doi:10.1038/nature24996

Genetically programmed chiral organoborane synthesis

S. B. Jennifer Kan[1]*, Xiongyi Huang[1]*, Yosephine Gumulya[1], Kai Chen[1] & Frances H. Arnold[1]

Recent advances in enzyme engineering and design have expanded nature's catalytic repertoire to functions that are new to biology[1–3]. However, only a subset of these engineered enzymes can function in living systems[4–7]. Finding enzymatic pathways that form chemical bonds that are not found in biology is particularly difficult in the cellular environment, as this depends on the discovery not only of new enzyme activities, but also of reagents that are both sufficiently reactive for the desired transformation and stable *in vivo*. Here we report the discovery, evolution and generalization of a fully genetically encoded platform for producing chiral organoboranes in bacteria. *Escherichia coli* cells harbouring wild-type cytochrome *c* from *Rhodothermus marinus*[8] (*Rma* cyt *c*) were found to form carbon–boron bonds in the presence of borane–Lewis base complexes, through carbene insertion into boron–hydrogen bonds. Directed evolution of *Rma* cyt *c* in the bacterial catalyst provided access to 16 novel chiral organoboranes. The catalyst is suitable for gram-scale biosynthesis, providing up to 15,300 turnovers, a turnover frequency of 6,100 h^{-1}, a 99:1 enantiomeric ratio and 100% chemoselectivity. The enantiopreference of the biocatalyst could also be tuned to provide either enantiomer of the organoborane products. Evolved in the context of whole-cell catalysts, the proteins were more active in the whole-cell system than in purified forms. This study establishes a DNA-encoded and readily engineered bacterial platform for borylation; engineering can be accomplished at a pace that rivals the development of chemical synthetic methods, with the ability to achieve turnovers that are two orders of magnitude (over 400-fold) greater than those of known chiral catalysts for the same class of transformation[9–11]. This tunable method for manipulating boron in cells could expand the scope of boron chemistry in living systems.

Boron-containing natural products are synthesized in the soil by the myxobacterium *Sorangium cellulosum* as antibiotics against Gram-positive bacteria[12]. In the sea, these molecules give the Jurassic red alga *Solenopora jurassica* its distinct pink colouration[13]; they are also produced by the bioluminescent bacterium *Vibrio harveyi* for cell–cell communication[14] (Extended Data Fig. 1). To prepare boron-containing biomolecules, living organisms produce small molecules that spontaneously react with boric acid available in the environment[15,16]. Although this non-enzymatic method for capturing boron is sufficient for the survival of an organism, it is limited by the inherent affinity of a substrate towards boric acid, and lacks tunability and generality for synthetic biology applications. Moreover, organisms that produce organoboranes (compounds that contain carbon–boron bonds) are unknown.

We envisioned that enzyme-catalysed borylation could provide living organisms with the ability to produce boron-containing products tailored to our needs. Such an enzyme is not known in nature, but we hypothesized that existing natural proteins might be repurposed and engineered to perform this task. In the past, we and others have exploited the promiscuity of natural and engineered haem proteins for various non-natural reactions[4,6,7,17]. The resulting enzymes are fully genetically encoded and carry out their synthetic functions in their bacterial expression hosts. Here, we focused on introducing boron motifs to organic molecules enantioselectively, as this would generate boron-containing carbon stereocentres, which are important structural features in functional organoboranes such as the US Food and Drug Administration (FDA)-approved chemotherapeutics Velcade and Ninlaro[18]. They are also versatile precursors for chemical derivatization through stereospecific carbon–boron to carbon–carbon or carbon–heteroatom bond conversion[19–21].

Although boron reagents applicable for carbon–boron bond formation in water are known[22,23], their biocompatibility, cell permeability, stability and reactivity in living systems, which contain an abundance of biomolecules, nucleic acids and metal ions, are uncertain. Nevertheless, as boron reagents designed for *in vivo* chemical biology applications are precedented[24–26], we reasoned that reagents suitable for biological borylation could be found. We identified borane–Lewis base complexes as potential candidates owing to their aqueous stability and reactivity towards carbenoid B–H insertion[9–11,27] (Extended Data Fig. 2), a mechanistic pathway we thought could be adapted for use in the biological environment owing to its orthogonality to the existing biochemistry of living systems.

We first set out to assess whether biological organoborane production might be feasible in a bacterial cell. *E. coli* BL21(DE3) cells harbouring wild-type cytochrome *c* from the Gram-negative, thermohalophilic bacterium *Rhodothermus marinus*[8] (*Rma* cyt *c*) were incubated with *N*-heterocyclic carbene borane[28,29] (NHC-borane) **1** and ethyl 2-diazopropanoate (Me-EDA) **2** in neutral buffer (M9-N minimal medium, pH 7.4). After incubation at room temperature, *in vivo* production of organoborane **3** was observed, with 120 turnovers (calculated with respect to the concentration of *Rma* cyt *c* expressed in *E. coli*, Fig. 1a, b) and an enantiomeric ratio (e.r.) of 85:15 (*R*/*S* isomer = 6, Fig. 1c). Because the pET22b/pEC86 expression system translocates *Rma* cyt *c* to the *E. coli* periplasm for post-translational maturation (during which the haem cofactor is covalently ligated to the cyt *c* apoprotein)[30], we assumed that borylation takes place in the periplasmic compartment. In the absence of *Rma* cyt *c*, *E. coli* yielded only a trace amount of borylation product with very low stereoselectivity (Extended Data Table 1). Both the substrates and the organoborane product were stable under these conditions. The haem cofactor alone could also promote the borylation reaction, although with no stereoselectivity. Other cytochrome *c* proteins, cytochromes P450, and globins also demonstrated carbon–boron bond-forming ability, but their selectivities were unsatisfactory (Extended Data Table 1).

To improve the performance of this whole-cell catalyst, we subjected the wild-type *Rma* cyt *c* (hereafter referred to as BOR^{WT}) to site-saturation mutagenesis, sequentially targeting active-site amino acid residues M100, V75 and M103, which are closest to the haem iron in BOR^{WT} (within 7 Å, Fig. 1d). Each single-site site-saturation mutagenesis library was cloned using the 22c-trick method[31], screened

[1]Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 East California Boulevard, MC 210-41, Pasadena, California 91125, USA.
*These authors contributed equally to this work.

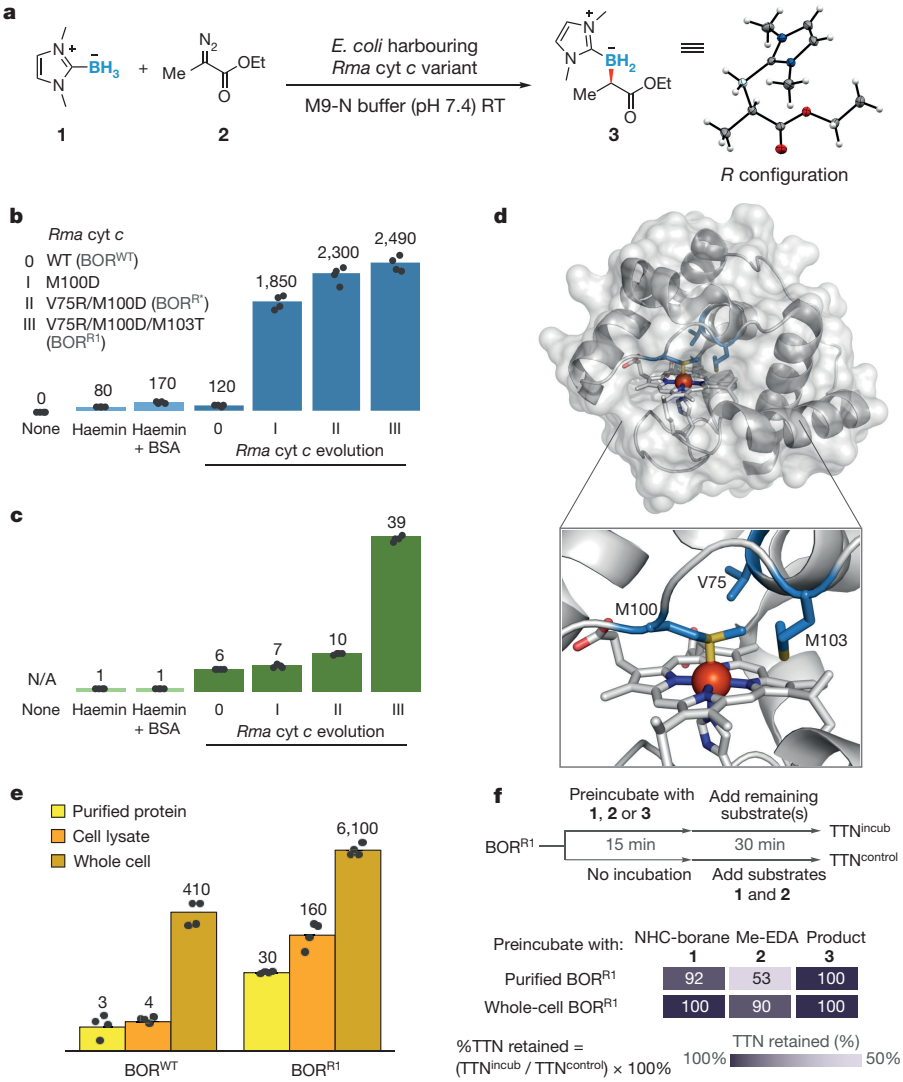


Figure 1 | Discovery, evolution and characterization of a bacterial catalyst for borylation. a, Reaction scheme shows a representative *in vivo* borylation reaction between NHC-borane **1** and diazo ester **2** to yield organoborane **3**. Standard substrate loading is 10 mM for both **1** and **2**. The absolute configuration of biosynthesized **3** was assigned as *R* by X-ray crystallography. **b**, **c**, Sequential site-saturation mutagenesis of *Rma* cyt *c* targeting active-site amino acid residues M100, V75 and M103 improved the turnover (**b**) and enantioselectivity (**c**) of bacterial production of organoborane **3**. Whole-cell *Rma* cyt *c* variants were compared using *E. coli* cells with an optical density at 600 nm (OD_{600}) of 15. Total turnover numbers (TTNs) were calculated with respect to the concentration of *Rma* cyt *c* expressed in *E. coli* and represent the total number of turnovers obtained from the catalyst under the stated reaction conditions. **d**, X-ray crystal structure of wild-type *Rma* cyt *c* (PDB: 3CP5). **e**, Turnover frequencies (h^{-1}, on log scale) of BOR^{WT} and BOR^{R1} as whole-cell catalysts, cell lysates, or purified proteins for the production of organoborane **3**. **f**, Purified and whole-cell BOR^{R1} were preincubated with NHC-borane **1**, Me-EDA **2**, or organoborane **3** before they were used as borylation catalysts to determine the inactivation effects of **1**–**3**. The numbers shown represent the %TTN retained after preincubation, and are relative to a control (no incubation) of the same type of catalyst (purified protein or whole cell). Bars and numbers above bars represent mean values averaged over four biological replicates. Individual data points are shown as overlays. BSA, bovine serum albumin; NHC, *N*-heterocyclic carbene.

as whole-cell catalysts in 96-well plates for improved borylation enantioselectivity, and the best variant was used to parent the next round of mutation and screening. With a single mutation M100D replacing the distal axial ligand, the first-generation biocatalyst exhibited a 16-fold improvement in turnover compared with the wild type (total turnover number (TTN) 1,850, Fig. 1b), with 88:12 e.r. (*R*/*S* isomer = 7; Fig. 1c). The M100D mutation also substantially improved carbene transfer reactivity for Si–H insertion catalysed by *Rma* cyt *c*[6]. This improvement in catalytic performance is probably due to the removal of the axial ligand from the haem iron, which opens a site primed for iron carbenoid formation and subsequent product formation[32]. Two subsequent rounds of mutagenesis and screening led to variant BOR^{R1} (V75R M100D M103T), which exhibited a turnover of 2,490 and an e.r. of 97.5:2.5 (*R*/*S* isomer = 39). This genetically programmed biological function is readily scalable from an analytical to a millimolar scale, with 0.5 mmol substrates, BOR^{R1} produced organoborane **3** in 97.5:2.5 e.r. and 75% isolated yield, with a TTN of 3,000. The absolute configuration of product **3** was unambiguously assigned as *R* by X-ray crystallography.

With an excellent borylating bacterium in hand, the properties and potential of the system were assessed. We characterized the initial rates of *in vivo* borylation and found that screening for improved enantioselectivity also led to an overall rate enhancement: whole-cell BOR^{R1} is 15 times faster than BOR^{WT}, with a turnover frequency of 6,100 h^{-1}. Notably, as purified protein or in cell lysate, both BOR^{R1} and BOR^{WT} are orders of magnitude slower than *in vivo* (Fig. 1e). When isolated BOR^{R1} protein and whole-cell BOR^{R1} were preincubated with Me-EDA **2** before the borylation reaction, the isolated

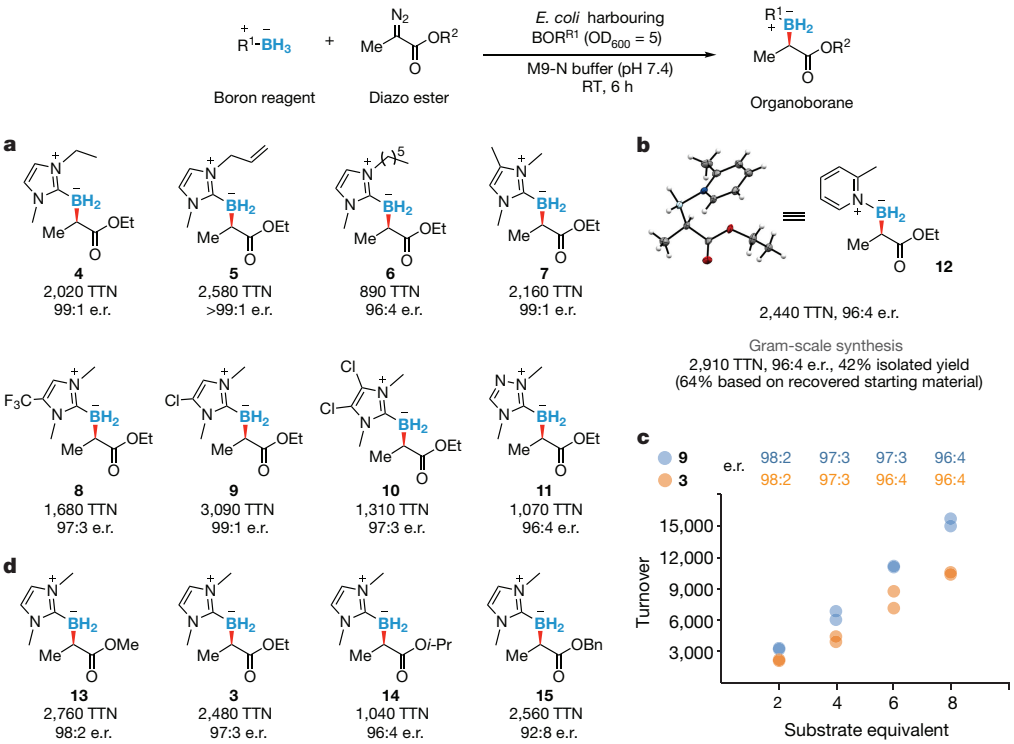


Figure 2 | Scope of chiral organoborane production in *E. coli*. **a**, **d**, Scope of boron reagent (**a**) and diazo ester (**d**) for borylation catalysed by *E. coli* harbouring BOR^{R1}. Standard substrate loading is 10 mM for both substrates. Reactions conducted in duplicate. **b**, Gram-scale synthesis (8.4 mmol) of organoborane **12** catalysed by whole-cell BOR^{R1} ($OD_{600} = 30$). The small scale preparation of **12** (2,440 TTN, 96:4 e.r.) is also reported for comparison. The absolute configuration of biosynthesized **12** was assigned as *R* by X-ray crystallography. **c**, Biosynthesis of organoboranes **9** (blue) and **3** (orange) catalysed by whole-cell BOR^{R1} ($OD_{600} = 15$). One substrate equivalent (8 mM final concentration of boron reagent and diazo ester) was added to the reaction every 75 min. Reactions conducted in biological duplicate. Bn, benzyl; e.r., enantiomeric ratio.

protein retained only around 50% of its activity, whereas whole-cell BOR^{R1} retained greater than 90% activity (Fig. 1f). NHC-borane **1** and organoborane product **3** did not inactivate the enzyme. Me-EDA probably inactivates BOR^{R1} through carbene transfer to the haem cofactor and/or the nucleophilic side chains of the protein, a mechanism we studied previously in detail for a cytochrome P450-based carbene transferase[33]. The intact periplasm apparently protects BOR^{R1} from inactivation by Me-EDA, and carbene transfer to yield the organoborane product is generally faster than protein inactivation pathway(s) under those conditions. Similar observations have been reported for other protein-based carbene transfer reaction systems[7,34]. Analysis of colony-forming units shows that *in vivo* organoborane production does not markedly reduce the viability of the *E. coli* (Extended Data Fig. 3).

Next, we explored the scope of boron reagents that could function in the cellular environment. Ten boron reagents were tested under turnover-optimized conditions: although the size, solubility and lipophilicity of these reagents varied, all were found to permeate the cell membrane and give the desired products with excellent selectivities and turnovers (Fig. 2a). Various substitutions on the NHC nitrogen are tolerated (**3**–**10**). The reaction is chemoselective in the presence of terminal olefins (**5**), which could function as a reaction handle suitable for downstream biological or bio-orthogonal derivatization. Sterically more demanding tetra- and penta-substituted NHCs are also accepted (**7**–**10**). As well as imidazole-based boron reagents, triazolylidene borane and picoline borane could also be used for *in vivo* borylation, yielding products **11** and **12** in 1,070 TTN and 2,440 TTN, respectively, with uniformly high selectivities (96:4 e.r.). On the gram scale, *in vivo* borylation produced 740 mg of picoline organoborane **12** with 2,910 TTN, 96:4 e.r. and 42% isolated yield (64% based on recovered starting material, Fig. 2b). The absolute configuration of **12** was assigned as *R* by X-ray crystallography. When substrates were added portion-wise at regular time intervals to *E. coli* expressing BOR^{R1} (we tested the sequential addition of up to eight equivalents of substrates over a period of 12 h, Fig. 2c; Extended Data Table 2), organoborane **3** was produced with 10,400 turnovers (50% yield, 96:4 e.r.), whereas organoborane **9** was obtained with 15,300 turnovers (73% yield, 96:4 e.r.). No substantial loss in activity or enantioselectivity was observed, demonstrating the potential of this bacterial catalyst for biosynthesis and incorporation into natural or engineered metabolic pathways.

Systematic modification of the diazo ester substituents from Et to Me, *i*-Pr or Bn revealed that the borylation ability of BOR^{R1} is not limited to Me-EDA (**3**, **13**–**15**, Fig. 2d). The relative insensitivity of the protein to the steric bulk of the ester might indicate that, in the putative iron carbenoid intermediate, this moiety is solvent-exposed rather than embedded within the active site. By re-randomizing the 103 position in BOR^{R1}, a residue we thought might modulate loop dynamics for improved binding of this substrate, the borylation turnover of **15** improved (from 2,560 to 4,200 TTN) using the triple-mutant V75R/M100D/M103D (BOR^{R2}, Fig. 3a). From the same site-saturation library, a borylation catalyst for trifluoromethyl-substituted diazo ester (CF_3-EDA) was also discovered (V75R/M100D/M103F, BOR^{R3}). Acceptor–acceptor diazo reagents such as CF_3-EDA are less reactive towards carbenoid formation because of their electron-deficient nature and have not been used before this for enzymatic carbene-transfer reactions. The present system tolerates this class of substrates and yielded product **16** with 95:5 e.r. and 1,560 TTN.

To further broaden the generality of this borylation platform, we re-examined the evolutionary landscape from BOR^{WT} to BOR^{R1} to search for promiscuous mutants that might unlock new reactivities. Double mutant V75P/M100D (BOR^{P*}) stood out as highly productive but poorly selective (69:31 e.r.) for Me-EDA borylation in the M100D/V75X site-saturation library. As proline-mediated helix kinks are known to induce structural and dynamic changes to proteins,

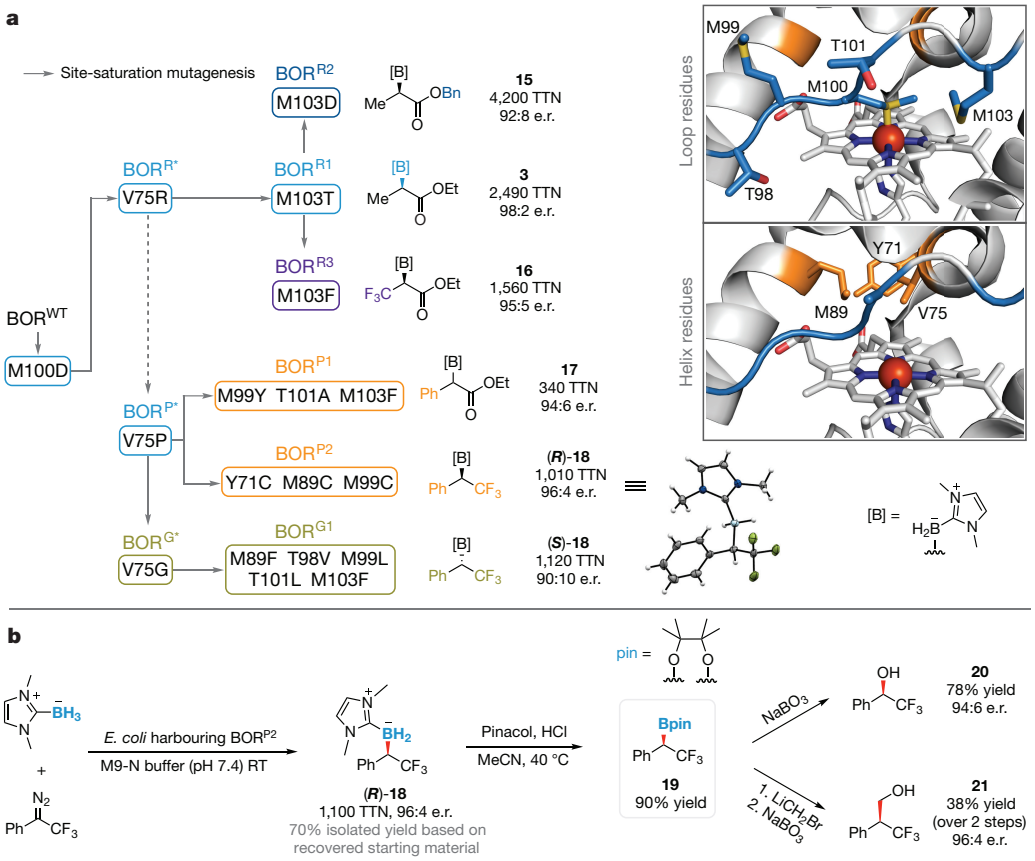


Figure 3 | Expanding the generality and utility of biological borylation. **a**, The generality of *in vivo* borylation was expanded through directed evolution to accommodate bulky substrates (**15**, **17**) and less reactive acceptor–acceptor diazo reagents (**16**), to move beyond diazo ester-based substrates ((***R***)-**18**, (***S***)-**18**), and to provide either enantiomer of the organoborane products ((***R***)-**18**, (***S***)-**18**). Reactions conducted in biological quadruplicate. Solid arrows represent site-saturation mutagenesis studies. BOR^{P*} was discovered in the M100D V75X site-saturation mutagenesis library for Me-EDA borylation. Amino acid residues targeted during directed evolution are depicted in the X-ray crystal structure of wild-type *Rma* cyt *c* (PDB: 3CP5). The absolute configuration of biosynthesized (***R***)-**18** was assigned by X-ray crystallography. **b**, Derivatization of biocatalytic product. Organoborane (***R***)-**18** was biosynthesized with *E. coli* harbouring BOR^{P2} ($OD_{600} = 30$) on a 1.3-mmol scale in 40% isolated yield (70% based on recovered starting material) for derivatization studies. Conversion to pinacol borane **19** was achieved with retention of the stereogenic carbon centre (stereoselectivity determined after derivatization to alcohol **20**). The yield reported for **19** was determined by ^{19}F NMR. We demonstrated the stereospecific transformation of **19** to alcohol **20** and Matteson homologation–oxidation product **21**.

we asked whether the V75P mutation might provide access to a unique reaction space. Ethyl 2-diazophenylacetate (Ph-EDA) is a bulky donor–acceptor diazo reagent inactive towards BOR^{WT}, but when added to *E. coli* harbouring BOR^{P*} with NHC-borane **1**, Ph-EDA was transformed to organoborane **17** in 100 TTN and 75:25 e.r. (Fig. 3a). By accumulating three additional loop mutations though directed evolution (M99Y, T101A and M103F, Extended Data Table 3), BOR^{P*} evolved into a synthetically useful catalyst (BOR^{P1}) for the borylation of Ph-EDA, supporting 340 turnovers with an e.r. of 94:6.

BOR^{P*} also allows us to move beyond diazo ester-based substrates and apply bacterial production to a different class of chiral organoboranes: although inactive towards BOR^{WT}, CF_3-substituted (diazomethyl)benzene (CF_3-DMB) reacted with NHC-borane **1** in the presence of BOR^{P*} to yield organoborane (***R***)-**18** *in vivo* with 74 turnovers and modest selectivity (79:21 e.r.). We enhanced this through three cysteine mutations at Y71, M89 and M99 (BOR^{P2}, Extended Data Table 3) to produce organoborane (***R***)-**18** in 96:4 e.r. and 1,010 TTN. Through X-ray crystallography, the absolute configuration of (***R***)-**18** was unambiguously assigned.

Finally, we asked whether the stereochemical preference of biological borylation could be switched. Towards this end, examination of the M100D V75X site-saturation library for CF_3-DMB borylation led us to identify a variant (V75G M100D; BOR^{G*}) having an inverted stereochemical preference to BOR^{P*} in the carbon–boron bond-forming step (31:69 e.r. for *R*/*S* isomer; 340 TTN). The selectivity of BOR^{G*} was further tuned through mutations M89F, T98V, M99L, T101L and M103F (BOR^{G1}, Extended Data Table 3) to yield organoborane (***S***)-**18** with 90:10 e.r. and 1,120 TTN.

Chiral α-trifluoromethylated organoboranes are useful synthetic building blocks that combine the unique properties of fluorinated motifs with the versatile synthetic applications of organoboranes[35]. However, methods for their asymmetric preparation are rare[11,36]. Our ability to biosynthesize both enantiomers of these molecules may have applications in pharmaceutical and agrochemical synthesis. For example, product (***R***)-**18** was converted to pinacol boronate **19** with retention of the stereogenic carbon centre (Fig. 3b). Through well-established stereospecific transformations[19–21], pinacol boronates can be diversified into a broad array of chiral compounds. We demonstrated the transformation of **19** to alcohol **20,** a motif found in compounds useful for the treatment of cancer[37] and neurodegenerative diseases[38], and the Matteson homologation–oxidation product **21**, both of which were obtained with good stereocontrol.

In conclusion, we present a platform for biological borylation, which can be tuned and configured through DNA manipulation. Microorganisms are powerful alternatives to chemical methods for producing pharmaceuticals, agrochemicals, materials and fuels. They are available by fermentation on a large scale and at low cost, and their genetically encoded synthetic abilities can be systematically modified and optimized. Borylation chemistry can now be added to the vast synthetic repertoire of biology.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 22 July; accepted 2 November 2017.
Published online 29 November 2017.

1. 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.* **54,** 3351–3367 (2015).
2. Hyster, T. K. & Ward, T. R. Genetic optimization of metalloenzymes: enhancing enzymes for non-natural reactions. *Angew. Chem. Int. Ed.* **55,** 7344–7357 (2016).
3. Hammer, S. C., Knight, A. M. & Arnold, F. H. Design and evolution of enzymes for non-natural chemistry. *Curr. Opin. Green Sustainable Chem.* **7,** 23–30 (2017)
4. 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).
5. Jeschek, M. *et al.* Directed evolution of artificial metalloenzymes for *in vivo* metathesis. *Nature* **537,** 661–665 (2016).
6. 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).
7. Tinoco, A., Steck, V., Tyagi, V. & Fasan, R. Highly diastereo- and enantioselective synthesis of trifluoromethyl-substituted cyclopropanes via myoglobin-catalyzed transfer of trifluoromethylcarbene. *J. Am. Chem. Soc.* **139,** 5293–5296 (2017).
8. Stelter, M. *et al.* A novel type of monoheme cytochrome *c*: biochemical and structural characterization at 1.23 A resolution of *Rhodothermus marinus* cytochrome *c*. *Biochemistry* **47,** 11953–11963 (2008).
9. Cheng, Q.-Q., Zhu, S.-F., Zhang, Y.-Z., Xie, X.-L. & Zhou, Q.-L. Copper-catalyzed B–H bond insertion reaction: a highly efficient and enantioselective C–B bond-forming reaction with amine–borane and phosphine–borane adducts. *J. Am. Chem. Soc.* **135,** 14094–14097 (2013).
10. Chen, D., Zhang, X., Qi, W.-Y., Xu, B. & Xu, M.-H. Rhodium(I)-catalyzed asymmetric carbene insertion into B–H bonds: highly enantioselective access to functionalized organoboranes. *J. Am. Chem. Soc.* **137,** 5268–5271 (2015).
11. Hyde, S. *et al.* Copper-catalyzed insertion into heteroatom–hydrogen bonds with trifluorodiazoalkanes. *Angew. Chem. Int. Ed.* **55,** 3785–3789 (2016).
12. Irschik, H., Schummer, D., Gerth, K., Höfle, G. & Reichenbach, H. The tartrolons, new boron-containing antibiotics from a myxobacterium, *Sorangium cellulosum*. *J. Antibiot.* **48,** 26–30 (1995).
13. Wolkenstein, K., Sun, H., Falk, H. & Griesinger, C. Structure and absolute configuration of Jurassic polyketide-derived spiroborate pigments obtained from microgram quantities. *J. Am. Chem. Soc.* **137,** 13460–13463 (2015).
14. Chen, X. *et al.* Structural identification of a bacterial quorum-sensing signal containing boron. *Nature* **415,** 545–549 (2002).
15. Elshahawi, S. I. *et al.* Boronated tartrolon antibiotic produced by symbiotic cellulose-degrading bacteria in shipworm gills. *Proc. Natl Acad. Sci. USA* **110,** E295–E304 (2013).
16. Dembitsky, V. M., Al Quntar, A. A. & Srebnik, M. Natural and synthetic small boron-containing molecules as potential inhibitors of bacterial and fungal quorum sensing. *Chem. Rev.* **111,** 209–237 (2011).
17. 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).
18. Das, B. C. *et al.* Boron chemicals in diagnosis and therapeutics. *Future Med. Chem.* **5,** 653–676 (2013).
19. Miyaura, N. & Suzuki, A. Palladium-catalyzed cross-coupling reactions of organoboron compounds. *Chem. Rev.* **95,** 2457–2483 (1995).
20. Leonori, D. & Aggarwal, V. K. Lithiation–borylation methodology and its application in synthesis. *Acc. Chem. Res.* **47,** 3174–3183 (2014).
21. Leonori, D. & Aggarwal, V. K. Stereospecific couplings of secondary and tertiary boronic esters. *Angew. Chem. Int. Ed.* **54,** 1082–1096 (2015).
22. Tehfe, M. A. *et al.* A water-compatible NHC-borane: Photopolymerizations in water and rate constants for elementary radical reactions. *ACS Macro Lett.* **1,** 92–95 (2012).
23. Handa, S., Wang, Y., Gallou, F. & Lipshutz, B. H. Sustainable Fe–ppm Pd nanoparticle catalysis of Suzuki-Miyaura cross-couplings in water. *Science* **349,** 1087–1091 (2015).
24. Chang, M. C. Y., Pralle, A., Isacoff, E. Y. & Chang, C. J. A selective, cell-permeable optical probe for hydrogen peroxide in living cells. *J. Am. Chem. Soc.* **126,** 15392–15393 (2004).
25. Halo, T. L., Appelbaum, J., Hobert, E. M., Balkin, D. M. & Schepartz, A. Selective recognition of protein tetraserine motifs with a cell-permeable, pro-fluorescent bis-boronic acid. *J. Am. Chem. Soc.* **131,** 438–439 (2009).
26. Kim, J. & Bertozzi, C. R. A bioorthogonal reaction of *N*-oxide and boron reagents. *Angew. Chem. Int. Ed.* **54,** 15777–15781 (2015).
27. Li, X. & Curran, D. P. Insertion of reactive rhodium carbenes into boron–hydrogen bonds of stable *N*-heterocyclic carbene boranes. *J. Am. Chem. Soc.* **135,** 12076–12081 (2013).
28. Curran, D. P. *et al.* Synthesis and reactions of *N*-heterocyclic carbene boranes. *Angew. Chem. Int. Ed.* **50,** 10294–10317 (2011).
29. Würtemberger-Pietsch, S., Radius, U. & Marder, T. B. 25 years of *N*-heterocyclic carbenes: activation of both main-group element–element bonds and NHCs themselves. *Dalton Trans.* **45,** 5880–5895 (2016).
30. Arslan, E., Schulz, H., Zufferey, R., Künzler, P & Thöny-Meyer, L. Overproduction of the *Bradyrhizobium japonicum* *c*-type cytochrome subunits of the cbb_3 oxidase in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **251,** 744–747 (1998).
31. Kille, S. *et al.* Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. *ACS Synth. Biol.* **2,** 83–92 (2013).
32. Mara, M. W. *et al.* Metalloprotein entatic control of ligand–metal bonds quantified by ultrafast X-ray spectroscopy. *Science* **356,** 1276–1280 (2017).
33. 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).
34. Hernandez, K. E. *et al.* Highly stereoselective biocatalytic synthesis of key cyclopropane intermediate to ticagrelor. *ACS Catal.* **6,** 7810–7813 (2016).
35. Argintaru, O. A., Ryu, D., Aron, I. & Molander, G. A. Synthesis and applications of α-trifluoromethylated alkylboron compounds. *Angew. Chem. Int. Ed.* **52,** 13656–13660 (2013).
36. Jiang, Q., Guo, T. & Yu, Z. Copper-catalyzed asymmetric borylation: Construction of a stereogenic carbon center bearing both CF_3 and organoboron functional groups. *J. Org. Chem.* **82,** 1951–1960 (2017).
37. Kanouni, T., Stafford, J. A., Veal, J. M. & Wallace, M. B. Histone demethylase inhibitors. WO 2014/151106 A1 (2014).
38. Scopes, D. Pyrrolo [3,2-E] [1,2,4] triazolo [1,5-A] pyrimidines derivatives as inhibitors of microglia activation. US patent 2012/0289523 A1 (2012).

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

Acknowledgements This work was supported in part by the National Science Foundation, Office of Chemical, Bioengineering, Environmental and Transport Systems SusChEM Initiative (grant CBET-1403077), the Gordon and Betty Moore Foundation through grant GBMF2809 to the Caltech Programmable Molecular Technology Initiative, and the Jacobs Institute for Molecular Engineering for Medicine at Caltech. X.H. is supported by a Ruth L. Kirschstein National Institutes of Health Postdoctoral Fellowship (F32GM125231). We thank O. F. Brandenberg, S. Brinkmann-Chen, T. Hashimoto, R. D. Lewis, and D. K. Romney for discussions and/or comments on the manuscript, and N. W. Goldberg and A. Zutshi for experimental assistance. We are grateful to S. Virgil, N. Torian, M. K. Takase and L. Henling for analytical support, and H. Gray for providing the pEC86 plasmid.

Author Contributions S.B.J.K. and X.H. designed the research with guidance from F.H.A. S.B.J.K., X.H., Y.G. and K.C. performed the experiments and analysed the data. S.B.J.K., X.H. and F.H.A. wrote the manuscript with input from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to F.H.A. (frances@cheme.caltech.edu).

Reviewer Information *Nature* thanks M. Fischbach and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Detailed experimental methods are available in the Supplementary Information.

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Materials. Plasmid pET22b(+) was used as a cloning vector, and cloning was performed using Gibson assembly[39]. The cytochrome *c* maturation plasmid pEC86[30] was used as part of a two-plasmid system to express prokaryotic cytochrome *c* proteins. Cells were grown using Luria-Bertani medium or HyperBroth (AthenaES) with 100 μg ml^{-1} ampicillin and 20 μg ml^{-1} chloramphenicol ($LB_{amp/chlor}$ or $HB_{amp/chlor}$). Cells without the pEC86 plasmid were grown with 100 μg ml^{-1} ampicillin (LB_{amp} or HB_{amp}). Electrocompetent *E. coli* cells were prepared following a published protocol[40]. T5 exonuclease, Phusion polymerase, and *Taq* ligase were purchased from New England Biolabs (NEB). M9-N minimal medium (abbreviated as M9-N buffer; pH 7.4) was used as a buffering system for whole cells, lysates and purified proteins unless otherwise specified. M9-N buffer was used without a carbon source; it contains 47.7 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.6 mM NaCl, 2.0 mM $MgSO_4$ and 0.1 mM $CaCl_2$.

Plasmid construction. All variants described in this paper were cloned and expressed using the pET22b(+) vector (Novagen). The gene encoding *Rma* cyt *c* (UniProt ID B3FQS5) was obtained as a single gBlock (IDT), codon-optimized for *E. coli*, and cloned using Gibson assembly[39] into pET22b(+) (Novagen) between restriction sites *Nde*I and *Xho*I in frame with an N-terminal pelB leader sequence (to ensure periplasmic localization and proper maturation; MKYLLPTAAAGLLLLAAQPAMA) and a C-terminal 6×His-tag. This plasmid was co-transformed with the cytochrome *c* maturation plasmid pEC86 into *E. cloni* EXPRESS BL21(DE3) cells (Lucigen).

Cytochrome *c* expression and purification. Purified cytochrome *c* proteins were prepared as follows. One litre $HB_{amp/chlor}$ in a 4 l flask was inoculated with an overnight culture (20 ml, $LB_{amp/chlor}$) of recombinant *E. cloni* EXPRESS BL21(DE3) cells containing a pET22b(+) plasmid encoding the cytochrome *c* variant, and the pEC86 plasmid. The culture was shaken at 37 °C and 200 rpm (no humidity control) until the OD_{600} was 0.7 (approximately 3 h). The culture was placed on ice for 30 min, and isopropyl β-*D*-1-thiogalactopyranoside (IPTG) and 5-aminolevulinic acid (ALA) were added to final concentrations of 20 μM and 200 μM, respectively. The incubator temperature was reduced to 20 °C, and the culture was shaken for 22 h at 200 rpm. Cells were collected by centrifugation (4000*g*, 15 min, 4 °C), and the cell pellet was stored at −20 °C until further use (at least 24 h). The cell pellet was resuspended in buffer containing 100 mM NaCl, 20 mM imidazole, and 20 mM Tris-HCl buffer (pH 7.5 at 25 °C) and cells were lysed by sonication (2 min, 2 s on, 2 s off, 40% duty cycle; Qsonica Q500 sonicator). Cell debris was removed by centrifugation for 20 min (5,000*g*, 4 °C). Supernatant was sterile-filtered through a 0.45 μm cellulose acetate filter and purified using a 1 ml Ni-NTA column (HisTrap HP, GE Healthcare) using an AKTA purifier FPLC system (GE Healthcare). The cytochrome *c* protein was eluted from the column by running a gradient from 20 to 500 mM imidazole over 10 column volumes. The purity of the collected cytochrome *c* fractions was analysed using SDS–PAGE. Pure fractions were pooled and concentrated using a 3 kDa molecular weight cut-off centrifugal filter and dialysed overnight into 0.05 M phosphate buffer (pH = 7.5) using 3 kDa molecular weight cut-off dialysis tubing. The dialysed protein was concentrated again, flash-frozen on dry ice, and stored at −20 °C. The concentration of cytochrome *c* was determined in triplicate using the haemochrome assay described below.

Cytochrome P450 and globin expression and purification. Purified P450s and globins were prepared differently from the cytochrome *c* proteins, and described as follows. One litre HB_{amp} in a 4 l flask was inoculated with an overnight culture (20 ml, LB_{amp}) of recombinant *E. cloni* EXPRESS BL21(DE3) cells containing a pET22b(+) plasmid encoding the P450 or globin variant. The culture was shaken at 37 °C and 200 rpm (no humidity control) until the OD_{600} was 0.7 (approximately 3 h). The culture was placed on ice for 30 min, and IPTG and ALA were added to final concentrations of 0.5 mM and 1 mM, respectively. The incubator temperature was reduced to 20 °C, and the culture was shaken for 20 h at 200 rpm. Cells were collected by centrifugation (4 °C, 15 min, 4,000*g*), and the cell pellet was stored at −20 °C until further use (at least 24 h). The cell pellet was resuspended in buffer containing 100 mM NaCl, 20 mM imidazole, and 20 mM Tris-HCl buffer (pH 7.5 at 25 °C). Haemin (30 mg ml^{-1}, 0.1 M NaOH; Frontier Scientific) was added to the resuspended cells such that 1 mg of haemin was added for every 1 g of cell pellet. Cells were lysed by sonication (2 min, 1 s on, 2 s off, 40% duty cycle; Qsonica Q500 sonicator). Cell debris was removed by centrifugation for 20 min (27,000*g*, 4 °C). Supernatant was sterile-filtered through a 0.45 μm cellulose acetate filter, and purified using a 1 ml Ni-NTA column (HisTrap HP, GE Healthcare) using an AKTA purifier FPLC system (GE healthcare). The P450 and globin proteins were eluted from the column by running a gradient from 20 mM to 500 mM imidazole over 10 column volumes. The purity of the collected protein fractions was analysed using SDS–PAGE. Pure fractions were pooled and concentrated using a 10 kDa molecular weight cut-off centrifugal filter and buffer-exchanged with 0.1 M phosphate buffer (pH = 8.0). The purified protein was flash-frozen on dry ice and stored at −20 °C. P450 and globin concentrations were determined in triplicate using published extinction coefficients and the haemochrome assay described below.

Haemochrome assay. A solution of sodium dithionite (10 mg ml^{-1}) was prepared in M9-N buffer. Separately, a solution of 1 M NaOH (0.4 ml) was mixed with pyridine (1 ml), followed by centrifugation (10,000*g*, 30 s) to separate the excess aqueous layer to give a pyridine–NaOH solution. To a cuvette containing 700 μl protein solution (purified protein or heat-treated lysate) in M9-N buffer, 50 μl of dithionite solution and 250 μl pyridine–NaOH solution were added. The cuvette was sealed with Parafilm, and the UV-Vis spectrum was recorded immediately. Cytochrome *c* concentration was determined using $\varepsilon_{550-535} = 22.1\ mM^{-1}\ cm^{-1}$ (ref. 41). Protein concentrations determined by the haemochrome assay were in agreement with those determined by the bicinchoninic acid assay (Thermo Fisher) using BSA for standard curve preparation.

Mutagenesis library construction. Cytochrome *c* site-saturation mutagenesis libraries were generated using a modified version of the 22-codon site-saturation method[31]. For each site-saturation library, oligonucleotides were ordered such that the coding strand contained the degenerate codon NDT, VHG or TGG. The reverse complements of these primers were also ordered. The three forward primers were mixed together in a 12:9:1 ratio, (NDT:VHG:TGG) and the three reverse primers were mixed similarly. Two PCRs were performed, pairing the mixture of forward primers with a pET22b(+) internal reverse primer, and the mixture of reverse primers with a pET22b(+) internal forward primer. The two PCR products were gel-purified, ligated together using Gibson assembly[39], and transformed into *E. cloni* EXPRESS BL21(DE3) cells.

Mutagenesis library screening in whole cells. Single colonies were picked with toothpicks off of $LB_{amp/chlor}$ agar plates, and grown in deep-well (2 ml) 96-well plates containing $LB_{amp/chlor}$ (400 μl) at 37 °C, 250 rpm shaking, and 80% relative humidity overnight. After 16 h, 30 μl aliquots of these overnight cultures were transferred to deep-well 96-well plates containing $HB_{amp/chlor}$ (1 ml) using a 12-channel EDP3-Plus 5–50 μl pipette (Rainin). Glycerol stocks of the libraries were prepared by mixing cells in $LB_{amp/chlor}$ (100 μl) with 50% v/v glycerol (100 μl). Glycerol stocks were stored at −78 °C in 96-well microplates. Growth plates were shaken for 3 h at 37 °C at 250 rpm, and 80% relative humidity. The plates were then placed on ice for 30 min. Cultures were induced by adding 10 μl of a solution, prepared in sterile deionized water, containing 2 mM IPTG and 20 mM ALA. The incubator temperature was reduced to 20 °C, and the induced cultures were shaken for 20 h (250 rpm, no humidity control). Cells were pelleted (4,000*g*, 5 min, 4 °C), resuspended in 380 μl M9-N buffer, and the plates containing the cell suspensions were transferred to an anaerobic chamber. To deep-well plates of cell suspensions were added NHC-borane substrate (10 μl per well, 400 mM in MeCN) and diazo reagent (10 μl per well, 400 mM in MeCN). The plates were sealed with aluminium sealing tape, removed from the anaerobic chamber, and shaken at 500 rpm for 6 h (24 h for reactions with Ph-EDA or CF_3-DMB because of their lower aqueous solubility). After quenching with hexanes:ethyl acetate (4:6 v/v, 0.6 ml), internal standard was added (20 μl of 20 mM 1,2,3-trimethoxybenzene in toluene). The plates were then sealed with sealing mats and shaken vigorously to thoroughly mix the organic and aqueous layers. The plates were centrifuged (4,000*g*, 5 min) and the organic layer (200 μl) was transferred to autosampler vials with vial inserts for gas chromatography–mass spectrometry (GC–MS) or chiral high performance liquid chromatography (HPLC) or supercritical fluid chromatography (SFC) analysis. Hits from library screening were confirmed by small-scale biocatalytic reactions.

Cell lysate preparation. Cell lysates were prepared as follows: *E. coli* cells expressing *Rma* cyt *c* variant were pelleted (4,000*g*, 5 min, 4 °C), resuspended in M9-N buffer and adjusted to the appropriate OD_{600}. Cells were lysed by sonication (2 min, 1 s on, 2 s off, 40% duty cycle; Qsonica Q500 sonicator), aliquoted into 2 ml microcentrifuge tubes, and the cell debris was removed by centrifugation for 10 min (14,000*g*, 4 °C). The supernatant was sterile-filtered through a 0.45 μm cellulose acetate filter, and the concentration of cytochrome *c* protein lysate was determined using the haemochrome assay. Using this protocol, the protein concentrations we typically observed for $OD_{600} = 15$ lysates were in the 8–15 μM range for wild-type *Rma* cyt *c* and 1–10 μM for other *Rma* cyt *c* variants.

Small-scale whole-cell bioconversion. In an anaerobic chamber, NHC-borane (10 μl, 400 mM in MeCN) and diazo reagent (10 μl, 400 mM in MeCN) were added to *E. coli* harbouring *Rma* cyt *c* variant (380 μl, adjusted to the appropriate OD_{600}) in a 2 ml crimp vial. The vial was crimp-sealed, removed from the anaerobic chamber, and shaken at 500 rpm at room temperature for 6 h (24 h for reactions

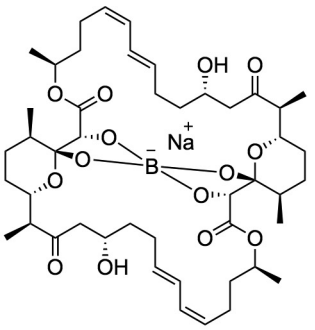
with Ph-EDA or CF_3-DMB). At the end of the reaction, the crimp vial was opened and the reaction was quenched with hexanes:ethyl acetate (4:6 v/v, 0.6 ml), followed by the addition of internal standard (20 μl of 20 mM 1,2,3-trimethoxybenzene in toluene). The reaction mixture was transferred to a microcentrifuge tube, vortexed (10 s, 3 times), then centrifuged (14,000*g*, 5 min) to completely separate the organic and aqueous layers (the vortex–centrifugation step was repeated if complete phase separation was not achieved). The organic layer (200 μl) was removed for GC–MS and chiral SFC or HPLC analysis. All biocatalytic reactions reported were performed in $n = 2$–4 technical replicates from at least two biological replicates. The TTNs reported are calculated with respect to *Rma* cyt *c* expressed in *E. coli* and represent the total number of turnovers obtained from the catalyst under the stated reaction conditions. For reactions using $OD_{600} = 15$ *E. coli* cells, the catalyst loadings were 0.0001–0.0015 mol% of enzymes with respect to the limiting reagent in the reaction. The $g_{\text{borylation product}}/g_{\text{dry cell weight}}$ ratios ranged from approximately 0.05 (wild-type) to approximately 2 (engineered variant).

Cell viability assay. The colony forming units (cfu) of whole-cell reactions (+ borylation) and controls without borylation reagents (− borylation) were determined with biological replicates according to the following procedures. Six 2 ml screw cap vials containing 380 μl suspension of *E. coli* harbouring BOR^{R1} ($OD_{600} = 15$) were transferred to an anaerobic chamber. To three of these vials were added NHC-borane **1** (10 μl, 400 mM in MeCN) and Me-EDA **2** (10 μl, 400 mM in MeCN). These vials were capped and shaken at 500 rpm in the anaerobic chamber (+ borylation). The remaining three vials were capped and shaken in the absence of reagents **1** and **2** (− borylation). After 2.5 h, all six vials were removed from the anaerobic chamber. Aliquots of cell suspension were removed from the vials and subjected to serial dilution to obtain stock solutions of 10^6, 10^7, and 10^8-fold dilution. 50 μl of each stock solution was plated on $LB_{amp/chlor}$ agar plates and incubated at 37 °C overnight. The cfu of the cell suspensions were calculated based on the colony counts of the 10^7-dilution plate. The cfu for each vial are shown in Extended Data Fig. 2.

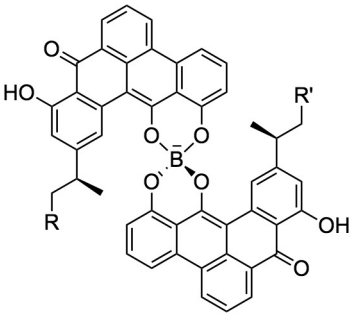
Biosynthesis of organoboranes 9 and 3 via serial substrate addition. Twelve 2 ml screw cap vials containing 400 μl suspension of cells harbouring *Rma* cyt *c* BOR^{R1} ($OD_{600} = 15$) and 100 μl of glucose (250 mM) were transferred to an anaerobic chamber. The 12 vials were grouped into 4 group sets to determine the yield, TTN and e.r. for reactions involving the stepwise addition of 2, 4, 6 or 8 equivalents of reagents. Each equivalent is 2.5 μl solution of NHC-BH_3 substrate in MeCN (2 M) and 2.5 μl Me-EDA solution in MeCN (2 M). The time interval between each equivalent was 75 min. All four group sets were shaken at 480 rpm in the anaerobic chamber until the completion of the addition and reaction for the last group set. The vials were then removed from the anaerobic chamber and quenched with 1 ml of ethyl acetate. The reaction mixture was transferred to a microcentrifuge tube, vortexed (10 s, 3 times), then centrifuged (14,000*g*, 5 min) to completely separate the organic and aqueous layers (the vortex–centrifugation step was repeated if complete phase separation was not achieved). The organic layer was removed. Another 1 ml of ethyl acetate was added for a second round of extraction and the organic solutions of two rounds of extraction were combined. The combined organic extracts were diluted (1.5-, 2-, and 5-fold dilution for experiments using 4, 6, and 8 equivalents of reagents, respectively). Internal standard (10 μl of 20 mM 1,2,3-trimethoxybenzene in toluene) was added to 300 μl of the diluted extract before the sample was subjected to GC–MS and chiral HPLC analysis to determine the yield, TTN, and e.r.

Data availability. Data supporting the findings of this study are available within the paper and its Supplementary Information, or are available from the corresponding author upon reasonable request. Crystallographic coordinates and structure factors have been deposited with the Cambridge Crystallographic Data Centre (https://www.ccdc.cam.ac.uk/) under reference numbers 1572198, 1572200 and 1572201 for organoboranes **3**, **18** and **12**, respectively.

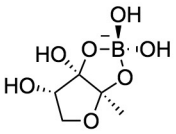
39. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).
40. Sambrook, J., Frisch, E. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1989).
41. Berry, E. A. & Trumpower, B. L. Simultaneous determination of hemes *a, b*, and *c* from pyridine hemochrome spectra. *Anal. Biochem.* **161,** 1–15 (1987).
42. Yang, J.-M. *et al.* Catalytic B–H bond insertion reactions using alkynes as carbene precursors. *J. Am. Chem. Soc.* **139,** 3784–3789 (2017).
43. Allen, T. H., Kawamoto, T., Gardner, S., Geib, S. J. & Curran, D. P. *N*-heterocyclic carbene boryl iodides catalyze insertion reactions of *N*-heterocyclic carbene boranes and diazoesters. *Org. Lett.* **19,** 3680–3683 (2017).
44. Wang, Z. J. *et al.* Improved cyclopropanation activity of histidine-ligated cytochrome P450 enables the enantioselective formal synthesis of levomilnacipran. *Angew. Chem. Int. Ed.* **53,** 6810–6813 (2014).



tartrolon B
antibiotic against Gram-positive bacteria

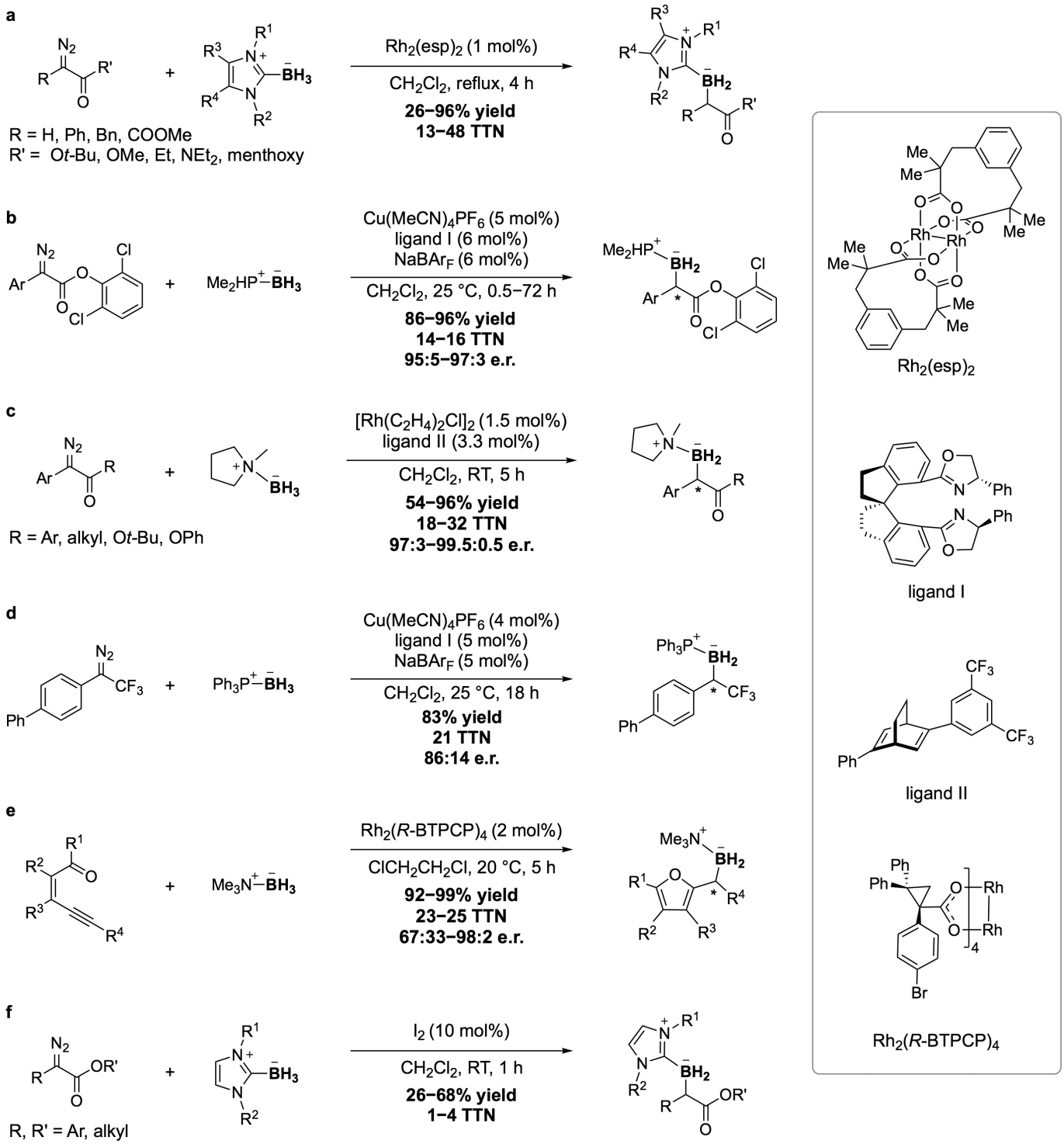


borolithochromes
pink pigments in Jurassic red alga *S. jurassica*

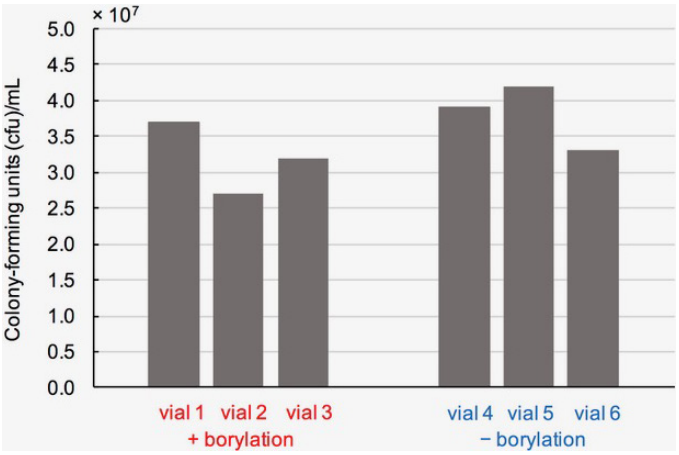


autoinducer-2
controls bacterial communication and bioluminescence in *V. harveyi*

Extended Data Figure 1 | Examples of boron-containing natural products.



Extended Data Figure 2 | Summary of known catalytic systems for metal–carbenoid insertion reactions of boranes. a, $Rh_2(esp)_2$-catalysed borylation of diazo esters with NHC-boranes[27]. **b**, $Cu(MeCN)_4PF_6$-catalysed borylation of diazo esters with phosphine-borane[9]. **c**, $[Rh(C_2H_4)_2Cl]_2$-catalysed borylation of diazo esters with amine-borane[10]. **d**, $Cu(MeCN)_4PF_6$-catalysed borylation of CF_3-substituted (diazomethyl)benzene with phosphine-borane[11]. **e**, $Rh_2(R\text{-BTPCP})_4$-catalysed borylation using alkynes as carbene precursors[42]. **f**, I_2-catalysed borylation of diazo esters with NHC-boranes[43].



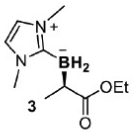
Extended Data Figure 3 | Effect of biological borylation on *E. coli* cell viability. Cell viability assay was performed in biological triplicate, see Methods section for experimental protocol.

Extended Data Table 1 | Preliminary borylation experiments with haem and haem proteins using NHC-borane (1) and Me-EDA (2) as substrates

Catalyst	TTN	e.r.
None	0	N/A
Haemin	80 ± 5	0
Haemin + BSA	170 ± 10	54:46
E. coli cell background	trace	55:45
R. marinus cyt *c*	120 ± 20	85:15
H. thermophilus cyt *c*	140 ± 10	55:45
P. ferrireducens protoglobin Y60V	NR	-
P411 CIS	trace	n.d.
BM3 P450 wild-type	NR	-
BM3 Hstar	trace	n.d.

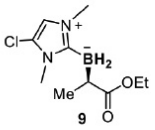
N/A, not applicable; NR, no product was detected; n.d., not determined. Experiments with cytochromes *c*, globin, or cytochromes P450 were performed using *E. coli* harbouring the corresponding protein ($OD_{600} = 15$). Reactions were performed in biological triplicate. TTNs reported represent mean values averaged over three biological replicates, and the errors are one standard deviation. Within the detection limit of the instrument, variability in e.r. was not observed. Unreacted starting materials were observed at the end of all reactions and no attempt was made to optimize these reactions. Experiments with haemin were performed using 100 μM haemin, 10 mM NHC-borane **1**, 10 mM Me-EDA **2**, 10 mM $Na_2S_2O_4$. Experiments with haemin and BSA were performed using 100 μM haemin in the presence of BSA (0.75 mg ml^{-1}) instead. Experiments to determine *E. coli* cell background reaction were performed with *E. cloni* EXPRESS BL21(DE3) cells containing a pET22b(+) plasmid encoding halohydrin dehalogenase from *Agrobacterium tumefaciens* (UniProt ID Q93D82) instead of *Rma* cyt *c*. *A. tumefaciens* halohydrin dehalogenase is inactive towards NHC-borane **1** and Me-EDA **2**. P411 CIS[4] and BM3 Hstar[44] are previously reported engineered BM3 P450 variants.

Extended Data Table 2 | Biosynthesis of organoboranes 3 and 9 via serial substrate addition



3

total equiv. of reagents	biological replicate 1			biological replicate 2		
	yield%	TTN	e.r.	yield%	TTN	e.r.
2	42	2270	97.5 : 2.5	40	2130	97.5 : 2.5
4	43	4560	97 : 3	37	3840	97 : 3
6	57	9000	96.5 : 3.5	43	6800	96.5 : 3.5
8	50	10500	96.5 : 3.5	48	10300	96.5 : 3.5

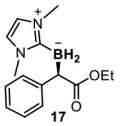


9

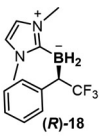
total equiv. of reagents	biological replicate 1			biological replicate 2		
	yield%	TTN	e.r.	yield%	TTN	e.r.
2	63	3300	97.5 : 2.5	59	3100	97.5 : 2.5
4	67	7000	97 : 3	55	5800	97 : 3
6	71	11100	96.5 : 3.5	69	10900	96.5 : 3.5
8	75	15800	96 : 4	72	14800	96 : 4

Experiments performed in biological duplicate. See Methods for experimental protocol.

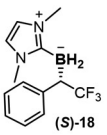
Extended Data Table 3 | Directed evolution of whole-cell *Rma* cyt *c* for improved enantioselectivity in the biosynthesis of organoboranes 17, (*R*)-18 and (*S*)-18



mutations	e.r. of **17**
M100D V75P	75 : 25
M100D V75P M99Y	81 : 19
M100D V75P M99Y T101A	89 : 11
M100D V75P M99Y T101A M103F	94 : 6



mutations	e.r. of **(*R*)-18**
M100D V75P	76 : 24
M100D V75P M89C	90 : 10
M100D V75P M89C Y71C	94 : 6
M100D V75P M89C Y71C M99C	96 : 4



mutations	e.r. of **(*S*)-18**
M100D V75G	73 : 27
M100D V75G M89F M103F	78 : 22
M100D V75G M89F M103F T101L	86 : 14
M100D V75G M89F M103F T101L M99L	88 : 12
M100D V75G M89F M103F T101L M99L T98V	90 : 10

nature research

Corresponding author(s): Frances H. Arnold

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Biocatalytic experiments were carried out in replicates (typically repeated two to four times) to ensure independent experiments are reproducible. This information is available in all relevant figure legends. No statistical methods were used to predetermine the sample size.

2. Data exclusions

Describe any data exclusions.

No data was excluded from analysis.

3. Replication

Describe whether the experimental findings were reliably reproduced.

All experimental findings were reliably reproduced using biological replicates.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

N.A.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

N.A.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	A statement indicating how many times each experiment was replicated
☒	☐	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☒	☐	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☒	☐	The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Data analysis: Microsoft excel HPLC: ChemStation for LC systems Rev. B.02.01-SR1, Agilent Technologies GCMS software: GCMS solution Version 2.72, Shimadzu Corporation SFC: ChemStation for LC 3D systems Rev. B.03.01, Agilent Technologies Chiral GC: GC ChemStation Rev. B.04.02, Agilent Technologies X-ray crystallography: The structure was solved by direct methods using SHELXS and refined against F2 on all data by full-matrix least squares with SHELXL-2014

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	N/A

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	N/A
b. Describe the method of cell line authentication used.	N/A
c. Report whether the cell lines were tested for mycoplasma contamination.	N/A
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	N/A

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	N/A

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	N/A