Oxidative cyclization of prodigiosin by an alkylglycerol monooxygenase-like enzyme

Tristan de Rond1*, Parker Stow1, Ian Egi2, Rebecca E Johnson1, Leanne Jade G Chan1,4, Garima Goyal3,4, Edward E K Baidoo3,4, Nathan J Hillson3,5, Christopher J Petzold3,4*, Richmond Sarpong1* & Jay D Keasling2,4,6,7*

Prodigiones, which are tripyrrole alkaloids displaying a wide array of bioactivities, occur as linear and cyclic congeners. Identification of an unclustered biosynthetic gene led to the discovery of the enzyme responsible for catalyzing the regiospecific C–H activation and cyclization of prodigiosin to cycloprodigiosin in *Pseudomonas rubra*. This enzyme is related to alkylglycerol monooxygenase and unrelated to RedG, the Rieske oxygenase that produces cyclized prodigiones in *Streptomyces*, implying convergent evolution. The prodigiones are a family of red tripyrrole natural products that display a broad range of promising medicinal properties, including antimalarial, anticancer, and immunosuppressive activities1–3. Notably, they have been shown to induce apoptosis in cancer cells while leaving nonmalignant cells unaffected3–5. Most prodigiones occur in linear and cyclic forms with respect to their aliphatic tails, which is apparent when comparing the structures of prodigiosin (1) to cycloprodigiosin (2) has been identified as the last common ancestor of *S. coelicolor* and *P. rubra*, implying convergent evolution. The gene neighborhood of pigN, whose precise roles remain uncharacterized, the function of each pig gene has been elucidated. The biosynthesis of 1 proceeds through a bifurcated pathway (Fig. 1a). PigBDE biosynthesize 2-methyl-3-amylypyrrole (MAP, 3). PigA and PigF–N form 4-methoxy-2,2′-bipyrrrole-5-carboxaldehyde (MBC, 4), and PigG condenses these two intermediates to yield 1. We expected the biosynthesis of 1 to proceed in a similar manner in *P. rubra*, given that its genome6,7 harbors a biosynthetic cluster showing high sequence identity with the *Serratia* ATCC39006 pig cluster (Fig. 1b and Supplementary Results, Supplementary Table 1).

To ensure that cyclization is the final step in cyclic prodigine biosynthesis in *P. rubra*, as is the case in *S. coelicolor* (Fig. 1c), we constructed an in-frame ΔpigE mutant of *P. rubra*. As expected, this mutant produced neither 1 nor 2, but upon feeding 1, we detected 2 using LC–MS (Supplementary Fig. 1). These results confirmed that the cyclization of 1 in *P. rubra*, like that in *S. coelicolor*, can occur as the final step in biosynthesis. Further support for this model is the observation that recombinant *Escherichia coli* expressing *P. rubra* pigBCDE produces 1 but not 2 upon feeding with synthetic 4 (Fig. 2a and Supplementary Fig. 2).

Bioinformatic analysis of the *P. rubra* genome5,7 revealed no RedG homologs, and the only genes encoding enzymes homologous to oxidases found in the pig cluster were those already ascribed to steps in the 1 biosynthesis pathway. Upon examining the *P. rubra* pig cluster for genes not present in *Serratia* (which does not produce 2), we noticed that although the two clusters show near-perfectly conserved synteny, the *P. rubra* cluster actually lacks a pigN gene (Fig. 1b). Although we found no close pigN homolog in the *P. rubra* genome, we did detect a homolog of *S. coelicolor* redF, which is thought to fulfill the role of pigN in *S. coelicolor*1,11,12 in the PRUB675 locus. For a more in-depth analysis of the relationship between these proteins, which are all in the DUF1295 protein family, see Supplementary Fig. 3. In *Serratia*, disruption of pigN impedes, but does not abolish, conversion of 4-hydroxy-2,2′-bipyrrrole-5-carboxaldehyde (HBC, 5) into 4. Accumulation of 5 results in diminished production of 1 and increased production of nor prodigiosin (6), which is formed by the PigC-catalyzed condensation of 5 with 3 (ref. 16). The same phenotype was observed for *P. rubra* ΔPRUB675, suggesting that PRUB675 fulfills the function of pigN in *P. rubra* (Supplementary Fig. 4).

The gene neighborhood of PRUB675 revealed that the gene appears to be part of a transcriptional unit with PRUB680, which bears homology to di-iron oxidases. Deleting PRUB680 in *P. rubra* abolished production of 2, but had no effect on 1 (Fig. 2a). Furthermore, heterologous expression of PRUB680 alongside pigBCDE in *E. coli* fed 4 led to the formation of 2. Direct bioconversion of 1 to 2 by recombinant *E. coli* could barely be observed, which has also been the case for the bioconversion of undecylprodigiosin to...
streptorubin B by Streptomyces expressing redG\(^2\). We suspect that 1
is unable to cross the E. coli cell wall effectively, whereas 4 can.

Bioinformatic analysis showed that PRUB680 shares no substan-
tial sequence similarity with RedG and is instead a member of the FA
hydroxylase family of integral membrane di-iron oxygenases. PRUB680
displays the characteristic eight-histidine motif that is essential for iron
binding and catalysis in this enzyme family (Supplementary Fig. 5)\(^{17,18}\).
Di-iron oxygenases are known to carry out a wide variety of C–H activa-
tion chemistries (Supplementary Figs. 6 and 7; Supplementary Table
2)\(^{17}\), but so far none have been reported to catalyze oxidative cycliza-
tion or C–C bond formation\(^19\).

The closest characterized homolog of PRUB680 is alkylglycerol
monooxygenase (AGMO), which is present only in metazoans and
some protists and forms an isolated eukaryotic branch of the FA
hydroxylase family. AGMO plays a central role in lipid homeostasis
by catalyzing the breakdown of ether lipids, a deficit of which leads
to the development of cataracts and disrupts spermatogenesis in
mice\(^20\). AGMO is distinct among di-iron oxygenases—most of which
are copper- or vanadium-containing proteins—atypical among enzymes
catalyzing novel C–H activation reactivity. Moreover, some of these
unexplored enzymes may be involved in C–H activating steps in the
biosynthesis of novel natural products. None of PRUB680’s closest
homologs are found in organisms known to produce prodiginines (Supplementary Fig. 7), suggesting that PRUB680 is a functional
outlier among enzymes that carry out other oxidative chemistry.
The genomic context of these homologs of PRUB680 gives no clear
indication regarding their functions (Supplementary Table 3).

Most characterized bacterial biosynthetic pathways are encoded
by genes that are physically clustered on the genome. However, in P. rubra, the genes encoding prodigine biosynthesis are split across
two loci, a situation we were alerted to by the absence of the strictly
conserved gene pigN. An analogous strategy may help to identify
biosynthetic enzymes in other organisms that are not clustered with
their respective pathways.

The exact role of PigN in prodigine biosynthesis is still
unknown. Given that PigB, which catalyzes the final step of 3
biosynthesis, is predicted to have two transmembrane helices
(Supplementary Fig. 10), and PigG has been found to localize to the
membrane when expressed heterologously\(^24\), we suspect that
the concluding steps of 1 biosynthesis occur at the membrane. PigN
has five predicted transmembrane helices (Supplementary Fig. 11)
and may act to recruit PigF to the membrane. The presence of
pigN (or redF) is strictly conserved among prodigine-producing
organisms despite the weak phenotype of ΔpigN merely decreasing
the ratio of 1 to 6, but not abolishing 1 production. P. rubra may
have acquired pigA–M in one horizontal gene transfer event while
acquiring PRUB675–680 independently, perhaps due to strong
selective pressure for a gene fulfilling the role of pigN.

In summary, we have shown that PRUB680, a membrane di-iron
monooxygenase-like enzyme, produces 2 by cyclization of 1, analogously
to the cyclization of undecylprodigiosin to form streptorubin B cata-
lized by RedG, a Rieske oxygenase-like enzyme. Despite sharing no
sequence similarity, both enzymes are predicted to employ histidine-ligated nonheme iron centers13,14 to catalyze the oxidative cyclization of prodigiones. PRUB680 bears strong homology to AGMO and has similar cofactor requirements, and therefore may serve as a prokaryotic model for the latter. Furthermore, the uncharacterized bacterial enzymes related to PRUB680 may provide a valuable source of novel C–H activation reactivity. PRUB680 itself may also prove useful as a biocatalyst to produce novel prodigiones. The fact that cyclic prodiginines biosynthesis evolved independently at least twice suggests that there exists a strong selective pressure to produce cyclic prodiginines. However, thus far, the ecological role of the prodiginines—and hence the adaptive advantage conferred by cyclic prodiginines—remains an enigma.

Received 1 February 2017; accepted 26 July 2017; published online 11 September 2017

Methods

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

References

1. Papreddy, K. et al. J. Med. Chem. 54, 5296–5306 (2011).
2. Stankovic, N., Senerovic, L., Ilic-Tomic, T., Vasiljevic, B. & Nikodinovic-Runic, J. Appl. Microbiol. Biotechnol. 98, 3841–3858 (2014).
3. Williamson, N.R. et al. Future Microbiol. 2, 605–618 (2007).
4. Montaner, B. & Pérez-Tomàs, R. Life Sci. 68, 2025–2036 (2001).
5. Yamamoto, C. et al. Hepatology 30, 894–902 (1999).
6. Jones, B.T., Hu, D.X., Savoie, B.M. & Thomson, R.J. J. Nat. Prod. 76, 1937–1945 (2013).
7. Lee, J.S. et al. Appl. Environ. Microbiol. 77, 4967–4973 (2011).
8. Sydor, P.K. et al. Nat. Chem. 3, 388–392 (2011).
9. Salem, S.M. et al. J. Am. Chem. Soc. 136, 4565–4574 (2014).
10. Withall, D.M., Haynes, S.W. & Challis, G.L. J. Am. Chem. Soc. 137, 7889–7897 (2015).
11. Kimata, S., Iizawa, M., Kawasaki, T. & Hayakawa, Y. J. Antibiot. (Tokyo) 70, 196–199 (2017).
12. Kancharla, P., Lu, W., Salem, S.M., Kelly, J.X. & Reynolds, K.A. J. Org. Chem. 79, 11674–11689 (2014).
13. Hu, D.X., Withall, D.M., Challis, G.L. & Thomson, R.J. Chem. Rev. 116, 7818–7853 (2016).
14. Williamson, N.R., Fineran, P.C., Leeper, F.J. & Salmon, G.P.C. Nat. Rev. Microbiol. 4, 887–899 (2006).
15. Xie, B.-B. et al. J. Bacteriol. 194, 1637–1638 (2012).
16. Williamson, N.R. et al. Mol. Microbiol. 56, 971–989 (2005).
17. Shanklin, J., Guy, J.E., Mishra, G. & Lindqvist, Y. J. Biol. Chem. 284, 18559–18563 (2009).
18. Bai, Y. et al. Nature 524, 252–256 (2015).
19. Tang, M.-C., Zou, Y., Watanabe, K., Walsh, C.T. & Tang, Y. Chem. Rev. 117, 5226–5333 (2017).
20. Watschinger, K. & Werner, E.R. JURMB Life 65, 366–372 (2013).
21. Watschinger, K. et al. Biochem. J. 443, 279–286 (2012).
22. Futai, M. J. Membr. Biol. 15, 15–28 (1974).
23. Mayer, M. et al. Pteridines 24, 105–109 (2013).
24. Chawrai, S.R., Williamson, N.R., Mahendiran, T., Salmon, G.P.C. & Leeper, F. J. Chem. Soc. 3, 447–454 (2012).
25. Barry, S.M. & Challis, G.L. ACS Catal. 3, 2362–2370 (2013).

Acknowledgments

We thank A. Deutschbauer for gifting E. coli WM3064. We thank I. Budin, J. Alonso-Gutierrez, M. Thompson, D. Fercher, J. Barajas, C. Eiben, C. Bailey, S. Richardson, M. Brown, M. Garber, S. Nies, C. Meadows, T.S. Lee, I. Katz, and S. Weschler for helpful discussions; E. de Ugarte for work on the graphical abstract; as well as S. Gardner, E. Coyne, and M. Agnitsch for everyday support. This work was supported by ERAStynBio (81861: “SynPath”) to J.D.K., the NIGMS (086374) to R.S., and the UC Berkeley SURF Rose Hills fellowship to P.S., as well as the DOE Joint BioEnergy Institute and the DOE Joint Genome Institute, both funded by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the US Department of Energy. The publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

Author contributions

T.d.R., R.E.J., R.S., and J.D.K. conceived of the study. T.d.R., P.S. and I.E. constructed plasmids and performed microbiological manipulations and extractions, R.E.J. performed synthetic organic chemistry. T.d.R., E.E.K.B., and C.J.P. performed analytical chemistry, L.J.G.C. and C.J.P. performed proteomic analysis, and G.G. and N.J.H. performed synthetic organic chemistry. T.d.R., R.E.J., R.S., and J.D.K. conceived of the study. T.d.R., P.S. and I.E. constructed plasmids and performed microbiological manipulations and extractions, R.E.J. performed synthetic organic chemistry. T.d.R., E.E.K.B., and C.J.P. performed analytical chemistry, L.J.G.C. and C.J.P. performed proteomic analysis, and G.G. and N.J.H. performed synthetic organic chemistry. All authors contributed to the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Any supplementary information, chemical compound information and source data are available in the online version of the paper.

Keywords: prodiginines, cyclic prodiginines, PRUB680, synthetic organic chemistry
ONLINE METHODS

Synthetic chemistry. Cycloprodigiosin (2) and MBC (4) were synthesized as described previously.

Bacterial cultivation. *E. coli* was propagated at 37 °C on LB agar or in LB medium. For the cultivation of *P. rubra* (at 30 °C), these media were supplemented with 10% v/v 180 g/L Instant Ocean Sea Salt (IO; Spectrum Brands, Blacksburg, VA) and autoclaved separately. Descriptions of bacterial strains employed in this study are provided in Supplementary Table 4.

Plasmid construction. DNA assembly protocols were designed using j5 and DeviceEditor software. Descriptions of plasmids employed in this study are provided in Supplementary Table 5. Assembly of DNA fragments (Supplementary Table 6) was performed using NEBuilder HiFi DNA Assembly Master Mix or NEB Golden Gate Assembly Mix (NEB) per the manufacturer’s directions. The 11-kb pigBCDE fragment was cloned behind a T7 promoter using the Zero Blunt TOPO PCR Cloning Kit (Thermo Fisher).

Targeted gene disruptions in *P. rubra*. We employed conjugative transfer of a suicide plasmid following literature precedent; however, counterselection with ScaB was not effective in our hands. This held true even with sacB under the control of promoters expressed highly in *P. rubra*, as determined by shot-gun proteomics (Supplementary Table 7). Instead, we replaced sacB with lacZ and identified double crossovers by blue-white screening (Supplementary Fig. 12). *E. coli* WM3064 was transformed with suicide vectors conferring both erythromycin and chloramphenicol resistance markers under the control of the *P. rubra* elongation factor G (*PRUB9669* on contig 67) promoter, the *P. rubra* 305 ribosomal protein S13 promoter (*PRUB13406* on contig 115); this is actually a polycistronic locus with a number of ribosomal proteins and an RNA polymerase subunit) driving lacZ, and −1 kb regions homologous to those upstream and downstream of the target. After overnight growth on LB agar with 25 μg/mL chloramphenicol, 100 μg/mL X-gal, and 300 μM diaminopimelic acid (DAP), a colony was patched directly on LB agar with 4% v/v 180 g/L IO and 300 μM DAP, with a wild-type *P. rubra* colony patched on top. After conjugating at 30 °C overnight, the patch was struck out for single colonies on LB agar with 10% v/v 180 g/L IO, 25 μg/mL erythromycin, and 500 μg/mL X-gal. Blue colonies (single-crossovers) were passaged until homogenous. These were then subcultured on the same growth media without erythromycin, and white colonies were isolated and confirmed to be sensitive to erythromycin. To distinguish double crossovers from revertants, colony PCR was performed by picking colonies into neat DMSO, diluting 1:10 with water and using that as template (1% v/v) with 5Prime HotMasterMix polymerase and primers as specified in Supplementary Table 6.

Prodigiosin production in *P. rubra*. *P. rubra* was grown in 50 mL 20% v/v LB, 10% v/v 180 g/L IO, and 70% de-ionized water in a 250-mL baffled shake flask. After 12 h of growth at 30 °C, 2 mL of the culture was extracted with 3 mL 1:1 chloroform:methanol. The organic layer was evaporated to dryness, dissolved in ethanol, diluted 1:1 in de-ionized water, and analyzed by LC–MRM–MS and LC–TOF–MS.

Heterologous prodigiosin production in *E. coli*. *E. coli* BLR (DE3) was transformed with plasmids containing pigBCDE and either PRUB680 or RFP. Overnight cultures were diluted 1:10 into LB medium supplemented with 50 μg/mL kanamycin and 25 μg/mL chloramphenicol and grown at 37 °C to an optical density at 600 nm (OD600) of 0.6, upon which they were induced with 100 μM IPTG at 30 °C for 16 h. 10 mL of cells were harvested by centrifugation (8,000g for 5 min), resuspended in 300 μL of LB medium, and spread onto agar plates with 0.1 mM IPTG and antibiotics as before. 10 μL of 1 mM MBC in 1:3 DMSO:water was spotted onto the plates, which were left to grow overnight at 30 °C. The pink halo (as can be seen in Supplementary Fig. 2) was scraped off and resuspended in 1 mL de-ionized water with 2% TFA by vigorous vortexing. The cell suspension was extracted with 2 mL of 1:1 v/v chloroform:methanol. The organic layer was evaporated to dryness dissolved in ethanol, diluted 1:1 in de-ionized water, and analyzed by LC–MRM–MS.

In vitro analysis of PRUB680 in inverted *E. coli* membrane vesicles. An overnight culture of *E. coli* BLR (DE3) containing pET28-PRUB680 was diluted 1:10 into 3 x 500 mL LB medium supplemented with 50 μg/mL kanamycin in 2-L baffled flasks and grown at 37 °C. When the cells reached an OD600 of 0.6, the flasks were cooled to 18 °C and induced with 100 mM IPTG for 3 h. The cells were harvested by centrifugation (5,000g for 10 min), washed with 30 mL spheroplasting buffer (30% sucrose, 200 mM Tris–HCl pH 8.0, 2 mM EDTA) and incubated, rocking for 30 min at room temperature in 30 mL spheroplasting buffer + 3 mg lysozyme. Spheroplasts were harvested by centrifugation (5,000g for 10 min), resuspended in 30 mL assay buffer (100 mM HEPES–KOH pH 7.8, 50 mM K2SO4, 1% v/v Sigma Protease Inhibitor Cocktail P8849), divided into 20 x 1.5 mL in 2-mL centrifuge tubes, and sonicated in a cup-horn sonicator (Qsonica Q700 with 431MPX horn; amplitude: 75%; 1 min on, 1 min off) for 45 min. of total ‘on’ time. The water bath temperature was maintained between 3 and 10 °C. The tubes were centrifuged at 12,000g for 10 min at 4 °C, and the supernatants were then combined and again centrifuged at 12,000g for 10 min at 4 °C. The supernatant was centrifuged at 120,000g for 30 min at 4 °C, and the orange pellet thoroughly resuspended in 22 mL assay buffer. For each reaction, 500 μL of the vesicle preparation was used. For the metal dependence experiments, EDTA was added to a final concentration of 4 mM, and the mixture was incubated at 4 °C for 5 min, followed by the addition of metal at a concentration of 5 mM and incubation at 4 °C for 5 min. For all experiments, 10 μM prodigiosin (Enzo Life Sciences, 100× stock solution prepared at 1 mM in 20% v/v ethanol in water) was added, the mixture transferred to a round-bottom glass tube (16 mm x 100 mm) at room temperature. The reaction was started by adding reducing cofactor (NADH, NADPH, (6R- or (6S)-tetrahydrobiopterin, or tetrahydrofolate, all from Sigma-Aldrich) at 250 μM, or, in the case of FMNH2, adding flavin mononucleotide (FMN) to the mixture preloaded with a cofactor generation system consisting of 20 mM glucose-6-phosphate, 250 μM NADP+, 1 unit/mL glucose-6-phosphate dehydrogenase, and 1 unit/mL NADPH:FMN oxidoreductase from *Photobacterium fischeri* (all from Sigma-Aldrich). Enzymatic generation of FMNH2 was necessary, because PRUB680 is inactivated by sodium dithionite. The *in situ* reduction of FMN to FMNH2 was verified by observing the loss of yellow color. Metal dependence experiments used 250 μM (6R)-tetrahydrobiopterin. After shaking at 200 r.p.m. at room temperature for 10 min, reactions were quenched using 2 mL 1:1 v/v chloroform:methanol. 500 μL de-ionized water was added, and the organic layer evaporated to dryness, dissolved in ethanol, diluted 1:1 with de-ionized water, and analyzed by LC–MRM–MS. Peak areas were calculated using Analyst 1.6.2. To calculate relative enzyme activity, cycloprodigiosin peak areas were normalized to the prodigiosin starting material peak areas to correct for extraction efficiency (<0.1% conversion had occurred under all conditions), and to a (6R)-tetrahydrobiopterin reaction to normalize for enzyme activity differences between vesicle preparations. All conditions shown in Figure 2b, except for FMNH2, were performed in parallel with the same vesicle preparation.

LC–MS acquisition and data analysis. LC–M RM–MS was performed on an AB Sciex 4000 QTRAP with an Agilent 1200 series LC system. 1 μL of sample was injected onto a Phenomenex Kinetex XB-C18 (3 mm x 100 mm) column. Mobile phase: A = 10 mM ammonium formate, brought to pH 4.5 with formic acid; B = methanol buffered identically to A. Method: 35% B for 5 min, ramp from 35% to 80% B in 30 min, 80% B for 8 min, ramp to 35% B in 2 min, re-equilibrate at 35% B for 15 min; all at a flow rate of 200 μL/min. A Turbo Spray ion source was used in positive ion mode (curtain gas: 20 L/min; temperature: 600 °C; voltage: 4,800 V; source gas: 50 L/min; entrance potential: 8 V; collision energy: 45; declustering potential: 45 V; column temperature: 50 °C; Q1 resolution: high; Q3 resolution: unit). The transitions monitored (324→252 for 1, 322→292 for 2) were based on published tandem MS spectra.

LC–TOF–HRMS was performed on an Agilent 1200 series Rapid Resolution HPLC system. Mobile phases were the same as above. 2 μL of sample was injected onto a Phenomenex Kinetex XB-C18 (3 mm x 50 mm) column. Mobile phase: 40% B for 8 min, ramp from 40% to 80% B in 20 min, 80% B for 4 min, ramp to 35% B in 1 min, re-equilibrate at 35% B for 7 min; all at a flow rate: 200 μL/min. An Agilent ESI source was used in the positive ion mode.
were obtained directly from Pfam version 3.0 (ref. 32). A maximum-likelihood phylogenetic tree was built with FastTree using default parameters33. Branches were assigned colors based on metadata from the UniProt database34. The tree was rendered using iTOL version 3 (ref. 35).

Data availability. Strains and plasmids developed for this study (Supplementary Tables 4 and 5), along with annotated sequences, have been deposited in the public instance of the Joint BioEnergy Institute Registry (https://public-registry.jbei.org/folders/260) and are physically available from the authors and/or Addgene (http://www.addgene.org) upon reasonable request. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

A Life Sciences Reporting Summary for this paper is available.

Semi-quantitative shotgun proteomics. Cell lysis and protein precipitation were performed using a chloroform–methanol extraction as previously described31. The protein pellet was resuspended in 100 mM (NH₄)HCO₃ with 20% methanol and the protein concentration was measured using the DC Protein Assay Kit (Bio-Rad). The protein was reduced with 5 mM TCEP for 30 min at room temperature, alkylated with 10 mM iodoacetamide for 30 min in the dark at room temperature, and digested with trypsin (1:50 w/w, trypsin:protein) overnight at 37 °C.

Peptide samples (20 μg) were analyzed on an Agilent 6550 iFunnel Q-TOF mass spectrometer coupled to an Agilent 1290 UHPLC system (Agilent Technologies). Peptides were loaded into an Ascentis Express Peptide ES-C18 column (100 mm × 2.1 mm i.d., 2.7-μm particle size; Sigma-Aldrich, St. Louis, MO, USA) operating at 60 °C and flowing at 0.400 mL/min. Mobile phase: A = 0.1% formic acid in water; B = 0.1% formic acid in acetonitrile. Method: 2% B for 2 min, ramp from 2% to 30% B over 30 min, ramp to 50% over 5 min, ramp to 80% over 1 min, hold at 80% B for 7 min, ramp to 2% B over 1 min, hold at 2% B for 4 min. An Agilent Dual Jet Stream Electrospray Ionization source was used in positive-ion mode. Gas temperature: 250 °C, drying gas: 14 L/min, nebulizer: 35 psig, sheath gas temp: 250 °C, sheath gas flow: 11 L/min, VCap: 4,500 V, nozzle voltage: 1,000 V, fragmentor: 180 V, and OCT 1 RF Vpp: 750 V.

The data were acquired with Agilent MassHunter Workstation version B.06.01 and searched against the P. rubra genome using Mascot version 2.3.02 (Matrix Science), then filtered and refined using Scaffold version 4.6.1 (Proteome Software, Inc.).

Bioinformatics. Pre-aligned Uniprot-RP75 (representative proteome clustered at 75% sequence identity) sequences for the FA hydroxylase (PF04116) family were obtained directly from Pfam version 3.0 (ref. 32). A maximum-likelihood phylogenetic tree was built with FastTree using default parameters33. Branches were assigned colors based on metadata from the UniProt database34. The tree was rendered using iTOL version 3 (ref. 35).

26. Johnson, R.E., de Rond, T., Lindsay, V.N.G., Keasling, J.D. & Sarpong, R. Org. Lett. 17, 3474–3477 (2015).
27. Dairi, K., Tripathy, S., Attardo, G. & Lavallée, J.-F. Tetrahedr. Lett. 47, 2605–2606 (2006).
28. Hillson, N.J., Rosengarten, R.D. & Keasling, J.D. ACS Synth. Biol. 1, 14–21 (2012).
29. Wang, P. et al. Microb. Cell Fact. 14, 11 (2015).
30. Alihosseini, F., Lango, J., Ju, K.-S., Hammock, B.D. & Sun, G. Biotechnol. Prog. 26, 352–360 (2010).
31. Batth, T.S., Keasling, J.D. & Petzold, C.J. Methods Mol. Biol. 944, 237–249 (2012).
32. Finn, R.D. et al. Nucleic Acids Res. 42, D222–D230 (2014).
33. Price, M.N., Dehal, P.S. & Arkin, A.P. Mol. Biol. Evol. 26, 1641–1650 (2009).
34. UniProt Consortium. Nucleic Acids Res. 43, D204–D212 (2015).
35. Letunic, I. & Bork, P. Nucleic Acids Res. 44, W242–W245 (2016).
# 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

### Sample size
Describe how sample size was determined.

For the enzyme kinetics in Figure 2b, N=3. This is enough to highlight the qualitative results shown in this figure.

### Data exclusions
Describe any data exclusions.

No data was excluded

### Replication
Describe whether the experimental findings were reliably reproduced.

All attempts at replication were successful

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

N/A

### 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.

## 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).

- 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.

All the software used is mentioned in the methods section.

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.

All materials are readily available from the authors.

Antibodies

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

N/A

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