

Enabling microbial syringol conversion through structure-guided protein engineering

Melodie M. Machovina^{a,1}, Sam J. B. Mallinson^{b,1}, Brandon C. Knott^{c,1}, Alexander W. Meyers^{d,1}, Marc Garcia-Borràs^{e,1}, Lintao Bu^c, Japheth E. Gado^{d,f}, April Oliver^a, Graham P. Schmidt^c, Daniel J. Hinchey^b, Michael F. Crowley^c, Christopher W. Johnson^d, Ellen L. Neidle^g, Christina M. Payne^f, Kendall N. Houk^{e,2}, Gregg T. Beckham^{d,h,2}, John E. McGeehan^{b,2}, and Jennifer L. DuBois^{a,2}

^aDepartment of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717; ^bCentre for Enzyme Innovation, School of Biological Sciences, Institute of Biological and Biomedical Sciences, University of Portsmouth, Portsmouth PO1 2UP, United Kingdom; ^cBiosciences Center, National Renewable Energy Laboratory, Golden, CO 80401; ^dNational Bioenergy Center, National Renewable Energy Laboratory, Golden, CO 80401; ^eDepartment of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095; ^fDepartment of Chemical Engineering, University of Kentucky, Lexington, KY 40506; ^gDepartment of Microbiology, University of Georgia, Athens, GA 30602; and ^hCenter for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN 37830

Contributed by Kendall N. Houk, May 11, 2019 (sent for review November 26, 2018; reviewed by Lindsay D. Eltis, Erika A. Taylor, and Binju Wang)

Microbial conversion of aromatic compounds is an emerging and promising strategy for valorization of the plant biopolymer lignin. A critical and often rate-limiting reaction in aromatic catabolism is *O*-aryl-demethylation of the abundant aromatic methoxy groups in lignin to form diols, which enables subsequent oxidative aromatic ring-opening. Recently, a cytochrome P450 system, GcoAB, was discovered to demethylate guaiacol (2-methoxyphenol), which can be produced from coniferyl alcohol-derived lignin, to form catechol. However, native GcoAB has minimal ability to demethylate syringol (2,6-dimethoxyphenol), the analogous compound that can be produced from sinapyl alcohol-derived lignin. Despite the abundance of sinapyl alcohol-based lignin in plants, no pathway for syringol catabolism has been reported to date. Here we used structure-guided protein engineering to enable microbial syringol utilization with GcoAB. Specifically, a phenylalanine residue (GcoA-F169) interferes with the binding of syringol in the active site, and on mutation to smaller amino acids, efficient syringol *O*-demethylation is achieved. Crystallography indicates that syringol adopts a productive binding pose in the variant, which molecular dynamics simulations trace to the elimination of steric clash between the highly flexible side chain of GcoA-F169 and the additional methoxy group of syringol. Finally, we demonstrate *in vivo* syringol turnover in *Pseudomonas putida* KT2440 with the GcoA-F169A variant. Taken together, our findings highlight the significant potential and plasticity of cytochrome P450 aromatic *O*-demethylases in the biological conversion of lignin-derived aromatic compounds.

demethylase | P450 | lignin | biorefinery

Lignin is a heterogeneous, recalcitrant biopolymer prevalent in plant cell walls, where it provides structure, defense against pathogens, and water and nutrient transport through plant tissue (1). Lignin is synthesized primarily from three aromatic building blocks (1, 2), making it the only abundant and renewable aromatic carbon feedstock available. Due to its recalcitrance, rot fungi and some bacteria have evolved powerful, oxidative enzymes that deconstruct lignin to smaller fragments (3, 4). Once broken down, the lignin oligomers can be assimilated as a carbon and energy source through at least four known aromatic-catabolic pathways (2, 5).

A critical reaction in the aerobic catabolism of lignin-derived compounds is *O*-aryl-demethylation, which occurs on methoxylated lignin-derived compounds to produce aromatic diols, such as catechol (1,2-dihydroxybenzene), protocatechuate (3,4-dihydroxybenzoate), and gallate (3,4,5-trihydroxybenzoate). Next, the aromatic rings are cleaved by intradiol or extradiol dioxygenases, and the products are funneled into central metabolism (6, 7). Harnessing this catabolic capability for transforming heterogeneous lignin streams into valuable chemicals is of keen

interest (7–12) and essential for economical lignocellulose conversion (11, 13, 14).

In most plants, lignin comprises primarily coniferyl (G) and sinapyl (S) alcohol monomers, which have one methoxy group and two methoxy groups on the aryl ring, respectively. Nearly all lignin-derived aromatics require *O*-demethylation of these methoxy groups as an essential step in their conversion to central intermediates. Significant effort has been dedicated to the discovery of enzymes that can demethylate the methoxy substituents of diverse aromatic compounds (15–26). Ornston et al. (18, 19) characterized the *O*-demethylation of vanillin (4-hydroxy-3-methoxybenzaldehyde) and vanillate (4-hydroxy-3-methoxybenzoate) analogs by the VanAB monooxygenase from *Acinetobacter baylyi* ADP1, which contains a Rieske nonheme iron center. The three-component LigX

Significance

Lignin is an abundant but underutilized heterogeneous polymer found in terrestrial plants. In current lignocellulosic biorefinery paradigms, lignin is primarily slated for incineration, but for a nonfood plant-based bioeconomy to be successful, lignin valorization is critical. An emerging concept to valorize lignin uses aromatic-catabolic pathways and microbes to funnel heterogeneous lignin-derived aromatic compounds to single high-value products. For this approach to be viable, the discovery and engineering of enzymes to conduct key reactions is critical. In this work, we have engineered a two-component cytochrome P450 enzyme system to conduct one of the most important reactions in biological lignin conversion: aromatic *O*-demethylation of syringol, the base aromatic unit of S-lignin, which is highly abundant in hardwoods and grasses.

Author contributions: E.L.N., C.M.P., K.N.H., G.T.B., J.E.M., and J.L.D. designed research; M.M.M., S.J.B.M., B.C.K., A.W.M., M.G.-B., L.B., J.E.G., A.O., G.P.S., D.J.H., M.F.C., C.W.J., and K.N.H. performed research; G.T.B., J.E.M., and J.L.D. contributed new reagents/analytic tools; M.M.M., S.J.B.M., B.C.K., A.W.M., M.G.-B., L.B., J.E.G., M.F.C., C.W.J., C.M.P., K.N.H., G.T.B., J.E.M., and J.L.D. analyzed data; and M.M.M., S.J.B.M., K.N.H., G.T.B., J.E.M., and J.L.D. wrote the paper.

Reviewers: L.D.E., University of British Columbia; E.A.T., Wesleyan University; and B.W., Xiamen University.

The authors declare no conflict of interest.

This open access article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

Data deposition: Protein crystal structure coordinates have been deposited in the Protein Data Bank, <https://www.rcsb.org/> (PDB ID codes 6HQK–6HQT).

¹M.M.M., S.J.B.M., B.C.K., A.W.M., and M.G.-B. contributed equally to this work.

²To whom correspondence may be addressed. Email: hok@chem.ucla.edu, gregg.beckham@nrel.gov, john.mcgeehan@port.ac.uk, or jennifer.dubois1@montana.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1820001116/-DCSupplemental.

monooxygenase system from *Sphingobium* sp. SYK-6, described by Masai et al. (22), also contains a Rieske nonheme iron component that is responsible for *O*-aryl-demethylation of a model biphenyl compound that mimics those in lignin. Masai et al. (20, 21) described two tetrahydrofolate-dependent enzymes, LigM and DesA, responsible for *O*-aryl-demethylation of vanillate and syringate (4-hydroxy-3,5-dimethoxybenzoate), respectively. Cytochrome P450 systems have also been reported to demethylate aromatic compounds, such as guaiacol, 4-methoxybenzoate, and guaethol (2-ethoxyphenol) (15, 17, 25, 26); however, the full gene sequences were either unreported or only recently identified (15, 17, 27), or the substrate was not of direct interest to lignin conversion (25, 26).

Our recent characterization of a two-component P450 enzyme system consisting of a reductase, GcoB, and a P450 oxidase, GcoA, demonstrated that it demethylates diverse aromatic compounds, including guaiacol (which can be derived from coniferyl alcohol and represents the aromatic functionality of G-lignin that must undergo demethylation), guaethol, anisole (methoxybenzene), 2-methylanisole, and 3-methoxycatechol (3MC) (28) with similar or greater efficiency than other *O*-aryl-demethylases described in the literature (22, 24, 29). However, GcoAB shows poor reactivity toward syringol, which can be derived from sinapyl alcohol via high-temperature reactions and represents the aromatic functionality of S-lignin that must undergo *O*-demethylation for further catabolism via ring-opening dioxygenases. G- and S-lignin are the major components of lignin in hardwoods and grasses (1). Due to their abundance, it is important to find enzymes that can act on the methoxy groups of both G- and S-lignin subunits. To date, there are no reports describing syringol *O*-demethylation or, more broadly, even its catabolism by microbes. Rather, the best-studied biological reaction of syringol is its 4–4 dimerization to form cerulignone (30–33).

Although our previous work showed that GcoA was not effective for syringol *O*-demethylation, crystallographic studies and molecular dynamics (MD) simulations indicated that a triad of active site phenylalanine residues is both highly mobile and important for positioning the substrate in its catalytically competent pose. In the present study, we hypothesized that substitution of GcoA-F169, which has the closest interaction with the bound substrate, may relax the specificity of the enzyme sufficiently to permit the *O*-demethylation of S-lignin type substrates. We tested that hypothesis using biochemical, structural, computational, and in vivo approaches. We demonstrate highly efficient in vitro and in vivo syringol turnover through structure-guided protein engineering, where the enzyme also retains highly efficient activity toward guaiacol.

Results

The Syringol Binding Mode Can Be Modulated by Active Site Engineering. Guaiacol assumes a productive orientation in the active site of GcoA, resulting in a shift in the spin state of the heme iron from low ($S = 1/2$) to high ($S = 5/2$), due to the actions of amino acid side chains that create a tight-fitting hydrophobic pocket. The closest contact is with GcoA-F169, which forms a hydrophobic interaction with the C6 carbon on the aromatic ring of guaiacol. Previous MD simulations suggest that this residue is highly mobile, predicting that the productive complex forms dynamically (28). Superposition of the cocrystal structures of GcoA with guaiacol and syringol reveals a shift in the positions of GcoA-F169 and the reactive syringol methoxy group relative to the heme (see below). Functionally, these shifts in the GcoA-F169 position permit binding of syringol, although binding in the productive conformation as measured by the shift in Fe(III) spin state is substantially diminished relative to binding of guaiacol (Table 1) (28).

We hypothesized that mutation of GcoA-F169 to a smaller residue (alanine) may relieve the apparent steric clash between it and the bound ligand in the active site, allowing syringol to adopt a productive conformation. Thus, we prepared the GcoA-F169A variant and measured its guaiacol and syringol binding properties (SI Appendix, Fig. S1). Both ligands were able to stimulate the Fe(III) spin state conversion at levels close to wild-type (WT) (Table 1), with K_D values in the low micromolar range. Notably, 3MC also bound and induced a spin state change in Fe(III), though with less affinity than guaiacol or syringol. We concluded that active site engineering could indeed lead to productive syringol binding and potentially to turnover by GcoA-F169 variants, and thus subsequently studied their reactivity with guaiacol, syringol, and 3MC.

GcoA-F169A Efficiently Demethylates both Guaiacol and Syringol with only Limited Uncoupling. Substrate analogs are known to stimulate the P450 reaction with NADH/O₂ without concomitant oxygenation of the organic substrate. This leads to uncoupling of the NADH and substrate oxidation reactions and to reduction of O₂ to either H₂O₂ or H₂O (34). Previous work has shown that syringol stimulates NADH consumption by WT GcoA (28), although without substantial syringol turnover. To address whether guaiacol and/or syringol would serve as substrates of GcoA-F169A, we monitored the disappearance of NADH (UV/vis) and aromatic substrate (HPLC) over time (Scheme 1). The rates of organic substrate and NADH consumption were robust and similar within error regardless of whether guaiacol or syringol

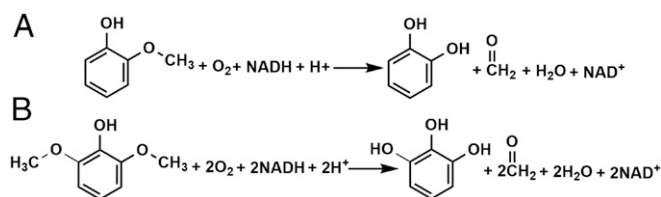
Table 1. Efficacy of GcoA-F169A relative to WT GcoA in binding and demethylating guaiacol and syringol

Parameter	WT GcoA			GcoA-F169A		
	Guaiacol	Syringol	3MC	Guaiacol	Syringol	3MC
K_D , μM^*	0.0065 ± 0.002	2.8 ± 0.5	3.7 ± 0.1	7.1 ± 0.1	1.7 ± 0.07	9.5 ± 0.2
%Fe(III) spin state conversion*	87 ± 1	56 ± 2	62 ± 0.01	72 ± 0.3	76 ± 0.7	65 ± 0.6
k_{cat} , $\text{s}^{-1}\dagger$	6.8 ± 0.02	—	n/a [‡]	11 ± 0.03	5.9 ± 0.01	n/a
K_M , mM^\dagger	60 ± 10	—	n/a	40 ± 6	10 ± 1	n/a
k_{cat}/K_M , $\text{mM}^{-1}\text{s}^{-1}\dagger$	110 ± 20	—	n/a	290 ± 40	600 ± 90	n/a

* K_D was measured by titrating 0–60 μM of substrate into a solution containing 2–6 μM WT or F169A GcoA in air-saturated buffer (25 mM Hepes, 50 mM NaCl, pH 7.5, 25 °C) and recording the ferric spin state change from the low-spin (417 nm) to high-spin (388 nm) species. The % spin state conversion was calculated by dividing the final high-spin species (maximum absorbance at 388 nm) by the starting low-spin species (maximum absorbance at 417 nm).

[†]The Michaelis constants are apparent, as the dioxygen and GcoB concentrations are not known to be saturating. The conditions used were 0.2 μM GcoAB, 100 $\mu\text{g}/\text{mL}$ catalase, 300 μM NADH, 210 μM O₂, and 0–300 μM substrate, 25 °C, 25 mM Hepes, 50 mM NaCl, pH 7.5.

[‡]n/a and dashes: Michaelis constants for 3MC could not be directly measured because of a high level of NADH uncoupling, indicated by “n/a”; see text. A substantial amount of syringol turnover was not observed for WT GcoA, again making it impossible to measure Michaelis parameters. This is indicated by dashes.



Scheme 1. *O*-demethylation of guaiacol to form catechol and formaldehyde (**A**) and syringol to form pyrogallol and two formaldehydes (**B**). The singly demethylated species, 3MC, is expected to form as an intermediate in reaction (**B**). See Fig. 1.

was used (*SI Appendix*, Table S1). This suggests that both guaiacol and syringol serve as substrates for GcoA-F169A.

Moreover, the oxidative *O*-demethylation of guaiacol appeared to be largely coupled to NADH consumption. When NADH and O_2 were present in excess of guaiacol, the measured stoichiometry of the GcoA-F169A-catalyzed reaction was very close to 1 molecule each of guaiacol and NADH consumed to 1 formaldehyde and 1 catechol produced (Fig. 1), without overconsumption of NADH ($103 \pm 7\%$ coupling efficiency; *SI Appendix*, Table S1).

Since both methoxy groups of syringol can potentially serve as substrates, we examined syringol turnover in several ways. Syringol (100 μ M) was first incubated with NADH (200 μ M) and excess dissolved O_2 (210 μ M), and the reaction with the GcoA-F169A mutant was allowed to go to completion. As with the guaiacol reaction, all the NADH and syringol were consumed (Fig. 1), implying that syringol undergoes two *O*-demethylations, producing 3MC and then pyrogallol. However, less formaldehyde ($170 \pm 10 \mu$ M) was produced than expected, suggesting some uncoupling of NADH/ O_2 consumption from the oxidative *O*-demethylation. Consistent with that hypothesis, $50 \pm 4 \mu$ M of the singly demethylated intermediate 3MC was observed at the end of the reaction, even though sufficient NADH/ O_2 were present to enable its complete conversion to pyrogallol. Notably, pyrogallol was not detected under any of the conditions used here, possibly due to its well-known instability in the presence of O_2 (35).

The stoichiometric analysis was next repeated with NADH and syringol present in equal concentrations ($\sim 200 \mu$ M each; 210 μ M O_2), conditions expected to permit at most one-half of the available methoxy groups to react. All of the NADH and $150 \pm 6 \mu$ M syringol were consumed, and $200 \pm 3 \mu$ M formaldehyde and $120 \pm 2 \mu$ M 3MC were generated (Fig. 1 and *SI Appendix*, Table S2). The accumulation of roughly one-half an equivalent of 3MC (relative to NADH) under these conditions suggested that the first *O*-demethylation of syringol must be faster than the second, and that the uncoupling reaction is likely stimulated by 3MC rather than by syringol. Consistent with those expectations, the rate of 3MC disappearance measured by HPLC was significantly slower than the disappearance of either guaiacol or syringol ($2.6 \pm 0.3 \mu$ M 3MC $s^{-1} \mu$ mol GcoA-F169A $^{-1}$ vs. $5.1 \pm 0.8 \mu$ M syringol $s^{-1} \mu$ mol GcoA-F169A $^{-1}$; *SI Appendix*, Table S1). Moreover, the faster consumption of NADH relative to 3MC suggested diminished reaction coupling ($64 \pm 10\%$ coupling; *SI Appendix*, Tables S1 and S2). In reactions containing 100 μ M 3MC and 200 μ M NADH, the majority of the initially available NADH was consumed, and $\sim 100 \mu$ M of formaldehyde was produced (Fig. 1). We hypothesized that the observed overconsumption of NADH was due to the uncoupled reaction, leading to H_2O_2 production. The production of H_2O_2 in the presence of excess NADH/ O_2 and limiting 3MC was confirmed using Amplex Red and horseradish peroxidase (HRP) (*SI Appendix*, Table S3). Consequently, accurate values for k_{cat} and K_M could not be measured using 3MC as a substrate (Table 1).

A broader survey of variants at the GcoA-F169 position (amino acids S, H, V, I, and L in addition to A) confirmed that GcoA-F169A exhibits the best catalytic performance in terms of both specific activity and reaction coupling, although other small side chains (S and V) also permitted reactivity with syringol, suggesting that these may permit syringol to assume a reactive conformation at the heme. Apparent steady-state kinetic parameters measured in air and at potentially subsaturating concentrations of GcoB (Table 1 and *SI Appendix*, Fig. S2 and Table S2) suggest that GcoA-F169A is a more effective catalyst toward the first methoxy group of syringol relative to guaiacol, with $k_{cat}/K_M[\text{syringol}]$ nearly double $k_{cat}/K_M[\text{guaiacol}]$. Moreover, GcoA-F169A has a slightly improved performance with guaiacol as a substrate relative to the WT enzyme.

Structural Analysis Reveals Productive Syringol Reorientation in GcoA-F169 Variants. Superposition of the structures of GcoA-ligand complexes indicated a significant rotation and translation of bound syringol relative to guaiacol (Fig. 2A), and we hypothesized that this could form the basis for the unproductive syringol uncoupling in the native enzyme. The comparative distances between the heme and the proximal methoxy carbon of guaiacol vs. syringol are within 0.4 Å, and even closer between the heme and methoxy oxygens (within 0.1 Å). In addition, there is no significant deviation from the plane of the aromatic rings between these ligands. In contrast, the angle of presentation of the methoxy to the heme diverges significantly in these complexes. Using the angle between the methoxy oxygen, heme iron, and terminal methoxy carbon atoms [O-Fe(III)-C] as a convenient readout of relative orientation, there is a 55% increase in angle from the guaiacol-bound structure (8.3°) compared with the syringol-bound structure (12.9°).

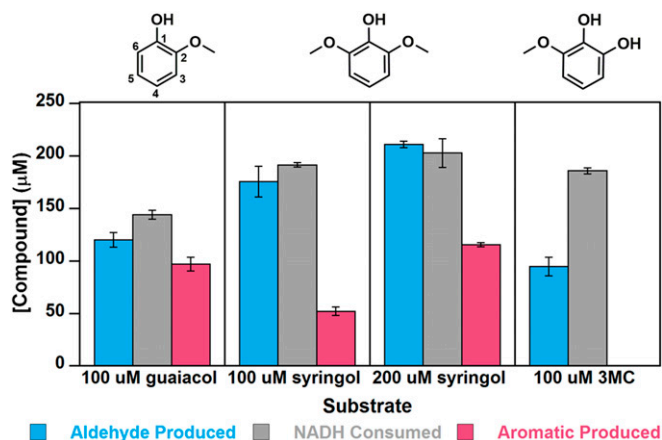


Fig. 1. Quantitative analyses of substrate consumption and product generation indicate nearly complete coupling of NADH/ O_2 consumption to substrate *O*-demethylation for guaiacol and progressively more uncoupling for syringol and 3MC. NADH (200 μ M) and guaiacol, syringol, or 3MC (100 or 200 μ M) were incubated in air with 0.2 μ M GcoA-F169A and GcoB (each with 25 mM Hepes, 50 mM NaCl, pH 7.5 at 25 °C and 210 μ M O_2). Reactants and products were quantified when the UV/vis spectrum ceased changing and the reaction was deemed complete. The total NADH consumed is compared with the amounts of formaldehyde and demethylated aromatic compound produced. Pyrogallol, the *O*-demethylated product of 3MC, is unstable in air under the conditions used in the assay and was not detected. Error bars represent ± 1 SD from three or more independent measurements. *P* values comparing NADH consumption and formaldehyde production were 0.035 for guaiacol, 0.20 for 100 μ M syringol, 0.41 for 200 μ M syringol, and 0.0035 for 3MC. For NADH consumption and aromatic product production, these *P* values were 0.011 for guaiacol, 0.00031 for 100 μ M syringol, and 0.0084 for 200 μ M syringol.

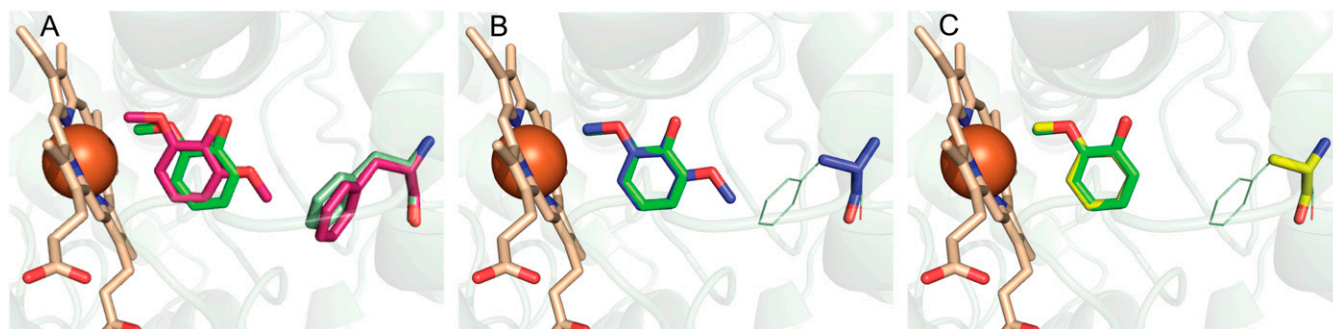


Fig. 2. Structure-guided active site engineering of GcoA. Superpositions of WT and GcoA-F169A ligand-bound structures of GcoA, the P450 monooxygenase component of GcoAB, are shown. The heme is colored in bronze stick. (A) The guaicol (green) and syringol (pink) complexes with WT GcoA are shown with the position of the GcoA-F169 residue highlighted. The translation and rotation of syringol compared with guaicol result in a shift of the target methoxy carbon away from the heme. The Fe(III) to guaicol methoxy carbon distance is 3.9 Å, and the Fe(III) to proximal syringol methoxy carbon distance is 4.3 Å. Data from PDB ID: 5NCB (28) and 5OMU (28). (B) The engineered GcoA-F169A-syringol complex (blue) enables positioning of the reactive methoxy group relative to the heme in a mode consistent with productive guaicol binding. GcoA-F169 from the guaicol-bound WT structure is shown in green lines. (C) Superposition of the WT (green) and GcoA-F169A (yellow) guaicol-bound complexes reveals that guaicol sits in an identical position in both crystal structures. GcoA-F169 from the guaicol-bound WT structure is shown in green lines.

To investigate this further, we generated multiple high-resolution GcoA-F169 variant cocrystal structures. A set of syringol-bound structures (*SI Appendix*, Table S4 and Fig. S4) provided direct insight into the minimal reduction in side chain bulk required to achieve the productive binding mode equivalent to that of guaicol. A stepwise trajectory of the bound syringol toward this optimum orientation with decreasing side chain bulk was observed in the superposition of four cocrystal structures (*SI Appendix*, Fig. S34). Specifically, GcoA-F169H creates an improved substrate orientation, further improved by GcoA-F169V, and essentially optimized in both the GcoA-F169S and -F169A proteins (*SI Appendix*, Fig. S34). Indeed, a comparison of the GcoA-F169A syringol structure with the WT guaicol structure revealed an almost perfect alignment relative to the aromatic rings of each substrate and a methoxy O-Fe-C angle of 8.6°, within 0.3° of that observed for guaicol (Fig. 2B).

Each protein variant also crystallized successfully with guaicol (*SI Appendix*, Table S5 and Fig. S5) and the structures showed that the orientation of the bound ligand remained consistent with that of the WT enzyme (*SI Appendix*, Fig. S3B). Even the largest reduction in side chain bulk, represented by the GcoA-F169A variant, retained the ideal reactive geometry for the natural substrate (Fig. 1C). Furthermore, comparison of the surrounding active site architecture confirmed no significant deviation from the WT. The resolution of these structures (1.66–2.17 Å) also provided sufficient electron density quality to explore changes in the hydration of the pocket (*SI Appendix*, Fig. S6). While the native enzyme excluded water from the active site pocket, we were interested to see whether this was maintained when a new cavity in the pocket was introduced. The syringol-bound mutants A, S, and V contain an additional ordered water in the active site (*SI Appendix*, Fig. S6), which may help maintain the substrate in a productive binding pose for catalysis. As expected, the bulkier GcoA-F169H mutant excluded water from the active site, as with the WT structure.

Syringol Clashes with both GcoA-F169 and the Substrate Access Lid in Simulations of WT GcoA. In the WT enzyme with syringol bound, active site crowding can be relieved in several ways. First, as already noted, syringol can shift toward the heme (Fig. 24); this effect is seen in MD simulations (80 ns) carried out on WT GcoA with either bound guaicol or syringol, although the effect is much subtler than in the crystal structures (*SI Appendix*, Fig. S7). A second effect is more pronounced in MD simulations, which show that GcoA-F169 is significantly more flexible and perturbed from the crystal structure position when syringol

rather than guaicol is bound at the WT active site (Fig. 3A, C, and E, *SI Appendix*, Fig. S8, and *Movies S1* and *S2*). This effect is complemented by the static picture given by the crystal structures (Fig. 24), showing that GcoA-F169 is “pushed away” from the substrate by a distance commensurate with the observed shift of the substrate.

Opening of the substrate access loop, a larger-scale phenomenon closely related to the movement of GcoA-F169, can also relieve active site crowding. All GcoA crystal structures reported to date present a closed active site “lid,” but previous MD simulations demonstrated the ability of the F/G helices (and their connecting loop) to move away from the active site, thus exposing the active site to solution, particularly in the apo form (28). Crystal packing may hinder lid opening in the GcoA-F169A structure; thus, the first two effects (shifting of substrate and GcoA-F169) are more pronounced in crystal structures. However, in MD of GcoA in solution, the active site loop is unconstrained, and the effect of the GcoA-F169 clash with syringol is observed less at the substrate and more on the enzyme. This includes the positioning and flexibility of GcoA-F169 (as mentioned above) and the substrate access lid (Fig. 3B, D, and F; *SI Appendix*, Fig. S9; and *Movies S3* and *S4*), which is significantly more prone to open with syringol bound in WT GcoA than with guaicol, as well as either substrate bound in MD simulations on the GcoA-F169A mutant. Open and closed access loops have been observed in P450 enzyme [P450cam (34) and BM3 (35, 36)] crystal structures. The open GcoA configurations that we observe in MD simulations are approximately one-half as open as the aforementioned open crystal structures and possibly not sufficiently open to allow substrate ingress and egress. However, this principle observed over 80 ns is likely to be more pronounced over the course of the full catalytic cycle. We also note that to date, efforts to crystallize GcoA in the apo state have proven unsuccessful; when achieved, this may reveal a more open configuration of the substrate access lid.

The above conclusions from 80-ns MD simulations are also supported by a deeper analysis of three independent 1-μs MD trajectories of WT GcoA with syringol and guaicol bound at the active site, which were originally presented in our previous study (figure S21 in ref. 28). GcoA-F169 is significantly perturbed from its crystal structure position and more mobile; this coincides with an increased propensity to open the substrate access lid, which is the only region of significant difference in flexibility (*SI Appendix*, Fig. S10).

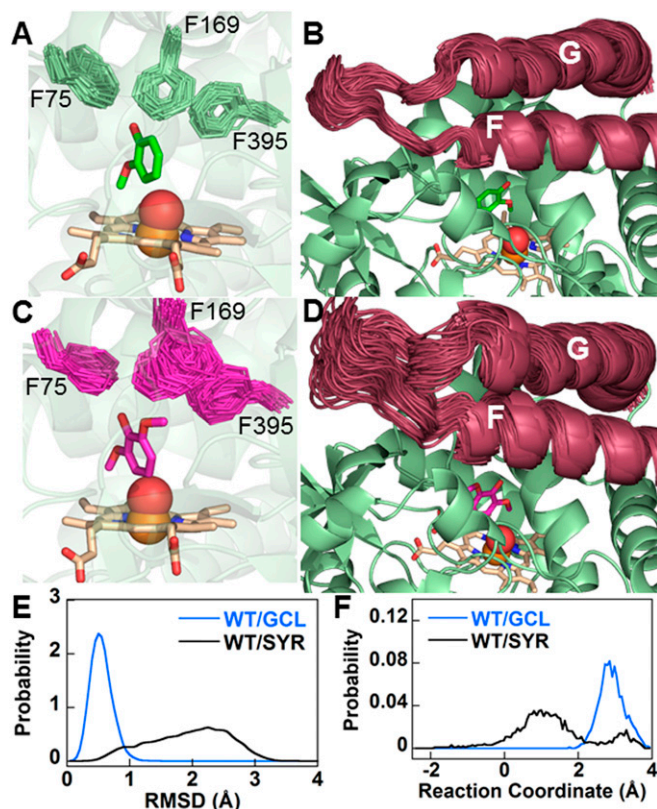


Fig. 3. GcoA-F169 in WT GcoA and the substrate access loop are significantly displaced with bound syringol. MD simulations with bound guaiacol indicate that GcoA-F169 (A) and the substrate access lid (B) are relatively stable (Movies S1 and S3). Introducing syringol results in increased flexibility of GcoA-F169 (C) and the substrate access loop (D) (Movies S2 and S4). In A–D, the position of each of the labeled Phe side chains (or, alternatively, the substrate access loop) is shown every 4 ns over the course of the 80-ns MD simulation. Substrate, the Phe side chains, and heme are shown in sticks, and the Fe atom and the O atom of a reactive heme-oxo intermediate are shown as spheres. Probability distributions are shown for the rmsd of the six ring carbons of GcoA-F169 from their crystal structure positions (E) and the reaction coordinates for opening/closing of the substrate access loop (as defined in *SI Appendix*; lower values indicate more open configurations) (F).

We also performed density functional theory (DFT) calculations on a truncated active site model, demonstrating that the *O*-demethylation of syringol proceeds via a similar pathway as previously described for guaiacol (28) (*SI Appendix*, Fig. S11 and Table S6). Optimized transition state geometries and free energy barriers for the rate-limiting hydrogen atom transfer are likewise very similar in the two cases. In addition, replica exchange thermodynamic integration (RETI) simulations were conducted to examine relative free energies associated with substrate binding and mutating GcoA-F169 in the closed state of the enzyme (*SI Appendix*, Figs. S12 and S13). These RETI simulations revealed additional substrate binding modes, made possible by the “softer” interactions between the substrate and enzyme as the simulations gradually “turn on” and “turn off” electrostatic and van der Waals components of intermolecular interactions, and quantify the effect of this binding flexibility on binding thermodynamics.

Sequence Position 169 in CYP255A Enzymes Is Highly Variable. GcoA belongs to the CYP255A family of cytochrome P450 enzymes (27). Conservation analyses of GcoA homologs revealed a notably variable position 169 among active site residues. Moreover, GcoA-F169 not only is the least conserved of the triad of phenylalanine residues in the active site, it is also among the least

conserved positions in the entire protein (Fig. 4 and *SI Appendix*, Fig. S14 and Table S7). From a multiple sequence alignment, we determined that alanine and phenylalanine are the most frequent residues used by CYP255A enzymes at position 169, with alanine present in the majority of sequences. Thus, the GcoA-F169A mutant, which showed enhanced turnover on guaiacol and syringol, is closer to the CYP255A consensus protein than the WT. It is interesting that although none of the GcoA homologs in our analyses exhibits a histidine at position 169, the GcoA-F169H mutant was stable and showed the highest specific activity on guaiacol. Next to GcoA-F169, A295 and T296 show the greatest variability of residues in the active site. Besides these, other residues within 6 Å of the center of mass of the guaiacol substrate generally show high conservation.

F169A Enables In Vivo Syringol Conversion by GcoA. Finally, we aimed to demonstrate in vivo conversion of syringol to pyrogallol using *Pseudomonas putida* KT2440, which has many native aromatic-catabolic pathways relevant to lignin conversion (8, 30, 37, 38). To accomplish this, we transformed plasmids expressing WT GcoA or the GcoA-F169A variant and GcoB into a strain that constitutively overexpresses PcaHG (*SI Appendix*, Tables S8 and S9), a native 3,4-protocatechuate dioxygenase from *P. putida* that converts pyrogallol into 2-pyrone-6-carboxylate, a more stable product (Fig. 5A) (39). ¹H NMR analysis of the culture medium revealed that, when cultured with 20 mM glucose and 1 mM syringol, peaks corresponding to syringol completely disappeared in cells expressing GcoA-F169A (AM157) after 6 h, which coincided with the appearance of 3MC, pyrogallol, and 2-pyrone 6-carboxylate (Fig. 5B and *SI Appendix*, Fig. S15). A standard for 2-pyrone 6-carboxylate was not available; however, we did verify the presence of 2-pyrone 6-carboxylate using LC-MS-MS (*SI Appendix*, Fig. S15). While small amounts of pyrogallol, 3MC, and 2-pyrone 6-carboxylate were also observed with the WT GcoAB (AM156), nearly 60% of the original syringol remained after 6 h. Neither pyrogallol nor 3MC was observed in the strain lacking GcoAB (AM155). These data indicate that the F169A variation enhances in vivo *O*-demethylation of syringol to 3MC and pyrogallol by GcoA.

Discussion and Conclusions

The creation of enzymes that overcome the challenge of lignin heterogeneity through increased substrate promiscuity is an attractive goal, but this might come at the cost of reduced activity toward the natural substrate. Unexpectedly, we found that GcoA-F169A not only binds both guaiacol and syringol in a

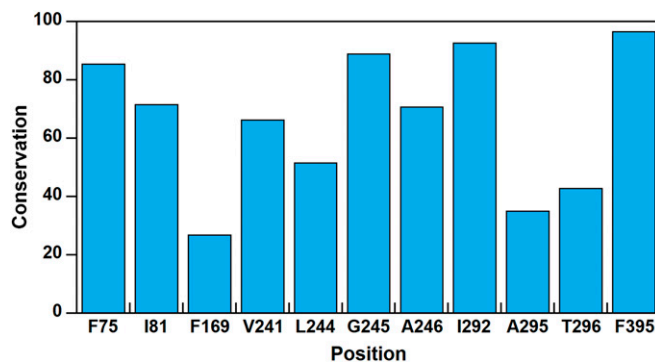


Fig. 4. Bioinformatic analysis of CYP255A sequences indicates variability in the 169th sequence position. Conservation of residues within 6 Å of guaiacol in GcoA, determined via an analysis of protein sequences from 482 GcoA homologs. (Details provided in *SI Appendix*.) Conservation scores are reported as percentiles. F169 is less conserved than 73% of the positions in GcoA. Data from PDB ID: 5NCB (28).

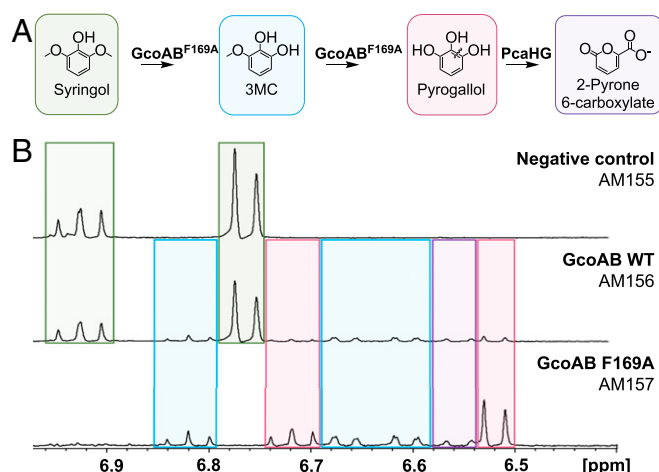


Fig. 5. GcoA-F169A converts syringol in vivo. (A) A pathway for in vivo syringol *O*-demethylation to pyrogallol and cleavage to 2-pyrone 6-carboxylate is proposed. (B) After 6 h, strains were analyzed for their ability to turn over syringol via ¹H NMR spectroscopy. Syringol (green) is completely converted to pyrogallol (pink), 3MC (blue), or 2-pyrone 6-carboxylate (purple) in AM157. The WT GcoA enzyme in AM156 shows only small amounts of conversion. AM155, which does not express GcoA, shows no conversion.

productive orientation analogous to guaiacol in WT GcoA (Fig. 2), but also is more catalytically efficient for *O*-demethylation of both guaiacol and syringol relative to WT, where *O*-demethylation is well coupled to NADH oxidation (Fig. 1 and Table 1). Along with this biochemical observation, the bioinformatics analysis shows that alanine is the most prevalent residue in the 169th sequence position in the CYP255A family. Considering these findings together, it is surprising that the WT GcoA does not possess an alanine at position 169 if guaiacol is the primary substrate, as was assumed in the original reports of GcoAB (17, 27, 28). Given that the GcoA-F169A mutation results in improved turnover of guaiacol, we speculate that either guaiacol is not the primary substrate or there has been little evolutionary pressure in *Amycolatopsis* sp. ATCC 39116 for improved turnover of guaiacol. Another potential explanation could be that syringol *O*-demethylation and subsequent ring cleavage lead to dead-end products that cannot be catabolized and may be toxic to the microbe; thus, GcoA-F169 could function to prevent natural syringol catabolism.

As a first step toward enabling syringol catabolism, in vivo experiments validated the in vitro studies by illustrating efficient *O*-demethylation of syringol and 3MC by the GcoA-F169 mutant. While 2-pyrone 6-carboxylate was detected, pyrogallol is a poor substrate of PcaHG, as most of the intermediate is lost to oxidation (SI Appendix, Fig. S15), and whether 2-pyrone-6-carboxylate can be catabolized further is unclear. Future work could focus on identifying or developing a dioxygenase capable of cleaving pyrogallol to a product that could be further metabolized. Coupled with GcoAB, the *meta*-cleavage pathway of *P. putida* mt-2 might enable complete assimilation of syringol if pyrogallol can be efficiently cleaved to 2-hydroxyruconate.

Cytochrome P450 systems represent one of the most versatile classes of enzymes, making them an ideal target for engineering enhanced activity and substrate promiscuity. Our system is a prime example. The mutation of a single residue resulted in efficient turnover of an S-lignin substrate, syringol, in addition to the native G-subunit substrate, guaiacol, which is not efficiently achieved in the WT enzyme. The plasticity of the GcoA active site may be amenable to even more modifications, allowing us to encompass other lignin monomers as substrates, such as four-

substituted compounds (e.g., vanillin, syringaldehyde). Indeed, in previous work, we showed that T296 sterically clashes with the C4 position of vanillin, preventing *O*-demethylation (28). Interestingly, the 296 position is also quite variable according to the bioinformatics analysis (Fig. 4). The results described here suggest that we may use a similar structure-guided approach to investigate the activity of several T296 variants on 4-substituted compounds. As a large number of lignin degradation products contain *para*-substituted R groups that are bulkier than hydrogen, creating an engineered bacterium in which a minimal number of genetic insertions lead to maximal lignin bioconversion is of keen interest for future work. More broadly, the evolutionary trajectory, substrate specificity, and catalytic efficiency of the CYP255A family of aromatic *O*-demethylases will be the subject of future work aimed at elucidating the relevance of this cytochrome P450 family for microbial lignin conversion.

Methods

Protein Expression and Purification. Mutagenesis was performed using primers listed in SI Appendix, with the Q5 polymerase and KLD enzyme mix (New England BioLabs) according to the manufacturer's protocol. Proteins were expressed as described previously (28).

Crystallization and Structure Determination. Crystallization, diffraction experiments, and structure solution were carried out as described previously (28).

Biochemical Characterization.

Heme quantification of GcoA-F169 mutants. Catalytically active heme bound to each GcoA mutant was determined as reported previously (28, 40) and described in detail in SI Appendix.

Determination of [FAD] and nonheme [Fe] in GcoB. The FAD and 2Fe-2S contents of GcoB were measured as reported previously (28, 41) and described in detail in SI Appendix.

Steady-state kinetics of GcoA-F169A. The *O*-demethylation reactions of guaiacol, syringol, and 3MC were continuously monitored using the NADH consumption assay as reported previously (28) and described in detail in SI Appendix.

Determination of substrate dissociation constants (K_D) with GcoA-F169A. The equilibrium binding constant, K_D , for GcoA-F169A and guaiacol, syringol, and 3MC was determined as described previously (28) and detailed in SI Appendix.

Formaldehyde determination. The [formaldehyde] produced on reaction with GcoA-F169A and substrates was determined using a colorimetric assay with tryptophan (28, 42); details are provided in SI Appendix.

HPLC for product identification and specific activity measurement. HPLC was used to verify the *O*-demethylated product of GcoA-F169A GcoA/GcoB with guaiacol, syringol, or 3MC. In addition, discontinuous HPLC was used to determine the specific activity of aromatic product disappearance. Details are provided in SI Appendix.

Detection of H_2O_2 via HRP and Amplex Red assay. A colorimetric assay involving HRP and Amplex Red was used to quantify H_2O_2 in the reaction between GcoA-F169A GcoA/GcoB, NADH and guaiacol, syringol, or 3MC, as described in SI Appendix.

MD, DFT, and Bioinformatics. MD simulations and DFT calculations were performed following a similar methodology as in our previous work (28). Full details of the computational methods and references are provided in SI Appendix. A total of 482 homologous CYP255A sequences were retrieved from a blastp search against GcoA. After multiple sequence alignment, conservation was analyzed from relative entropy calculations for each site. Further details are provided in SI Appendix.

In Vivo Syringol Utilization. Strains used for shake flask experiments were grown overnight in LB media and resuspended the next day in M9 minimal media with 20 mM glucose, as described in SI Appendix. Cells were grown until they reached an OD₆₀₀ of approximately 1, at which point syringol was added at a final concentration of 1 mM. ¹H NMR spectroscopy was used to analyze syringol consumption.

Data Deposition. The atomic coordinates and structure factors have been deposited in the Protein Data Bank, <https://www.rcsb.org> (PDB ID codes 6HQK, 6HQL, 6HQM, 6HQN, 6HQO, 6HQP, 6HQQ, 6HQR, 6HQS, and 6HQT).

ACKNOWLEDGMENTS. We thank Josh Vermaas for his assistance with the simulation setup and Mark Allen and the staff at Diamond Light Source for their support and contributions to X-ray crystallography. We also thank Brenna Black and Bill Michener for their assistance with analytics. This material is also based on work supported by (while C.M.P. is serving at) the National Science Foundation (NSF). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. Support for this work was provided by the NSF (Grants MCB-1715176, M.M.M. and J.L.D.; CHE-1361104, to K.N.H.; and CBET-1552355, to C.M.P.) and the Biotechnology and Biological Sciences Research Council (Grants BB/P011918/1 and BB/L001926/1, to J.E.M.). This work was partially authored by the Alliance for Sustainable Energy, LLC, the manager and operator of the National Renewable Energy Laboratory for the US Department of Energy (DOE) (Contract DE-AC36-08GO28308). B.C.K., A.W.M., L.B., G.P.S., M.F.C., C.W.J., and G.T.B. received funding for the MD simulations and in vivo experiments from the US DOE, Office of Energy Efficiency and Renewable Energy (EERE), Bioenergy Tech-

nologies Office. M.M.M. received funding support from the US DOE, Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research program, which is administered by the Oak Ridge Institute for Science and Education for the DOE under Contract DE-SC0014664. G.T.B. and J.E.M. received funding for the structural biology efforts from The Center for Bioenergy Innovation, a US DOE Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. M.G.-B. is the recipient of a Ramón Areces Foundation postdoctoral fellowship. Computer time was provided by Extreme Science and Engineering Discovery Environment, supported by NSF Grant ACI-1548562 via allocation MCB-090159 to G.T.B. at the Pittsburgh Supercomputing Center (Bridges) and the Texas Advanced Computing Center (Stampede2), by the National Renewable Energy Laboratory Computational Sciences Center supported by the DOE Office of EERE under Contract DE-AC36-08GO28308, and by the UCLA Institute for Digital Research and Education. K.N.H. received financial support from the National Institute for General Medical Sciences (Grant GM-124480).

- W. Boerjan, J. Ralph, M. Baucher, Lignin biosynthesis. *Annu. Rev. Plant Biol.* **54**, 519–546 (2003).
- G. Fuchs, M. Boll, J. Heider, Microbial degradation of aromatic compounds—from one strategy to four. *Nat. Rev. Microbiol.* **9**, 803–816 (2011).
- T. D. Bugg, M. Ahmad, E. M. Hardiman, R. Singh, The emerging role for bacteria in lignin degradation and bio-product formation. *Curr. Opin. Biotechnol.* **22**, 394–400 (2011).
- D. Floudas *et al.*, The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* **336**, 1715–1719 (2012).
- D. Salvachúa *et al.*, Lignin depolymerization by fungal secretomes and a microbial sink. *Green Chem.* **18**, 6046–6062 (2016).
- E. Masai, Y. Katayama, M. Fukuda, Genetic and biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds. *Biosci. Biotechnol. Biochem.* **71**, 1–15 (2007).
- G. T. Beckham, C. W. Johnson, E. M. Karp, D. Salvachúa, D. R. Vardon, Opportunities and challenges in biological lignin valorization. *Curr. Opin. Biotechnol.* **42**, 40–53 (2016).
- J. G. Linger *et al.*, Lignin valorization through integrated biological funneling and chemical catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 12013–12018 (2014).
- T. D. Bugg, R. Rahmanpour, Enzymatic conversion of lignin into renewable chemicals. *Curr. Opin. Chem. Biol.* **29**, 10–17 (2015).
- O. Y. Abdelaziz *et al.*, Biological valorization of low molecular weight lignin. *Bio-technol. Adv.* **34**, 1318–1346 (2016).
- R. Rinaldi *et al.*, Paving the way for lignin valorisation: Recent advances in bio-engineering, biorefining and catalysis. *Angew. Chem. Int. Ed. Engl.* **55**, 8164–8215 (2016).
- W. Schutyser *et al.*, Chemicals from lignin: An interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem. Soc. Rev.* **47**, 852–908 (2018).
- A. J. Ragauskas *et al.*, Lignin valorization: Improving lignin processing in the biorefinery. *Science* **344**, 1246843 (2014).
- Z. Sun, B. Fridrich, A. de Santi, S. Elangovan, K. Barta, Bright side of lignin depolymerization: Toward new platform chemicals. *Chem. Rev.* **118**, 614–678 (2018).
- A. Dardas *et al.*, The demethylation of guaiacol by a new bacterial cytochrome P-450. *Arch. Biochem. Biophys.* **236**, 585–592 (1985).
- U. Karlson *et al.*, Two independently regulated cytochromes P-450 in a Rhodococcus rhodochrous strain that degrades 2-ethoxyphenol and 4-methoxybenzoate. *J. Bacteriol.* **175**, 1467–1474 (1993).
- L. D. Eltis, U. Karlson, K.N. Timmis, Purification and characterization of cytochrome P450RR1 from Rhodococcus rhodochrous. *Eur. J. Biochem.* **213**, 211–216 (1993).
- A. Segura, P. V. Bünz, D. A. D'Argenio, L. N. Ornston, Genetic analysis of a chromosomal region containing vanA and vanB, genes required for conversion of either ferulate or vanillate to protocatechuate in *Acinetobacter*. *J. Bacteriol.* **181**, 3494–3504 (1999).
- B. Morawski, A. Segura, L. N. Ornston, Substrate range and genetic analysis of *Acinetobacter* vanillate demethylase. *J. Bacteriol.* **182**, 1383–1389 (2000).
- E. Masai *et al.*, A novel tetrahydrofolate-dependent O-demethylase gene is essential for growth of *Sphingomonas paucimobilis* SYK-6 with syringate. *J. Bacteriol.* **186**, 2757–2765 (2004).
- T. Abe, E. Masai, K. Miyauchi, Y. Katayama, M. Fukuda, A tetrahydrofolate-dependent O-demethylase, LigM, is crucial for catabolism of vanillate and syringate in *Sphingomonas paucimobilis* SYK-6. *J. Bacteriol.* **187**, 2030–2037 (2005).
- T. Yoshikata *et al.*, Three-component O-demethylase system essential for catabolism of a lignin-derived biphenyl compound in *Sphingobium* sp. strain SYK-6. *Appl. Environ. Microbiol.* **80**, 7142–7153 (2014).
- A. Harada *et al.*, The crystal structure of a new O-demethylase from *Sphingobium* sp. strain SYK-6. *FEBS J.* **284**, 1855–1867 (2017).
- A. C. Kohler, M. J. L. Mills, P. D. Adams, B. A. Simmons, K. L. Sale, Structure of aryl O-demethylase offers molecular insight into a catalytic tyrosine-dependent mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E3205–E3214 (2017).
- S. G. Bell *et al.*, The crystal structures of 4-methoxybenzoate bound CYP199A2 and CYP199A4: Structural changes on substrate binding and the identification of an anion binding site. *Dalton Trans.* **41**, 8703–8714 (2012).
- S. G. Bell *et al.*, Investigation of the substrate range of CYP199A4: Modification of the partition between hydroxylation and desaturation activities by substrate and protein engineering. *Chemistry* **18**, 16677–16688 (2012).
- M. Tumen-Velasquez *et al.*, Accelerating pathway evolution by increasing the gene dosage of chromosomal segments. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 7105–7110 (2018).
- S. J. B. Mallinson *et al.*, A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioreconversion. *Nat. Commun.* **9**, 2487 (2018).
- S. C. Farrow, P. J. Facchini, Dioxygenases catalyze O-demethylation and O,O-demethylation with widespread roles in benzylisoquinoline alkaloid metabolism in opium poppy. *J. Biol. Chem.* **288**, 28997–29012 (2013).
- D. Salvachúa, E. M. Karp, C. T. Nimlos, D. R. Vardon, G. T. Beckham, Towards lignin consolidated bioprocessing: Simultaneous lignin depolymerization and product generation by bacteria. *Green Chem.* **17**, 4951–4967 (2015).
- S. M. Bölicke, W. Ternes, Isolation and identification of oxidation products of syringol from brines and heated meat matrix. *Meat Sci.* **118**, 108–116 (2016).
- O. E. Adedokun, I. R. Green, M. le Roes-Hill, S. G. Burton, Enzymatic modification of 2,6-dimethoxyphenol for the synthesis of dimers with high antioxidant capacity. *Process Biochem.* **47**, 1926–1932 (2012).
- Y. Y. Wan, Y. M. Du, T. S. Miyakoshi, Enzymatic catalysis of 2,6-dimethoxyphenol by laccases and products characterization in organic solutions. *Sci. China B Chem.* **51**, 669–676 (2008).
- Y.-T. Lee, R. F. Wilson, I. Rupniewski, D. B. Goodin, P450cam visits an open conformation in the absence of substrate. *Biochemistry* **49**, 3412–3419 (2010).
- I. F. Sevrioukova, H. Li, H. Zhang, J. A. Peterson, T. L. Poulos, Structure of a cytochrome P450-redox partner electron-transfer complex. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1863–1868 (1999).
- D. C. Haines, D. R. Tomchick, M. Machius, J. A. Peterson, Pivotal role of water in the mechanism of P450BM-3. *Biochemistry* **40**, 13456–13465 (2001).
- D. R. Vardon *et al.*, Adipic acid production from lignin. *Energy Environ. Sci.* **8**, 617–628 (2015).
- J. Nogales *et al.*, Unravelling the gallic acid degradation pathway in bacteria: The gal cluster from *Pseudomonas putida*. *Mol. Microbiol.* **79**, 359–374 (2011).
- Y. Saeki, M. Nozaki, S. Senoh, Cleavage of pyrogallol by non-heme iron-containing dioxygenases. *J. Biol. Chem.* **255**, 8465–8471 (1980).
- F. P. Guengerich, M. V. Martin, C. D. Sohl, Q. Cheng, Measurement of cytochrome P450 and NADPH-cytochrome P450 reductase. *Nat. Protoc.* **4**, 1245–1251 (2009).
- S. G. Bell, A. B. Tan, E. O. Johnson, L. L. Wong, Selective oxidative demethylation of veratric acid to vanillic acid by CYP199A4 from *Rhodopseudomonas palustris* HaA2. *Mol. Biosyst.* **6**, 206–214 (2010).
- J. Chrastil, J. T. Wilson, A sensitive colorimeter method for formaldehyde. *Anal. Biochem.* **63**, 202–207 (1975).