In *Advances in Protein Chemistry and Structural Biology*; Academic Press, 2020; Vol. 122, pp 1–31.
(33) Martínez-Martínez, M.; Coscolín, C.; Santiago, G.; Chow, J.; Stogios, P. J.; Bargiela, R.; Gertler, C.; Navarro-Fernández, J.; Bollinger, A.; Thies, S.; et al. Determinants and Prediction of Esterase Substrate Promiscuity Patterns. *ACS Chem. Biol.* **2018**, *13* (1), 225–234.
(34) Xiang, R. Machine Learning Based Prediction of Esterases' Promiscuity. **2020**.
(35) Roda, S.; Fernandez-Lopez, L.; Cañadas, R.; Santiago, G.; Ferrer, M.; Guallar, V. Computationally Driven Rational Design of Substrate Promiscuity on Serine Ester Hydrolases. *ACS Catal.* **2021**, *11* (6), 3590–3601.
(36) Ashworth, J.; Havranek, J. J.; Duarte, C. M.; Sussman, D.; Monnat, R. J., Jr; Stoddard, B. L.; Baker, D. Computational Redesign of Endonuclease DNA Binding and Cleavage Specificity. *Nature* **2006**, *441* (7093), 656–659.
(37) Korkegian, A.; Black, M. E.; Baker, D.; Stoddard, B. L. Computational Thermostabilization of an Enzyme. *Science* **2005**, *308* (5723), 857–860.
(38) Chevalier, B. S.; Kortemme, T.; Chadsey, M. S.; Baker, D.; Monnat, R. J.; Stoddard, B. L. Design, Activity, and Structure of a Highly Specific Artificial Endonuclease. *Mol. Cell* **2002**, *10* (4), 895–905.
(39) Wijma, H. J.; Floor, R. J.; Jekel, P. A.; Baker, D.; Marrink, S. J.; Janssen, D. B. Computationally Designed Libraries for Rapid Enzyme Stabilization. *Protein Eng. Des. Sel.* **2014**, *27* (2), 49–58.
(40) Heinisch, T.; Pellizzoni, M.; Dürrenberger, M.; Tinberg, C. E.; Köhler, V.; Klehr, J.; Häussinger, D.; Baker, D.; Ward, T. R. Improving the Catalytic Performance of an Artificial Metalloenzyme by Computational Design. *J. Am. Chem. Soc.* **2015**, *137* (32), 10414–10419.
(41) Wijma, H. J.; Floor, R. J.; Bjelic, S.; Marrink, S. J.; Baker, D.; Janssen, D. B. Enantioselective Enzymes by Computational Design and in Silico Screening. *Angew. Chem. Int. Ed Engl.* **2015**, *54* (12), 3726–3730.
(42) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Rothlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F.; et al. De Novo Computational Design of Retro-Aldol Enzymes. *Science* **2008**, *319* (5868), 1387–1391.
(43) Richter, F.; Leaver-Fay, A.; Khare, S. D.; Bjelic, S.; Baker, D. De Novo Enzyme Design

Using Rosetta3. *PLoS One* **2011**, *6* (5), e19230.
(44) Kim, H. J.; Ruszczycky, M. W.; Choi, S.-H.; Liu, Y.-N.; Liu, H.-W. Enzyme-Catalysed [4+2] Cycloaddition Is a Key Step in the Biosynthesis of Spinosyn A. *Nature* **2011**, *473* (7345), 109–112.
(45) Miao, Y.; Metzner, R.; Asano, Y. Kemp Elimination Catalyzed by Naturally Occurring Aldoxime Dehydratases. *Chembiochem* **2017**, *18* (5), 451–454.
(46) Khersonsky, O.; Lipsh, R.; Avizemer, Z.; Ashani, Y.; Goldsmith, M.; Leader, H.; Dym, O.; Rogotner, S.; Trudeau, D. L.; Prilusky, J.; et al. Automated Design of Efficient and Functionally Diverse Enzyme Repertoires. *Mol. Cell* **2018**, *72* (1), 178–186.e5.
(47) Khersonsky, O.; Fleishman, S. J. Why Reinvent the Wheel? Building New Proteins Based on Ready-Made Parts. *Protein Sci.* **2016**, *25* (7), 1179–1187.
(48) Goldenzweig, A.; Fleishman, S. J. Principles of Protein Stability and Their Application in Computational Design. *Annu. Rev. Biochem.* **2018**, *87*, 105–129.
(49) Trudeau, D. L.; Edlich-Muth, C.; Zarzycki, J.; Scheffen, M.; Goldsmith, M.; Khersonsky, O.; Avizemer, Z.; Fleishman, S. J.; Cotton, C. A. R.; Erb, T. J.; et al. Design and in Vitro Realization of Carbon-Conserving Photorespiration. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (49), E11455–E11464.
(50) Weinstein, J.; Khersonsky, O.; Fleishman, S. J. Practically Useful Protein-Design Methods Combining Phylogenetic and Atomistic Calculations. *Curr. Opin. Struct. Biol.* **2020**, *63*, 58–64.
(51) VanDrisse, C. M.; Lipsh-Sokolik, R.; Khersonsky, O.; Fleishman, S. J.; Newman, D. K. Computationally Designed Pyocyanin Demethylase Acts Synergistically with Tobramycin to Kill Recalcitrant Biofilms. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118* (12). https://doi.org/10.1073/pnas.2022012118.
(52) Petrović, D.; Risso, V. A.; Kamerlin, S. C. L.; Sanchez-Ruiz, J. M. Conformational Dynamics and Enzyme Evolution. *J. R. Soc. Interface* **2018**, *15* (144). https://doi.org/10.1098/rsif.2018.0330.
(53) Carvalho, A. T. P.; Barrozo, A.; Doron, D.; Kilshtain, A. V.; Major, D. T.; Kamerlin, S. C. L. Challenges in Computational Studies of Enzyme Structure, Function and Dynamics. *J. Mol. Graph. Model.* **2014**, *54*, 62–79.
(54) Risso, V. A.; Martinez-Rodriguez, S.; Candel, A. M.; Krüger, D. M.; Pantoja-Uceda, D.; Ortega-Muñoz, M.; Santoyo-Gonzalez, F.; Gaucher, E. A.; Kamerlin, S. C. L.; Bruix, M.; et al. De Novo Active Sites for Resurrected Precambrian Enzymes. *Nat. Commun.* **2017**, *8*, 16113.
(55) Amrein, B. A.; Bauer, P.; Duarte, F.; Janfalk Carlsson, Å.; Naworyta, A.; Mowbray, S. L.; Widersten, M.; Kamerlin, S. C. L. Expanding the Catalytic Triad in Epoxide Hydrolases and Related Enzymes. *ACS Catal.* **2015**, *5* (10), 5702–5713.
(56) Hong, N.-S.; Petrović, D.; Lee, R.; Gryn'ova, G.; Purg, M.; Saunders, J.; Bauer, P.; Carr, P. D.; Lin, C.-Y.; Mabbitt, P. D.; et al. The Evolution of Multiple Active Site Configurations in a Designed Enzyme. *Nat. Commun.* **2018**, *9* (1), 3900.
(57) Shen, R.; Crean, R. M.; Johnson, S. J.; Kamerlin, S. C. L.; Hengge, A. C. Single Residue on the WPD-Loop Affects the pH Dependency of Catalysis in Protein Tyrosine Phosphatases. *JACS Au* **2021**, No. jacsau.1c00054. https://doi.org/10.1021/jacsau.1c00054.
(58) Crean, R. M.; Gardner, J. M.; Kamerlin, S. C. L. Harnessing Conformational Plasticity to Generate Designer Enzymes. *J. Am. Chem. Soc.* **2020**, *142* (26), 11324–11342.
(59) Wijma, H. J.; Fürst, M. J. L. J.; Janssen, D. B. A Computational Library Design Protocol for Rapid Improvement of Protein Stability: FRESCO. *Methods Mol. Biol.* **2018**, *1685*,

69–85.
(60) van Beek, H. L.; Wijma, H. J.; Fromont, L.; Janssen, D. B.; Fraaije, M. W. Stabilization of Cyclohexanone Monooxygenase by a Computationally Designed Disulfide Bond Spanning Only One Residue. *FEBS Open Bio* **2014**, *4*, 168–174.
(61) Floor, R. J.; Wijma, H. J.; Colpa, D. I.; Ramos-Silva, A.; Jekel, P. A.; Szymański, W.; Feringa, B. L.; Marrink, S. J.; Janssen, D. B. Computational Library Design for Increasing Haloalkane Dehalogenase Stability. *Chembiochem* **2014**, *15* (11), 1660–1672.
(62) Wu, B.; Wijma, H. J.; Song, L.; Rozeboom, H. J.; Poloni, C.; Tian, Y.; Arif, M. I.; Nuijens, T.; Quaedflieg, P. J. L. M.; Szymanski, W.; et al. Versatile Peptide C-Terminal Functionalization via a Computationally Engineered Peptide Amidase. *ACS Catal.* **2016**, *6* (8), 5405–5414.
(63) Li, R.; Wijma, H. J.; Song, L.; Cui, Y.; Otzen, M.; Tian, Y. 'e; Du, J.; Li, T.; Niu, D.; Chen, Y.; et al. Computational Redesign of Enzymes for Regio- and Enantioselective Hydroamination. *Nat. Chem. Biol.* **2018**, *14* (7), 664–670.
(64) Municoy, M.; González-Benjumea, A.; Carro, J.; Aranda, C.; Linde, D.; Renau-Mínguez, C.; Ullrich, R.; Hofrichter, M.; Guallar, V.; Gutiérrez, A.; et al. Fatty-Acid Oxygenation by Fungal Peroxygenases: From Computational Simulations to Preparative Regio- and Stereoselective Epoxidation. *ACS Catal.* **2020**, *10* (22), 13584–13595.
(65) Viña‐Gonzalez, J.; Jimenez‐Lalana, D.; Sancho, F.; Serrano, A.; Martinez, A. T.; Guallar, V.; Alcalde, M. Structure‐guided Evolution of Aryl Alcohol Oxidase from Pleurotus Eryngii for the Selective Oxidation of Secondary Benzyl Alcohols. *Adv. Synth. Catal.* **2019**, *361* (11), 2514–2525.
(66) Serrano, A.; Sancho, F.; Viña-González, J.; Carro, J.; Alcalde, M.; Guallar, V.; Martínez, A. T. Switching the Substrate Preference of Fungal Aryl-Alcohol Oxidase: Towards Stereoselective Oxidation of Secondary Benzyl Alcohols. *Catal. Sci. Technol.* **2019**, *9* (3), 833–841.
(67) Santiago, G.; Martínez-Martínez, M.; Alonso, S.; Bargiela, R.; Coscolín, C.; Golyshin, P. N.; Guallar, V.; Ferrer, M. Rational Engineering of Multiple Active Sites in an Ester Hydrolase. *Biochemistry* **2018**, *57* (15), 2245–2255.
(68) Hosseini, A.; Brouk, M.; Lucas, M. F.; Glaser, F.; Fishman, A.; Guallar, V. Atomic Picture of Ligand Migration in Toluene 4-Monooxygenase. *J. Phys. Chem. B* **2015**, *119* (3), 671–678.
(69) Hernández-Ortega, A.; Ferreira, P.; Merino, P.; Medina, M.; Guallar, V.; Martínez, A. T. Stereoselective Hydride Transfer by Aryl-Alcohol Oxidase, a Member of the GMC Superfamily. *Chembiochem* **2012**, *13* (3), 427–435.
(70) Seelig, B.; Szostak, J. W. Selection and Evolution of Enzymes from a Partially Randomized Non-Catalytic Scaffold. *Nature* **2007**, *448* (7155), 828–831.
(71) Rufo, C. M.; Moroz, Y. S.; Moroz, O. V.; Stöhr, J.; Smith, T. A.; Hu, X.; DeGrado, W. F.; Korendovych, I. V. Short Peptides Self-Assemble to Produce Catalytic Amyloids. *Nat. Chem.* **2014**, *6* (4), 303–309.
(72) Moroz, Y. S.; Dunston, T. T.; Makhlynets, O. V.; Moroz, O. V.; Wu, Y.; Yoon, J. H.; Olsen, A. B.; McLaughlin, J. M.; Mack, K. L.; Gosavi, P. M.; et al. New Tricks for Old Proteins: Single Mutations in a Nonenzymatic Protein Give Rise to Various Enzymatic Activities. *JACS* **2015**, *137* (47), 14905–14911.
(73) Díaz-Caballero, M.; Navarro, S.; Nuez-Martínez, M.; Peccati, F.; Rodríguez-Santiago, L.; Sodupe, M.; Teixidor, F.; Ventura, S. pH-Responsive Self-Assembly of Amyloid Fibrils for Dual Hydrolase-Oxidase Reactions. *ACS Catalysis* **2020**, *11* (2), 595–607.

(74) Jeong, W. J.; Yu, J.; Song, W. J. Proteins as Diverse, Efficient, and Evolvable Scaffolds for Artificial Metalloenzymes. *Chem. Commun.* **2020**, *56* (67), 9586–9599.
(75) Leveson-Gower, R. B.; Mayer, C.; Roelfes, G. The Importance of Catalytic Promiscuity for Enzyme Design and Evolution. *Nature Reviews Chemistry* **2019**, *3* (12), 687–705.
(76) Roelfes, G. LmrR: A Privileged Scaffold for Artificial Metalloenzymes. *Acc. Chem. Res.* **2019**, *52* (3), 545–556.
(77) Bos, J.; Browne, W. R.; Driessen, A. J. M.; Roelfes, G. Supramolecular Assembly of Artificial Metalloenzymes Based on the Dimeric Protein LmrR as Promiscuous Scaffold. *J. Am. Chem. Soc.* **2015**, *137* (31), 9796–9799.
(78) Der, B. S.; Edwards, D. R.; Kuhlman, B. Catalysis by a de Novo Zinc-Mediated Protein Interface: Implications for Natural Enzyme Evolution and Rational Enzyme Engineering. *Biochemistry* **2012**, *51* (18), 3933–3940.
(79) Palomo, J. M. Artificial Enzymes with Multiple Active Sites. *Current Opinion in Green and Sustainable Chemistry* **2021**, *29*, 100452.
(80) Large, B.; G. Baranska, N.; L. Booth, R.; S. Wilson, K.; Duhme-Klair, A.-K. Artificial Metalloenzymes: The Powerful Alliance between Protein Scaffolds and Organometallic Catalysts. *Current Opinion in Green and Sustainable Chemistry* **2021**, *28*, 100420.
(81) Bos, J.; Fusetti, F.; Driessen, A. J. M.; Roelfes, G. Enantioselective Artificial Metalloenzymes by Creation of a Novel Active Site at the Protein Dimer Interface. *Angew. Chem. Int. Ed Engl.* **2012**, *51* (30), 7472–7475.
(82) Lin, Y.-W.; Nagao, S.; Zhang, M.; Shomura, Y.; Higuchi, Y.; Hirota, S. Rational Design of Heterodimeric Protein Using Domain Swapping for Myoglobin. *Angew. Chem. Int. Ed Engl.* **2015**, *54* (2), 511–515.
(83) Farid, T. A.; Kodali, G.; Solomon, L. A.; Lichtenstein, B. R.; Sheehan, M. M.; Fry, B. A.; Bialas, C.; Ennist, N. M.; Siedlecki, J. A.; Zhao, Z.; et al. Elementary Tetrahelical Protein Design for Diverse Oxidoreductase Functions. *Nat. Chem. Biol.* **2013**, *9* (12), 826–833.
(84) Roy, A.; Sarrou, I.; Vaughn, M. D.; Astashkin, A. V.; Ghirlanda, G. De Novo Design of an Artificial bis[4Fe-4S] Binding Protein. *Biochemistry* **2013**, *52* (43), 7586–7594.
(85) Tebo, A. G.; Pecoraro, V. L. Artificial Metalloenzymes Derived from Three-Helix Bundles. *Curr. Opin. Chem. Biol.* **2015**, *25*, 65–70.
(86) Filice, M.; Romero, O.; Gutiérrez-Fernández, J.; de Las Rivas, B.; Hermoso, J. A.; Palomo, J. M. Synthesis of a Heterogeneous Artificial Metallolipase with Chimeric Catalytic Activity. *Chem. Commun.* **2015**, *51* (45), 9324–9327.
(87) Zhou, Z.; Roelfes, G. Synergistic Catalysis in an Artificial Enzyme by Simultaneous Action of Two Abiological Catalytic Sites. *Nature Catalysis* **2020**, *3* (3), 289–294.
(88) Palomo, J. M. Nanobiohybrids: A New Concept for Metal Nanoparticles Synthesis. *Chem. Commun.* **2019**, *55* (65), 9583–9589.
(89) Filice, M.; Losada-Garcia, N.; Perez-Rizquez, C.; Marciello, M.; Morales, M. del P.; Palomo, J. M. Palladium-Nanoparticles Biohybrids in Applied Chemistry. *Applied Nano* **2020**, *2* (1), 1–13.
(90) Zhao, L.; Cai, J.; Li, Y.; Wei, J.; Duan, C. A Host-Guest Approach to Combining Enzymatic and Artificial Catalysis for Catalyzing Biomimetic Monooxygenation. *Nat. Commun.* **2020**, *11* (1), 2903.
(91) Filice, M.; Marciello, M.; Morales, M. del P.; Palomo, J. M. Synthesis of Heterogeneous Enzyme-Metal Nanoparticle Biohybrids in Aqueous Media and Their Applications in C-C Bond Formation and Tandem Catalysis. *Chem. Commun.* **2013**, *49* (61), 6876–6878.

(92) Benavente, R.; Lopez-Tejedor, D.; Del Puerto Morales, M.; Perez-Rizquez, C.; Palomo, J. M. The Enzyme-Induced Formation of Iron Hybrid Nanostructures with Different Morphologies. *Nanoscale* **2020**, *12* (24), 12917–12927.
(93) Benavente, R.; Lopez-Tejedor, D.; Palomo, J. M. Synthesis of a Superparamagnetic Ultrathin $FeCO_3$ Nanorods-Enzyme Bionanohybrid as a Novel Heterogeneous Catalyst. *Chem. Commun.* **2018**, *54* (49), 6256–6259.
(94) Domingo-Cabasés, M.; Rodà, S.; Santiago, G.; Guallar, V. PluriZymes: New Enzymes for New Times. In *Book of abstracts*; Barcelona Supercomputing Center, 2019; pp 81–82.
(95) Rodà, S.; Santiago, G.; Guallar, V. Towards PET Degradation Engineering. In *Book of abstracts*; Barcelona Supercomputing Center, 2019; pp 67–68.
(96) Lecina, D.; Gilabert, J. F.; Guallar, V. Adaptive Simulations, towards Interactive Protein-Ligand Modeling. *Sci. Rep.* **2017**, *7* (1), 8466.
(97) Chaudhury, S.; Lyskov, S.; Gray, J. J. PyRosetta: A Script-Based Interface for Implementing Molecular Modeling Algorithms Using Rosetta. *Bioinformatics* **2010**, *26* (5), 689–691.
(98) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **2007**, *9* (3), 90–95.
(99) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; et al. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLoS Comput. Biol.* **2017**, *13* (7), e1005659.
(100) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *The Journal of Chemical Physics* **1983**, *79* (2), 926–935.
(101) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65* (3), 712–725.
(102) Andersen, H. C. Molecular Dynamics Simulations at Constant Pressure And/or Temperature. *The Journal of Chemical Physics* **1980**, *72* (4), 2384–2393.
(103) Chow, K.-H.; Ferguson, D. M. Isothermal-Isobaric Molecular Dynamics Simulations with Monte Carlo Volume Sampling. *Computer Physics Communications* **1995**, *91* (1-3), 283–289.
(104) Åqvist, J.; Wennerström, P.; Nervall, M.; Bjelic, S.; Brandsdal, B. O. Molecular Dynamics Simulations of Water and Biomolecules with a Monte Carlo Constant Pressure Algorithm. *Chemical Physics Letters* **2004**, *384* (4-6), 288–294.


PELE sampling
Mutants Design
Evaluation
Selection
H/D/G
L/E/R
Protein Improvement

TOC Graphic