Chemistry–A European Journal
Research Article
doi.org/10.1002/chem.202201895
Chemistry Europe
European Chemical Societies Publishing
www.chemeurj.org

Enabling Aromatic Hydroxylation in a Cytochrome P450 Monooxygenase Enzyme through Protein Engineering

Tom Coleman,[a] Joel Z. H. Lee,[a] Alicia M. Kirk,[b] Daniel Z. Doherty,[a] Matthew N. Podgorski,[a] Dilshi K. Pinidiya,[a] John B. Bruning,[c] James J. De Voss,[b] Elizabeth H. Krenske,*[b] and Stephen G. Bell*[a]

Abstract: The cytochrome P450 (CYP) family of heme monooxygenases catalyse the selective oxidation of C–H bonds under ambient conditions. The CYP199A4 enzyme from *Rhodopseudomonas palustris* catalyses aliphatic oxidation of 4-cyclohexylbenzoic acid but not the aromatic oxidation of 4-phenylbenzoic acid, due to the distinct mechanisms of aliphatic and aromatic oxidation. The aromatic substrates 4-benzyl-, 4-phenoxy- and 4-benzoyl-benzoic acid and methoxy-substituted phenylbenzoic acids were assessed to see if they could achieve an orientation more amenable to aromatic oxidation. CYP199A4 could catalyse the efficient benzylic oxidation of 4-benzylbenzoic acid. The methoxy-substituted phenylbenzoic acids were oxidatively demethylated with low activity. However, no aromatic oxidation was observed with any of these substrates. Crystal structures of CYP199A4 with 4-(3′-methoxyphenyl)benzoic acid demonstrated that the substrate binding mode was like that of 4-phenylbenzoic acid. 4-Phenoxy- and 4-benzoyl-benzoic acid bound with the ether or ketone oxygen atom hydrogen-bonded to the heme aqua ligand. We also investigated whether the substitution of phenylalanine residues in the active site could permit aromatic hydroxylation. Mutagenesis of the F298 residue to a valine did not significantly alter the substrate binding position or enable the aromatic oxidation of 4-phenylbenzoic acid; however the F182L mutant was able to catalyse 4-phenylbenzoic acid oxidation generating 2′-hydroxy-, 3′-hydroxy- and 4′-hydroxy metabolites in a 83:9:8 ratio, respectively. Molecular dynamics simulations, in which the distance and angle of attack were considered, demonstrated that in the F182L variant, in contrast to the wild-type enzyme, the phenyl ring of 4-phenylbenzoic acid attained a productive geometry for aromatic oxidation to occur.

Introduction

The oxidation of aliphatic and aromatic C–H bonds are extremely challenging chemical reactions.[1] Remarkably, these reactions are achieved under ambient conditions by enzymes including copper monooxygenases, non-heme iron systems and the heme-dependent cytochrome P450 (CYP) monooxygenases.[2] In the cytochrome P450 enzymes a highly reactive ferryl compound I (Cpd I) species is the reactive intermediate (Scheme 1).[2a,3] The enzyme inserts an oxygen atom into an aliphatic C–H bond through a rebound mechanism first proposed by Groves and McClusky.[4] In this pathway the Cpd I intermediate abstracts a hydrogen atom from the substrate resulting in a substrate radical and a Fe^{IV}-OH species (compound II, Cpd II).[3c,4–5] The substrate radical remains closely associated with Cpd II where it recombines with the hydroxyl radical to yield the alcohol product.[5] It has been proposed that when the substrate C–H bond is optimally positioned relative to the heme and Cpd I these two steps can occur in a dynamically coupled process and can effectively bypasses the formation of a discrete radical species.[6] The strong C–H bonds of aromatic molecules are not able to be abstracted by Cpd I. Aromatic oxidation reaction are hypothesised to occur by electrophilic attack of Cpd I on the π system of the aromatic species (*ipso* substitution) or arene oxide (or oxepin) formation (the NIH mechanism; Scheme 1).[7]

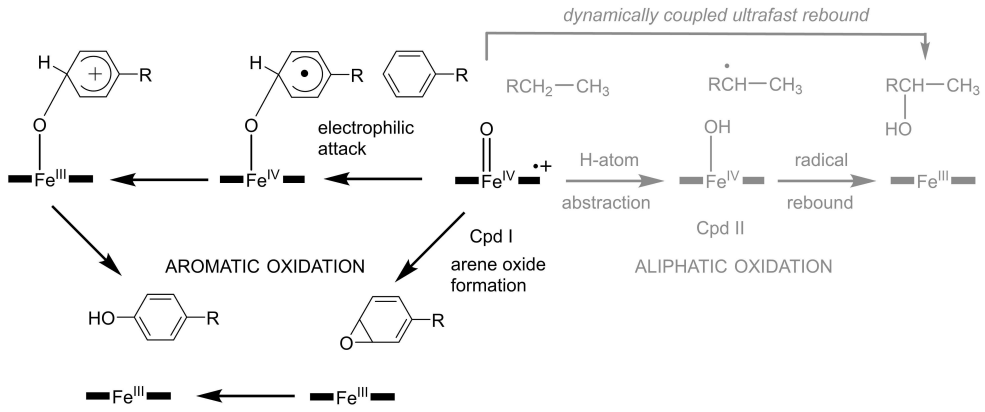
The cytochrome P450 enzyme CYP199A4 from *Rhodopseudomonas palustris* strain HaA2 can oxidatively demethylate 4-methoxybenzoic acid to yield 4-hydroxybenzoic acid. Subsequent studies have demonstrated that this enzyme has a high affinity for *para*-substituted benzoic acids and that the oxidation reaction occurs at the *para* substituent.[7–8] For example, CYP199A4 is able to selectively demethylate 4-methylamino-benzoic acid and catalyse sulfoxide formation in 4-methylthiobenzoic acids.[9] Alkyl-substituted benzoic acids, such

[a] *Dr. T. Coleman, J. Z. H. Lee, D. Z. Doherty, M. N. Podgorski, D. K. Pinidiya, Dr. S. G. Bell*
Department of Chemistry, University of Adelaide
Adelaide, SA, 5005 (Australia)
E-mail: stephen.bell@adelaide.edu.au

[b] *A. M. Kirk, Prof. J. J. De Voss, Dr. E. H. Krenske*
School of Chemistry and Molecular Bioscience
University of Queensland
St Lucia, Qld, 4072 (Australia)
E-mail: e.krenske@uq.edu.au

[c] *Dr. J. B. Bruning*
School of Biological Sciences, University of Adelaide
Adelaide, SA 5005 (Australia)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202201895



Scheme 1. The product formation pathways involved in aliphatic hydroxylation (grey) and aromatic hydroxylation (black) by a P450 enzyme. In the aliphatic hydroxylation pathway, the hydroxylation product arises from a radical rebound mechanism. In aromatic hydroxylation, the reaction can proceed by formation of an arene oxide (NIH mechanism) or electrophilic attack of Cpd I on the aromatic π system (ipso substitution).

as 4-ethyl- and 4-isopropyl-benzoic acid, are desaturated to an alkene alongside the hydroxylation reaction at the benzylic C–H bond by CYP199A4.[8c] The enzyme can also stereoselectively epoxidise 4-vinylbenzoic acid yielding (*S*)-4-(oxiran-2-yl)-benzoic acid in 99% enantiomeric excess.[10] The loss of reducing equivalents due to uncoupling side reactions in many of these CYP199A4 oxidation reactions with *para*-substituted benzoic acids is low resulting in high levels of product formation.[8e,9–11] The high selectivity of this enzyme for the *para*-substituted benzoic acids over their *meta*-substituted analogues was demonstrated to be due to the orientation of the C–H bonds, the sulfur or the vinyl substituent relative to Cpd I.[12]

Recently, we demonstrated that 4-cyclohexylbenzoic acid was oxidised by CYP199A4 to 4-(1′-hydroxycyclohexyl)benzoic acid and 4-[(*E*)-2′-hydroxycyclohexyl]benzoic acid. However, 4-phenylbenzoic acid was not hydroxylated.[13] Biochemical assays and analysis of the X-ray crystal structures highlighted that both substrates bound within the active site of CYP199A4 in a similar manner. Molecular dynamics simulations showed that the C–H bonds of the cyclohexyl ring could achieve a suitable orientation to be attacked by Cpd I but that the aromatic ring of the phenyl group could not resulting in no aromatic oxidation. It has been reported by others that the closely related enzyme, CYP199A2 from *R. palustris* strain CGA009, can catalyse the aromatic oxidation of several substrates.[14] Similar studies have been undertaken on other P450 systems to highlight the requirements for efficient aromatic oxidation.[15]

To understand why aromatic oxidation in this system does not occur as readily as other monooxygenase reactions, and to further investigate the mechanistic requirements for aromatic oxidation by cytochrome P450 enzymes in general, we chose a selection of substrates including 4-benzyl-, 4-phenoxy- and 4-benzoyl-benzoic acids, for assessment with CYP199A4 (Figure 1). The first two of these would be expected to be more flexible than 4-phenylbenzoic acid and therefore might be able to attain positions within the active site more amenable to aromatic oxidation. 4-(3′-Methoxyphenyl)benzoic acid and 4-(4′-methoxyphenyl)benzoic acid were also assessed to determine if these methoxy-substituted substrates would react more readily than the unsubstituted 4-phenylbenzoic acid species. We also investigated if the substitution of selected phenylalanine amino acid residues in the active site could alter the binding orientation of the substrate to enable aromatic hydroxylation of 4-phenylbenzoic acid. Molecular dynamics simulations, in which the distance and angle of attack were considered, were undertaken to assess if 4-phenylbenzoic acid attained a productive geometry for aromatic oxidation in these mutants.

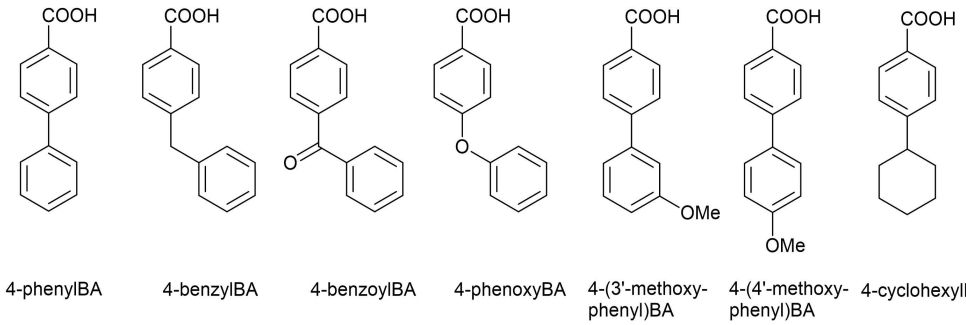


Figure 1. Compounds studied with CYP199A4 during this work. BA: benzoic acid.

Results and Discussion

Substrate binding

Substrate binding was monitored using UV-vis spectroscopy. 4-Benzylbenzoic acid induced a change in the spectrum indicative of a 90% shift from the low-spin to the high-spin state of the ferric heme when added to CYP199A4 (Figure 2, Table 1). It bound with a lower affinity (K_d = 5.2 ± 0.25 μM) than 4-methoxybenzoic acid, 4-cyclohexylbenzoic acid and 4-phenylbenzoic acid (Table 1). However, its binding affinity was superior to that of 4-*tert*-butylbenzoic acid, and the biochemical data demonstrates that this substrate could bind within the active site of CYP199A4.[12]

Addition of 4-benzoylbenzoic acid to CYP199A4 shifted the Soret maximum from 419 nm to 421 nm. The difference spectrum was like those obtained for ligands classified as inducing a type II shift, displaying a minimum (trough) at ~415 nm and a maximum (peak) at ~430 nm (Table 1, Figure 2 and Figure S1 in the Supporting Information). This substrate bound with a dissociation constant of 2.3 ± 0.4 μM. The carbonyl group, which is the only difference in the substrate structure compared to 4-benzylbenzoic acid, is most likely interacting with the distal aqua ligand of the heme. Alternatively, it may have displaced the aqua ligand and is interacting directly with the iron. 4-Phenoxybenzoic acid, in which the methylene bridge of 4-benzylbenzoic acid has been replaced with an oxygen atom, induced virtually no perturbation in the UV-vis spectrum of CYP199A4 (Figure 2). Therefore, whether this substrate binds within the active site of CYP199A4 requires further clarification. 4-(3′-Methoxyphenyl)benzoic acid generated a ≥95% shift to high-spin in CYP199A4, but with 4-(4′-methoxyphenyl)benzoic acid a much lower shift of ~10% was observed (Figure S2). 4-(3′-Methoxyphenyl)benzoic acid bound with significantly higher affinity than the 4′-methoxy isomer (0.6 vs. 100 μM, Table 1, Figure S1). Overall, the data highlight that the majority these aromatic substrates, except for 4-(4′-methoxyphenyl)benzoic acid were a good fit for the active site of CYP199A4.

Substrate turnover and product distributions

4-Benzylbenzoic acid was oxidised by CYP199A4 to yield a single major product with a high product formation rate (522 min^{-1}) and coupling efficiency (Table 1). The HPLC of the whole-cell oxidation reaction showed a second peak which coeluted with 4-benzoylbenzoic acid (Figure 3). The major metabolite was isolated by semi-prep HPLC and characterised by NMR spectroscopy (Figure S3). The NMR spectrum of the

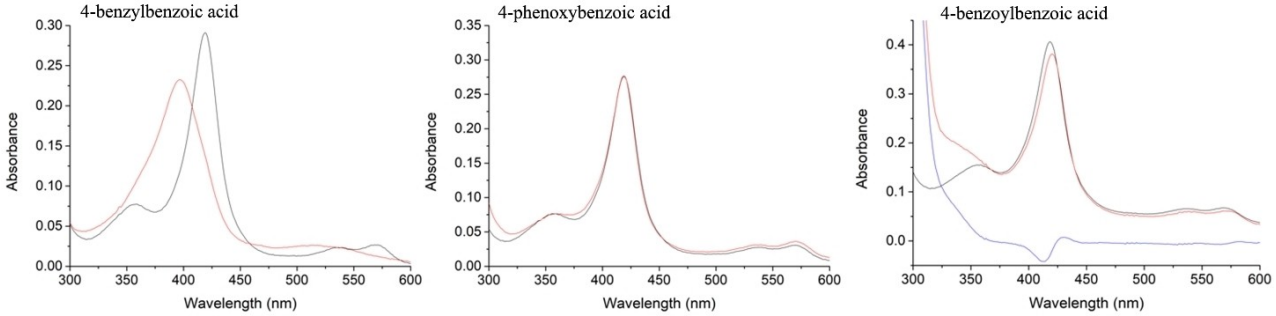


Figure 2. UV-vis analysis of the spin state of CYP199A4 with selected aromatic *para*-substituted benzoic acids. The substrate-free form is shown in black and the substrate-bound form in red. Where a type II shift was observed, the difference spectrum is shown in blue.

Table 1. Substrate binding and in-vitro turnover data for CYP199A4 and selected substrates.

Substrate BA benzoic acid	HS [%]	K_d [μM]	NADH[a] min^{-1}	PFR min^{-1}	Coupling[b] [%]
4-benzylBA	90	5.2 ± 0.25	549 ± 6	522 ± 61	95 ± 11
4-benzoylBA	type II[c]	2.3 ± 0.4	9.2 ± 1	–[e]	–[e]
4-phenoxyBA	< 5	–[d]	10.7 ± 1	–[e]	–[e]
4-(3′-methoxyphenyl)BA	≥ 95	0.6 ± 0.1	703 ± 65	–[f]	–[f]
4-(4′-methoxyphenyl)BA	10	100 ± 10	62 ± 2	–[f]	–[f]
4-phenylBA[13]	90	1.7 ± 0.1	902 ± 34	–[e]	–[e]
4-cyclohexylBA[13]	≥ 95	0.45 ± 0.05	169 ± 11	52 ± 6	33 ± 2
4-*tert*-butylBA[12]	≥ 95	39 ± 2	227 ± 4	227 ± 32	100 ± 13
4-methoxyBA[8c]	≥ 95	0.28 ± 0.01	1340 ± 30	1220 ± 120	91 ± 10

The NADH oxidation rate is the frequency of NADH oxidation, including product formation and uncoupling reactions. PFR is the product formation rate. The reactions were performed using a HaPuR:HaPux:CYP199A4 concentration ratio of 1:10:1 (0.5 μM CYP199A4 enzyme, 50 mM Tris, pH 7.4). Rates are reported as mean ± S.D. ($n \geq 3$) and given in nmol.nmol-CYP^{-1}.min^{-1} (abbreviated to min^{-1}). [a] The NADH leak rate was measured to be ~9.0 min^{-1}. [b] The coupling efficiency is defined as the percentage of NADH utilised for the formation of oxidation products. [c] A shift on the Soret band to 421 nm was observed (the type II difference spectrum presented in Figure 2). [d] Not able to be determined due to minimal spectral changes upon addition of substrate. [e] No detectable product formation was observed. [f] Very low levels of product were formed.

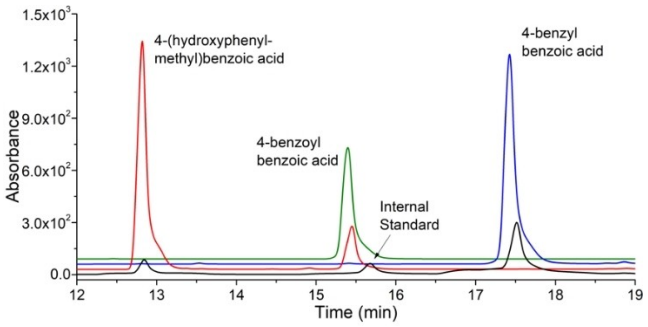
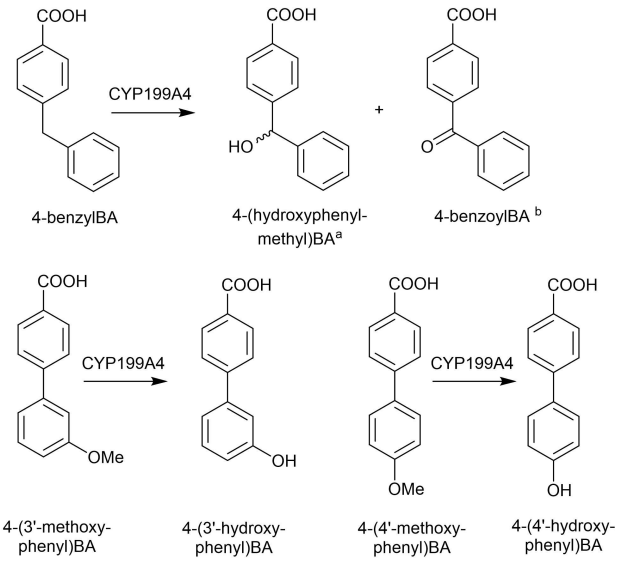


Figure 3. HPLC analysis of the CYP199A4-catalysed oxidation of 4-benzylbenzoic acid. The CYP199A4 in-vitro turnover (black), substrate control (blue), whole-cell oxidation turnover (red) and 4-benzoylbenzoic acid product control (green) are shown. The internal standard in the in-vitro oxidation, 9-hydroxyfluorene, is at 15.6 min (not present in the whole-cell oxidation). The substrate eluted at 17.5 min, the 4-benzoylbenzoic acid further oxidation product at 15.5 min, and the 4-[hydroxy(phenyl)methyl]benzoic acid product at 12.8 min. For clarity the chromatograms have been offset along the *y*-axis.

product contained the same number of aromatic protons as the substrate and this along with the ^{13}C NMR enabled unambiguous assignment as 4-[hydroxy(phenyl)methyl]benzoic acid (Scheme 2 and Figure S3). The reaction proceeded with moderate enantioselectivity; ~50% *ee* (Figure S4a).The other peak, which appeared in the whole-cell oxidation system was confirmed as 4-benzoylbenzoic acid by ^{1}H NMR (Figure S3).

Neither 4-benzoylbenzoic acid nor 4-phenoxybenzoic acid were oxidised by CYP199A4 (Figure S4b–d). In line with their respective inhibitory type II and low ($\leq$5%) spin state shifts the NADH oxidation activity of both with the CYP199A4 enzyme was like that of the leak rate (Table 1). The oxidation of the substrate 4-(3′-methoxyphenyl)benzoic acid by CYP199A4 occurred with a high NADH oxidation rate (703 $\pm$ 65 min^{-1}; Table 1) but only low levels of a single product was formed (Table 1, Figure S4e and f). The NADH oxidation rate of CYP199A4 catalysed turnover of 4-(4′-methoxyphenyl)benzoic acid was significantly slower (62.3 min^{-1}) than the 3-methoxy isomer. Low levels of a single product were also formed from the CYP199A4 catalysed oxidation of this substrate ($\leq$1% coupling efficiency; Figure S4g). For both methoxyphenyl benzoic acids the products were identified by GC-MS analysis of the trimethylsilyl derivatised reactions (Figure S5). In each, the MS demonstrated that the product arose from oxidative demethylation of the methoxy group (the mass of the product was greater than the product by 58 after derivatisation with BSTFA/TMSCl to add trimethylsilyl groups at the acid and alcohol moieties; Figure S5). The aromatic π system is not attacked by Cpd I in these enzyme/substrate combinations. However, aliphatic C–H bond abstraction of the methoxy group can occur. The results we observe here for these enzyme/substrate combinations arise, not from attack of the aromatic pi system by Cpd I, but from aliphatic C–H bond abstraction of either the benzylic methylene or methoxy group. In the case of the latter, however, the methoxy groups seem to achieve a viable orientation for H-abstraction in only a small number of instances given the very low levels of metabolites observed.



Scheme 2. Products formed from the CYP199A4-catalysed oxidation of selected substrate. BA: benzoic acid. [a] 4-Benzoylbenzoic acid was formed as a further oxidation product during whole-cell oxidation reactions.

Crystal structures of substrate bound forms of CYP199A4

To better understand the substrate binding and the activity data obtained with these aromatic compounds and CYP199A4, we attempted to obtain crystal structures of the enzyme bound with the different substrates. We were able to obtain crystal structures of CYP199A4 bound with 4-phenoxy-, 4-benzoyl- and 4-(3′-methoxyphenyl)-benzoic acids (Table S1, PDB IDs 7TND, 8D39 and 7JXB, respectively). Each crystallised in $P2_1$ symmetry with a single subunit. In each substrate-bound CYP199A4 complex, the overall protein fold was very similar to each other and the previously solved CYP199A4 structures (Figure S6, in each case, the rmsd was $<$0.3 Å compared to the 4-phenylbenzoic acid bound CYP199A4 structure; PDB 7JW5). For all three structures there was electron density in the active site that could be modelled by the appropriate substrate molecule. The benzoate moiety of each of the substrates were found in comparable positions to those of previously solved substrate-bound crystal structures (Figures 4 and 5). The location of the active site amino acids, the capping chloride anion and the active site water molecule, which interacts with the carboxylate moiety of the substrate, were also in similar positions, unless otherwise stated.

The 4-(3′-methoxyphenyl)benzoic acid structure (PDB 7JXB; resolution 1.66 Å) shows that the phenylbenzoic acid component of this substrate overlaps closely with that of 4-phenylbenzoic acid. As with 4-phenyl- and 4-cyclohexyl-benzoic acids the F298 residue shifts to accommodate the steric bulk of the large *para*-methoxyphenyl moiety (see Figure S7a).[13] The methoxy substituent of this moiety is positioned in a pocket, close to the F298 and F182 residues of the active site and points away from the heme (Figure 4).

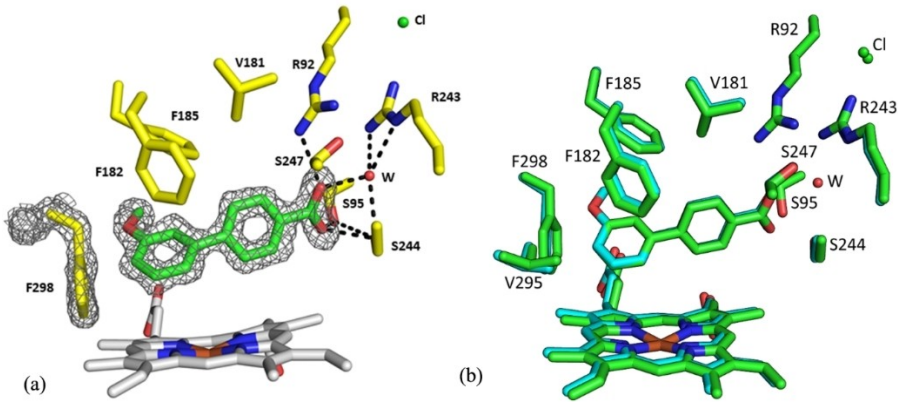


Figure 4. a) Feature-enhanced map of the active site of the 4-(3′-methoxyphenyl)benzoic acid-bound structure of CYP199A4 (PDB ID: 7JXB). The enzyme residues are in yellow, the heme is displayed in grey with the iron centre in brown, the substrate is green, and the red and green spheres are a water molecule and chloride ion, respectively. The electron density of the substrate is shown contoured as a grey mesh ($\sigma = 1.2$) using a reduced bias $2mF_o-DF_c$ composite omit map. b) Superimposed active sites of CYP199A4 with 4-phenylbenzoic acid (green, PDB ID: 7JW5) and 4-(3′-methoxyphenyl)benzoic acid (cyan) bound.

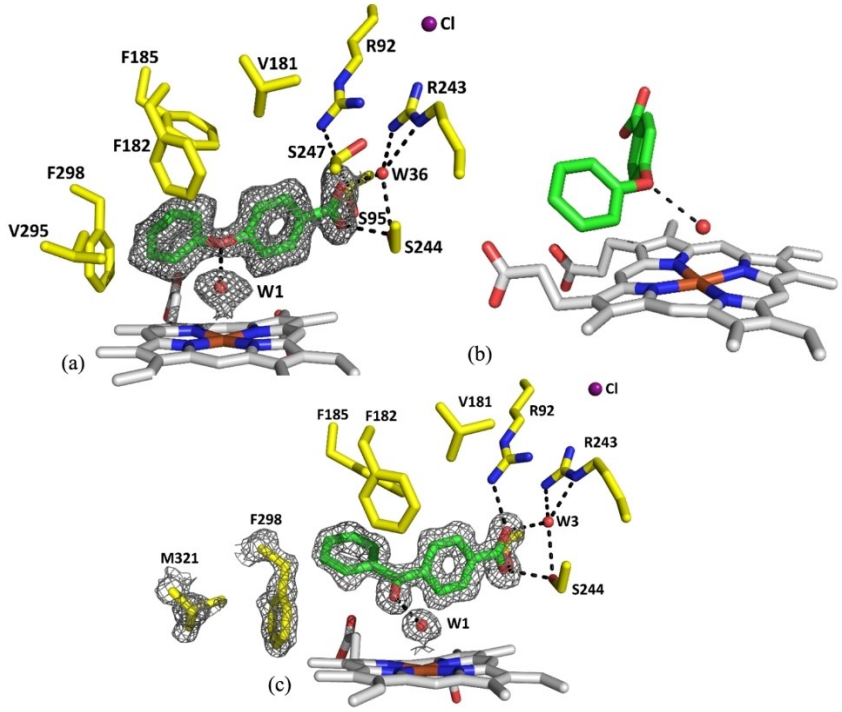


Figure 5. a) The active site of the 4-(phenoxy)benzoic acid-bound structure of CYP199A4 (PDB ID: 7TND, resolution 1.82 Å). The substrate and heme-bound water ligand are shown as $2mF_o-DF_c$ feature-enhanced map in a grey mesh (contoured at 1.0 σ, 1.5-Å carve). The colour scheme is the same as in Figure 4, except that the chlorine ion is purple. b) Interaction of the phenoxy ether oxygen and the heme-bound water (W1). This view highlights the bent nature of the phenoxy ligand. c) The crystal structure of wild-type CYP199A4 bound to 4-benzoylbenzoic acid, determined at 1.27-Å resolution. A $2mF_o-DF_c$ composite omit map of 4-benzoylbenzoic acid, the heme-bound aqua ligand, F298 and M321 is shown as grey mesh contoured at 1.0σ (1.5 Å carve radius). Both F298 and M321 are displaced from their usual positions to accommodate the bulky substrate (see Figure S7 for further details). The heme-bound aqua ligand is retained when 4-benzoylbenzoic acid is bound and is hydrogen-bonded to the ketone oxygen.

The aromatic Cα, Cβ and Cγ carbon atoms of the methoxyphenyl ring in the 4-(3′-methoxyphenyl)benzoic acid are as close to modelled Cpd I oxygen atom as those of the 4-cyclohexyl- and 4-phenyl-benzoic acid substrates (2.7 to 3.9 Å; see Table S2 for details). Previously we used molecular dynamics to demonstrate that 4-phenylbenzoic acid cannot attain a suitable orientation relative to Cpd I for efficient aromatic oxidation.[13] Given that no aromatic hydroxylation occurs with 4-(3′-methoxyphenyl)benzoic acid this implies that as with 4-phenylbenzoic acid, other factors than distance must be important in determining whether aromatic hydroxylation can occur. Given the similarity between these structures the

Fe=O–C angles for the carbons of the aromatic ring relative to heme iron-oxo group might not be ideal for aromatic oxidation (Table S2). The demethylation of the methoxy group observed for 4-(3′-methoxyphenyl)benzoic acid is not readily explained by its crystal structure. The observed activity is most likely to be explained by an alternative conformation of substrate within the active site, which is not observed in the crystal structure. This suggests some mobility of the substrate with respect to rotation of the phenylmethoxy ring. It is also in agreement with the very low levels of oxidised metabolites that are formed.

4-Phenoxybenzoic acid induces a $<5\%$ shift from low-spin to high-spin when it is added to CYP199A4. This implies that this substrate may not bind to CYP199A4 or that if it does bind the heme-bound water ligand is not displaced. The crystal structure of CYP199A4 with this substrate clearly demonstrates the latter is the case. The ether oxygen of 4-phenoxybenzoic acid is hydrogen-bonded to the heme-bound aqua ligand (W1; Fe–OH_2 distance 2.45 Å, Table S2) with an O–O distance of 2.8 Å. This interaction explains why the water ligand is not displaced and the minimal spectroscopic changes observed on substrate binding.

In this substrate the electron density of both aromatic rings could be clearly resolved. However, the resolution of the phenoxy aromatic ring was lower (Figure 4) and the B factors indicated that this part of the ligand was more mobile than the benzoate moiety (Table S3a). However, we note that the B factors of this substrate where lower than those for 4-phenylbenzoic acid (Table S3b). This substrate bound with the phenoxy aromatic ring in the pocket normally occupied by Phe298 (Figure 5). The distance of the closest aromatic carbons of the 4-phenoxybenzoic acid substrate were further from the heme than those of the phenylbenzoic acid analogues (>4.8 Å, Table S2, Figure S7).

We were unable to obtain a crystal structure of CYP199A4 with 4-benzylbenzoic acid but if this substrate bound in the active site in a similar fashion to 4-phenoxybenzoic acid we would expect the hydrophobic nature of the methylene group to displace the 6th aqua ligand of the heme. This methylene group would also be ideally located for C–H bond abstraction which, when coupled to its more reactive nature, would explain the regioselectivity for oxidation at this carbon. The crystal structure of CYP199A4 bound with 4-benzoylbenzoic demonstrated that the ketone moiety points towards the heme (Figure 5c). It interacts with, the 6th aqua ligand which is not displaced, and this would result in the altered low-spin state of the ferric heme as observed in UV-vis spectrum. In the 4-benzoylbenzoic acid structure, F298 and M321 have been shifted further than with the other substrates to accommodate the sterically bulky substrate (Figure 5c). F185 also shifts subtly to accommodate 4-benzoylbenzoic acid (Figure S7c–e).

Engineering CYP199A4 for aromatic oxidation

Given that we did not observe aromatic oxidation by the wild-type (WT) CYP199A4 enzyme with any of the substrates we decided to assess if we could achieve this through protein engineering. The residues which interact with the benzoate portion of the substrate molecules are known to be critical for substrate binding.[8a,c,16] Therefore, we focused on mutations of residues which are in contact with the phenyl moiety of 4-phenylbenzoic acid and in the crystal structure with 4-(3′-methoxyphenyl)benzoic described above.[13] We chose the phenylalanine residues at position 182 and 298 (Figures 4 and 5). F182L was chosen as the equivalent mutation in CYP199A2 has been reported to enhance aromatic hydroxylation of cinnamic acid derivatives.[14a] F298V was chosen to generate enough space for the bulkier ligands to be accommodated in the region normally occupied by the mobile side chain of F298. We generated, produced and purified the F298V and F182L mutants of CYP199A4 (see the Supporting Information for details). The F298V mutant could be heterologously produced in similar quantities to the WT enzyme but the F182L mutation resulted in a significant decrease in the levels of soluble, active protein that was generated using *Escherichia coli*. These mutations were chosen to generate additional space within the substrate binding pocket and to alter the position of the substrate in the enzyme. This could enable a greater degree of freedom for rotation of phenyl rings either of which might permit the substrate to attain a position suitable for aromatic oxidation. Both mutants were tested with 4-phenylbenzoic acid and 4-cyclohexylbenzoic acid as a control. The UV-vis spectra of these mutants demonstrated that both substrates bound to the F298V mutant (both resulting in $\geq 85\%$ high-spin species; Figure S8). However, the binding of 4-phenylbenzoic acid to the F182L variant showed a reduction in the amount of the high-spin species ($\sim 10\%$) yet 4-cyclohexylbenzoic acid induced a 90% shift to the high-spin form in this mutant. This suggests a different binding position for 4-phenylbenzoic acid in this mutant, which is less able to displace the heme-bound aqua ligand.

The F298V mutant was unable to oxidise 4-phenylbenzoic acid despite a high rate of NADH consumption (NADH oxidation rate of 1030 ± 90 min^{-1}; Figure S9). In the control experiments the F298V variant was able to oxidise 4-cyclohexylbenzoic acid with similar activity to the WT enzyme (NADH oxidation rate of 162 ± 55 min^{-1}, coupling efficiency $31 \pm 9\%$; Figure S10). There was a moderate change in product selectivity compared to the WT enzyme. Both the WT and the mutant generated 4-(2-hydroxycyclohexyl)benzoic acid as the major metabolite but the F298V variant generated more of the minor hydroxylation product at the benzylic carbon (34 vs 24%; Figure S10).

We were able to solve the crystal structures of the F298V mutant with both substrates (PDB 7TNF and 7TNU, respectively, Table S4). While the overall structures were like those of the WT enzyme (Figure S11) some changes were observed in the active site (Figures 6, S12 and S13). In both structures of the mutant the modification of F298 to the less bulky residue valine enables additional waters to enter the active site close to the threonine 395 residue (Figure 6). The F298V mutation subtly altered the position of 4-phenylbenzoic acid within the active site (Figure 6a). Overall, this substrate was slightly further from the heme in the mutant when compared to the WT enzyme (Table S5). In contrast, the position of 4-cyclohexylbenzoic acid

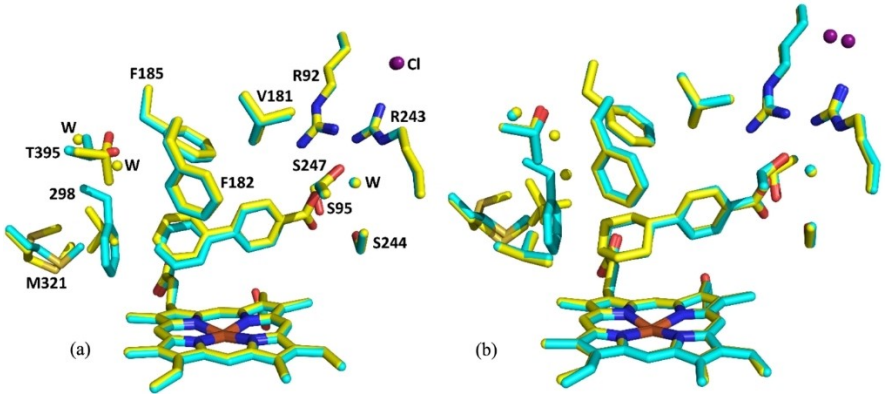
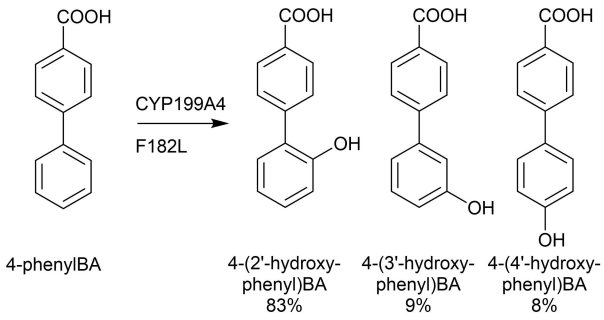


Figure 6. Overlaid structures of WT and F298V CYP199A4 bound with 4-phenylbenzoic and 4-cyclohexylbenzoic acids. a) 4-Phenylbenzoic acid bound to WT CYP199A4 (cyan, PDB ID: 7JW5) and F298V CYP199A4 (yellow, PDB ID: 7TNF). b) 4-Cyclohexylbenzoic acid bound to WT CYP199A4 (cyan, PDB ID: 6C2D) and F298V CYP199A4 (yellow, PDB ID: 7TNU). Water molecules are shown as yellow or cyan spheres.

is essentially identical in both the WT and F298V structures (Figure 6b and Table S5).

The results with the F182L mutant were more dramatic. Despite a lower NADH oxidation rate ($75 \pm 0.3\ min^{-1}$) for the in-vitro oxidation of 4-phenylbenzoic acid three potential product metabolites were identified by HPLC and GC-MS. All three had MS consistent with phenol formation (Figures S9 and S14). The two minor metabolites could be identified as 4-(3′-hydroxyphenyl)benzoic acid (9%) and 4-(4′-hydroxyphenyl)benzoic acid (8%) by matching their HPLC and GC retention times, and their mass spectra with those of authentic standards. The MS also matched those obtained above for the oxidative demethylation of 4-(3′-methoxyphenyl)benzoic acid and 4-(4′-methoxyphenyl)benzoic acid, respectively (Figure S5 and Figure S14). The major metabolite was therefore assigned as 4-(2′-hydroxyphenyl)benzoic acid (83%) by comparing its retention times and MS to that of an authentic product standard (Scheme 3, Figure S14). The coupling efficiency of the reaction was 30% resulting in a product formation rate of $26 \pm 3\ min^{-1}$.

Due to the low levels of product and the reduced production amount of the F182L variant we were unable to obtain significant levels of these metabolites for further characterisation using a whole-cell oxidation system.[8a] The F182L variant was also capable of oxidising 4-cyclohexylbenzoic acid resulting in a reduction in the amount of total metabolite formation and in the proportion of 4-(2-hydroxycyclohexyl)benzoic acid generated. Benzylic hydroxylation gave the major product with a couple of other minor metabolites being observed (Figure S15). Importantly, we have been able to demonstrate that aromatic oxidation is possible with the F182L mutant of the CYP199A4 enzyme. This suggests that in this variant the substrate must be able to adopt conformations which enables this activity. This agrees with others who studied the equivalent mutation (F185 L) in the CYP199A2 enzyme or the oxidation of cinnamic acid and naphthoic acid derivatives.[14,17]



Scheme 3. Products formed from oxidation of 4-phenylbenzoic acid catalysed by the CYP199A4 F182L mutant. BA: benzoic acid.

Molecular dynamics of 4-phenylbenzoate bound to the CYP199A4 F182L mutant

Molecular dynamics (MD) simulations were employed to investigate the conformational changes in substrate binding with the CYP199A4 F182L mutant. To achieve this, in-silico mutation of the 4-phenylbenzoic acid bound CYP199A4 WT crystal structure (PDB ID: 7JW5) was performed using *UCSF ChimeraX*.[13] The F182 benzyl side chain was converted to an isobutyl group through deletion of three CH and addition of five hydrogens. Two initial starting structures (A1 and A2) were generated with slightly different conformations of the L182 side chain. MD simulations were performed as described previously with duplicate runs beginning from each of the starting conformations.[13] An additional 200 ns of MD simulations was also performed for each run. A conformational change was observed between 50–250 ns across all 4 runs after which the two initial starting structures had converged: the L182 side chain adopted the same conformation with subsequent reorientation of the F298 side chain. This is illustrated in Figure 7.

Note the new conformation of the L182 side chain is consistent with that predicted by a rotamer search in *UCSF*

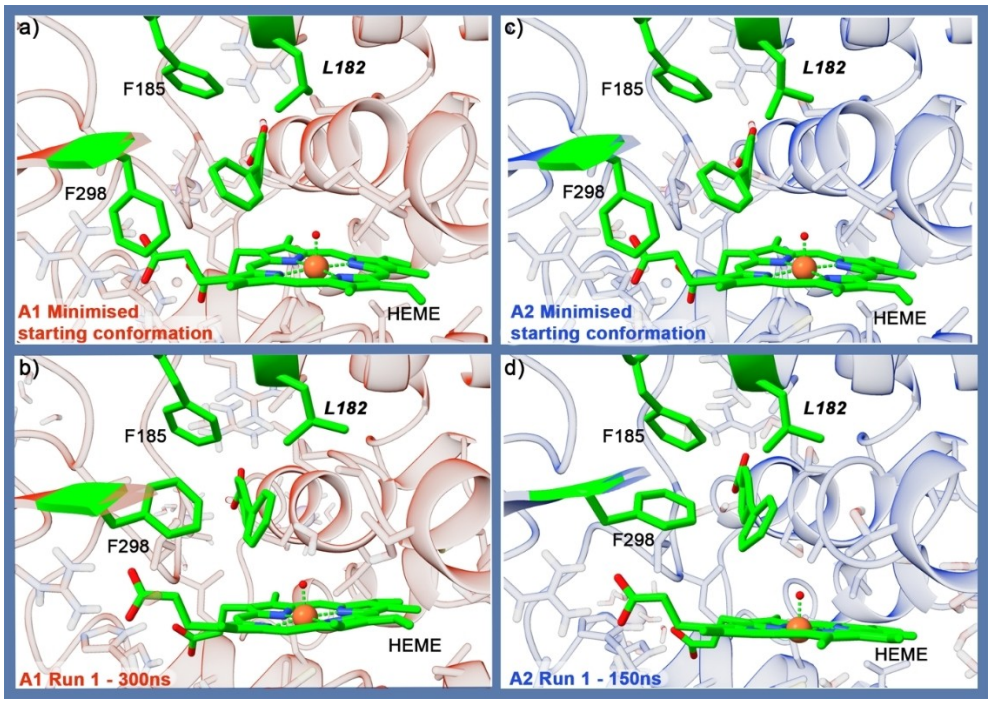


Figure 7. a) Minimised starting conformation A1 for in-silico F182L mutation and b) conformation after 300 ns in Run 1 of MD simulations showing the change in 4-phenylbenzoate biphenyl dihedral and position of the F298 side chain. c) Alternative starting conformation A2 and d) subsequent conformation after 150 ns MD simulations in Run 1.

ChimeraX (Figure S16). A change in the F298 N–CA–CB–CG dihedral angle was indicative of the overall conformational change (Figure S17). The movement of the F298 side chain back into the active site places it in a position like the crystal structures with smaller substrates like 4-vinylbenzoate bound as shown in Figure 8. The biphenyl dihedral of the substrate also shifts by approximately 80° with the conformational change. The system was therefore, considered equilibrated after 250 ns. For all runs, the RMSD remained relatively level suggesting a high stability of the overall enzyme conformation (Figure S18) with deviations of $\leq$2.5 Å from the WT crystal structure CYP199A4 with 4-phenylbenzoic acid bound. The RMSF indicated flexibility in the loop regions and D helix, consistent with that observed for the WT enzyme (Figure S18).

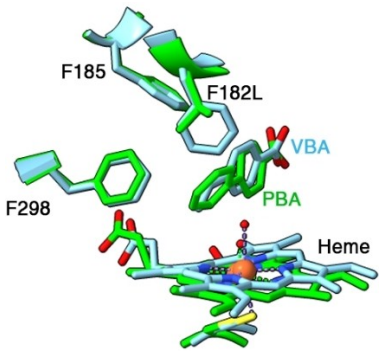


Figure 8. Snapshot from A2 Run 2 450 ns of MD simulations of 4-phenylbenzoate with the F182L mutant (green) overlayed with the crystal structure of 4-vinylbenzoic acid bound to CYP199A4 WT (PDB ID: 7KCS; blue).

Snapshots from 250–700 ns of the MD were analysed along four key geometric parameters (describing distance and angle of attack) which were previously determined useful for understanding substrate positioning for P450-catalysed aromatic oxidation.[13] These geometric parameters can be used to assess whether potential sites of oxidation are proximate to the Fe=O with an alignment in which the aromatic π-cloud is accessible for attack by the Fe=O. Figure 9 shows the distribution of C=O distance and Fe–O–C angle. It appears that with the F182L mutant Cβ, Cγ, and Cδ are all near the Fe=O approaching within a C–O distance of 3 Å and at suitable Fe–O–C angles compared to the DFT transition-state structures.

Note, as reported previously,[13] the DFT transition-state structures (Figure S19) are taken to represent "ideal" geometries of which there are two common types of arrangements (i.e., TS1 and TS2). Figure 10 shows the distribution of C_Y–C_X–O angles and C_Z–C_Y–C_X–O dihedral angles across the MD snapshots, where C_X is the site of oxidation being either Cβ, Cγ, or Cδ. It is these distributions that differ most notably compared to those calculated for the WT enzyme.

For the F182L mutant, it seems that the additional space in the active site and the absence of a favourable T-shaped interaction between the phenyl rings of the substrate and F182 allows the substrate phenyl ring to rotate to achieve conformations where the aromatic π-cloud is accessible by the Fe=O for oxidation at Cβ, Cγ, and Cδ. This is indicated by the measured angles in the snapshots approaching the values of the DFT

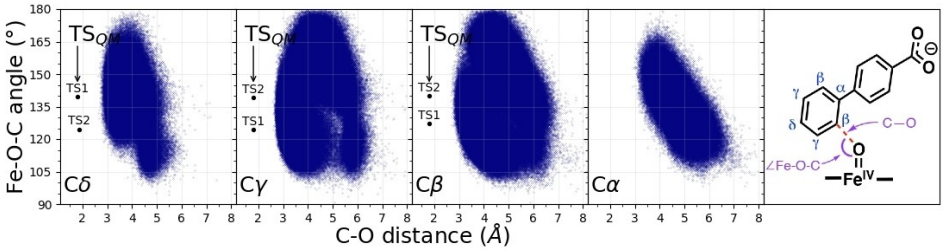


Figure 9. Distributions of C–O distances [Å] and Fe–O–C angles [°] for the phenyl carbon atoms during 250–700 ns of MD simulations of 4-phenylbenzoate bound to Cpd I of CYP199A4 F182L. Black dots represent the QM ideal TS geometries: Cβ TS1 (1.81 Å, 140°), TS2 (1.83 Å, 127°); Cγ TS1 (1.78 Å, 139°), TS2 (1.82 Å, 124°); Cδ TS1 (1.85 Å, 140°), TS2 (1.80 Å, 125°).

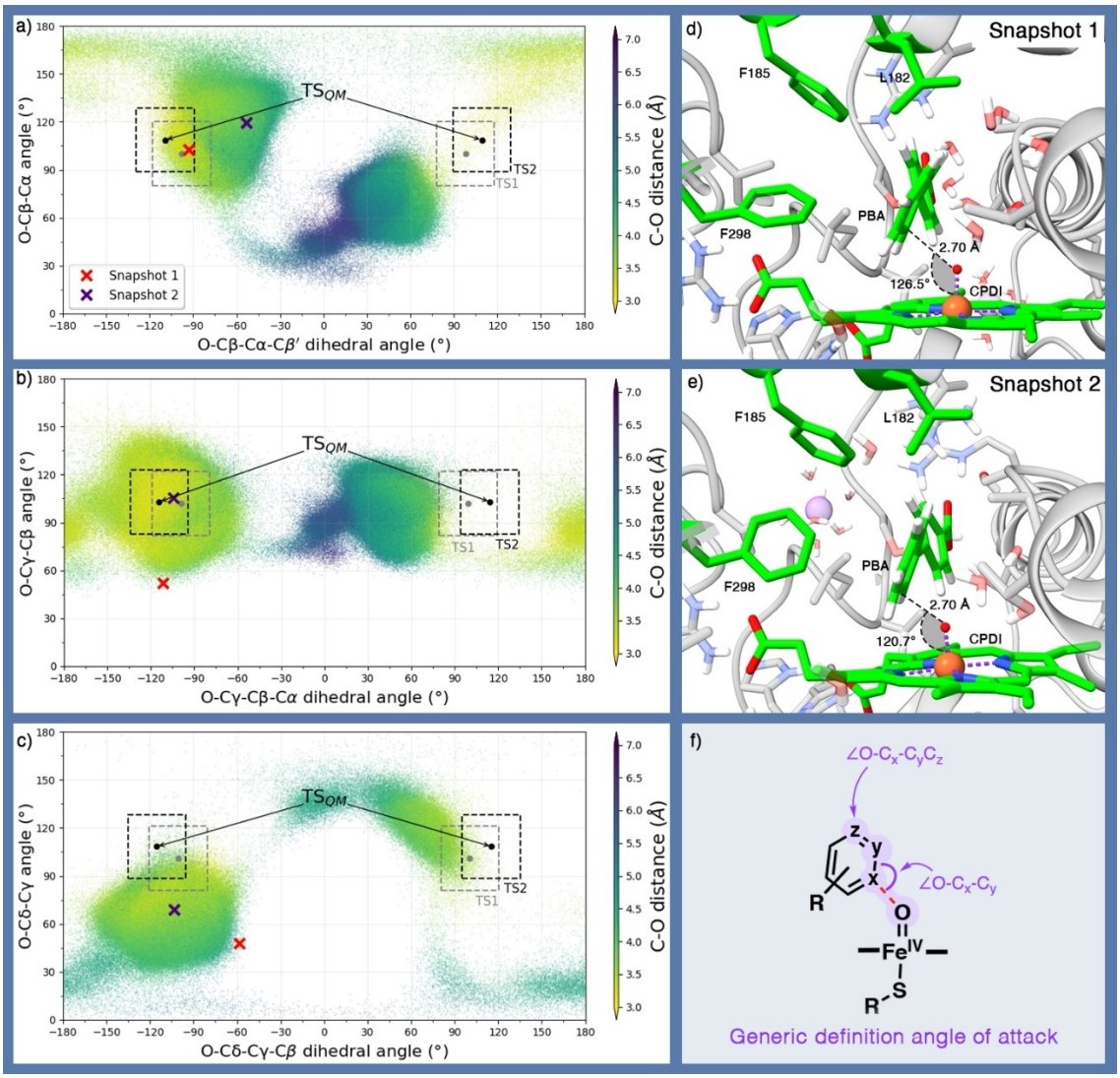


Figure 10. Distribution of angles of attack of Fe=O towards a) Cβ as defined by the O–Cβ–Cα angle [°] and O–Cβ–Cα–Cβ′ dihedral [°], b) Cγ as defined by the O–Cγ–Cβ angle [°] and O–Cγ–Cβ–Cα dihedral [°] and c) Cδ as defined by the O–Cδ–Cγ angle [°] and O–Cδ–Cγ–Cβ dihedral [°] in MD simulations of 4-phenylbenzoate bound to Cpd I of CYP199A4 F182L. Black and grey dots represent the QM ideal TS geometries for TS1 [Cβ (±109°, 109°)/Cγ (±114°, 103°)/Cδ (±115°, 108°)] and TS2 [Cβ (±98°, 100°)/Cγ (±99°, 102°)/Cδ (±100°, 101°)], respectively, with ±20° margins outlined by rectangles. Representative snapshots from the MD simulations are shown for conformations approaching the TSs for oxidation at d) Cβ and e) Cγ with the corresponding locations marked with an X on the distribution maps. f) Generic definition of angles of attack depicted.

transition-state structures to within ±20° as demarcated by dashed black and grey boxes. Two snapshots are shown to illustrate MD snapshots where the conformation appears suitable for oxidation at Cβ or Cγ respectively.

Table 2. Percentage of snapshots from 250–700 ns of MD simulations with catalytically active poses as defined by C–O distance <4 Å, Fe–O–C, C–C–O and C–C–C–O angles within $\pm 20°$ of DFT transition-state structure TS1 or TS2. Snapshots considered catalytically active for multiple sites of oxidation simultaneously have been divided equally between those two sites.

Site of oxidation	Catalytically active poses in run [%] A1 Run 1	A2 Run 1	A1 Run 2	A2 Run 2	Average % ± SD	Predicted ratio (error)
Cβ (*ortho*)	7	3	4	15	7 ± 6	10 (8)
Cγ (*meta*)	86	50	46	72	60 ± 20	86 (34)
Cδ (*para*)	2	5	3	2	3 ± 1	4 (2)
Total	95	58	53	89	70 ± 20	

To investigate if the MD simulations could provide useful information to explain the regioselectivity of aromatic hydroxylation, the percentage of catalytically active poses observed during the MD simulations were quantified. Table 2 summaries the percentage of snapshots considered catalytically active according to the following criteria: C–O distance <4 Å, and Fe–O–C angle, C_Y–C_X–O angle and C_Z–C_Y–C_X–O dihedral angle within $\pm 20°$ of the DFT transition-state values. Note that a small percentage of snapshots were considered suitable for oxidation at two sites. For example, 3 ± 2% of snapshots that appear suitable for oxidation at Cγ are also considered suitable for oxidation at Cβ. These snapshots were divided equally between the two sites. Refer to the Table S6 for a breakdown of percentages of catalytically active poses over time. Based on the average percentage of catalytically active poses for oxidation at each site, the predicted ratio of Cβ/Cγ/Cδ oxidation from the MD simulations would be 10:86:4 which is not in agreement with the experimentally determined ration of 83:9:8. However, oxidation at Cβ is intrinsically preferred based on the DFT transition-state energies which give a 6.3 kcal mol^{-1} lower barrier for attack at Cβ compared to Cγ (Figure S19). Therefore, it would be expected that significant proportions of oxidation at Cβ would occur if both sites are accessible for reaction and oxidation at Cβ remained kinetically more favourable than oxidation at Cγ.

Conclusion

The CYP199A4-catalysed oxidation of 4-benzylbenzoic acid was regioselective for hydroxylation at Cα with no oxidation occurring at the aromatic ring. 4-Phenoxybenzoic acid, in which the benzylic methylene group of 4-benzylbenzoic acid is replaced with an oxygen atom, was not oxidised by CYP199A4. The crystal structure provided evidence that 4-phenoxybenzoic acid did bind in the active site of CYP199A4 and that the aqua ligand was not displaced. The ether oxygen of the substrate interacts with this aqua ligand. The observation that 4-benzoylbenzoic acid binding to CYP199A4 induces a type II shift could be explained as in the crystal structure the carbonyl oxygen is interacting with the heme water ligand.

The crystal structure of CYP199A4 bound with 4-(3′-methoxyphenyl)benzoic acid was obtained and showed that the methoxy substituent, which is the site of oxidation, pointed away from the heme. The carbons of the *para*-aromatic substituent were a similar distance from the heme iron to the aliphatic C–H bonds of previously studied alkylbenzoic acids (e.g., 4-cyclohexylbenzoic acid) and 4-phenylbenzoic acid. As no aromatic oxidation was detected, these results provide further evidence that oxidation of these aromatic substrates is not well supported by the WT CYP199A4 enzyme. This substrate bound to CYP199A4 with high affinity, and the enzyme could undergo the catalytic cycle, but no detectable product from aromatic oxidation was generated. This must result from the positioning of the substrate in the active site relative to the reactive compound I intermediate. The related substrate 4-(4′-methoxyphenyl)benzoic acid bound to CYP199A4 with much lower affinity, thus suggesting that its shape is less compatible with the active site. Despite this it was also oxidatively demethylated by the enzyme with low activity. There was, however, no evidence for aromatic oxidation.

Aromatic oxidation is known to be supported by P450 enzymes. The high bond strength of aromatic C–H bonds means that this reaction occurs according to a different mechanism from aliphatic hydroxylation (Scheme 1).[7a,c,18] Mutagenesis of the F298 residue did not enable the aromatic oxidation of 4-phenylbenzoic acid. While small differences were observed in the binding mode, as evidenced by the crystal structures and the UV-vis spectra, no aromatic oxidation was observed. Therefore, the required orientation of the aromatic π-system of the substrate relative to Cpd I of the enzyme must not be attained in this system. The F182L mutation must be able to alter the binding orientation of 4-phenylbenzoic acid to enable aromatic oxidation to occur. MD simulations, which consider the trajectory of attack of the Fe=O moiety on the π system demonstrated that, in the F182L mutant, the aromatic ring of 4-phenylbenzoic acid was able to attain the correct geometry and distance with respect to the iron-oxo to enable its oxidation. Further studies would be required to obtain a more stable CYP199 family heme monooxygenase catalyst suitable for the scale-up of the aromatic oxidation observed here.

Experimental Section

General: General reagents and organic substrates were from Sigma-Aldrich, TCI, Fluorochem or VWR. Product standards were from Carbosynth (UK). Buffer components, NADH and isopropyl β-D-thiogalactopyranoside (IPTG) were from Astral Scientific, Australia. The expression, purification and characterisation of WT CYP199A4, HaPux and HaPuR have been described elsewhere unless otherwise stated.[8a–c] The CYP199A4 protein concentration was calculated using $\varepsilon_{419} = 119$ mM^{-1} cm^{-1} as described previously.[8a–c] Prior to use, glycerol was removed from CYP199A4 using a 5 mL gel filtration column (PD-10, GE Healthcare). The F182L and F298V mutants of CYP199A4 were prepared using standard methods (see the Supporting Information for more details) and were produced and purified using the same method as the WT enzyme.

UV-vis spectra and spectroscopic activity assays were recorded at 30 ± 0.5 °C on an Agilent CARY-60 spectrophotometer. Analytical HPLC was performed on an Agilent 1260 Infinity Pump equipped

with an autoinjector or a Shimadzu LC-20AD equipped with a SIL-20 A autosampler, CTO-20 A, SPD-20 A UV detector and CBM-20Alite communications module was used. Each used an Agilent Eclipse Plus C_{18} or a Phenomenex Kinetex 5u XB–C18 100 A column (250 mm x 4.6 mm, 5 μm). A gradient of 20–95% acetonitrile in water (both containing trifluoroacetic acid, TFA, 0.1%) was run at a flow rate of 1 mL min^{-1} over 30 min. The chromatogram was monitored at 254 nm. Preparative HPLC was performed using a Gilson 322 pump system equipped with a Phenomenex Kinetex 5u EVO C18 100 A AXIA column (15 cm ×21.2 mm, 5 μm) and a fraction collector. Further preparative HPLC was conducted using a Schimadzu LC-20AR pump system equipped with a SIL-20AC HT autosampler and SPD–M20 A diode array detector. An isocratic concentration of 40% acetonitrile in water (TFA, 0.1%) was run at 8 mL min^{-1} and monitored using UV detection at 240 and 254 nm. Those fractions containing a single product (≥95%) were combined and freeze-dried using an Alpha 2–4 LDplus freeze-dryer (Christ, Germany), followed by characterisation using NMR and GC-MS.

GC-MS data were collected on a Shimadzu GC-17 A using a QP5050 A GC-MS detector or a Shimadzu GC-2010 gas chromatograph equipped with an autoinjector and a GCMS-QP2010S detector. Both instruments contained a DB-5 MS fused silica column (30 m x 0.25 mm, 0.25 μm). The injector and interface were maintained at a constant temperature of 250 and 280 °C. The oven temperature on the GC-17 A was held at 120 °C for 3 min and then increased at 6 °C min^{-1} up to 230 °C, where it was held for 6 min. On the GC-2010 for the first 3 min the column temperature was held at 120 °C and then increased to 240 °C at a rate of 7.5 °C min^{-1}. The column was held at this temperature for 6 min.

Substrate binding: spin state determination and binding titrations: The high-spin heme content was determined using an enzyme concentration of 2.2–3.5 μM in 50 mM Tris, pH 7.4, with the addition of 0.5–1 μL aliquots of substrate (from a 100 mM stock) until the spectrum did not change. The percentage shift was estimated (to ca. ±5%) as described previously.[8e,19]

For dissociation constant determination CYP199A4 was diluted to 0.5–2.0 μM in 50 mM Tris, pH 7.4, in a total volume of 2.5 mL. The substrate (0.5–2 μL) was added using a Hamilton syringe from a 1, 10 or 100 mM stock solutions in ethanol or DMSO. The Soret band peak-to-trough difference (ΔA) in absorbance was recorded between 700 and 250 nm. Additional aliquots of substrate were added until ΔA did not change further. The dissociation constants, K_d, were obtained by fitting ΔA against total substrate concentration [S] to the tight binding quadratic equation [Eq. (1)]:[20]

$$\frac{\Delta A}{\Delta A_{max}} = \frac{([E]+[S]+K_d)-\sqrt{\{([E]+[S]+K_d)^2-4[E][S]\}}}{2[E]} \quad (1)$$

ΔA_{max} is the maximum absorbance difference and [E] is the enzyme concentration. For examples (Table 1) where the K_d was greater than five times the concentration of the enzyme it was obtained by fitting the difference in absorbance against the substrate concentration to the hyperbolic function [Eq. (2)]:

$$\Delta A = \frac{\Delta A_{max} \times [S]}{K_d[S]} \quad (2)$$

Activity assays: In vitro NADH turnover rate assays were performed with mixtures (1.2 mL) containing 50 mM Tris, pH 7.4, 5 μM HaPux, 0.5 μM HaPuR (the electron transfer proteins) and 0.5 μM CYP199A4. The buffer solution was oxygenated before addition of the proteins. The resulting mixture was equilibrated at 30 °C for 2 min and NADH was added to ca. 320 μM (final $A_{340}=2.00$). The reaction was initiated by the addition of the substrate to a final concentration of 1 mM (from a 100 mM stock solution in ethanol or DMSO). The rate of NADH oxidation was determined by recording the absorbance at 340 nm and using the extinction coefficient, $\varepsilon_{340}=6.22$ $mM^{-1}cm^{-1}$, to determine the resulting gradient. The product formation rate and coupling efficiency were calculated after quantitating the amount of product in the turnover.

Analysis of metabolites: Several of the products of enzyme turnover were identified by co-elution of authentic standards using HPLC analysis. After the NADH had been consumed in substrate oxidation incubations, 132 μL of the reaction mixture was mixed with 2 μL of an internal standard solution (10 mM 9-hydroxyflourene in ethanol). This was mixed with 66 μL of acetonitrile before analysis by HPLC. The HPLC retention times for the substrates and products are given in the appropriate chromatograms in the Supporting Information.

GC-MS of the trimethylsilyl chloride (TMSCl) derivatised forms was used for confirmation of product identification. For gas chromatography analysis, 990 μL of the reaction mixture was mixed with 10 μL of an internal standard solution (20 mM 9-hydroxyfluorene) and 20 μL of 3 M HCl. The mixture was extracted three times with 400 μL of ethyl acetate and the organic extracts were combined and dried over $MgSO_4$. Solvent was evaporated under a stream of dinitrogen and the sample dissolved in 200 μL anhydrous acetonitrile. Excess derivatisation agent (35 μL, *N*,*O*-bis(trimethylsilyl)trifluoroacetamide with trimethylsilyl chloride; BSTFA + TMSCl, 99:1, Sigma) was added and the mixture was left at 37 °C for 2 h. These reaction mixtures were used directly for GC-MS analysis. The retention times for the derivatives are given in the appropriate chromatogram in the Supporting Information.

Calibrations of fixed amounts of product standards were used to quantitate the level of the metabolites in the turnovers by HPLC or GC-MS. When an authentic sample was unavailable, they were synthesised using whole-cell oxidation systems or the coupling was estimated based on the closest available compound.

Whole-cell oxidation reactions: When product identification required the synthesis of the metabolite, whole-cell turnovers were used to scale up the reaction. pETDuet-HaPux-HaPuR and pRSFDuet-HaPux-CYP199A4 plasmids were transformed into BL21(DE3) and grown in 200 mL lysogeny broth, containing the required antibiotics and 1.5 mL of trace elements (0.74 g $CaCl_2 \cdot H_2O$, 0.18 g $ZnSO_4 \cdot 7H_2O$, 0.132 g $MnSO_4 \cdot 4H_2O$, 20.1 g Na_2EDTA, 16.7 g $FeCl_3 \cdot 6H_2O$, 0.10 g $CuSO_4 \cdot 5H_2O$, 0.25 g $CoCl_2 \cdot 6H_2O$), at 37 °C overnight. The incubator temperature was lowered to 18 °C for 1 h before the addition of 0.25 mM IPTG (from a stock of 0.5 M in water). The culture was grown for a further 24 h, and then harvested by centrifugation (5000 *g*, 10 min). The harvested cells were resuspended in *E. coli* minimal medium (EMM; K_2HPO_4 7 g, KH_2PO_4 3 g, Na_3citrate 0.5 g, $(NH_4)_2SO_4$ 1 g, $MgSO_4$ 0.1 g, 20% glucose (20 mL) and glycerol (1% *v*/*v*) per litre and split into 200 mL aliquots in 2 L baffled flasks. Substrate (2 mM) was added to the whole-cell reaction and the mixture was shaken at 160 rpm at 30 °C for 20 h.

To isolate the products, the supernatant (200 mL) was acidified, extracted in ethyl acetate (3 x 100 mL), washed with brine (100 mL) and dried with $MgSO_4$. The organic extracts were pooled and the solvent was removed under reduced pressure. The mixture of products was separated using preparative HPLC. The solvent was removed by freeze-drying. NMR spectra were acquired on an Agilent DD2 spectrometer operating at 500 MHz for ^1H and 126 MHz for ^{13}C. A combination of ^1H and ^{13}C experiments were used to determine the structures of the products. These purified

products were then evaluated by analytical HPLC and GC-MS to assign each to an observed metabolite peak.

Protein crystallisation and X-ray crystallography: For crystallisation, the CYP199A4 enzymes were further purified by size exclusion chromatography (ENrich SEC 650, 10×300 mm) at a flow rate of 1 mL min^{-1} with Tris (pH 7.4, 50 mM) as the elution buffer and then concentrated to 30–40 mg mL^{-1} in 50 mM Tris, pH 7.4. Substrate was added to the protein from a 100 mM stock solution in EtOH/DMSO to a final concentration of 1 or 5 mM. Crystals were obtained using the hanging-drop vapour diffusion method at 16 °C using 1–1.2 μL of protein with 1–1.2 μL of reservoir solution and equilibrated with 500 μL of the same reservoir solution. Rectangular plate-shaped crystals of approx. 300×150×20 μm were obtained in 1 week using a reservoir solution containing 0.2 M magnesium acetate tetrahydrate, 20–23% *w/v* PEG-3350 and 0.1 M Bis·Tris pH 5.25–5.5. Crystals were harvested using a Microloop or Micromount (MiTeGen), then cryoprotected by immersion in Parabar 10312 (Paratone-N, Hampton Research) or NVH Immersion Oil (Cargille Labs) and flash frozen in liq. N_2. X-ray diffraction data was collected at 100 K on the MX1 and MX2 beamlines at the Australian Synchrotron.[21]

All diffraction data were indexed and integrated using iMosflm;[22] scaled, merged and R-free flags were added using Aimless,[23] both available in the CCP4 suite of programs.[24] Alternatively, automatic data processing was carried out at the Australian Synchrotron MX1 and MX2 beamlines using *xdsme* and Aimless.[23,25] The phase problem was solved using the molecular replacement method with Phaser in CCP4,[26] using 4-methoxybenzoic acid-bound CYP199A4 (PDB ID: 5UVB, with the 4-methoxybenzoic acid and heme ligands removed) or 4-cyclopropylbenzoic acid-bound CYP199A4 (PDB ID: 5UVB) as the initial search model. Electron density maps were obtained after initial model building and the model was rebuilt using Coot.[27] Structural refinements were performed over multiple iterations using Phenix Refine available in the Phenix suite of programs.[28] Composite omit maps or feature-enhanced maps were generated using the composite omit or feature-enhanced maps program in Phenix. Data collection and structural refinement statistics are summarised in Tables S1 and S4. The coordinates for the crystal structures have been deposited in the Protein Data Bank (accession codes presented in Table S1 for the WT enzyme structures and Table S4 for the F298V mutant).

MD simulations and QM calculations: MD simulations were performed with the substrate bound to the active site of the heme iron(IV)-oxo form (Cpd I) of F182L CYP199A4 using the same methodology as that described previously for the WT CYP199A4 system with an additional preparatory step, as follows.[13] After the initial preparation of the 4-phenylbenzoate bound WT CYP199A4 crystal structure, in-silico mutation of the F182 side chain from phenyl to iso-butyl was achieved using *UCSF ChimeraX*.[29] Two initial starting conformations were generated. After equilibration, MD simulations were performed with Amber 18[30] for 700 ns in duplicate from each starting conformation for a total of four runs.

Supplementary to the QM transition state structures reported previously for oxygen addition at Cβ and Cγ of 4-phenylbenzoic acid, the transition state for oxygen attack at Cδ was calculated using the same method as reported previously,[13] that is, at the B3LYP–D3BJ/6-311+G(d,p)(H,C,N,O,S) LANL2DZ(Fe)//B3LYP/6-31G(d)(H,C,N,O,S)LANL2DZ(Fe) level of theory in Gaussian 16[31] using a truncated heme-thiolate for Cpd I.

Acknowledgements

This work was supported by ARC grant DP140103229 (to J.J.D.V. and S.G.B.) and DP180103047 (E.H.K.). S.G.B. acknowledges the ARC for Future Fellowships (FT140100355).The authors also acknowledge the Australian government for Research Training Program Scholarships (PhD to T.C., H.H.Z.L, A.M.K. and M.N.P.). We also acknowledge the Westpac Bicentennial Foundation (Westpac Future Leaders Scholarship to A.M.K.), the University of Adelaide for a Constance Fraser Scholarship; and the CSIRO Synthetic Biology Future Science Platform for a Scholarship (J.H.Z.L.). This research was supported by an AINSE Ltd. Postgraduate Research Award (PGRA) to M.N.P. Computational resources were provided by the University of Queensland Research Computing Centre. We would like to thank the scientists at the MX1 beamline at the Australian Synchrotron for help with data collection. We acknowledge ANSTO for financial support and in providing the facility used in this work. Open Access publishing facilitated by The University of Adelaide, as part of the Wiley - The University of Adelaide agreement via the Council of Australian University Librarians.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: cytochromes · enzyme mechanisms · metalloenzymes · molecular dynamics · X-ray crystallography

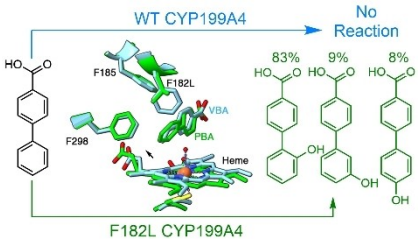
[1] a) X. Engelmann, I. Monte-Perez, K. Ray, *Angew. Chem. Int. Ed.* **2016**, *55*, 7632–7649; *Angew. Chem.* **2016**, *128*, 7760–7778; b) J. F. Hartwig, M. A. Larsen, *ACS Cent. Sci.* **2016**, *2*, 281–292; c) T. Newhouse, P. S. Baran, *Angew. Chem. Int. Ed.* **2011**, *50*, 3362–3374; *Angew. Chem.* **2011**, *123*, 3422–3435; d) Y. Qiu, S. Gao, *Nat. Prod. Rep.* **2016**, *33*, 562–581.
[2] a) P. R. Ortiz de Montellano, *Chem. Rev.* **2010**, *110*, 932–948; b) T. L. Poulos, *Chem. Rev.* **2014**, *114*, 3919–3962.
[3] a) B. Meunier, S. P. de Visser, S. Shaik, *Chem. Rev.* **2004**, *104*, 3947–3980; b) F. Ogliaro, S. P. de Visser, S. Cohen, P. K. Sharma, S. Shaik, *J. Am. Chem. Soc.* **2002**, *124*, 2806–2817; c) J. Rittle, M. T. Green, *Science* **2010**, *330*, 933–937; d) S. Shaik, S. Cohen, Y. Wang, H. Chen, D. Kumar, W. Thiel, *Chem. Rev.* **2010**, *110*, 949–1017; e) P. K. Sharma, S. P. de Visser, S. Shaik, *J. Am. Chem. Soc.* **2003**, *125*, 8698–8699.
[4] J. T. Groves, G. A. McClusky, *J. Am. Chem. Soc.* **1976**, *98*, 859–861.
[5] X. Huang, J. T. Groves, *J. Biol. Inorg. Chem.* **2017**, *22*, 185–207.
[6] M. R. Sarkar, S. D. Houston, G. P. Savage, C. M. Williams, E. H. Krenske, S. G. Bell, J. J. De Voss, *J. Am. Chem. Soc.* **2019**, *141*, 19688–19699.
[7] a) M. Asaka, H. Fujii, *J. Am. Chem. Soc.* **2016**, *138*, 8048–8051; b) D. M. Jerina, J. W. Daly, B. Witkop, *J. Am. Chem. Soc.* **1968**, *90*, 6523–6525; c) J. E. Stok, S. Chow, E. H. Krenske, C. Farfan Soto, C. Matyas, R. A. Poirier, C. M. Williams, J. J. De Voss, *Chemistry* **2016**, *22*, 4408–4412.
[8] a) S. G. Bell, A. B. Tan, E. O. Johnson, L. L. Wong, *Mol. BioSyst.* **2010**, *6*, 206–214; b) S. G. Bell, W. Yang, A. B. Tan, R. Zhou, E. O. Johnson, A. Zhang, W. Zhou, Z. Rao, L. L. Wong, *Dalton Trans.* **2012**, *41*, 8703–8714; c) S. G. Bell, R. Zhou, W. Yang, A. B. Tan, A. S. Gentleman, L. L. Wong, W. Zhou, *Chemistry* **2012**, *18*, 16677–16688; d) R. R. Chao, J. J. De Voss, S. G.

Bell, *RSC Adv.* **2016**, *6*, 55286–55297; e) T. Coleman, R. R. Chao, J. B. Bruning, J. De Voss, S. G. Bell, *RSC Adv.* **2015**, *5*, 52007–52018.
[9] T. Coleman, S. H. Wong, M. N. Podgorski, J. B. Bruning, J. J. De Voss, S. G. Bell, *ACS Catal.* **2018**, *8*, 5915–5927.
[10] T. Coleman, A. M. Kirk, R. R. Chao, M. N. Podgorski, J. S. Harbort, L. R. Churchman, J. B. Bruning, P. V. Bernhardt, J. R. Harmer, E. H. Krenske, J. J. De Voss, S. G. Bell, *ACS Catal.* **2021**, *11*, 1995–2010.
[11] T. Coleman, J. E. Stok, M. N. Podgorski, J. B. Bruning, J. J. De Voss, S. G. Bell, *J. Biol. Inorg. Chem.* **2020**, *25*, 853–966.
[12] M. N. Podgorski, T. Coleman, R. R. Chao, J. J. De Voss, J. B. Bruning, S. G. Bell, *J. Inorg. Biochem.* **2020**, *203*, 110913.
[13] T. Coleman, A. M. Kirk, J. H. Z. Lee, D. Z. Doherty, J. B. Bruning, E. H. Krenske, J. J. De Voss, S. G. Bell, *ACS Catal.* **2022**, *12*, 1258–1267.
[14] a) T. Furuya, Y. Arai, K. Kino, *Appl. Environ. Microbiol.* **2012**, *78*, 6087–6094; b) T. Furuya, K. Kino, *Biosci. Biotechnol. Biochem.* **2009**, *73*, 2796–2799; c) T. Furuya, K. Kino, *Appl. Microbiol. Biotechnol.* **2010**, *85*, 1861–1868.
[15] a) X. Wu, Y. Chen, X. Wang, W. Wei, Y. Liang, *J. Org. Chem.* **2021**, *86*, 13768–13773; b) R. Leth, B. Ercig, L. Olsen, F. S. Jørgensen, *J. Chem. Inf. Model.* **2019**, *59*, 743–753.
[16] a) R. R. Chao, I. C. K. Lau, J. J. De Voss, S. G. Bell, *ChemCatChem* **2016**, *8*, 3626–3635; b) T. Coleman, R. R. Chao, J. De Voss, S. G. Bell, *Biochim. Biophys. Acta Proteins Proteomics* **2016**, *1864*, 667–675; c) R. R. Chao, I. C. Lau, T. Coleman, L. R. Churchman, S. A. Child, J. H. Z. Lee, J. B. Bruning, J. J. De Voss, S. G. Bell, *Chemistry* **2021**, *27*, 14765–14777.
[17] K. Haslinger, K. L. J. Prather, *Microb. Cell Fact.* **2020**, *19*, 26. https://doi.org/10.1186/s12934-020-01300-9.
[18] a) S. Shaik, P. Milko, P. Schyman, D. Usharani, H. Chen, *J. Chem. Theory Comput.* **2011**, *7*, 327–339; b) C. J. Whitehouse, N. H. Rees, S. G. Bell, L. L. Wong, *Chemistry* **2011**, *17*, 6862–6868; c) C. M. Bathelt, A. J. Mulholland, J. N. Harvey, *J. Phys. Chem. A* **2008**, *112*, 13149–13156; d) C. M. Bathelt, L. Ridder, A. J. Mulholland, J. N. Harvey, *Org. Biomol. Chem.* **2004**, *2*, 2998–3005; e) P. Rydberg, U. Ryde, L. Olsen, *J. Phys. Chem. A* **2008**, *112*, 13058–13065.
[19] N. K. Maddigan, S. G. Bell, *Arch. Biochem. Biophys.* **2017**, *615*, 15–21.
[20] J. W. Williams, J. F. Morrison, *Methods Enzymol.* **1979**, *63*, 437–467.
[21] a) D. Aragão, J. Aishima, H. Cherukuvada, R. Clarken, M. Clift, N. P. Cowieson, D. J. Ericsson, C. L. Gee, S. Macedo, N. Mudie, S. Panjikar, J. R. Price, A. Riboldi-Tunnicliffe, R. Rostan, R. Williamson, T. T. Caradoc-Davies, *J. Synchrotron Radiat.* **2018**, *25*, 885–891; b) N. P. Cowieson, D. Aragao, M. Clift, D. J. Ericsson, C. Gee, S. J. Harrop, N. Mudie, S. Panjikar, J. R. Price, A. Riboldi-Tunnicliffe, R. Williamson, T. Caradoc-Davies, *J. Synchrotron Radiat.* **2015**, *22*, 187–190.
[22] T. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, A. G. Leslie, *Acta Crystallogr. Sect. D* **2011**, *67*, 271–281.
[23] P. R. Evans, G. N. Murshudov, *Acta Crystallogr. Sect. D* **2013**, *67*, 1204–1214.
[24] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. Sect. D* **2011**, *67*, 235–242.
[25] P. Legrand, *XDSME: XDS Made Easier*, GitHub repository, **2017**, https://github.com/legrandp/xdsme, DOI 10.5281/zenodo.837885.
[26] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
[27] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D* **2010**, *66*, 486–501.
[28] P. D. Adams, G. B. P. V. Afonine, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. Sect. D* **2010**, *66*, 213–221.
[29] E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, T. E. Ferrin, *Protein Sci.* **2021**, *30*, 70–82.
[30] D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R. Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. J. Mermelstein, K. M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman in *AMBER 18*, *Vol.* University of California, San Francisco, **2018**.
[31] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox in *Gaussian 16 (Rev. B01)*, *Vol.* Gaussian, Inc., Wallingford CT, **2016**.

Manuscript received: June 20, 2022
Accepted manuscript online: August 31, 2022
Version of record online: ■■, ■■■

RESEARCH ARTICLE

Structural and biochemical data demonstrate that 4-phenylbenzoic acid binds within the active site of the cytochrome P450 monooxygenase CYP199A4, but no oxidation occurs. Protein engineering and molecular dynamic simulations establish that aromatic oxidation is enabled by a subtle shift of the substrate relative to the heme, thus improving our understanding of this important reaction.



Dr. T. Coleman, J. Z. H. Lee, A. M. Kirk, D. Z. Doherty, M. N. Podgorski, D. K. Pinidiya, Dr. J. B. Bruning, Prof. J. J. De Voss, Dr. E. H. Krenske, Dr. S. G. Bell**

1 – 14

Enabling Aromatic Hydroxylation in a Cytochrome P450 Monooxygenase Enzyme through Protein Engineering