Radical-Mediated Regiodivergent C(sp^3)–H Functionalization of *N*-Substituted Indolines via Enzymatic Carbene Transfer

Bo M. Couture[1], Ru Cui[1], Jia-Min Chu[2], Yong Zhang[2,*], and Rudi Fasan[1,3,**]

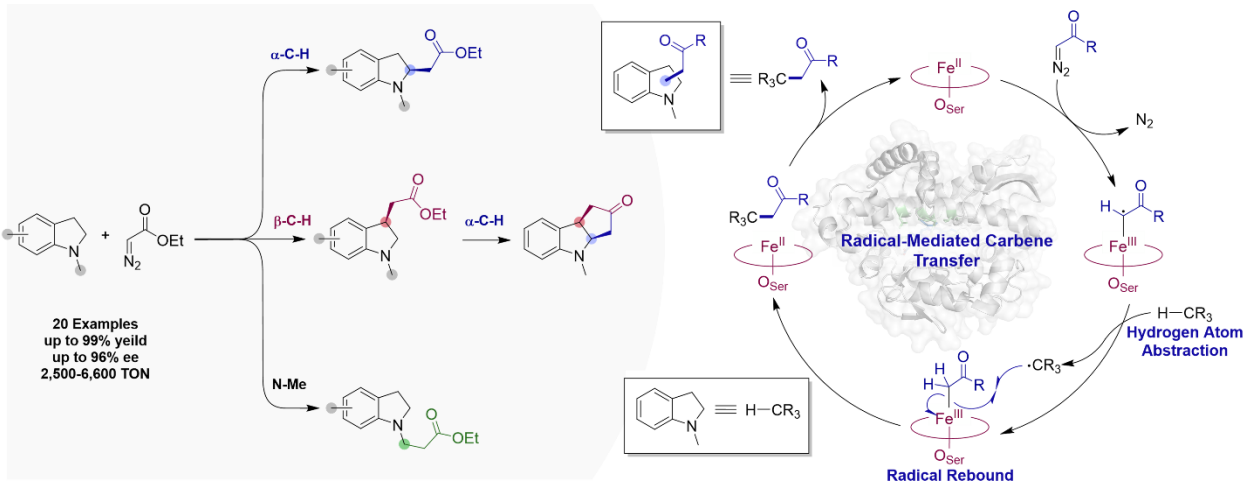
[1]*Department of Chemistry and Biochemistry, University of Texas at Dallas, Richardson, TX, 75080, United States*

[2]*Department of Chemistry and Chemical Biology, Stevens Institute of Technology, Hoboken, NJ, 07030, United States*

[3]Lead Contact: Rudi Fasan (rudi.fasan@utdallas.edu)

[*]Yong Zhang: yzhang37@stevens.edu (Y. Z.)

[**]Rudi Fasan: rudi.fasan@utdallas.edu (R. F.)



Summary

Indolines, ubiquitous structural motifs found in pharmaceuticals and natural products, pose a challenge for modification via selective C(sp^3)–H functionalization. Herein, we report the regio- and stereoselective C(sp^3)–H functionalization of *N*-substituted indolines to produce both α- and β-functionalized indolines via carbene transfer chemistry with engineered iron-based CYP119 catalysts. Further, these catalysts enable selective functionalization of exocyclic C(sp^3)–

H bond in *N*-methyl indolines and the combination of enzyme-mediated α- and β-C(sp^3)–H functionalization in a biocatalytic cascade to yield polycyclic indoline-containing scaffolds, akin to many drug frameworks. Finally, computational and experimental mechanistic studies provide evidence for a radical-mediated C–H functionalization pathway, providing first insights into the mechanism of P450-catalyzed C(sp^3)–H carbene insertion. Altogether, this work provides a direct and tunable strategy for the synthesis of functionalized indolines as key building blocks for medicinal chemistry and natural product synthesis and sheds light into the mechanism of P450-catalyzed C(sp^3)–H functionalization via carbene transfer.

Key Words

Asymmetric Synthesis, Biocatalysis, C-H Functionalization, Carbene Transfer, Cytochrome P450, Hemoprotein, Metalloenzyme, Protein Engineering

Introduction

Given the privileged nature of functionalized indolines in pharmaceuticals and their ubiquitous presence within biological systems (**Figure 1a**),[1-4] many efforts have been made toward the synthesis of such scaffolds. Existing methodologies largely rely on either reduction of the indole counterpart[5-7] or catalytic intramolecular cyclization reactions[8-11] to produce functionalized indolines. Comparatively, direct C–H functionalization offers an attractive approach for the synthesis and diversification of indoline-based scaffolds, although methods for their regio- and stereoselective direct C–H functionalization have still been elusive.[12]

Biocatalysis has been attracting increasing attention as a strategy to address important challenges in chemical synthesis.[13-21] In the area of C–H functionalization, various classes of oxidizing enzymes, including cytochromes P450, flavin-dependent halogenases and non-heme iron-dependent dioxygenases and halogenases, have proven useful for the selective functionalization of aromatic and aliphatic C–H bonds in small molecules.[13-28] Furthermore, through protein engineering, regiodivergent selectivity has been achieved, for example, for the halogenation of aromatic compounds using evolved flavin-dependent halogenases[22, 29-30] or for late-stage C(sp^3)-H hydroxylation of complex natural products using engineered P450 enzymes and halogenases (**Figure 1B**),[31-38] thus creating new opportunities for the diversification and/or chemoenzymatic synthesis of these molecules. More recently, important progress has been made

also in the development of biological catalysts for C–H functionalization via 'non-native' chemistry such as carbene transfer catalysis.[23-24, 39-50] Among them, iron-based cytochrome P450s have constituted attractive systems for achieving selective C(sp^3)–H functionalization in the presence of diazoester-based carbene donor reagents.[23, 45-47, 49, 51] Despite this progress, and in stark contrast to the scope and regiodivergent selectivity achieved via these enzymes' native reactivity (**Figure 1B**), these systems and methodologies have been largely limited to a narrow set of substrates and to functionalization of a single C(sp^3)–H site in the target substrate.[23, 45-47, 49, 51]

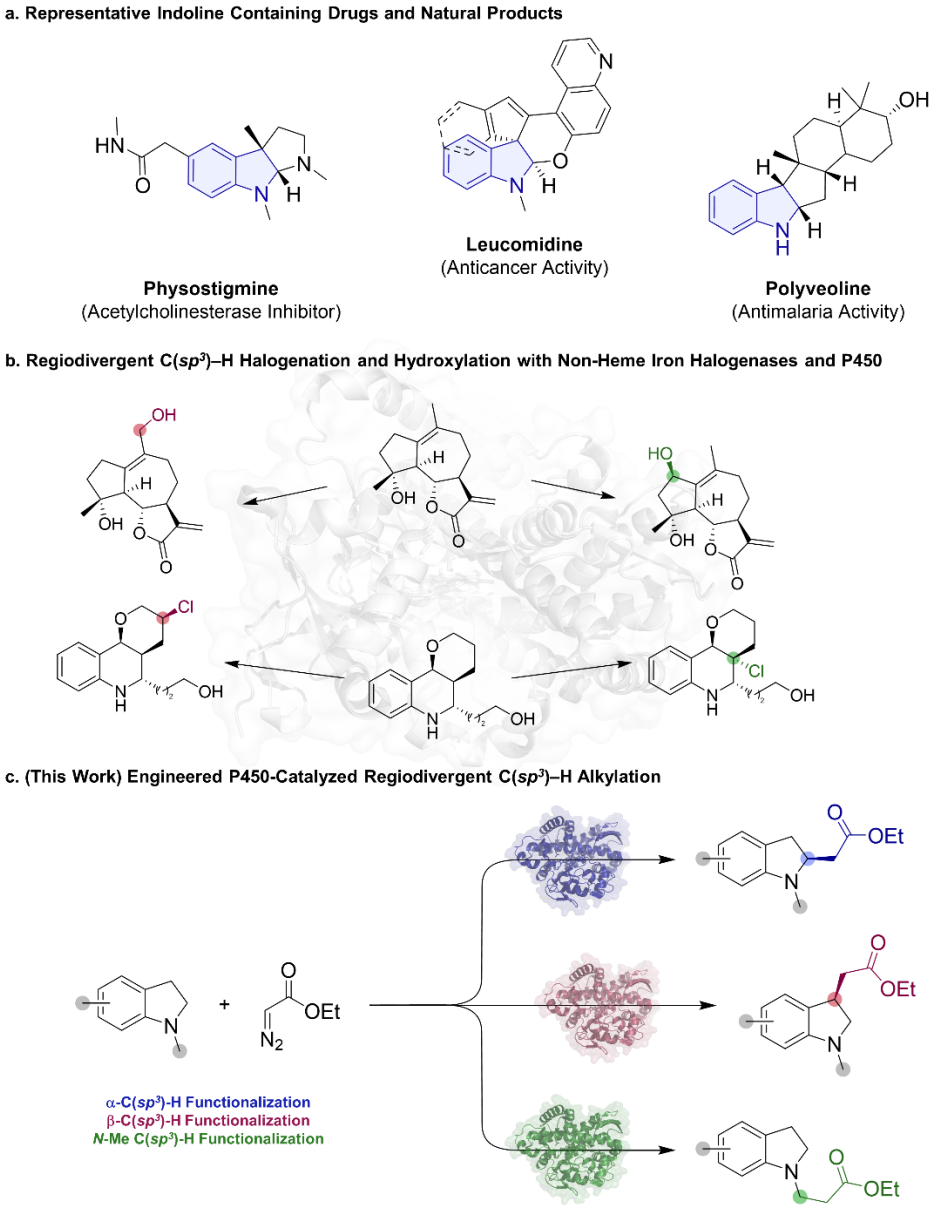


Figure 1. (a) Representative indoline containing drugs and natural products. **(b)** P450-catalyzed Regiodivergent C(sp3)–H hydroxylation. **(c)** Engineered P450-catalyzed regiodivergent C(sp^3)–H alkylation.

Herein, we describe the development of a biocatalytic strategy for the direct C(sp^3)–H functionalization of indolines, a structural subclass found in many bioactive molecules, via P450-catalyzed carbene transfer. While being compatible with a broad range of indoline substrates and diazo reagents, this methodology is also shown to allow regiodivergent access to three distinct C(sp^3)–H bonds in the substrate in a highly regio- and enantioselective manner (**Figure 1c**). As such, this strategy can provide efficient and direct strategy to the synthesis and diversification of these medicinally important scaffolds, whose synthesis is neither straightforward[52] not readily accessible using metal-catalyzed carbene transfer chemistry. The synthetic utility of the present biocatalysts and methodology is further exemplified through the synthesis of a polycyclic indoline-based core structure akin to that found in many pharmacologically active molecules.[3, 53-55] Mechanistic studies provide first-time insights into the mechanism of the present reaction and hemoprotein-catalyzed C(sp3)–H carbene insertion.

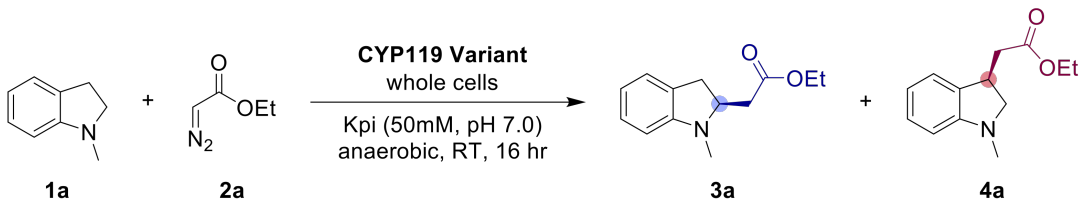
Results and Discussion

CYP119 biocatalyst for α-C–H functionalization of *N*-Methyl Indoline with EDA

Motivated by our recent success in developing engineered CYP119 variants for C(sp^3)–H bond functionalization of *N*-aryl-pyrrolidines via carbene transfer with diazoacetone and ethyl diazoacetate,[23] we sought to develop a biocatalytic method for the C(sp^3)–H functionalization of *N*-substituted indolines, which are privileged scaffolds in medicinal chemistry. Of note, this transformation poses a unique challenge in terms of regioselectivity given the presence of two C(sp^3)–H sites with similar reactivities, i.e., at the α and β position with respect to the nitrogen atom.[56] To this end, we began our investigation by screening a small panel (~100) of engineered CYP119 variants derived from our previous work,[23] including CYP119 (F153G, A209G, T213G, V254A, C317S) (a.k.a. **PS137**). The latter variant features an expanded active site as a result of four space-creating active site mutations and it was previously found to exhibit a pronounced substrate promiscuity toward C–H functionalization of *N*-aryl-pyrrolidines via carbene transfer.

These CYP119-based variants also harbor a non-native axial serine ligation, a mutation shown to be beneficial for catalysis of non-native carbene transfer reactions in this[23] and other P450 scaffolds.[57]

Table 1. Intermolecular C–H functionalization of *N*-methyl Indoline (1a) with EDA (2a) using hemoprotein CYP119 and variants thereof.[a]



1a 2a 3a 4a

No.	Catalyst	Yield (3a)[b]	3a:4a	TON (3a)[c]	e.r. (3a)[d]
1	Hemin	0	0	-	-
2	CYP119 (WT)	0	0	-	-
3	[**PS137**] CYP119 (F153A, A209G, T213G, V254A, C317S)	17%	66:34	210	53:47
4	CYP119 (F153G, T213A, V254I, C317S)	26%	70:30	840	58:42
5	CYP119 (F153G, T213A, V254L, C317S)	20%	76:24	1,150	60:40
6	CYP119 (F153G, T213A, V254F, C317S)	73%	87:13	4,180	62:38
7	CYP119 (F153G, T213A, V254Y, C317S)	66%	84:16	3,210	78:12
8	[**PS168**] CYP119 (F153G, T213A, V254W, C317S)	92%	92:8	5,270	91:9
9[e]	[**PS168**] CYP119 (F153G, T213A, V254W, C317S)	42%	92:8	8,930	91:9
10[f]	[**PS168**] CYP119 (F153G, T213A, V254W, C317S)	92%	92:8	460	91:9
11[g]	[**PS168**] CYP119 (F153G, T213A, V254W, C317S)	92%	92:8	460	91:9

[a] Standard reaction conditions: protein expressing C41(DE3) *E. coli* cells (OD_{600} = 60), 10 mM **1a**, 80 mM EDA (**2a**), in KPi buffer (50 mM, pH 7), room temperature, 16 hours, anaerobic chamber. [b] Assay yields as determined by GC using calibration curves with isolated product **3a**. [c] TON of major product as calculated based on the protein concentration measured from cell lysate. [d] Enantiomeric ratio (e.r.) of the major product as determined by chiral HPLC. [e] OD_{600} = 15. [f] Using 20 μM purified protein and 10 mM $Na_2S_2O_4$. [g] Using 20 μM lyophilized purified protein and 10 mM $Na_2S_2O_4$. N.d. = not determined.

The C(sp^3)–H functionalization of *N*-methyl indoline (**1a**) in the presence of ethyl diazoacetate (**2a**, EDA) was investigated as model reaction for this work (**Table 1**). While wild-type CYP119 shows no activity, the promiscuous variant **PS137** was found capable of converting *N*-methyl indoline (**1a**) in 17% yield (210 turnovers or TON) to yield a mixture of the α-amine C(sp^3)–H functionalization product **3a** and β-C(sp^3)–H functionalization product **4a** in approximately 2:1 ratio (**Table 1,** Entry 3). Along with the desired C–H functionalization products, the **PS137**-catalyzed reaction was accompanied with the formation of various demethylation and desaturation/N-H insertion by-products (**SI Table S5**), which have implications with respect to the mechanism of this transformation as discussed later. Importantly, the isolated cofactor hemin along with various organometallic carbene transfer catalysts[58], such as $Rh_2(OAc)_4$, $Ru(BPY)_2$, Co(TPP), $Cu(OTf)_2$, and Fe(TPP), showed no product formation (**SI Table S1**), highlighting the peculiar role of the protein matrix in enhancing the enzyme's reactivity toward this challenging transformation. In addition, the ability of **PS137** to target each of the three distinct C(sp^3)–H bonds in the indoline substrate (i.e., benzylic and the endo and exocyclic α-amino C-H bonds) held promise toward tuning the enzyme's regioselectivity via protein engineering.

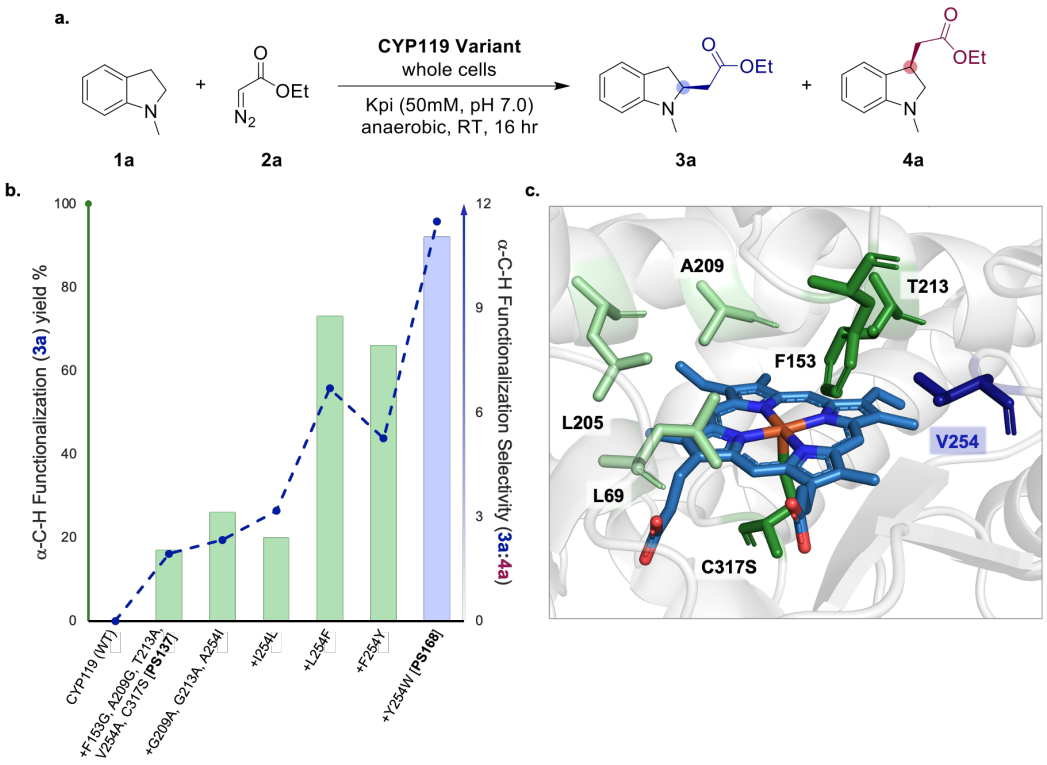


Figure 2. Directed evolution of CYP119 catalysts for regioselective C–H functionalization of *N*-methyl Indoline (1a) with EDA (2a). (**a**) General scheme for C–H functionalization of *N*-methyl Indoline (**1a**) with EDA (**2a**) to form the α-functionalization product **3a** and β-functionalization product **4a**. (**b**) Directed evolution of CYP119 catalysts for regioselective C–H functionalization of *N*-methyl Indoline (**1a**) with EDA (**2a**). Yields as determined under standard reaction conditions with EDA (**Table 1**). (**c**) X-ray crystal structure of CYP119 from *Sulfolobus solfataricus* (PDB 1IO7).[59] Active site residues targeted for mutagenesis are highlighted in dark green, conserved active site residues are highlighted in light green, active site residue V254 is highlighted in dark blue, and the heme cofactor is shown in teal.

Encouraged by these results, we aimed to identify CYP119 catalysts capable of favoring the formation of the desirable α-amine C(sp^3)–H functionalization product **3a** with higher regioselectivity as well as higher chemoselectivity against formation of the undesired demethylation/unsaturation byproducts. To this end, we extended our screening to an in-house library of CYP119-derived variants generated in previous evolution campains[23], targeting the partial mutagenesis of active site residues F153, L205, A209, and V254 using (mostly) apolar amino acids of variable size (Ala, Phe, Ile, Leu, Pro, Ser, Thr, Val). The enzyme variants were

expressed in *E. coli* C41(DE3) and screened as whole cell reactions in multi-well plates. These experiments revealed a set of structurally related variants, i.e., CYP119 (F153G, T213A, V254X, C317S, where X is Ile, Leu or Phe), that show a clear beneficial effect of increased steric bulk at the level of position 254 (Phe>Leu>Ile) toward increasing product yield as well as favoring formation of the α-amine C(sp^3)–H functionalization product **3a** over **4a** (**Table 1**, Entries 4-6), with the V254F containing variant showing the highest levels of catalytic activity and regioselectivity among them (73% yield, 4,180 TON and 87:13 r.r.; **Table 1**, Entry 6). These variants also exhibited appreciable enantiopreference for formation of the *S*-enantiomer (20-24% *ee*; **Table 1**, entry 4-6). Based on this insightful structure-activity data, we chose to investigate the effect of larger aromatic substituents, i.e., tyrosine and tryptophan, at the 254 position. Among them, and in line with the aforementioned trend, the V254W-containing variant CYP119 (F153G, T213A, V254W, C317S), referred to as **PS168**, showed further enhanced catalytic activity (4,180 → 5,270 TON) and regioselectivity (87:13 → 92:8 r.r.) for formation of **3a**. In addition, **PS168** showed improved enantioselectivity in the reaction (78:12 → 91:9 e.r.) and it catalyzes nearly quantitative conversion of the indoline substrate to the α-C–H functionalized product **3a** with no formation of undesirable by-products (**SI Table S5**). Altogether, these results showed that the modulation of steric encumbrance at the 254 site, which faces the 'eastern' side of the heme pocket in CYP119 (**Figure 2c**), is critical toward favoring attack of the α-C–H site vs. β-C–H site by the heme carbene as well as enforcing enantiopreference for the (*S*)-**3a** isomer. The superior performance of aromatic residues in this regard (Trp > Tyr ≈ Phe >> Ile/Leu/Val) also suggests the potential involvement of favorable π-π interactions between them and the indoline substrate.

Additional experiments indicated that up to 8,900 total turnovers (TTN) could be obtained for the **PS168**-catalyzed α-C–H functionalization reaction using whole cells under catalyst limiting conditions (OD_{600} = 15) (**Table 1**, entry 9). Furthermore, this reaction was determined to proceed with equally high yields and enantioselectivity using purified protein at 0.2 mol% as catalyst (**Table 1**, entry 10) and similar results could be obtained using the same CYP119 variant in lyophilized form (**Table 1**, entry 11), thus demonstrating the robustness of this biocatalyst to lyophilization and long-term storage, which are favorable attributes for preparative scale and industrial applications.

Activity of Cobalt-substituted CYP119 variants

In previous studies, we found that replacement of the heme cofactor in myoglobin with cobalt-protoporphyrin IX (Co-ppIX) engendered this hemoprotein with C–H carbene insertion reactivity in the presence of phthalan as the substrate[51]. In the interest of investigating the effect of a similar metal substitution in the present CYP119-based biocatalysts, selected variants were selected for further characterization in Co-substituted form. Conveniently, these variants were produced via recombinant expression in the presence of Co-ppIX[51], using an improved protocol reported by the Buller group.[60] Interestingly, all of the variants were found to be catalytically active toward the C–H functionalization of *N*-methyl indoline (**1a**) with EDA (**2a**), showing comparable or slightly inferior activity and regioselectivity compared to the iron-containing counterparts (**SI Table S2**). Although the performance of these metallo-substituted variants did not exceed that of the iron-containing counterparts, the functionality of these enzymes as carbene transferases is notable and it could prove useful for other types of non-native transformations.

C–H Functionalization of Indoline-based Substrates

Focusing on the best biocatalyst identified for α-C–H functionalization of **1a**, **PS168**, we next explored the substrate scope of this enzyme using a range of aryl and *N*-substituted indoline compounds (**1a-1o**). Notably, each of these substrates underwent α-C–H functionalization in the presence of EDA with excellent regio- and chemoselectivity, i.e., showing no formation of the potentially competing β-C–H functionalization product and byproducts, respectively. However, these reactions were characterized by variable yields and enantioselectivity (**Figure 3c**), indicating a certain degree of substrate specificity as observed in other P450-catalyzed native and non-native reactions.[23, 61-62] To address this limitation, we employed a substrate versus enzyme library approach, where a set of selected CYP119-derived carbene transferases from the **PS168** lineage and other generations were screened against the substrate panel in a high-throughput manner (~2,000 substrate/enzyme combinations). From these experiments, a subset of evolved CYP119 variants was shown to catalyze the α-C–H functionalization of each indoline-based substrates with high activity (2,900-6,540 TON) and good to excellent enantioselectivity (up to 96% *ee*; **Figure 3a**).

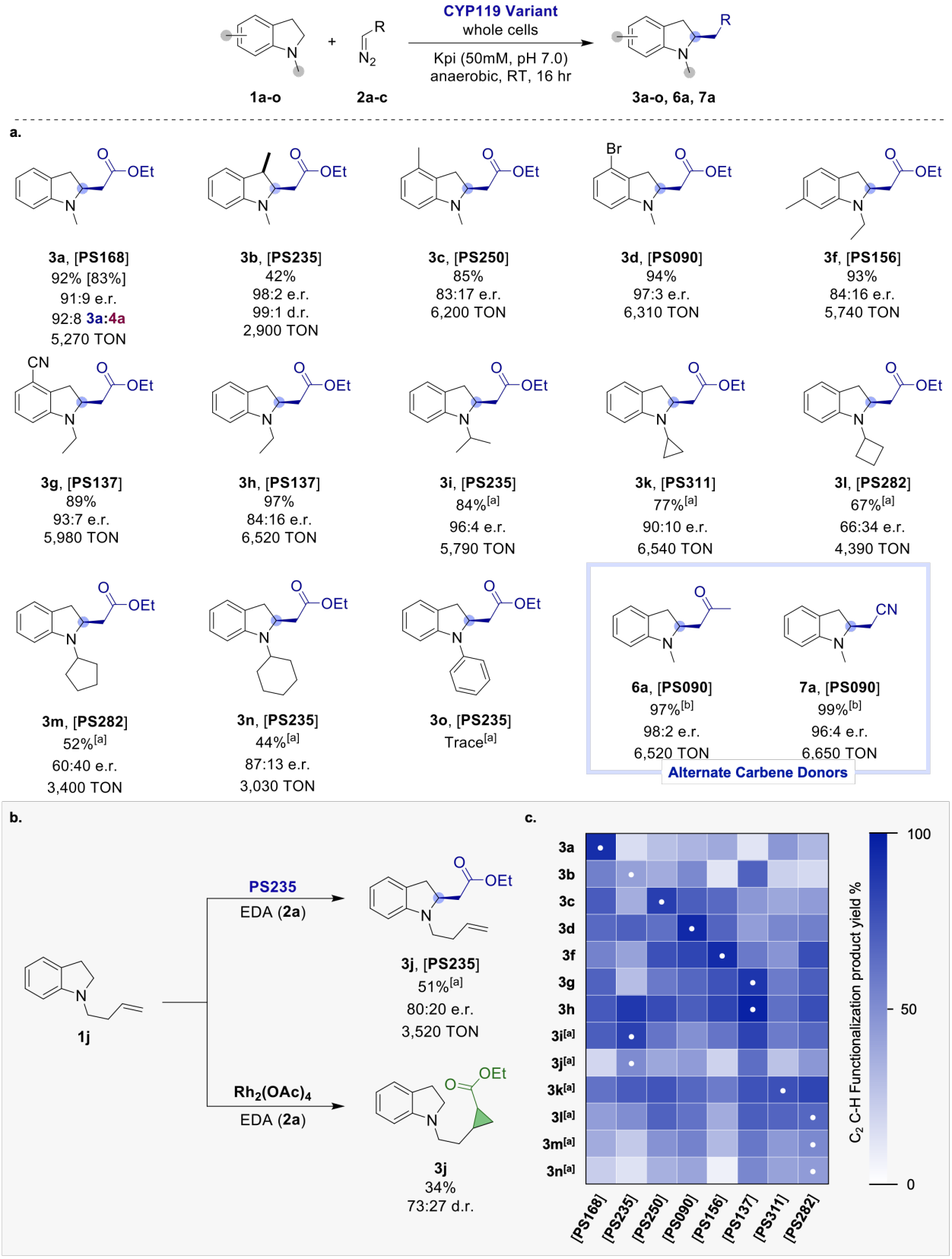


Figure 3. Activity and selectivity of CYP119 biocatalysts for α-C–H functionalization of indoline-based substrates derivatives (1a-o) with EDA. (**a**) Yields, TON, and enantioselectivity were determined from whole cell reactions under standard reaction conditions with EDA as described in **Table 1**. Analytical yields were determined using GC/LC using calibration curve prepared with isolated products [Isolated yield in brackets]. [a] Using 240 mM EDA (**2a**) at 40 °C. [b] Using 240 mM diazoacetone (**2b**) or diazoacetonitrile (**2c**). (**b**) Chemodivergent reactivity of CYP119 **PS253** and $Rh_2(OAc)_4$ (10 mol%) with 1-(but-3-en-1-yl)indoline (**1j**) and EDA (**2a**). (**c**) Heat map depicting intermolecular C–H functionalization of *N*-substituted Indolines (**1a-n**) with EDA (**2a**) using hemoprotein CYP119 variants. White dots (•) indicate the best variant for the corresponding substrate. Details about the variants are provided in **SI Figure S1**.

In particular, both electron-withdrawing (**3c** and **3f**) and electron-donating (**3d** and **3g**) substitutions on the aryl ring were well tolerated, resulting in high levels of activity and enantioselectivity (85-94% yield, 5,740-6,310 TON, 17:84-97:3 e.r.). Racemic 3-methyl substituted indoline derivative **1b** was efficiently converted into the desired product **3b** with both excellent diastereo- and enantioselectivity (98:2 e.r., 99:1 d.r.). This demonstrates the ability of the enzyme to induce kinetic resolution of racemic starting materials, a valuable trait in the context of stereoselective catalysis. In addition to methyl group as *N*-substituent, a variety of linear (**3f-h**, **3j**), branched (**3i**), and cyclic (**1k-n**) alkyl groups were found to be tolerated at this position, with a slight decrease in yield and/or enantioselectivity for bulkier *N*-cycloalkyl substitutions (i.e., **3m-n** vs. **3h-k**). In contrast, *N*-aryl-indolines such as **3o** were converted only with trace activity. Enzymatic conversion of *N*-alkyl indolines was initially affected by the low solubility of these compounds in aqueous media even in the presence of organic cosolvent (10% v/v EtOH). However, improved yields for these reactions (e.g., 27→44% yield **3n**) could be achieved by raising the reaction temperature to 40°C, a readily applicable condition thanks to the thermostability of the CYP119 variants. Notably, substrate **1j** could be converted to the desired α-C–H functionalization product **3j** with high regio- and chemoselectivity and without affecting the terminal olefinic group, highlighting the ability of the CYP119 based catalyst to favor the more challenging C–H carbene insertion reaction over cyclopropanation, i.e., unlike synthetic catalysts used for carbene transfer reactions (**Figure 3b**).

Although individual CYP119 variants were selected to optimize yield and enantioselectivity for each substrate (**Figure 3a**), each CYP119 variant is capable of catalyzing the desired reaction (**Figure 3c**). These factors show the generality of our engineered CYP119 variants towards the α-C–H functionalization of indoline-based derivatives. This study also allowed us to further hypothesize the role of steric bulk in promoting favorable binding conformations to target α-C–H functionalization compared to β-C–H functionalization. Wherein, when using bulkier *N*-alkyl substitutions, none of the previously observable β-C–H functionalization product is formed (**3f-o**). We could also observe a negative correlation between substrate-induced steric bulk and catalyst-induced bulk, as catalysts that contain bulkier residues within the active site showed lower activity toward bulkier substrates, i.e. **PS168** and **PS156** containing a V254W and A209W mutation showed basal activity towards substrate **1n**, bearing a *N*-cyclohexyl moiety.

Finally, efficient and selective α-C–H functionalization of *N*-methyl indoline (**1a**) could be achieved using both diazoacetone (**2b**) and diazoacetonitrile (**2c**) as carbene donors, yielding the respective products (**6a** and **7a**) in quantitative yields (97-99%) and with high catalytic activity (>6,000 TON), regioselectivity (100:0 r.r.) and enantioselectivity (92-96% ee) (**Figure 3a**). Altogether, these results demonstrate the generality of this CYP119-based methodology for α-C–H functionalization of a broad range of *N*-substituted indolines with high catalytic efficiency and selectivity (**Figure 3**).

Regiodivergent CYP119 catalysts for indoline C–H functionalization

During initial catalyst development for achieving selective α-C–H functionalization of *N*-methyl-indoline, it was noted that early generations of enzymes bearing more open active sites showed appreciable regioselective toward formation of the β-C–H functionalization product (**4a**) (e.g., **PS137**: 66:34 **3a**:**4a**; **Table 1**, entry 3). Building upon this finding, we sought to develop a regiocomplementary catalyst capable of selectively targeting the β-position (**Figure 4a**). To this end, we selected variants from the in-house CYP119 library that contained bulky residues (Phe, Tyr, Trp) at positions L205, L69, and A205, which are located on the opposite site of the active site compared to the Val254 residue shown to be instrumental in tuning regioselectivity to the α-C–H functionalization product (**Figure 2b/c**).

Utilizing this approach, we identified CYP119 (F153G, L205W, T213A, V254A, C317S), named **PS235**, as an efficient and selective biocatalyst for functionalization of the β-C–H position

in *N*-methyl-indoline (**1a**), producing **4a** in high yield and TON (78% yield, 5,380 TON) as well as high enantio- and regioselectivity (94:6 e.r., 16:84 **3a**:**4a; Figure 4d**). As anticipated, the active site mutations in **PS235** include a bulky substitution (i.e., L205W) (**Figure 4c**) on the opposite side of the enzyme active site compared to a similar bulky substitution (i.e., V254W) harbored by the α-C–H selective variant **PS168** (**Figure 4b**). These features along with the identity of the tryptophan residue involved in these mutations support the proposed role of both steric effects as well as favorable substrate-protein π-π interactions mediated by the tryptophan residue in controlling the orientation of the substrate to favor functionalization of either the α- or the β-C(sp^3)–H bond in the two regiodivergent biocatalysts.

Given the value of 2,3-difunctionalized indolines in medicinal chemistry (**Figure 1**) and given our success in the kinetic resolution of racemic **1b** with 96% *ee* and 98 *de* (**3b**, **Figure 3a**), we challenged the β-C–H selective variant **PS235** with racemic *N*-methyl-2-methyl-indoline (**1p**). Importantly, substitution at the α-position was tolerated by the enzyme and the racemic substrate could be converted to the optically active product **4b** in good diastereomeric and enantiomeric excess (79:21 e.r., 67:33 d.r.), albeit low catalytic efficiency (12% yield). Furthermore, using a different CYP119 variant (**PS137**), this transformation could be carried out with further increased enantio- and diastereoselectivity (92:8 e.r., 81:19 d.r.), illustrating the potential value of this biocatalytic systems for kinetic resolution applications.

During the exploration of the evolved variant **PS235**, we observed that the regioselectivity of this enzyme could be directed to the functionalization of the *N*-methyl C–H bond in the presence of substitutions at the C_5 (**5e**) and C_6 (**5f**) position on the aryl ring (**Figure 4d**). These reactions were found to proceed with full regiocontrol and excellent catalytic activity (89-92% yield, 6,140-6,340 TON), complementing the scope of the α- and β-C–H functionalization reactions catalyzed by the CYP119 catalysts. Altogether, these results demonstrated the capability of the present methodology to target as many as three different C(sp^3)–H sites in a substrate for C–H functionalization via carbene transfer, a feature unprecedented for carbene transferases and rarely achieved with synthetic carbene transfer catalysts.[63-64]

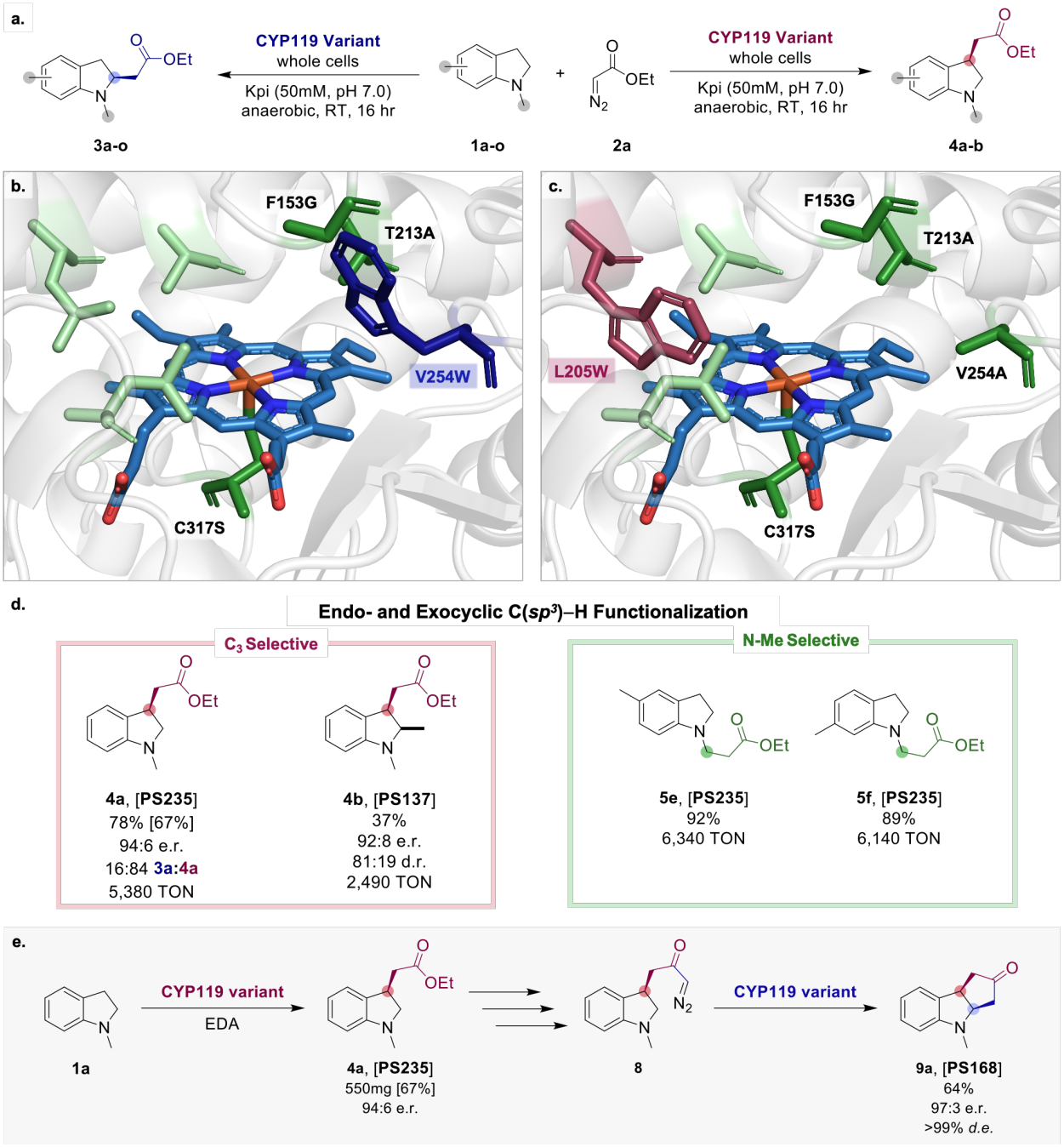


Figure 4. Directed evolution of CYP119 catalysts for regiodivergent C–H functionalization of *N*-methyl Indoline (1a) with EDA (2a). (**a**) Regiodivergent pathway for C–H functionalization of *N*-methyl Indoline (**1a**) with EDA to form α-alkylated product **3a** and β–alkylated product **4a**. (**b-c**) Comparison of modeled active sites for α-C–H selective variant **PS168** (**b**) and β-C–H bond selective **PS235** (**c**). The heme cofactor is shown as a stick model in teal, mutated active site residues (vs. wild-type) are highlighted in dark green, and the TRP residues at position 245

(**PS168**) or 205 (**PS235**) is highlighted in dark blue or dark red, respectively. (**d**) Benzylic and *N*-methyl C–H functionalization products. Yields, TON, and enantioselectivity were determined from whole cell reactions under standard reaction conditions as described in **Table 1**. (**e**) Secondary functionalization pathway for the intramolecular C–H functionalization of diazo tethered *N*-methyl indole (**8**) into product **9a**.

Polycyclic Indolines via Tandem Enzyme-Catalyzed C(sp^3)-H Carbene Insertion

As described above, our investigation of the substrate scope of the engineered CYP119 variants established that both α-C–H and β-C–H selective biocatalysts can tolerate substitutions in the adjacent position in the indoline substrate targeted for functionalization (i.e., **3b** and **4b**). In addition, with product **6a**, we established that diazoketones are accepted as carbene donors by these catalysts. Based on these results, we envisioned the possibility to construct a tricyclic indoline-based scaffold via a tandem CYP119-mediated C–H functionalization strategy, in which the enzymatic β-C–H insertion product **4a** is converted into the corresponding diazo ketone (**8**) and then subjected to an intramolecular C–H insertion to form the desired polycyclic compound (**Figure 4e**). The latter is akin to the indoline-based core structure found in various natural products and pharmaceuticals (**Figure 1a**). Toward this goal, we performed a preparative scale (500 mg, 3.75 mmol) reaction of *N*-methyl indoline (**1a**) in the presence of EDA (**2a**) was carried out using the β-C–H selective CYP119 variant **PS235**, to afford **4a** in 67% isolated yield. The ester group of the enantioenriched **4a** was then chemically converted to the respective diazoketone **8** in 45% yield over 3 steps using established chemistry (see SI for details). Gratifyingly, the diazoketone intermediate **8** could be effectively cyclized by the α-C–H selective variant, **PS168**, to afford the desired polycyclic product **9a** in high yield (64%) and high enantio- and diastereomeric excess (>500:1 d.r.; 97:3 e.r.; **Figure 4e**). These results further highlight the generality and utility of the engineering CYP119 library for the synthesis of biologically relevant synthons.

Mechanistic Insights into CYP119-Catalyzed Indoline C–H Functionalization

As noted earlier, the C–H functionalization reaction with *N*-methyl-indoline and EDA catalyzed by the early, unoptimized CYP119 catalysts is accompanied by the formation of three distinct by-products, which were determined to correspond to *N*-methyl-indole (**1a-DS**), C_3-

functionalized *N*-methyl indole (**4a-DS**), and the *N*-alkylated-indoline **13a** (**SI Table S5**, Entry 2). While these side-reactions could be suppressed in the presence of the optimized biocatalysts and reaction conditions, they bear important mechanistic implications. Indeed, as summarized in **Figure 5**, all of these side products can be explained based on desaturation reactions involving substrate-derived radical intermediates, which supports a radical, stepwise mechanism for the present C–H carbene insertion reaction also in line with the results from computational analyses described further below.

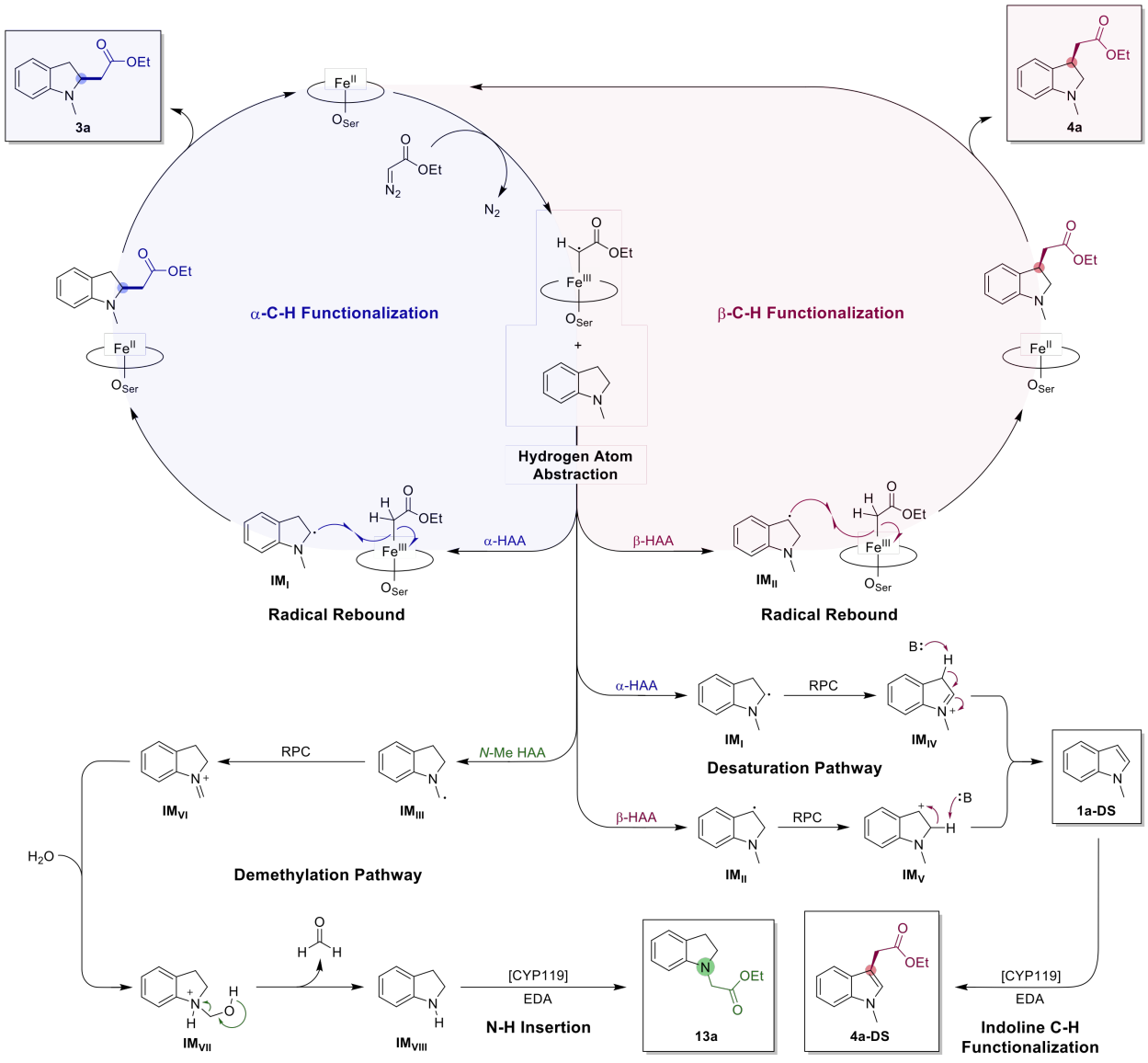


Figure 5. Proposed stepwise radical-mediated mechanism for the C–H functionalization of *N*-methyl indoline (**1a**) in the presence of engineered CYP119 catalysts.

For the productive pathway leading to the C–H functionalized products, the reactive heme-carbene species derived from reaction of ferrous protein with EDA is proposed to abstract a hydrogen atom from either the α- or β-C–H site with respect to the indoline amine, with regioselectivity being controlled by the enzyme's active site configuration (**Figure 4b/c**) and resulting in the C-centered radical intermediates $\mathbf{IM_I}$ or $\mathbf{IM_{II}}$, respectively. Radical rebound with the iron-carbenoid species yields the respective C–H functionalization products (e.g., **3a** or **4a**). For the side-reaction leading to the desaturation product **1a-DS**, the radical intermediate $\mathbf{IM_I}$ or $\mathbf{IM_{II}}$ can undergo radical polar crossover (RPC), likely through single-electron transfer to the heme cofactor, to give the respective carbocation $\mathbf{IM_{IV}}$ or $\mathbf{IM_V}$, which upon deprotonation, yield *N*-methyl-indole **1a-*DS*** (**Figure 5**). The latter can be then converted by the enzyme to **4a-DS** via a indole C(3)-H functionalization reaction with EDA akin to that previously reported by our group[44] and others[48, 65] for other engineered hemoproteins.

On the other hand, formation of the *N*-alkylated-indoline product **13a** can be rationalized based on a first step involving enzymatic *N*-demethylation of *N*-methyl-indoline **1a**, followed by N–H carbene insertion of the resulting indoline to give **13a** (**Figure 5**). Consistent with our results with **5e** and **5f**, hydrogen atom abstraction at the level of *N*-methyl group is also accessible to this biocatalyst, resulting in the carbon-based radical $\mathbf{IM_{III}}$ which, upon RPC, is expected to produce the iminium intermediate ($\mathbf{IM_{VI}}$). The latter can then undergo hydrolysis to form indoline, which can then give rise to **13a** via N–H carbene insertion with EDA, a known reaction for engineered hemoproteins.[66-68]

Various lines of experimental evidence support the proposed reaction pathways. Since the model substrate *N*-methyl-indoline was found to be susceptible to (slow) desaturation to indole in the presence of air and further modification of the CYP119 product could occur by action of other enzymes in whole cell reactions, control experiments were first performed to confirm the enzymatic origin of the observed side-products. To this end, time-course experiments were carried out in the presence of *N*-cyclopropyl-indole (**1k**), which is stable toward oxidative desaturation, and the unselective variant **PS282** (in purified form) as the catalyst. Under catalytic (air-free) conditions, formation of the C–H insertion product **3k** is accompanied by accumulation of desaturated by-products **1k-*DS*** and **4k-*DS*** in approximately 73:9:18 ratio over 90 minutes, with corresponding initial formation rates (TOF) of 28, 5.7, and 2.5 turnovers/minute for **3k**, **1k-*DS***, and **4k-*DS***, respectively (**SI Figure S7**). In contrast, no formation of either desaturation byproduct

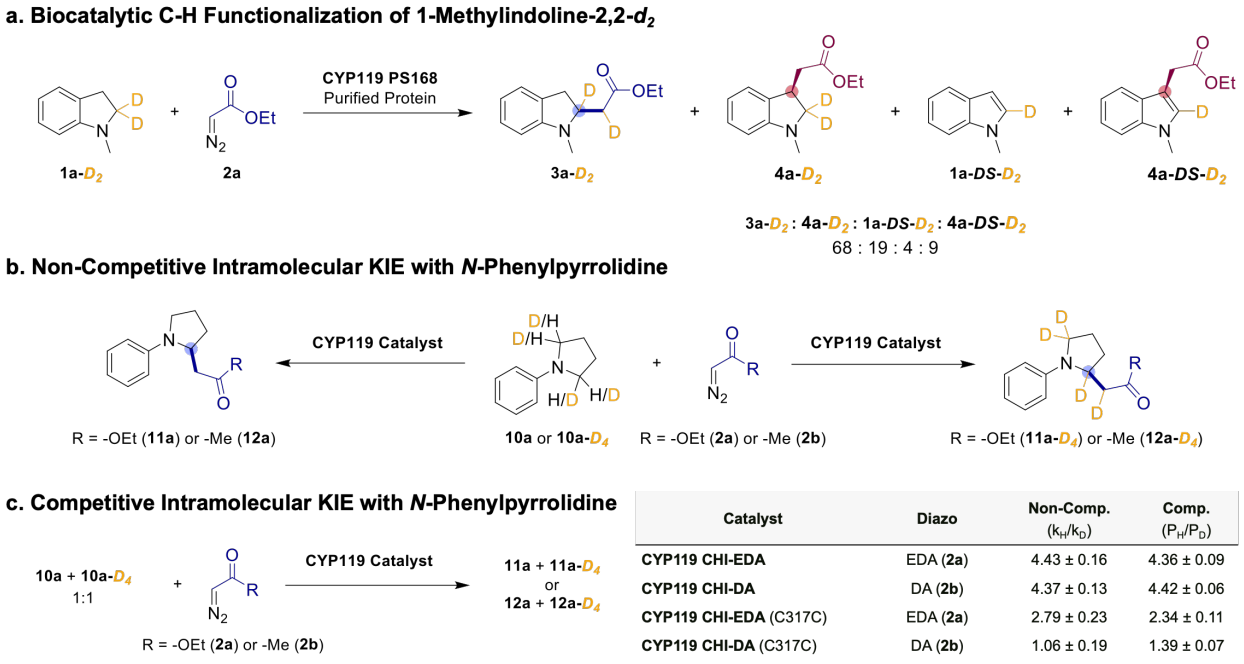
was observed in the absence of the enzyme or EDA within the same conditions (**SI Figure S7**), clearly indicating that these species are enzymatic products. To further investigate the sequence of reactions leading to the C_3-functionalized *N*-alkyl-indole product (i.e., **4a-DS** from **1a** or **4k-DS** from **1k**), products **4a** and **1a-*DS*** were used as substrates under standard catalytic conditions in the presence of **PS282**. While **4a** was fully preserved in the reaction mixture, **1a-*DS*** was consumed to give rise to the C_3-functionalized product **4a-*DS*** (**SI Figure S8**). Although the mechanism of this step was not investigated, previous studies with engineered myoglobin support a stepwise mechanism involving a zwitterionic intermediate.[44, 69]

N-Demethylation of *N*-methyl-indoline via the proposed mechanism in **Figure 5** implies the release of formaldehyde as byproduct. To test this, the enzymatic reaction mixture was added with the formaldehyde trapping agent *O*-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine (PFBHA, **14**), followed by GC analysis to detect the corresponding formaldehyde adduct (oxime). As anticipated, detectable amount of the PFBHA-derived oxime was detected in this reaction and this species was found to be proportional to the concentration of the demethylated/*N*-alkylated product **13a** (**SI Figure S7**). In separate experiments, we further determined that indoline is readily converted by the enzyme to the *N*-alkylated product **13a** under standard catalytic conditions in the presence of EDA (**SI Figure S8**). Collectively, these results support the proposed mechanistic pathways for formation of the experimentally observed products and byproducts in this reaction (**Figure 5**).

While the application of reducing conditions (i.e., sodium dithionite or intracellular environment) in the current protocol entails the involvement of ferrous hemoprotein as carbene transfer catalyst, we further explored the importance of the redox state of the protein in this reaction, also in light of recent reports on the activity of iron(III)-metalloporphyrins for carbene transfer reactions.[70] In the absence of the reductant, the **PS168**-catalyzed reaction with **1a** and EDA proceeds with reduced catalytic activity (45→32% yield; **SI Table S6**, entry 1-2), but identical regio- and enantioselectivity as compared to that under reducing conditions. Further experiments were conducted to discern whether the former activity stems from the ferric protein or if the hemoprotein is reduced *in situ* by the diazo compound, as previously observed by our group for certain axial substituted myoglobin-based carbene transferases.[71] Accordingly, the same reactions were carried out in the presence of CO-saturated buffer, with CO being expected to bind with high affinity to only the ferrous form of the protein, thus inhibiting its reactivity. Under these conditions,

with or without an external reductant, no carbene transfer activity was observed (**SI Table S6**, entry 3-4), indicating that ferrous CYP119 is the catalytic species responsible for formation of the C–H functionalization products (**Figure 5**).

Finally, further mechanistic insights were gained using the deuterium labeled substrate *N*-methyl-2-D_2-indoline (**1a-*D*$_2$**). In the presence of the α-C–H selective variant **PS168**, enzymatic transformation of **1a-*D*$_2$** in the presence of EDA (**2a**) showed a notable change in the overall product distribution (vs. reaction with **1a**) to favor formation of the β-C–H functionalized product **4a-*D*$_2$**, along with the *N*-methyl-indole byproducts **1a-DS-*D*$_2$** and **4a-DS-*D*$_2$** (92:8 →68:19:4:9; **Figure 6**). Insightfully, and in line with the overall mechanism of **Figure 5**, these results indicate that (*i*) the H atom abstraction (HAA) step is regioselectivity determining and (*ii*) that the carbon radical intermediate derived from β-HAA represents a pathway branching point toward formation of the desaturation byproduct (indole). Indeed, as the energy barrier for α-HAA in increased by the H→D substitution, it is conceivable that the reaction pathway is partially diverged to favor the β-HAA manifold, thereby leading to **4a-*D*$_2$** via the productive pathway, or to **1a-DS-*D*$_2$**, via the unproductive, side-reaction. Since no desaturation byproduct is observed in **PS168**-catalyzed reaction with **1a**, it can be further evinced that the desaturation products **1a-DS-*D*$_2$** (and **4a-DS-*D*$_2$**) primarily derive from the β-HAA pathway (vs. α-HAA path), possibly due to a slower radical rebound step and/or more favorable conformation for RPC.



Catalyst	Diazo	Non-Comp. (k_H/k_D)	Comp. (P_H/P_D)
CYP119 CHI-EDA	EDA (**2a**)	4.43 ± 0.16	4.36 ± 0.09
CYP119 CHI-DA	DA (**2b**)	4.37 ± 0.13	4.42 ± 0.06
CYP119 CHI-EDA (C317C)	EDA (**2a**)	2.79 ± 0.23	2.34 ± 0.11
CYP119 CHI-DA (C317C)	DA (**2b**)	1.06 ± 0.19	1.39 ± 0.07

Figure 6. P450 CYP119-catalyzed KIE experiments. (**a**) Intermolecular non-competitive KIE experiment with *N*-methylindoline-2,2-d_2 (**1a-*D*$_2$**) and EDA (**2a**). (**b**) Intermolecular competitive and non-competitive KIE experiments with *N*-phenylpyrrolidine (**10a**) or *N*-phenylpyrrolidine-2,2,5,5-d_4 (**10-*D*$_4$**) and EDA (**2a**) or DA (**2b**). (**c**) Intermolecular competitive and non-competitive KIE experiments with *N*-phenylpyrrolidine (**10a**) or *N*-phenylpyrrolidine-2,2,5,5-d_4 (**10-*D*$_4$**) and EDA (**2a**) or DA (**2b**). Elaborated KIE data in **SI Figure S8/9**.

Mechanistic Studies

To further elucidate the mechanism of this P450-catalyzed C–H carbene insertion, we also investigated our previously reported C(sp^3)–H functionalization reaction of aryl pyrrolidines with EDA and diazoacetone, as these and the present reaction share the same catalytic system (i.e., serine-ligated CYP119).[23] Importantly, and unlike the indoline reaction, kinetic isotope effect (KIE) values can be more readily measured for the pyrrolidine reactions due to the absence of side products and pathway branching points. Accordingly, KIE values were calculated from both competitive and non-competitive intermolecular KIE experiments using deuterated and non-deuterated *N*-phenyl pyrrolidine (**10a-*D*$_4$** and **10a**) in the presence of both EDA (**2a**) and diazoacetone (**2b**) and in combination with the respective optimized CYP119 catalysts (**CHI-EDA** and **CHI-DA**). In both cases, a relatively large primary KIE value of 4.36-4.43 was measured for both competitive and non-competitive (parallel) kinetic experiments (**Figure 6b/c**). The similar

values in both experimental settings indicate that the C–H bond cleavage step is part of the rate-determining step. In addition, the nature of the carbene donor (i.e., diazoester vs. diazoacetone), and thus of the corresponding heme-carbene, has no noticeable effect on the kinetic role of the HAA step. To further investigate the role of the non-native serine axial ligand, the same KIE experiments were also carried out using cystine-ligated versions of the aforementioned CYP119 variants, i.e., **CHI-EDA** (S317C) and **CHI-DA** (S317C) (**Figure 6b/c**). Interestingly, much smaller KIE values were obtained in both cases compared to the serine-ligated counterparts. For the reaction with EDA, KIE values of 2.34-2.79 were obtained from competitive and non-competitive reactions against a value of 4.36-4.43 measured with **CHI-EDA**. Similarly, the reactions of the cysteine-ligated **CHI-DA** (S317C) enzyme with diazoacetone yielded moderate to no KIE of 1.39 (compet.) and 1.06 (non-compet.), against a much larger KIE of 4.37-4.42 observed for the serine-ligated counterpart. These differences indicate a notable change in the kinetic impact of C–H bond cleavage as a result of the change in heme axial coordination environment, further highlighting the often critical role of the axial ligand for influencing reactivity and the mechanism of hemoprotein-based carbene transferases.[72-74]

A quantum chemical study was performed to further examine the mechanism of C–H carbene insertion catalyzed by the engineered P450s described here. DFT calculations were carried out using the previous method that accurately predicted various experimental structures and reactivities of heme carbenes,[75-81] using [Fe(Por)(MeO^-)] to mimic the active site core part of the biocatalyst as done previously,[79-81] where Por is a non-substituted porphyrin and MeO^- is the model for the Ser ligand.[82-85] Since conformations and spin states may influence the reaction results,[77-78, 80, 86] we first conducted a detailed study of these effects involving both the substrate and heme catalyst to select the most favorable conformations and spin states of the species along the reaction pathways, as described in more detail in the Supporting Information.

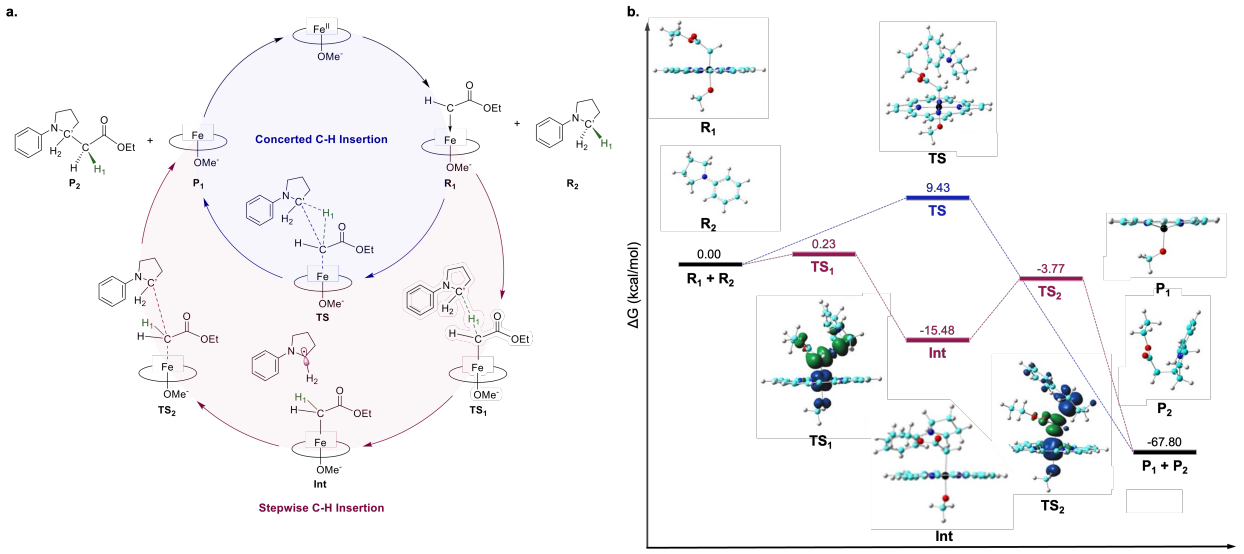


Figure 7. (**a**) Concerted and stepwise pathways for heme catalyzed C–H insertion of *N*-phenylpyrrolidine. Oval represents porphyrin. (**b**) Schematic free energy diagram for heme catalyzed C–H insertion. Spin densities of the transition states are shown. Atom color scheme: C-cyan, N-blue, O- red, H-grey. (spin density contour value: ± 0.004 au)

As shown in **Figure 7a**, both concerted and stepwise mechanisms were investigated. The concerted reaction pathway is similar to that reported previously on C–H insertions catalyzed by a different iron porphyrin carbene[77] and it features a simultaneous hydride transfer (as evidenced by a significant negative charge transfer from substrate to carbene, -0.441 e), C–C' bond formation, and Fe–C bond breaking. In contrast, the stepwise reaction pathway entails a hydrogen atom transfer (HAT) from the substrate to the heme carbene to form a carbon-center radical intermediate, followed by radical rebound to form the new C–C' bond, while breaking the Fe–C bond (**Figure 7b**). In the present system, a hydrogen atom transfer feature can be seen from the increase in the C'–H_1 bond length from 1.095 Å in **R_2** to 1.354 Å in **TS_1** and then to 3.399 Å in **Int**, while C–H_1 bond length decreases from 1.338 Å in **TS_1** to 1.086 Å in **Int** (See **SI Table S20**). The most favorable spin state calculated for **TS_1** is the open-shell singlet, where the radical is equally shared between C (-0.496 e) and C' (-0.445 e), whereas the hydrogen atom carries spin densities in the opposite direction (0.017 e), indicating partial transfer of the hydrogen atom. In this electronic state, Fe is in the ferric form with spin densities of +0.995 e (**SI Table S22**). In the intermediate **Int**, after donating the hydrogen atom, the substrate has the radical with spin densities of 0.995 e,

see **SI Table S22**). A subsequent radical rebound lead to the formation of the C–C' bond in **TS_2**, which has relatively lower energy compared to the hydrogen atom transfer step (-3.77 kcal/mol, see **Figure 7b**). In this step, again both C and C' show radical feature, but with opposite spin directions (-0.438 e and 0.641 e respectively), ready for a radical coupling to facilitate the formation of the final product. While partial C–C' bond formation can be seen by its distance shortening from 4.436 Å in **Int** to 2.543 Å in **TS_2**, a concomitant Fe–C bond elongation of ~0.3 Å in this step also indicates partial cleavage of that bond, which proceeds to the final release of the product.

Overall, our calculations indicate that the stepwise reaction pathway exhibits a significantly lower energy barrier (>9 kcal/mol difference) compared to the concerted reaction pathway (See **Figure 7b**), suggesting that the stepwise radical mechanism is preferred. Furthermore, the calculated KIE values for the concerted C–H carbene insertion step via **TS** and for the first step of stepwise pathway (**TS_1**) are 2.14 and 4.63, respectively, the latter being closer to the experimentally determined KIE values of 4.3-4.6 for this reaction (**Figure 6**). To further understand the origin of the reactivity differences between these two mechanisms, the geometric data were examined in more detail. As seen in **SI Table S11**, the overall structures of **TS** and **TS_1** are similar except that **TS** has a significantly longer Fe–L (1.987 Å) and shorter Fe–C (1.946 Å) bonds compared to **TS_1** (Fe–L: 1.912Å, Fe–C: 1.982 Å). On the one hand, the shorter Fe–L distance in **TS_1** results in a stronger *trans* effect that pushes the carbene moiety away from iron center (longer Fe–C bond), thus facilitating attack on the substrate. On the other hand, breaking of the shorter Fe–C bond (1.946 Å) in the concerted **TS** is associated with a higher energy cost compared to cleavage of the Fe–C bond in the radical pathway **TS_2**, which is longer (2.345 Å). Altogether, these structural features contribute to favor the stepwise radical mechanism for this biocatalytic transformation.

Conclusion

In summary, we have developed a new, efficient biocatalytic strategy for the enantioselective C(sp^3)–H functionalization of indoline scaffolds via P450-mediated carbene transfer. This method is amenable to the transformation of a variety of variously substituted indolines with diazoacetate and its scope extends to include other carbene donor reagents, such as diazoacetonitrile and diazoketone. Importantly, this work also demonstrates the possibility of achieving regiodivergent selectivity in enzyme-catalyzed C–H carbene insertion, through the

regio- and enantioselective functionalization of up to three distinct C(sp^3)–H sites in an indoline substrate (**Figure 1c**). While regiodivergent P450-catalyzed C(sp^3)–H oxidation has been previously achieved through the native monooxygenase reactivity of these enzymes,[31-32, 34, 36] regiodivergent C(sp^3)–H functionalization via carbene transfer has represented a major challenge[23, 45-47, 49, 51] and has been so far largely elusive to engineered biological catalysts. On the one hand, this capability can provide rapid access to enantioenriched indoline-derived constitutional isomers, which can be valuable building blocks for medicinal chemistry and whose synthesis would require multi-step syntheses.[52] On the other hand, as demonstrated through the chemoenzymatic synthesis of **9a**, these regiodivergent carbene transferases can be leveraged to afford more complex, polycyclic indoline-based scaffolds akin to those found in various bioactive molecules.[3, 53-55] Regiodivergent selectivity is also of particular interest in the construction of compound libraries for drug discovery campaigns.[87-88]

Our mechanistic studies provide valuable, first-time insights into the mechanism of hemoprotein-catalyzed carbene C(sp^3)–H insertion, and collectively support the involvement of a radical, stepwise pathway, akin to the mechanism of native P450-catalyzed hydroxylation reactions[89] and non-native C–H amination reactions via nitrene transfer catalyzed by engineered P450s and other hemoproteins.[90-92] Overall, this work expands the methodological toolbox for realizing selective C(sp^3)–H functionalization via enzyme-mediated carbene transfer and this along with the mechanistic understanding of productive and non-productive pathways in these reactions are expected to guide future development of new and improved biocatalysts for this important class of transformations.

Experimental Section

Resource availability

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Rudi Fasan (rudi.fasan@utdallas.edu).

Materials availability

All reagents in this study are commercially available or can be prepared as indicated in the supplemental information.

Data and code availability

All supporting data and findings are available in the manuscript or supplemental information. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Acknowledgements

This work was supported by the U.S. National Institute of Health grant GM098628 (R.F.). R.F. acknowledges chair endowment support from the Robert A. Welch Foundation (Chair, AT-0051). Y.Z. acknowledges financial support from the NSF grant CHE-2054897 and technical support from Xinyi Zhao and Emily McGuire for preliminary calculations. The authors are grateful to the UTD Center for High-Throughput Reaction Discovery & Synthesis supported by grant RR230018 from the Cancer Prevention and Research Institute of Texas.

Author Contributions

B.M.C. conducted experiments and prepared the manuscript and supplemental information. R.C. aided in product isolation and reaction optimization. R.F. supervised the project and contributed to the preparation of the manuscript. J.-M.C. and Y.Z. are responsible for the computational calculations and computational section of the manuscript.

Declaration of Interests

The authors declare no competing interests.

References:

1. Taylor, R. D.; MacCoss, M.; Lawson, A. D., Rings in drugs: Miniperspective. *Journal of medicinal chemistry* **2014,** *57* (14), 5845-5859.
2. Song, J.; Chen, D.-F.; Gong, L.-Z., Recent progress in organocatalytic asymmetric total syntheses of complex indole alkaloids. *National Science Review* **2017,** *4* (3), 381-396.
3. Silva, T. S.; Rodrigues Jr, M. T.; Santos, H.; Zeoly, L. A.; Almeida, W. P.; Barcelos, R. C.; Gomes, R. C.; Fernandes, F. S.; Coelho, F., Recent advances in indoline synthesis. *Tetrahedron* **2019,** *75* (14), 2063-2097.
4. Wei, H.; Li, B.; Wang, N.; Ma, Y.; Yu, J.; Wang, X.; Su, J.; Liu, D., Development and Application of Indolines in Pharmaceuticals. *ChemistryOpen* **2023,** *12* (2), e202200235.
5. Chandrasekhar, S.; Basu, D.; Reddy, C. R., Palladium-catalyzed reduction of N-(tert-butoxycarbonyl) indoles by polymethylhydrosiloxane. *Synthesis* **2007,** *2007* (10), 1509-1512.
6. Wang, L.; Shao, Y.; Liu, Y., Nucleophilic Addition of Grignard Reagents to 3–Acylindoles: Stereoselective Synthesis of Highly Substituted Indoline Scaffolds. *Organic Letters* **2012,** *14* (15), 3978-3981.
7. Nieto, J. M.; Lupton, K. H., Indole and Indoline Scaffolds in Antimicrobials: Overview, Synthesis and Recent Advances in Antimicrobial Research. *Current Medicinal Chemistry* **2021,** *28* (24), 4828-4844.
8. Boger, D. L.; Coleman, R. S., Intramolecular Diels-Alder reactions of 1, 2-diazines: general indoline synthesis. Studies on the preparation of the central and right-hand segments of CC-1065. *The Journal of Organic Chemistry* **1984,** *49* (12), 2240-2245.
9. Presset, M.; Pignon, A.; Paul, J.; Le Gall, E.; Léonel, E.; Martens, T., Synthesis of Indolines by a Zn-Mediated Mannich Reaction/Pd-Catalyzed Amination Sequence. *The Journal of Organic Chemistry* **2017,** *82* (6), 3302-3310.
10. Li, Y.; Chang, Y.; Li, Y.; Cao, C.; Yang, J.; Wang, B.; Liang, D., Iron-Catalyzed exo-Selective Synthesis of Cyanoalkyl Indolines via Cyanoisopropylarylation of Unactivated Alkenes. *Advanced Synthesis & Catalysis* **2018,** *360* (13), 2488-2492.
11. Liang, D.; Ge, D.; Lv, Y.; Huang, W.; Wang, B.; Li, W., Silver-catalyzed radical arylphosphorylation of unactivated alkenes: synthesis of 3-phosphonoalkyl indolines. *The Journal of Organic Chemistry* **2018,** *83* (8), 4681-4691.
12. Guillemard, L.; Kaplaneris, N.; Ackermann, L.; Johansson, M. J., Late-stage C-H functionalization offers new opportunities in drug discovery. *Nat. Rev. Chem.* **2021,** *5* (8), 522-545.
13. Fraley, A. E.; Sherman, D. H., Halogenase engineering and its utility in medicinal chemistry. *Bioorg Med Chem Lett* **2018,** *28* (11), 1992-1999.
14. Latham, J.; Brandenburger, E.; Shepherd, S. A.; Menon, B. R. K.; Micklefield, J., Development of Halogenase Enzymes for Use in Synthesis. *Chem Rev* **2018,** *118* (1), 232-269.
15. Zhang, R. K.; Huang, X.; Arnold, F. H., Selective CH bond functionalization with engineered heme proteins: new tools to generate complexity. *Curr Opin Chem Biol* **2019,** *49*, 67-75.
16. Zwick, C. R.; Renata, H., Harnessing the biocatalytic potential of iron- and α-ketoglutarate-dependent dioxygenases in natural product total synthesis. *Natural Product Reports* **2020,** *37* (8), 1065-1079.
17. Ren, X.; Fasan, R., Engineered and artificial metalloenzymes for selective C–H functionalization. *Current Opinion in Green and Sustainable Chemistry* **2021,** *31*, 100494.

18. Wang, W.; Taber, D. F.; Renata, H., Practical Enzymatic Production of Carbocycles. *Chemistry – A European Journal* **2021,** *27* (46), 11773-11794.
19. Pyser, J. B.; Chakrabarty, S.; Romero, E. O.; Narayan, A. R. H., State-of-the-Art Biocatalysis. *ACS Central Science* **2021,** *7* (7), 1105-1116.
20. Couture, B.; Chattopadhyay, A.; Fasan, R., Biocatalytic Carbene and Nitrene Transfer Reactions. In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*, Elsevier: 2023.
21. Reetz, M. T.; Qu, G.; Sun, Z., Engineered enzymes for the synthesis of pharmaceuticals and other high-value products. *Nature Synthesis* **2024,** *3* (1), 19-32.
22. Hayashi, T.; Ligibel, M.; Sager, E.; Voss, M.; Hunziker, J.; Schroer, K.; Snajdrova, R.; Buller, R., Evolved aliphatic halogenases enable regiocomplementary C− H functionalization of a pharmaceutically relevant compound. *Angewandte Chemie International Edition* **2019,** *58* (51), 18535-18539.
23. Ren, X.; Couture, B. M.; Liu, N.; Lall, M. S.; Kohrt, J. T.; Fasan, R., Enantioselective Single and Dual α-C–H Bond Functionalization of Cyclic Amines via Enzymatic Carbene Transfer. *Journal of the American Chemical Society* **2023,** *145* (1), 537-550.
24. Zhang, J.; Maggiolo, A. O.; Alfonzo, E.; Mao, R.; Porter, N. J.; Abney, N. M.; Arnold, F. H., Chemodivergent C(sp3)–H and C(sp2)–H cyanomethylation using engineered carbene transferases. *Nature Catalysis* **2023,** *6* (2), 152-160.
25. Renata, H.; Shimizu, E.; Zwick, C. R., Regiodivergent biocatalytic hydroxylation of l-Glutamine facilitated by characterization of Non-Heme dioxygenases from Non-Ribosomal peptide biosyntheses. *Tetrahedron* **2021,** *90*, 132190.
26. Craven, E. J.; Latham, J.; Shepherd, S. A.; Khan, I.; Diaz-Rodriguez, A.; Greaney, M. F.; Micklefield, J., Programmable late-stage C−H bond functionalization enabled by integration of enzymes with chemocatalysis. *Nature Catalysis* **2021,** *4* (5), 385-394.
27. Zetzsche, L. E.; Chakrabarty, S.; Narayan, A. R. H., Development of a P450 Fusion Enzyme for Biaryl Coupling in Yeast. *ACS Chemical Biology* **2022,** *17* (11), 2986-2992.
28. Espinoza, R. V.; Haatveit, K. C.; Grossman, S. W.; Tan, J. Y.; McGlade, C. A.; Khatri, Y.; Newmister, S. A.; Schmidt, J. J.; Garcia-Borràs, M.; Montgomery, J.; Houk, K. N.; Sherman, D. H., Engineering P450 TamI as an Iterative Biocatalyst for Selective Late-Stage C-H Functionalization and Epoxidation of Tirandamycin Antibiotics. *ACS Catal* **2021,** *11* (13), 8304-8316.
29. Andorfer, M. C.; Park, H. J.; Vergara-Coll, J.; Lewis, J. C., Directed evolution of RebH for catalyst-controlled halogenation of indole C–H bonds. *Chemical science* **2016,** *7* (6), 3720-3729.
30. Shepherd, S. A.; Karthikeyan, C.; Latham, J.; Struck, A.-W.; Thompson, M. L.; Menon, B. R. K.; Styles, M. Q.; Levy, C.; Leys, D.; Micklefield, J., Extending the biocatalytic scope of regiocomplementary flavin-dependent halogenase enzymes. *Chemical Science* **2015,** *6* (6), 3454-3460.
31. Kille, S.; Zilly, F. E.; Acevedo, J. P.; Reetz, M. T., Regio-and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution. *Nature Chemistry* **2011,** *3* (9), 738-743.
32. Zhang, K.; Shafer, B. M.; Demars, M. D.; Stern, H. A.; Fasan, R., Controlled oxidation of remote sp3 C–H bonds in artemisinin via P450 catalysts with fine-tuned regio-and stereoselectivity. *Journal of the American Chemical Society* **2012,** *134* (45), 18695-18704.

33. Greule, A.; Stok, J. E.; De Voss, J. J.; Cryle, M. J., Unrivalled diversity: the many roles and reactions of bacterial cytochromes P450 in secondary metabolism. *Natural product reports* **2018,** *35* (8), 757-791.
34. Alwaseem, H.; Frisch, B. J.; Fasan, R., Anticancer activity profiling of parthenolide analogs generated via P450-mediated chemoenzymatic synthesis. *Bioorganic & Medicinal Chemistry* **2018,** *26* (7), 1365-1373.
35. Lukowski, A. L.; Ellinwood, D. C.; Hinze, M. E.; DeLuca, R. J.; Du Bois, J.; Hall, S.; Narayan, A. R., C–H hydroxylation in paralytic shellfish toxin biosynthesis. *Journal of the American Chemical Society* **2018,** *140* (37), 11863-11869.
36. Zhang, X.; King-Smith, E.; Dong, L.-B.; Yang, L.-C.; Rudolf, J. D.; Shen, B.; Renata, H., Divergent synthesis of complex diterpenes through a hybrid oxidative approach. *Science* **2020,** *369* (6505), 799-806.
37. Alwaseem, H.; Giovani, S.; Crotti, M.; Welle, K.; Jordan, C. T.; Ghaemmaghami, S.; Fasan, R., Comprehensive Structure–Activity Profiling of Micheliolide and its Targeted Proteome in Leukemia Cells via Probe-Guided Late-Stage C–H Functionalization. *ACS Central Science* **2021,** *7* (5), 841-857.
38. Iizaka, Y.; Arai, R.; Takahashi, A.; Ito, M.; Sakai, M.; Fukumoto, A.; Sherman, D. H.; Anzai, Y., Engineering sequence and selectivity of late-stage C-H oxidation in the MycG iterative cytochrome P450. *Journal of Industrial Microbiology and Biotechnology* **2022,** *49* (1), kuab069.
39. Turner, N. J., Directed evolution drives the next generation of biocatalysts. *Nature chemical biology* **2009,** *5* (8), 567-573.
40. Reetz, M. T., Laboratory evolution of stereoselective enzymes: a prolific source of catalysts for asymmetric reactions. *Angewandte Chemie International Edition* **2011,** *50* (1), 138-174.
41. Dydio, P.; Key, H. M.; Nazarenko, A.; Rha, J. Y.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F., An artificial metalloenzyme with the kinetics of native enzymes. *Science* **2016,** *354* (6308), 102-106.
42. Brandenberg, O. F.; Fasan, R.; Arnold, F. H., Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions. *Curr Opin Biotechnol* **2017,** *47*, 102-111.
43. Andorfer, M. C.; Lewis, J. C., Understanding and improving the activity of flavin-dependent halogenases via random and targeted mutagenesis. *Annual review of biochemistry* **2018,** *87*, 159-185.
44. Vargas, D. A.; Tinoco, A.; Tyagi, V.; Fasan, R., Myoglobin-Catalyzed C-H Functionalization of Unprotected Indoles. *Angew Chem Int Ed Engl* **2018,** *57* (31), 9911-9915.
45. Gu, Y.; Natoli, S. N.; Liu, Z.; Clark, D. S.; Hartwig, J. F., Site-Selective Functionalization of (sp(3))C-H Bonds Catalyzed by Artificial Metalloenzymes Containing an Iridium-Porphyrin Cofactor. *Angew Chem Int Ed Engl* **2019,** *58* (39), 13954-13960.
46. Zhang, J.; Huang, X.; Zhang, R. K.; Arnold, F. H., Enantiodivergent α-Amino C-H Fluoroalkylation Catalyzed by Engineered Cytochrome P450s. *J Am Chem Soc* **2019,** *141* (25), 9798-9802.
47. Zhang, R. K.; Chen, K.; Huang, X.; Wohlschlager, L.; Renata, H.; Arnold, F. H., Enzymatic assembly of carbon–carbon bonds via iron-catalysed sp3 C–H functionalization. *Nature* **2019,** *565* (7737), 67-72.
48. Brandenberg, O. F.; Chen, K.; Arnold, F. H., Directed Evolution of a Cytochrome P450 Carbene Transferase for Selective Functionalization of Cyclic Compounds. *Journal of the American Chemical Society* **2019,** *141* (22), 8989-8995.

49. Zhou, A. Z.; Chen, K.; Arnold, F. H., Enzymatic Lactone-Carbene C–H Insertion to Build Contiguous Chiral Centers. *ACS Catalysis* **2020,** *10* (10), 5393-5398.
50. Rumo, C.; Stein, A.; Klehr, J.; Tachibana, R.; Prescimone, A.; Häussinger, D.; Ward, T. R., An Artificial Metalloenzyme Based on a Copper Heteroscorpionate Enables sp3 C–H Functionalization via Intramolecular Carbene Insertion. *Journal of the American Chemical Society* **2022,** *144* (26), 11676-11684.
51. Sreenilayam, G.; Moore, E. J.; Steck, V.; Fasan, R., Metal Substitution Modulates the Reactivity and Extends the Reaction Scope of Myoglobin Carbene Transfer Catalysts. *Adv Synth Catal* **2017,** *359* (12), 2076-2089.
52. Chhabra, N.; Aseri, M. L.; Padmanabhan, D., A review of drug isomerism and its significance. *International journal of applied and basic medical research* **2013,** *3* (1), 16.
53. Ruiz-Sanchis, P.; Savina, S. A.; Albericio, F.; Alvarez, M., Structure, Bioactivity and Synthesis of Natural Products with Hexahydropyrrolo [2, 3-b] indole. *Chemistry–A European Journal* **2011,** *17* (5), 1388-1408.
54. Zi, W.; Zuo, Z.; Ma, D., Intramolecular dearomative oxidative coupling of indoles: a unified strategy for the total synthesis of indoline alkaloids. *Accounts of Chemical Research* **2015,** *48* (3), 702-711.
55. Griffiths, B. M.; Burl, J. D.; Wang, X., Bioinspired Discovery of Chemical Reactions and Biological Probes. *Synlett* **2016,** *27* (14), 2039-2042.
56. Liu, F.; Su, M., Chapter 8 - Indole and indoline scaffolds in drug discovery. In *Privileged Scaffolds in Drug Discovery*, Yu, B.; Li, N.; Fu, C., Eds. Academic Press: 2023; pp 147-161.
57. Coelho, P. S.; Wang, Z. J.; Ener, M. E.; Baril, S. A.; Kannan, A.; Arnold, F. H.; Brustad, E. M., A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins in vivo. *Nat Chem Biol* **2013,** *9* (8), 485-7.
58. Doyle, M. P.; Duffy, R.; Ratnikov, M.; Zhou, L., Catalytic carbene insertion into C− H bonds. *Chemical reviews* **2010,** *110* (2), 704-724.
59. Park, S.-Y.; Yamane, K.; Adachi, S.-i.; Shiro, Y.; Weiss, K. E.; Maves, S. A.; Sligar, S. G., Thermophilic cytochrome P450 (CYP119) from Sulfolobus solfataricus: high resolution structure and functional properties. *Journal of Inorganic Biochemistry* **2002,** *91* (4), 491-501.
60. Weaver, B. R.; Perkins, L. J.; Fernandez Candelaria, F. O.; Burstyn, J. N.; Buller, A. R., Molecular Determinants of Efficient Cobalt-Substituted Hemoprotein Production in E. coli. *ACS Synthetic Biology* **2023,** *12* (12), 3669-3679.
61. Key, H. M.; Dydio, P.; Liu, Z.; Rha, J. Y.; Nazarenko, A.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F., Beyond Iron: Iridium-Containing P450 Enzymes for Selective Cyclopropanations of Structurally Diverse Alkenes. *ACS Cent Sci* **2017,** *3* (4), 302-308.
62. Knight, A. M.; Kan, S. B. J.; Lewis, R. D.; Brandenberg, O. F.; Chen, K.; Arnold, F. H., Diverse Engineered Heme Proteins Enable Stereodivergent Cyclopropanation of Unactivated Alkenes. *ACS Central Science* **2018,** *4* (3), 372-377.
63. Garlets, Z. J.; Wertz, B. D.; Liu, W.; Voight, E. A.; Davies, H. M. L., Regio- and Stereoselective Rhodium(II)-Catalyzed C–H Functionalization of Cyclobutanes. *Chem* **2020,** *6* (1), 304-313.
64. Boni, Y. T.; Cammarota, R. C.; Liao, K.; Sigman, M. S.; Davies, H. M. L., Leveraging Regio- and Stereoselective C(sp3)–H Functionalization of Silyl Ethers to Train a Logistic Regression Classification Model for Predicting Site-Selectivity Bias. *Journal of the American Chemical Society* **2022,** *144* (34), 15549-15561.

65. Hock, K. J.; Knorrscheidt, A.; Hommelsheim, R.; Ho, J.; Weissenborn, M. J.; Koenigs, R. M., Tryptamine Synthesis by Iron Porphyrin Catalyzed C-H Functionalization of Indoles with Diazoacetonitrile. *Angew Chem Int Ed Engl* **2019,** *58* (11), 3630-3634.
66. Wang, Z. J.; Peck, N. E.; Renata, H.; Arnold, F. H., Cytochrome P450-catalyzed insertion of carbenoids into N–H bonds. *Chemical Science* **2014,** *5* (2), 598-601.
67. Sreenilayam, G.; Fasan, R., Myoglobin-catalyzed intermolecular carbene N–H insertion with arylamine substrates. *Chemical Communications* **2015,** *51* (8), 1532-1534.
68. Steck, V.; Sreenilayam, G.; Fasan, R., Selective Functionalization of Aliphatic Amines via Myoglobin-catalyzed Carbene N-H Insertion. *Synlett* **2020,** *31* (3), 224-229.
69. Vargas, D. A.; Khade, R. L.; Zhang, Y.; Fasan, R., Biocatalytic Strategy for Highly Diastereo- and Enantioselective Synthesis of 2,3-Dihydrobenzofuran-Based Tricyclic Scaffolds. *Angew Chem Int Ed Engl* **2019,** *58* (30), 10148-10152.
70. Lee, W.-C. C.; Wang, D.-S.; Zhu, Y.; Zhang, X. P., Iron(III)-based metalloradical catalysis for asymmetric cyclopropanation via a stepwise radical mechanism. *Nature Chemistry* **2023,** *15* (11), 1569-1580.
71. Moore, E. J.; Fasan, R., Effect of proximal ligand substitutions on the carbene and nitrene transferase activity of myoglobin. *Tetrahedron* **2019,** *75* (16), 2357-2363.
72. Chen, K.; Zhang, S. Q.; Brandenberg, O. F.; Hong, X.; Arnold, F. H., Alternate Heme Ligation Steers Activity and Selectivity in Engineered Cytochrome P450-Catalyzed Carbene-Transfer Reactions. *J Am Chem Soc* **2018,** *140* (48), 16402-16407.
73. Tinoco, A.; Wei, Y.; Bacik, J. P.; Carminati, D. M.; Moore, E. J.; Ando, N.; Zhang, Y.; Fasan, R., Origin of high stereocontrol in olefin cyclopropanation catalyzed by an engineered carbene transferase. *ACS Catal* **2019,** *9* (2), 1514-1524.
74. Carminati, D. M.; Fasan, R., Stereoselective Cyclopropanation of Electron-Deficient Olefins with a Cofactor Redesigned Carbene Transferase Featuring Radical Reactivity. *ACS Catalysis* **2019,** *9* (10), 9683-9697.
75. Khade, R. L.; Fan, W.; Ling, Y.; Yang, L.; Oldfield, E.; Zhang, Y., Iron Porphyrin Carbenes as Catalytic Intermediates: Structures, Mossbauer and NMR Spectroscopic Properties, and Bonding. *Angew. Chem. Int. Ed.* **2014,** *53*, 7574-7578.
76. Khade, R. L.; Zhang, Y., Catalytic and Biocatalytic Iron Porphyrin Carbene Formation: Effects of Binding Mode, Carbene Substituent, Porphyrin Substituent, and Protein Axial Ligand. *J. Am. Chem. Soc.* **2015,** *137* (24), 7560-7563.
77. Khade, R. L.; Zhang, Y., C-H Insertions by Iron Porphyrin Carbene: Basic Mechanism and Origin of Substrate Selectivity. *Chemistry - A Euopean Journal* **2017,** *23*, 17654-17658
78. Wei, Y.; Tinoco, A.; Steck, V.; Fasan, R.; Zhang, Y., Cyclopropanations via Heme Carbenes: Basic Mechanism and Effects of Carbene Substituent, Protein Axial Ligand, and Porphyrin Substitution. *J. Am. Chem. Soc.* **2018,** *140*, 1649-1662
79. Tinoco, A.; Wei, Y.; Bacik, J. P.; Moore, E. J.; Ando, N.; Zhang, Y.; Fasan, R., Origin of high stereocontrol in olefin cyclopropanation catalyzed by an engineered carbene transferase. *ACS Catal.* **2019,** *9*, 1514-1524
80. Khade, R., L.; Chandgude, A. L.; Fasan, R.; Zhang, Y., Mechanistic Investigation of Biocatalytic Heme Carbenoid Si-H Insertions. *ChemCatChem* **2019,** *11*, 3101-3108.
81. Vargas, D. A.; Khade, R. L.; Zhang, Y.; Fasan, R., Biocatalytic strategy for highly diastereo- and enantioselective synthesis of 2,3-dihydrobenzofuran based tricyclic scaffolds. *Angew. Chem. Int. Ed.* **2019,** *58*, 10148-10152.

82. Li, Z.; Burnell, D. J.; Boyd, R. J., Computational Study of Engineered Cytochrome P450-Catalyzed C–H Amination: The Origin of the Regio- and Stereoselectivity. *The Journal of Physical Chemistry Part B* **2017,** *121* (48), 10859-10868.
83. Wang, J.; Gao, H.; Yang, L.; Gao, Y. Q., Role of Engineered Iron-haem Enzyme in Reactivity and Stereoselectivity of Intermolecular Benzylic C–H Bond Amination. *ACS Catal.* **2020,** *10* (9), 5318-5327.
84. Li, X.; Dong, L.; Liu, Y., Theoretical Study of Iron Porphyrin Nitrene: Formation Mechanism, Electronic Nature, and Intermolecular C–H Amination. *Inorg. Chem.* **2020,** *59* (3), 1622-1632.
85. Athavale, S. V.; Gao, S.; Liu, Z.; Mallojjala, S. C.; Hirschi, J. S.; Arnold, F. H., Biocatalytic, Intermolecular C−H Bond Functionalization for the Synthesis of Enantioenriched Amides. *Angew. Chem. Int. Ed.* **2021,** *60* (47), 24864-24869.
86. Wei, Y.; Conklin, M.; Zhang, Y., Biocatalytic Intramolecular C−H aminations via Engineered Heme Proteins: Full Reaction Pathways and Axial Ligand Effects. *Chemistry – A European Journal* **2022,** *28* (59), e202202006.
87. Gerry, C. J.; Wawer, M. J.; Clemons, P. A.; Schreiber, S. L., DNA Barcoding a Complete Matrix of Stereoisomeric Small Molecules. *J Am Chem Soc* **2019,** *141* (26), 10225-10235.
88. Bassi, G.; Favalli, N.; Vuk, M.; Catalano, M.; Martinelli, A.; Trenner, A.; Porro, A.; Yang, S.; Tham, C. L.; Moroglu, M.; Yue, W. W.; Conway, S. J.; Vogt, P. K.; Sartori, A. A.; Scheuermann, J.; Neri, D., A Single-Stranded DNA-Encoded Chemical Library Based on a Stereoisomeric Scaffold Enables Ligand Discovery by Modular Assembly of Building Blocks. *Adv Sci* **2020,** *7* (22).
89. Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I., Structure and chemistry of cytochrome P450. *Chem Rev* **2005,** *105* (6), 2253-2277.
90. Singh, R.; Kolev, J. N.; Sutera, P. A.; Fasan, R., Enzymatic C(sp(3))-H Amination: P450-Catalyzed Conversion of Carbonazidates into Oxazolidinones. *ACS Catal.* **2015,** *5* (3), 1685-1691.
91. Yang, Y.; Cho, I.; Qi, X. T.; Liu, P.; Arnold, F. H., An Enzymatic Platform for the Asymmetric Amination of Primary, Secondary and Tertiary C(sp(3))—H Bonds. *Nat. Chem.* **2019,** *11* (11), 987-993.
92. Roy, S.; Vargas, D. A.; Ma, P.; Sengupta, A.; Zhu, L.; Houk, K. N.; Fasan, R., Stereoselective Construction of β-, γ-, and δ-Lactam Rings via Enzymatic C-H Amidation. *Res Sq* **2023**.