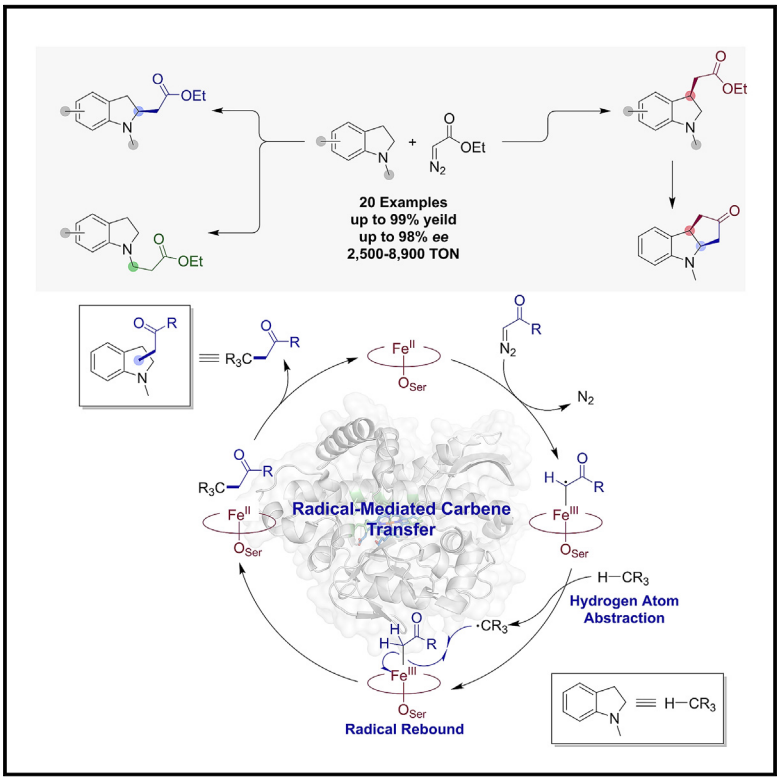
Chem Catalysis

Article

Radical-mediated regiodivergent C(sp^3)–H functionalization of *N*-substituted indolines via enzymatic carbene transfer

Graphical abstract



Highlights

- Biocatalytic approach for stereoselective C(sp^3)–H carbene insertion in indolines
- Tunable regioselectivity toward three distinct C(sp^3)–H sites via protein engineering
- Synthesis of polycyclic indoline-containing building blocks via biocatalytic cascade
- First mechanistic insights into hemoprotein-catalyzed C(sp^3)–H carbene insertion

Authors

Bo M. Couture, Ru Cui, Jia-Min Chu, Zhuofan Shen, Sagar D. Khare, Yong Zhang, Rudi Fasan

Correspondence

yzhang37@stevens.edu (Y.Z.), rudi.fasan@utdallas.edu (R.F.)

In brief

This work reports the development of an efficient biocatalytic strategy for the stereoselective and regiodivergent C(sp^3)–H functionalization of N-substituted indolines, a privileged scaffold in bioactive molecules, using engineered CYP119-based "carbene transferases." A biocatalytic cascade combining two engineered biocatalysts with divergent regioselectivity enabled the stereoselective synthesis of polycyclic indoline-based scaffolds found in natural products. Finally, computational and experimental mechanistic studies offer first-time insights into the mechanism of hemoprotein-catalyzed carbene C(sp^3)–H insertion and protein-controlled site selectivity in this class of transformations.

Couture et al., 2024, Chem Catalysis *4*, 101133
November 21, 2024
https://doi.org/10.1016/j.checat.2024.101133

CellPress

Chem Catalysis

CellPress

Article

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] Zhuofan Shen,[3] Sagar D. Khare,[3,4] Yong Zhang,[2,*] and Rudi Fasan[1,5,*]
[1]Department of Chemistry and Biochemistry, University of Texas at Dallas, Richardson, TX 75080, USA
[2]Department of Chemistry and Chemical Biology, Stevens Institute of Technology, Hoboken, NJ 07030, USA
[3]Department of Chemistry and Chemical Biology, Rutgers, the State University of New Jersey, Piscataway, NJ 08854, USA
[4]Institute for Quantitative Biomedicine, Rutgers, the State University of New Jersey, Piscataway, NJ 08854, USA
[5]Lead contact
*Correspondence: yzhang37@stevens.edu (Y.Z.), rudi.fasan@utdallas.edu (R.F.)
https://doi.org/10.1016/j.checat.2024.101133

THE BIGGER PICTURE Selective C(sp^3)–H functionalization is a powerful approach for the synthesis and diversification of organic molecules. While regiodivergent enzymatic C(sp^3)–H oxidation has been achieved with monooxygenase enzymes, regiodivergent C(sp^3)–H functionalization via "abiological" carbene transfer chemistry has been elusive to biocatalysis and challenging for chemocatalysts. Herein, we report an efficient biocatalytic strategy for the highly stereoselective and regiodivergent C(sp^3)–H carbene insertion in indoline substrates, which are ubiquitous motifs in drugs and natural products. These regiodivergent "carbene transferases" were further leveraged in a biocatalytic cascade to afford polycyclic indoline-based scaffolds. Our studies also provide first-time insight into the mechanism of hemoprotein-catalyzed C(sp^3)–H insertion and protein-mediated regiocontrol in these reactions, which can guide future development of biological catalysts for this important class of transformations.

SUMMARY

Indolines are ubiquitous structural motifs occurring in pharmaceuticals and natural products. Here, we report a strategy for regio- and stereoselective C(sp^3)–H functionalization of *N*-substituted indolines via carbene transfer chemistry mediated by engineered CYP119-based catalysts. These systems offer high enantioselectivity and high catalytic efficiency, as well as regiodivergent selectivity, furnishing an efficient and convenient route for diversification of these important scaffolds via direct C(sp^3)–H functionalization. Selective functionalization of exocyclic C(sp^3)–H bond in *N*-methyl indolines was also achieved, and a biocatalytic cascade combining enzyme-mediated α- and β-C(sp^3)–H functionalization yielded a polycyclic indoline-containing motif found in drugs. Mechanistic and computational studies support a radical-mediated C–H functionalization pathway and provide insights into protein-mediated regiodivergent selectivity. Altogether, this work offers a direct and tunable strategy to access functionalized indolines as key building blocks for medicinal chemistry and natural product synthesis and provides first insights into the mechanism of P450-catalyzed C(sp^3)–H carbene insertion.

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 been elusive.[12]

Biocatalysis has attracted 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, unspecific peroxygenases, flavin-dependent halogenases, and non-heme iron-dependent dioxygenases and halogenases, have proved useful for the selective

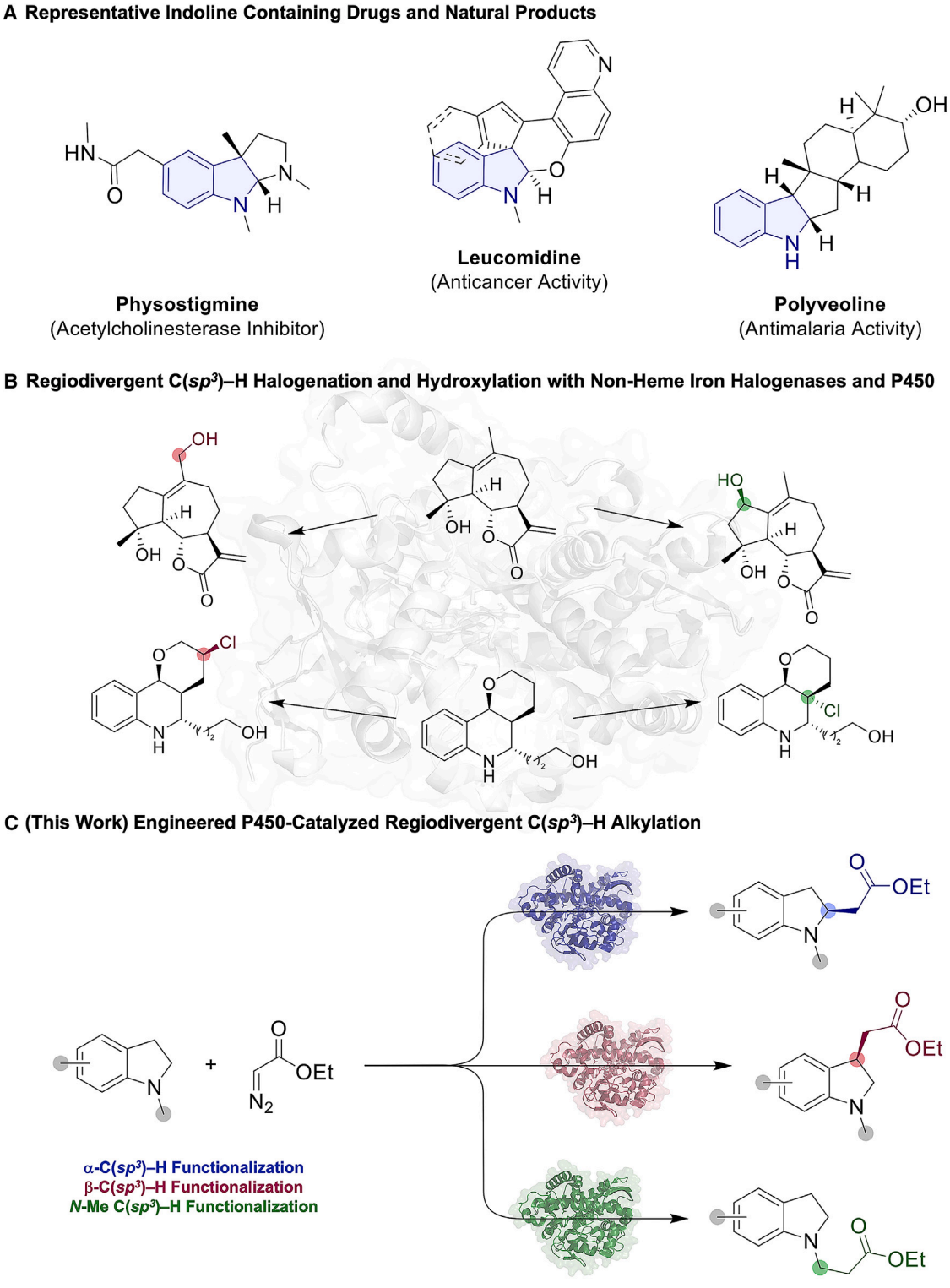


Figure 1. Biological significance of indolines and enzyme-catalyzed regiodivergent C(sp^3)–H functionalization

(A) Representative indoline-containing drugs and natural products.

(B) P450-catalyzed regiodivergent C(sp^3)–H hydroxylation.

(C) Engineered P450-catalyzed regiodivergent C(sp^3)–H alkylation.

functionalization of aromatic and aliphatic C–H bonds in small molecules.[13,14,16,18,21–29] Furthermore, through protein engineering, regiodivergent selectivity has been achieved, for example, for the halogenation of aromatic compounds using evolved flavin-dependent halogenases[22,30,31] or for late-stage C(sp^3)–H hydroxylation of complex natural products using engineered P450 enzymes and halogenases (Figure 1B),[32–39] thus creating new opportunities for the diversification and/or chemoenzymatic synthesis of these molecules. More recently, important progress has also been made in the development of biological catalysts for C–H functionalization via "non-native" chemistry such as carbene transfer catalysis.[40–50] 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.[46,47,50–52] 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.[45–47,50–52]

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. In addition to being compatible with a broad range of indoline substrates and diazo reagents, this methodology is 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[53] nor 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.[3,54–56] Mechanistic studies provide first-time insights into the mechanism of the present reaction and hemoprotein-catalyzed C(sp^3)–H carbene insertion.

RESULTS AND DISCUSSION

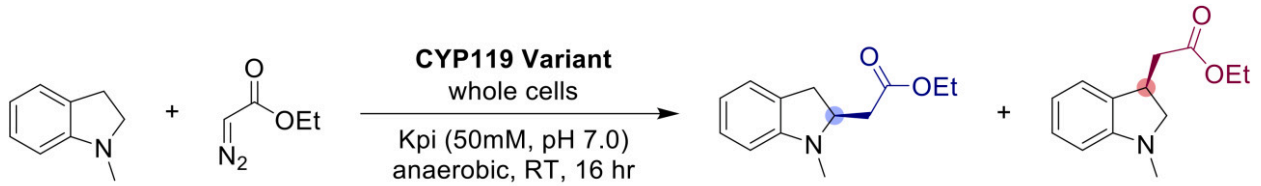
CYP119 biocatalyst for α-C–H functionalization of *N*-methyl indoline with EDA

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,[50] 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.[57] To this end, we began our investigation by screening a small panel (~100) of engineered CYP119 variants derived from our previous work,[50] including CYP119 (F153G, A209G, T213G, V254A, C317S) (called **CYP119-137**). 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[50] and other P450 scaffolds.[58]

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 **CYP119-137** 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 4). Along with the desired C–H functionalization products, the **CYP119-137**-catalyzed reaction was accompanied with the formation of various demethylation and desaturation/N–H insertion by-products (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,[59] such as $Rh_2(OAc)_4$, $Ru(BPY)_2$, Co(TPP), $Cu(OTf)_2$, and Fe(TPP), showed no product formation (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 **CYP119-137** 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.

Encouraged by these results, we aimed to identify CYP119 catalysts capable of favoring the formation of the 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 campaigns[50] 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 *Escherichia 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 5–7). Among them, 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 7; Figure 2B). These variants also exhibited appreciable enantiopreference for formation of the *S*-enantiomer (20%–24% *ee*; Table 1, entry 5–7). 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 **CYP119-168**, showed further enhanced catalytic activity (4,180 → 5,270 TON) and regioselectivity (87:13 → 92:8 r.r.) for formation of

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



1a 2a 3a 4a

Entry	Catalyst[a]	Yield (3a)[b]	3a:4a	TON (3a)[c]	e.r. (3a)[d]
1	hemin	0	0	–	–
2	"empty" *E. coli* cells	0	0	–	–
3	CYP119 (WT)	0	0	–	–
4	[**CYP119-137**] CYP119 (F153A, A209G, T213G, V254A, C317S)	17%	66:34	210	53:47
5	CYP119 (F153G, T213A, V254I, C317S)	26%	70:30	840	58:42
6	CYP119 (F153G, T213A, V254L, C317S)	20%	76:24	1,150	60:40
7	CYP119 (F153G, T213A, V254F, C317S)	73%	87:13	4,180	62:38
8	CYP119 (F153G, T213A, V254Y, C317S)	66%	84:16	3,210	78:12
9	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	92%	92:8	5,270	91:9
10[e]	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	42%	92:8	8,930	91:9
11[f]	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	92%	92:8	460	91:9
12[g]	[**CYP119-168**] 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 for product **3a** as calculated based on the protein concentration measured in cell lysate using the CO-binding assay.
[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 $Na_2S_2O_4$.

3a. In addition, **CYP119-168** 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 byproducts (Table S5).

Additional experiments indicated that up to 8,900 total turnovers (TTN) could be obtained for the **CYP119-168**-catalyzed α-C–H functionalization reaction using whole cells under catalyst limiting conditions (optical density 600 [OD_{600}] = 15) (Table 1, entry 10). 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 11) and similar results could be obtained using the same CYP119 variant in lyophilized form (Table 1, entry 12), thus demonstrating the robust nature of this biocatalyst to lyophilization and long-term storage, which are desirable attributes for preparative scale and industrial applications.

Given our prior findings on the C–H carbene insertion reactivity of cobalt-substituted myoglobins using phthalan as the substrate,[51] we investigated the effect of a similar metal substitution in the present CYP119-based biocatalysts, which were prepared via recombinant expression in the presence of Co-ppIX.[51,60] Interestingly, all of these Co-substituted variants showed activity toward the C–H functionalization of *N*-methyl indoline (**1a**) with EDA (**2a**) (Table S2). Although the activity and regioselectivity of these metallo-substituted variants were only comparable or slightly inferior compared to those 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**, **CYP119-168**, 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.[50,61,62] To address this limitation, we employed a substrate versus enzyme library approach, where a set of selected CYP119-derived carbene transferases from the **CYP119-168** 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).

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, 83:17–97:3 e.r.). Racemic 3-methyl-substituted indoline derivative **1b** was efficiently converted into the desired product **3b** with both excellent diastereo- and enantioselectivity (99:1 d.r., 98:2 e.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 (**3k–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). 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, unlike organometallic catalysts used for carbene transfer reactions (Figure 3B).

Whereas different CYP119 variants can be selected for optimal 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 toward the α-C–H functionalization of indoline-based derivatives. These results also revealed the role of steric bulk in promoting favorable binding conformations to target α-C–H functionalization compared to β-C–H functionalization. In the presence of bulkier *N*-alkyl substitutions, no β-C–H functionalization product is formed (**3f-o**). We also noted an inverse correlation between steric bulk in the substrate and in the catalyst active site, as catalysts that contain bulkier residues within the active site tend to exhibit lower activity toward bulkier substrates. For example, **CYP119-168** and **CYP119-156**, which contain a V254W and A209W mutation, showed basal activity toward substrate **1n**, bearing an *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.

Regiodivergent CYP119 catalysts for indoline C–H functionalization

During initial catalyst development for achieving selective α-C–H functionalization of *N*-methyl-indoline, we 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., **CYP119-137**: 66:34 **3a**:**4a**; Table 1, entry 4). 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 (Figures 2B and 2C).

Through this approach, we identified CYP119 (F153G, L205W, T213A, V254A, C317S), named **CYP119-235**, 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 **CYP119-235** 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 **CYP119-168** (Figure 4B).

Given the value of 2,3-difunctionalized indolines in medicinal chemistry (Figure 1) and 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 **CYP119-235** 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 (67:33 d.r., 79:21 e.r.), albeit with low efficiency (12% yield). Using a different CYP119 variant (**CYP119-137**), 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.

Upon exploration of the evolved variant **CYP119-235**, 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.[64,65]

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

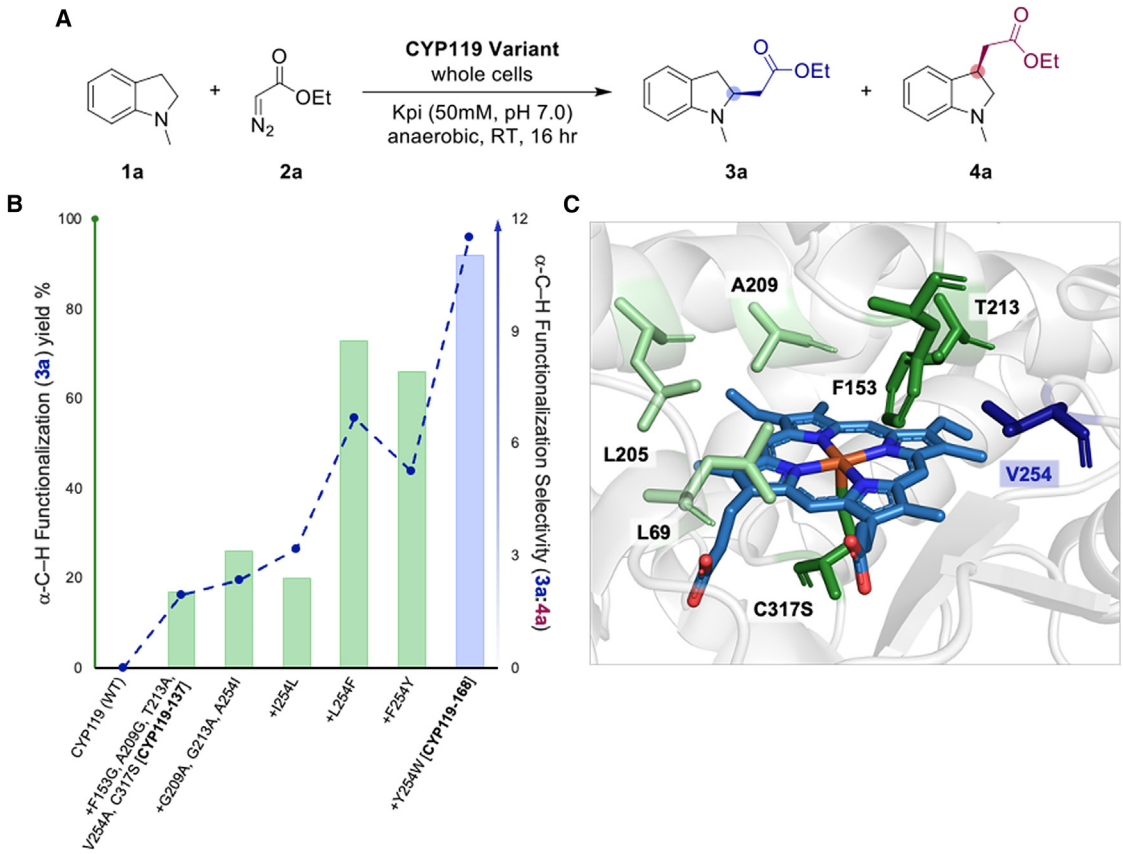


Figure 2. Directed evolution of CYP119 catalysts for regioselective C–H functionalization of *N*-methyl indoline (1a) with EDA (2a)

(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**.

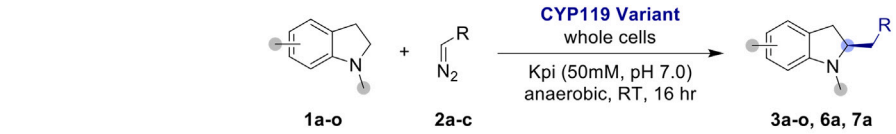
(B) Reconstructed directed evolution of CYP119 catalysts for regioselective C–H functionalization of *N*-methyl indoline (**1a**) with EDA (**2a**). Yields as determined under standard reaction conditions with EDA (Table 1).

(C) X-ray crystal structure of CYP119 from *Sulfolobus solfataricus* (PDB: 1IO7).[63] 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.

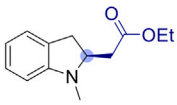
β-C–H selective biocatalysts can tolerate substitutions in the adjacent position in the indoline substrate targeted for functionalization (i.e., **3b** and **4b**). In addition, with product **6a**, we established that diazoketones are accepted as carbene donors by these catalysts. Based on these results, 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 **9a** (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**) using the β-C–H selective CYP119 variant **CYP119-235** 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 three steps using established chemistry (see supplemental information for details). Gratifyingly, the diazoketone intermediate **8** could be then effectively cyclized by the α-C–H selective variant, **CYP119-168**, 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 highlight utility of the engineered CYP119 library for the synthesis of biologically relevant synthons.

Mechanistic insights into CYP119-catalyzed indoline C–H functionalization

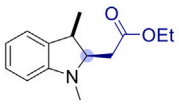
As noted earlier, the C–H functionalization reaction with *N*-methyl-indoline and EDA catalyzed by the early, unoptimized CYP119 catalysts is accompanied by the formation of three distinct by-products, which were determined to correspond to *N*-methyl-indole (**1a-DS**), C_3-functionalized *N*-methyl indole (**4a–DS**), and the *N*-alkylated-indoline **13a** (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 suggests a radical, stepwise mechanism for the present C–H carbene insertion reaction also in line with the results from computational analyses described further below.



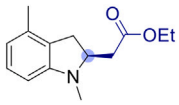
A



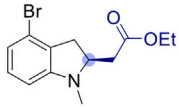
3a, [**CYP119-168**]
92% [83%]
91:9 e.r.
92:8 **3a**:**4a**
5,270 TON



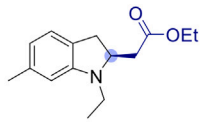
3b, [**CYP119-235**]
42%
98:2 e.r.
99:1 d.r.
2,900 TON



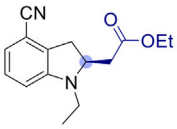
3c, [**CYP119-250**]
85%
83:17 e.r.
6,200 TON



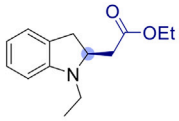
3d, [**CYP119-090**]
94%
97:3 e.r.
6,310 TON



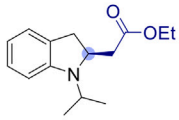
3f, [**CYP119-156**]
93%
84:16 e.r.
5,740 TON



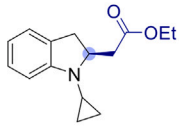
3g, [**CYP119-137**]
89%
93:7 e.r.
5,980 TON



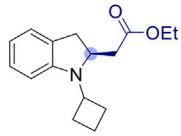
3h, [**CYP119-137**]
97%
84:16 e.r.
6,520 TON



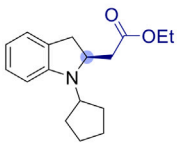
3i, [**CYP119-235**]
84%[a]
96:4 e.r.
5,790 TON



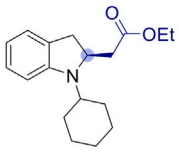
3k, [**CYP119-311**]
77%[a]
90:10 e.r.
6,540 TON



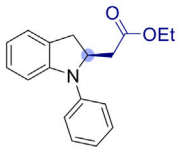
3l, [**CYP119-282**]
67%[a]
66:34 e.r.
4,390 TON



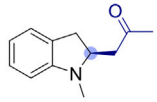
3m, [**CYP119-282**]
52%[a]
60:40 e.r.
3,400 TON



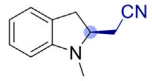
3n, [**CYP119-235**]
44%[a]
87:13 e.r.
3,030 TON



3o, [**CYP119-235**]
Trace[a]



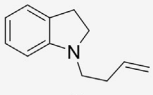
6a, [**CYP119-090**]
97%[b]
98:2 e.r.
6,520 TON



7a, [**CYP119-090**]
99%[b]
96:4 e.r.
6,650 TON

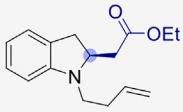
Alternate Carbene Donors

B



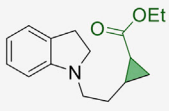
1j

CYP119-235
EDA (**2a**)



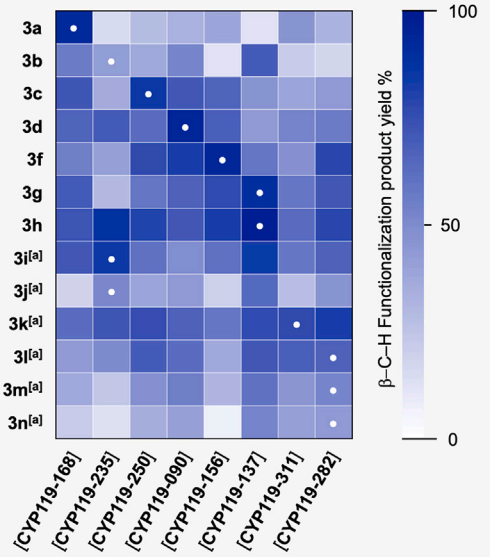
3j, [**CYP119-235**]
51%[a]
80:20 e.r.
3,520 TON

$Rh_2(OAc)_4$
EDA (**2a**)



3j
34%
73:27 d.r.

C



(legend on next page)

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 N–H, with regioselectivity being controlled by the enzyme's active site configuration (*vide infra*) and resulting in the C-centered radical intermediates **IM$_{I}$** or **IM$_{II}$**, respectively. Radical rebound with the iron-carbenoid species yields the corresponding C–H functionalization products (e.g., **3a** or **4a**). For the side reaction leading to the desaturation product **1a-DS**, we envision the radical intermediate **IM$_{I}$** or **IM$_{II}$** may undergo radical polar crossover (RPC), e.g., via single-electron transfer to the heme cofactor, to give the respective carbocation **IM$_{IV}$** or **IM$_{V}$**, which, upon deprotonation, yield *N*-methyl-indole **1a-*DS*** (Figure 5). Alternatively, the latter can be produced from the radical intermediate **IM$_{I}$** (or **IM$_{II}$**) via a second HAA mediated by the Fe(III)-alkyl species (*vide infra*). Regardless of the nature of this step, the ensuing desaturation product *N*-methyl-indole (**1a-*DS***) can be then converted by the enzyme to **4a-DS** via an indole C(3)–H functionalization reaction with EDA akin to that previously reported by our group[44] and others[66,67] 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 **IM$_{III}$** which, upon RPC, is expected to produce the iminium intermediate (**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.[68–70]

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 **CYP119-282** (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 min, with corresponding initial formation rates (TOF) of 28, 5.7, and 2.5 turnovers per minute for **3k**, **1k-*DS***, and **4k-*DS***, respectively (Figure S2). In contrast, no formation of either desaturation byproduct was observed in the absence of the enzyme or EDA within the same conditions (Figure S2), 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 **CYP119-137**. While **4a** was fully preserved in the reaction mixture, **1a-*DS*** was consumed to give rise to the C_3-functionalized product **4a-*DS*** (Figure S8). Although the mechanism of this step was not investigated, previous studies with engineered myoglobin support a stepwise mechanism involving a zwitterionic intermediate.[44,71]

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, a 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** (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 (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.[72] In the absence of the reductant, the **CYP119-168**-catalyzed reaction with **1a** and EDA proceeds with reduced catalytic activity (45% → 32% yield; 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.[73] 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 (Table S6, entry 3–4), indicating that ferrous CYP119 is the catalytic species responsible for formation of the C–H functionalization products.

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 yields are indicated in brackets. TON as determined based on P450 concentration in cell lysate. (a) Using 240 mM EDA (**2a**) at 40°C (b) Using 240 mM diazoacetone (**2b**) or diazoacetonitrile (**2c**).

(B) Chemodivergent reactivity of CYP119 **CYP119-253** and $Rh_2(OAc)_4$ (10 mol %) with 1-(but-3-en-1-yl)indoline (**1j**) and EDA (**2a**).

(C) Heatmap 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 Figure S1.

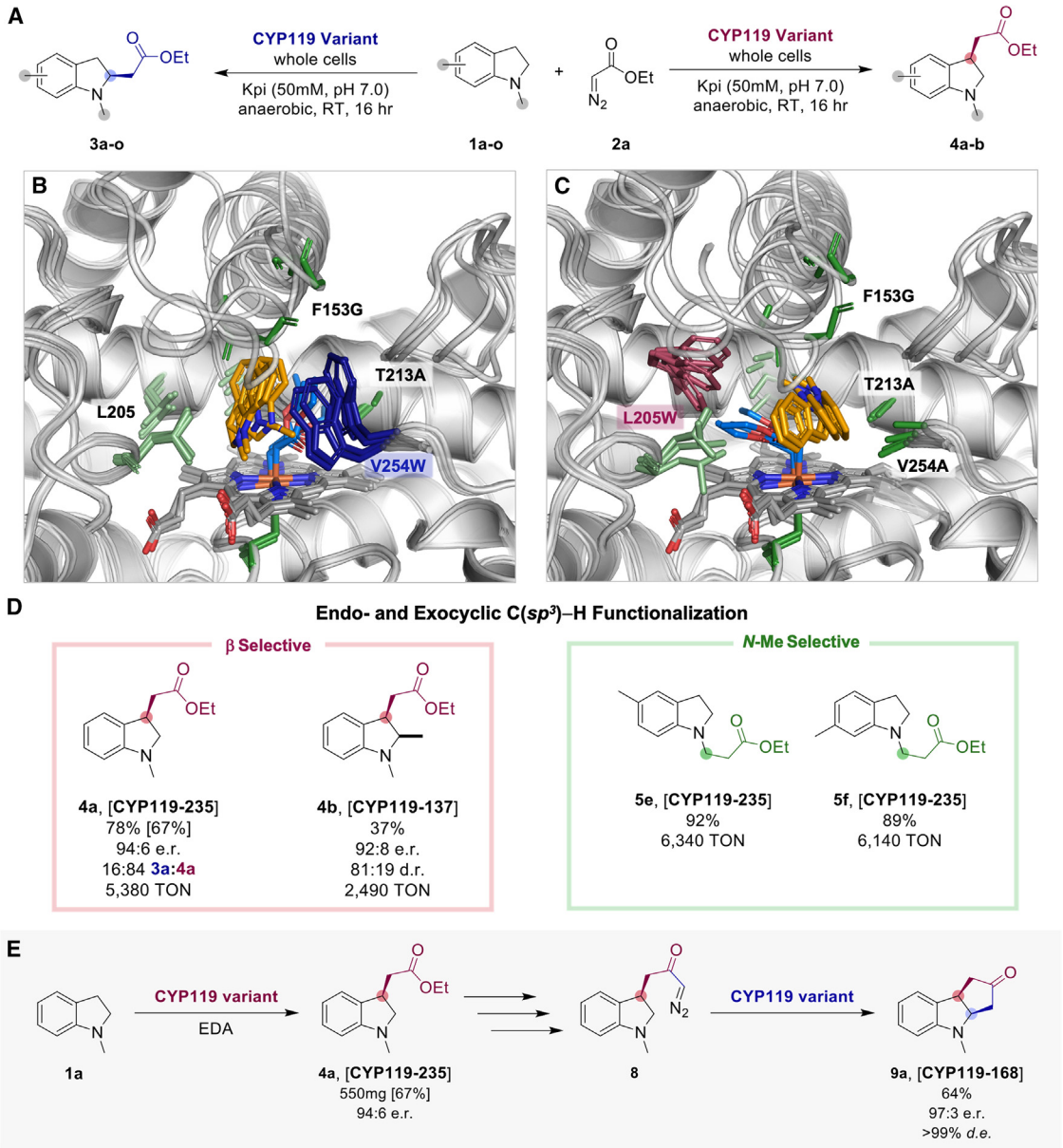


Figure 4. CYP119-catalyzed regiodivergent C–H functionalization of *N*-methyl indoline (1a) with EDA (2a)

(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 and C) Active site views of representative Rosetta models for **CYP119-168** in complex with DFT-calculated α-pro(*S*)-C–H transition state (B) and of **CYP119-235** in complex with DFT-calculated β-pro(*R*)-C–H transition state (C). The heme cofactor is shown as a line model, while mutated active site residues (green), *N*-methyl indoline (orange), and heme-bound carbene (blue) are shown as stick models. See Figure S10 for additional data.

(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**.

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 **CYP119-168**, 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 (1) the H atom abstraction (HAA) step is regioselectivity determining, and (2) 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 is increased

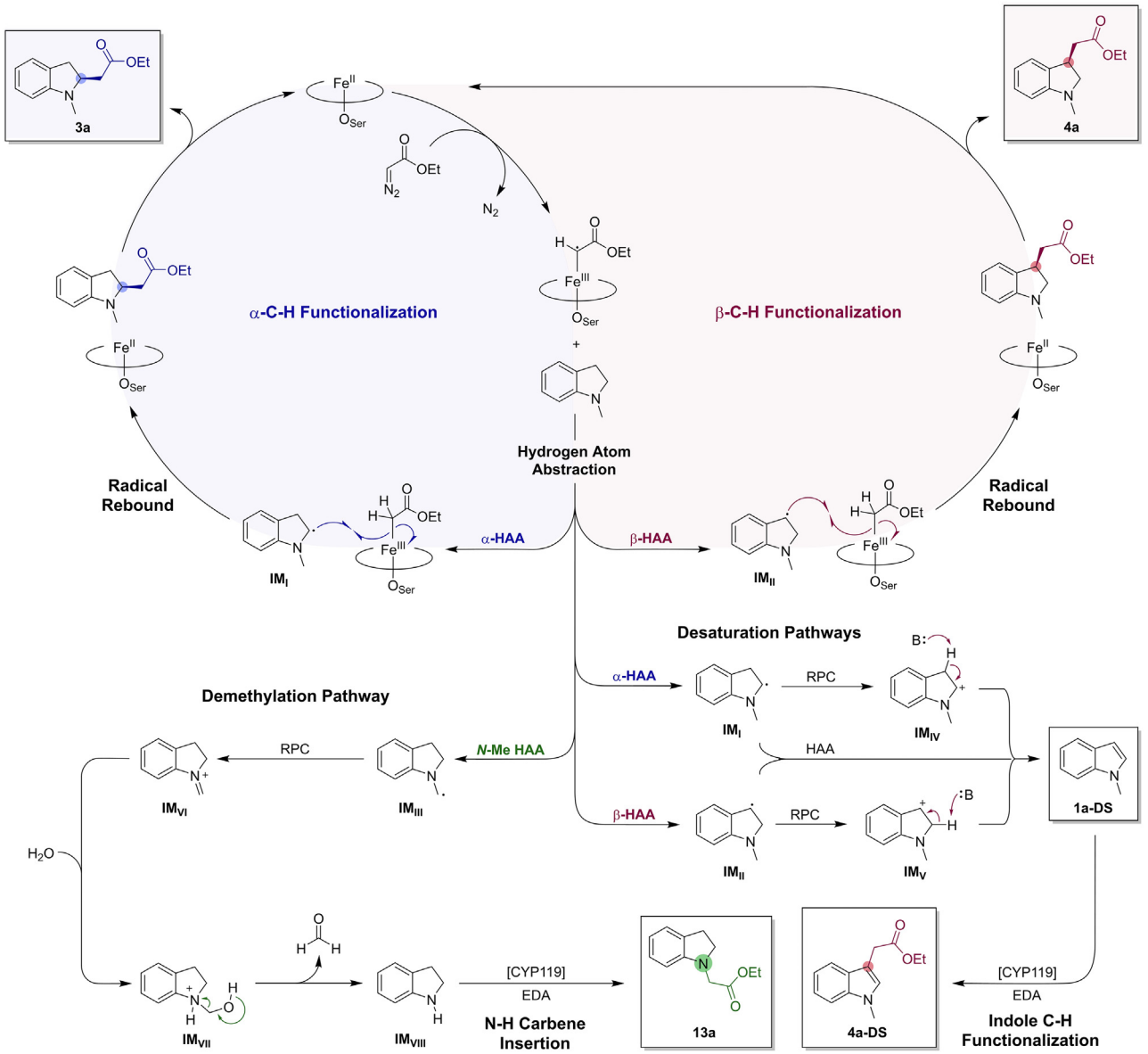


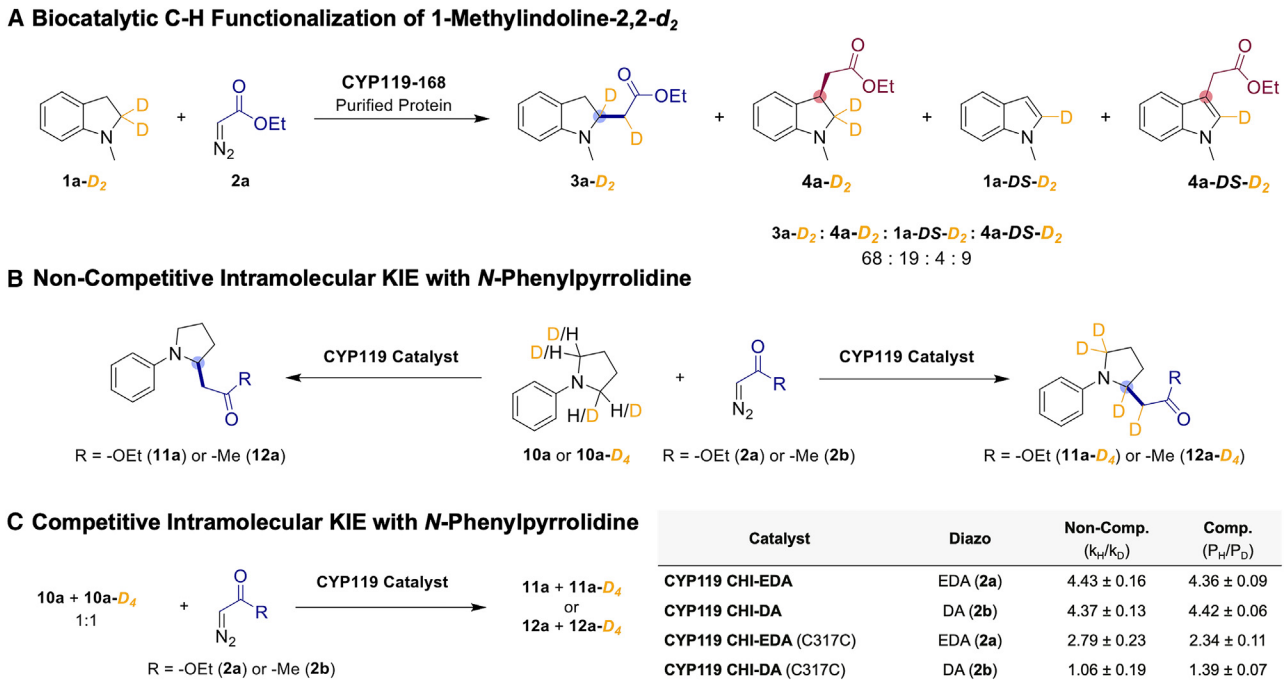
Figure 5. Proposed stepwise radical-mediated mechanism for the C–H functionalization of *N*-methyl indoline (1a) in the presence of engineered CYP119 catalysts

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 **CYP119-168**-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 pathway), possibly due to a slower radical rebound step and/or more favorable conformation for RPC.

It is worth noting that these desaturation reactions are reminiscent of those described for certain cytochromes P450 in non-classical oxidation reactions under aerobic conditions.[74–76] While the cationic mechanism proposed in Figure 5 is plausible based on computational studies of these analogous reactions,[77] an alternative desaturation mechanism entailing a second HAA on the radical intermediate **IM$_I$** or **IM$_{II}$** by the Fe(III)-alkyl species may be also operative, in analogy to desaturation reactions catalyzed by other metalloenzymes.[78] Further studies will thus be required to elucidate this aspect in more detail.

Mechanistic studies

To further illuminate the mechanism of the present P450-catalyzed C–H carbene insertion, we investigated our previously reported C(sp^3)–H functionalization reaction of aryl pyrrolidines



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. Kinetic isotope effect 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 Figures S3 and S4.

with EDA and diazoacetone, as these and the present reaction share the same catalytic system (i.e., serine-ligated CYP119).[50] 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 (Figures 6B and 6C). 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 cysteine-ligated versions of the aforementioned CYP119 variants, i.e., **CHI-EDA** (S317C) and **CHI-DA** (S317C) (Figures 6B and 6C). 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.[79–81]

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,[80,82–87] using [Fe(Por)(MeO$^-$)] to mimic the active site core part of the biocatalyst as done previously,[80,86,87] where Por is a non-substituted porphyrin and MeO$^-$ is the model for the Ser ligand.[88–91] Since conformations and spin states may influence the reaction results,[84–86,92] 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 supplemental information.

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[84] and it features a simultaneous hydride transfer (as shown 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 $\mathbf{R_2}$ to 1.354 Å in $\mathbf{TS_1}$ and then to 3.399 Å in **Int**, while C−H_1 bond length decreases from 1.338 Å in $\mathbf{TS_1}$ to 1.086 Å in **Int** (see Table S20). The most favorable spin state calculated for $\mathbf{TS_1}$ 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 (0.090 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 (Table S22). In the intermediate **Int**, after donating the hydrogen atom, the substrate C′ has the radical with spin densities of 0.995 e, see Table S22). A subsequent radical rebound lead to the formation of the C−C′ bond in $\mathbf{TS_2}$, which has 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 $\mathbf{TS_2}$, 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 ($\mathbf{TS_1}$) 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 Table S11, the overall structures of **TS** and $\mathbf{TS_1}$ are similar except that **TS** has a significantly longer Fe−L (1.987 Å) and shorter Fe−C (1.946 Å) bonds compared to $\mathbf{TS_1}$ (Fe−L, 1.912 Å; Fe−C, 1.982 Å). On the one hand, the shorter Fe−L distance in $\mathbf{TS_1}$ 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 $\mathbf{TS_2}$, which is longer (2.345 Å). Altogether, these structural features contribute to favor the stepwise radical mechanism for this biocatalytic transformation.

In addition, we computationally studied the rate-determining steps of the C–H insertion reaction with EDA and *N*-methyl indoline **1a** as the substrate, using the most favorable spin states for concerted and radical pathways as derived from the analyses above. As shown in Table S19, these two transition states have barriers similar to those calculated above for the reaction with the pyrrolidine substrate **10a** (within ∼1 kcal/mol differences). Therefore, these results also showed preference for the stepwise radical reaction pathway in the presence of **1a** as the substrate, featuring a lower barrier by 7.44 kcal/mol compared to the concerted pathway. Overall, these data further support a radical reaction mechanism for the C–H functionalization reaction investigated in this work.

Analysis of regio- and stereocontrol in the regiodivergent biocatalysts

Further computational studies were performed to gain insights into the origin of the divergent regioselectivity (i.e., α-C-H vs. β-C–H) of the engineered **CYP119-168** and **CYP119-235** biocatalysts. To this end, the most energetically favorable binding pose of the transition states (TS) for the regio- and stereoselectivity-determining hydrogen atom abstraction (HAA) step were modeled in the two enzymes. Briefly, transition states corresponding to near-attack conformations (NACs) for the insertion of the heme-bound carbene into the pro-(*S*) and pro-(*R*) C–H bond at the α- and β-C–H site of the *N*-methyl indoline substrate were calculated and optimized by DFT (see supplemental information for further details and results). These analyses showed that the α-regio-isomers TS are about 5–9 kcal mol^{-1} lower in energy than the β-regio-isomers TS, indicating a generally higher reactivity of the α-C–H site vs. β-C–H site and no stereoselectivity, in the absence of the protein environment (Table S23). Models of the **CYP119-168** and **CYP119-235** variants were generated using Rosetta[93] based on available crystal structures of this enzyme in both ligand-free form and in complex with imidazole-based ligands (Figure S9). As done previously with other systems,[94] the DFT-optimized TS regio-, stereo-, and conformational isomers were these superimposed onto the **CYP119-168** and **CYP119-235** structures and the resulting protein-TS complexes were optimized using Rosetta to obtain the corresponding energies (Tables S24 and S25; see supplemental information for Rosetta modeling details).

Inspection of most representative, low-energy models of these complexes revealed that mutations F153G and T213A, which are shared by both variants, both contribute to significantly enlarge the active site cavity to accommodate the bulky TS complex. The location and orientation of the *N*-methyl indoline substrate, however, were found to differ drastically in the two enzymes. In the α-C–H selective variant **CYP119-168**, the indoline substrate occupies a space near residue Leu205, with an orientation that favors carbene attack (from the *re*-face) to the α-pro(*S*)-C–H bond (Figure 4B), consistent with the regio and stereoselectivity of this biocatalyst. In **CYP119-235**, this arrangement is disfavored by mutation L205W, which inserts a bulky indole ring into the cavity occupied by the indoline molecule in **CYP119-168** (Figure S10C). Conversely, in the β-C–H selective variant **CYP119-235,** the *N*-methyl-indoline substrate occupies the cavity between Ala213 and Ala254 (Figure 4C), which is greatly enlarged by mutations T213A and V254A, respectively, compared to the wild-type enzyme. In this complex, the *N*-methyl-indoline ring is oriented so that carbene attack (from the *si*-face) is directed toward the

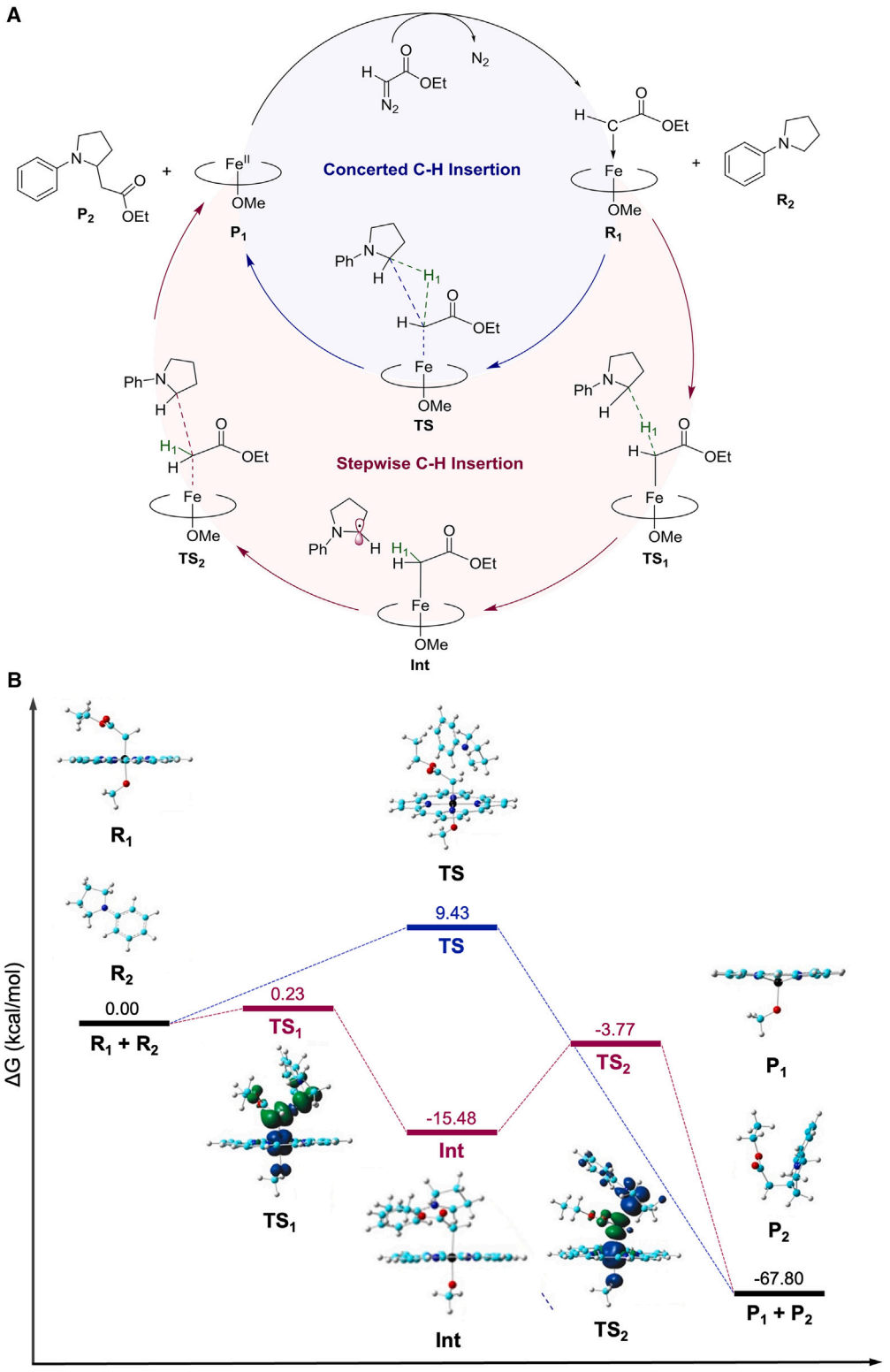


Figure 7. Computational analysis of concerted vs. radical-mediated pathways

(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; Fe, black (spin-density contour value: ±0.004 a.u.).

β-pro(*R*)-C–H bond of the substrate (Figure 4C), in line with the selectivity of this variant. In **CYP119-168**, this arrangement is disfavored by the presence of a bulky Trp residue (vs. Ala in **CYP119-235**) in position 254 (Figure S10B).

Overall, these analyses reveal how mutation of Thr213 and of the "gating" residue Phe153 (Figure S9) both contribute to make CYP119 active site more accessible to carbene/indoline complex, thus enhancing its non-native carbene transferase reactivity. On the other hand, modulation of steric encumbrance at the 254 and 205 positions, which are located at opposing sides of the heme pocket, plays a critical role toward dictating the α- vs. β-C–H selectivity in the C–H alkylation reaction catalyzed by the two regiodivergent biocatalysts. These findings also provide a rationale for explaining the experimentally observed beneficial effect of increasing steric bulk in position 254 toward favoring α-C–H regioselectivity and (*S*)-stereoselectivity (Trp > Tyr ≈ Phe ≫ Ile/Leu/Val; Table 1).

Conclusions

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 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,[32,33,35,37] regiodivergent C(sp^3)–H functionalization via carbene transfer has represented a major challenge[45–47,50–52] and so far has been 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.[53] 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.[3,54–56] Regiodivergent selectivity is also of particular interest in the construction of compound libraries for drug discovery campaigns.[95,96]

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[97] and non-native C–H amination reactions via nitrene transfer catalyzed by engineered P450s and other hemoproteins.[98–100] Our molecular modeling studies clarify the role of key active site mutations in controlling the α- vs. β-C–H selectivity and the enantiopreference of the two regiodivergent biocatalysts developed for this transformation. Overall, this work expands the methodological toolbox for realizing selective C(sp^3)–H functionalization via enzyme-mediated carbene transfer. This advance, along with the insights gained into the productive and non-productive pathways in these reactions, and molecular basis of protein-controlled regioselectivity, are expected to guide future development of new and improved biocatalysts for this important class of transformations.

EXPERIMENTAL PROCEDURES

Details regarding the experimental procedures can be found in the supplemental experimental procedures.

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 described in the supplemental information.

Data and code availability

There is no dataset or code associated with this publication. All supporting data and findings are available in the manuscript or supplemental information.

ACKNOWLEDGMENTS

This work was supported by the US National Institutes of Health grant GM098628 (R.F.). R.F. acknowledges chair endowment support from the Robert A. Welch Foundation (chair, AT-0051-20221212). S.D.K. acknowledges financial support from the NSF grant CHE-2054897. 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 (CPRIT).

AUTHOR CONTRIBUTIONS

R.F. and B.M.C. conceived the project. B.M.C. conducted the bulk of the experiments and experimental mechanistic studies, under guidance and supervision of R.F. R.C. contributed to reaction optimization and product isolation, under guidance and supervision of R.F. J.-M.C. performed DFT analyses of the mechanism under guidance and supervision of Y.Z. Z.S. performed computational analysis of the enzyme regio/stereocontrol under guidance and supervision of S.D.K. B.M.C. and R.F. wrote the manuscript with input and additions from all the other co-authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.checat.2024.101133.

Received: May 7, 2024
Revised: August 12, 2024
Accepted: September 9, 2024
Published: November 21, 2024

REFERENCES

1. Taylor, R.D., MacCoss, M., and Lawson, A.D.G. (2014). Rings in drugs. J. Med. Chem. *57*, 5845–5859. https://doi.org/10.1021/jm4017625.

2. Song, J., Chen, D.F., and Gong, L.Z. (2017). Recent progress in organocatalytic asymmetric total syntheses of complex indole alkaloids. Natl. Sci. Rev. *4*, 381–396. https://doi.org/10.1093/nsr/nwx028.

3. Silva, T.S., Rodrigues, M.T., Santos, H., Zeoly, L.A., Almeida, W.P., Barcelos, R.C., Gomes, R.C., Fernandes, F.S., and Coelho, F. (2019). Recent advances in indoline synthesis. Tetrahedron *75*, 2063–2097. https://doi.org/10.1016/j.tet.2019.02.006.

4. Wei, H., Li, B., Wang, N., Ma, Y., Yu, J., Wang, X., Su, J., and Liu, D. (2023). Development and Application of Indolines in Pharmaceuticals. ChemistryOpen *12*, e202200235. https://doi.org/10.1002/open.202200235.

5. Chandrasekhar, S., Basu, D., and Reddy, C.R. (2007). Palladium-catalyzed reduction of -(-butoxycarbonyl)indoles by polymethylhydrosiloxane. Synthesis-Stuttgart *2007*, 1509–1512. https://doi.org/10.1055/s-2007-966029.

6. Wang, L., Shao, Y., and Liu, Y. (2012). Nucleophilic addition of Grignard reagents to 3-acylindoles: stereoselective synthesis of highly substituted indoline scaffolds. Org. Lett. *14*, 3978–3981. https://doi.org/10.1021/ol301750b.

7. Nieto, M.J., and Lupton, H.K. (2021). Indole and Indoline Scaffolds in Antimicrobials: Overview, Synthesis and Recent Advances in Antimicrobial Research. Curr. Med. Chem. *28*, 4828–4844. https://doi.org/10.2174/0929867327666201102114923.

8. Boger, D.L., and Coleman, R.S. (1984). 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. J. Org. Chem. *49*, 2240–2245. https://doi.org/10.1021/jo00186a032.

9. Presset, M., Pignon, A., Paul, J., Le Gall, E., Léonel, E., and Martens, T. (2017). Synthesis of Indolines by a Zn-Mediated Mannich Reaction/Pd-Catalyzed Amination Sequence. J. Org. Chem. *82*, 3302–3310. https://doi.org/10.1021/acs.joc.7b00013.

10. Li, Y., Chang, Y., Li, Y., Cao, C., Yang, J., Wang, B., and Liang, D. (2018). Iron-Catalyzed exo-Selective Synthesis of Cyanoalkyl Indolines via Cyanoisopropylarylation of Unactivated Alkenes. Adv. Synth. Catal. *360*, 2488–2492. https://doi.org/10.1002/adsc.201800296.

11. Liang, D., Ge, D., Lv, Y., Huang, W., Wang, B., and Li, W. (2018). Silver-Catalyzed Radical Arylphosphorylation of Unactivated Alkenes: Synthesis of 3-Phosphonoalkyl Indolines. J. Org. Chem. *83*, 4681–4691. https://doi.org/10.1021/acs.joc.8b00450.

12. Guillemard, L., Kaplaneris, N., Ackermann, L., and Johansson, M.J. (2021). Late-stage C-H functionalization offers new opportunities in drug discovery. Nat. Rev. Chem *5*, 522–545. https://doi.org/10.1038/s41570-021-00300-6.

13. Fraley, A.E., and Sherman, D.H. (2018). Halogenase engineering and its utility in medicinal chemistry. Bioorg. Med. Chem. Lett. *28*, 1992–1999. https://doi.org/10.1016/j.bmcl.2018.04.066.

14. Latham, J., Brandenburger, E., Shepherd, S.A., Menon, B.R.K., and Micklefield, J. (2018). Development of Halogenase Enzymes for Use in Synthesis. Chem. Rev. *118*, 232–269. https://doi.org/10.1021/acs.chemrev.7b00032.

15. Zhang, R.K., Huang, X., and Arnold, F.H. (2019). Selective CH bond functionalization with engineered heme proteins: new tools to generate complexity. Curr. Opin. Chem. Biol. *49*, 67–75. https://doi.org/10.1016/j.cbpa.2018.10.004.

16. Zwick, C.R., and Renata, H. (2020). Harnessing the biocatalytic potential of iron- and alpha-ketoglutarate-dependent dioxygenases in natural product total synthesis. Nat. Prod. Rep. *37*, 1065–1079. https://doi.org/10.1039/c9np00075e.

17. Ren, X., and Fasan, R. (2021). Engineered and Artificial Metalloenzymes for Selective C-H Functionalization. Curr. Opin. Green Sustain. Chem. *31*, 100494. https://doi.org/10.1016/j.cogsc.2021.100494.

18. Wang, W., Taber, D.F., and Renata, H. (2021). Practical Enzymatic Production of Carbocycles. Chemistry *27*, 11773–11794. https://doi.org/10.1002/chem.202101232.

19. Pyser, J.B., Chakrabarty, S., Romero, E.O., and Narayan, A.R.H. (2021). State-of-the-Art Biocatalysis. ACS Cent. Sci. *7*, 1105–1116. https://doi.org/10.1021/acscentsci.1c00273.

20. Couture, B., Chattopadhyay, A., and Fasan, R. (2024). Biocatalytic Carbene and Nitrene Transfer Reactions. In Comprehensive Chirality (Elsevier), pp. 111–142.

21. Reetz, M.T., Qu, G., and Sun, Z. (2024). Engineered enzymes for the synthesis of pharmaceuticals and other high-value products. Nat. Synth. *3*, 19–32. https://doi.org/10.1038/s44160-023-00417-0.

22. Hayashi, T., Ligibel, M., Sager, E., Voss, M., Hunziker, J., Schroer, K., Snajdrova, R., and Buller, R. (2019). Evolved Aliphatic Halogenases Enable Regiocomplementary C-H Functionalization of a Pharmaceutically Relevant Compound. Angew. Chem. Int. Ed. Engl. *58*, 18535–18539. https://doi.org/10.1002/anie.201907245.

23. Renata, H., Shimizu, E., and Zwick, C.R. (2021). Regiodivergent Biocatalytic Hydroxylation of L-Glutamine Facilitated by Characterization of Non-Heme Dioxygenases from Non-Ribosomal Peptide Biosyntheses. Tetrahedron *90*, 132190. https://doi.org/10.1016/j.tet.2021.132190.

24. Craven, E.J., Latham, J., Shepherd, S.A., Khan, I., Diaz-Rodriguez, A., Greaney, M.F., and Micklefield, J. (2021). Programmable late-stage C-H bond functionalization enabled by integration of enzymes with chemocatalysis. Nat. Catal. *4*, 385–394. https://doi.org/10.1038/s41929-021-00603-3.

25. 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., et al. (2021). Engineering P450 TamI as an Iterative Biocatalyst for Selective Late-Stage C-H Functionalization and Epoxidation of Tirandamycin Antibiotics. ACS Catal. *11*, 8304–8316. https://doi.org/10.1021/acscatal.1c01460.

26. Münch, J., Püllmann, P., Zhang, W., and Weissenborn, M.J. (2021). Enzymatic Hydroxylations of sp3-Carbons. ACS Catal. *11*, 9168–9203. https://doi.org/10.1021/acscatal.1c00759.

27. Knorrscheidt, A., Soler, J., Hünecke, N., Püllmann, P., Garcia-Borràs, M., and Weissenborn, M.J. (2021). Accessing Chemo- and Regioselective Benzylic and Aromatic Oxidations by Protein Engineering of an Unspecific Peroxygenase. ACS Catal. *11*, 7327–7338. https://doi.org/10.1021/acscatal.1c00847.

28. Zetzsche, L.E., Chakrabarty, S., and Narayan, A.R.H. (2022). Development of a P450 Fusion Enzyme for Biaryl Coupling in Yeast. ACS Chem. Biol. *17*, 2986–2992. https://doi.org/10.1021/acschembio.2c00690.

29. Monterrey, D.T., Menés-Rubio, A., Keser, M., Gonzalez-Perez, D., and Alcalde, M. (2023). Unspecific peroxygenases: The pot of gold at the end of the oxyfunctionalization rainbow? Curr. Opin. Green Sustainable Chem. *41*, 100786. https://doi.org/10.1016/j.cogsc.2023.100786.

30. Andorfer, M.C., Park, H.J., Vergara-Coll, J., and Lewis, J.C. (2016). Directed Evolution of RebH for Catalyst-Controlled Halogenation of Indole C-H Bonds. Chem. Sci. *7*, 3720–3729. https://doi.org/10.1039/C5SC04680G.

31. Shepherd, S.A., Karthikeyan, C., Latham, J., Struck, A.W., Thompson, M.L., Menon, B.R.K., Styles, M.Q., Levy, C., Leys, D., and Micklefield, J. (2015). Extending the biocatalytic scope of regiocomplementary flavin-dependent halogenase enzymes. Chem. Sci. *6*, 3454–3460. https://doi.org/10.1039/c5sc00913h.

32. Kille, S., Zilly, F.E., Acevedo, J.P., and Reetz, M.T. (2011). Regio- and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution. Nat. Chem. *3*, 738–743. https://doi.org/10.1038/nchem.1113.

33. Zhang, K., Shafer, B.M., Demars, M.D., second, Stern, H.A., and Fasan, R. (2012). Controlled oxidation of remote sp3 C-H bonds in artemisinin

via P450 catalysts with fine-tuned regio- and stereoselectivity. J. Am. Chem. Soc. *134*, 18695–18704. https://doi.org/10.1021/ja3073462.

34. Greule, A., Stok, J.E., De Voss, J.J., and Cryle, M.J. (2018). Unrivalled diversity: the many roles and reactions of bacterial cytochromes P450 in secondary metabolism. Nat. Prod. Rep. *35*, 757–791. https://doi.org/10.1039/c7np00063d.

35. Alwaseem, H., Frisch, B.J., and Fasan, R. (2018). Anticancer activity profiling of parthenolide analogs generated via P450-mediated chemoenzymatic synthesis. Bioorg. Med. Chem. *26*, 1365–1373. https://doi.org/10.1016/j.bmc.2017.08.009.

36. Lukowski, A.L., Ellinwood, D.C., Hinze, M.E., DeLuca, R.J., Du Bois, J., Hall, S., and Narayan, A.R.H. (2018). C-H Hydroxylation in Paralytic Shellfish Toxin Biosynthesis. J. Am. Chem. Soc. *140*, 11863–11869. https://doi.org/10.1021/jacs.8b08901.

37. Zhang, X., King-Smith, E., Dong, L.B., Yang, L.C., Rudolf, J.D., Shen, B., and Renata, H. (2020). Divergent synthesis of complex diterpenes through a hybrid oxidative approach. Science *369*, 799–806. https://doi.org/10.1126/science.abb8271.

38. Alwaseem, H., Giovani, S., Crotti, M., Welle, K., Jordan, C.T., Ghaemmaghami, S., and Fasan, R. (2021). Comprehensive Structure-Activity Profiling of Micheliolide and its Targeted Proteome in Leukemia Cells via Probe-Guided Late-Stage C-H Functionalization. ACS Cent. Sci. *7*, 841–857. https://doi.org/10.1021/acscentsci.0c01624.

39. Iizaka, Y., Arai, R., Takahashi, A., Ito, M., Sakai, M., Fukumoto, A., Sherman, D.H., and Anzai, Y. (2022). Engineering sequence and selectivity of late-stage C-H oxidation in the MycG iterative cytochrome P450. J. Ind. Microbiol. Biotechnol. *49*, kuab069. https://doi.org/10.1093/jimb/kuab069.

40. Turner, N.J. (2009). Directed evolution drives the next generation of biocatalysts. Nat. Chem. Biol. *5*, 567–573. https://doi.org/10.1038/nchembio.203.

41. Reetz, M.T. (2011). Laboratory evolution of stereoselective enzymes: a prolific source of catalysts for asymmetric reactions. Angew. Chem. Int. Ed. Engl. *50*, 138–174. https://doi.org/10.1002/anie.201000826.

42. Dydio, P., Key, H.M., Nazarenko, A., Rha, J.Y.E., Seyedkazemi, V., Clark, D.S., and Hartwig, J.F. (2016). An artificial metalloenzyme with the kinetics of native enzymes. Science *354*, 102–106. https://doi.org/10.1126/science.aah4427.

43. Brandenberg, O.F., Fasan, R., and Arnold, F.H. (2017). Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions. Curr. Opin. Biotechnol. *47*, 102–111. https://doi.org/10.1016/j.copbio.2017.06.005.

44. Vargas, D.A., Tinoco, A., Tyagi, V., and Fasan, R. (2018). Myoglobin-Catalyzed C-H Functionalization of Unprotected Indoles. Angew. Chem. Int. Ed. Engl. *57*, 9911–9915. https://doi.org/10.1002/anie.201804779.

45. Gu, Y., Natoli, S.N., Liu, Z., Clark, D.S., and Hartwig, J.F. (2019). Site-Selective Functionalization of (sp(3))C-H Bonds Catalyzed by Artificial Metalloenzymes Containing an Iridium-Porphyrin Cofactor. Angew. Chem. Int. Ed. Engl. *58*, 13954–13960. https://doi.org/10.1002/anie.201907460.

46. Zhang, J., Huang, X., Zhang, R.K., and Arnold, F.H. (2019). Enantiodivergent alpha-Amino C-H Fluoroalkylation Catalyzed by Engineered Cytochrome P450s. J. Am. Chem. Soc. *141*, 9798–9802. https://doi.org/10.1021/jacs.9b04344.

47. Zhang, R.K., Chen, K., Huang, X., Wohlschlager, L., Renata, H., and Arnold, F.H. (2019). Enzymatic assembly of carbon-carbon bonds via iron-catalysed sp(3) C-H functionalization. Nature *565*, 67–72. https://doi.org/10.1038/s41586-018-0808-5.

48. Rumo, C., Stein, A., Klehr, J., Tachibana, R., Prescimone, A., Häussinger, D., and Ward, T.R. (2022). An Artificial Metalloenzyme Based on a Copper Heteroscorpionate Enables sp(3) C-H Functionalization via Intramolecular Carbene Insertion. J. Am. Chem. Soc. *144*, 11676–11684. https://doi.org/10.1021/jacs.2c03311.

49. Zhang, J., Maggiolo, A.O., Alfonzo, E., Mao, R., Porter, N.J., Abney, N., and Arnold, F.H. (2023). Chemodivergent C(sp(3))-H and C(sp(2))-H Cyanomethylation Using Engineered Carbene Transferases. Nat. Catal. *6*, 152–160. https://doi.org/10.1038/s41929-022-00908-x.

50. Ren, X., Couture, B.M., Liu, N., Lall, M.S., Kohrt, J.T., and Fasan, R. (2023). Enantioselective Single and Dual alpha-C-H Bond Functionalization of Cyclic Amines via Enzymatic Carbene Transfer. J. Am. Chem. Soc. *145*, 537–550. https://doi.org/10.1021/jacs.2c10775.

51. Sreenilayam, G., Moore, E.J., Steck, V., and Fasan, R. (2017). Metal Substitution Modulates the Reactivity and Extends the Reaction Scope of Myoglobin Carbene Transfer Catalysts. Adv. Synth. Catal. *359*, 2076–2089. https://doi.org/10.1002/adsc.201700202.

52. Zhou, A.Z., Chen, K., and Arnold, F.H. (2020). Enzymatic Lactone-Carbene C-H Insertion to Build Contiguous Chiral Centers. ACS Catal. *10*, 5393–5398. https://doi.org/10.1021/acscatal.0c01349.

53. Chhabra, N., Aseri, M.L., and Padmanabhan, D. (2013). A review of drug isomerism and its significance. Int. J. Appl. Basic Med. Res. *3*, 16–18. https://doi.org/10.4103/2229-516X.112233.

54. Ruiz-Sanchis, P., Savina, S.A., Albericio, F., and Álvarez, M. (2011). Structure, bioactivity and synthesis of natural products with hexahydropyrrolo[2,3-b]indole. Chemistry *17*, 1388–1408. https://doi.org/10.1002/chem.201001451.

55. Zi, W., Zuo, Z., and Ma, D. (2015). Intramolecular dearomative oxidative coupling of indoles: a unified strategy for the total synthesis of indoline alkaloids. Acc. Chem. Res. *48*, 702–711. https://doi.org/10.1021/ar5004303.

56. Wang, X., Griffiths, B., and Burl, J. (2016). Bioinspired Discovery of Chemical Reactions and Biological Probes. Synlett *27*, 2039–2042. https://doi.org/10.1055/s-0035-1561638.

57. Liu, F., and Su, M. (2023). Indole and indoline scaffolds in drug discovery. In Privileged Scaffolds in Drug Discovery, B. Yu, N. Li, and C. Fu, eds. (Academic Press), pp. 147–161.

58. Coelho, P.S., Wang, Z.J., Ener, M.E., Baril, S.A., Kannan, A., Arnold, F.H., and Brustad, E.M. (2013). A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins in vivo. Nat. Chem. Biol. *9*, 485–487. https://doi.org/10.1038/nchembio.1278.

59. Doyle, M.P., Duffy, R., Ratnikov, M., and Zhou, L. (2010). Catalytic carbene insertion into C-H bonds. Chem. Rev. *110*, 704–724. https://doi.org/10.1021/cr900239n.

60. Weaver, B.R., Perkins, L.J., Fernandez Candelaria, F.O., Burstyn, J.N., and Buller, A.R. (2023). Molecular Determinants of Efficient Cobalt-Substituted Hemoprotein Production in E. coli. ACS Synth. Biol. *12*, 3669–3679. https://doi.org/10.1021/acssynbio.3c00481.

61. Key, H.M., Dydio, P., Liu, Z., Rha, J.Y.E., Nazarenko, A., Seyedkazemi, V., Clark, D.S., and Hartwig, J.F. (2017). Beyond Iron: Iridium-Containing P450 Enzymes for Selective Cyclopropanations of Structurally Diverse Alkenes. ACS Cent. Sci. *3*, 302–308. https://doi.org/10.1021/acscentsci.6b00391.

62. Knight, A.M., Kan, S.B.J., Lewis, R.D., Brandenberg, O.F., Chen, K., and Arnold, F.H. (2018). Diverse Engineered Heme Proteins Enable Stereodivergent Cyclopropanation of Unactivated Alkenes. ACS Cent. Sci. *4*, 372–377. https://doi.org/10.1021/acscentsci.7b00548.

63. Park, S.Y., Yamane, K., Adachi, S.i., Shiro, Y., Weiss, K.E., Maves, S.A., and Sligar, S.G. (2002). Thermophilic cytochrome P450 (CYP119) from Sulfolobus solfataricus: high resolution structure and functional properties. J. Inorg. Biochem. *91*, 491–501. https://doi.org/10.1016/s0162-0134(02)00446-4.

64. Garlets, Z.J., Wertz, B.D., Liu, W., Voight, E.A., and Davies, H.M.L. (2020). Regio- and Stereoselective Rhodium(II)-Catalyzed C-H Functionalization of Cyclobutanes. Chem *6*, 304–313. https://doi.org/10.1016/j.chempr.2019.12.014.

65. Boni, Y.T., Cammarota, R.C., Liao, K., Sigman, M.S., and Davies, H.M.L. (2022). Leveraging Regio- and Stereoselective C(sp(3))-H Functionalization

of Silyl Ethers to Train a Logistic Regression Classification Model for Predicting Site-Selectivity Bias. J. Am. Chem. Soc. *144*, 15549–15561. https://doi.org/10.1021/jacs.2c04383.

66. Hock, K.J., Knorrscheidt, A., Hommelsheim, R., Ho, J., Weissenborn, M.J., and Koenigs, R.M. (2019). Tryptamine Synthesis by Iron Porphyrin Catalyzed C-H Functionalization of Indoles with Diazoacetonitrile. Angew. Chem. Int. Ed. Engl. *58*, 3630–3634. https://doi.org/10.1002/anie.201813631.
67. Brandenberg, O.F., Chen, K., and Arnold, F.H. (2019). Directed Evolution of a Cytochrome P450 Carbene Transferase for Selective Functionalization of Cyclic Compounds. J. Am. Chem. Soc. *141*, 8989–8995. https://doi.org/10.1021/jacs.9b02931.
68. Wang, Z.J., Peck, N.E., Renata, H., and Arnold, F.H. (2014). Cytochrome P450-Catalyzed Insertion of Carbenoids into N-H Bonds. Chem. Sci. *5*, 598–601. https://doi.org/10.1039/C3SC52535J.
69. Sreenilayam, G., and Fasan, R. (2015). Myoglobin-catalyzed intermolecular carbene N-H insertion with arylamine substrates. Chem. Commun. *51*, 1532–1534. https://doi.org/10.1039/c4cc08753d.
70. Steck, V., Sreenilayam, G., and Fasan, R. (2020). Selective Functionalization of Aliphatic Amines via Myoglobin-catalyzed Carbene N-H Insertion. Synlett *31*, 224–229. https://doi.org/10.1055/s-0039-1690007.
71. Vargas, D.A., Khade, R.L., Zhang, Y., and Fasan, R. (2019). Biocatalytic Strategy for Highly Diastereo- and Enantioselective Synthesis of 2,3-Dihydrobenzofuran-Based Tricyclic Scaffolds. Angew. Chem. Int. Ed. Engl. *58*, 10148–10152. https://doi.org/10.1002/anie.201903455.
72. Lee, W.C.C., Wang, D.S., Zhu, Y., and Zhang, X.P. (2023). Iron(III)-based metalloradical catalysis for asymmetric cyclopropanation via a stepwise radical mechanism. Nat. Chem. *15*, 1569–1580. https://doi.org/10.1038/s41557-023-01317-8.
73. Moore, E.J., and Fasan, R. (2019). Effect of proximal ligand substitutions on the carbene and nitrene transferase activity of myoglobin. Tetrahedron *75*, 2357–2363. https://doi.org/10.1016/j.tet.2019.03.009.
74. Rettie, A.E., Boberg, M., Rettenmeier, A.W., and Baillie, T.A. (1988). Cytochrome P-450-catalyzed desaturation of valproic acid in vitro. Species differences, induction effects, and mechanistic studies. J. Biol. Chem. *263*, 13733–13738. https://doi.org/10.1016/S0021-9258(18)68302-4.
75. Guengerich, F.P. (2018). Mechanisms of Cytochrome P450-Catalyzed Oxidations. ACS Catal. *8*, 10964–10976. https://doi.org/10.1021/acscatal.8b03401.
76. Ren, X., O'Hanlon, J.A., Morris, M., Robertson, J., and Wong, L.L. (2016). Synthesis of Imidazolidin-4-ones via a Cytochrome P450-Catalyzed Intramolecular C-H Amination. ACS Catal. *6*, 6833–6837. https://doi.org/10.1021/acscatal.6b02189.
77. Kumar, D., De Visser, S.P., and Shaik, S. (2004). Oxygen economy of cytochrome P450: what is the origin of the mixed functionality as a dehydrogenase-oxidase enzyme compared with its normal function? J. Am. Chem. Soc. *126*, 5072–5073. https://doi.org/10.1021/ja0318737.
78. Cooper, H.L.R., Mishra, G., Huang, X., Pender-Cudlip, M., Austin, R.N., Shanklin, J., and Groves, J.T. (2012). Parallel and competitive pathways for substrate desaturation, hydroxylation, and radical rearrangement by the non-heme diiron hydroxylase AlkB. J. Am. Chem. Soc. *134*, 20365–20375. https://doi.org/10.1021/ja3059149.
79. Chen, K., Zhang, S.Q., Brandenberg, O.F., Hong, X., and Arnold, F.H. (2018). Alternate Heme Ligation Steers Activity and Selectivity in Engineered Cytochrome P450-Catalyzed Carbene-Transfer Reactions. J. Am. Chem. Soc. *140*, 16402–16407. https://doi.org/10.1021/jacs.8b09613.
80. Tinoco, A., Wei, Y., Bacik, J.P., Carminati, D.M., Moore, E.J., Ando, N., Zhang, Y., and Fasan, R. (2019). Origin of high stereocontrol in olefin cyclopropanation catalyzed by an engineered carbene transferase. ACS Catal. *9*, 1514–1524. https://doi.org/10.1021/acscatal.8b04073.
81. Carminati, D.M., and Fasan, R. (2019). Stereoselective Cyclopropanation of Electron-Deficient Olefins with a Cofactor Redesigned Carbene Transferase Featuring Radical Reactivity. ACS Catal. *9*, 9683–9697. https://doi.org/10.1021/acscatal.9b02272.
82. Khade, R.L., Fan, W., Ling, Y., Yang, L., Oldfield, E., and Zhang, Y. (2014). Iron Porphyrin Carbenes as Catalytic Intermediates: Structures, Mossbauer and NMR Spectroscopic Properties, and Bonding. Angew. Chem. Int. Ed. *53*, 7574–7578. https://doi.org/10.1002/anie.201402472.
83. Khade, R.L., and Zhang, Y. (2015). Catalytic and Biocatalytic Iron Porphyrin Carbene Formation: Effects of Binding Mode, Carbene Substituent, Porphyrin Substituent, and Protein Axial Ligand. J. Am. Chem. Soc. *137*, 7560–7563. https://doi.org/10.1021/jacs.5b03437.
84. Khade, R.L., and Zhang, Y. (2017). C-H Insertions by Iron Porphyrin Carbene: Basic Mechanism and Origin of Substrate Selectivity. Chemistry - A Euopean Journal *23*, 17654–17658. https://doi.org/10.1002/chem.201704631.
85. Wei, Y., Tinoco, A., Steck, V., Fasan, R., and Zhang, Y. (2018). Cyclopropanations via Heme Carbenes: Basic Mechanism and Effects of Carbene Substituent, Protein Axial Ligand, and Porphyrin Substitution. J. Am. Chem. Soc. *140*, 1649–1662. https://doi.org/10.1021/jacs.7b09171.
86. Khade, R.,L., Chandgude, A.L., Fasan, R., and Zhang, Y. (2019). Mechanistic Investigation of Biocatalytic Heme Carbenoid Si-H Insertions. ChemCatChem *11*, 3101–3108. https://doi.org/10.1002/cctc.201801755.
87. Vargas, D.A., Khade, R.L., Zhang, Y., and Fasan, R. (2019). Biocatalytic strategy for highly diastereo- and enantioselective synthesis of 2,3-dihydrobenzofuran based tricyclic scaffolds. Angew. Chem. Int. Ed. *58*, 10148–10152. https://doi.org/10.1002/anie.201903455.
88. Li, Z., Burnell, D.J., and Boyd, R.J. (2017). Computational Study of Engineered Cytochrome P450-Catalyzed C–H Amination: The Origin of the Regio- and Stereoselectivity. J. Phys. Chem. B *121*, 10859–10868. https://doi.org/10.1021/acs.jpcb.7b10256.
89. Wang, J., Gao, H., Yang, L., and Gao, Y.Q. (2020). Role of Engineered Iron-haem Enzyme in Reactivity and Stereoselectivity of Intermolecular Benzylic C–H Bond Amination. ACS Catal. *10*, 5318–5327. https://doi.org/10.1021/acscatal.0c00248.
90. Li, X., Dong, L., and Liu, Y. (2020). Theoretical Study of Iron Porphyrin Nitrene: Formation Mechanism, Electronic Nature, and Intermolecular C–H Amination. Inorg. Chem. *59*, 1622–1632. https://doi.org/10.1021/acs.inorgchem.9b02216.
91. Athavale, S.V., Gao, S., Liu, Z., Mallojjala, S.C., Hirschi, J.S., and Arnold, F.H. (2021). Biocatalytic, Intermolecular C−H Bond Functionalization for the Synthesis of Enantioenriched Amides. Angew. Chem. Int. Ed. Engl. *60*, 24864–24869. https://doi.org/10.1002/anie.202110873.
92. Wei, Y., Conklin, M., and Zhang, Y. (2022). Biocatalytic Intramolecular C−H aminations via Engineered Heme Proteins: Full Reaction Pathways and Axial Ligand Effects. Chem. Eur J. *28*, e202202006. https://doi.org/10.1002/chem.202202006.
93. Leaver-Fay, A., Tyka, M., Lewis, S.M., Lange, O.F., Thompson, J., Jacak, R., Kaufman, K., Renfrew, P.D., Smith, C.A., Sheffler, W., et al. (2011). Chapter nineteen - Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. In Methods Enzymol., Methods in Enzymology, *487*, M.L. Johnson and L. Brand, eds. (Academic Press), pp. 545–574.
94. Nam, D., Tinoco, A., Shen, Z., Adukure, R.D., Sreenilayam, G., Khare, S.D., and Fasan, R. (2022). Enantioselective Synthesis of alpha-Trifluoromethyl Amines via Biocatalytic N-H Bond Insertion with Acceptor-Acceptor Carbene Donors. J. Am. Chem. Soc. *144*, 2590–2602. https://doi.org/10.1021/jacs.1c10750.
95. Gerry, C.J., Wawer, M.J., Clemons, P.A., and Schreiber, S.L. (2019). DNA Barcoding a Complete Matrix of Stereoisomeric Small Molecules. J. Am. Chem. Soc. *141*, 10225–10235. https://doi.org/10.1021/jacs.9b01203.

96. Bassi, G., Favalli, N., Vuk, M., Catalano, M., Martinelli, A., Trenner, A., Porro, A., Yang, S., Tham, C.L., Moroglu, M., et al. (2020). A Single-Stranded DNA-Encoded Chemical Library Based on a Stereoisomeric Scaffold Enables Ligand Discovery by Modular Assembly of Building Blocks. Adv. Sci. *7*, 2001970. https://doi.org/10.1002/advs.202001970.

97. Denisov, I.G., Makris, T.M., Sligar, S.G., and Schlichting, I. (2005). Structure and chemistry of cytochrome P450. Chem. Rev. *105*, 2253–2277. https://doi.org/10.1021/cr0307143.

98. Singh, R., Kolev, J.N., Sutera, P.A., and Fasan, R. (2015). Enzymatic C(sp(3))-H Amination: P450-Catalyzed Conversion of Carbonazidates into Oxazolidinones. ACS Catal. *5*, 1685–1691. https://doi.org/10.1021/Cs5018612.

99. Yang, Y., Cho, I., Qi, X., Liu, P., and Arnold, F.H. (2019). An Enzymatic Platform for the Asymmetric Amination of Primary, Secondary and Tertiary C(sp(3))—H Bonds. Nat. Chem. *11*, 987–993. https://doi.org/10.1038/s41557-019-0343-5.

100. Roy, S., Vargas, D.A., Ma, P., Sengupta, A., Zhu, L., Houk, K.N., and Fasan, R. (2024). Stereoselective Construction of beta-gamma-and delta-Lactam Rings via Enzymatic C-H Amidation. Nature Catal *7*, 65–76. https://doi.org/10.1038/s41929-023-01068-2.