A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioconversion

Sam J.B. Mallinson1, Melodie M. Machovina2,3, Rodrigo L. Silveira2,4, Marc Garcia-Borràs5, Nathan Gallup2,5, Christopher W. Johnson2, Mark D. Allen1, Munir S. Skaf4, Michael F. Crowley6, Ellen L. Neidle7, Kendall N. Houk5, Gregg T. Beckham2, Jennifer L. DuBois3 & John E. McGeehan1

Microbial aromatic catabolism offers a promising approach to convert lignin, a vast source of renewable carbon, into useful products. Aryl-O-demethylation is an essential biochemical reaction to ultimately catabolize coniferyl and sinapyl lignin-derived aromatic compounds, and is often a key bottleneck for both native and engineered bioconversion pathways. Here, we report the comprehensive characterization of a promiscuous P450 aryl-O-demethylase, consisting of a cytochrome P450 protein from the family CYP255A (GcoA) and a three-domain reductase (GcoB) that together represent a new two-component P450 class. Though originally described as converting guaiacol to catechol, we show that this system efficiently demethylates both guaiacol and an unexpectedly wide variety of lignin-relevant monomers. Structural, biochemical, and computational studies of this novel two-component system elucidate the mechanism of its broad substrate specificity, presenting it as a new tool for a critical step in biological lignin conversion.
Lignin is a heterogeneous, aromatic biopolymer found in abundance in plant cell walls where it is used for defense, structure, and nutrient and water transport. Given its prevalence in plant tissues, lignin is the largest reservoir of renewable, aromatic carbon found in nature. The ubiquitous availability of lignin in the environment, coupled to its inherent structural heterogeneity and complexity, has led to the evolution of microbial strategies to break lignin polymers down to smaller fragments using powerful oxidative enzymes secreted by rot fungi and some bacteria. These lignin oligomers can be further assimilated as carbon and energy sources, through at least four known catabolic paradigms.

The most well-understood aromatic catabolic mechanism, mainly studied in aerobic soil bacteria, relies on the use of non-heme iron-dependent dioxygenases to oxidatively ring-open structurally diverse, lignin-derived aromatic compounds. These dioxygenases act on central intermediate substrates, such as catechol, protocatechuate, and gallate, either in an intra- or extradiol manner. Lignin is primarily based on coniferyl (G) and sinapyl (S) alcohol subunits, which exhibit one or two methoxy groups on the aromatic ring, respectively. Nearly all lignin-derived compounds must therefore be O-demethylated to diols before they can be oxidatively cleaved to generate ring-opened compounds, which are ultimately routed to central carbon metabolism (Fig. 1). More recently, the same aromatic-catabolic pathways have been invoked as a potential means to convert lignin to useful products in biorefineries. O-demethylation is therefore a critical reaction for assimilating lignin-derived carbon in both natural carbon cycling as well as in emerging biotechnology applications.

The importance of O-demethylation has motivated substantial efforts toward the discovery and characterization of enzymes capable of demethylating the methoxy substituents of diverse lignin-derived substrates. For example, Ornston et al. described the VanAB O-demethylase in Acinetobacter baylyi ADP1, which converts vanillate to the central intermediate, protocatechuate, via a Rieske non-heme iron monooxygenase mechanism. VanAB, which is common in many aromatic-catabolic soil bacteria, is active on vanillate analogs, but to our knowledge, has not been reported to be active on other lignin-derived compounds. Masai and colleagues first described LigX from Sphingobium sp. SYK-6, a model bacterium for aromatic catabolism. LigX also employs a Rieske non-heme iron monooxygenase mechanism to demethylate a biphenyl compound representing a common lignin linkage. Masai et al. additionally reported, in SYK-6, two tetrahydrofuran-dependent O-demethylases, LigM and DesA. LigM primarily demethylates vanillate and 3-O-methylgallate, whereas DesA principally demethylates syringate with very weak activity on vanillate.

Earlier reports from Eltis et al. and Bell et al. described cytochrome P450-based demethylation of aromatic compounds, though either the full gene sequences were not reported until recently, or the para-substituted substrate (4-methoxybenzoate) was of limited interest for the lignin degradation problem. Similarly, Dardas et al. found evidence of a P450 in Moraxella GU2 responsible for the O-demethylation of guaiacol and guaethol; however, neither the gene sequence nor identity of the P450 or its reductase partner was isolated.

The relatively narrow substrate specificities elucidated to date for aryl-O-demethylation, coupled to the potentially broad distribution of structurally distinct, methoxylated lignin products found in nature, prompted us to search for alternative mechanisms for this key reaction. Because G-unit monomers constitute a majority of plant-derived lignin, we initially focused on O-demethylation of guaiacol (2-methoxyphenol), which in turn represents the simplest G-unit monomer derivable from lignin. As reported in a companion study, we isolated a cytochrome P450-reductase gene pair, gcoAB, from Amycolatopsis sp. ATCC 39116 (encoding proteins with accession numbers WP_020419855.1 and WP_020419854.1). Introduction of this pair via plasmid-based expression into Pseudomonas putida KT2440, a robust aromatic-catabolic bacterium, was sufficient to confer growth on guaiacol. Here, we report a comprehensive structural, biochemical, and computational description of this new cytochrome P450-based mechanism for aryl-O-demethylation. Unlike other known tetrahydrofuran- or non-heme iron-dependent demethylases, which are fairly substrate specific, the P450-reductase pair characterized here (GcoAB) demethylates diverse aromatic substrates, potentially providing an important advantage in both natural and biotechnological contexts. The results presented here suggest a remarkably flexible active site that may promote promiscuous substrate usage.

**Results**

**GcoA crystal structures suggest broad substrate specificity.** The X-ray crystal structures of GcoA (in complex with guaiacol) and GcoB were determined to resolutions of 1.4 Å and 1.7 Å, respectively (Figs. 2 and 3, Supplementary Figs. 1–7, and Supplementary Table 2). The GcoA structure reveals a typical P450 single-domain architecture with a central heme adjacent to a buried active site, captured with the substrate access loop in the closed position (Fig. 2a, b). GcoA possesses a broadly hydrophobic pocket with the two oxygen atoms of the substrate coordinated by backbone carbonyl and amide nitrogen groups from residues Val241 and Gly245, respectively. A series of hydrophobic amino acids is responsible for positioning the aromatic ring, including a triad of phenylalanine residues lining the active site.
cytochrome P450 reductase or a ferredoxin and ferredoxin reductase, cytochrome P450s require two partner proteins, usually a cytochrome P450 reductase.

GcoB has an unusual three-domain structure. Most bacterial cytochrome P450s require two partner proteins, usually a cytochrome P450 reductase or a ferredoxin and ferredoxin reductase, that transfer electrons from a carrier such as NAD(P)H to the cytochrome. Analysis of the GcoB sequence, however, suggests it contains all of the necessary domains in a single polypeptide, with an N-terminal 2Fe-2S ferredoxin domain followed by an FAD and NAD(P) binding region with homology to ferredoxin-NADPH reductase (FNR) type oxi-do-reductases.

Structurally, the compact, N-terminal 2Fe-2S domain of GcoB (Fig. 3a, b) bears strong homology to putidaredoxin (Pdx), a ferredoxin that transfers electrons to P450cam in the three-component camphor hydroxylase system from P. putida (Supplementary Fig. 6–8). The C-terminal region consists of an FAD-binding domain, containing 6 beta-strands and a single alpha helix, followed by an NADH-binding domain. These C-terminal domains show structural homology to FAD-type cytochrome P450 reductase (CPR) domains in which the C-terminal portion primarily stabilizes the isoalloxazine moiety of FAD while the N-terminal domain coordinates the diphosphate bridge between the flavin and adenosine groups (Fig. 3c, d).

A structural comparison of the NADH-binding domain with the NADPH binding domain from a related CPR is strongly predictive of a binding preference for NADH over NADPH. While the individual domains are structurally similar to other CPR proteins, the overall domain architecture is not, and instead it is highly conserved with reductase proteins such as BenC27, the benzoate 1,2-dioxygenase reductase from A. baylyi ADP1, which supplies electrons to Rieske-type aromatic ring-hydroxylating dioxygenases. This suggests an unexpected convergence in the organization of the reductase partners for these evolutionarily distinct oxygenases.

GcoA and GcoB form a dimer complex in solution. GcoA and GcoB operate as a multi-domain complex; guaiacol is processed exclusively in GcoA, with electrons supplied by the GcoB redox machinery. When expressed and purified individually, each protein was shown to be monomeric in solution using a combination of hydrodynamic methods (Supplementary Table 3). Size-exclusion chromatography revealed a strong interaction between GcoA and GcoB (Supplementary Fig. 9) which was confirmed by analytical ultracentrifugation, indicating that GcoAB is a heterodimer in solution. GcoA has a characteristic basic pocket on the proximal face, previously identified in other P450 systems28–30 as the docking surface for the associated reductase partner (Supplementary Fig. 10). Similarly, the surface of GcoB is predicted to have an acidic patch that interfaces with the
The catalytic cofactors in GcoA and GcoB were spectroscopically characterized prior to describing the reactivity of the enzyme pair (Supplementary Fig. 12–15). The UV/visible spectrum of GcoA exhibits a sharp heme Soret peak at 417 nm and α- and β-bands (Q-bands) at 537 and 567 nm. GcoB has an absorbance maximum at 454 nm, indicative of oxidized FAD, and peaks at 423 nm and 480 nm most likely due to the 2Fe-2S cluster. Total heme, FAD, and 2Fe-2S occupancies for the protein monomers were determined at 0.9 (active heme = 0.8 via a CO-binding assay), 0.7, and 0.8 equivalents, respectively, using measured extinction coefficients: GcoA-heme $\varepsilon_{417\text{ nm}} = 114 \pm 4 \text{ M}^{-1} \text{ cm}^{-1}$; GcoB-FAD $\varepsilon_{454\text{ nm}} = 26.6 \pm 0.2 \text{ M}^{-1} \text{ cm}^{-1}$; GcoB-2Fe-2S $\varepsilon_{423\text{ nm}} = 25.2 \pm 0.1 \text{ M}^{-1} \text{ cm}^{-1}$.

Reduced GcoB exhibits a rhombic EPR signature with temperature saturation behavior typical for a 2Fe-2S cluster.

The reductase activity of GcoB was monitored using NADH or NADPH in a cytochrome c reduction assay (Supplementary Fig. 16–20). In agreement with our comparative structural analysis (Supplementary Fig. 7), the reaction with NADH was over 50-fold faster than with NADPH ($k_{\text{cat}} = 44 \pm 1 \text{ s}^{-1}$, $K_M = 16 \pm 0.2 \text{ M}$, 25 °C, pH 7.5; all activity measurements referenced per active cofactor). Demethylation of guaiacol ($K_D = 6 \text{ M}$; Table 1) was then monitored over time via NADH disappearance at 340 nm under steady-state conditions (Fig. 1b). The resulting $k_{\text{cat}}$ (6.8 ± 0.5 s$^{-1}$) was approximately sixfold less than the $k_{\text{cat}}$ for the GcoB reduction reaction alone, suggesting that the overall rate of the two-enzyme catalytic cycle is limited by steps involving GcoA. The value for $k_{\text{cat}}$ is similar to or greater in magnitude than $k_{\text{cat}}$ for other known, non-P450-type O-aryl-demethylases acting upon their preferred substrates, including LigM (5.8 ± 0.25 s$^{-1}$), LigX (6.1 ± 0.2 s$^{-1}$), or PODA (0.034 s$^{-1}$) (Supplementary Table 4).

GcoAB efficiently demethylates multiple aromatic substrates. The catalytic cofactors in GcoA and GcoB were spectroscopically characterized prior to describing the reactivity of the enzyme pair (Supplementary Fig. 12–15). The UV/visible spectrum of GcoA exhibits a sharp heme Soret peak at 417 nm and α- and β-bands (Q-bands) at 537 and 567 nm. GcoB has an absorbance maximum at 454 nm, indicative of oxidized FAD, and peaks at 423 nm and 480 nm most likely due to the 2Fe-2S cluster. Total heme, FAD, and 2Fe-2S occupancies for the protein monomers were determined at 0.9 (active heme = 0.8 via a CO-binding assay), 0.7, and 0.8 equivalents, respectively, using measured extinction coefficients: GcoA-heme $\varepsilon_{417\text{ nm}} = 114 \pm 4 \text{ M}^{-1} \text{ cm}^{-1}$; GcoB-FAD $\varepsilon_{454\text{ nm}} = 26.6 \pm 0.2 \text{ M}^{-1} \text{ cm}^{-1}$; GcoB-2Fe-2S $\varepsilon_{423\text{ nm}} = 25.2 \pm 0.1 \text{ M}^{-1} \text{ cm}^{-1}$.

Reduced GcoB exhibits a rhombic EPR signature with temperature saturation behavior typical for a 2Fe-2S cluster.

The reductase activity of GcoB was monitored using NADH or NADPH in a cytochrome c reduction assay (Supplementary Fig. 16–20). In agreement with our comparative structural analysis (Supplementary Fig. 7), the reaction with NADH was over 50-fold faster than with NADPH ($k_{\text{cat}} = 44 \pm 1 \text{ s}^{-1}$, $K_M = 16 \pm 0.2 \text{ M}$, 25 °C, pH 7.5; all activity measurements referenced per active cofactor). Demethylation of guaiacol ($K_D = 6 \text{ M}$; Table 1) was then monitored over time via NADH disappearance at 340 nm under steady-state conditions (Fig. 1b). The resulting $k_{\text{cat}}$ (6.8 ± 0.5 s$^{-1}$) was approximately sixfold less than the $k_{\text{cat}}$ for the GcoB reduction reaction alone, suggesting that the overall rate of the two-enzyme catalytic cycle is limited by steps involving GcoA. The value for $k_{\text{cat}}$ is similar to or greater in magnitude than $k_{\text{cat}}$ for other known, non-P450-type O-aryl-demethylases acting upon their preferred substrates, including LigM (5.8 ± 0.25 s$^{-1}$), LigX (6.1 ± 0.2 s$^{-1}$), or PODA (0.034 s$^{-1}$) (Supplementary Table 4).
The products were then analyzed. The total NADH consumed is compared above to the amounts of aldehyde and de-alkylated aromatic compound produced. Error bars represent ±1 standard deviation from three or more independent measurements.

Ten structurally diverse O-methylated aromatic compounds were screened for activity with GcoA. Seven of these induced measurable NADH consumption (Fig. 4) and $K_D$ values spanning 70 nM to 37 μM (Table 1). All nonetheless yielded $K_M$(O-methyl-aryl) values comparable to $K_M$(guaialcol), indicating that all form catalytically productive complexes with GcoA. Only three (vanililate, ferulate, and veratrole) had no detectable binding or catalytic interaction with GcoA.

We next examined whether NADH consumption was coupled to substrate demethylation. Aldehyde and the demethylated aromatic products were quantified and compared to the total concentration of NADH consumed following quenching of a reaction containing saturating amounts of each substrate (Table 1, Supplementary Fig. 19). For all substrates tested except guaethol, in which acetaldehyde is produced, formaldehyde was the expected product. Five of the seven substrates that produced aldehyde (guaialcol, guaethol, 3-methoxy-catechol), -CH$_3$ (2-methyl-anisole), and -2-methyl-anisole) did so in a ~1:1 ratio with NADH consumed (Fig. 4). Syringol and vanillin stimulated NADH turnover but produced less than stoichiometric amounts of formaldehyde. We hypothesized that these two substrates bind in the active site in the same location as guaialcol, displacing water and stimulating the spin-state change that permits reduction of the heme iron$^{26}$. However, the reaction with O$_2$ is uncoupled to substrate demethylation in some proportion of turnovers, likely leading to H$_2$O$_2$ release. For all 7 substrates, the expected hydroxylated-aryl product was detected and its identity confirmed by matching its HPLC retention time with known standards. However, we observed variable stability of these hydroxylated compounds in air; hence, the quantity of aldehyde product is likely a better indicator of the extent to which NADH consumption is coupled to aldehyde production.

Examining the structures of the 10 tested substrates in light of the crystallographic data suggests an emerging structure-activity relationship (Supplementary Fig. 5). First, the efficiency of guaialcol as a substrate shows that the enzyme is capable not only of accommodating the larger ethoxy group (Fig. 2d) but also of catalyzing de-ethylation. Second, there appears to be a varying degree of flexibility in the permissible substituents around the aryl ring. C1 (guaiacol numbering, Fig. 4) can accommodate -OH (guaialcol, guaethol, 3-methoxy-catechol), -CH$_3$ (2-methyl-anisole), or -H (anisole) with comparable efficiency constants [$k_{cat}$/ $K_M$(O-methyl-aryl)]. However, veratrole, which has -OCH$_3$ at this position, is a non-substrate/non-binder, suggesting that steric constraints supplied by the nearby a-helix (Fig. 2c) might limit the size of substrates here. Similarly, the carbon ortho to the -OH of guaiacol can have either -H or -OH (3-methoxy-catechol) substituents without substantial penalty to the efficiency constant, though -OCH$_3$ (syringol) again leads to a partial substrate, partial uncoupler. Finally, substituting the -H at the C4 guaiacol position, which is closest to the side chain of Thr296, with a formyl group (vanillin) likewise yields a partial substrate, partial uncoupler; however, carboxylic acid (vanililate and ferulate) substituents preclude binding and/or demethylation. The partial demethylation observed for syringol and vanillin, despite the fact that both assume a guaiacol-like binding mode in the crystal structures (Fig. 2), suggests that dynamic factors may be important for understanding the efficiency of the GcoA-substrate interaction.

Enzyme opening and closing are key steps in GcoA catalysis. Some P450 enzymes are known to undergo conformational opening and closing motions during their catalytic cycles$^{37}$. While the structures of GcoA obtained in this study were crystallized in an apparent closed state, molecular dynamics (MD) simulations indicated that an in silico-generated apo form of GcoA can spontaneously open on the time scale of 1 μs, with the structural changes occurring mainly in the F and G helices (Fig. 5a, b, Supplementary Fig. 21 and Supplementary Movies 1-2). The overall conformation of the MD-generated GcoA open structure closely resembles that of the well-characterized open form of the P450 from Bacillus megaterium BM3, (PDB ID: 2HPD)$^{38}$ (Supplementary Fig. 22).

We subsequently employed umbrella sampling (US) to obtain the free energy profile for the opening-closing motion of the
enzyme in the apo, guaiacol (substrate), and catechol (product) forms. The free energy was computed along a reaction coordinate defined as the difference between the root mean square deviation (RMSD) from the open structure and the RMSD from the closed structure. The free energy profile of GcoA:apo is relatively flat and exhibits two minima, one associated with the open state and another with the closed state (Fig. 5c). The free energy of the closed state is, within error, equivalent to that of the open state, with a free energy barrier <4.5 kcal/mol. This indicates that GcoA can easily transition between the open and closed states when a substrate is absent, which may help to explain the difficulty in obtaining a crystal structure of the apo-GcoA.

In contrast, the free energy profile associated with the opening-closing motion of either GcoA:guaiacol or GcoA:catechol exhibits only one minimum, associated with the closed state, suggesting that the enzyme will close when the substrate or product is bound. We note, however, that the free energy cost for GcoA:guaiacol or GcoA:catechol is significantly higher (5.1 kcal/mol) and associated with a transition state that is 2-3 Å RMSD from the crystallographic structures and corresponds to a RMSD >3 Å, as their RMSD never reaches values >3 Å, in contrast to the RMSD in GcoA:apo, which can reach values greater than 4 Å. In the configuration shown in Fig. 5e, which occurs in the crystal structures and corresponds to an RMSD <2 Å, the three Phe residues are close to each other and interact directly with guaiacol. This configuration tends to occur when GcoA is closed in the simulations (Fig. 5d, Supplementary Fig. 21). In the configuration in Fig. 5f, which corresponds to the 2-3 Å RMSD range and occurs when GcoA is partially open (with the reaction coordinate around 0), the side chain of Phe169 deviates from its crystallographic position, but the three Phe residues still interact with the substrate and exclude water from the active site. In the configuration in Fig. 5g, which shows the most open state of GcoA:apo (when the reaction coordinate assumes values less than 0 and at an RMSD value >3 Å), the Phe side chains move even further apart from each other, expanding the binding site and allowing the substrate to enter the active site without significant changes in the active site geometry.
allowing water to enter the enzyme. Overall, the presence of guaiacol or catechol in the binding site keeps the three Phe residues arranged around the ligand and prevents the full closed-to-open transition of the active site.

Simulations of GcoA:guaiethol, GcoA:syringol, and GcoA:vanilline indicate different effects of these substrates on the conformational flexibility of GcoA (Supplementary Fig. 21) and help explain the structure-activity relationships in this enzyme. GcoA:guaiethol exhibits less flexibility than GcoA:guaiacol, as observed by the limited opening-closing motions and binding site expansion. Given an almost fivefold reduction in $k_{\text{cat}}$ for guaiethol compared to guaiacol, it appears that such flexibility is required for optimal substrate turnover. GcoA:syringol and GcoA:vanilin, on the other hand, are more flexible than GcoA:guaiacol and more prone to opening-closing transitions and expansion of the active site. This indicates that syringol and vanilin, which bind to the active site, stimulate NADH turnover, but are not stoichiometrically demethylated (Table 1), are less effective in maintaining the enzyme in the closed state than guaiacol and guaiethol. This suggests that successful protein engineering for alternative substrates will require careful consideration to balance conformational flexibility with productive binding and catalysis, and these data provide a route to help define the optimum window.

**Proposed reaction mechanism for guaiacol O-demethylation.** Having identified that GcoA contributes the rate limiting step in this 2-protein system, density functional theory (DFT) calculations were used to investigate the mechanism for guaiacol O-demethylation. The putative enzymatic reaction is shown (Fig. 6). The DFT calculations using a truncated model system identified two possible reaction pathways (path A and path B) that GcoA could catalyze, which rely to two different approaches of the guaiacol substrate to the Fe–O active species (Fig. 6) (see Supplementary Table 5 and Supplementary Data 1 for DFT energies and optimized geometries).

Path A leads to the formation of hemiacetal (3) through a hydrogen atom transfer (HAT) rate-limiting step, followed by a rapid OH rebound. Hemiacetal 3 can then degrade in solution to form the O-demethylated catechol product and formaldehyde. Conversely, path B would form a stable acetal (5) in two sequential HAT reactions: a rate-limiting C-H abstraction followed by a subsequent O-H abstraction to generate a biradical intermediate (4) that cyclizes in a barrierless process to form (5) (Fig. 6).

The main difference between the two rate-limiting transition states (TSs) is the conformation that the guaiacol substrate adopts with respect to the Fe–O active species. TS1-a$^1$ corresponds to the HAT transition state when the substrate/Fe = O orientation is similar to the one observed in the substrate-bound crystal structure (see Figs. 2c and 6b); TS1-b$^2$ is the lower energy HAT TS (1.5 kcal/mol lower than TS1-a$^1$). In this case, the substrate orientation allows the guaiacol hydroxyl group to interact by H-bond with the Fe = O, stabilizing the TS but also permitting the second OH H-abstraction (Fig. 6). The direct comparison of the two rate-limiting TSs proved the intrinsic preference of the substrate to react following path B over path A. Nevertheless, the strong preference of the substrate to bind in the specific orientation found in the crystal structure as observed during the course of the MD simulations (Fig. 5f and Supplementary Fig. 27), indicates that path A will be followed although it is energetically less favorable.

Alternatively, open-shell singlet biradical intermediate (4) in path B could form the less stable zwiterionic closed-shell electronic configuration (1.6 kcal/mol higher in energy than the biradical) and further react with a water molecule to generate the hemiacetal (3). The absence of water molecules in the active site environment when guaiacol is bound as observed from our MD simulations (Supplementary Fig. 28) argues against this possibility.

**Discussion**

Recent efforts from multiple groups have attempted to harness aromatic catabolism for productivity utilizing lignin.$^{8-10,18,39-44}$ As a single microbe is unlikely to have the full complement of necessary catabolic enzymes for lignin bioconversion, a key component of such synthetic biology strategies is the introduction of foreign catabolic genes to expand substrate specificities of the host microbe. Bacterial enzymes that catalyze the demethylation of lignin-derived aryl-methoxy substrates are of particular interest, as the demethylation reaction presents a bottleneck for the conversion of lignin into desirable products. Currently, Rieske non-heme iron monooxygenases$^{14,15,18}$ and tetrahydrofolate-dependent O-demethylases,$^{16,17}$ offer two well-known paradigms for aryl-O-demethylation. This study presents a detailed characterization of a third, cytochrome P450-based enzymatic strategy that could fill a critical gap for engineering applications.
From a metabolic engineering standpoint, the GcoAB system offers a number of potential advantages. First, the native substrate of GcoA, guaiacol, is a major breakdown product of plant lignin. Demethylation of guaiacol yields catechol, which can be ring-opened via either intra- or extra-diol cleavage catechol dioxygenases. Second, compared to other known O-aryl-demethyleases, the substrate preferences of GcoA are intrinsically broad, admitting a variety of guaiacol analogs that are also known lignin breakdown products. Third, we anticipate a P450 system to be amenable to further tuning using directed evolution techniques. A prior report of a closely related cytochrome P450 that can demethylate 4-methoxybenzoate suggests that the GcoA active site may be modified to admit larger, more hydrophilic, lignin-derived substrates such as ferulate or vanilllate. Indeed, genes encoding putative homologs to the two-component GcoA and GcoB system described here are predicted in the genomes of several bacterial species belonging to the genera Rhodococcus, Streptomyces, and Gordonia, among others, and the substrate preferences for this diverse group remain unclear, but offer a promising platform for further exploration and engineering. Moreover, work from Bell et al. revealed an unpaired Rhodopseudomonad cytochrome P450 can also demethylate 4-methoxybenzoate and be productively engineered to accommodate 4-ethylbenzoate. While retaining a classical P450 fold, this CYP199A4 system exhibits an alternative binding mode in terms of both substrate positioning relative to the heme, and steric selectivity with an alternative set of aromatic residues lining the active site pocket, further demonstrating the diversity within this class of enzymes. Fourth, a heme-based P450 may offer a simpler alternative for aromatic demethylation compared to tetrahydrofolate-dependent O-demethylases, given the relative ubiquity of P450s and robust heme biosynthetic pathways in potential bacterial hosts. Finally, distinct from most P450 systems, the GcoB reductase is encoded as a single polypeptide rather than two.

Close examination of the GcoA-guaiacol active site shows that substrate binding involves interactions between the peptide backbone and the substrate hydroxyl (ring C1) and methoxy groups (C2). The ring C3 position has a relatively close (3.8 Å) backbone and the substrate hydroxyl (ring C1) and methoxy two. From a metabolic engineering standpoint, the GcoAB system offers a number of potential advantages. First, the native substrate of GcoA, guaiacol, is a major breakdown product of plant lignin. Demethylation of guaiacol yields catechol, which can be ring-opened via either intra- or extra-diol cleavage catechol dioxygenases. Second, compared to other known O-aryl-demethyleases, the substrate preferences of GcoA are intrinsically broad, admitting a variety of guaiacol analogs that are also known lignin breakdown products. Third, we anticipate a P450 system to be amenable to further tuning using directed evolution techniques. A prior report of a closely related cytochrome P450 that can demethylate 4-methoxybenzoate suggests that the GcoA active site may be modified to admit larger, more hydrophilic, lignin-derived substrates such as ferulate or vanilllate. Indeed, genes encoding putative homologs to the two-component GcoA and GcoB system described here are predicted in the genomes of several bacterial species belonging to the genera Rhodococcus, Streptomyces, and Gordonia, among others, and the substrate preferences for this diverse group remain unclear, but offer a promising platform for further exploration and engineering. Moreover, work from Bell et al. revealed an unpaired Rhodopseudomonad cytochrome P450 can also demethylate 4-methoxybenzoate and be productively engineered to accommodate 4-ethylbenzoate. While retaining a classical P450 fold, this CYP199A4 system exhibits an alternative binding mode in terms of both substrate positioning relative to the heme, and steric selectivity with an alternative set of aromatic residues lining the active site pocket, further demonstrating the diversity within this class of enzymes. Fourth, a heme-based P450 may offer a simpler alternative for aromatic demethylation compared to tetrahydrofolate-dependent O-demethylases, given the relative ubiquity of P450s and robust heme biosynthetic pathways in potential bacterial hosts. Finally, distinct from most P450 systems, the GcoB reductase is encoded as a single polypeptide rather than two.

Close examination of the GcoA-guaiacol active site shows that substrate binding involves interactions between the peptide backbone and the substrate hydroxyl (ring C1) and methoxy groups (C2). The ring C3 position has a relatively close (3.8 Å) interaction with the porphyrin-y-meso carbon that bridges the propionate-substituted pyrrole rings. However, the remaining ring positions are not directly occluded by backbone/porphyrin atoms. Reactivity studies showed that guaiacol analogs with substitutions at C4, C5, and C6 remained substrates with comparable efficiencies (kcat/Km) to guaiacol itself. Even substitutions at C1 and C2 (ethoxy for methoxy-) are permitted. Comparison of the structures of the guaiacol, guaethol, vanilllan, and syringol ligand-bonded GcoA structures shows that all assume a similar orientation of the guaiacol substrate. The spectra of GcoA and GcoB were measured in 25 mM HEPES, 50 mM NaCl at pH 7.5 with a Lambda 25 spectrofluorometer (Perkin Elmer) over a wavelength range of 200–600 nm at 1 nm intervals in a quartz cuvette (Hellma Analytics).

**Heme quantification.** Catalytically active heme bound to GcoA was determined via a spectrophotometric/CO-binding assay. CO gas was bubbled into a cuvette containing 0.94–2.5 μM GcoA (Pierce BCA assay). Excess sodium dithionite (~1 mg) was added to reduce the heme iron. The spectrum was recorded over a period of several minutes as a peak at ~450 nm gradually appeared, attributed to the catalytically competent, ferrous CO-bound heme. A spectrum for a control containing only dithionite-reduced GcoA was measured, and a difference spectrum computed. Absorptions at 420, 450, and 490 nm were recorded to calculate the amount of active GcoA (P450) or inactive GcoA (P420) nm. The equations used to compute the concentrations of catalytically competent and inactive heme are shown below. Reported values are the average of three or more measurements.

Here ΔA450 and ΔA420 are the differences between the reference and sample spectrum at absorbances 450 and 420 nm, respectively.

The pyridine hemochrome assay was additionally used to assess the total heme content in GcoA. Volume of 200 μL of a 6 μM GcoA solution was added to 797 μL of 50 mM NaOH, 20% pyridine and 3 μL K3Fe(CN)6 and a UV/vis spectrum measured. Absorptions (2–5 mg) sodium dithionite was used to reduce the heme and the absorbance at 556 nm was compared to the oxidized spectrum (ΔΔA). The Beer-Lambert law was used to calculate the amount of heme present, using ε556 = 28.4
mM–1 cm–1. Reported values are averaged from three or more measurements. An extinction coefficient for GcoA-bound heme, using the [GcoA-heme] determined from the CO binding assay, was estimated via the slope of a line relating absorbance at 417 nm (Soret peak) to [GcoA-heme].

**Determination of [FAD] and non-heme [Fe] in GcoB.** FAD was released from GcoB by denaturing 200 µM of a protein (0.024 µM) solution with 5 µl saturated ammonium sulfate (7% v/v H2SO4), similar to studies with related cytochrome P450s.47 Precipitated protein was centrifuged and the UV/vis spectrum of the FAD-containing supernatant was measured. The absorbance at 454 nm, εFAD = 11.3 mM–1 cm–1, and total protein concentration determined by the Bradford assay were used to determine [FAD] bound to GcoB. An extinction coefficient for GcoB-bound FAD was estimated via the slope of a line relating absorbance at 454 nm to [GcoB:FAD].

The Fe–S content of GcoB was assessed both by quantifying non-heme Fe(II) and spectrophotometrically characterizing the cluster as a whole (below). GcoB was denatured as described above. Volume of 50 µl of supernatant was added to 25 µl of 5% w/v sodium ascorbate to reduce the iron. Volume of 100 µl of bathophenanthrolin disulfonate (0.1% w/v in dH2O) was added and the sample was incubated for 1 h. The resulting Fe(II) complex was quantified via its absorbance at 535 nm (εFeIII = 22.14 mM–1 cm–1, determined using FeSO4 standards). An extinction coefficient for GcoB-bound 2Fe-2S cluster was estimated via the slope of a line relating absorbance at 423 nm to [GcoB:2Fe-2S].

**EPR and UV/vis spectroscopic characterization of FeS clusters.** A 150 µM sample of GcoB was brought into a MBraun chamber and exchanged into anaerobic 50 mM Tris, 200 mM NaCl, 5% glycerol, pH 7.0. The sample was then reduced anaerobically with 10 mM sodium dithionite and loaded into an EPR tube. The tube was capped prior to removing from the chamber and frozen in liquid N2. X-band EPR spectra were recorded on a Bruker 500 spectrometer equipped with a Super High Q (SHQ) resonator, in-cavity cryogen-free system (ColdEdge Technologies), and Mercury/TC temperature controller (Oxford). Spin quantifications were determined by comparison to copper standards at 75, 100, and 125 µM via double integration of the spectra. The effect of substrate addition on the ratios of Fe–S clusters was determined by comparison to a control sample.

**Reductase activity measurement of GcoB.** GcoB activity was measured using a continuous colorimetric assay involving cytochrome c as a colorimetric electron acceptor.45 A total of 4.8 mM GcoB (referred to [FAD]) and 42 µM cytochrome c (from equine heart) were dissolved in buffer (25 mM HEPES, 50 mM NaCl, pH 7.5), 250 µM FAD was then added to initiate the NADH- and GcoB-dependent reduction of cytochrome c. UV/vis spectra (Varian Cary 4000, Agilent) were recorded in the scanning kinetics mode. The increase in absorbance at 550 nm due to reduced cytochrome c was monitored over time, and the specific activity (nmol reduced cytochrome c min–1 nmol GcoB–1) was calculated using:

$$\Delta A_{550} = \Delta A_{550} \text{Abs}_{\text{max}} \cdot 0.021/\text{ml reaction} = \text{specific activity}.$$  (4)

For determining steady-state kinetic constants, the above protocol was used as a function of [NADH]. A total of 4.6 mM GcoB (referred to [FAD]) and 43 µM cytochrome c were dissolved in buffer (25 mM HEPES, 50 mM NaCl, pH 7.5), 25 °C, and the reaction was initiated with the addition of 2.5–200 µM NADH. Initial velocities (vi) were determined from linear fits to the initial portion of the progress curve, plotted as a function of [NADH], and fit to the Michaelis-Menten Eq. (5) using the KaleidaGraph software:

$$v_i = \frac{v_{\text{max}} [S]}{K_{\text{M}} + [S]}.$$

**Steady-state kinetics analysis of GcoAB.** The demethylation of guaiacol and substrate analogues was continuously monitored under steady-state conditions. A total of 0.2 µM each of GcoA and GcoB were dissolved in air-saturated buffer (25 mM HEPES, 50 mM NaCl in a cuvette at pH 7.5, 25 °C). A total of 100 µg/ml catalese was added to each reaction to capture any H2O2 formed during the uncoupled reaction. A saturating amount of NADH (≥250 µM) was added and a background rate of NADH oxidation in air (200 µM O2) recorded via continuous scanning of the UV/vis spectrum. A total of 20–300 µM guaiacol (preferred substrate) or an alternate substrate from a 2.5 mM stock dissolved in DMSO was added and the reaction was monitored via measurement of UV/vis spectra for several minutes. The initial velocity was determined by the characteristics of the absorbance change as a function of time. A plot of vi vs [guaiacol] was fit to Eq. (2) to obtain kcat, KmGcoA, and KmGcoB. For specific activity determination, the above method was used but with saturating (300 µM) guaiacol. The linear portion of [NADH] vs time was fit and referenced to the amount of GcoA used (0.2 µM). Reported values are the average of ≥3 measurements and reported errors are standard deviations.

For vanillic, whose UV/vis spectrum overlaps with that of NADH, fluorescence was used to monitor NADH disappearance using a FluoroMax3 instrument (Horiba, Jobin Yvon). A standard curve for NADH (0–350 µM) was generated by exciting the sample (25 mM HEPES, 50 mM NaCl pH 7.5) at 340 nm and monitoring the emission at 548 nm (vanillin did not excite or emit at this wavelength). The intensity vs [NADH] was plotted and fit to a 4-parameter logistic equation:

$$\text{Intensity} = \frac{a - d}{1 + (\frac{c}{x})^n} + d.$$  (6)

where a is the theoretical response at [NADH] = 0, b is the slope of the curve at the inflection point, c is [NADH] at the inflection point, and d is the theoretical response at infinite [NADH]. Reactions with vanillin were performed in the same manner described above.

**Determination of substrate dissociation constants with GcoA.** A total of 0–60 µM of substrate analogues, in 0.25 or 0.5 µM aliquots, were titrated into a cuvette containing 1–6 mM GcoA in 25 mM HEPES, 50 mM NaCl, pH 7.5. The spectrum after each substrate addition was recorded, beginning with no substrate bound. The solution reached equilibrium before the next addition. A difference spectrum was made to illustrate the shift from a low-spin apo-heme complex to the high-spin substrate-bound complex (spectral shift from 417 nm to 388 nm). The resulting difference spectra showed a peak at 388 nm, and a trough at 420 nm. The absorbance at 388 nm was plotted as a function of substrate, yielding a quadratic curve that was fit to Eq. (7) to determine the Ki.

$$\Delta A_{\text{obs}} = \Delta A_{\text{abs}} \frac{E_0 + K_F - \sqrt{E_0 + K_F + K_D}}{E_0 + K_F} - A_0$$  (7)

where E0, Ef, and Eabs are the ligand concentrations, total protein (substrate) concentration, the equilibration dissociation constant, and the maximum absorbance, respectively. Reported values are the average of 2 or more measurements.

Since the UV/vis spectra of vanillic acid and acetaldehyde overlap, fluorescence quenching was monitored to determine the Ki. A 4 µM GcoA (25 mM HEPES, 50 mM NaCl, pH 7.5) was excited at 283 nm and the absorbance read at 548 nm. The reaction was monitored to determine the steady-state rate of fluorescence quenching.

$$\%F_{\text{quenched}} = \frac{F_{\text{max}} - F_{\text{steady}}}{F_{\text{max}}} \times 1000.$$  (8)

**Aldehyde product determination.** For the quantification of formaldehyde (demethylation) or acetaldehyde (de-ethylation) production, the reaction was monitored via UV/vis or fluorescence (in the case of vanillic acid) using the same set of conditions as outlined above for specific activity determination. After eight minutes, aliquots of the sample were removed and carried onto the respective aldehyde detection assay. For reactions that produced formaldehyde (e.g., all substrates except guaiacol), a colorimetric tetrophenyl assay was used, described previously.50 Briefly, 200 µl of the reaction was quenched by adding 200 µl of 0.1% triphenyl solution in 50% ethanol, 200 µl 90% sulfuric acid and 40 µl of 1% FeCl3. The solution was then incubated in a heating block for 90 min at 70 °C. After cooling, the absorbance was read at 575 nm and the [formaldehyde] calculated by comparing to a standard curve made with 0–320 µM formaldehyde. A reaction with just the assay components was treated in the same manner as the reaction samples and used as a baseline. Control reactions included everything but GcoA/B, catalase, and substrate. Reported values are the average of ≥3 measurements.

[Acetaldehyde] produced during the dealkylation of guaiacol was also determined by using a colorimetric assay and a generated acetaldehyde standard curve. A kit from BioAssays was used. Briefly, 20 µl of the reacted sample was transferred to a 96-well plate and 80 µl of the working reagent, consisting of NAD/MTT and aldehyde dehydrogenase, was added. The reaction was incubated for 30 min at room temperature and the absorbance read at 565 nm. Aldehyde dehydrogenase and NAD react with acetaldehyde to produce acetic acid and NADH. The NADH can then reduce MTT, resulting in the absorbance at 565 nm. Control reactions included samples without substrate and/or aldehyde dehydrogenase. Reported values are the average of ≥3 measurements.

**HPLC product identification.** Analyte analysis of samples was performed on an Agilent 1200 LC system (Agilent Technologies, Santa Clara, CA) equipped with a G1315A quadrupole array detector (DAD). Each sample and standard was injected at a volume of 10 µl onto a Phenomenex Luna C18(2) column 5 µm, 4.6 × 150 mm column (Phenomenex, Torrance, CA). The column temperature was maintained at
at time $t = 5$, (A) = 50% and (B) = 50% at $t = 35$ min; (A) = 1% and (B) = 99% at $t = 35.01$ min; (A) = 99% and (B) = 1% at $t = 37.01$ min; (A) = 99% and (B) = 1% at $t = 47.00$ min. The flow rate was held constant at 0.6 ml min$^{-1}$ resulting in a run time of 47 min. Calibration curve concentration for each analyte varied between the ranges of 2.5–300 ηg L$^{-1}$. In state and conformational freedom from 1.1 to 1.8. For the protein, the f3L5B Amber force field$^{62}$ was employed along with the TIP3P water model$^{63}$. The simulations were performed using periodic boundary conditions, with short-range interactions truncated at a cutoff radius of 8.0 Å and a reaction field (PMF) for long-range interactions$^{64}$. The equations of motion were integrated with a time-step of 2.0 fs, with bonds involving hydrogen atoms constrained at their equilibrium values using SHAKE. The temperature was kept constant at 300 K using the Langevin thermostat with a collision frequency of 1.0 ps$^{-1}$. The pressure was controlled at 1.0 bar only during the initial equilibration steps (described below) with the Berendsen barostat using a relaxation time of 2.0 ps.

**Dynamics**

**Dynamics version 5.26.60** software package (Protein Solutions). Samples were measured at 37.5 °C and the volumes of the protein and water were estimated using the empirical equation for a globular protein: $V = M \times R_\text{g}^2 \times 1.5$, where $M$ is the molecular mass of the protein in kilodaltons and $R_\text{g}$ is the hydrodynamic radius of the protein in nm.

**Analytical ultracentrifugation**

Velocity analytical ultracentrifugation was performed using a Beckman XL-A analytical ultracentrifuge with an An50-Ti rotor. Double-cylinder Epon cells were used with 1.2 cm path lengths and quartz window assemblies. Protein concentration was measured at 37.5 °C and the volume of the protein and water were estimated using the empirical equation for a globular protein: $V = M \times R_\text{g}^2 \times 1.5$, where $M$ is the molecular mass of the protein in kilodaltons and $R_\text{g}$ is the hydrodynamic radius of the protein in nm.

**MD simulation and DFT systems setup**

Molecular dynamics simulations were performed with G90$^{60}$. A truncated model containing the porphyrin pyrrole core, Fe center and a methanethiol to mimic cysteine as Fe-axial ligand was used. Geometry optimizations and frequency calculations were performed using unrestricted B3LYP (UB3LYP)$^{68}$ with the LANL2DZ basis set for iron and 6-31G(d) on all other atoms. Transition states with one small negative force constant corresponding to the desired transformation. Enthalpies and entropies were calculated for 1 atm and 298.15 K. A correction to the harmonic oscillator approximation, as discussed by Truhlar and co-workers, was also applied to the entropy calculations by raising all frequencies below 100 cm$^{-1}$ to $k = 10$ kcal mol$^{-1}$ A$^{-2}$. Simulation in an additional window centered at 2.2 Å with $k = 10$ kcal mol$^{-1}$ A$^{-2}$ was conducted to assure enough overlap between neighboring windows. Therefore, a total of 29 windows were used for the PMF calculation. Within each window, 260 ns of restrained MD simulation was carried out after a 10-ns equilibration (not considered in the PMF calculation), totaling 270 ns window. Initial configurations for the different windows of the GcoA:apo US simulations were taken from the unbiased simulation where we observed the closing/open transition. For the GcoA:guaiacol at the GcoA:acetaldehyde systems, we started from the closed structure and followed a scheme where we used the equilibrated configuration of the previous window (window $i$) to start the simulation of the next window (window $i+1$). The PMFs were obtained as the average of PMFs calculated for blocks of 10 ns. The Weighted Gramian Analytical Method$^{79}$ was employed to reweight the biased histograms obtained with US MD.

**Density functional theory calculations**

DFT calculations were performed using Gaussian 09.$^{80}$. A truncated model containing the porphyrin pyrrole core, Fe center and a methanethiol to mimic cysteine as Fe-axial ligand was used. Geometry optimizations and frequency calculations were performed using unrestricted B3LYP (UB3LYP)$^{68}$ with the LANL2DZ basis set for iron and 6-31G(d) on all other atoms. Transition states with one negative force constant corresponding to the desired transformation. Enthalpies and entropies were calculated for 1 atm and 298.15 K. A correction to the harmonic oscillator approximation, as discussed by Truhlar and co-workers, was also applied to the entropy calculations by raising all frequencies below 100 cm$^{-1}$ to $k = 10$ kcal mol$^{-1}$ A$^{-2}$. Simulation in an additional window centered at 2.2 Å with $k = 10$ kcal mol$^{-1}$ A$^{-2}$ was conducted to assure enough overlap between neighboring windows. Therefore, a total of 29 windows were used for the PMF calculation. Within each window, 260 ns of restrained MD simulation was carried out after a 10-ns equilibration (not considered in the PMF calculation), totaling 270 ns window. Initial configurations for the different windows of the GcoA:apo US simulations were taken from the unbiased simulation where we observed the closing/open transition. For the GcoA:guaiacol at the GcoA:acetaldehyde systems, we started from the closed structure and followed a scheme where we used the equilibrated configuration of the previous window (window $i$) to start the simulation of the next window (window $i+1$). The PMFs were obtained as the average of PMFs calculated for blocks of 10 ns. The Weighted Gramian Analytical Method$^{79}$ was employed to reweight the biased histograms obtained with US MD.

**Data availability**

Coordinates and associated structure factors have been deposited with the PDB (www.pdb.org) under accession codes 5NCB, 5OMR, 5OMS, 5OMZ.
NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-04878-2 | www.nature.com/naturecommunications

5. Fuchs, G., Boll, M. & Heider, J. Microbial degradation of aromatic compounds
7. Masai, E., Katayama, Y. & Fukuda, M. Genetic and biochemical investigations
6. Vaillancourt, F. H., Bolin, J. T. & Eltis, L. D. The ins and outs of ring-cleaving
10. Bugg, T. D. H. & Rahmanpour, R. Enzymatic conversion of lignin into
13. Eltis, L. D., Karlson, U. & Timmis, K. N. Puri
14. Segura, A., Bünz, P. V., D
15. Morawski, B., Segura, A. & Ornston, L. N. Substrate range and genetic analysis of
11. Ragauskas, A. J. et al. Lignin valorization: improving lignin processing in the
22. Bell, S. G. et al. The crystal structures of 4-methoxybenzoate bound
23. Bell, S. G. et al. Investigation of the substrate range of CYP199A4: modi
25. Hannemann, F., Bichet, A., Even, K. M. & Bernhardt, R. Cytochrome P450 systems—biological variations of electron transport chains. Biochim. Biophys. Acta 1770, 330–344 (2007).
26. Wang, M. et al. Three-dimensional structure of NADPH–cytochrome P450 reductase: prototype for FMN- and FAD-containing enzymes. Proc. Natl Acad. Sci. USA 94, 8411–8416 (1997).
27. Karlsson, A. et al. X-ray crystal structure of benzoate 1, 2-dioxygenase from Acinetobacter sp. strain ADP1. J. Mol. Biol. 318, 261–272 (2002).
28. Svrčíková, I. F., Li, H., Zhang, H., Peterson, J. A. & Poulos, T. L. Structure of a cytochrome P450-redox partner electron-transfer complex. Proc. Natl Acad. Sci. USA 96, 1863–1868 (1999).
29. Tripathi, S., Li, H. & Poulos, T. L. Structural basis for effector control and redox partner recognition in cytochrome P450. Science 340, 1227–1230 (2013).
30. Hasemann, C. A., Kurumbl, R. G., Boddupalli, S. S., Peterson, J. A. & Deisenhofer, J. Structure and function of cytochromes P450: a comparative analysis of three crystal structures. Structure 3, 41–62 (1995).
31. Huang, W. C., Ellis, J. I. & Roberts, G. C. Redox-linked domain movements in the catalytic cycle of cytochrome p450 reductase. Structure 21, 1581–1589 (2013).
32. Kelly, S. L. & Kelly, D. E. Microbial cytochromes P450: biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us? Philos. Trans. R. Soc. Lond. B Biol. Sci. 368, 20120476 (2013).
33. Guengerich, F. P. & Munro, A. A. Unusual cytochrome P450 enzymes and reactions. J. Biol. Chem. 288, 17065–17073 (2013).
34. Eby, D. M., Beharry, Z. M., Coutler, E. D., Kurtz, D. M. Jr & Neidle, E. L. Characterization and evolution of anthranilate 1, 2-dioxygenase from E. coli. J. Bacteriol. 183, 109–118 (2001).
35. Guengerich, F. P., Martin, M. V., Sohl, C. D. & Cheng, Q. Measurement of cytochrome P450 and NADPH-cytochrome P450 reductase. Nat. Protoc. 4, 1245–1251 (2009).
36. Rupp, H., Rao, K. K., Hall, D. O. & Cammack, R. Electron spin relaxation of iron–sulphur proteins studied by microwave power saturation. Biochim. Biophys. Acta 537, 255–260 (1979).
37. Dubey, K. D., Wang, B. J. Shah, S. Molecular dynamics and QM/MM calculations predict the substrate-induced gating of cytochrome P450 BM3 and the regio- and stereoselectivity of fatty acid hydroxylase. J. Am. Chem. Soc. 138, 837–845 (2016).
38. Ravichandran, K., Boddupalli, S., Haasermann, C., Peterson, J. & Deisenhofer, J. Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450s. Science 261, 731–736 (1993).
39. Kosa, M. & Ragauskas, A. J. lignin to lipid biocconversion by oleaginous Rhodococcus. Green Chem. 15, 2070–2074 (2013).
40. Sainsbury, P. D. et al. Breaking down lignin to high-value chemicals: the conversion of lignocellulose to vanillin in a gene deletion mutant of Rhodococcus jostii RHA1. ACS Chem. Biol. 8, 2151–2156 (2013).
41. Vardon, D. R. et al. Adipic acid production from lignin. Energy Environ. Sci. 8, 617–628 (2015).
42. Microty, Z., Gomis, M., Mines, P., Law, P. & Bugg, T. D. Biocatalytic conversion of lignin to aromatic dicarboxylic acids in Rhodococcus jostii RHA1 by re
43. Jung, S. F., Lautchi, R. & Arnold, F. H. Cytochrome P450: taming a wild type enzyme. Curr. Opin. Biotechnol. 22, 809–817 (2011).
44. Berry, E. A. & Trumpower, B. L. Simultaneous determination of hemes a, b, and c from pyridine hemochrom spectra. Anal. Biochem. 161, 1–15 (1987).
45. Swoboda, B. E. P. The relationship between molecular conformation and the binding of flavin–adenine dinucleotide in glucose oxidase. Biochim. Biophys. Acta 175, 365–379 (1969).
46. Bell, S. G., Tan, A. B., Johnson, E. O. & Wong, L. L. Selective oxidative demethylation of veratic acid to vanillic acid by CYP199A4 from Rhodopseudomonas palustris HA2. Mol. Biol. Ecol. 20, 224–225 (2014).
47. Noll, S. & Schweiger, A. EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. J. Magn. Reson. 178, 42–55 (2006).
48. Črástil, J. A. W. & John, T. A sensitive colorimetric method for formaldehyde. Anal. Biochem. 69, 202–207 (1975).
49. Winn, M. D. et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242 (2011).
50. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221 (2010).
51. Latz, T. M., Shah, B. D., E3205, K. M. & Pelletier, S. L. Analytical Ultraconcentration in Biochemistry and Polymer Science (The Royal Society of Chemistry: Cambridge, UK, 1992).
