JACS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Caltech Library

Subscriber access provided by Caltech Library

Article

Directed Evolution of a Cytochrome P450 Carbene Transferase for Selective Functionalization of Cyclic Compounds

Oliver F. Brandenberg, Kai Chen, and Frances H. Arnold

J. Am. Chem. Soc., **Just Accepted Manuscript** • DOI: 10.1021/jacs.9b02931 • Publication Date (Web): 09 May 2019

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Directed Evolution of a Cytochrome P450 Carbene Transferase for Selective Functionalization of Cyclic Compounds

Oliver F. Brandenberg[1,†], Kai Chen[1], Frances H. Arnold[1*]

[1] Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA.

[*] Corresponding author

KEYWORDS Biocatalysis, Cytochrome P450, Carbene Transfer, Indole Alkylation, Cyclopropanation.

ABSTRACT: Transfers of carbene moieties to heterocycles or cyclic alkenes to obtain C(sp^2)–H alkylation or cyclopropane products are valuable transformations for synthesis of pharmacophores and chemical building blocks. Through their readily tunable active site geometries, hemoprotein "carbene transferases" could provide an alternative to traditional transition metal catalysts by enabling heterocycle functionalizations with high chemo-, regio- and stereocontrol. However, carbene transferases accepting heterocyclic substrates are scarce; the few enzymes capable of heterocycle or cyclic internal alkene functionalization described to date are characterized by low turnovers or depend on artificially introduced, costly iridium–porphyrin cofactors. We addressed this challenge by evolving a cytochrome P450 for highly efficient carbene transfer to indoles, pyrroles, and cyclic alkenes. We first developed a spectrophotometric high-throughput screening assay based on 1-methylindole C_3-alkylation that enabled rapid analysis of thousands of P450 variants and comprehensive directed evolution via random and targeted mutagenesis. This effort yielded a P450 variant with 11 amino acid substitutions and a large deletion of the non-catalytic P450 reductase domain, which chemoselectively C_3-alkylates indoles with up to 470 turnovers per minute and 18,000 total turnovers. We subsequently used this optimized alkylation variant for parallel evolution towards more challenging heterocycle carbene functionalizations, including C_2/C_3 regioselective pyrrole alkylation, enantioselective indole alkylation with ethyl-2-diazopropanoate, and cyclic internal alkene cyclopropanation. The resulting set of efficient biocatalysts showcases the tunability of hemoproteins for highly selective functionalization of cyclic targets and the power of directed evolution to enhance the scope of new-to-nature enzyme catalysts.

Introduction

Cyclic compounds, particularly heterocycles, are key components in biologically active compounds, from the building blocks of life (*e.g.* nucleic acids, amino acids and carbohydrates) to natural products with pharmacologically relevant activities (*e.g.* vitamins, alkaloids and antibiotics). Accordingly, heterocycles have been playing an increasingly significant role in the modern repertoire of medicinal chemistry,[1] with molecules containing at least one heterocyclic fragment comprising nearly 60 % of FDA-approved pharmaceuticals.[2] Because heterocycles are powerful handles to regulate polarity, lipophilicity and hydrogen bonding properties, extensive efforts are directed at tuning these structures to manipulate the pharmacological, pharmacokinetic and toxicological features of drug candidates.[3–5] Therefore, development of selective and efficient methods for functionalization of heterocycles will streamline synthesis of medicinal compounds and pharmacophores.

One major way to derivatize heterocyclic fragments is to directly functionalize C–H bonds with new chemical entities.[6] Alternatively, one can derivatize ring structures by functionalizing existing C–C bonds[7]. Nature employs different enzymatic machineries to accomplish these two processes. For instance, methyltransferases install methyl moieties on the C–H bonds of DNA bases;[8] flavin-dependent halogenases are responsible for installing halogens on heterocycle-based natural products, such as indoles, pyrroles and azoles;[9] and (mono)oxygenases like cytochrome P450s or flavoenzymes are involved in oxidation of heterocyclic substrates to induce or lose aromaticity.[10–12]

To expand nature's capacity for functionalization of (hetero)cyclic molecules, we wished to use recently developed "carbene transferase" enzymes to functionalize carbon–hydrogen and carbon–carbon double bonds in heterocycles. Carbene transferases are engineered heme-dependent proteins capable of transferring carbene moieties to different target compounds.[13] Reactivities developed in recent years include carbene transfer to olefins,[14–16] alkynes,[17] and X–H bonds (X = C, N, S, Si, B)[18–23] to generate chemical structures often not found in nature. These enzymatic transformations grant biocatalytic access to important pharmaceuticals including levomilnacipran, ticagrelor, tasimelteon, tranylcypromine and others.[24–26]

However, enzymatic carbene transfer for heterocycle derivatization faces several challenges. Carbene transfer has been achieved mostly for sterically unhindered functional groups, such as terminal olefins. Thus it remains to be determined whether sterically more demanding, internal carbon–carbon double bonds in heterocycles can be efficiently functionalized. A second challenge for enzymatic carbene transfer to heterocycles is selectivity. Synthetic chemists have used transition metal catalysts based on Rh, Ru, Pd, Cu and others to selectively transfer carbenes to heteroarenes, such as indoles and pyrroles,

yielding either cyclopropanes or C(sp^2)–H alkylation products.[27–32] Mimicking this activity would require enzymes to have exquisite control over chemo-, regio- and stereo-selectivity. Here, we set out to direct the evolution of enzyme catalysts with high activity and selectivity in heterocycle carbene functionalization reactions.

Towards this goal, we used a two-stage engineering approach. We first evolved a cytochrome P450 variant for C_3 alkylation of indoles, using a high-throughput spectrophotometric screen that allowed us to evaluate several thousand enzyme variants per generation. This enabled a mutagenesis strategy targeting the entire P450 protein sequence, which delivered a variant containing 11 mutations as well as a truncation of a large part of the non-catalytic reductase domain and which chemoselectively and efficiently alkylates indoles. Hypothesizing that this enzyme, evolved for alkylation of a heterocyclic substrate, would provide a launch pad for further exploration of challenging heterocycle functionalizations, in the second stage we performed parallel directed evolution to engineer a set of P450 catalysts capable of regio-divergent pyrrole alkylations, enantioselective indole alkylations with α-disubstituted carbenes, and stereoselective cyclic alkene cyclopropanations.

Results and Discussion

Leveraging a high-throughput screening assay to evolve a cytochrome P450 for indole alkylation

We began by testing a range of hemoproteins for carbene transfer (using ethyl diazoacetate (EDA) as carbene precursor) to a selection of heterocyclic substrates (Supplementary Table 1). Cytochrome P450-BM3 variant P411-CIS, a serine-ligated variant previously evolved for styrene cyclopropanation,[14] stood out as highly active for C_3-alkylation of 1-methylindole, catalyzing this reaction with 14% yield and 690 TTN. To our delight, we found that the resulting alkylated product **3a** showed a pronounced red-shift in UV absorbance compared to 1-methylindole. This difference in absorbance properties allowed us to develop a simple UV-Vis spectrophotometric screening assay, in which the catalytic activities of enzyme variants for the alkylation reaction (performed in whole *Escherichia coli* (*E. coli*) cells) could be measured by UV absorbance at 305 nm in a 96-well plate reader (Figure 1A, Supplementary Figure 1).

P411-CIS is a 1048-residue protein consisting of a heme domain containing the active site (residues 1-470) and a reductase domain that shuttles electrons from NADPH to the heme domain and is essential for P450-BM3's native oxo-transfer chemistry (residues 471-1048, subdivided into FMN and FAD domains; Figure 1B).[35] Having a screen that enables analysis of several thousand enzyme variants, we introduced mutations throughout the heme and reductase domains using both random mutagenesis and site-saturation mutagenesis and screened the resulting libraries for improved 1-methylindole alkylation activity (Supplementary Table 2).

After six rounds of directed evolution for formation of **3a**, of which the last four rounds were conducted under aerobic conditions, we obtained a P411 variant we refer to as P411-HF (Heterocycle Functionalization) (Figure 1C). This variant contains eight mutations in the heme domain, three in the FMN domain and a stop codon at the end of the FMN domain (Q674*), which deletes the entire FAD domain (Figure 1B, Supplementary Figure 2). Of these mutations, five were introduced by site-saturation mutagenesis (V87T, H92F, L181G, I263M, T438C),

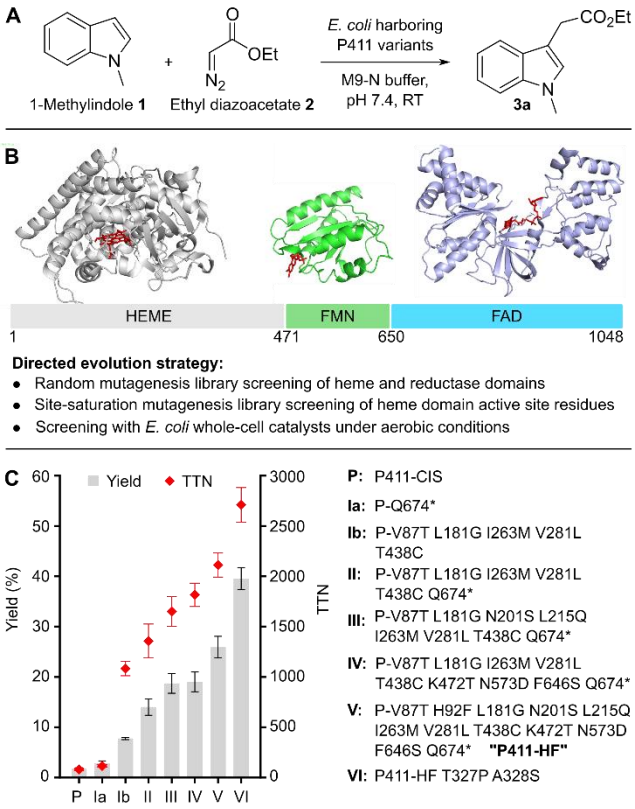


Figure 1: Evolution of P411-CIS for 1-methylindole C_3 alkylation. (A) Screening reaction used to evolve P411 variants for higher indole alkylation activity. (B) Domain structure of P450-BM3. As a crystal structure of the entire P450-BM3 protein is lacking, individual domain structures with approximate domain boundaries are shown (pdb entries: heme domain 4h23[14]; FMN domain 1bvy[33]; FAD domain 4dqk[34]); co-factors are shown in red. (C) Evolutionary lineage from P411-CIS to P411-HF T327P A328S. Data were obtained from two independent assays with six replicates per variant, run in 96-well plates under aerobic conditions with 10 mM **1** and 15 mM **2** at room temperature (RT) for 16h.

and seven (N201S, L215Q, V281L, K472T, N573D, F646S, Q674*) were introduced by random mutagenesis. A final round of double site-saturation mutagenesis library screening on P411-HF revealed mutations T327P and A328S. This enzyme delivered the alkylated product with up to 2700 TTN in aerobic 96-well plate screening reactions (Figure 1C).

P411-HF: A catalyst for chemoselective alkylation of unprotected indoles

P411-HF is a highly active alkylation enzyme: analytical whole-cell reactions with P411-HF under anaerobic conditions with 30 mM 1-methylindole and 1.5 equivalents of EDA delivered alkylation product **3a** in 95% yield with almost 12,000 TTN (Figure 2A, entry 2). Under catalyst-limiting conditions, P411-HF delivered more than 18,000 TTN (Figure 2A, entry 4). Under aerobic conditions, variant P411-HF T327P A328S proved superior to P411-HF and delivered the alkylated product in 74% yield and with 1670 TTN (Figure 2A, entry 6; P411-HF delivered **3a** in 46% yield under these conditions). Thus, even with four rounds of directed evolution under aerobic conditions,

P411-HF is about an order of magnitude more effective under anaerobic conditions. Furthermore, P411-HF delivers almost two orders of magnitude higher TTN than recently reported variants of the hemoproteins Mb or YfeX catalyzing indole alkylations (170 to 240 TTN),[36,37] and up to three orders of magnitude higher TTN than Rh, Cu, Pd or Fe-based transition metal catalysts reported for the same type of reaction.[30,31,38–40]

To investigate catalyst lifetime, we tested sequential additions of 10 mM **1** and **2** substrate equivalents to whole-cell reactions in 30-min intervals. This did not increase product formation, which indicates that the enzyme is inactivated during the reaction (Supplementary Figure 3A). Furthermore, a 10-minute EDA pre-incubation of P411-HF whole-cell catalyst resulted in a 26% loss in catalytic activity (Supplementary Figure 3B). Presumably, alkylation of enzyme residues and/or the porphyrin cofactor limits enzyme lifetime, which may possibly be addressed by further directed evolution.[41]

P411-HF anaerobic whole-cell reactions reached completion after 2 h. The high initial turnover frequency (TOF) of 470 turnovers min^{-1} represents a 35-fold increase over the parent variant P411-CIS (Figure 2B). Under aerobic conditions, P411-HF T327P A328S showed an initial rate of 21 turnovers min^{-1}, compared to 3 turnovers min^{-1} for P411-CIS (Figure 2C). The initial rates of P411-HF compare well to previously disclosed Mb variants, which exhibited ~10 turnovers min^{-1} in the alkylation of indole with EDA under anaerobic conditions.[36]

The most striking mutation found during the directed evolution was the truncation of the FAD domain; introduction of the stop codon by Q674* reduces the size of the protein catalyst by 35% (from 1048 to 673 residues) and increases reaction yield by ca. 1.7-fold, driven primarily by an increase in protein expression level (ca. 1.6-fold) with a slight increase in TTN (Figure 1B, Supplementary Figure 4A). Comparable positive effects of deleting the FAD domain have subsequently been observed for other P411-catalyzed carbene transfer reactions: FAD truncation discovered in the course of this study proved beneficial for carbene C(sp^3)–H insertion[18] as well as cyclopropanation reactions (Supplementary Figure 4B).[42] Notably, FAD domain deletion significantly destabilized the protein, as whole-cell alkylation assays performed with a protein denaturant pre-treatment revealed up to 33% lower activity of P411-CIS Q674* compared to P411-CIS (Supplementary Figure 4C).[43,44] Thus, FAD domain deletion exerts a destabilizing effect, but contributes to higher P411 expression levels and carbene transfer activity. Mounting evidence suggests that native P450-BM3 adopts a homodimeric quaternary structure in solution via FAD domain interactions, and that the dimeric form is required for oxo-transfer activity.[34,45–48] The data presented here suggest that for non-natural carbene-transfer reactions a different, possibly monomeric, catalyst structure may be suitable or even beneficial. Attempts to obtain crystal structures of the FAD-truncated P411-HF variant have unfortunately not been successful.

Next, we compared 1-methylindole alkylation activities of P411-CIS and P411-HF as purified proteins and with different reducing agents (Supplementary Figure 5). We found that purified P411-CIS catalyzes 1-methylindole alkylation with 165 TTN with sodium dithionite as reductant, 135 TTN with NADPH as reductant, but only 8 TTN in the absence of reductant. To our surprise, P411-HF showed 1890 TTN with NADPH (corresponding to 94% yield), 1750 TTN with sodium dithionite, and an almost equally high 1670 TTN without any reductant present. Thus, P411-HF maintains a catalytically competent

A

Entry	Catalyst	Conditions	**1** (mM)	**2** (mM)	Yield **3a** (%)	TTN
1	P411-CIS	Anaerobic OD_{600}=30	30	45	13	1660
2	P411-HF	Anaerobic OD_{600}=30	30	45	95	11900
3	P411-CIS	Anaerobic OD_{600}=10	30	45	5	2010
4	P411-HF	Anaerobic OD_{600}=10	30	45	51	18700
5	P411-CIS	Aerobic OD_{600}=30	10	15	4	140
6	P411-HF T327P A328S	Aerobic OD_{600}=30	10	15	74	1670

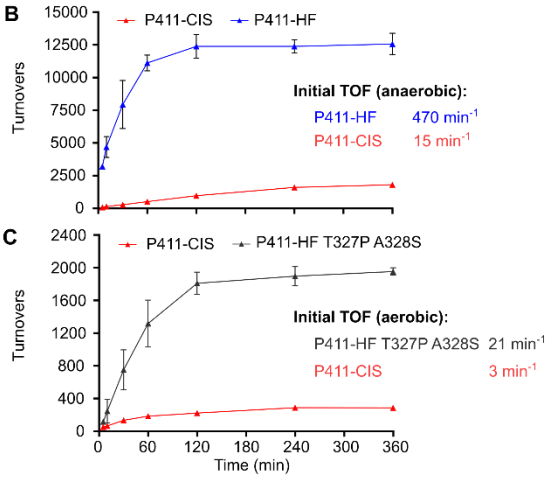


Figure 2: Characterization of P411-HF. (A) Comparison of P411-CIS and P411-HF in analytical (400 µL) biocatalytic whole-cell reactions for formation of **3a** under aerobic or anaerobic conditions and different substrate and catalyst loadings. (B,C) Kinetics of whole cell-reactions for formation of **3a** under anaerobic (B) and aerobic (C) conditions. Reactions were run with cells at OD_{600}=30 with 30 mM **1** and 45 mM **2** (anaerobic, B), or 10 mM **1** and 15 mM **2** (aerobic, C). Initial turnover frequencies were calculated from samples analyzed 10 minutes after starting the reaction. Data in (A) to (C) were obtained from two to three independent assays with two replicates per variant.

state of the heme cofactor even in the absence of reductant, in stark contrast to P411-CIS. While native P450 oxo-transfer chemistry requires a stoichiometric supply of electrons, P411 carbene transfer catalysis functions without an electron supply as long as the iron cofactor is in the ferrous state.[14] Therefore, truncating the reductase domain can be expected to be less consequential for carbene transfer catalysis. The extent to which FAD domain truncation contributes to maintaining the enzyme in the catalytically active ferrous state remains to be determined.

Lastly, we tested the substrate scope of P411-HF and its suitability for synthesis of alkylated indoles. With 1-methylindole, preparative-scale reactions (0.6 mmol) delivered **3a** with 83 % isolated yield (108 mg) (Figure 3). To our delight, we found that P411-HF selectively alkylates unprotected indoles at the C_3 position. No N–H insertion products were observed, and alkylated products were isolated in good yields across a range of substituted, unprotected indoles (Figure 3). We observed a clear reactivity trend, with only trace amount of alkylation product obtained for C_4-substituted indoles, comparatively low yields for C_5-substituted indoles, and good yields for indoles with substi-

tutions at C_6 or C_7. Presumably, substitutions at C_4 or C_5 interfere with productive indole binding in the active site. Also, only low activity was observed for 1,2-dimethylindole; alkylation of this substrate was improved by the M263A mutation, which likely creates more space in the active site above the heme cofactor and enables productive binding of this sterically hindered 1,2-disubstituted indole.

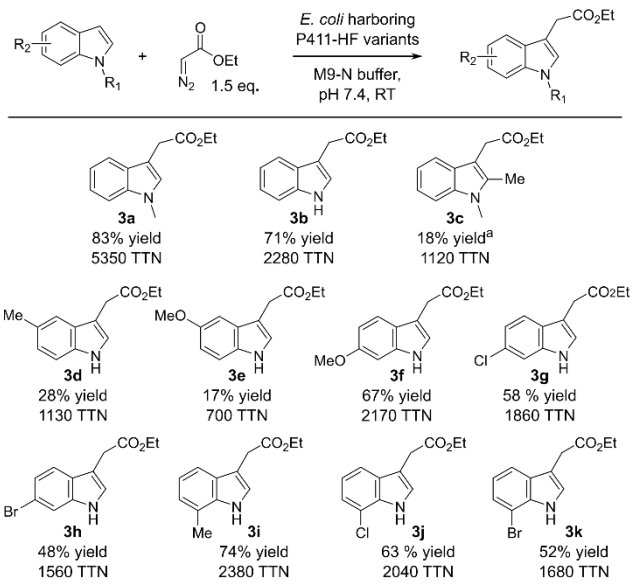


Figure 3: P411-HF indole alkylation substrate scope. Preparative-scale reactions were performed at 0.6 mmol scale (1-methylindole) or 0.3 mmol scale (all other substrates) to yield the C_3-alkylated products. Isolated yields are given; TTN are based on the isolated yield. [a]For 1,2-dimethylindole, reactions were performed with variant P411-HF M263A and HPLC yield and TTN are given.

Utilizing P411-HF as a starting point for additional cyclic alkene functionalization reactions

While evolving for new functions, enzymes may gain activities for additional reactions that were not actively selected for. However, such promiscuous activities may in turn serve as evolutionary starting points to evolve novel functions. [49–51] For example, we found that variant P411-CIS, previously evolved from P450-BM3 for styrene cyclopropanation,[14,15] showed more than 30-fold higher activity for 1-methylindole alkylation than P450-BM3 (Supplementary Table 1). Thus, P411-CIS was chosen as starting point for evolving indole alkylation activity, ultimately yielding variant P411-HF (Figure 4A). In turn, we speculated that P411-HF, engineered to accommodate a heterocyclic structure near the catalytic iron center, may provide a good starting point for further evolution to unlock additional carbene transfer reactions to (hetero)cyclic alkenes. To test this hypothesis, we performed three evolutionary "case studies", focusing on regioselective pyrrole alkylation, enantioselective indole alkylation with an α-disubstituted carbene precursor, and cyclic alkene cyclopropanation.

First, we attempted functionalizing pyrroles at either the C_2 or C_3 position, a considerable challenge for small-molecule catalysts.[27] Transition metal catalysis typically relies on electronic bias between the two positions and predominantly leads to functionalization of C_2–H bonds.[52] Alternatively, bulky substituents on the nitrogen were shown to shift regioselectivity towards the C_3 position.[53] Enzyme catalysis can take advantage of precise substrate orientation within the active site, and thus potentially drive site selectivity independent of electronic or steric bias.[54] Accordingly, we speculated that the P411-HF active site could be tuned by directed evolution to favor alkylation at either the C_2 or C_3 position.

We found that P411-HF catalyzes carbene transfer to 1-methylpyrrole with 180 TTN, showing a 3:1 selectivity for the electronically favored C_2-alkylated product (Figure 4A, Supplementary Figure 6, Supplementary Table 3). Subsequent site-saturation mutagenesis and screening revealed a new variant with mutation M263T, providing a 1:1 ratio of both products and confirming that active site mutations can indeed shift the selectivity. Further evolution yielded two divergent P411 lineages that selectively yield the C_2 or C_3 alkylation products: variant P411-HF M263T G181E A82T makes the C_3-alkylated product with 705 TTN and 9:91 C_2:C_3 ratio, whereas variant P411-HF M263T T87W G181K yields the C_2-alkylated product with 485 TTN and 98:2 C_2:C_3 ratio (Figure 4A). Thus, the P411-HF active site can be sculpted with only three rounds of mutagenesis and screening to exert precise control over pyrrole heterocycle alkylation regioselectivity. The resulting variants may provide starting points for evolving enzymes to selectively alkylate other pyrrole derivatives.

Next, we asked whether P411-HF could serve as a starting point to evolve for enantioselective indole alkylation using an α-disubstituted carbene precursor (*e.g.*, ethyl-2-diazopropanoate, Me-EDA), thereby delivering chiral alkylated indole core **5** found in bioactive natural products.[56] Alkylation of 1-methylindole with Me-EDA represented a difficult biocatalytic reaction: no activity was found for either free hemin or for P411-CIS, and variant P411-HF M263Y delivered the product with only 2 TTN. This low activity, however, was sufficient for GC-MS based screening, and we evolved this variant over seven rounds of site-saturation mutagenesis and screening to obtain variant P411-HF T87A Y263E A328Y L437M A268G T327P, which delivers **5** with 775 TTN (Figure 4B, Supplementary Figure 7, Supplementary Table 4). While this represents a >300-fold increase in catalytic activity, enantioselectivity remained at ca. 50% over the course of evolution. The limited enantioselectivity may be grounded in the enzymatic reaction mechanism: nucleophilic attack by the indole C_3 carbon at the iron-carbene center followed by protonation is generally considered the mechanistic pathway of indole alkylation.[36,38,57] This process would require the enzyme to have perfect stereocontrol over the protonation step in order to yield high enantioselectivity. We speculate that further engineering of the P411-HF active site could possibly fine-tune hydrogen- bonding networks between the reaction intermediate and the protein, which might allow better control of the protonation step and lead to higher stereoselectivity. The evolved catalyst is nonetheless among the few P411's efficiently catalyzing transfer of an α-disubstituted carbene.

Finally, we reasoned that P411-HF may provide a good starting point to unlock cyclopropanation reactions of (hetero)cyclic internal alkenes. We first tested cyclopropanation using indene and EDA and found that P411-HF yielded the *cis*-configured cyclopropane **6a** with 1880 TTN, 95:5 dr and 94% *ee* (Figure 4C). This was significantly higher than P411-CIS, showing 305 TTN, 69:31 *dr* and 78% *ee* (Supplementary Figure 8, Supplementary Table 5), supporting the hypothesis that evolution for

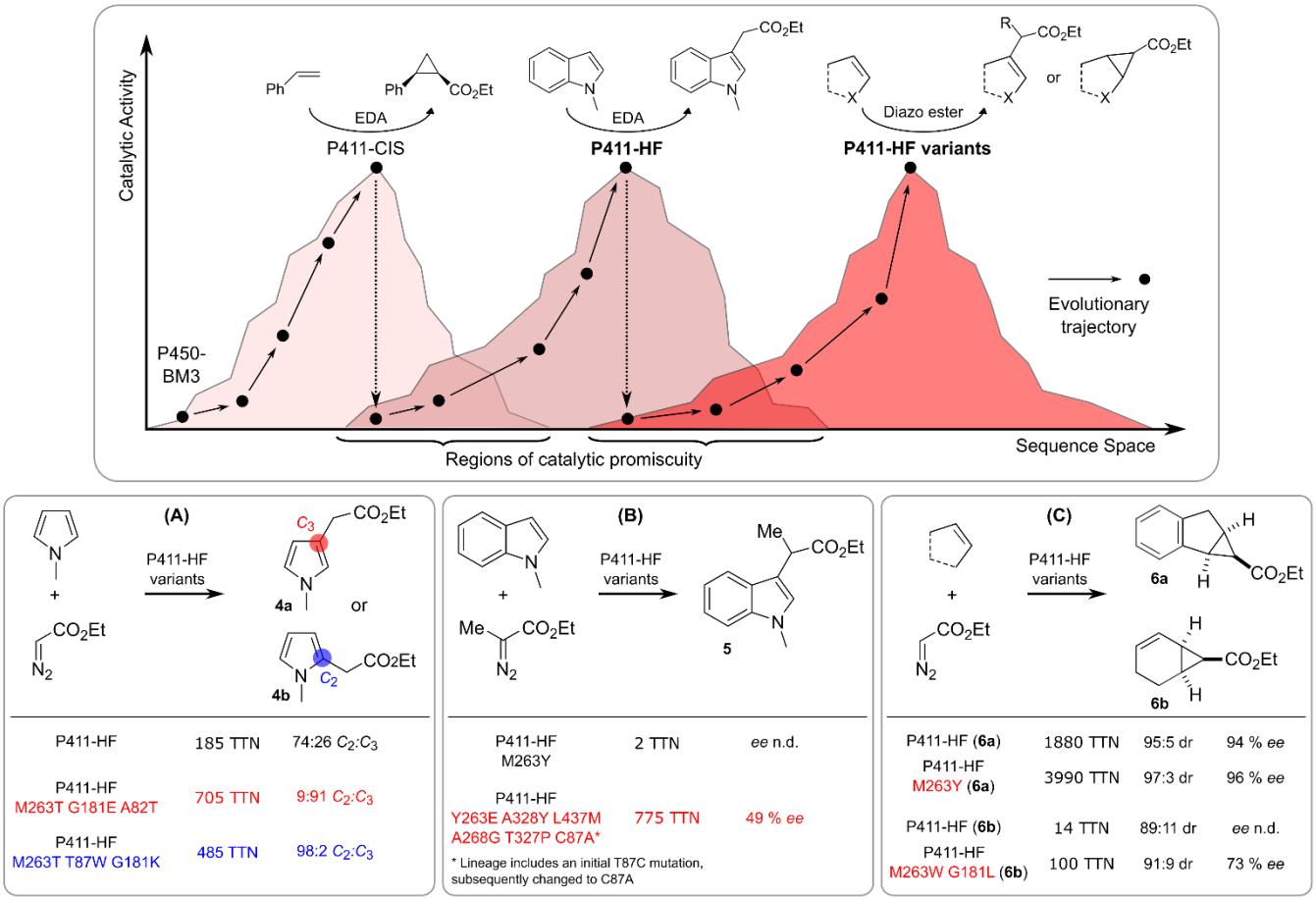


Figure 4: Evolving P411-HF for selective (hetero)cycle functionalization reactions. Top: Schematic representation of the relation between protein sequence and catalytic function. A given enzyme sequence may exhibit additional 'promiscuous' functions that can be optimized by directed evolution.[55] We set out to test whether promiscuous activities of P411-HF, gained during evolution for 1-methylindole alkylation, could serve as starting points to evolve further (hetero)cycle functionalization reactions. (A) Evolution of P411-HF for regioselective 1-methylpyrrole alkylation. (B) Evolution of P411-HF for enantioselective 1-methylindole alkylation with ethyl-2-diazopropanoate. (C) Evolution of P411-HF for stereoselective cyclic alkene cyclopropanation. Data in (A) to (C) were derived from two independent experiments performed in duplicate; see Supplementary Tables 3 to 6 and Supplementary Figures 6 to 8 for evolutionary lineages and experimental details.

indole alkylation in parallel raised activity for cyclopropanation of sterically similar substrates. One additional round of site-saturation mutagenesis and screening delivered P411-HF M263Y, providing **6a** with almost 4000 TTN, 97:3 dr and 96% ee (Figure 4C). Comparable hemoprotein cyclopropanation activity on internal alkene substrates was achieved previously only with iridium porphyrin-substituted hemoprotein variants,[58] and this enzymatic reaction delivers a potentially useful cyclopropane core recently explored as pharmacophore.[59]

Similar results were obtained on a less-activated cyclic alkene, 1,3-cyclohexadiene (Figure 4C, Supplementary Figure 8, Supplementary Table 6). Again, P411-HF showed enhanced cyclopropanation activity over its parent P411-CIS, and two additional rounds of site-saturation mutagenesis and screening yielded P411-HF M263W G181L, delivering the *cis*-configured cyclopropane **6b** with 100 TTN, 91:1 *dr* and 73% *ee*. Lower TTNs (< 100) were observed on heterocyclic substrates, including 2,3-benzofuran, 2,3-dihydrofuran and benzothiophene, while no activity was observed with aliphatic internal alkenes 2,4-hexadiene, 2-octene, cyclopentene, cyclohexene or cycloheptene, indicating that these substrates still present a challenge for the current P411-HF carbene transferase lineage. Nonetheless, the results presented here demonstrate that enzymes containing the native iron-heme cofactor are capable of internal alkene cyclopropanation, with activity and selectivity tunable by directed evolution.

Concluding Remarks

Enzyme catalysis represents a potentially sustainable alternative to chemical synthesis. Unlocking its promise, however, requires that enzymes perform equal or better than state-of-the-art synthetic methods. In this study, we engineered a set of cytochrome P411 variants for carbene transfer to heterocycles and internal alkenes, synthetically important substrate classes largely absent from the repertoire of carbene transferases described to date. We tackled this problem with a two-tiered engineering strategy, first subjecting a cytochrome P411 variant to a comprehensive directed evolution effort, which delivered the highly active indole C_3-alkylation variant P411-HF. Further engineering of P411-HF towards more challenging heterocycle and internal alkene functionalizations then delivered a suite of P411 variants for a diverse set of heterocycle functionalization

reactions. Only a few rounds of site-saturation mutagenesis and screening were required to achieve regioselective pyrrole alkylation, enantioselective indole alkylation, and stereoselective cyclic alkene cyclopropanation. The evolved enzymes catalyze hundreds to thousands of turnovers, are fully genetically encoded, and function with an earth-abundant iron cofactor. These variants complement existing transition metal catalysts for selective heterocycle functionalizations, provide potential starting points for future directed evolution endeavors, and expand the catalytic repertoire of the carbene transferase family of enzymes.

ASSOCIATED CONTENT

Supporting Information. Experimental procedures, supplementary data, calibration curves, and characterization of alkylation and cyclopropanation products. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

* Frances H. Arnold
frances@cheme.caltech.edu

Present Addresses

†Present address: Howard Hughes Medical Institute, Department of Human Genetics, University of California Los Angeles, Los Angeles, California 90095, USA.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

This work was supported by a grant from the National Science Foundation, Division of Molecular and Cellular Biosciences to F.H.A. (NSF MCB-1513007). O.F.B. acknowledges support from the Deutsche Forschungsgemeinschaft (DFG grant BR 5238/1-1) and the Swiss National Science Foundation (SNF grant P300PA-171225). K.C. thanks the Resnick Sustainability Institute at Caltech for fellowship support.

ACKNOWLEDGMENTS

The authors are truly grateful to Dr. Christopher Prier for help with initial experimental design and comments on the manuscript. We also thank Dr. David Romney and Anders Knight for helpful comments and discussions.

REFERENCES

(1) Gomtsyan, A. Heterocycles in Drugs and Drug Discovery. *Chem Heterocycl Comp* **2012**, *48* (1), 7–10. https://doi.org/10.1007/s10593-012-0960-z.

(2) Vitaku, E.; Smith, D. T.; Njardarson, J. T. Analysis of the Structural Diversity, Substitution Patterns, and Frequency of Nitrogen Heterocycles among U.S. FDA Approved Pharmaceuticals. *J. Med. Chem.* **2014**, *57* (24), 10257–10274. https://doi.org/10.1021/jm501100b.

(3) St. Jean, D. J.; Fotsch, C. Mitigating Heterocycle Metabolism in Drug Discovery. *J. Med. Chem.* **2012**, *55* (13), 6002–6020. https://doi.org/10.1021/jm300343m.

(4) Delost, M. D.; Smith, D. T.; Anderson, B. J.; Njardarson, J. T. From Oxiranes to Oligomers: Architectures of U.S. FDA Approved Pharmaceuticals Containing Oxygen Heterocycles. *J. Med. Chem.* **2018**, *61* (24), 10996–11020. https://doi.org/10.1021/acs.jmedchem.8b00876.

(5) Davison, E. K.; Brimble, M. A. Natural Product Derived Privileged Scaffolds in Drug Discovery. *Curr. Opin. Chem. Biol.* **2019**, *52*, 1–8. https://doi.org/10.1016/j.cbpa.2018.12.007.

(6) Topczewski, J. J.; Cabrera, P. J.; Saper, N. I.; Sanford, M. S. Palladium-Catalysed Transannular C–H Functionalization of Alicyclic Amines. *Nature* **2016**, *531* (7593), 220–224. https://doi.org/10.1038/nature16957.

(7) Malapit, C. A.; Howell, A. R. Recent Applications of Oxetanes in the Synthesis of Heterocyclic Compounds. *J. Org. Chem.* **2015**, *80* (17), 8489–8495. https://doi.org/10.1021/acs.joc.5b01255.

(8) Kumar, S.; Cheng, X.; Klimasauskas, S.; Mi, S.; Posfai, J.; Roberts, R. J.; Wilson, G. G. The DNA (Cytosine-5) Methyltransferases. *Nucleic Acids Res* **1994**, *22* (1), 1–10.

(9) Andorfer, M. C.; Lewis, J. C. Understanding and Improving the Activity of Flavin-Dependent Halogenases via Random and Targeted Mutagenesis. *Annu. Rev. Biochem.* **2018**, *87* (1), 159–185. https://doi.org/10.1146/annurev-biochem-062917-012042.

(10) Li, Q.-S.; Schwaneberg, U.; Fischer, P.; Schmid, R. D. Directed Evolution of the Fatty-Acid Hydroxylase P450 BM-3 into an Indole-Hydroxylating Catalyst. *Chemistry – A European Journal* **2000**, *6* (9), 1531–1536. https://doi.org/10.1002/(SICI)1521-3765(20000502)6:9<1531::AID-CHEM1531>3.0.CO;2-D.

(11) Kim, D.; Guengerich, F. P. Cytochrome P450 Activation of Arylamines and Heterocyclic Amines. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45* (1), 27–49. https://doi.org/10.1146/annurev.pharmtox.45.120403.100010.

(12) Zheng, Q.; Fang, H.; Liu, W. Post-Translational Modifications Involved in the Biosynthesis of Thiopeptide Antibiotics. *Org. Biomol. Chem.* **2017**, *15* (16), 3376–3390. https://doi.org/10.1039/C7OB00466D.

(13) Brandenberg, O. F.; Fasan, R.; Arnold, F. H. Exploiting and Engineering Hemoproteins for Abiological Carbene and Nitrene Transfer Reactions. *Curr. Opin. Biotech.* **2017**, *47*, 102–111. https://doi.org/10.1016/j.copbio.2017.06.005.

(14) Coelho, P. S.; Wang, Z. J.; Ener, M. E.; Baril, S. A.; Kannan, A.; Arnold, F. H.; Brustad, E. M. A Serine-Substituted P450 Catalyzes Highly Efficient Carbene Transfer to Olefins *in Vivo*. *Nat. Chem. Biol.* **2013**, *9* (8), 485–487. https://doi.org/10.1038/nchembio.1278.

(15) Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H. Olefin Cyclopropanation via Carbene Transfer Catalyzed by Engineered Cytochrome P450 Enzymes. *Science* **2013**, *339* (6117), 307–310. https://doi.org/10.1126/science.1231434.

(16) Bordeaux, M.; Tyagi, V.; Fasan, R. Highly Diastereoselective and Enantioselective Olefin Cyclopropanation Using Engineered Myoglobin-Based Catalysts. *Angew. Chem. Int. Ed.* **2015**, *54* (6), 1744–1748. https://doi.org/10.1002/anie.201409928.

(17) Chen, K.; Huang, X.; Kan, S. B. J.; Zhang, R. K.; Arnold, F. H. Enzymatic Construction of Highly Strained Carbocycles. *Science* **2018**, *360* (6384), 71–75. https://doi.org/10.1126/science.aar4239.

(18) Zhang, R. K.; Chen, K.; Huang, X.; Wohlschlager, L.; Renata, H.; Arnold, F. H. Enzymatic Assembly of Carbon–

Carbon Bonds via Iron-Catalysed Sp 3 C–H Functionalization. *Nature* **2019**, *565* (7737), 67. https://doi.org/10.1038/s41586-018-0808-5.

(19) Wang, Z. J.; Peck, N. E.; Renata, H.; Arnold, F. H. Cytochrome P450-Catalyzed Insertion of Carbenoids into N–H Bonds. *Chem. Sci.* **2013**, *5* (2), 598–601. https://doi.org/10.1039/C3SC52535J.

(20) Chen, K.; Zhang, S.-Q.; Brandenberg, O. F.; Hong, X.; Arnold, F. H. Alternate Heme Ligation Steers Activity and Selectivity in Engineered Cytochrome P450-Catalyzed Carbene-Transfer Reactions. *J. Am. Chem. Soc.* **2018**, *140* (48), 16402–16407. https://doi.org/10.1021/jacs.8b09613.

(21) Kan, S. B. J.; Huang, X.; Gumulya, Y.; Chen, K.; Arnold, F. H. Genetically Programmed Chiral Organoborane Synthesis. *Nature* **2017**, *552* (7683), 132–136. https://doi.org/10.1038/nature24996.

(22) Kan, S. B. J.; Lewis, R. D.; Chen, K.; Arnold, F. H. Directed Evolution of Cytochrome c for Carbon–Silicon Bond Formation: Bringing Silicon to Life. *Science* **2016**, *354* (6315), 1048–1051. https://doi.org/10.1126/science.aah6219.

(23) Tyagi, V.; Bonn, R. B.; Fasan, R. Intermolecular Carbene S–H Insertion Catalysed by Engineered Myoglobin-Based Catalysts. *Chem. Sci.* **2015**, *6* (4), 2488–2494. https://doi.org/10.1039/C5SC00080G.

(24) Wang, Z. J.; Renata, H.; Peck, N. E.; Farwell, C. C.; Coelho, P. S.; Arnold, F. H. Improved Cyclopropanation Activity of Histidine-Ligated Cytochrome P450 Enables the Enantioselective Formal Synthesis of Levomilnacipran. *Angew. Chem. Int. Ed. 126* (26), 6928–6931. https://doi.org/10.1002/ange.201402809.

(25) Bajaj, P.; Sreenilayam, G.; Tyagi, V.; Fasan, R. Gram-Scale Synthesis of Chiral Cyclopropane-Containing Drugs and Drug Precursors with Engineered Myoglobin Catalysts Featuring Complementary Stereoselectivity. *Angew. Chem. Int. Ed.* **2016**, *55* (52), 16110–16114. https://doi.org/10.1002/anie.201608680.

(26) Hernandez, K. E.; Renata, H.; Lewis, R. D.; Kan, S. B. J.; Zhang, C.; Forte, J.; Rozzell, D.; McIntosh, J. A.; Arnold, F. H. Highly Stereoselective Biocatalytic Synthesis of Key Cyclopropane Intermediate to Ticagrelor. *ACS Catal.* **2016**, *6* (11), 7810–7813. https://doi.org/10.1021/acscatal.6b02550.

(27) Davies, H. M. L.; Hedley, S. J. Intermolecular Reactions of Electron-Rich Heterocycles with Copper and Rhodium Carbenoids. *Chem. Soc. Rev.* **2007**, *36* (7), 1109–1119. https://doi.org/10.1039/B607983K.

(28) Reiser, O. Catalytic Conversion of Furans and Pyrroles to Natural Products and Analogues Utilizing Donor-Acceptor Substituted Cyclopropanes as Key Intermediates. *Isr. J. Chem.* **2016**, *56* (6–7), 531–539. https://doi.org/10.1002/ijch.201500103.

(29) Xu, H.; Li, Y.-P.; Cai, Y.; Wang, G.-P.; Zhu, S.-F.; Zhou, Q.-L. Highly Enantioselective Copper- and Iron-Catalyzed Intramolecular Cyclopropanation of Indoles. *J. Am. Chem. Soc.* **2017**, *139* (23), 7697–7700. https://doi.org/10.1021/jacs.7b03086.

(30) DeAngelis, A.; Shurtleff, V. W.; Dmitrenko, O.; Fox, J. M. Rhodium(II)-Catalyzed Enantioselective C−H Functionalization of Indoles. *J. Am. Chem. Soc.* **2011**, *133* (6), 1650–1653. https://doi.org/10.1021/ja1093309.

(31) Gao, X.; Wu, B.; Huang, W.-X.; Chen, M.-W.; Zhou, Y.-G. Enantioselective Palladium-Catalyzed C□H Functionalization of Indoles Using an Axially Chiral 2,2′-Bipyridine Ligand. *Angew. Chem. Int. Ed.* **2015**, *54* (41), 11956–11960. https://doi.org/10.1002/anie.201504483.

(32) Chan, W.-W.; Yeung, S.-H.; Zhou, Z.; Chan, A. S. C.; Yu, W.-Y. Ruthenium Catalyzed Directing Group-Free C2-Selective Carbenoid Functionalization of Indoles by α-Aryldiazoesters. *Org. Lett.* **2010**, *12* (3), 604–607. https://doi.org/10.1021/ol9028226.

(33) Sevrioukova, I. F.; Li, H.; Zhang, H.; Peterson, J. A.; Poulos, T. L. Structure of a Cytochrome P450–Redox Partner Electron-Transfer Complex. *PNAS* **1999**, *96* (5), 1863–1868.

(34) Joyce, M. G.; Ekanem, I. S.; Roitel, O.; Dunford, A. J.; Neeli, R.; Girvan, H. M.; Baker, G. J.; Curtis, R. A.; Munro, A. W.; Leys, D. The Crystal Structure of the FAD/NADPH-Binding Domain of Flavocytochrome P450 BM3. *FEBS J.* **2012**, *279* (9), 1694–1706. https://doi.org/10.1111/j.1742-4658.2012.08544.x.

(35) C. Whitehouse, C. J.; G. Bell, S.; Wong, L.-L. P450 BM3 (CYP102A1): Connecting the Dots. *Chem. Soc. Rev.* **2012**, *41* (3), 1218–1260. https://doi.org/10.1039/C1CS15192D.

(36) Vargas, D. A.; Tinoco, A.; Tyagi, V.; Fasan, R. Myoglobin-Catalyzed C−H Functionalization of Unprotected Indoles. *Angew. Chem. Int. Ed.* **2018**, *57* (31), 9911–9915. https://doi.org/10.1002/anie.201804779.

(37) Hock, K. J.; Knorrscheidt, A.; Hommelsheim, R.; Ho, J.; Weissenborn, M. J.; Koenigs, R. M. Tryptamine Synthesis by Iron Porphyrin Catalyzed C−H Functionalization of Indoles with Diazoacetonitrile. *Angew. Chem. Int. Ed.* **2019**, *58* (11), 3630–3634. https://doi.org/10.1002/anie.201813631.

(38) Cai, Y.; Zhu, S.-F.; Wang, G.-P.; Zhou, Q.-L. Iron-Catalyzed C□H Fuctionalization of Indoles. *Adv. Syn. Catal.* **2011**, *353* (16), 2939–2944. https://doi.org/10.1002/adsc.201100334.

(39) Delgado-Rebollo, M.; Prieto, A.; Pérez, P. J. Catalytic Functionalization of Indoles by Copper-Mediated Carbene Transfer. *ChemCatChem* **2014**, *6* (7), 2047–2052. https://doi.org/10.1002/cctc.201400097.

(40) Gao, X.; Wu, B.; Yan, Z.; Zhou, Y.-G. Copper-Catalyzed Enantioselective C–H Functionalization of Indoles with an Axially Chiral Bipyridine Ligand. *Org. Biomol. Chem.* **2016**, *14* (35), 8237–8240. https://doi.org/10.1039/C6OB01516F.

(41) Renata, H.; Lewis, R. D.; Sweredoski, M. J.; Moradian, A.; Hess, S.; Wang, Z. J.; Arnold, F. H. Identification of Mechanism-Based Inactivation in P450-Catalyzed Cyclopropanation Facilitates Engineering of Improved Enzymes. *J. Am. Chem. Soc.* **2016**, *138* (38), 12527–12533. https://doi.org/10.1021/jacs.6b06823.

(42) Brandenberg, O. F.; Prier, C. K.; Chen, K.; Knight, A. M.; Wu, Z.; Arnold, F. H. Stereoselective Enzymatic Synthesis of Heteroatom-Substituted Cyclopropanes. *ACS Catal.* **2018**, *8* (4), 2629–2634. https://doi.org/10.1021/acscatal.7b04423.

(43) Ignatova, Z.; Gierasch, L. M. Monitoring Protein Stability and Aggregation in Vivo by Real-Time Fluorescent Labeling. *PNAS* **2004**, *101* (2), 523–528. https://doi.org/10.1073/pnas.0304533101.

(44) Ghaemmaghami, S.; Oas, T. G. Quantitative Protein Stability Measurement *in Vivo*. *Nat. Struc. Mol. Biol.* **2001**, *8* (10), 879–882. https://doi.org/10.1038/nsb1001-879.

(45) Black, S. D.; Martin, S. T. Evidence for Conformational Dynamics and Molecular Aggregation in Cytochrome P450 102 (BM-3). *Biochemistry* **1994**, *33* (40), 12056–12062. https://doi.org/10.1021/bi00206a007.

(46) Zhang, H.; Yokom, A. L.; Cheng, S.; Su, M.; Hollenberg, P. F.; Southworth, D. R.; Osawa, Y. The Full-Length Cytochrome P450 Enzyme CYP102A1 Dimerizes at Its Re-

ductase Domains and Has Flexible Heme Domains for Efficient Catalysis. *J. Biol. Chem.* **2018**, *293* (20), 7727–7736. https://doi.org/10.1074/jbc.RA117.000600.

(47) Kitazume, T.; Haines, D. C.; Estabrook, R. W.; Chen, B.; Peterson, J. A. Obligatory Intermolecular Electron-Transfer from FAD to FMN in Dimeric P450BM-3. *Biochemistry* **2007**, *46* (42), 11892–11901. https://doi.org/10.1021/bi701031r.

(48) Neeli, R.; Girvan, H. M.; Lawrence, A.; Warren, M. J.; Leys, D.; Scrutton, N. S.; Munro, A. W. The Dimeric Form of Flavocytochrome P450 BM3 Is Catalytically Functional as a Fatty Acid Hydroxylase. *FEBS Lett.* **2005**, *579* (25), 5582–5588. https://doi.org/10.1016/j.febslet.2005.09.023.

(49) Jensen, R. A. Enzyme Recruitment in Evolution of New Function. *Annu. Rev. Microbiol.* **1976**, *30* (1), 409–425. https://doi.org/10.1146/annurev.mi.30.100176.002205.

(50) O'Brien, P. J.; Herschlag, D. Catalytic Promiscuity and the Evolution of New Enzymatic Activities. *Chemistry & Biology* **1999**, *6* (4), R91–R105. https://doi.org/10.1016/S1074-5521(99)80033-7.

(51) Khersonsky, O.; Tawfik, D. S. Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. *Annu. Rev. Biochem.* **2010**, *79* (1), 471–505. https://doi.org/10.1146/annurev-biochem-030409-143718.

(52) Maryanoff, B. E. Reaction of Dimethyl Diazomalonate and Ethyl 2-Diazoacetoacetate with N-Methylpyrrole. *J. Org. Chem.* **1982**, *47* (15), 3000–3002. https://doi.org/10.1021/jo00136a038.

(53) Maryanoff, B. E. Carbenoid Chemistry. Reaction of Pyrrole Derivatives with Ethyl Diazoacetate. *J. Org. Chem.* **1979**, *44* (24), 4410–4419. https://doi.org/10.1021/jo01338a033.

(54) Andorfer, M. C.; Grob, J. E.; Hajdin, C. E.; Chael, J. R.; Siuti, P.; Lilly, J.; Tan, K. L.; Lewis, J. C. Understanding Flavin-Dependent Halogenase Reactivity via Substrate Activity Profiling. *ACS Catal.* **2017**, *7* (3), 1897–1904. https://doi.org/10.1021/acscatal.6b02707.

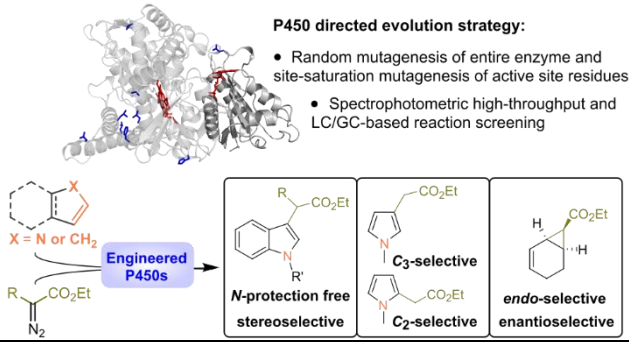
(55) Renata, H.; Wang, Z. J.; Arnold, F. H. Expanding the Enzyme Universe: Accessing Non-Natural Reactions by Mechanism-Guided Directed Evolution. *Angew. Chem. Int. Ed.* **2015**, *54* (11), 3351–3367. https://doi.org/10.1002/anie.201409470.

(56) Goto, T.; Natori, Y.; Takeda, K.; Nambu, H.; Hashimoto, S. Catalytic Enantioselective C–H Functionalization of Indoles with α-Diazopropionates Using Chiral Dirhodium(II) Carboxylates: Asymmetric Synthesis of the (+)-α-Methyl-3-Indolylacetic Acid Fragment of Acremoauxin A. *Tetrahedron: Asymmetry* **2011**, *22* (8), 907–915. https://doi.org/10.1016/j.tetasy.2011.05.011.

(57) Davies, H. M. L.; Spangler, J. E. Reactions of Indoles with Metal-Bound Carbenoids. In *Adv. Heterocyc. Chem.*; Elsevier, 2013; Vol. 110, pp 43–72. https://doi.org/10.1016/B978-0-12-408100-0.00002-1.

(58) Key, H. M.; Dydio, P.; Liu, Z.; Rha, J. Y.-E.; Nazarenko, A.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F. Beyond Iron: Iridium-Containing P450 Enzymes for Selective Cyclopropanations of Structurally Diverse Alkenes. *ACS Cent. Sci.* **2017**, *3* (4), 302–308. https://doi.org/10.1021/acscentsci.6b00391.

(59) Ji, Y.-Y.; Lin, S.-D.; Wang, Y.-J.; Su, M.-B.; Zhang, W.; Gunosewoyo, H.; Yang, F.; Li, J.; Tang, J.; Zhou, Y.-B.; Yu, L.-F. Tying up Tranylcypromine: Novel Selective Histone Lysine Specific Demethylase 1 (LSD1) Inhibitors. *Eur. J. Med. Chem.* **2017**, *141*, 101–112. https://doi.org/10.1016/j.ejmech.2017.09.073.


P450 directed evolution strategy:
• Random mutagenesis of entire enzyme and site-saturation mutagenesis of active site residues
• Spectrophotometric high-throughput and LC/GC-based reaction screening
X = N or CH2
Engineered P450s
N-protection free stereoselective
C3-selective
C2-selective
endo-selective enantioselective