

**Biocatalysis**

# Reversing the Enantioselectivity of Enzymatic Carbene N–H Insertion Through Mechanism-Guided Protein Engineering\*

 Carla Calvó-Tusell<sup>†</sup>, Zhen Liu<sup>†,\*</sup>, Kai Chen, Frances H. Arnold, and Marc Garcia-Borràs\*

**Abstract:** We report a computationally driven approach to access enantiodivergent enzymatic carbene N–H insertions catalyzed by P411 enzymes. Computational modeling was employed to rationally guide engineering efforts to control the accessible conformations of a key lactone-carbene (LAC) intermediate in the enzyme active site by installing a new H-bond anchoring point. This H-bonding interaction controls the relative orientation of the reactive carbene intermediate, orienting it for an enantioselective *N*-nucleophilic attack by the amine substrate. By combining MD simulations and site-saturation mutagenesis and screening targeted to only two key residues, we were able to reverse the stereoselectivity of previously engineered *S*-selective P411 enzymes. The resulting variant, **L5\_FL-B3**, accepts a broad scope of amine substrates for N–H insertion with excellent yields (up to >99%), high efficiency (up to 12300 TTN), and good enantiocontrol (up to 7:93 *er*).

## Introduction

New-to-nature biocatalytic reactions have emerged recently as powerful tools for organic synthesis and drug development.<sup>[1–4]</sup> Hemeproteins have been repurposed and engineered to catalyze abiological transformations involving metal carbenes, nitrenes, and other reactive intermediates for constructing valuable chiral molecules.<sup>[5–12]</sup> Although many of these enzymatic systems are optimized to obtain excellent stereoselectivities, achieving enantiodivergent biocatalytic reactions to access each stereoisomer of a chiral product has been challenging and has only been achieved in a few cases (for selected examples, see refs. [7,8,13–19]). Typically, it required intensive protein engineering effort to invert the enantio-preference of an enzyme or employed a different protein scaffold.<sup>[7,19]</sup> Less labor-intensive approaches to alter the stereoselectivity of these new-to-nature enzymes are highly desirable.

Chiral amines are important molecules for organic chemistry and biochemistry.<sup>[20,21]</sup> Expanding the repertoire of C–N bond forming enzymes is valuable for preparing those structures.<sup>[6,22]</sup> We previously identified a series of dual-functional “carbene transferases” based on P411 enzymes (cytochromes P450 with serine as the heme-ligating residue) that induce nucleophilic attack of amines to a lactone-carbene (LAC) intermediate and promote subsequent stereoselective protonation to afford chiral amines with high efficiency and enantioselectivity (Figure 1a).<sup>[23,24]</sup> The final variants **L6\_FL** and **L7\_FL** were found to perform highly efficient carbene N–H insertion toward the *S*-enantiomer. Along the enzyme lineage, a dramatic boost in enantioselectivity (from 40:60 *er* with **L5\_FL** to 97:3 *er* with **L6\_FL** for *N*-methyl aniline) was observed after the introduction of a single mutation, A264S. Computational modeling revealed that the newly introduced serine residue at position 264 establishes hydrogen-bonding interactions between the residue side chain and the lactone ester group. Those interactions stabilize the LAC intermediate to explore a major orientation, which is key to the stereocontrol. Taking advantage of these mechanistic insights, we aimed to explore a computation-assisted directed evolution approach for obtaining *R*-selective variants. In this study, we show that by combining the power of experimental protein engineering with structural and mechanistic insights obtained from computational modeling, we could rationally reshape the active site of **L6\_FL** to switch the orientation of the LAC intermediate and access the opposite *R*-enantiomer of the product (Figure 1b). This work illustrates a strategy to access enantiodivergent abiological transformations through

[\*] C. Calvó-Tusell,<sup>†</sup> Dr. M. Garcia-Borràs  
 Institut de Química Computacional i Catalisi and Departament de Química, Universitat de Girona  
 C/M. Aurèlia Capmany, 69, 17003 Girona (Spain)  
 E-mail: marc.garcia@udg.edu

Dr. Z. Liu,<sup>†</sup> Dr. K. Chen, Prof. Dr. F. H. Arnold  
 Division of Chemistry and Chemical Engineering, California Institute of Technology  
 1200 E California Blvd., Pasadena, CA 91125 (USA)

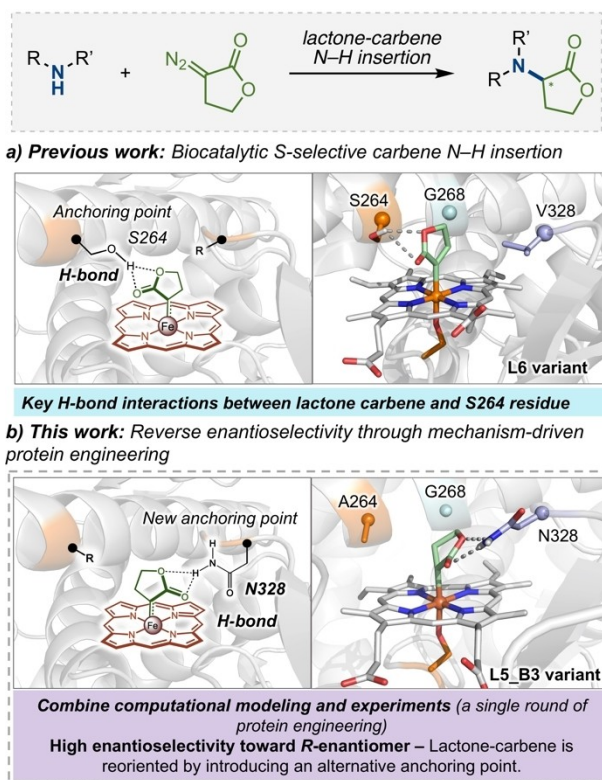
Dr. Z. Liu<sup>†</sup>  
 National Institute of Biological Sciences  
 Beijing, 102206 (China)  
 E-mail: liuzhen@nibs.ac.cn

Dr. K. Chen  
 Innovative Genomics Institute, University of California Berkeley, CA (USA)

[<sup>†</sup>] These authors contributed equally to this work.

[\*\*] A previous version of this manuscript has been deposited on a preprint server (<https://doi.org/10.26434/chemrxiv-2022-f02xh>).

© 2023 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.



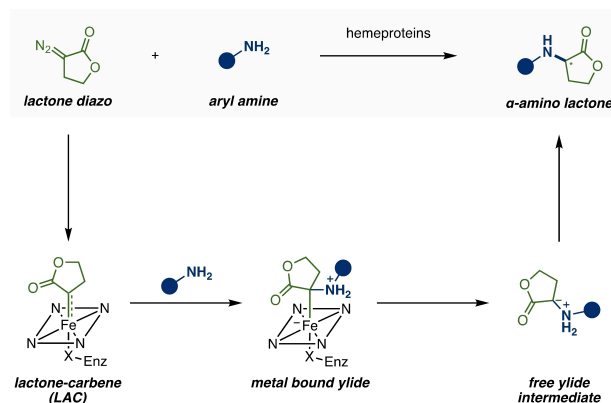
**Figure 1.** Mechanism-guided strategy for engineering P411 variants with reversed enantioselectivity in carbene transfer.

a mechanism-driven approach, based on rationally controlling the key intermediates formed in the enzyme active site during catalysis.<sup>[25]</sup> Furthermore, this enzymatic platform represents a rare example of catalytic systems achieving enantiodivergent control through fine-tuning the H-bond interactions between the catalyst and reactive intermediates.

## Results and Discussion

### Mechanistic Insights from Computational Truncated Models

The intrinsic reaction mechanism of P411-catalyzed lactone-carbene N–H insertion was initially investigated<sup>[23]</sup> which provided essential insights to guide subsequent computational modeling and protein engineering (Scheme 1 and Figure 2). 4-Methoxy aniline **1a** and lactone diazo **2** were selected as the model substrates. Following the previous work from Shaik's group on the mechanism of P450 (cysteine ligated) catalyzed carbene N–H insertion using Density Functional Theory (DFT) calculations based on a truncated heme catalyst model,<sup>[24]</sup> we performed similar DFT calculations using **1a** as the amine substrate reacting with the lactone-carbene species.<sup>[23]</sup> The computational model, i.e. *theozyme* (truncated enzyme), included a methanol molecule H-bonded to the ester group of LAC in order to mimic the potential H-bond donors in the enzyme active site (e.g., S264 in the previously demonstrated **L6\_FL** and



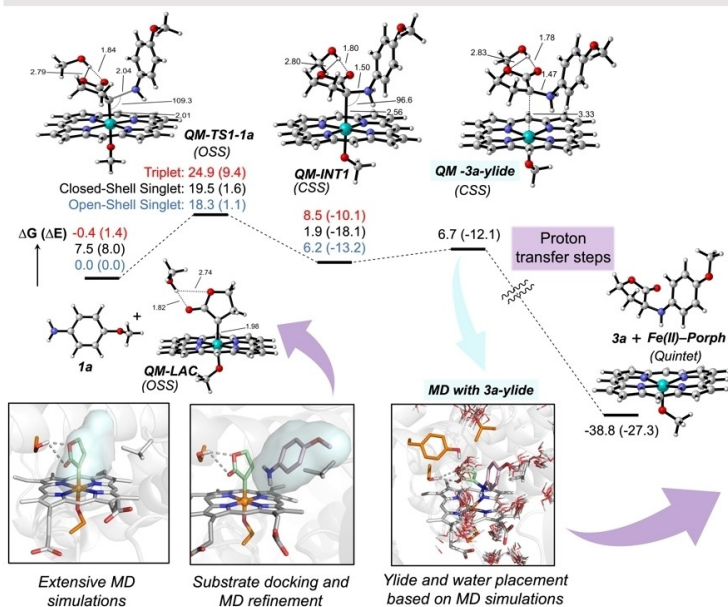
**Scheme 1.** General depiction of the reaction mechanism for heme-protein-catalyzed lactone carbene N–H insertion.

**L7\_FL** enzyme variants). Calculations showed that a plausible mechanism involves a first N-nucleophilic attack by the amine to the electrophilic iron lactone-carbene center (Figure 2a), forming an ylide intermediate covalently bound to the iron. The dissociation of the ylide from the iron was found to be energetically slightly uphill and barrierless. These results are in line with previous computational studies by the Shaik group on N–H insertion considering the acyclic ethyl diazoacetate (EDA) as the carbene precursor.<sup>[24]</sup>

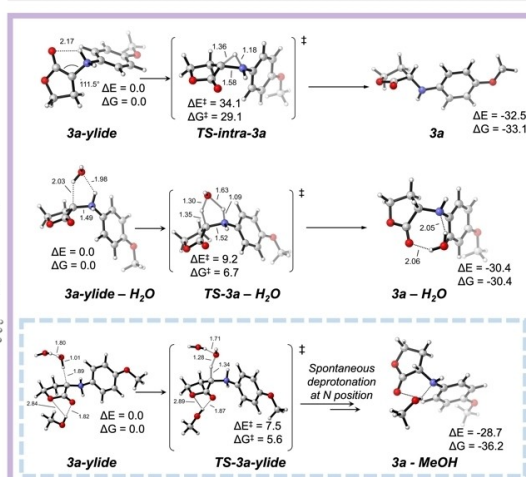
Different from the acyclic carbene system, which is proposed to involve the formation of an enol via the direct intramolecular proton transfer from the protonated nitrogen to the oxygen of the carbonyl group,<sup>[24]</sup> our calculations indicated that the ylide intermediate could directly dissociate from the iron center. With the LAC, the 5-membered transition state required for the enol formation from the ylide intermediate becomes disfavored as compared to the acyclic carbene system due to the geometric strain induced by the lactone ring (see Figure 2c3). Additionally, the H-bond interaction between the carbonyl group and the external H-bond donor (e.g., methanol in our model or S264 in the enzyme active site) is found to further disfavor the formation of this enol intermediate (see Figure 2c2). Calculations also showed that the direct proton transfer from the protonated amine to the vicinal carbon via a three-membered transition state is energetically highly disfavored.

Based on these observations, we proposed that the ylide rearrangement and the protonation of the carbon center should be facilitated by water molecules. Molecular Dynamics (MD) simulations carried out with the **L6\_FL** variant having the ylide intermediate bound in a conformation that mimics the just dissociated complex characterized from DFT calculations revealed that a few water molecules are present in the active site. These water molecules approach the ylide intermediate bearing interactions with serine 264 from the top of the lactone ring, opposite face to the heme cofactor (see Figure 2).<sup>[23]</sup> Considering this arrangement of water molecules around the ylide intermediate in the active site, truncated model systems were built and DFT calculations were carried out. Model calculations indicated that these water molecules could effectively protonate the carbon

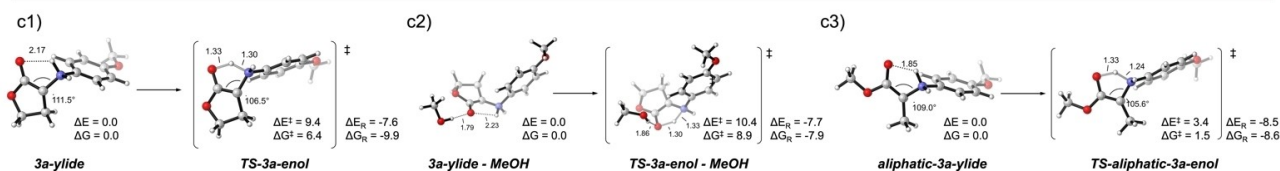
## a) Plausible reaction mechanism:



## b) Possible proton transfer mechanisms:



## c) Enol formation from ylide:



**Figure 2.** DFT *theozyme* calculations. Plausible reaction mechanism based on DFT calculations using a truncated computational model. Data from ref. [23]. a) DFT energy profile for lactone-carbene N–H insertion involving model substrate **1a**, from ref. [23]. A truncated model that includes a methanol molecule to mimic P411-L6 active site S264 residue based on substrate-bound MD simulations is used. Results obtained for energetically accessible electronic states are reported, and the lowest in energy optimized geometries for each stationary point are shown. b) Optimized model transition states (TSs) for stereoselective **3a** formation from **3a**-ylide. Computational models were built based on the conformations explored by the **3a**-ylide when formed in P411-L6 active site and the arrangement of water molecules around the ylide as observed from MD simulations.<sup>[23]</sup> c) Computational modeling of intramolecular ylide-enol tautomerization for: c1) **3a**-ylide as the model substrate; c2) **3a**-ylide as the model substrate and considering the H-bond interactions between the lactone carbonyl and a methanol molecule that mimics active site S264 residue in P411-L6; c3) acyclic-**3a**-ylide as the model substrate.

center in a fast proton transfer step, right after the ylide dissociates from the iron and before this reactive intermediate leaves the enzyme active site (see Figure 2b). Therefore, the protonation step is proposed to be stereoselective, taking place from a specific face of the ylide intermediate and mediated by a water molecule. This step ultimately depends on the first *N*-nucleophilic attack, which determines which enantiotopic face of the ylide is oriented opposite to the heme cofactor and exposed to these active site water molecules.

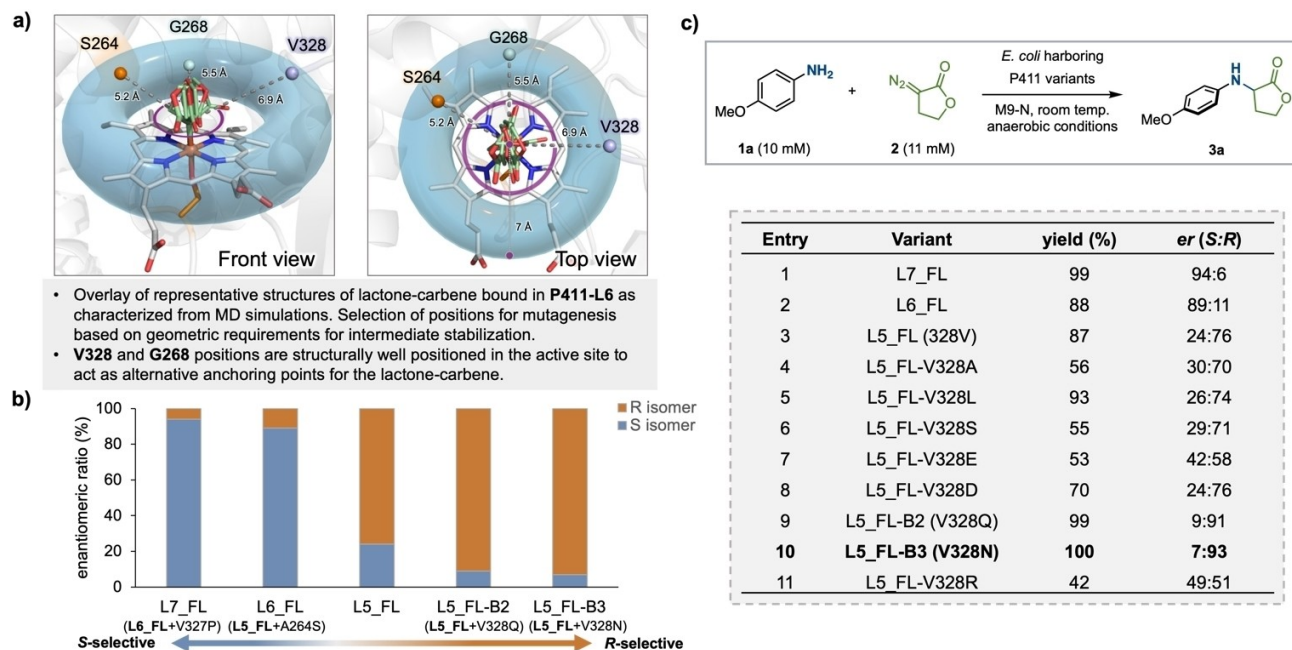
### Mechanism-Guided Protein Engineering: Controlling the Orientation of the Carbene Intermediate

With all these mechanistic and structural insights obtained from computational modeling, we sought to rationally engineer new enzyme variants to access opposite selectivities. To do so, we proposed to invert the orientation that

the LAC can explore in the enzyme active site, to force the *N*-nucleophilic attack to take place from the opposite enantiotopic face of the lactone ring, which should eventually lead to the opposite product enantiomer following a rapid stereoselective water-assisted protonation.

To alter the LAC orientation in the enzyme active site, we hypothesized that one could (1) replace the serine at position 264 in *S*-selective P411-L6/L7\_FL variant with a non-polar residue to disrupt the original H-bond interaction; and (2) introduce a H-bond donor residue at the opposite side of the catalytic pocket that could serve as a new anchoring point for the LAC and invert its orientation in the enzyme active site. By analyzing the computational models generated for the LAC intermediate bound in the poorly selective **L5** and the selective **L6** variants (FAD domain truncated versions of **L5\_FL** and **L6\_FL**), we identified two positions that could accommodate alternative anchoring points to invert the LAC orientation: positions 268 and 328 (Figure 3a). These positions were selected based on geo-





**Figure 3.** a) Structural analysis of P411-L6 active site with lactone-carbene (LAC) bound as characterized from MD simulations (see ref. [23]), and identification of positions for alternative anchoring of LAC intermediate. b) Identification of *R*-selective variant through site-saturation mutagenesis (SSM) and screening at 268 and 328 sites. c) Comparison of residues at 328 position by site-directed mutagenesis. The experiments were performed using *E. coli* that expressed P411 enzymes with 10 mM substrate **1a**, 11 mM **2**, and 25 mM D-glucose in M9-N buffer (pH 7.4) at room temperature under anaerobic conditions.

metric requirements, their spatial disposition on the equatorial plane and their appropriate distance with respect to the LAC intermediate, in order to ensure an effective interaction between the new H-bond donor and the LAC ester group. We first applied a distance threshold of 7.0 Å to select amino acid side chains that are found to directly interact with the LAC intermediate in **L5** and **L6** variants during MD simulations. The selection was limited to positions on the equatorial plane of the LAC ring. Then, we analyzed specific MD snapshots in which the LAC intermediate explores the targeted conformation that we were aiming to stabilize. From there, we identified positions 268 and 328 as suitable hosts for polar side chains with the appropriate directionality to stabilize this particular conformation of the LAC intermediate. To minimize disruption of the active site environment evolved for efficient and selective *N*-nucleophilic addition and protonation steps, **L5\_FL** bearing a non-polar alanine residue at position 264 was used as the starting point for the engineering campaign.

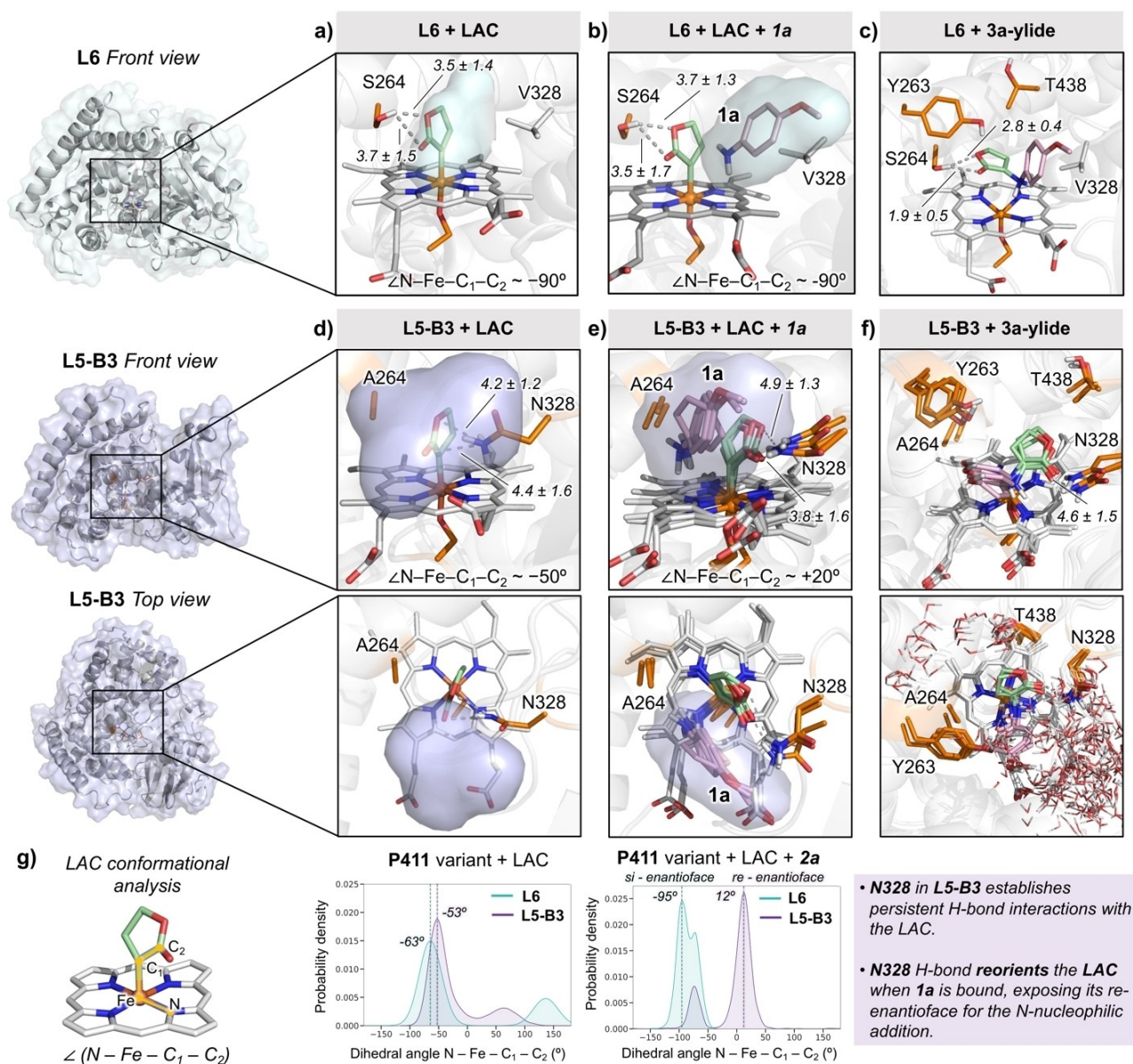
**L7\_FL** and **L6\_FL** could catalyze the reaction between 4-methoxy aniline **1a** and lactone diazo **2**, giving product **3a** in 94:6 and 89:11 *er*, respectively, favoring the formation of the *S*-enantiomer (Figure 3b).<sup>[23,26]</sup> Performing site-saturation mutagenesis (SSM) at the 328 and 268 positions using **L5\_FL** as the parent and screening led to two *R*-selective variants. In both variants, **L5\_FL-B2** and **L5\_FL-B3**, V328 was mutated to a protic residue, Q and N, respectively, which flipped the enantioselectivities to 9:91 and 7:93 *er*. Site-directed mutagenesis at 328 demonstrated that shorter polar residues (serine, Figure 3c entry 6) or charged amino

acids (glutamic acid, aspartic acid or arginine, Figure 3c entries 7, 8 and 11) led to significantly lower enantioselectivities and yields. No selectivity-enhancing mutations were found at position 268 (Figure 3b), suggesting that polar residues at this position cannot establish an effective H-bond with the LAC intermediate while allowing a selective nucleophilic attack by the amine substrate.

#### Molecular Basis for *R*-Selectivity in **L5\_FL-B3**

With the best *R*-selective variant, **L5\_FL-B3**, we performed molecular dynamics (MD) simulations to unravel the role of mutation V328 N in driving the *R*-selective carbene N-H insertion. First, the **L5-B3** variant was modeled considering the heme domain with the LAC bound. As expected, the newly introduced V328N residue establishes persistent H-bond interactions with the ester group of the LAC (Figure 4d and S3), and it is placed at the opposite side in the active site as compared to S264 in the **L6** variant (Figure 1a and Figure 4a). The relative orientation of the LAC with respect to the heme (described by the  $\angle(\text{N-Fe-C}_1\text{-C}_2)$  dihedral angle) is similar in **L5-B3** and **L6** (Figure 4a,d,g). Nevertheless, the interaction between the lactone and the newly introduced N328 side chain only makes accessible the *si* face of the LAC for the *N*-nucleophilic attack (Figure 4b,e,g), which is opposite to the **L6** variant.<sup>[23]</sup>

Next, the amine substrate **1a** bound in the **L5-B3** active site in the presence of the LAC was modeled, using restrained MD simulations to mimic near-attack conforma-



**Figure 4.** Computational modeling of L6 and L5-B3 variants based on MD simulations. a) Representative snapshot obtained from L6 variant MD simulations describing the major conformation explored by the LAC bound. The  $\angle(\text{N-Fe-C}_1\text{-C}_2)$  dihedral angle describes the relative orientation explored by the carbene (see Supporting Information for additional details and replicas). The blue surface describes the available space in the active site cavity near the LAC intermediate. b) Representative snapshot from L6 variant restrained-MD simulations describing the major near-attack conformation explored by the amine **1a** for the N-nucleophilic attack to the LAC intermediate. c) Representative snapshot obtained from restrained-MD simulations with **3a-ylide** formed in L6 active-site. d) Representative snapshot obtained from L5-B3 variant MD simulations describing the major conformation explored by the LAC bound. The purple surface describes the available space in the active site cavity near the LAC intermediate. e) Overlay of representative snapshots from L5-B3 variant restrained-MD simulations describing the major near-attack conformation explored by the amine **1a** for the N-nucleophilic attack to the LAC intermediate. f) Overlay of 3 representative snapshots obtained from restrained-MD simulations with **3a-ylide** formed in L5-B3 active-site. Water molecules shown are taken from 25 random structures selected across the 100 ns MD trajectory. g) Probability density plots describing the conformations explored by the LAC when bound in L5-B3 and L6<sup>[23]</sup> active sites, in the absence or presence of **1a** substrate, estimated from accumulated MD trajectories. Key distances and angles are given in Å and degrees [°].

tions for the N-nucleophilic attack (see Supporting Information for details, see also ref. [27]). These simulations showed that the substrate bound in a catalytically relevant mode for the N-nucleophilic addition induces a slight reorientation of the LAC (rotation along the Fe-C bond from ca.  $-50^\circ$  to ca.

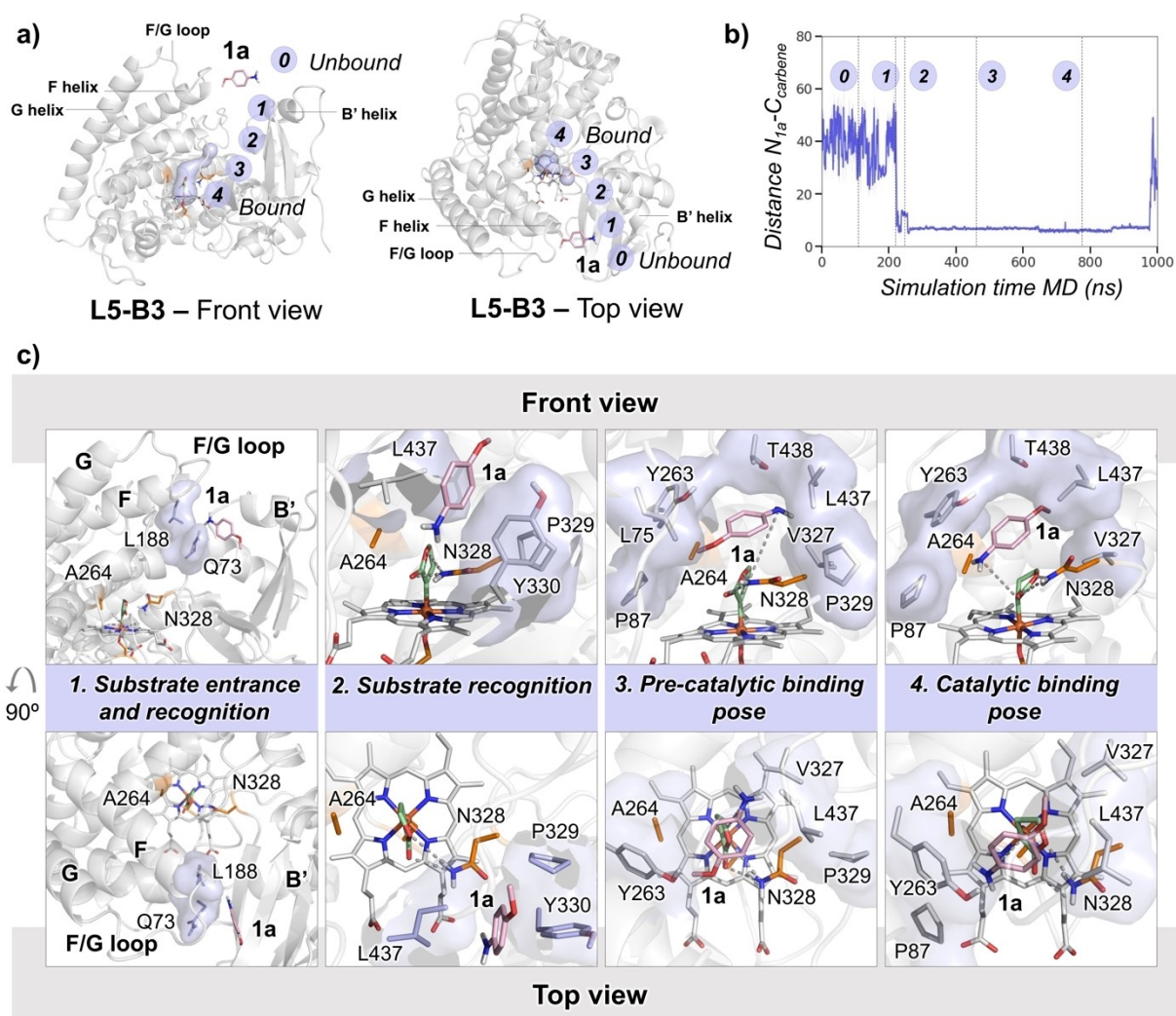
$+15^\circ$ , Figure 4d,e) that keeps the H-bond between the lactone ester group and the amide of N328. This H-bond interaction was previously shown to be also important for enhancing the electrophilicity of the LAC.<sup>[23,27,28]</sup> Consequently, this binding mode of **1a** and LAC is expected to be



more reactive than alternative ones lacking this H-bond interaction, thus biasing the reaction to happen from this characterized near-attack conformation. The amine substrate occupies the available space near the enantiotopic *si* face of the lactone and A264 (Figure 4d,e). This binding mode of the substrate is further stabilized by hydrophobic interactions with residues L75, P87, L437, and T438 (Figure 5c). All these factors synergistically favor the selective *N*-nucleophilic attack to the *si* face of the LAC ring. This is possible because the N328 side chain possesses the appropriate polarity and length to interact with the LAC via an H-bond in the absence but also in the presence of the substrate.

We then modeled the ylide intermediate in the **L5-B3** active site formed from the characterized near attack conformation, using restrained MD simulations to study how the ylide is accommodated in the active site when it

dissociates from the iron (see Supporting Information for details). Simulations indicate that the ylide intermediate, once formed, can maintain the hydrogen bond with the N328 side chain. This helps stabilize the ylide in the active site within a major binding mode where the lactone and amine aromatic rings occupy similar positions as in the substrate-bound complex (Figure 4f), without significant fluctuations or conformational changes (see Figure S10). In line with the previous observations for the **L6** system,<sup>[23]</sup> the general hydrophobicity of **L5-B3** active site is retained and only a few water molecules can access the active site from a predefined water channel (see Figure S9), from the top side of the lactone ring. Similar to the **L6** variant, it is proposed that these water molecules can rapidly protonate the ylide at the C position from the top face (*si* face) of the lactone ring, thus forming the *R*-enantiomer of the product.



**Figure 5.** Spontaneous binding process of amine **1a** in **L5-B3** with LAC bound, as characterized from unbiased MD simulations. a) Schematic representation of the characterized substrate binding pathway. b) Binding process as described by the distance between the **1a** amine nitrogen atom ( $N_{1a}$ ) and the LAC central carbon atom ( $C_{\text{carbene}}$ ) along the MD simulation time. The  $N_{1a}$ – $C_{\text{carbene}}$  distance goes from large values ( $> 50 \text{ \AA}$ ) when the substrate is in the bulk solvent to small values (ca.  $5 \text{ \AA}$ ) when the substrate explores catalytically relevant binding modes for the *N*-nucleophilic attack. c) Selected snapshots from the spontaneous binding pathway MD trajectory shown in 5b. See Figure S13 for further details.

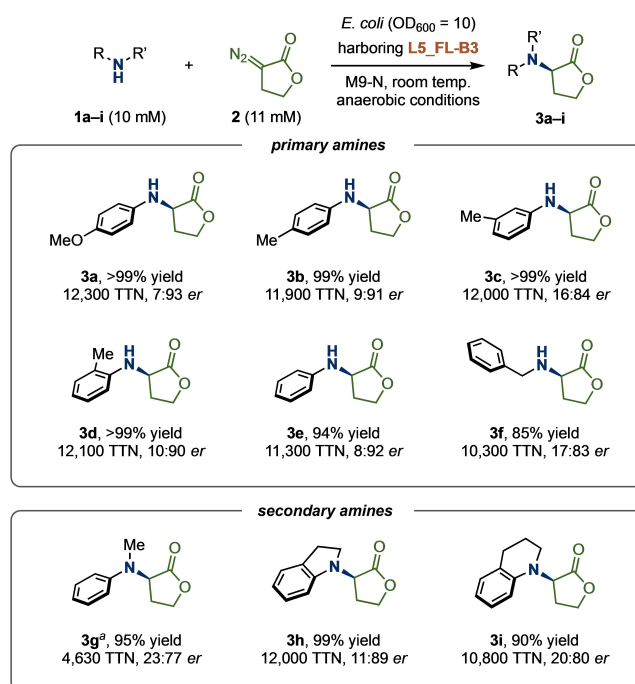
Finally, the spontaneous binding<sup>[29]</sup> pathway of amine **1a** from the bulk solvent to the active sites of **L5-B3** and **L6** variants with the LAC bound was characterized using extensive MD simulations (Figure 5, **S12** and **S13**; see Supporting Information for details). These simulations were performed without imposing any restriction or potential bias. Starting with 4 molecules of the amine **1a** randomly placed in the bulk solvent, a total of 10 independent MD replicas for each variant were propagated for 250 ns, observing one spontaneous substrate binding event for **L5-B3** and two for **L6**, in which the substrate accesses into the protein scaffold. These successful binding trajectories were then propagated up to 1000 ns. We observed that, in both **L5-B3** and **L6** cases, the amine substrate accesses the active site from the top side of the P411 protein. This corresponds to the substrate entrance channel previously characterized for the parent P450<sub>BM3</sub> and related P450 enzymes,<sup>[30–32]</sup> which is located between the F/G helices and FG loop region and the B' helix (Figure 5a). Therefore, the substrate binding pathway is not altered by evolution or the presence of the reactive carbene species. Further structural analysis of the binding pathway in **L5-B3** revealed that some of the mutations introduced in the previous engineering effort, A330Y and Q437L, participate in substrate recognition during binding (Figure 5c). These two residues are located in the substrate entrance channel. Specifically, Q437L is found to act as a gate that, once the substrate accesses the binding pocket in a pre-catalytic binding mode, closes the entrance channel and keeps the hydrophobicity of the active site by limiting the access of water molecules into it (Figure 5c, and Figure S14). Additionally, A330Y contributes to the precise positioning of the water molecules in the active site for stereoselective protonation. Both factors would be important for the enantioselective protonation of the ylide intermediate. Notably, the catalytically relevant binding poses characterized from these spontaneous substrate binding simulations (Figure 5c) are equivalent to the ones previously observed from the restrained MD simulations (see above, Figure 4e). These results further validate the utility of the computational approach used to study the catalytically relevant substrate binding modes based on sequential (restrained-)MD simulations (a first set of MD simulations with LAC bound, which are followed by substrate docking calculations and refined by restrained-MD simulations. See Supporting Information for details).

Finally, we also explored using MD simulations on why charged amino acids at 328 position led to significant decrease in yield and selectivity. We modeled **L5-V328R** variant with LAC bound in the active site and found that although R328 side chain can establish H-bond interactions with the LAC carbonyl group, the largest size of the side chain creates a highly packed active site environment for the LAC, thus hindering the approach of the amine to the carbene for an efficient *N*-nucleophilic attack (see Figure S15). Additionally, the charged side chain disrupts the polarity of the active site, altering the positioning of water molecules in the catalytic pocket, which is expected to have a huge impact in the final stereoselective proton transfer step. These results highlight that N328 (and Q328) have the

appropriate size and polarity to interact with the LAC, while allowing its efficient interaction with the substrate and facilitate the subsequent protonation step.

### Substrate Scope of the Newly Engineered Variant

After obtaining mechanistic insights on the *R*-selective variant, we next explored the substrate scope using the full-length **L5\_FL-B3** variant (Figure 6). To our delight, **L5\_FL-B3** accepted a broad range of amine substrates bearing a variety of substitution patterns (**3a–i**). All reactions were completed in excellent yields (up to >99%), high total turnover numbers (TTN, 4630–12300), and moderate to good enantioselectivities (up to 7:93 *er*).<sup>[26]</sup> It is notable that an aliphatic amine is also a competent substrate for this transformation (**3f**), which is a challenging process in synthetic chemistry.<sup>[33]</sup> It is worth mentioning that **L5\_FL** was only mildly selective for **3a** (24:76 *er*), and almost non-selective for **3d** (46.5:53.5 *er*) and **3g** (39.5:60.5 *er*) (see Supporting Information of ref.<sup>[23]</sup>), indicating the key role of V328 N for achieving high *R*-selectivity. Together with the previously demonstrated variants, **L6\_FL** and **L7\_FL**, we are now able to achieve both enantiomers of the lactone-carbene *N*–H insertion products from a diverse range of aromatic and aliphatic amines.<sup>[23]</sup>



**Figure 6.** Substrate scope. Yields and *er* were determined by HPLC analysis. Reactions were done in triplicate. The experiments were performed using *E. coli* (OD<sub>600</sub> = 10) that expressed **L5\_FL-B3** with 10 mM substrate **1a**, 11 mM **2**, and 25 mM D-glucose in M9-N buffer (pH 7.4) at room temperature under anaerobic conditions. See Supporting Information for more details. <sup>a</sup> Reaction performed using whole *E. coli* cells at OD<sub>600</sub> = 20.

## Conclusion

In summary, we have developed an enantiodivergent enzymatic platform for carbene N–H insertion chemistry. A highly efficient, *R*-selective P411 variant, **L5\_FL-B3**, was identified in a single round of protein engineering through a computation-assisted mechanism-guided approach. This variant complements the previously engineered *S*-selective mutants.<sup>[23]</sup>

Computational modeling was used to investigate the key LAC intermediates formed in the active site. These models served as starting points to search and characterize key positions for controlling the orientation of the LAC intermediate via H-bond interactions. The relative orientation of the LAC in the active site determines which enantiotopic face of the lactone-carbene is accessible for a selective *N*-nucleophilic attack by the amine substrate, prior to a final enantioselective protonation. MD simulations were employed to elucidate the origin of enantioselectivity and high activity of **L5\_FL-B3**, and to characterize the amine binding process. We also showed that **L5\_FL-B3** could accept a broad scope of substrates with excellent yields (up to >99% yield, 12300 TTN) and good enantiocontrol (up to 7:93 *er*).

This work demonstrates that it is possible to geometrically control reactive carbene intermediates formed in enzyme active sites to modulate the selectivity of carbene transfer reactions. Beyond our example, there have been many more biocatalytic transformations, natural or non-natural, recruiting similar hydrogen bonds in enzyme active sites to drive stereoselectivity,<sup>[34–38]</sup> but very few have demonstrated protein engineering to introduce a different hydrogen bond-anchoring point to reaction intermediates could alter the stereo- or site-selectivity. We hope our study will inspire more mechanism-driven protein engineering efforts, aiming to control key biocatalytic intermediates formed in enzyme active sites to enhance activity and control selectivity.

## Acknowledgements

This work was supported by the Spanish MICINN (Ministerio de Ciencia e Innovación) PID2019-111300GA-I00 project (M.G.B.), the Ramón y Cajal program via the RYC2020-028628-I fellowship (M.G.B.), the Generalitat de Catalunya (2021SGR00623 project, M.G.B.) and the NSF Division of Molecular and Cellular Biosciences (grant 2016137 to F.H.A.). K.C. thanks the Life Sciences Research Foundation for funding support. We thank Dr. Sabine Brinkmann-Chen, Dr. Yang Yang, Dr. Ferran Feixas, Ziyang Qin, and Dr. Cooper S. Jamieson for helpful discussions and comments on the manuscript. Open Access funding provided thanks to the CRUE-CSIC agreement with Wiley.

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are available in the Supporting Information of this article.

**Keywords:** Biocatalysis · Carbene Transferases · Computational Modeling · Fleeting Intermediates · Protein Engineering

- [1] K. Chen, F. H. Arnold, *Nat. Catal.* **2020**, *3*, 203–213.
- [2] D. Yi, T. Bayer, C. P. S. Badenhorst, S. Wu, M. Doerr, M. Höhne, U. T. Bornscheuer, *Chem. Soc. Rev.* **2021**, *50*, 8003–8049.
- [3] E. L. Bell, W. Finnigan, S. P. France, A. P. Green, M. A. Hayes, L. J. Hepworth, S. L. Lovelock, H. Niikura, S. Osuna, E. Romero, K. S. Ryan, N. J. Turner, S. L. Flitsch, *Nat. Rev. Methods Primers* **2021**, *1*, 46.
- [4] J. B. Pyser, S. Chakrabarty, E. O. Romero, A. R. H. Narayan, *ACS Cent. Sci.* **2021**, *7*, 1105–1116.
- [5] Z. Liu, F. H. Arnold, *Curr. Opin. Biotechnol.* **2021**, *69*, 43–51.
- [6] Y. Yang, F. H. Arnold, *Acc. Chem. Res.* **2021**, *54*, 1209–1225.
- [7] Q. Zhou, M. Chin, Y. Fu, P. Liu, Y. Yang, *Science* **2021**, *374*, 1612–1616.
- [8] A. L. Chandgude, X. Ren, R. Fasan, *J. Am. Chem. Soc.* **2019**, *141*, 9145–9150.
- [9] P. Dydio, H. M. Key, A. Nazarenko, J. Y.-E. Rha, V. Seyedkazemi, D. S. Clark, J. F. Hartwig, *Science* **2016**, *354*, 102–106.
- [10] M. Pott, M. Tinzl, T. Hayashi, Y. Ota, D. Dunkelmann, P. R. E. Mittel, D. Hilvert, *Angew. Chem. Int. Ed.* **2021**, *60*, 15063–15068.
- [11] K. J. Hock, A. Knorrscheidt, R. Hommelsheim, J. Ho, M. J. Weissenborn, R. M. Koenigs, *Angew. Chem. Int. Ed.* **2019**, *58*, 3630–3634.
- [12] L. E. Zetzsche, J. A. Yazarians, S. Chakrabarty, M. E. Hinze, L. A. M. Murray, A. L. Lukowski, L. A. Joyce, A. R. H. Narayan, *Nature* **2022**, *603*, 79–85.
- [13] J. Zhang, X. Huang, R. K. Zhang, F. H. Arnold, *J. Am. Chem. Soc.* **2019**, *141*, 9798–9802.
- [14] A. Z. Zhou, K. Chen, F. H. Arnold, *ACS Catal.* **2020**, *10*, 5393–5398.
- [15] X. Ren, N. Liu, A. L. Chandgude, R. Fasan, *Angew. Chem. Int. Ed.* **2020**, *59*, 21634–21639.
- [16] V. Steck, D. M. Carminati, N. R. Johnson, R. Fasan, *ACS Catal.* **2020**, *10*, 10967–10977.
- [17] D. Nam, A. Tinoco, Z. Shen, R. D. Adukure, G. Sreenilayam, S. D. Khare, R. Fasan, *J. Am. Chem. Soc.* **2022**, *144*, 2590–2602.
- [18] X. Ren, A. L. Chandgude, D. M. Carminati, Z. Shen, S. D. Khare, R. Fasan, *Chem. Sci.* **2022**, *13*, 8550–8556.
- [19] A. M. Knight, S. B. J. Kan, R. D. Lewis, O. F. Brandenburg, K. Chen, F. H. Arnold, *ACS Cent. Sci.* **2018**, *4*, 372–377.
- [20] D. Ghislieri, N. J. Turner, *Top. Catal.* **2014**, *57*, 284–300.
- [21] T. C. Nugent, *Chiral Amine Synthesis: Methods, Developments and Applications*, Wiley, Hoboken, **2010**.
- [22] N. J. Turner, L. Humphreys, *Biocatalysis in Organic Synthesis*, The Royal Society Of Chemistry, London, **2018**.
- [23] Z. Liu, C. Calvó-Tusell, A. Z. Zhou, K. Chen, M. Garcia-Borràs, F. H. Arnold, *Nat. Chem.* **2021**, *13*, 1166–1172.
- [24] D. A. Sharon, D. Mallick, B. Wang, S. Shaik, *J. Am. Chem. Soc.* **2016**, *138*, 9597–9610.
- [25] M. T. Reetz, M. Garcia-Borràs, *J. Am. Chem. Soc.* **2021**, *143*, 14939–14950.
- [26] Reaction under aerobic conditions, using substrate **1a** (4-methoxy aniline) and lactone diazo compound with **L7-FL** variant, showed similar reaction outcomes (including *er* and



- yield) to those under anaerobic conditions, indicating that the enzymatic platform is not sensitive to air.
- [27] M. Garcia-Borràs, S. B. J. Kan, R. D. Lewis, A. Tang, G. Jimenez-Osés, F. H. Arnold, K. N. Houk, *J. Am. Chem. Soc.* **2021**, *143*, 7114–7123.
- [28] R. L. Khade, Y. Zhang, *J. Am. Chem. Soc.* **2015**, *137*, 7560–7563.
- [29] For spontaneous substrate binding simulations strategy used in this study, see: C. Calvó-Tusell, M. A. Maria-Solano, S. Osuna, F. Feixas, *J. Am. Chem. Soc.* **2022**, *144*, 7146–7159.
- [30] C. G. Acevedo-Rocha, A. Li, L. D'Amore, S. Hoebenreich, J. Sanchis, P. Lubrano, M. P. Ferla, M. Garcia-Borràs, S. Osuna, M. T. Reetz, *Nat. Commun.* **2021**, *12*, 1621.
- [31] N. Ahalawat, J. Mondal, *J. Am. Chem. Soc.* **2018**, *140*, 17743–17752.
- [32] S. J. B. Mallinson, M. M. Machovina, R. L. Silveira, M. Garcia-Borràs, N. Gallup, C. W. Johnson, M. D. Allen, M. S. Skaf, M. F. Crowley, E. L. Neidle, K. N. Houk, G. T. Beckham, J. L. Dubois, J. E. McGeehan, *Nat. Commun.* **2018**, *9*, 2487.
- [33] M. L. Li, J. H. Yu, Y. H. Li, S. F. Zhu, Q. L. Zhou, *Science* **2019**, *366*, 990–994.
- [34] Y. Fu, H. Chen, W. Fu, M. Garcia-Borràs, Y. Yang, P. Liu, *J. Am. Chem. Soc.* **2022**, *144*, 13344–13355.
- [35] B. A. Sandoval, S. I. Kurtoic, M. M. Chung, K. F. Biegasiewicz, T. K. Hyster, *Angew. Chem. Int. Ed.* **2019**, *58*, 8714–8718.
- [36] M. A. Emmanuel, N. R. Greenberg, D. G. Oblinsky, T. K. Hyster, *Nature* **2016**, *540*, 414–417.
- [37] K. F. Biegasiewicz, S. J. Cooper, M. A. Emmanuel, D. C. Miller, T. K. Hyster, *Nat. Chem.* **2018**, *10*, 770–775.
- [38] K. F. Biegasiewicz, S. J. Cooper, X. Gao, D. G. Oblinsky, J. H. Kim, S. E. Garfinkle, L. A. Joyce, B. A. Sandoval, G. D. Scholes, T. K. Hyster, *Science* **2019**, *364*, 1166–1169.

Manuscript received: March 17, 2023

Accepted manuscript online: June 1, 2023

Version of record online: July 20, 2023