HHS Public Access
Author manuscript
Nature. Author manuscript; available in PMC 2019 June 19.

Published in final edited form as:
Nature. 2019 January ; 565(7737): 67–72. doi:10.1038/s41586-018-0808-5.

Enzymatic assembly of carbon–carbon bonds via iron-catalysed *sp*³ C–H functionalization

Ruijie K. Zhang[1], **Kai Chen**[1], **Xiongyi Huang**[1], **Lena Wohlschlager**[1,2], **Hans Renata**[1,3], and **Frances H. Arnold**[1,*]

[1]Division of Chemistry and Chemical Engineering MC 210-41, California Institute of Technology, Pasadena, CA 91125, USA.

[2]Current address: Institute of Food Technology, University of Natural Resources and Life Sciences, Vienna, Austria

[3]Current address: Department of Chemistry, The Scripps Research Institute, 130 Scripps Way, Jupiter, Florida 33458, USA.

Though abundant in organic molecules, carbon–hydrogen (C–H) bonds are typically considered unreactive and unavailable for chemical manipulation. Recent advances in C–H functionalization technology have begun to transform this logic, while emphasizing the challenge and importance of selective installation of sp^3 C–alkyl groups into a hydrocarbon framework[1,2]. Here we describe the first iron-based catalysts for enantio-, regio-, and chemo-selective intermolecular alkylation of sp^3 C–H bonds through carbene C–H insertion. The catalysts, derived from a cytochrome P450 enzyme whose native cysteine axial ligand has been substituted for serine ("cytochrome P411"), are fully genetically encoded and produced in bacteria, where they can be tuned by directed evolution for activity and selectivity. That these proteins activate iron, the most abundant transition metal, to perform this challenging chemistry provides a desirable alternative to noble metal catalysts, which have dominated the field of C–H functionalization[1,2]. The laboratory-evolved enzymes functionalize diverse substrates containing benzylic, allylic, or α-amino C–H bonds with high turnover and exquisite selectivity. Furthermore, these highly efficient enzymes have enabled the development of concise routes to several natural products. The demonstration that these enzymes mediate sp^3 C–H alkylation using their native iron-haem cofactor unlocks a vast natural haem protein diversity for this abiological transformation and will

*Correspondence and requests for materials should be addressed to F.H.A. frances@cheme.caltech.edu.

Author contributions
R.K.Z. designed the overall research with F.H.A. providing guidance. R.K.Z. and H.R. designed and conducted the initial screening of haem proteins; R.K.Z. and L.W. performed the directed evolution experiments. R.K.Z., K.C., and X.H. designed and performed the substrate scope studies. R.K.Z. and F.H.A. wrote the manuscript with input from all authors.

Data Availability
Full experimental details, information about enzyme variants, Supplementary Figures S1–S15, and Supplementary Tables S1–S13 are available in the Supplementary Information. Any additional information is available from the corresponding author upon request.

Author Information
A provisional patent application has been filed through the California Institute of Technology based on the results presented here.

Supplementary Information is available in the online version of the paper.

facilitate the development of new enzymatic C–H functionalization reactions for applications in chemistry and synthetic biology.

Biological systems use a limited set of chemical strategies to form carbon–carbon (C–C) bonds during construction of organic molecules[3]. Whereas many of these approaches rely on the manipulation of functional groups, certain enzymes, including members of the radical *S*-adenosylmethionine (SAM) family, can perform alkylation of sp^3 C–H bonds. This has been an especially versatile strategy for structural diversification, as seen by its essential role in the biosynthesis of structurally varied natural products and cofactors[4–6]. Known biological machineries for this transformation, however, are limited to enzymes that transfer a methyl group[5,6] or conjugate an activated radical acceptor substrate[4,7] to specific molecules, with methylation as a common mode for sp^3 C–alkyl installation by radical SAM enzymes (Fig. 1a).

We sought to introduce a new enzymatic strategy for the alkylation of sp^3 C–H bonds. For our design, we drew inspiration from the most widely used biological C–H functionalization transformation, C–H oxygenation. Enzymes such as the cytochromes P450 accomplish C–H oxygenation using a haem cofactor; their activities rely on activation of molecular oxygen for the controlled generation of a high-energy iron-oxo intermediate capable of selective insertion into a substrate C–H bond[8]. Analogously, we anticipated that the combination of a haem protein and a diazo compound would generate a protein-enclosed iron-carbene species and that this carbene could participate in a selective C–H insertion reaction with a second substrate (Fig. 1b). While it has been shown that haem proteins are capable of performing carbene transfer processes such as cyclopropanation and heteroatom–hydrogen bond insertions[9–11], their functionalization of sp^3 C–H bonds remained elusive.

Metal-carbene sp^3 C–H insertion in small-molecule catalysis, especially intermolecular and stereoselective versions of this reaction, typically relies on transition metal complexes based on rhodium[12], iridium[13], and others[14–16]. Artificial metalloproteins for carbene C–H insertion have been created by introducing an iridium-porphyrin into variants of apo haem proteins[17]. Though rare, there are a few examples of iron-carbene sp^3 C–H insertion. The iron-catalysed examples employ elevated temperatures (e.g. 80 °C)[18], are stoichiometric[19], or are restricted to intramolecular reactions[20], indicating a high activation energy barrier for C–H insertion with an iron-carbene. However, because the protein framework of an enzyme can impart significant rate enhancements to reactions[21] and even confer activity to an otherwise unreactive cofactor[22], we surmised that directed evolution could reconfigure a haem protein to overcome the barrier for the iron-carbene C–H insertion reaction and acquire this new function (Fig. 1b).

In initial studies, we tested a panel of seventy-eight haem proteins which included variants of cytochromes P450, cytochromes *c*, and globin homologs. The haem proteins in whole *Escherichia coli* (*E. coli*) cells were combined with *p*-methoxybenzyl methyl ether (**1a**) and ethyl diazoacetate (**2**) at room temperature under anaerobic conditions; the resulting reactions were analysed for formation of C–H alkylation product **3a** (Fig. 2a., see Supplementary Information for the complete list of tested haem proteins). We found haem proteins from two superfamilies that showed low levels of this promiscuous activity, establishing the possibility of creating C–H alkylation enzymes with very different protein architectures. An engineered variant of cytochrome $P450_{BM3}$ from *Bacillus megaterium* with an axial cysteine-to-serine mutation (cytochrome "P411"), P-4 A82L (ref. 22), provided **3a** with 13 total turnovers (TTN). In addition, nitric oxide dioxygenase from *Rhodothermus marinus* containing the Y32G mutation (*Rma* NOD Y32G) catalysed the reaction with 7 TTN. A second alkane substrate, 4-ethylanisole (**1i**), was also accepted by the nascent C–H alkylation enzymes, albeit with lower turnover numbers (Supplementary Table S2). The haem cofactor alone (iron protoporphyrin IX) or in the presence of bovine serum albumin were inactive (Supplementary Tables S1 and S2).

With P411 P-4 A82L as the starting template, sequential rounds of site-saturation mutagenesis and screening in whole *E. coli* cells were performed to identify increasingly active and enantioselective biocatalysts for C–H alkylation. Amino acid residues chosen for mutagenesis included those which line the active site pocket, reside on loops and other flexible regions of the protein, or possess a nucleophilic side chain[23]. Improved variants were subsequently evaluated in reactions using clarified *E. coli* lysate with *p*-methoxybenzyl methyl ether (**1a**) and 4-ethylanisole (**1i**) (Fig. 2b and Supplementary Fig. S1). Five rounds of mutagenesis and screening yielded variant P411-gen6, which furnished product **3a** with 60 TTN. Unlike the native monooxygenase activity, the C–H alkylation process does not require reducing equivalents from the FAD and FMN domains of the enzyme. Surmising that these domains may not be needed for the C–H alkylation reaction, we performed systematic truncations of P411-gen6 to determine the minimally sufficient domain(s) for retaining catalytic activity. Curiously, removal of the FAD domain, containing 37% of the amino acids in the full-length protein, created an enzyme with higher C–H alkylation activity: P411ΔFAD-gen6 delivers **3a** with 100 TTN, a 1.7-fold increase in TTN compared with P411-gen6 (Supplementary Fig. S2). This indicates that the FAD domain may have (negative) allosteric effects on C–H alkylation activity. Further studies with these truncated enzymes revealed that they could be used in whole *E. coli* cells, in clarified *E. coli* cell lysate, and as purified proteins (Supplementary Table S3). Eight additional rounds of mutagenesis and screening yielded P411-CHF (P411ΔFAD *C*–*H F*unctionalization enzyme, full list of changes provided in the Supplementary Information).

P411-CHF displays 140-fold improvement in activity over P-4 A82L and delivers **3a** with excellent stereoselectivity (2020 TTN, 96.7 : 3.3 e.r. using clarified *E. coli* lysate). Subsequent studies showed that the stereoselectivity could be improved by conducting the reaction at lower temperature (e.g. 4 °C) with no significant change to TTN (Supplementary Table S4). Enzymatic C–H alkylation can be performed on millimole scale: using 1.0 mmol substrate **1a**, *E. coli* harbouring P411-CHF at 4 °C furnished **3a** in 82% isolated yield, 1060 TTN, and 98.0 : 2.0 e.r. (Fig. 2b). Preliminary mechanistic investigations were pursued to interrogate the nature of the C–H insertion step. Independent initial rates measured for reactions with substrate **1a** or deuterated substrate **1a-*d*$_2$** revealed a normal kinetic isotope effect (KIE) of 5.1 for C–H alkylation catalysed by P411-CHF, suggesting that C–H insertion is rate-determining (Supplementary Fig. S5).

Using *E. coli* harbouring P411-CHF, we assayed a range of benzylic substrates for coupling with ethyl diazoacetate (Fig. 3). Both electron-rich and electron-deficient functionalities on the aromatic ring are well-tolerated (**3a**–**3e**, **3h**); cyclic substrates are also suitable coupling partners (**3f**, **3g**). Functionalization of alkyl benzenes is successful at secondary benzylic *sp*^{3} C–H bonds (**3i**–**3l**). Notably, in the biotransformation of substrate **1l** containing both tertiary and secondary benzylic C–H bonds, P411-CHF preferentially functionalizes the secondary position despite its higher C–H bond dissociation energy (BDE). The carbene intermediate derived from ethyl diazoacetate belongs to the acceptor-only class. In contrast to the more widely-used donor/acceptor carbenes, acceptor-only intermediates are more electrophilic, and as a result selective reactions with this carbene class are still a major challenge for small-molecule catalysts[13,16]. Our results show that P411-CHF can control this highly reactive intermediate to furnish the desired *sp*^{3} C–H alkylation products and do so with high enantioselectivity.

Enzymes can exhibit excellent reaction selectivity arising from their ability to form multiple interactions with substrates and intermediates throughout a reaction cycle. We hypothesized that the protein scaffold could be tuned to create complementary enzymes which access different reaction outcomes available to a substrate. When P411-CHF was challenged with 4-allylanisole (**1m**), a substrate which can undergo both C–H alkylation and cyclopropanation, we observed that C–H alkylation product **3m** dominates, with selectivity > 25:1 (Fig. 3b, Supplementary Fig. S6). In contrast, a related full-length P411 variant P-I263F, containing thirteen mutations in the haem domain relative to P411-CHF, catalysed only the formation of cyclopropane product **3m'**. Additionally, despite the established reactivity of silanes with iron-carbene[10], P411-CHF delivered C–H alkylation product **3h** when substrate **1h** was employed in the reaction (Si–H insertion product **3h'** was also observed but its formation may not be catalysed by P411-CHF, Supplementary Fig. S7). Reaction with P-I263F, in contrast, provided only the Si–H insertion product. These examples demonstrate an exceptional feature of macromolecular enzymes: different products can be obtained simply by changing the amino acid sequence of the protein catalyst.

Enzymatic C–H alkylation is not limited to functionalization of benzylic C–H bonds. Structurally dissimilar molecules containing allylic or propargylic C–H bonds are excellent substrates for this chemistry (Fig. 4a). In contrast to **1a**–**1m**, which contain a rigid benzene ring, compounds **4a**–**4c** and **4e** feature flexible linear alkyl chains. Their successful enantioselective alkylation suggests that the enzyme active site can accommodate substrate conformational flexibility while enforcing a favoured substrate orientation relative to the carbene intermediate. To demonstrate the utility of this biotransformation, we applied the methodology to the formal synthesis of lyngbic acid (Fig. 4a). Marine cyanobacteria incorporate this versatile biomolecule into members of the malyngamide family of natural products; likewise, total synthesis approaches to these natural products typically access lyngbic acid as a strategic intermediate *en route* to the target molecules[24]. Using *E. coli* harbouring P411-CHF, intermediate **5a** was produced on 2.4 mmol scale in 86% isolated yield, 2810 TTN, and 94.7 : 5.3 e.r.. Subsequent hydrogenation and hydrolysis provided (*R*)-(+)-**6** in quantitative yield, which can be elaborated to (*R*)-(+)-lyngbic acid by decarboxylative alkenylation[25].

As part of our substrate scope studies, we challenged P411-CHF with alkyl amine compounds. Compounds of this type are typically challenging for C–H functionalization methods because the amine functionality may coordinate to and inhibit the catalyst or create the opportunity for undesirable side reactions (e.g. ylide formation and its associated rearrangements)[26]. Using **7a** or **7b**, substrates which have both benzylic C–H bonds and α-amino C–H bonds, P411-CHF delivered the corresponding *β*-amino ester product with high efficiency (**8a** and **8b**, Fig. 4b). Notably, benzylic C–H insertion was not observed (with **7a**, Supplementary Fig. S9) or significantly suppressed (with **7b**, Supplementary Fig. S10), despite the typically lower BDEs of benzylic C–H bonds compared to α-amino C–H bonds. Additionally, *N*-aryl pyrrolidines (**7c**–**7e**) served as excellent substrates and were selectively alkylated at the α-amino sp^3 position. Using P411-CHF, the sp^3 C–H alkylation of **7c** outcompetes a Friedel-Crafts type reaction on the aryl ring, which is a favourable process with other carbene-transfer systems[27, 28]. Furthermore, alkylation product **8d** offers a conceivable strategy for the synthesis of *β*-homoproline, a motif which has been investigated for medicinal chemistry applications[29].

Given that P411-CHF alkylates both primary and secondary α-amino C–H bonds, we interrogated whether the enzyme could be selective for one of these positions. Employing *N*-methyl tetrahydroquinoline **7f** as the alkane substrate, P411-CHF afforded *β*-amino ester products with 1050 TTN and a 9 : 1 ratio of regioisomers (C2 : C1, and 73.0 : 27.0 e.r. for **8f**) (Fig. 4b). As the tetrahydroquinoline ring is a privileged structural motif in natural products and bioactive molecules[30], its selective functionalization could provide a concise strategy for the synthesis of alkaloids. To improve the selectivity for alkylation of **7f**, we tested variants along the evolutionary lineage from P-4 A82L to P411-CHF. We found that P411-gen5 had even better regioselectivity and delivered product with the opposite stereo-preference. In a 3.0 mmol scale reaction, *E. coli* harbouring P411-gen5 delivered **8f** in 85% yield with excellent selectivity (1310 TTN, > 50 : 1 r.r., 8.9 : 91.1 e.r.). In only a few steps, the enzymatic product was successfully transformed to alkaloid (*R*)-(+)-cuspareine[30] (Fig. 4b).

Finally, we probed the introduction of different alkyl groups. Using different diazo reagents, enzymatic C–H alkylation can diversify one alkane substrate, such as **7a**, to several products (**10a**–**10c** in Fig. 4c and Supplementary Fig. S11). The diazo substrate scope extends beyond ester-based reagents: Weinreb amide diazo compound **9c** and diazoketone **9d** were found to participate in enzymatic C–H alkylation to furnish products **10c** and **10d**, respectively. Additional substitution at the α-position of the carbene, however, is generally not well-tolerated by P411-CHF and current related enzymes. With the exception of **10b**, reactions using disubstituted carbene reagents failed to yield appreciable amounts of desired products (Supplementary Fig. S11).

This study demonstrates that a cytochrome P450 can acquire the ability to construct C–C bonds from sp^3 C–H bonds and that activity and selectivity can be greatly enhanced using directed evolution. Nature provides a huge collection of possible alternative starting points for expanding the scope of this reaction even further and for achieving other selectivities. The cytochrome P450 superfamily can access an immense set of organic molecules for its native oxygenation chemistry; we envision that P411-derived enzymes and other natural haem protein diversity can be leveraged to generate families of C–H alkylation enzymes that emulate the scope and selectivity of nature’s C–H oxygenation catalysts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by the National Science Foundation (NSF), Division of Molecular and Cellular Biosciences (grant MCB-1513007). R.K.Z. acknowledges support from the NSF Graduate Research Fellowship (grant DGE-1144469) and the Donna and Benjamin M. Rosen Bioengineering Center. X.H. is supported by a Ruth L. Kirschstein National Institutes of Health Postdoctoral Fellowship (grant F32GM125231). L.W. received support from the Austrian Marshall Plan Foundation. We thank A. Z. Zhou for experimental assistance; N. W. Goldberg, S. C. Hammer, K. E. Hernandez, Z. Jia, A. M. Knight, G. Kubik, R. D. Lewis, C. K. Prier, D. K. Romney, and J. Zhang for helpful discussions; S. Virgil, M. Shahgholi, and D. VanderVelde for analytical support; B. Stoltz for use of polarimeter and gas chromatography equipment.

References

1. Hartwig JF & Larsen MA Undirected, homogeneous C–H bond functionalization: Challenges and opportunities. ACS Cent. Sci 2, 281–292 (2016). [PubMed: 27294201]
2. Saint-Denis TG, Zhu R-Y, Chen G, Wu Q-F & Yu J-Q Enantioselective C(sp3)–H bond activation by chiral transition metal catalysts. Science 359, doi: 10.1126/science.aao4798 (2018).
3. Frey PA & Hegeman AD Enzymatic Reaction Mechanisms (Oxford University Press, New York, 2007), chap. 14.
4. Yokoyama K & Lilla EA C–C bond forming radical SAM enzymes involved in the construction of carbon skeletons of cofactors and natural products. Nat. Prod. Rep 35, 660–694 (2018). [PubMed: 29633774]
5. Bauerle MR, Schwalm EL & Booker SJ Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation. J. Biol. Chem 290, 3995–4002 (2015). [PubMed: 25477520]
6. McLaughlin MI & van der Donk WA Stereospecific radical-mediated B_{12}-dependent methyl transfer by the fosfomycin biosynthesis enzyme Fom3. Biochemistry 57, 4967–4971 (2018). [PubMed: 29969250]
7. Shisler KA & Broderick JB Glycyl radical activating enzymes: Structure, mechanism, and substrate interactions. Arch. Biochem. Biophys 546, 64–71 (2014). [PubMed: 24486374]
8. Poulos TL Heme enzyme structure and function. Chem. Rev 114, 3919–3962 (2014). [PubMed: 24400737]
9. Coelho PS, Brustad EM, Kannan A & Arnold FH Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. Science 339, 307–310 (2013). [PubMed: 23258409]
10. Kan SBJ, Lewis RD, Chen K & Arnold FH Directed evolution of cytochrome c for carbon–silicon bond formation: Bringing silicon to life. Science 354, 1048–1051 (2016). [PubMed: 27885032]
11. Brandenberg OF, Fasan R & Arnold FH Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions. Curr. Opin. Biotechnol 47, 102–111 (2017). [PubMed: 28711855]
12. Liao K, Pickel TC, Boyarskikh V, Bacsa J, Musaev DG & Davies HML Site-selective and stereoselective functionalization of non-activated tertiary C–H bonds. Nature 551, 609–613 (2017). [PubMed: 29156454]
13. Weldy NM, Schafer AG, Owens CP, Herting CJ, Varela-Alvarez A, Chen S, Niemeyer Z, Musaev DG, Sigman MS, Davies HML & Blakey SB Iridium(III)-bis(imidazolinyl)phenyl catalysts for enantioselective C–H functionalization with ethyl diazoacetate. Chem. Sci 7, 3142–3146 (2016). [PubMed: 29997805]
14. Wang Y, Wen X, Cui X & Zhang XP Enantioselective radical cyclization for construction of 5-membered ring structures by metalloradical C−H alkylation. J. Am. Chem. Soc 140, 4792–4796 (2018). [PubMed: 29584958]
15. Caballero A, Despagnet-Ayoub E, Díaz-Requejo MM, Díaz-Rodríguez A, González-Núñez ME, Mello R, Muñoz BK, Ojo W-S, Asensio G, Etienne M & Pérez PJ Silver-catalyzed C−C bond

formation between methane and ethyl diazoacetate in supercritical CO_2. Science 332, 835–838 (2011). [PubMed: 21566191]

16. Wu W-T, Yang Z-P & You S-L in *Asymmetric Functionalization of C–H Bonds*, ed. You S-L (Royal Society of Chemistry, Cambridge, UK, 2015), chap. 1.
17. Key HM, Dydio P, Clark DS & Hartwig JF Abiological catalysis by artificial haem proteins containing noble metals in place of iron. Nature 534, 534–537 (2016). [PubMed: 27296224]
18. Mbuvi HM & Woo LK Catalytic C–H insertions using iron(III) porphyrin complexes. Organometallics 27, 637–645 (2008).
19. Li Y, Huang J-S, Zhou Z-Y, Che C-M, & You X-Z Remarkably stable iron porphyrins bearing nonheteroatom-stabilized carbene or (alkoxycarbonyl)carbenes: Isolation, X-ray crystal structures, and carbon atom transfer reactions with hydrocarbons. J. Am. Chem. Soc 124, 13185–13193 (2002). [PubMed: 12405847]
20. Griffin JR, Wendell CI, Garwin JA & White MC Catalytic C(sp3)–H alkylation via an iron carbene intermediate. J. Am. Chem. Soc 139, 13624–13627 (2017). [PubMed: 28898063]
21. Herschlag D & Natarajan A Fundamental challenges in mechanistic enzymology: Progress toward understanding the rate enhancements of enzymes. Biochemistry 52, 2050–2067 (2013). [PubMed: 23488725]
22. Prier CK, Zhang RK, Buller AR, Brinkmann-Chen S & Arnold FH Enantioselective, intermolecular benzylic C−H amination catalysed by an engineered iron-haem enzyme. Nat. Chem 9, 629–634 (2017). [PubMed: 28644476]
23. Renata H, Lewis RD, Sweredoski MJ, Moradian A, Hess S, Wang ZJ & Arnold FH Identification of mechanism-based inactivation in P450-catalyzed cyclopropanation facilitates engineering of improved enzymes. J. Am. Chem. Soc 138, 12527–12533 (2016). [PubMed: 27573353]
24. Zhang J-T, Qi X-L, Chen J, Li B-S, Zhou Y-B & Cao X-P Total synthesis of malyngamides K, L, and 5”-*epi*-C and absolution configuration of malyngamide L. J. Org. Chem 76, 3946–3959 (2011). [PubMed: 21495715]
25. Edwards JT, Merchant RR, McClymont KS, Knouse KW, Qin T, Malins LR, Vokits B, Shaw SA, Bao D-H, Wei F-L, Zhou T, Eastgate MD & Baran PS Decarboxylative alkenylation. Nature 545, 213–218 (2017). [PubMed: 28424520]
26. He J, Hamann LG, Davies HML & Beckwith REJ Late-stage C–H functionalization of complex alkaloids and drug molecules via intermolecular rhodium-carbenoid insertion. Nat. Commun 6, 5943 (2015). [PubMed: 25581471]
27. Yang J-M, Cai Y, Zhu S-F & Zhou Q-L Iron-catalyzed arylation of α-aryl-α-diazoesters. Org. Biomol. Chem 14, 5516–5519 (2016). [PubMed: 26805776]
28. Xu B, Li M-L, Zuo X-D, Zhu S-F & Zhou Q-L Catalytic asymmetric arylation of α-aryl-α-diazoacetates with aniline derivatives. J. Am. Chem. Soc 137, 8700–8703 (2015). [PubMed: 26121223]
29. Cardillo G, Gentilucci L, Qasem AR, Sgarzi F & Spampinato S Endomorphin-1 analogues containing β-Proline are μ-opioid receptor agonists and display enhanced enzymatic hydrolysis resistance. J. Med. Chem 45, 2571–2578 (2002). [PubMed: 12036366]
30. Sridharan V, Suryavanshi PA & Menéndez JC Advances in the chemistry of tetrahydroquinolines. Chem. Rev 111, 7157–7259 (2011). [PubMed: 21830756]

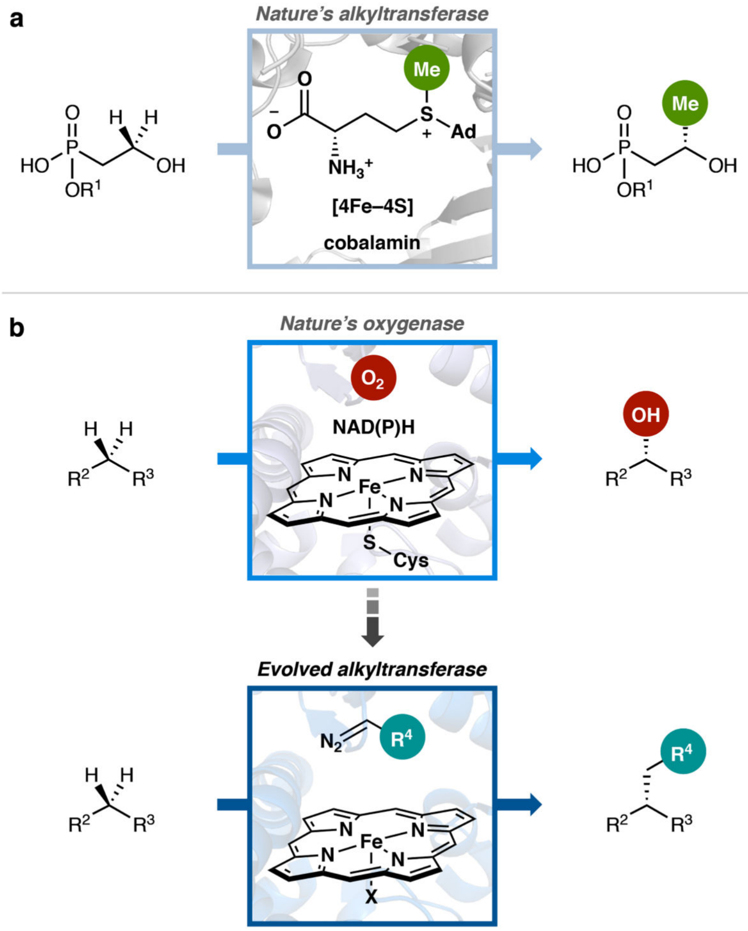
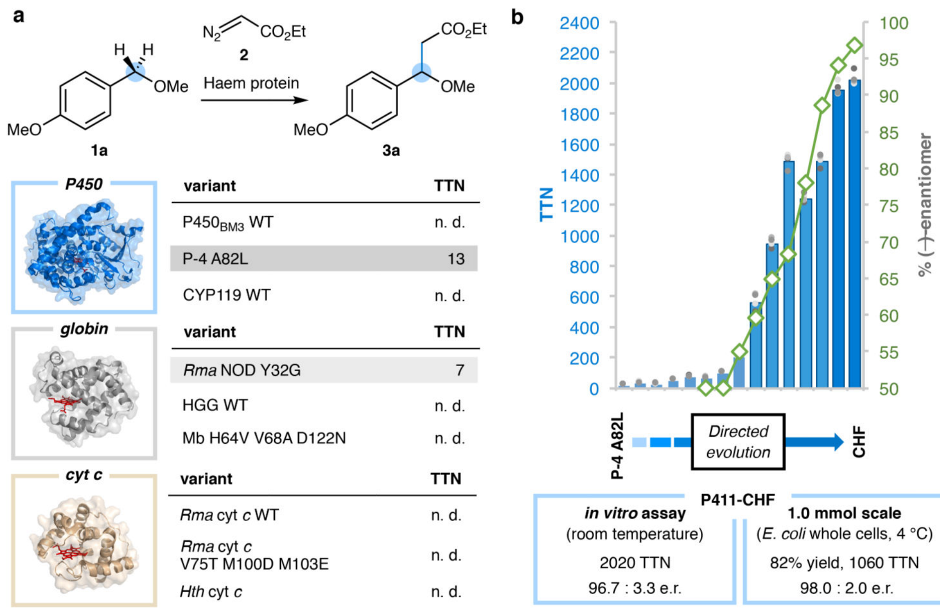


Figure 1 |. Enzymatic C–H functionalization systems.

a, Methylation catalysed by cobalamin-dependent radical SAM enzymes, as illustrated by Fom3 in fosfomycin biosynthesis[6]. **b,** Oxygenation catalysed by cytochrome P450 monooxygenase (top) and envisioned alkylation reaction achieved under haem protein catalysis (bottom). Structural illustrations are adapted from Protein Data Bank (PDB) ID code 5UL4 (radical SAM enzyme) and PDB 2IJ2 (cytochrome $P450_{BM3}$). Ad, adenosyl; Cys, cysteine; R, organic group; X, amino acid.



variant	TTN
P450$_{BM3}$ WT	n. d.
P-4 A82L	13
CYP119 WT	n. d.

variant	TTN
Rma NOD Y32G	7
HGG WT	n. d.
Mb H64V V68A D122N	n. d.

variant	TTN
Rma cyt *c* WT	n. d.
Rma cyt *c* V75T M100D M103E	n. d.
Hth cyt *c*	n. d.

Figure 2 |. Haem protein-catalysed *sp*^{3} C–H alkylation.

a, Select subset of haem proteins tested for promiscuous C–H alkylation activity. Structural illustrations are of representative superfamily members with the haem cofactor shown as red sticks: cytochrome P450$_{BM3}$ (PDB 2IJ2), sperm whale myoglobin (PDB 1A6K), and *Rma* cytochrome *c* (PDB 3CP5). TTN, total turnover number; n. d., not detected; WT, wild type; Mb, sperm whale myoglobin, HGG, Hell's Gate globin; cyt *c*, cytochrome *c*; *Hth*, *Hydrogenobacter thermophilus*. **b,** Directed evolution of a cytochrome P411 for enantioselective C–H alkylation (reaction shown in (**a**)). Bars represent average TTNs from reactions performed in quadruplicate; each TTN data point is shown as a grey dot. Enantioselectivity results are represented by green diamonds. Unless otherwise indicated, reaction conditions were haem protein in *E. coli* whole cells (OD_{600} = 30, (**a**)) or in clarified *E. coli* lysate (**b**), 10 mM substrate **1a**, 10 mM ethyl diazoacetate, 5 vol% EtOH in M9-N buffer at room temperature under anaerobic conditions for 18 hours. Reactions performed with lysate contain 1 mM $Na_2S_2O_4$. TTN is defined as the amount of indicated product divided by total haem protein as measured by the hemochrome assay. See Supplementary Information for the complete list of haem proteins tested and detailed experimental procedures.

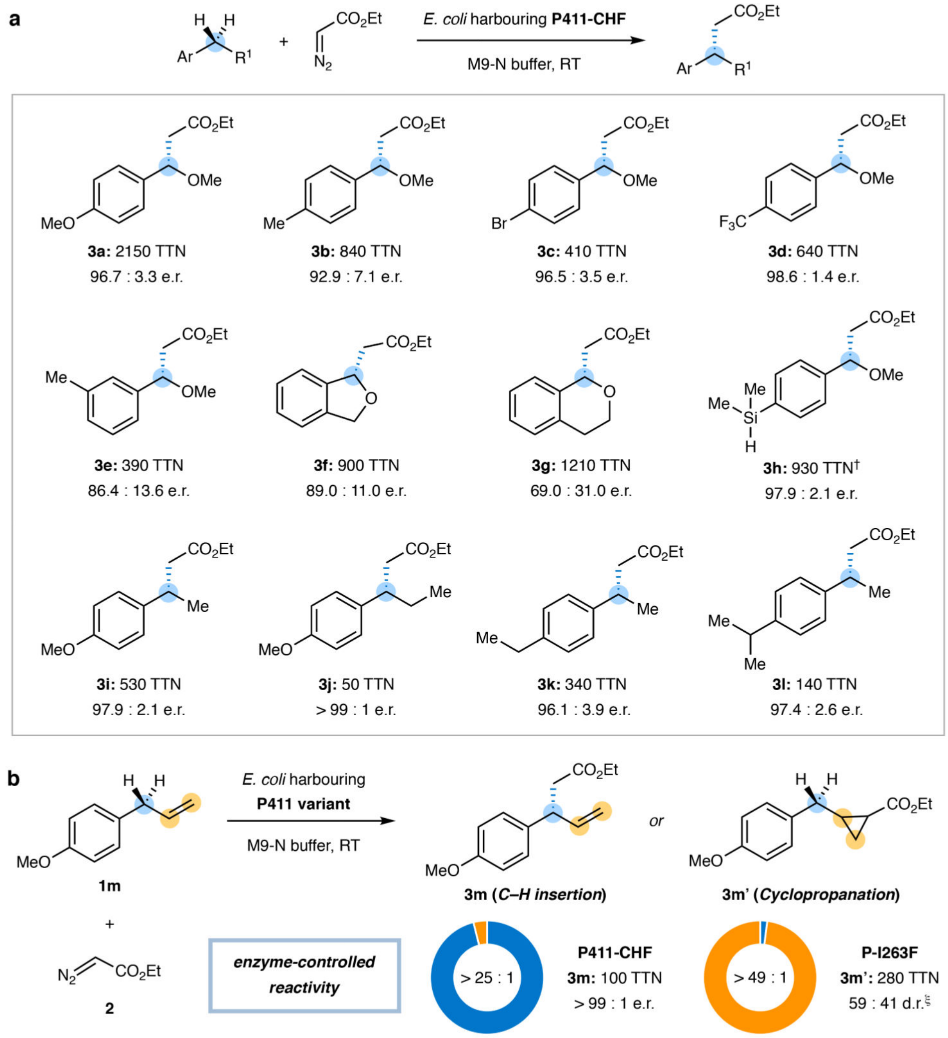


Figure 3 |. Substrate scope for benzylic C–H alkylation with P411-CHF.

a, Experiments were performed using *E. coli* expressing cytochrome P411-CHF (OD_{600} = 30) with 10 mM substrate **1a**–**1l** and 10 mM ethyl diazoacetate at room temperature (RT) under anaerobic conditions for 18 hours; each reported TTN is the average of quadruplicate reactions. See Supplementary Fig. S12 for the full list of alkane substrates. †Si–H insertion product **3h'** is also observed (Supplementary Fig. S7). **b,** Reaction selectivity for carbene C–H insertion or cyclopropanation can be controlled by the protein scaffold. Experiments were performed as in (**a**) using the indicated P411 variant. ξd.r. is given as *cis* : *trans*; e.r. was not determined.

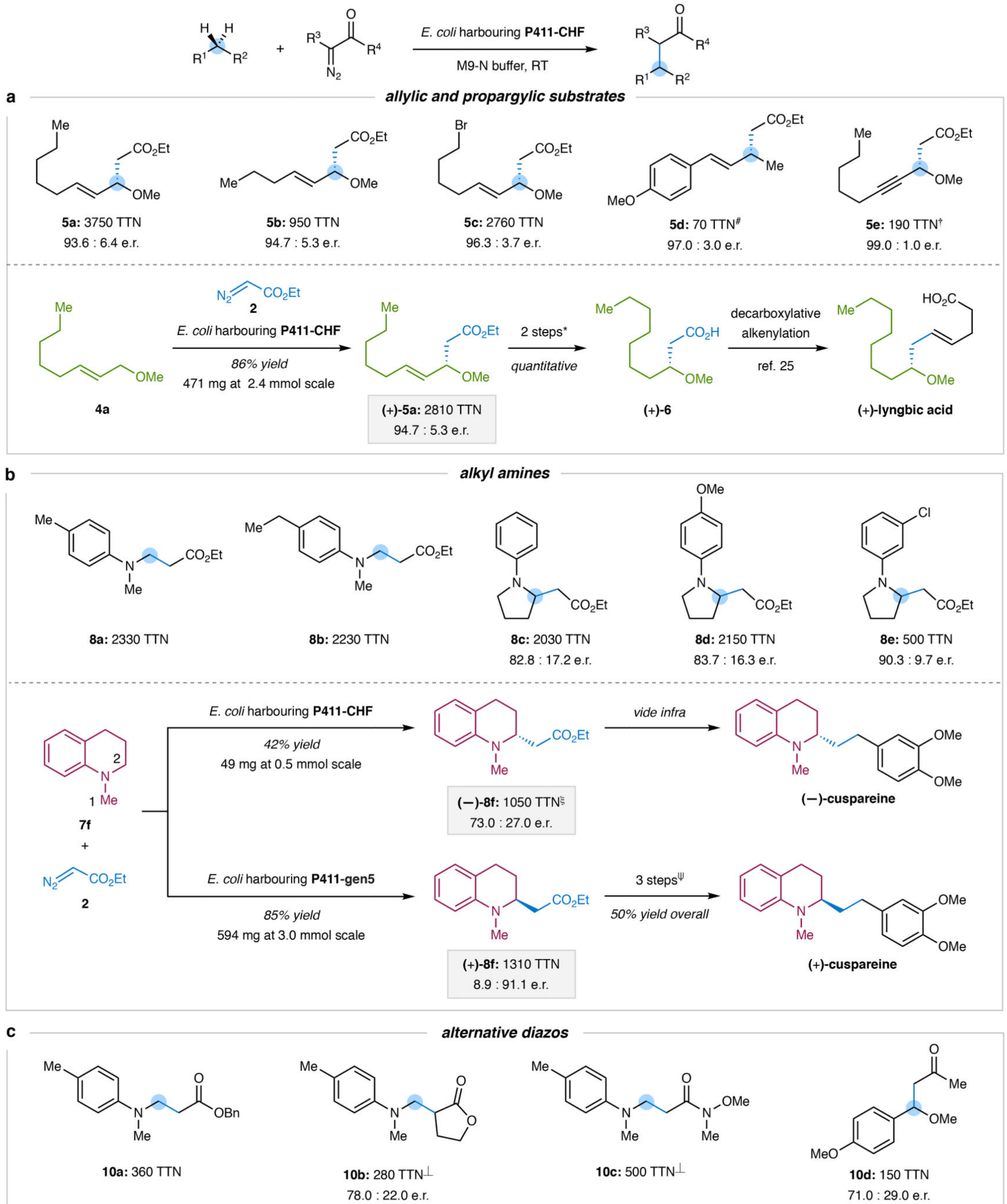


Figure 4 |. Application of P411 enzymes for sp^3 C–H alkylation.

a, Allylic and propargylic C–H alkylation. Unless otherwise indicated, experiments were performed using *E. coli* expressing cytochrome P411-CHF with 10 mM substrate **4a**–**4e** and 10 mM ethyl diazoacetate; each reported TTN is the average of quadruplicate reactions. #TTN was calculated based on isolated yield from a reaction performed at 0.25 mmol scale. †Cyclopropene product was also observed (Supplementary Fig. S8). *****Hydrogenation, followed by hydrolysis. **b,** Enzymatic alkylation of substrates containing α-amino C–H bonds. Unless otherwise indicated, experiments were performed at 0.5 mmol scale using *E.*

coli expressing cytochrome P411-CHF with substrates **7a–7f** and ethyl diazoacetate; TTNs were calculated based on isolated yields of products shown. [ξ]Isolated in 9 : 1 r.r. for **8f** : **8f’**. eReduction, halogen exchange, and Suzuki-Miyaura cross-coupling. **c,** Enzymatic C–H alkylation with alternative diazo reagents. Unless otherwise indicated, reactions were performed at 0.5 mmol scale using *E. coli* expressing cytochrome P411-CHF with coupling partner **1a** or **7a** and diazo compounds **9a–9d**; TTNs were calculated based on isolated yields of products shown. [⊥]Variant P411-IY T327I was used. See Supplementary Information for the complete list substrates (Fig. S12 and Fig. S13), information about enzyme variants, and full experimental details.