LETTER

https://doi.org/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] & Frances H. Arnold[1]*

Although 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 importance of and challenges associated with selective alkylation at a sp^3 carbon[1,2]. Here we describe iron-based catalysts for the enantio-, regio- and chemoselective intermolecular alkylation of sp^3 C–H bonds through carbene C–H insertion. The catalysts, derived from a cytochrome P450 enzyme in which the 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 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 excellent selectivity. Furthermore, they have enabled the development of concise routes to several natural products. The use of the native iron-haem cofactor of these enzymes to mediate sp^3 C–H alkylation suggests that diverse haem proteins could serve as potential catalysts for this abiological transformation, and will 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 the 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 is a versatile strategy for structural diversification, as seen by its essential role in the biosynthesis of structurally varied natural products and cofactors[4–6]. However, known biological machineries for this transformation are limited to enzymes that transfer a methyl group[5,6] or conjugate a radical acceptor substrate[4,7] to specific molecules, with methylation as a common mode of 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 the activation of molecular oxygen for the controlled generation of a high-energy iron-oxo intermediate that is 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). Although 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 is yet to be achieved.

Metal carbene sp^3 C–H insertion in small-molecule catalysis, in particular intermolecular and stereoselective versions of this reaction, typically relies on transition-metal complexes based on rhodium[12], iridium[13] and other metals[14–16]. Artificial metalloproteins for carbene C–H insertion have been created by introducing an iridium-porphyrin into variants of apo haem proteins[17]. Although rare, there are a few examples of iron carbene sp^3 C–H insertion. The iron-catalysed examples use high temperatures (for example 80 °C)[18], are stoichiometric[19] or are restricted to intramolecular reactions[20], which suggests there is a high activation energy barrier to the insertion of an iron carbene into a C–H bond. However, because the protein framework of an enzyme can impart substantial 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

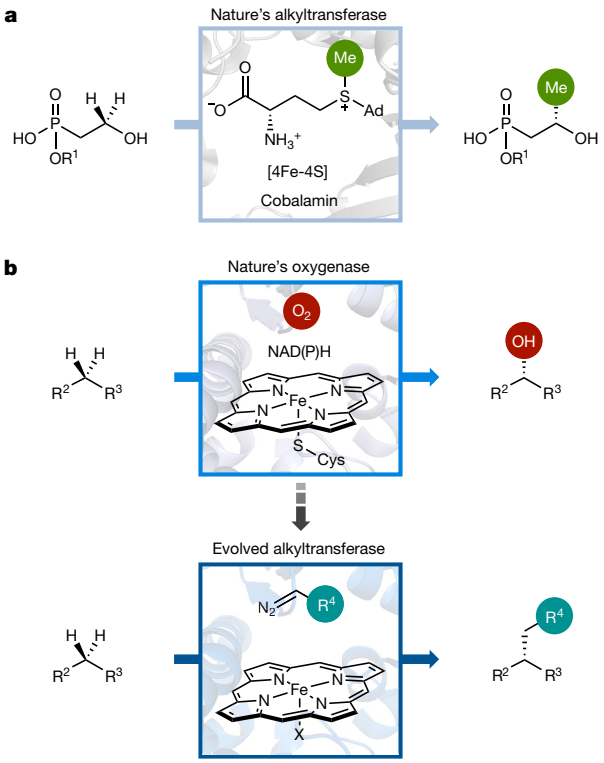

Fig. 1 | Enzymatic C–H functionalization systems. a, Methylation catalysed by cobalamin-dependent radical SAM enzymes, as illustrated by Fom3 in fosfomycin biosynthesis[6]. **b**, Oxygenation catalysed by cytochrome P450 monooxygenase (top) and the proposed alkylation reaction achieved under haem protein catalysis (bottom). Structural illustrations are adapted from Protein Data Bank (PDB) ID 5UL4 (radical SAM enzyme) and PDB 2IJ2 (cytochrome $P450_{BM3}$). Ad, adenosyl; Cys, cysteine; R, organic group; X, amino acid.

[1]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA. [2]Present address: Institute of Food Technology, University of Natural Resources and Life Sciences, Vienna, Austria. [3]Present address: Department of Chemistry, The Scripps Research Institute, Jupiter, FL, USA. *e-mail: frances@cheme.caltech.edu

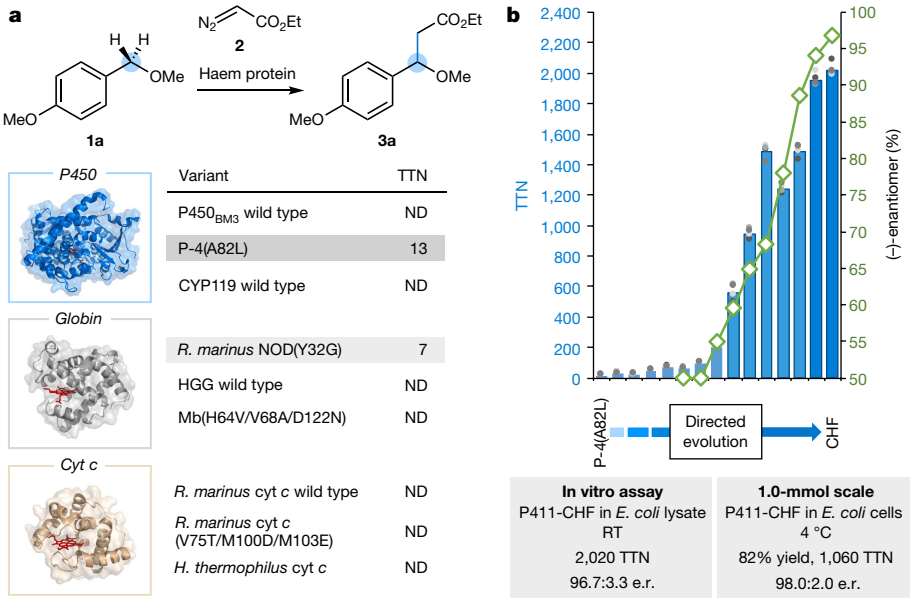

Fig. 2 | Haem-protein-catalysed sp^3 C–H alkylation. a, A selected subset of haem proteins that were tested for promiscuous C–H alkylation activity. Structural illustrations are of the following superfamily members with the haem cofactor shown as red sticks: cytochrome $P450_{BM3}$ (top, PDB 2IJ2), sperm whale myoglobin (middle, PDB 1A6K) and *R. marinus* cytochrome *c* (bottom, PDB 3CP5). cyt *c*, cytochrome *c*; HGG, Hell's Gate globin; *H. thermophilus*, *Hydrogenobacter thermophilus*; Mb, sperm whale myoglobin; ND, not detected. **b**, Directed evolution of a cytochrome P411 for enantioselective C–H alkylation (reaction shown in **a**). Bars represent mean TTN values averaged over four reactions (performed from two independent cell cultures, each used for duplicate reactions); each TTN data point is shown as a grey dot. Enantioselectivity data are represented by green diamonds. Unless otherwise indicated, reaction conditions were as follows: haem protein in *E. coli* whole cells (optical density at 600 nm, OD_{600}, of 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 (RT) under anaerobic conditions for 18 h; see Supplementary Information for conditions of the 1.0-mmol reaction in **b**. Reactions performed with lysate contain 1 mM $Na_2S_2O_4$. TTN is defined as the amount of indicated product divided by haem protein as measured by the haemochrome assay. See Supplementary Information for the complete list of haem proteins tested and detailed experimental procedures.

the iron carbene C–H insertion reaction and acquire this new function (Fig. 1b).

In initial studies, we tested a panel of 78 haem proteins that included variants of cytochromes P450, cytochromes *c* and globin homologues. The haem proteins in whole *Escherichia 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 the formation of C–H alkylation product **3a** (Fig. 2a; see Supplementary Information for the complete list of haem proteins tested). Haem proteins from two superfamilies were found to show low levels of this promiscuous activity, establishing the possibility of creating C–H alkylation enzymes with very different protein architectures. Among the proteins tested were variants of cytochrome $P450_{BM3}$ from *Bacillus megaterium* in which the axial cysteine ligand is substituted for serine, known as cytochrome P411s[31]. We found that one of these variants, P-4(A82L)[22], which differs from the wild type by 18 mutations, provided **3a** with a total turnover number (TTN) of 13. In addition, nitric oxide dioxygenase from *Rhodothermus marinus* with a tyrosine-to-glycine mutation at position 32 (*R. marinus* 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 2). The haem cofactor alone (iron protoporphyrin IX) or in the presence of bovine serum albumin was inactive (Supplementary Tables 1 and 2).

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 that 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. 1). 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 these enzymes. 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. Notably, removal of the FAD domain, which contains 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. 2). This indicates that the FAD domain may have (negative) allosteric effects on the 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 3). Eight additional rounds of mutagenesis and screening yielded P411-CHF (P411ΔFAD C–H functionalization enzyme; for the full list of changes, see Supplementary Information).

P411-CHF displays a 140-fold improvement in activity over P-4(A82L) and delivers **3a** with excellent stereoselectivity (2,020 TTN, 96.7: 3.3 enantiomeric ratio (e.r.) using clarified *E. coli* lysate). Subsequent studies showed that the stereoselectivity could be improved by conducting the reaction at lower temperature (for example, 4 °C) without substantial change to TTN (Supplementary Table 4). Enzymatic C–H alkylation can be performed on a millimole scale: using 1.0 mmol substrate **1a**, *E. coli* harbouring P411-CHF at 4 °C furnished **3a** in 82% isolated yield, 1,060 TTN, and 98.0: 2.0 e.r. (Fig. 2b). Preliminary mechanistic investigations were pursued to investigate 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 of 5.1 for C–H alkylation catalysed by P411-CHF, suggesting that C–H insertion is rate-determining (Supplementary Fig. 5).

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

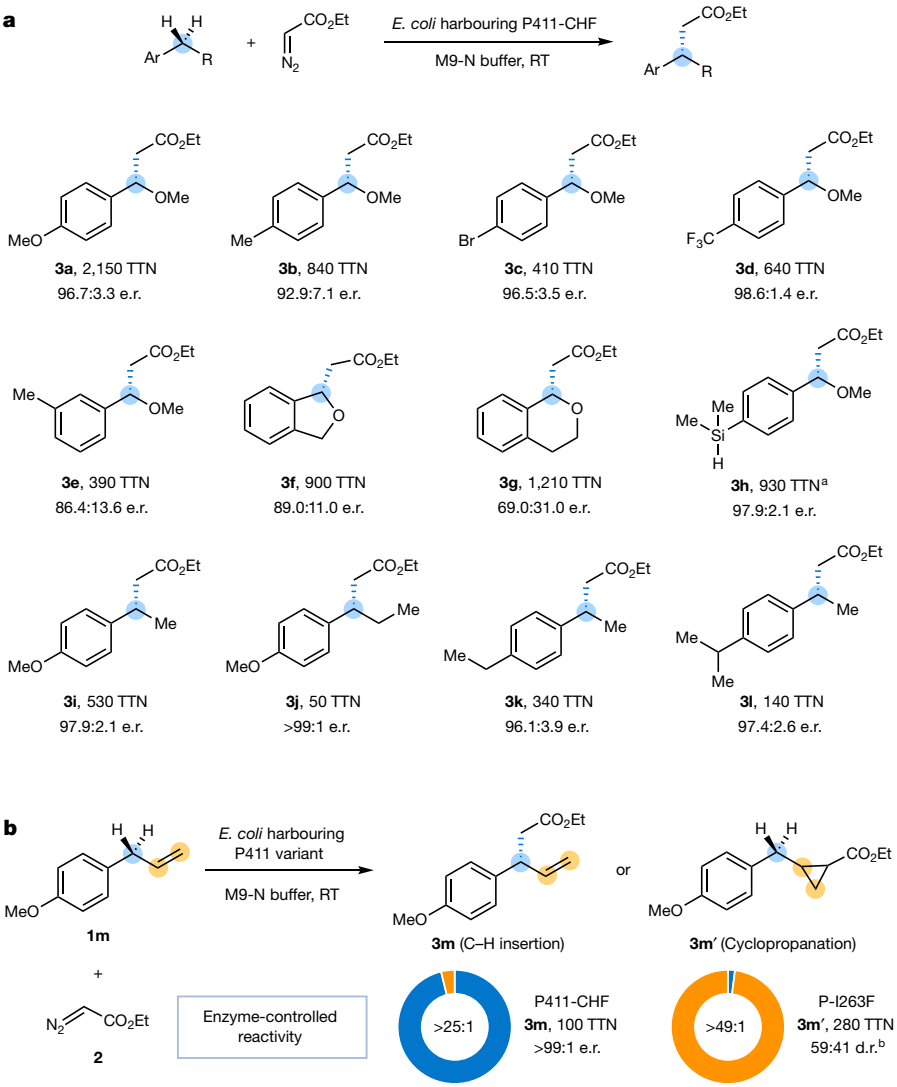

Fig. 3 | Substrate scope for benzylic C–H alkylation with P411-CHF. **a**, Experiments were performed using *E. coli* expressing cytochrome P411-CHF (OD_{600} = 30) with 10 mM substrate **1a**–**1l** and 10 mM ethyl diazoacetate at room temperature under anaerobic conditions for 18 h; reported TTNs are the average of four reactions (performed from two independent cell cultures, each used for duplicate reactions). See Supplementary Fig. 12 for the full list of alkane substrates. [a]Si–H insertion product **3h′** is also observed (Supplementary Fig. 7). **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. [b]Diastereomeric ratio (d.r.) is given as *cis*:*trans*; e.r. was not determined.

ring are well-tolerated (**3a**–**3e**, **3h**); cyclic substrates are also suitable coupling partners (**3f**, **3g**). The functionalization of alkyl benzenes at secondary benzylic sp^3 C–H bonds was found to be successful (**3i**–**3l**). Notably, in the biotransformation of substrate **1l**, which contains both tertiary and secondary benzylic C–H bonds, P411-CHF preferentially functionalizes the secondary position despite its higher C–H bond dissociation energy. The carbene intermediate derived from ethyl diazoacetate belongs to the acceptor-only class. Compared 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 does 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 proposed that the protein scaffold could be tuned to create complementary enzymes that can access different reaction outcomes available to a substrate. When P411-CHF was challenged with 4-allylanisole (**1m**)—a substrate that can undergo both C–H alkylation and cyclopropanation—we observed that C–H alkylation product **3m** dominated, with selectivity > 25:1 (Fig. 3b, Supplementary Fig. 6). By contrast, a related full-length P411 variant P-I263F, containing 13 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 carbenes[10], P411-CHF delivered C–H alkylation product **3h** when substrate **1h** was used in the reaction (Si–H insertion product **3h′** was also observed, but its formation may not be catalysed by P411-CHF, Supplementary Fig. 7). Reaction with P-I263F, by 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 the 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.

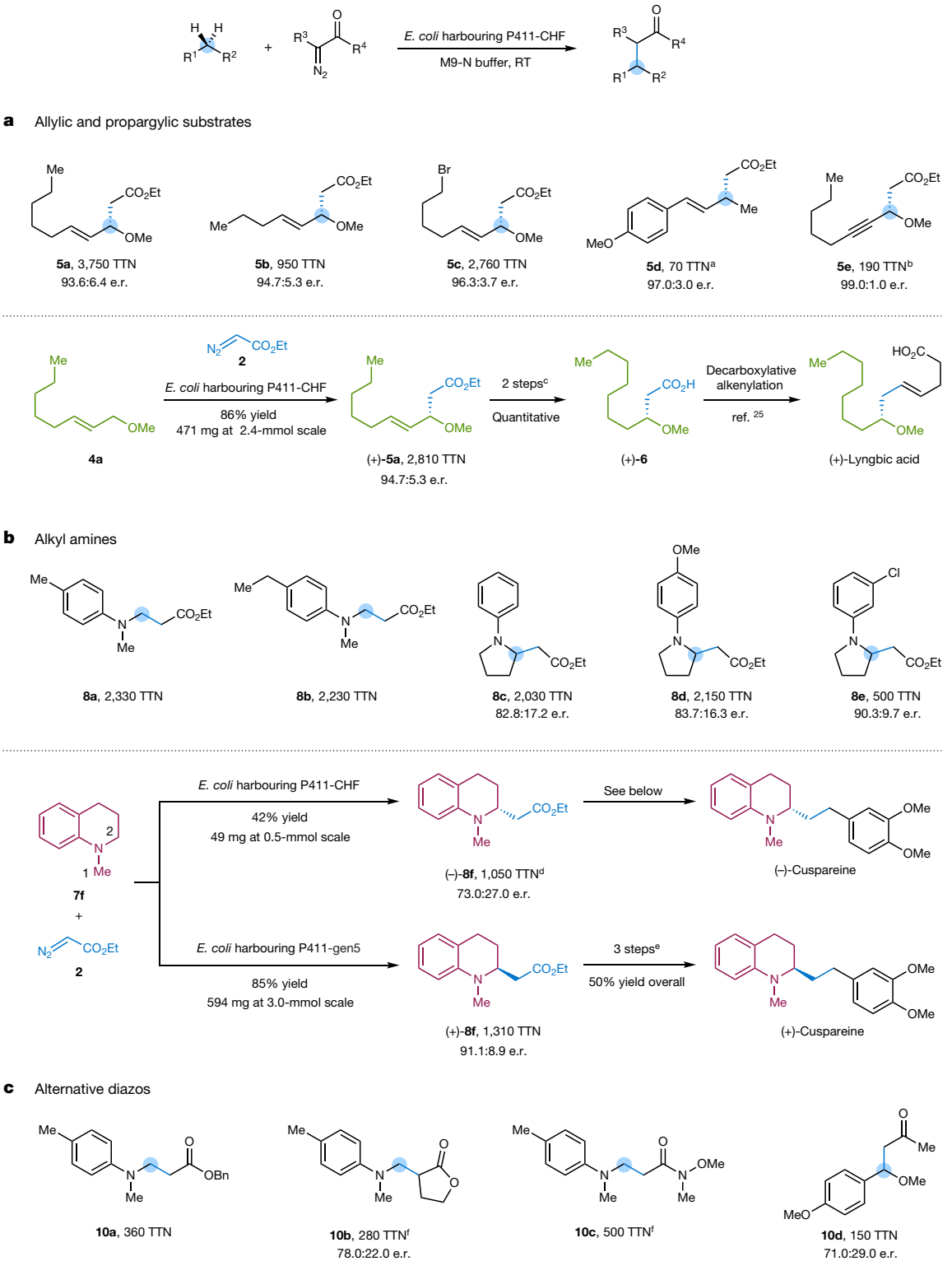

Fig. 4 | Application of P411 enzymes for sp^3 C–H alkylation. a, Allylic and propargylic C–H alkylation. Unless otherwise indicated, experiments were performed using *E. coli* expressing cytochrome P411-CHF with 10 mM substrate **4a**–**4e** and 10 mM ethyl diazoacetate; reported TTNs are the average of four reactions (performed from two independent cell cultures, each used for duplicate reactions). [a]TTN was calculated on the basis of isolated yield from a reaction performed on a 0.25-mmol scale. [b]The cyclopropene product was also observed (Supplementary Fig. 8). [c]Hydrogenation, followed by hydrolysis. **b**, Enzymatic alkylation of substrates containing α-amino C–H bonds. Unless otherwise indicated, reactions were performed on a 0.5-mmol scale using *E. coli* expressing cytochrome P411-CHF with substrates **7a**–**7f** and ethyl diazoacetate; TTNs were calculated on the basis of isolated yields of products shown. [d]Isolated in 9:1 r.r. for **8f**:**8f′** determined by ^{1}H nuclear magnetic resonance analysis; TTN is reported for the sum of the regioisomers. [e]Reduction, halogen exchange and Suzuki–Miyaura cross-coupling. **c**, Enzymatic C–H alkylation with alternative diazo reagents. Unless otherwise indicated, reactions were performed on a 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 on the basis of isolated yields of products shown. [f]Variant P411-IY(T327I) was used. See Supplementary Information for the complete list of substrates (Supplementary Figs. 12, 13), information about enzyme variants, and full experimental details.

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 malyngamides 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 a 2.4-mmol scale in 86% isolated yield, 2,810 TTN, and 94.7:5.3 e.r. Subsequent hydrogenation and hydrolysis provided (*R*)-(+)-**6** in quantitative yield, which can be transformed to (*R*)-(+)-lyngbic acid by decarboxylative alkenylation[25].

As part of our investigation into the substrate scope of the reaction, we challenged P411-CHF with alkyl amine compounds. Compounds of this type are typically challenging substrates for C–H functionalization methods because the amine functionality may coordinate to and inhibit the catalyst or undergo undesirable side reactions (for example, ylide formation and its associated rearrangements)[26]. Using **7a** or **7b**, substrates that 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 either not observed (with **7a**, Supplementary Fig. 9) or was considerably suppressed (with **7b**, Supplementary Fig. 10), despite the typically lower bond dissociation energies of benzylic C–H bonds compared to α-amino C–H bonds. Additionally, *N*-aryl pyrrolidines (**7c**–**7e**) were found to be 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 that has been investigated for medicinal chemistry applications[29].

Given that P411-CHF alkylates both primary and secondary α-amino C–H bonds, we investigated whether the enzyme could be selective for one of these positions. Using *N*-methyl tetrahydroquinoline **7f** as the alkane substrate, P411-CHF afforded β-amino ester products with 1,050 TTN and a 9:1 ratio of regioisomers (C2:C1, and 73.0:27.0 e.r. for (−)-**8f**) (Fig. 4b). The tetrahydroquinoline ring is a prevalent structural motif in natural products and bioactive molecules[30], and its selective functionalization could provide a concise strategy for the synthesis of alkaloids. To improve the selectivity for the alkylation of **7f**, we tested variants along the evolutionary lineage from P-4(A82L) to P411-CHF. We found that, compared with P411-CHF, P411-gen5 showed even better regioselectivity and the opposite stereochemical preference for C–C bond formation. In a reaction on 3.0-mmol scale, *E. coli* harbouring P411-gen5 delivered (+)-**8f** in 85% yield with excellent selectivity (1,310 TTN, >50:1 regiomeric ratio (r.r.), 91.1:8.9 e.r.). In only a few steps, the enzymatic product was successfully transformed to the alkaloid (*R*)-(+)-cuspareine[30] (Fig. 4b).

Finally, we explored 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. 11). 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 the current related enzymes. With the exception of **10b**, reactions using disubstituted carbene reagents did not yield appreciable amounts of desired products (Supplementary Fig. 11).

This study demonstrates that a cytochrome P450 can acquire the ability to construct C–C bonds from sp^3 C–H bonds, and that the activity and selectivity of the reaction 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 foresee 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.

Data availability

All relevant data are provided in Supplementary Information. Any additional information is available from the corresponding author on request.

Received: 18 July 2018; Accepted: 1 November 2018;
Published online: 19 December 2018

1. Hartwig, J. F. & Larsen, M. A. Undirected, homogeneous C–H bond functionalization: challenges and opportunities. *ACS Cent. Sci.* **2**, 281–292 (2016).
2. Saint-Denis, T. G., Zhu, R.-Y., Chen, G., Wu, Q.-F. & Yu, J.-Q. Enantioselective C(sp^3)–H bond activation by chiral transition metal catalysts. *Science* **359**, eaao4798 (2018).
3. Frey, P. A. & Hegeman, A. D. *Enzymatic Reaction Mechanisms* Ch. 14 (Oxford Univ. Press, New York, 2007).
4. Yokoyama, K. & Lilla, E. A. 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).
5. Bauerle, M. R., Schwalm, E. L. & Booker, S. J. Mechanistic diversity of radical *S*-adenosylmethionine (SAM)-dependent methylation. *J. Biol. Chem.* **290**, 3995–4002 (2015).
6. McLaughlin, M. I. & van der Donk, W. A. Stereospecific radical-mediated B_{12}-dependent methyl transfer by the fosfomycin biosynthesis enzyme Fom3. *Biochemistry* **57**, 4967–4971 (2018).
7. Shisler, K. A. & Broderick, J. B. Glycyl radical activating enzymes: structure, mechanism, and substrate interactions. *Arch. Biochem. Biophys.* **546**, 64–71 (2014).
8. Poulos, T. L. Heme enzyme structure and function. *Chem. Rev.* **114**, 3919–3962 (2014).
9. Coelho, P. S., Brustad, E. M., Kannan, A. & Arnold, F. H. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science* **339**, 307–310 (2013).
10. 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* **354**, 1048–1051 (2016).
11. Brandenberg, O. F., Fasan, R. & Arnold, F. H. Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions. *Curr. Opin. Biotechnol.* **47**, 102–111 (2017).
12. Liao, K. et al. Site-selective and stereoselective functionalization of non-activated tertiary C–H bonds. *Nature* **551**, 609–613 (2017).
13. Weldy, N. M. et al. Iridium(III)-bis(imidazolinyl)phenyl catalysts for enantioselective C–H functionalization with ethyl diazoacetate. *Chem. Sci.* **7**, 3142–3146 (2016).
14. Wang, Y., Wen, X., Cui, X. & Zhang, X. P. Enantioselective radical cyclization for construction of 5-membered ring structures by metalloradical C–H alkylation. *J. Am. Chem. Soc.* **140**, 4792–4796 (2018).
15. Caballero, A. et al. Silver-catalyzed C–C bond formation between methane and ethyl diazoacetate in supercritical CO_2. *Science* **332**, 835–838 (2011).
16. Wu, W.-T., Yang, Z.-P. & You, S.-L. in *Asymmetric Functionalization of C–H Bonds* (ed. You, S.-L.) Ch. 1 (Royal Society of Chemistry, Cambridge, 2015).
17. Key, H. M., Dydio, P., Clark, D. S. & Hartwig, J. F. Abiological catalysis by artificial haem proteins containing noble metals in place of iron. *Nature* **534**, 534–537 (2016).
18. Mbuvi, H. M. & Woo, L. K. 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).
20. Griffin, J. R., Wendell, C. I., Garwin, J. A. & White, M. C. Catalytic C(sp^3)–H alkylation via an iron carbene intermediate. *J. Am. Chem. Soc.* **139**, 13624–13627 (2017).
21. Herschlag, D. & Natarajan, A. Fundamental challenges in mechanistic enzymology: progress toward understanding the rate enhancements of enzymes. *Biochemistry* **52**, 2050–2067 (2013).
22. Prier, C. K., Zhang, R. K., Buller, A. R., Brinkmann-Chen, S. & Arnold, F. H. Enantioselective, intermolecular benzylic C–H amination catalysed by an engineered iron-haem enzyme. *Nat. Chem.* **9**, 629–634 (2017).
23. Renata, H. et al. Identification of mechanism-based inactivation in P450-catalyzed cyclopropanation facilitates engineering of improved enzymes. *J. Am. Chem. Soc.* **138**, 12527–12533 (2016).
24. Zhang, J.-T. et al. Total synthesis of malyngamides K, L, and 5″-*epi*-C and absolute configuration of malyngamide L. *J. Org. Chem.* **76**, 3946–3959 (2011).
25. Edwards, J. T. et al. Decarboxylative alkenylation. *Nature* **545**, 213–218 (2017).
26. He, J., Hamann, L. G., Davies, H. M. L. & Beckwith, R. E. J. Late-stage C–H functionalization of complex alkaloids and drug molecules via intermolecular rhodium-carbenoid insertion. *Nat. Commun.* **6**, 5943 (2015).
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).

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).
29. Cardillo, G., Gentilucci, L., Qasem, A. R., 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).
30. Sridharan, V., Suryavanshi, P. A. & Menéndez, J. C. Advances in the chemistry of tetrahydroquinolines. *Chem. Rev.* **111**, 7157–7259 (2011).
31. Coelho, P. S. et al. A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins *in vivo*. *Nat. Chem. Biol.* **9**, 485–487 (2013).

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 discussions; S. Virgil, M. Shahgholi and D. VanderVelde for analytical support; and B. Stoltz for use of polarimeter and gas chromatography equipment.

Reviewer information *Nature* thanks B. de Bruin, N. Turner and T. Ward for their contribution to the peer review of this work.

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.

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

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0808-5.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to F.H.A.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.