HHS Public Access

Author manuscript

J Am Chem Soc. Author manuscript; available in PMC 2023 January 17.

Published in final edited form as:

J Am Chem Soc. 2023 January 11; 145(1): 537–550. doi:10.1021/jacs.2c10775.

Enantioselective Single and Dual α-C—H Bond Functionalization of Cyclic Amines via Enzymatic Carbene Transfer

Xinkun Ren[1,3], **Bo M. Couture**[1], **Ningyu Liu**[1], **Manjinder S. Lall**[2], **Jeffrey T. Kohrt**[2], **Rudi Fasan**[1,*]

[1]Department of Chemistry, University of Rochester, Rochester, NY 14627, United States.

[2]Pfizer Inc., Medicine and Design, Groton, CT 06340, United States.

[3]Current address: College of Engineering and Applied Sciences, Nanjing University, Nanjing, Jiangsu Province 210023, China.

Abstract

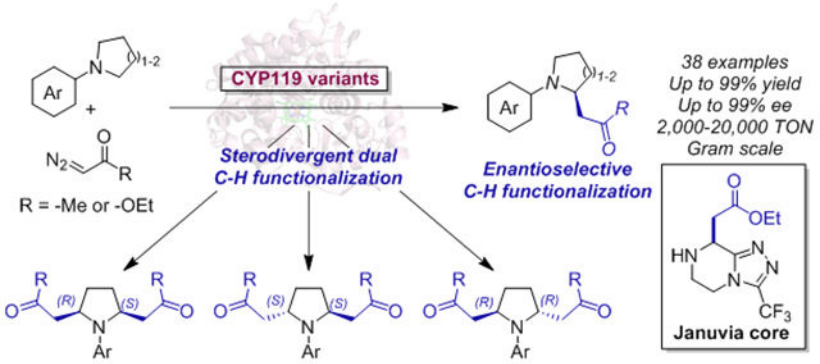
Cyclic amines are ubiquitous structural motifs found in pharmaceuticals and biologically active natural products, making methods for their elaboration via direct C—H functionalization of considerable synthetic value. Herein, we report the development of an iron-based biocatalytic strategy for enantioselective α-C—H functionalization of pyrrolidines and other saturated N-heterocycles via a carbene transfer reaction with diazoacetone. Currently for organometallic catalysts, this transformation can be accomplished in high yields, high catalytic activity and high stereoselectivity (up to 99:1 e.r. and 20,350 TON) using engineered variants of cytochrome P450 CYP119 from *Sulfolobus solfataricus*. This methodology was further extended to enable enantioselective α-C—H functionalization in the presence of ethyl diazoacetate as carbene donor (up to 96:4 e.r. and 8,920 TON), and the two strategies were combined to achieve a one-pot as well as a tandem dual C—H functionalization of the cyclic amine substrate with enzyme-controlled diastereo- and enantiodivergent selectivity. This biocatalytic approach is amenable to gram-scale synthesis and can be applied to drug scaffolds for late-stage C—H functionalization. This work provides an efficient and tunable method for direct asymmetric α-C—H functionalization of saturated *N*-heterocycles which should offer new opportunities for the synthesis, discovery, and optimization of bioactive molecules.

Graphical Abstract

*Corresponding Authors: rfasan@ur.rochester.edu.

Supporting Information

Supporting information includes supplementary Tables and Figures, chiral GC and SFC chromatograms, synthetic procedures, compound characterization data, NMR spectra, crystallographic data.



Introduction

Saturated *N*-containing heterocycles (e.g., pyrrolidine, piperidine, morpholine) are key components of many pharmaceuticals and biologically active natural products (Figure 1).[1] Given their value as 'priviledged' scaffolds in medicinal chemistry, major efforts have been devoted to the development of strategies for the functionalization of these compounds. Among them, C(sp^3)—H functionalization strategies resulting in the formation of new chiral carbon-carbon bonds are of particular interest.[2, 3] Current chemical strategies for α-C(sp^3)—H functionalization of saturated *N*-heterocycles include multistep sequences in which strong bases are used to generate reactive α-amino anions, which can undergo transition metal-mediated alkylation/arylation (Figure 2a).[4-9] Other notable approaches involve the use of oxidative α-C—H functionalization to generate an intermediate iminium ion, which is captured by carbon-based nucleophiles (Figure 2b);[10-14], directing group mediated C(sp^3)—H functionalization[15-17], and photoredox strategies strategies (Figure 2c).[18-22] Despite this progress, these protocols require multistep sequences, rare transition metals, and/or they lack of stereoselectivity.

Enzymes are attractive alternatives for C—H functionalization because of their inherent chemoselectivity, sustainability, and potential to be optimized via protein engineering for tuning activity and stereoselectivity.[23-28] While natural enzymes capable of forging new carbon-carbon bonds via C(sp^3)—H functionalization are rare and largely limited to specific substrates (e.g., methyl transfer reactions catalyzed by *S*-adenosylmethionine dependent enzymes)[29-33], recent advances in protein engineering have expanded the repertoire of enzyme-catalyzed C—H functionalization via carbene transfer chemistry (Figure 2d-e).[23, 34] Traditionally, C—H functionalization via metal–carbenoid insertion has been pursued through small molecule organometallic catalysts, including complexes with rhodium,[35-37] iridium,[38, 39] and other metals.[40-43] However, the realization of asymmetric intermolecular C(sp^3)—H carbene insertions has represented a major challenge, with viable strategies being largely limited to rhodium-based systems and 'donor-acceptor' carbene transfer reagents.[44] Building upon key advances in the development of engineered hemoproteins as 'carbene transferases' for olefin cyclopropanation and carbene heteroatom-hydrogen insertion,[25] the reaction scope of these systems have been recently extended to the functionalization of C—H bonds.[23, 34] Using artificial metalloenzymes containing an metallo-substituted hemproteins, we and the Hartwig group reported the C(sp^3)—H

alkylation in phthalan substrates (Figure 2d)[45-47] and C—H functionalization in indoles[48] using ethyl diazoacetate (EDA) as carbene precursor.[45-47] More recently, Arnold and coworkers reported an engineered biocatalyst derived from $P450_{BM3}$ ('P411-CHF') for the enantioselective insertion of EDA into benzylic and allylic C(sp^3)—H bonds (Figure 2e).[49] This biocatalyst was later optimized for the α-C—H alkylation of secondary anilines in the presence of a diazolactone reagent or trifluorodiazoethane.[50, 51] Despite this progress, the range of available strategies for biocatalytic C(sp^3)—H functionalization via carbene transfer remains scarce. Here, we describe the development of a versatile biocatalytic platform for the efficient and enantioselective α-C(sp^3)—H functionalization of cyclic amines using diazoacetone and ethyl diazoacetate as carbene donors (Figure 2). This approach offers a simple, scalable, and sustainable route to the preparation of enantioenriched α-functionalized cyclic amines amenable to further diverstification using different chemistries. In addition, we show that these methodologies can be combined to afford difunctionalized products with diastereodivergent selectivity (Figure 2) and applied to late-stage functionalization of a drug precursor, which highlights their value for asymmetric synthesis and medicinal chemistry, respectively.

Results and Discussion

Development of CYP119 catalysts for α-C—H functionalization of *N*-phenylpyrrolidine with diazoacetone

As an initial goal, we targeted the development of a biocatalyst that can promote the C—H functionalization of *N*-substituted pyrrolidine substrate **2a** in the presence of diazoacetone (**1a**), a reaction previously unreported for both chemical and biological catalysts (Table 1). While diazoketones are versatile yet underexplored reagents for carbene transfer reaction, we envisioned this reaction would enable C—H functionalization of the cyclic amine with concomitant installation of a keto group, which is amenable to further diversification via known chemistries.[52] In initial efforts, we screened an in-house library of engineered myoglobin variants that were previously shown to have high activity for a variety of carbene transfer reactions, including the C—H functionalization of indoles[48] and cyclopropanation with benzyldiazoketones.[52] However, none of these myoglobin variants show activity in the reaction of **1a** with **2a** either as purified proteins or in whole cells. These results prompted us to consider other hemoprotein scaffolds. Given the success with applying engineered $P450_{BM3}$ variants for benzylic C—H functionalization with ethyl diazoacetate (EDA) but cognizant of their poor reactivity with diazoacetone,[49] we directed our attention to other members of the cytochrome P450 enzyme family, namely $P450_{cam}$ (camphor hydroxylase)[53], the thermophilic P450 CYP119 (Figure 3a) from *Sulfolobus solfataricus*[54], and the explosive-degrading P450 XplA[55], which was recently shown to possess non-native C—H amination reactivity via nitrene transfer.[56] However, none of these enzyme show detectable activity in the target reaction either as purified proteins and in whole cells (Table 1; SI Table S1-S2). The highly evolved P411-CHF[49], along with transition metal catalysts known to catalyze carbene C—H insertion reactions[57] (SI Table S1) also failed to produce the desired C—H insertion product, further highlighting the challenges associated with this transformation.

Given the lack of activity of wild-type P450 XplA, $P450_{cam}$, and CYP119, we generated a set of active site 'alanine-scanning' libraries in which amino acid residues lying in close proximity to heme cofactor are substituted for alanine (5 sites for XplA, 11 sites for $P450_{cam}$, and 6 sites for CYP119; SI Table S2), with the goal of systematically varying the active site shape of these enzymes and identify mutations that could favor interaction with the non-native substrate. The corresponding variants were expressed in *E. coli* C41(DE3) and tested for their activity in the reaction with **2a** and diazoacetone (**1a**) as whole cells. While the large majority of these enzyme variants showed no activity in the reaction (SI Table S2), CYP119 variant T213A was found to exhibit basal activity (0.4% yield) toward formation of the C—H functionalization product **3a**/**4a** (Table 1, entry 6). Using CYP119(T213A) as starting point, we next evaluated the effects of mutating the heme axial ligand as a means to tune the reactivity of the enzyme, since this approach has proven valuable in the context of other hemoprotein-catalyzed carbene transfer reactions.[46, 58-60] Accordingly, the heme-coordinating Cys317 residue in CYP119(T213A) was mutated to each of the other proteinogenic nucleophilic amino acid residues such as His, Ser, Thr, Tyr, Arg, Lys, Asp, and Glu. Among these variants, CYP119(T213A, C317S) showed slightly improved activity compared to the parent enzyme, producing **3a**/**4a** in 2% yield and 69:31 enantiomeric ratio (Table 1, Entry 7). The CO-bound ferrous form of this serine-ligated CYP119 variant displays a Soret peak at 406 nm (Figure 3b). A similar blue-shift in the Soret band has been observed upon an axial Cys→Ser substitution in $P450_{BM3}$[58], although the λ_{max} of the ferrous CO-bound form of these enzymes clearly depends on the type of P450 under investigation (e.g., 411 nm for $P450_{BM3}$,[58] 416 nm for $P450_{cam}$(C353S) variants in Table S2).

To improve its reactivity toward C—H functionalization of **2a**, CYP119(T213A, C317S) was then subjected to iterative rounds of active site mutagenesis and screening (Figure 3c). Specifically, active site residues L69, F153, L205, A209, and V254 were individually randomized using a combination of partial amino acid alphabets (KBG, WDC, MHG degenerate codons in 1:1:1 ratio) comprising (mostly) uncharged residues of variable size (Gly, Ala, Ser, Val, Leu, Trp, Thr, Asn, Gln, Cys, Pro, Ile, Met, Phe, Thr). The libraries were expressed in *E. coli* C41(DE3) and screened as whole cells in 96-well plates. Using this strategy, a dramatic improvement in both activity and stereoselectivity of the enzyme for the synthesis of **3a** could be achieved after three rounds of directed evolution, as summarized in Figure 3c. In particular, accumulation of a beneficial mutation at position 254 (V254A) led to a tripled mutant variant (called 'CHI-g1') that displays improved yield (2→14%) and enantioselectivity (94:6 vs. 69:31 e.r.) compared to CYP119(T213A, C317S) (Table 1, Entry 8). Using CHI-g1 as the parent, introduction of another space-creating mutation at the level of active site residue 209 (Ala209→Gly) gave a variant (called 'CHI-g2') featuring excellent enantioselectivity (99.5:0.5 er) as well as dramatically (>15-fold) improved TON for the conversion of **2a** into **3a** (200→3,100 TON; Table 1, Entry 9). Despite the high TON, the product yield obtained in whole cell reactions using this variant was still moderate (23%). CHI-g2 was thus subjected to another round of KBG/WDC/MHG-based mutagenesis at the yet unaltered active site positions (i.e. L69, F153, L205) leading to identification of an improved variant carrying a F153G mutation, called 'CHI-DA', (Table 1, Entry 10). Using whole cell reactions at a cell density (OD_{600}) of 40, CYP119 variant CHI-DA delivers **3a** in

quantitative yield as well as with excellent enantioselectivity (>99% *ee*), supporting 12,900 turnovers (Table 1, Entry 10). The same yield and enantioselectivity were also obtained for reactions with purified protein at 0.2 mol% (Table 1, Entry 11). The (*S*)-absolute configuration of **3a** was assigned based on crystallographic analysis of the related product **3e** (SI Figure S4).

From time course experiments, the CHI-DA catalyzed conversion of **2a** into **3a** was determined to proceed with an initial turnover frequency (TOF) of 150 turnovers/min and to reach completion within 12 hours (SI Figure S1). Furthermore, under catalyst-limited conditions (OD_{600} = 10) CHI-DA was determined to catalyze the C-H functionalization of **2a** with over 20,000 turnovers (Table 1, Entry 12). Thus, in addition to offering excellent enantioselectivity, the catalytic activity (TON) of this evolved CYP119-based carbene transferase is one to two orders of magnitude higher than those previously achieved with engineered P411s on other C—H carbene insertion reactions (Figure 2e).[49-51]

Substrate scope of CHI-DA with pyrrolidine derivatives

To explore the substrate scope of CHI-DA, this enzyme was challenged with a panel of variously substituted *N*-aryl pyrrolidine substrates (**2b**-**2l**) in the presence of diazoacetone (**1a**) (SI Table S3). Among these compounds, *para*-substituted *N*-phenyl-pyrrolidine derivatives such as **2b** (*p*-Me), **2e** (*p*-F), and **2f** (*p*-Cl) were efficiently converted (5,150-9,340 TON) into the desired products **3b**, **3e**, and **3f** in good to high yields (51-93%) and enantioselectivity (86:14 to 91:9 e.r.) (SI Table S3). The *para*-methoxy derivative **3g** was also afforded in good yield and high TON (5,240) albeit with more moderate enantioselectivity (75:25 e.r.). In contrast, lower levels of activity and/or enantioselectivity were observed in the presence of the *ortho*- and *meta*-substituted *N*-phenyl-pyrrolidine derivatives (SI Table S3), indicating a lower tolerance of the enzyme to substitutions other than at the para position. To overcome this limitation, the most active members from the enzyme libraries derived from CHI-g1, CHI-g2, and CHI-g3 were screened against these substrates, resulting in the identification of catalysts with improved activity and enantioselectivity for the synthesis of **3c** (*m*-Me), **3h** (*m*-Br), **3i** (*o*-CN), as well as the pyridyl derivative **3j** (Figure 4). Using CHI-g2(L205V), the latter product (**3j**) could be obtained in 3-fold higher yield (59% vs. 20%), 10-fold higher TON, and higher enantioselectivity (87:13 vs. 60:40 er) compared to CHI-DA. In general, CHI-g2(L205V) shows significantly improved activity and enantioselectivity across multiple substrates not well accepted by CHI-DA (e.g., **3h**: 32% vs. 5% yield and 94:6 vs. 85:15 e.r.; Figure 4 and Table S3), thus offering a useful complementarity to the latter enzyme in terms of substrate scope. A racemic substrate bearing α methyl group to the pyrrolidine nitrogen (**2k**) was also prepared and tested. Notably, the corresponding C—H insertion product **3k** was obtained in 86:14 diastereomeric ratio using either CHI-DA or CHI-g2 (Figure 4), suggesting the potential utility of these enzymes also for kinetic resolution of racemic substrates upon further optimization of the biocatalyst for this application. In contrast, no enantioselectivity could be achieved with **3d** (Figure 4). In addition to the pyrrolidine-based substrates, efficient and highly enantioselective α-C—H functionalization could be achieved in the presence of other heterocyclic substrates such as piperidine, to give **3m** in 46% yield (14,210 TON; 52:48 e.r.), and phthalan, whose transformation delivered **3l** in 67%

yield (22,100 TON) and excellent enantioselectivity (99:1 e.r.) using CHI-g2(A13G) as the catalyst (Figure 4). Collectively, these CYP119-based biocatalysts are capable of processing a broad range of *N*-aryl pyrrolidines and other saturated heterocycles with high TON and, in most cases, high stereoselectivity (Figure 4).

C—H functionalization of *N*-phenyl-pyrrolidine with EDA

Given the high activity and selectivity of the engineered CYP119 catalysts in the α-C—H functionalization with diazoacetone, we extended our investigations to the C—H functionalization of *N*-phenyl-pyrrolidine in the presence of ethyl diazoacetate (EDA; **1b**). While this reaction was previously investigated using P411-CHF, its scope was limited to a few substrates (3).[49] On the other hand, other hemoproteins and heme enzymes showed no activity (SI Table S4). Following a similar strategy as described above for the diazoacetone reaction, the CYP119-based active-site alanine scanning library was initially screened in whole cells by targeting C—H functionalization of **2a** with EDA as a model reaction. Unlike the diazoacetone counterpart, wild-type CYP119 and most of the alanine mutants showed detectable activity in this reaction (3-6% yields; 130-250 TON; SI Table S5) with the L69A, T213A, or V254A single mutation each providing a 2-fold improvement in yield compared to the wild type enzyme. Combination of the most beneficial alanine mutation (T213A) with the axial ligand mutation C317S led to a 3-fold improvement in yield (6→18%) and 6-fold improvement in TON (135→812), furnishing **6a** with moderate enantioselectivity (71:29 e.r.; Table 2, entry 4). Using CYP119(T213A, V254A) as background, systematic mutagenesis of the heme coordinating cysteine residue (Cys317) established that, while all of these axial ligand variants remain functional, the Cys→Ser mutation is again most beneficial for supporting this carbene transfer reaction (SI Figure S3). Notably, the corresponding triple mutant variant, CYP119(T213A, V254A, C317S) (called ‘CHI-g1’) was found to catalyze the formation of **6a** in quantitative yield (99%) and slightly improved enantioselectivity (79:21 e.r.) compared to CYP119(T213A, C317S) (Table 2, entry 5). To further enhance the enantioselectivity for this reaction, the CHI-g1 derived libraries generated via KBG/WDC/MHG-mutagenesis of F153, L205, A209 and G210 were screened in the presence of **2a** and EDA. This process led to the identification of an improved catalyst for this reaction, CYP119(F153G, T213A, V254A, C317S) (called ‘CHI-EDA’), which is capable of delivering the desired C—H insertion product **6a** in 89% yield, 87:13 e.r. with 8,920 TON in whole-cell reactions (Table 2, entry 6). Similar results (99% yield; 87:13 e.r.) were obtained using this enzyme in purified form (Table 2, entry 7). From a preparative-scale reaction with CHI-EDA expressing cells, 182 mg of **6a** was readily obtained in 82% isolated yield, demonstrating the scalability of this reaction. Notably, this biocatalytic system was found to furnish **6a** in good yield (51%) with TON of 5,090 in the presence of air after two hours reaction time (Table 2, entry 9). This activity corresponds to a mere 2-fold decrease in activity compared to that measured under anaerobic conditions (Table 2, entry 9 vs. 8), further highlighting the efficiency of this biocatalyst. To our knowledge, this is the first example of a biocatalytic C—H carbene insertion reaction *achieved under aerobic conditions.*

While CHI-EDA catalyzes the C—H alkylation of **2a** with (*S*)-enantioselectivity, enantiodivergent selectivity in this transformation could be achieved using CYP119(A209W,

T213G, V254A, C317S), which produces the (*R*)-configured product **7a** in 71% yield and 81:19 e.r. (Table 2, entry 10). In addition, kinetic experiments showed that the CHI-EDA-catalyzed reaction with **2a** and EDA is remarkably fast, proceeding with an initial turnover frequency (TOF) of 2,900 turnovers per minute and reaching completion within only two hours (SI Figure S2). This reaction is about 20-fold faster than the CHI-DA-catalyzed reaction with **2a** and diazoacetone (**1a**), which proceeds with a TOF of 150 turnovers min^{-1}) and reaches completion in 12 hours (SI Figure S1).

C-H functionalization of cyclic amines with EDA

Upon challenging CHI-EDA with other *N*-aryl-pyrrolidine derivatives (**2b**-**2n**), lower levels of activity were generally observed compared to **6a** (e.g., 27-34% yield for **6e** and **6f**), indicating a more pronounced substrate specificity for this biocatalyst and reaction as opposed to broader substrate scope of the CYP119 biocatalysts for the reactions with diazoacetone (e.g., CHI-DA; SI Table S3). This behaviour is reminiscent of that of engineered cytochrome P450s in native (monooxygenation) reactions or non-native reactions.[62] This finding prompted us to pursue a substrate vs. library approach for identifying better catalysts for these target substrates. Accordingly, the latter were screened in parallel against the most active CYP119 variants identified during the evolution of CHI-DA (Figure 3c). Among them, the CHI-g2-derived variants proved most effective for the enantioselective C—H functionalization of substituted *N*-aryl-pyrrolidines, giving **6b**, **6c**, **6f**, **6h**-**6k**, and **6m** in up to 97% yield and 81:19 enantiomeric ratio (Figure 5). On the other hand, products **6d**, **6e**, **6g**, and **6n** were most efficiently afforded using CHI-g3 derived variants, with up to 95% yield and 96:4 e.r. (Figure 5). Notably, using their respective optimal catalysts, **6e** *(p-F)* and **6i** *(o-CN)* were obtained in nearly quantitative yields with over 12,000-18,000 TON, which corresponds to 5-10 fold higher TON values than previously reported for P411-CHF in a related reaction.[49] Furthermore, along with the previously mentioned enantiodivergent synthesis of **7a** using CHI-g3 (Table 2), catalysts useful for the synthesis of **7b, 7f** and **7m** with inverted enantioselectivity were also discovered (e.g., 31:69 e.r. for **7b** (*p*-Me) with CHI-g2(G210T) vs. 76:24 e.r. with CHI-g2(G210T); Figure 5).

Compared to the diazoacetone-driven reactions, the CYP119-catalyzed reactions in the presence of EDA was determine to proceed with lower activity but significantly higher enantioselectivity with N-phenyl-piperidine as the substrate (**6m**, 53% yield, 85:15 e.r.), and they could be further extended to the α-C—H functionalization of a morpholine substrate (**6n,** 23% yields, 87:13 e.r.; Figure 5), thus demonstrating the scope of this methodology across other important types of cyclic amine scaffolds. Furthermore, both stereoisomers of the piperidine product **6m** and **7m** were obtained in enantioenriched form (85:15 and 28:72 e.r.) using stereodivergent variants (Figure 5). Notably, these reactions can be readily scaled up to obtain the desired C—H insertion products on a semi-preparative scale (60-100 mg).

Stereodivergent dual C—H functionalization with diazoacetone and EDA

During the screening of the engineered CYP119 libraries, we noticed the presence of highly active variants capable of catalyzing a double C—H insertion in *N*-phenylpyrrolidine with both diazoacetone and EDA. These double insertion products were isolated and determined

by NMR spectroscopy to correspond to 2,5-disubstituted products as both *cis* and *trans*-isomers (**5a-c** and **8a-c**; Figure 6a). We further noticed that the ratio between the *cis* and *trans*-isomers varied among different CYP119 variants, suggesting that the stereoselectivity of the double insertion reaction can be tuned via protein engineering. Accordingly, we sought to extend the scope of the present methods to the stereodivergent dual C—H functionalization through both a one-pot/single diazo reagent reaction and via a tandem process with two different diazo reagents.

Toward the former goal, the engineered CYP119 libraries were screened for the functionalization of **2a** with EDA in the presence of excess diazo compound (4 equiv.). From this screening, diastereo- and enantiodivergent CYP119 variants capable of selectively yielding each of the three possible stereoisomeric products, namely *cis*-(*S,R*)-**8a**, *trans*-(*S,S*)-**8b**, and *trans*-(*R,R*)-**8c**, were identified (Figure 6b-c). For formation of meso-compound *cis*-(*S,R*)-**8a**, CYP119(T213A, C317S) (named 'DCHI-EDA$_{cis}$') showed excellent diastereoselectivity (99:0.5:0.5 d.r.) along with high activity (61% yield) (Figure 6c). On the other hand, the two *trans*-enantiomers *trans*-(*S,S*)-**8b**, and *trans*-(*R,R*)-**8c** could be obtained with excellent diasteroselectivity and enantioselectivity (99% *ee*) using two related variants, i.e. CYP119(F153Y, T213G, V254A, C317S) (called 'DCHI-EDA$_{trans\text{-}SS}$') and CYP119(F153A, T213G, V254A, C317S) (named 'DCHI-EDA$_{trans\text{-}RR}$') (Figure 6c). Interestingly, a complete inversion in the enantiopreference in these enzymes can be ascribed to a single mutation in position 153 (Ala vs. Tyr). Albeit with lower diastereoselectivity than DCHI-EDA$_{trans\text{-}SS}$, CHI-EDA shows also high activity and excellent enantioselectivity for formation of *trans*-(*S,S*)-**8b** in the presence of excess EDA (99.5:0.5 e.r., Figure 6c).

These findings then prompted us to pursue a strategy for achieving a stereoselective dual C—H functionalization of the cyclic amine substrate via a combination of the two methodologies for C—H insertion with diazoacetone and with EDA described above. To this end, a gram-scale synthesis of the diazoacetone insertion product **3a** was carried out via a whole-cell biotransformation of **2a** with diazoacetone with *E. coli* cells expressing variant CHI-DA, resulting in the isolation of 1.24 gram of **3a** in high enantiopurity (>99% *ee*) and 64% isolated yield (Figure 6d). The enzymatic product was then applied to screen the CYP119 libraries for variants capable of catalyzing the α-C—H functionalization of this substrate to produce the difunctionalized products **9** and **10** with diastereodivergent selectivity (Figure 6d). While multiple variants accepted **3a** to produce the desired difunctionalized products, CYP119(L205T, A209G, T213A, V254A, C317S) was found to catalyze the efficient and selective formation of the *cis* product **9** (69:31 d.r., 63% yield), whereas CYP119(T213G, C317S) offered complementary diastereoselectivity in this reaction by favoring the formation of *trans*-product **10** (72:28 d.r., 76% yield). Both products were obtained in high enantiopurity (>99% *ee*) owing to the excellent enantioselectivity stereocontrol of the prior CHI-DA catalyzed step.

In addition to representing first examples of a dual enzyme-catalyzed C-H carbene insertion on a single substrate, these results also illustrate the value of the present strategies toward enabling the synthesis of stereoisomeric and enantioenriched compounds decorated with multiple functional groups (e.g., ester/keto group) that are readily amenable for further

functionalization. Among other applications, these types of compounds can be valuable building blocks for generating stereoisomeric libraries in drug discovery campaigns.[63, 64]

Late-stage C—H functionalization of advanced pharmaceutical intermediate

In the interest of deteterming whether the present methodologies could be extended to late-stage functionalization of drug scaffolds, we targeted the core structure of the antidiabetic drug sitagliptin (Januvia®) (Figure 7) for C—H functionalization. In addition to the presence of a metal coordinating group (triazole), this molecule features a fused piperazine ring that contains multiple α-C—H bonds of similar reactivity, thus presenting a challenge in terms of regioselectivity for late-stage functionalization via chemical means. To identify catalysts for its modification, the MOM-protected compound **11a** was screened against the library of engineered CYP119 variants in the presence of EDA. Gratifyingly, CYP(T213A, C317S) was found to catalyze the C(sp^3)—H functionalization of the drug core molecule to give **12** in 79 : 21 e.r. and 25% yield after removal of the MOM protecting group (Figure 7). Notably, the same enzyme was able to accept the unprotected core (**11b**) to produce the dual N—H/C—H insertion product **13** with similar enantioselectivity (74 : 26 e.r.), along with the N—H insertion product **14** (Figure 7). Altogether, these results provide a proof-of-principle demonstration of the value of the present biocatalytic strategy for late-stage functionalization of drug scaffolds.

Conclusion

In summary, we have reported the development of a highly enantioselective, biocatalytic strategy for the asymmetric α-C—H functionalization of pyrrolidines with diazoacetone as carbene donor reagent. In particular, CYP119-derived biocatalysts were evolved to catalyze this challenging reaction with high efficiency and high enantioselectivity across multiple substrates (up to 99:1 e.r.), supporting over 20,000 catalytic turnovers. The latter corresponds to the highest catalytic activity for an enzyme-catalyzed intermolecular carbene C—H insertion reaction reported to date, and it compares well with the highest TON values reported for native monooxygenation reactions catalyzed by engineered P450s.[65-73] These biocatalytic reactions can be performed in whole cells and are readily scalable, as illustrated by gram-scale synthesis of **3a** (Figure 6). This methodology was further extended to the enantioselective α-C—H functionalization of a range of cyclic amine substrates, including pyrrolidine, piperidine, and morpholine scaffolds, using EDA. Furthermore, by combining the two methodologies, it was possible to accomplish both a one-pot and a tandem dual α—C—H functionalization with diastereo- and enantiodivergent selectivity, providing access to valuable polyfunctionalized building blocks in different stereochemical configurations. Finally, this approach was successfully applied to the selective late-stage C—H functionalization of the core structure of a drug molecule (sitagliptin), providing a direct path for its functionalization or diversification. Complementing other biocatalytic strategies,[74-78] these biocatalysts are expected to expand opportunities for the synthesis and diversification of bioactive molecules containing saturated *N*-heterocycles and we anticipate that iron-based CYP119 derived catalysts can be leveraged for other types of synthetically useful carbene C—H insertion reactions.

Experimental Procedures

General information

All chemicals were purchased from commercial sources or provided by Pfizer and used without further purification. Substrates **2b-2d** were synthesized according to the reported procedure. Unless stated otherwise, all column purification were performed on a Biotage Selekt Flash Chromatography, eluted with ethyl acetate: hexane, 0-15% gradient, detected by UV absorption at 254nm. Thin Layer Chromatography (TLC) was carried out using Merck Millipore TLC silica gel 60 F254 glass plates. 1H, 13C, and 19F NMR spectra were measured on a Bruker DPX-400 instrument (operating at 400 MHz for 1H, 100 MHz for 13C, and 375 MHz for 19F) or a Bruker DPX-500 instrument (operating at 500 MHz for 1H and 125 MHz for ^{13}C).

Protein Expression

Wild-type and engineered CYP119 variants were expressed in *E. coli* C41(DE3) cells as follows. After transformation, cells were grown in LB medium (ampicillin, 100 mg/L) at 37°C (200 rpm) until OD_{600} reached 1 to 1.2. Cells were then induced with 0.25 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) and 0.3 mM δ-aminolevulinic acid (δ-ALA). After induction, cultures were shaken at 180 rpm and 27°C and harvested after 20 hours by centrifugation at 4,000 rpm at 4°C. The cells were resuspended in 20 mL of Ni-NTA Lysis Buffer (50 mM KPi, 250 mM NaCl, 10 mM histidine, pH 8.0). Resuspended cells were frozen and stored at −80°C until purification. Cell suspensions were thawed at room temperature, lysed by sonication, and clarified by centrifugation (14,000 rpm, 50 min, 4°C). The clarified lysate was transferred to a Ni-NTA column equilibrated with Ni-NTA Lysis Buffer. The resin was washed with 50 mL of Ni-NTA Lysis Buffer and then 50 mL of Ni-NTA Wash Buffer (50 mM KPi, 250 mM NaCl, 20 mM histidine, pH 8.0). Proteins were eluted with Ni-NTA Elution Buffer (50 mM KPi, 250 mM NaCl, 250 mM histidine, pH 7.0). After elution from the Ni-NTA column, the protein was buffer exchanged against 50 mM KPi buffer (pH 7.0) using 10 kDa Centricon filters. The concentration of the CYP119 variants was determined from CO-binding assays (difference spectra) using ε_{406} = 100 $mM^{-1}cm^{-1}$ as the extinction coefficient.

Protein Engineering

Protein evolution was conducted through iterative rounds of site saturation mutagenesis based on active site residues which showed different product profiles with their alanine-scanning libraries. In each round, the mutagenesis was conducted using the Quickchange method.[79] A mixture of DNA degenerate primers (KBG:WDC:MHG = 1:1:1) encoding (mostly) uncharged amino acid residues of variable size (Gly, Ala, Ser, Val, Leu, Trp, Thr, Asn, Gln, Cys, Pro, Ile, Met, Phe, Thr) were used and the PCR products were transformed into *E. coli* DH5a cells after digestion with DpnI restriction enzyme. The colonies were collected in LB medium (ampicillin, 100 mg L^{-1}) and plasmids were extracted by QIAprep Spin Miniprep Kit (Cat No.27104). Library coverage was then assessed by DNA sequencing to confirm the incorporation of desired mutations. The library of CYP119 variants was then transformed into *E. coli* DH5α cells and the proteins expressed in 96-well plates under the conditions described above. After expression, the cells were pelleted by centrifugation and

resuspended in KPi buffer (50 mM, pH 7). The reactions were initiated by adding substrate into each well of the plate in an anerobic chamber, followed by shaking for 12 hours. The reactions mixtures were extracted with DCM and analyzed by chiral GC-FID or chiral SFC. The CYP119 variant that showed improved activity and enantioselectivity was sequenced and used as template for next round of mutagenesis and protein evolution.

Enzymatic Reactions

Analytical scale enzymatic reactions with purified proteins were carried out at a 500 μL-scale using the CYP119 variant (or other protein), cyclic amine substrate, diazoacetone (or ethyl diazoacetate) and sodium dithionite $Na_2S_2O_4$ at the concentrations indicated in the Tables and legends. In a typical procedure, a solution of $Na_2S_2O_4$ in potassium phosphate buffer (50 mM, pH 7.0) was degassed by bubbling argon into the mixture for 3 min in a sealed vial. A buffered solution containing the CYP119 variant was carefully degassed in a similar manner in a separate vial. The two solutions were then mixed via cannula transfer. Reactions were initiated by addition of N-phenylpyrrolidine derivative (from a 0.5 M stock solution in ethanol), followed by the addition of diazoacetone/ethyl diazo acetate (from a 0.5 M stock solution in ethanol) with a syringe, and the reaction mixture was stirred for 16 hours at room temperature, under positive argon pressure. For whole cell experiments, reactions were carried out at a 500 μL-scale using *E. coli.* whole cells expressing the CYP119 variant, cyclic amine substrate, and diazoacetone (or ethyl diazoacetate) at the concentrations indicated in the Tables and legends. In a typical procedure, a sealed vial containing whole cells was degassed with argon for 3 min. The reactions were initiated by addition of N-phenylpyrrolidine derivative (from a 0.5 M stock solution in ethanol), followed by the addition of diazo compound (from a 0.5 M stock solution in ethanol) with a syringe. The reaction mixture was stirred for 16 hours at room temperature under positive argon pressure. The TON for the whole-cell reactions were calculated based on CYP119 concentration in the reaction mixture as measured via UV-vis spectroscopy after cell lysis.

Reaction Analysis

The reactions were analyzed by adding 25 μL of internal standard (benzodioxole, 50 mM in methanol) to a 500 μL aliquot of the reaction mixture, followed by extraction with 500 μL dichloromethane (DCM) and centrifugation at 14,000 rpm. The organic layer was collected and analyzed by GC for yield, and chiral SFC for enantioselectivity. The TON for the whole-cell reactions were calculated based on CYP119 concentration in the reaction mixture as measured via UV-vis using the CO-binding assay (ε_{406} = 100 $mM^{-1}cm^{-1}$) after cell lysis. Calibration curves of the different products were constructed using authentic standards from the whole cell reactions (General Procedure A). Enantioselectivity was determined by SFC using a chiral column as described in analytical methods section.

General Procedure A: Whole-cell biocatalytic reactions for C-H functionalization on a preparative scale

These reactions were carried out on a 40 mL-scale using C41(DE3) *E. coli* cells expressing the CYP119 variant, 10 mM cyclic amine substrate, 20 mM diazo reagent (diazoacetone or EDA). In a typical procedure, the substrate (0.4 mmol in 1 mL of ethanol) was added slowly to a 125 mL Erlenmeyer flask containing a stirring suspension of CYP119-expressing cells

(OD_{600} = 40 in KPi, pH 7) in an anaerobic chamber. After stirring for 5 minutes, 1.6 mL of 500 mM diazo solution (2 equiv.) in ethanol were added into the Erlenmeyer flask, then sealed with a rubber septum. Reaction mixture stirred at room temperature overnight. The reaction mixtures were extracted with ethyl acetate (100 mL x 3) and the combined organic layers were dried over $MgSO_4$ and concentrated under reduced pressure. The TON for the whole-cell reactions were calculated based on CYP119 concentration in the reaction mixture as measured via UV-vis spectroscopy using the CO-binding assay (ε_{406} = 100 $mM^{-1}cm^{-1}$) after cell lysis. The crude product was purified by flash column chromatography using silica gel and ethyl acetate/hexanes as the eluent to isolate the product. The purified product was characterized by NMR, GC-MS, and chiral SFC for stereoselectivity determination and they were used as authentic standards for the construction of the calibration curves (TON and % conversion determination). Functionalization of the sitagliptin core was carried out in a similar manner using CYP119(T213A, C317S) expressing *E. coli* C41(DE3) cells (OD_{600} = 120), 2.5 mM **11a** or **11b**, 40 mM EDA, in KPi buffer (50 mM, pH 7), room temperature, 16 hours, in anaerobic chamber. Product yields and characterization data are provided as Supporting Information.

General Procedure B: Whole-cell biocatalytic reactions for dual C-H functionalization with EDA on a preparative scale

Same procedure as General Procedure A with the only difference that 4 equivalent of the EDA (3.2 mL of 0.5 M stock solution) were used instead of 2 equivalents. The crude product was purified by flash column chromatography using silica gel and ethyl acetate/hexanes as the eluent to isolate the product. The purified product was characterized by NMR, GC-MS, and chiral SFC for stereoselectivity determination and they were used as authentic standards for the construction of the calibration curves (TON and % conversion determination). Product yields and characterization data for **8a-c** are provided as Supporting Information.

Biocatalytic synthesis of difunctionalized products 9a and 10a

These reactions were carried out in a two-step process. In a first step, *N*-phenylpyrrolidine (**2a**) (1.40 g in 2.5 mL of ethanol) was added slowly to a 125 mL Erlenmeyer flask containing a stirring suspension of *E. coli* cells expressing CHI-DA (190 mL, OD_{600} = 40 in KPi, pH 7) in an anaerobic chamber. After stirring for 5 minutes, 15 mL of 1 M diazoacetone (**1a**) (2 equiv.) solution in ethanol were added dropwise into the Erlenmeyer flask, then sealed with a rubber septum. Reaction mixture stirred at room temperature overnight. The reaction mixture was extracted with ethyl acetate (3 x 100 mL) and the combined organic layers were dried over $MgSO_4$ and concentrated under reduced pressure. The TON for the whole-cell reactions were calculated based on CYP119 concentration in the reaction mixture as measured via UV-vis spectroscopy using the CO-binding assay (ε_{406} = 100 $mM^{-1}cm^{-1}$) after cell lysis. The crude product was purified by flash column chromatography using silica gel and ethyl acetate/hexanes (1:4) as the eluent to afford enantiopure **3a** (>99% *ee*) in 64% isolated yield (1.93 g). For the second step, **3a** (0.4 mmol in 1 mL of ethanol) was added slowly to a 125 mL Erlenmeyer flask containing 40 mL solution of *E. coli* cells (OD_{600} = 40 in KPi, pH 7) expressing the desired CYP119 variant (as in Figure 7) in an anaerobic chamber. After stirring for 5 minutes, 1.6 mL of 500 mM

EDA (**1b**) (2 equiv.) solution in ethanol were added to the cell suspension and then the flask was sealed with a rubber septum. The reaction mixture was stirred at room temperature overnight. The reaction mixtures were extracted with ethyl acetate (100 mL x 3) and the combined organic layers were dried over $MgSO_4$ and concentrated under reduced pressure. The TON for the whole-cell reactions were calculated based on CYP119 concentration in the reaction mixture as measured via UV-vis spectroscopy using the CO-binding assay (ε_{406} = 100 $mM^{-1}cm^{-1}$) after cell lysis. The crude product was purified by flash column chromatography using silica gel and ethyl acetate/hexanes (1:4) as the eluent to afford **9a** in 63% isolated yield and **10a** in 76% isolated yield. The purified product was characterized by NMR, GC-MS, and chiral SFC for stereoselectivity determination and they were used as authentic standards for the construction of the calibration curves (TON and % conversion determination).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by the National Institute of General Medical Sciences (NIGMS) R01 grant GM098628 (R.F.). The authors are grateful to Pfizer for providing financial support and compounds to study the modification of medicinal compounds. We thank Drs. Douglas K. Spracklin and R. Scott Obach (Pfizer) for valuable discussions, Dr. James Bradow (Pfizer) for determining the enantiomeric ratio of compound **6i**, Dr. William Brennessel (U. Rochester) for assistance with crystallographic analyses. MS and X-ray instrumentation are supported by U.S. National Science Foundation grants CHE-0946653 and CHE-1725028 and the U.S. National Institute of Health grants S10OD030302 and S10OD025242.

REFERENCES

(1). Vitaku E, Smith DT, and Njardarson JT Analysis of the Structural Diversity, Substitution Patterns, and Frequency of Nitrogen Heterocycles among US FDA Approved Pharmaceuticals, J. Med. Chem 2014, 57, 10257–10274. [PubMed: 25255204]

(2). Campos KR Direct sp(3) C-H bond activation adjacent to nitrogen in heterocycles, Chem. Soc. Rev 2007, 36, 1069–1084. [PubMed: 17576475]

(3). Mitchell EA, Peschiulli A, Lefevre N, Meerpoel L, and Maes BUW Direct alpha-Functionalization of Saturated Cyclic Amines, Chem. Eur. J 2012, 18, 10092–10142. [PubMed: 22829434]

(4). Chen WJ, Ma LL, Paul A, and Seidel D Direct alpha-C-H bond functionalization of unprotected cyclic amines, Nat. Chem 2018, 10, 165–169. [PubMed: 29359746]

(5). Campos KR, Klapars A, Waldman JH, Dormer PG, and Chen CY Enantioselective, palladium-catalyzed alpha-arylation of N-Boc-pyrrolidine, J. Am. Chem. Soc 2006, 128, 3538–3539. [PubMed: 16536525]

(6). Beak P, Kerrick ST, Wu SD, and Chu JX Complex-Induced Proximity Effects - Enantioselective Syntheses Based on Asymmetric Deprotonations of N-Boc-Pyrrolidines, J. Am. Chem. Soc 1994, 116, 3231–3239.

(7). McGrath MJ, and O'Brien P Catalytic asymmetric deprotonation using a ligand exchange approach, J. Am. Chem. Soc 2005, 127, 16378–16379. [PubMed: 16305208]

(8). Seel S, Thaler T, Takatsu K, Zhang C, Zipse H, Straub BF, Mayer P, and Knochert P Highly Diastereoselective Arylations of Substituted Piperidines, J. Am. Chem. Soc 2011, 133, 4774–4777. [PubMed: 21388211]

(9). Cordier CJ, Lundgren RJ, and Fu GC Enantioconvergent Cross-Couplings of Racemic Alkylmetal Reagents with Unactivated Secondary Alkyl Electrophiles: Catalytic Asymmetric Negishi alpha-

Alkylations of N-Boc-pyrrolidine, J. Am. Chem. Soc 2013, 135, 10946–10949. [PubMed: 23869442]

(10). Feng KB, Quevedo RE, Kohrt JT, Oderinde MS, Reilly U, and White MC Late-stage oxidative C(sp(3))-H methylation, Nature 2020, 580, 621-+. [PubMed: 32179876]

(11). Li ZP, and Li CJ CuBr-catalyzed efficient alkynylation of sp(3) C-H bonds adjacent to a nitrogen atom, J. Am. Chem. Soc 2004, 126, 11810–11811. [PubMed: 15382913]

(12). Girard SA, Knauber T, and Li CJ The Cross-Dehydrogenative Coupling of C-sp3-H Bonds: A Versatile Strategy for C-C Bond Formations, Angew. Chem. Int. Ed 2014, 53, 74–100.

(13). Novaes LFT, Ho JSK, Mao K, Liu K, Tanwar M, Neurock M, Villemure E, Terrett JA, and Lin S Exploring Electrochemical C(sp(3))-H Oxidation for the Late-Stage Methylation of Complex Molecules, J. Am. Chem. Soc 2022, 144, 1187–1197. [PubMed: 35015533]

(14). Chang Z, Huang J, Wang S, Chen G, Zhao H, Wang R, and Zhao D Copper catalyzed late-stage C(sp(3))-H functionalization of nitrogen heterocycles, Nat. Commun 2021, 12, 4342. [PubMed: 34267229]

(15). Pastine SJ, Gribkov DV, and Sames D sp(3) C-H bond arylation directed by amidine protecting group: alpha-arylation of pyrrolidines and piperidines, J. Am. Chem. Soc 2006, 128, 14220–14221. [PubMed: 17076471]

(16). Spangler JE, Kobayashi Y, Verma P, Wang DH, and Yu JQ alpha-Arylation of Saturated Azacycles and N-Methylamines via Palladium(II)-Catalyzed C(sp(3))-H Coupling, J. Am. Chem. Soc 2015, 137, 11876–11879. [PubMed: 26322957]

(17). Jain P, Verma P, Xia GQ, and Yu JQ Enantioselective amine alpha-functionalization via palladium-catalysed C-H arylation of thioamides, Nat. Chem 2017, 9, 140–144. [PubMed: 28282045]

(18). McNally A, Prier CK, and MacMillan DWC Discovery of an alpha-Amino C-H Arylation Reaction Using the Strategy of Accelerated Serendipity, Science 2011, 334, 1114–1117. [PubMed: 22116882]

(19). Shaw MH, Shurtleff VW, Terrett JA, Cuthbertson JD, and MacMillan DWC Native functionality in triple catalytic cross-coupling: sp(3) C-H bonds as latent nucleophiles, Science 2016, 352, 1304–1308. [PubMed: 27127237]

(20). Beatty JW, and Stephenson CRJ Amine Functionalization via Oxidative Photoredox Catalysis: Methodology Development and Complex Molecule Synthesis, Accounts Chem. Res 2015, 48, 1474–1484.

(21). Vasilopoulos A, Krska SW, and Stahl SS. C(sp(3))-H methylation enabled by peroxide photosensitization and Ni-mediated radical coupling, Science 2021, 372, 398–303. [PubMed: 33888639]

(22). Shu XM, Zhong D, Lin YM, Qin X, and Huo HH Modular Access to Chiral alpha-(Hetero)aryl Amines via Ni/Photoredox-Catalyzed Enantioselective Cross-Coupling, J. Am. Chem. Soc 2022, 144, 8797–8806. [PubMed: 35503417]

(23). Ren X, and Fasan R Engineered and Artificial Metalloenzymes for Selective C-H Functionalization, Curr. Opin. Green Sustain. Chem 2021, 31, 100494. [PubMed: 34395950]

(24). Andorfer MC, and Lewis JC Understanding and Improving the Activity of Flavin-Dependent Halogenases via Random and Targeted Mutagenesis, Annu. Rev. Biochem 2018, 87, 159–185. [PubMed: 29589959]

(25). Brandenberg OF, Fasan R, and Arnold FH Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions, Curr. Opin. Biotech 2017, 47, 102–111. [PubMed: 28711855]

(26). Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, and Robins K Engineering the third wave of biocatalysis, Nature 2012, 485, 185–194. [PubMed: 22575958]

(27). Reetz MT Laboratory Evolution of Stereoselective Enzymes: A Prolific Source of Catalysts for Asymmetric Reactions, Angew. Chem. Int. Ed 2011, 50, 138–174.

(28). Turner NJ Directed evolution drives the next generation of biocatalysts, Nat. Chem. Biol 2009, 5, 568–574.

(29). Yokoyama K, and Lilla EA C-C bond forming radical SAM enzymes involved in the construction of carbon skeletons of cofactors and natural products, Nat. Prod. Rep 2018, 35, 660–694. [PubMed: 29633774]

(30). Bauerle MR, Schwalm EL, and Booker SJ Mechanistic Diversity of Radical S-Adenosylmethionine (SAM)-dependent Methylation, J. Biol. Chem 2015, 290, 3995–4002. [PubMed: 25477520]

(31). McLaughlin MI, and van der Donk WA Stereospecific Radical-Mediated B-12-Dependent Methyl Transfer by the Fosfomycin Biosynthesis Enzyme Fom3, Biochemistry 2018, 57, 4967–4971. [PubMed: 29969250]

(32). Tang MEQY, Grathwol CW, Aslan-Uzel AS, Wu SK, Link A, Pavlidis IV, Badenhorst CPS, and Bornscheuer UT Directed Evolution of a Halide Methyltransferase Enables Biocatalytic Synthesis of Diverse SAM Analogs, Angew. Chem. Int. Ed 2021, 60, 1524–1527.

(33). Herbert AJ, Shepherd SA, Cronin VA, Bennett MR, Sung RA, and Micklefield J Engineering Orthogonal Methyltransferases to Create Alternative Bioalkylation Pathways, Angew. Chem. Int. Ed 2020, 59, 14950–14956.

(34). Zhang RJK, Huang XY, and Arnold FH Selective C-H bond functionalization with engineered heme proteins: new tools to generate complexity, Curr. Opin. Chem. Biol 2019, 49, 67–75. [PubMed: 30343008]

(35). Liao K, Yang YF, Li Y, Sanders JN, Houk KN, Musaev DG, and Davies HML Design of catalysts for site-selective and enantioselective functionalization of non-activated primary C-H bonds, Nat. Chem 2018, 10, 1048–1055. [PubMed: 30082883]

(36). Fu JT, Ren Z, Bacsa J, Musaev DG, and Davies HML Desymmetrization of cyclohexanes by site- and stereoselective C-H functionalization, Nature 2018, 564, 395-+. [PubMed: 30568203]

(37). Liao KB, Negretti S, Musaev DG, Bacsa J, and Davies HML Site-selective and stereoselective functionalization of unactivated C-H bonds, Nature 2016, 533, 230–234. [PubMed: 27172046]

(38). Suematsu H, and Katsuki T Iridium(III) Catalyzed Diastereo- and Enantioselective C-H Bond Functionalization, J. Am. Chem. Soc 2009, 131, 14218-+. [PubMed: 19757773]

(39). Weldy NM, Schafer AG, Owens CP, Herting CJ, Varela-Alvarez A, Chen S, Niemeyer Z, Musaev DG, Sigman MS, Davies HML, and Blakey SB Iridium(III)-bis(imidazolinyl)phenyl catalysts for enantioselective C-H functionalization with ethyl diazoacetate, Chem. Sci 2016, 7, 3142–3146. [PubMed: 29997805]

(40). Diaz-Requejo MM, Belderrain TR, Nicasio MC, Trofimenko S, and Perez PJ Intermolecular copper-catalyzed carbon-hydrogen bond activation via carbene insertion, J. Am. Chem. Soc 2002, 124, 896–897. [PubMed: 11829584]

(41). Griffin JR, Wendell CI, Garwin JA, and White MC Catalytic C(sp(3))-H Alkylation via an Iron Carbene Intermediate, J. Am. Chem. Soc 2017, 139, 13624–13627. [PubMed: 28898063]

(42). Che CM, Lo VKY, Zhou CY, and Huang JS Selective functionalisation of saturated C-H bonds with metalloporphyrin catalysts, Chem. Soc. Rev 2011, 40, 1950–1975. [PubMed: 21387046]

(43). Zhu SF, and Zhou QL Iron-catalyzed transformations of diazo compounds, Nat. Sci. Rev 2014, 1, 580–603

(44). Davies HML, and Morton D Guiding principles for site selective and stereoselective intermolecular C-H functionalization by donor/acceptor rhodium carbenes, Chem. Soc. Rev 2011, 40, 1857–1869. [PubMed: 21359404]

(45). Dydio P, Key HM, Nazarenko A, Rha JY, Seyedkazemi V, Clark DS, and Hartwig JF An artificial metalloenzyme with the kinetics of native enzymes, Science 2016, 354, 102–106. [PubMed: 27846500]

(46). Sreenilayam G, Moore EJ, Steck V, and Fasan R Metal substitution modulates the reactivity and extends the reaction scope of myoglobin carbene transfer catalysts, Adv. Synth. Cat 2017, 359, 2076–2089.

(47). Gu Y, Natoli SN, Liu Z, Clark DS, and Hartwig JF Site-Selective Functionalization of (sp(3))C-H Bonds Catalyzed by Artificial Metalloenzymes Containing an Iridium-Porphyrin Cofactor, Angew. Chem. Int. Ed 2019, 58, 13954–13960.

(48). Vargas DA, Tinoco A, Tyagi V, and Fasan R Myoglobin-Catalyzed C-H Functionalization of Unprotected Indoles, Angew. Chem. Int. Ed 2018, 57, 9911–9915.

(49). Zhang RK, Chen K, Huang X, Wohlschlager L, Renata H, and Arnold FH Enzymatic assembly of carbon-carbon bonds via iron-catalysed sp(3) C-H functionalization, Nature 2019, 565, 67–72. [PubMed: 30568304]

(50). Zhou AZ, Chen K, and Arnold FH Enzymatic Lactone-Carbene C-H Insertion to Build Contiguous Chiral Centers, ACS Catal. 2020, 10, 5393–5398.

(51). Zhang JE, Huang XY, Zhang RJK, and Arnold FH Enantiodivergent alpha-Amino C-H Fluoroalkylation Catalyzed by Engineered Cytochrome P450s, J. Am. Chem. Soc 2019, 141, 9798–9802. [PubMed: 31187993]

(52). Nam D, Steck V, Potenzino RJ, and Fasan R A Diverse Library of Chiral Cyclopropane Scaffolds via Chemoenzymatic Assembly and Diversification of Cyclopropyl Ketones, J. Am. Chem. Soc 2021, 143, 2221–2231. [PubMed: 33497207]

(53). Bell SG, Harford-Cross CF, and Wong LL Engineering the CYP101 system for in vivo oxidation of unnatural substrates, Protein Eng. 2001, 14, 797–802. [PubMed: 11739899]

(54). McLean MA, Maves SA, Weiss KE, Krepich S, and Sligar SG Characterization of a cytochrome P450 from the acidothermophilic archaea Sulfolobus solfataricus, Biochem. Bioph. Res. Commun 1998, 252, 166–172.

(55). Rylott EL, Jackson RG, Edwards J, Womack GL, Seth-Smith HMB, Rathbone DA, Strand SE, and Bruce NC An Explosive-Degrading Cytochrome P450 Activity and its Targeted Application for the Phytoremediation of RDX, Nat. Biotechnol 2006, 24, 216–219. [PubMed: 16429147]

(56). Steck V, Kolev JN, Ren X, and Fasan R Mechanism-Guided Design and Discovery of Efficient P450-Derived C—H Amination Biocatalysts, J. Am. Chem. Soc 2020, 142, 10343–10357. [PubMed: 32407077]

(57). Doyle MP, Duffy R, Ratnikov M, and Zhou L Catalytic Carbene Insertion into C-H Bonds, Chem. Rev 2010, 110, 704–724. [PubMed: 19785457]

(58). Coelho PS, Wang ZJ, Ener ME, Baril SA, Kannan A, Arnold FH, and Brustad EM A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins in vivo, Nat. Chem. Biol 2013, 9, 485–487. [PubMed: 23792734]

(59). Wei Y, Tinoco A, Steck V, Fasan R, and 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. [PubMed: 29268614]

(60). Moore EJ, and Fasan R Effect of proximal ligand substitutions on the carbene and nitrene transferase activity of myoglobin, Tetrahedron 2019, 75, 2357–2363. [PubMed: 31133770]

(61). Park SY, Yamane K, Adachi S, Shiro Y, Weiss KE, Maves SA, and Sligar SG Thermophilic cytochrome P450 (CYP119) from Sulfolobus solfataricus: high resolution structure and functional properties, J. Inorg. Biochem 2002, 91, 491–501. [PubMed: 12237217]

(62). Key HM, Dydio P, Liu ZN, Rha JYE, Nazarenko A, Seyedkazemi V, Cark DS, and Hartwig JF Beyond Iron: Iridium-Containing P450 Enzymes for Selective Cyclopropanations of Structurally Diverse Alkenes, ACS Centr. Sci 2017, 3, 302–308.

(63). Gerry CJ, Wawer MJ, Clemons PA, and Schreiber SL DNA Barcoding a Complete Matrix of Stereoisomeric Small Molecules, J. Am. Chem. Soc 2019, 141, 10225–10235. [PubMed: 31184885]

(64). Bassi G, Favalli N, Vuk M, Catalano M, Martinelli A, Trenner A, Porro A, Yang S, Tham CL, Moroglu M, Yue WW, Conway SJ, Vogt PK, Sartori AA, Scheuermann J, and Neri D A Single-Stranded DNA-Encoded Chemical Library Based on a Stereoisomeric Scaffold Enables Ligand Discovery by Modular Assembly of Building Blocks, Advanced Science 2020, 7, 2001970. [PubMed: 33240760]

(65). Fasan R, Chen MM, Crook NC, and Arnold FH Engineered Alkane-Hydroxylating Cytochrome P450(BM3) Exhibiting Nativelike Catalytic Properties, Angew. Chem. Int. Ed 2007, 46, 8414–8418.

(66). Ilie A, Harms K, and Reetz MT P450-Catalyzed Regio- and Stereoselective Oxidative Hydroxylation of 6-Iodotetralone: Preparative-Scale Synthesis of a Key Intermediate for Pd-Catalyzed Transformations, J. Org. Chem 2018, 83, 7504–7508. [PubMed: 29313346]

(67). Kolev JN, Zaengle JM, Ravikumar R, and Fasan R Enhancing the Efficiency and Regioselectivity of P450 Oxidation Catalysts by Unnatural Amino Acid Mutagenesis, Chembiochem 2014, 15, 1001–1010. [PubMed: 24692265]

(68). Kuehnel K, Maurer SC, Galeyeva Y, Frey W, Laschat S, and Urlacher VB Hydroxylation of Dodecanoic Acid and (2R,4R,6R,8R)-Tetramethyldecanol on a Preparative Scale Using an NADH-Dependent CYP102A1 Mutant, Adv. Synth. Cat 2007, 349, 1451–1461.

(69). Ren XK, O'Hanlon JA, Morris M, Robertson J, and Wong LL Synthesis of Imidazolidin-4-ones via a Cytochrome P450-Catalyzed Intramolecular C—H Amination, ACS Catal. 2016, 6, 6833–6837.

(70). Roiban GD, Agudo R, and Reetz MT Cytochrome P450 Catalyzed Oxidative Hydroxylation of Achiral Organic Compounds with Simultaneous Creation of Two Chirality Centers in a Single C—H Activation Step, Angew. Chem. Int. Ed 2014, 53, 8659–8663.

(71). Sarkar MR, Dasgupta S, Pyke SM, and Bell SG Selective Biocatalytic Hydroxylation of Unactivated Methylene C—H Bonds in Cyclic Alkyl Substrates, Chem. Commun 2019, 55, 5029–5032.

(72). Shoji O, Yanagisawa S, Stanfield JK, Suzuki K, Cong ZQ, Sugimoto H, Shiro Y, and Watanabe Y Direct Hydroxylation of Benzene to Phenol by Cytochrome P450BM3 Triggered by Amino Acid Derivatives, Angew. Chem. Int. Ed 2017, 56, 10324–10329.

(73). Venkataraman H, Verkade-Vreeker MCA, Capoferri L, Geerke DP, Vermeulen NPE, and Commandeur JNM Application of Engineered Cytochrome P450 Mutants as Biocatalysts for the Synthesis of Benzylic and Aromatic Metabolites of Fenamic Acid NSAIDs, Bioorg. Med. Chem 2014, 22, 5613–5620. [PubMed: 24999003]

(74). Simon RC, Grischek B, Zepeck F, Steinreiber A, Belaj F, and Kroutil W Regio-and stereoselective monoamination of diketones without protecting groups, Angew. Chem. Int. Ed 2012, 51, 6713–6716.

(75). Pavlidis IV, Weiss MS, Genz M, Spurr P, Hanlon SP, Wirz B, Iding H, and Bornscheuer UT Identification of (S)-selective transaminases for the asymmetric synthesis of bulky chiral amines, Nat. Chem 2016, 8, 1076–1082. [PubMed: 27768108]

(76). Aleku GA, France SP, Man H, Mangas-Sanchez J, Montgomery SL, Sharma M, Leipold F, Hussain S, Grogan G, and Turner NJ A reductive aminase from Aspergillus oryzae, Nat. Chem 2017, 9, 961–969. [PubMed: 28937665]

(77). Hussain S, Leipold F, Man H, Wells E, France SP, Mulholland KR, Grogan G, and Turner NJ An (R)-Imine Reductase Biocatalyst for the Asymmetric Reduction of Cyclic Imines, Chemcatchem 2015, 7, 579–583. [PubMed: 27547270]

(78). Zawodny W, Montgomery SL, Marshall JR, Finnigan JD, Turner NJ, and Clayden J Chemoenzymatic Synthesis of Substituted Azepanes by Sequential Biocatalytic Reduction and Organolithium-Mediated Rearrangement, J. Am. Chem. Soc 2018, 140, 17872–17877. [PubMed: 30521324]

(79). Weiner MP, Costa GL, Schoettlin W, Cline J, Mathur E, and Bauer JC Site-Directed Mutagenesis of Double-Stranded DNA by the Polymerase Chain-Reaction, Gene 1994, 151, 119–123. [PubMed: 7828859]

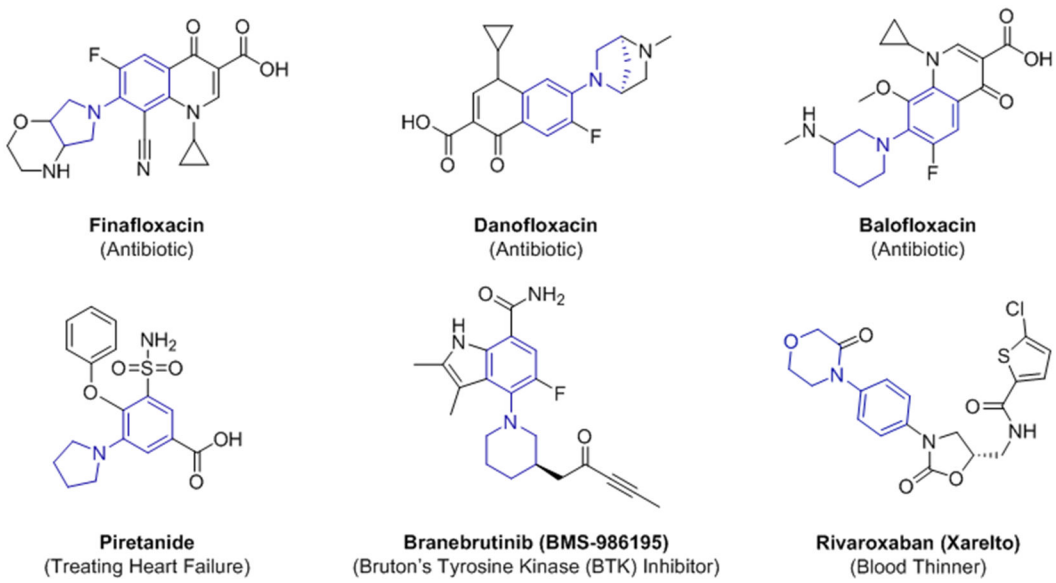


Figure 1.

Representative drugs and bioactive natural products containing N-aryl pyrrolidines, piperidines, and other saturated *N*-heterocycles.

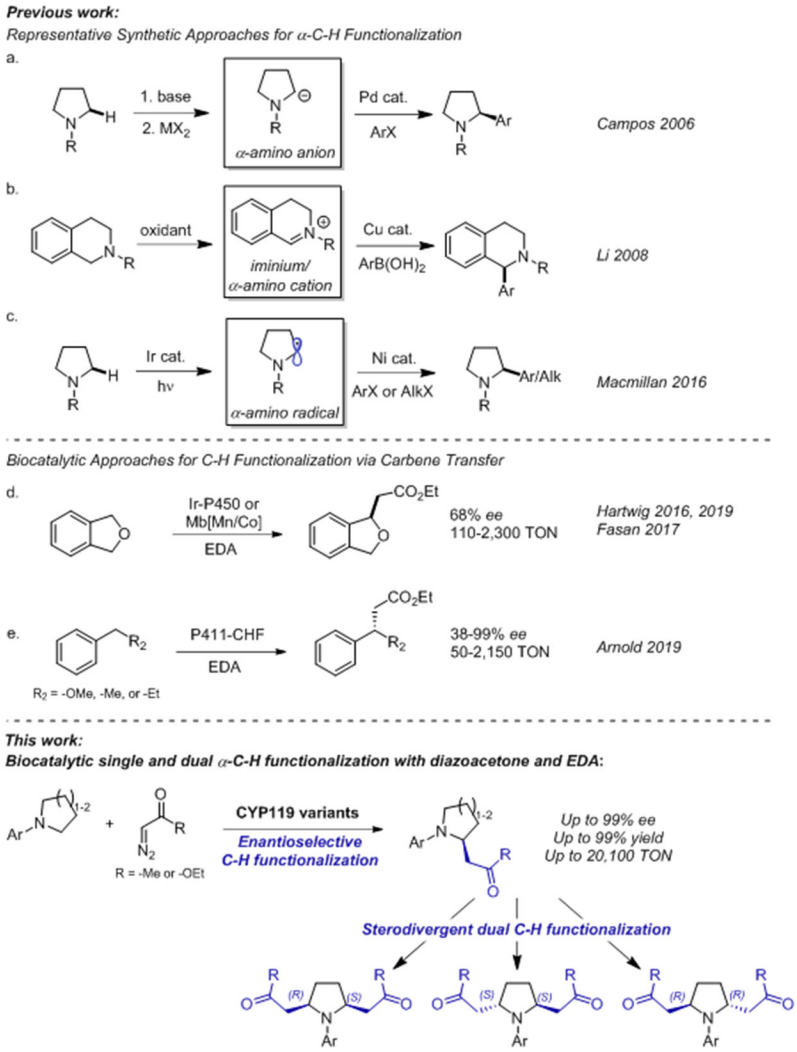


Figure 2.

Representative chemocatalytic methods for α-C—H functionalization (a-c). Previous (d-e) and present biocatalytic strategy for C—H functionalization via enzyme-catalyzed carbene transfer.

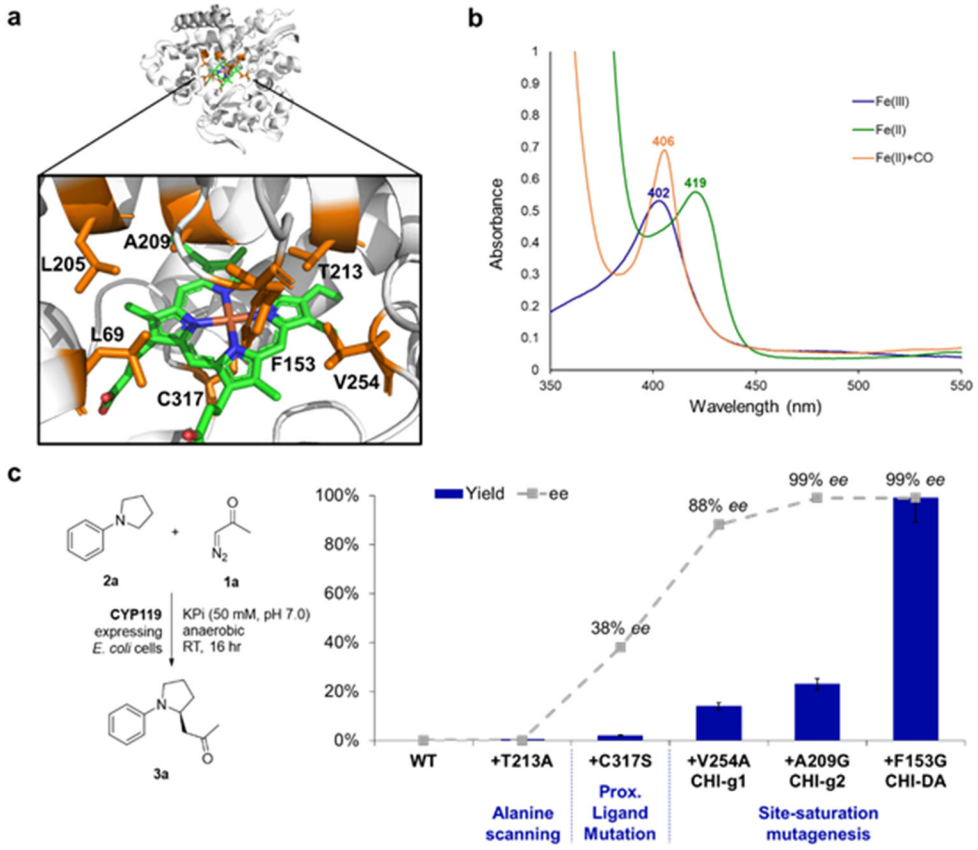


Figure 3. Structure, spectral properties, and directed evolution of CYP119 catalysts for enantioselective C—H functionalization of *N*-phenylpyrrolidine with diazoacetone.

(**a**) X-ray crystal structure of CYP119 from *Sulfolobus solfataricus* (PDB 1IO7[61]). Amino acid residues targeted for mutagenesis are highlighted in orange, and the heme cofactor is shown in green. (**b**) UV-Vis absorption spectra for CYP119(T213A, C317S) in its ferric (Fe(III), blue), ferrous state (Fe(II), green), and CO-bound form (Fe(II)+CO; orange). (**c**) Directed evolution of CYP119 catalysts for enantioselective C—H functionalization of *N*-phenylpyrrolidine (**2a**) with diazoacetone (**1a**). Yields and % *ee* as determined under standard reaction conditions with diazoacetone (Table 1).

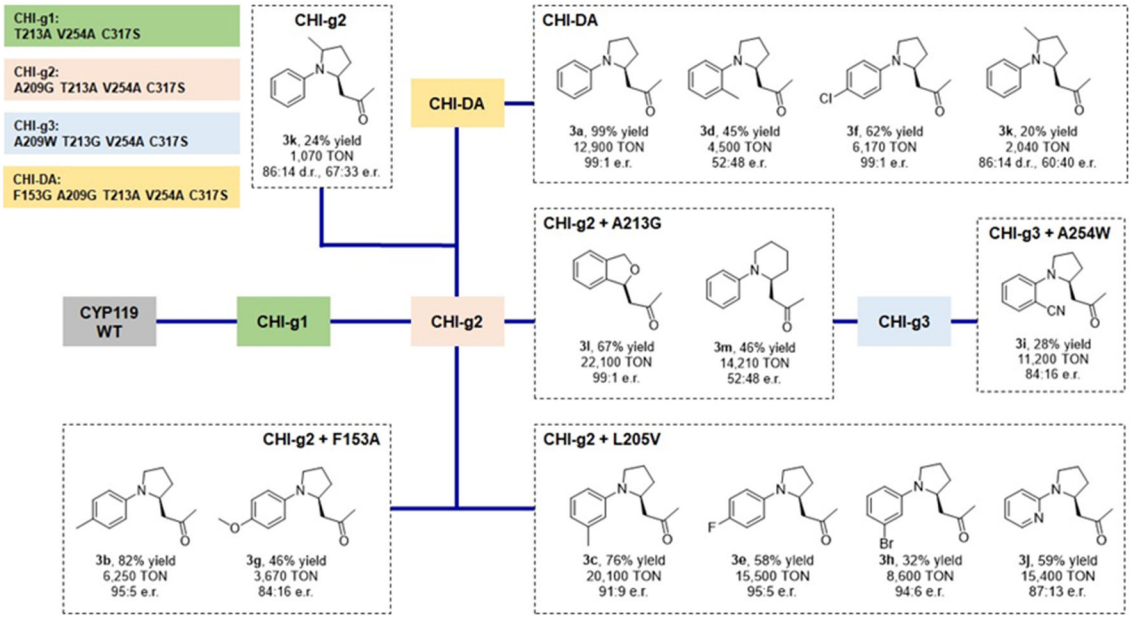


Figure 4. Activity and selectivity of CYP119 based biocatalysts for α-C—H functionalization of *N*-aryl-pyrrolidines with diazoacetone.

Yields, TON, and enantioselectivity as determined from whole cell reactions under standard reaction conditions with diazoacetone, as described in Table 1. The lines indicate the relationship among the enzyme variants along the evolutionary lineage.

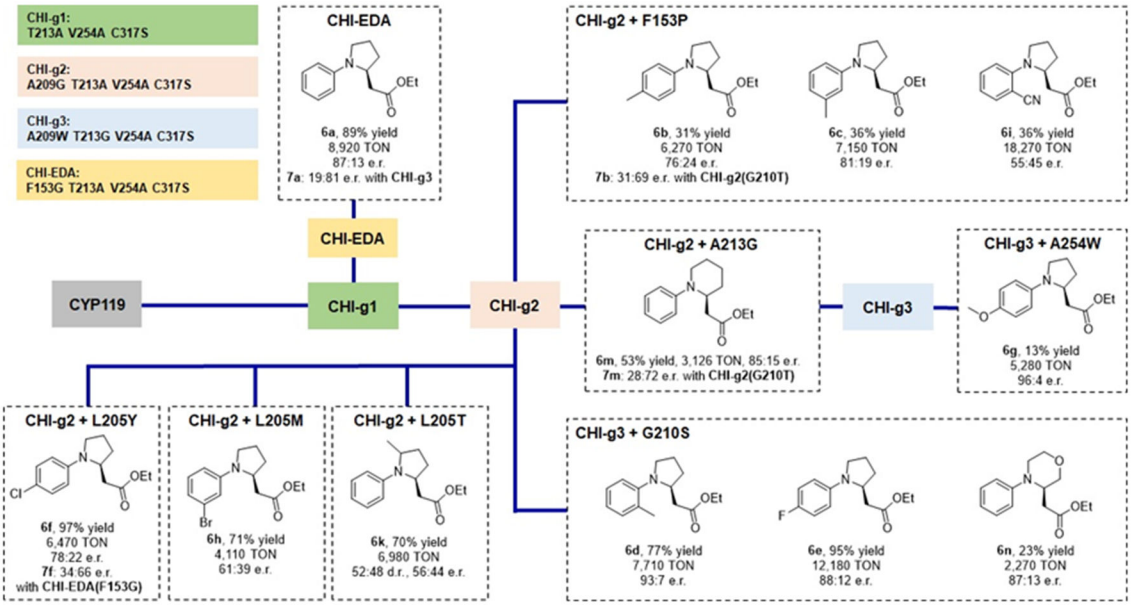


Figure 5.

Activity and selectivity of CYP119 biocatalysts for α-C—H functionalization of cyclic amines with EDA. Yields, TON, and enantioselectivity were determined from whole cell reactions under standard reaction conditions with EDA as described in Table 2. The lines indicate the relationship among the enzyme variants along the evolutionary lineage.

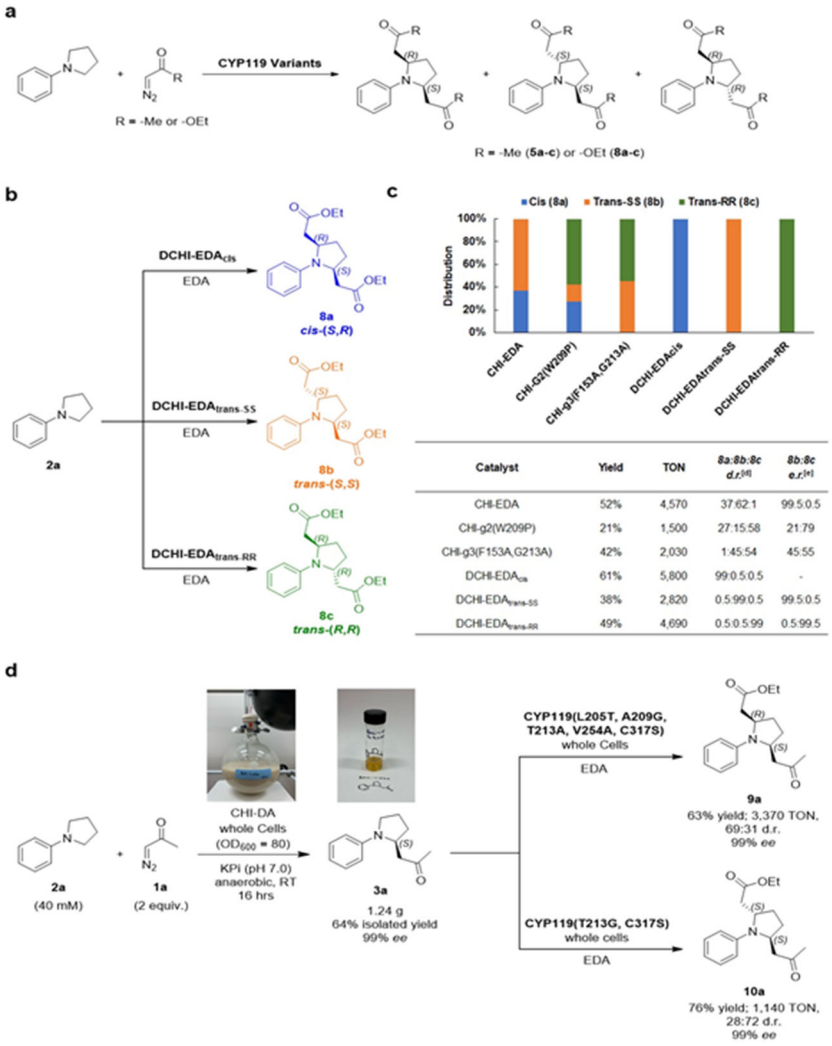


Figure 6. Stereodivergent dual α-C-H functionalization.

(**a**) General scheme for dual C—H carbene insertion with diazoacetone and ethyl diazoacetate in *N*-phenylpyrrolidine. (**b-c**) One-pot dual C—H functionalization of **2a** with EDA using CYP119 with stereodivergent selectivity. Whole cell reactions were carried under standard reaction conditions with EDA as in Table 2, but using 40 mM EDA. The graph reports the relative distribution of products **8a-c** for the three DCHI-EDA variants along with three other representative CYP119 variants. Assay yields (GC) and TON values reported in the table correspond to the double C—H functionalization product. (**d**) Tandem dual C—H functionalization of **2a** with diazoacetone and EDA. The EDA reaction was carried out using standard reaction conditions as described in Table 2. Diastereomeric and enantiomeric ratios were determined by chiral GC and SFC.

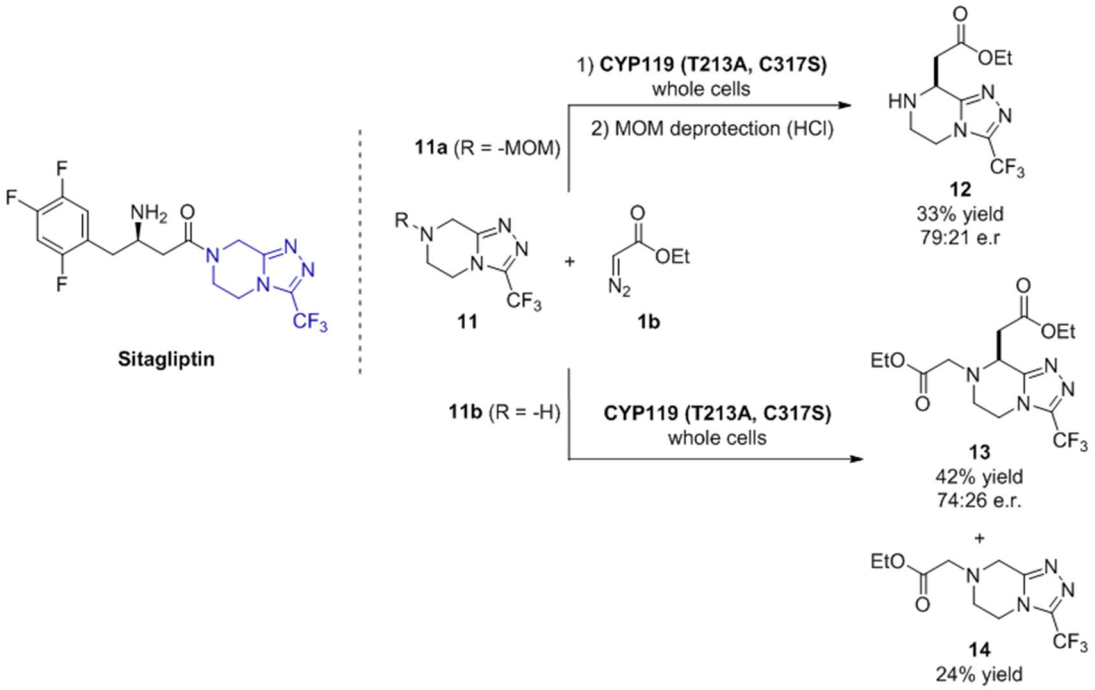
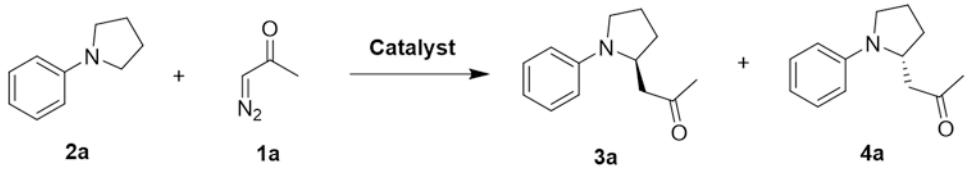


Figure 7.

Late-stage enzymatic C—H functionalization of sitagliptin core with engineered CYP119 catalysts.

Table 1.

Intermolecular C—H functionalization of *N*-phenylpyrrolidine (2a) with diazoacetone using hemoproteins and variants thereof.[a]



2a 1a 3a 4a

Entry	Catalyst	Yield[b]	TON[c]	e.r.[d] (3a:4a)
1	Hemin	0%	0	-
2	Mb (WT)	0%	0	-
3	Mb(H64G, V68A)	0%	0	-
4	P411-CHF	0%	0	-
5	CYP119 (WT)	0%	0	-
6	CYP119 (T213A)	0.4%	35	n.d.
7	CYP119(T213A, C317S)	2%	170	69:31
8	CYP119 (T213A, V254A, C317S) **CHI-g1**	14%	200	94:6
9	CYP119 (A209G, T213A, V254A, C317S) **CHI-g2**	23%	3,100	99.5:0.5
10	CYP119 (F153G, A209G, T213A, V254A, C317S) **CHI-DA**	99%	12,900	99.5:0.5
11[e]	CYP119 (F153G, A209G, T213A, V254A, C317S) **CHI-DA**	99%	500	99.5:0.5
12[f]	CYP119 (F153G, A209G, T213A, V254A, C317S) **CHI-DA**	53%	20,350	99.5:0.5

[a] Standard reaction conditions: protein expressing C41(DE3) *E. coli* cells, OD_{600} = 40, 10 mM **2a**, 20 mM diazoacetone (**1a**), in KPi buffer (50 mM, pH 7), room temperature, 16 hours, in anaerobic chamber.

[b] Assay yields as determined by GC using calibration curves with isolated product.

[c] TON as calculated based on the protein concentration measured from cell lysate.

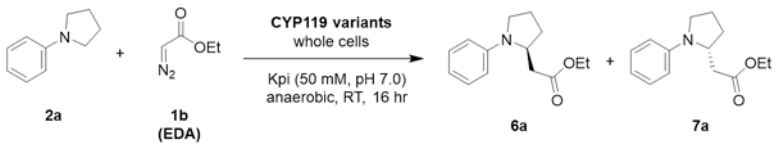
[d] Enantiomeric ratio (e.r.) for **3a**:**4a** as determined by chiral SFC.

[e] Using 20 μM purified protein and 10 mM $Na_2S_2O_4$.

[f] OD_{600} = 10. N.d. = not determined.

Table 2.

Activity and selectivity of representative engineered CYP119 variants for the intermolecular C—H functionalization of N-phenylpyrrolidine with EDA.[a]



Entry	Enzyme Variant	Yield[b]	TON[c]	e.r. (6a:7a)
1	CYP119 (WT)	3%	76	n.d.
2	CYP119(C317S)	5%	110	n.d.
3	CYP119(T213A	6%	135	n.d.
4	CYP119(T213A, C317S	18%	812	71:29
5	CYP119(T213A, V254A, C317S) (**CHI-g1**)	99%	4,400	79:21
6	CYP119(F153G, T213A, V254A, C317S) (**CHI-EDA**)	89% (64%)[d]	8,920	87:13
7[e]	CYP119(F153G, T213A, V254A, C317S)	99%	500	87:13
8[f]	CYP119(F153G, T213A, V254A, C317S)	87%	8,710	87:13
9[f,g]	CYP119(F153G, T213A, V254A, C317S)	51%	5,090	87:13
10	CYP119(A209W, T213G, V254A, C317S	71%	6,750	19:81

[a] Standard reaction conditions: CYP119 expressing C41(DE3) *E. coli* cells, OD_{600} = 40, 10 mM **1a**, 20 mM EDA, in KPi buffer (50 mM, pH 7), room temperature, 16 hours, in anaerobic chamber.

[b] Assay yields as determined by GC using calibration curves generated with isolated product.

[c] TON as calculated based on the protein concentration of the cell lysate.

[d] Isolated yield.

[e] Using 20 μM purified protein and 10 mM $Na_2S_2O_4$.

[f] Reaction time: 2 hours.

[g] Under aerobic conditions. N.d. = not determined.