ACES Asian Chemical Editorial Society

doi.org/10.1002/asia.202200986

CHEMISTRY AN ASIAN JOURNAL Research Article

www.chemasianj.org

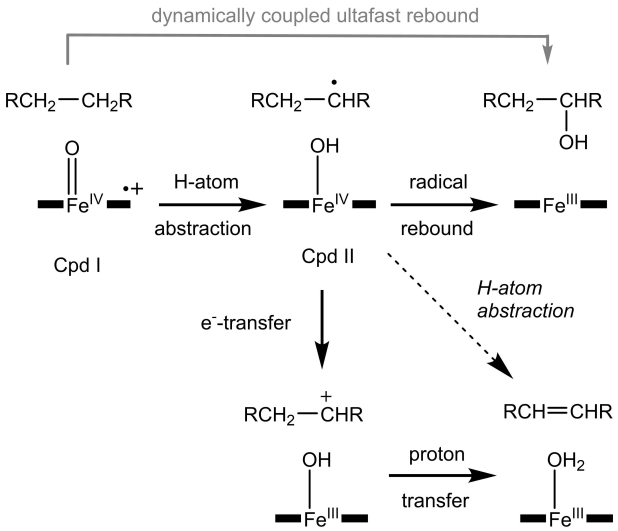
Exploring the Factors which Result in Cytochrome P450 Catalyzed Desaturation Versus Hydroxylation

Tom Coleman,[a] Daniel Z. Doherty,[a] Ting Zhang,[b] Matthew N. Podgorski,[a] Ruihong Qiao,[b] Joel H. Z. Lee,[a] John B. Bruning,[c] James J. De Voss,[d] Weihong Zhou,[b] and Stephen G. Bell*[a]

Abstract: The cytochrome P450 family of monooxygenase enzymes have essential biological roles involving the selective oxidation of carbon-hydrogen bonds. They can also catalyze other important metabolic reactions including desaturation to form alkenes. Currently the factors that control the partition between P450 hydroxylation and desaturation pathways are poorly defined. The CYP199A4 enzyme from the bacterium *Rhodopseudomonas palustris* HaA2 catalyzes the oxidation of 4-ethyl- and 4-isopropylbenzoic acids with hydroxylation and desaturation occurring in significant quantities. Here we demonstrate that 4-cyclopropylbenzoic acid is regioselectively hydroxylated by CYP199A4 at the benzylic carbon. In contrast, the oxidation of 4-*n*-propylbenzoic acid by CYP199A4 results in three major metabolites: an alkene from desaturation and two hydroxylation products at the benzylic (Cα) and Cβ carbons in similar quantities. Extending the length of the alkyl substituent resulted in 4-*n*-butylbenzoic acid being oxidized at the benzylic position (45%) and desaturated (55%). In contrast, 4-isobutylbenzoic generated very little alkene (5%) but was hydroxylated at the benzylic position (54%) and at the tertiary Cβ position (41%). The oxidation of 4-*n*-propylbenzoic acid by the F298 V mutant of CYP199A4 occurred with no hydroxylation at Cβ and a significant increase in metabolites arising from desaturation (73%). The X-ray crystal structures of CYP199A4 with each substrate revealed that they bind in the active site with the alkyl substituent positioned over the heme. However, the longer alkylbenzoic acids were bound in a different conformation as was 4-*n*-propylbenzoic acid in the F298 V mutant. Overall, the changes in metabolite distribution could be ascribed to bond strength differences and the position of the alkyl group relative to the heme.

Introduction

In nature, the cytochrome P450 (CYP) family of heme monooxygenase enzymes catalyzes the insertion of an oxygen atom into unactivated C–H bonds.[1] The identity of the active intermediate, the so-called compound I (Cpd I), in the majority of reactions catalyzed by these enzymes is widely accepted to be a ferryl-oxo heme radical cation.[1a,2] Hydroxylation is proposed to occur via the radical rebound mechanism (Scheme 1).[3] Cpd I abstracts a hydrogen atom from the substrate, resulting in the formation of a Fe(IV)–OH species



Scheme 1. The product formation pathways involved in aliphatic hydroxylation and desaturation. The solid arrows indicate the formation of hydroxylation product via radical rebound, and the desaturation product via a cationic intermediate; the dashed arrow indicates an alternative mechanism whereby a second hydrogen atom is abstracted from the radical species.

[a] *Dr. T. Coleman, D. Z. Doherty, M. N. Podgorski, J. H. Z. Lee, Dr. S. G. Bell*
Department of Chemistry
University of Adelaide
Adelaide, SA, 5005 (Australia)
E-mail: stephen.bell@adelaide.edu.au
[b] *T. Zhang, R. Qiao, Prof. W. Zhou*
College of Life Sciences
and The State Key Laboratory of Medicinal Chemical Biology
Nankai University
Tianjin 300071 (P. R. China)
[c] *Dr. J. B. Bruning*
School of Biological Sciences
University of Adelaide
Adelaide, SA, 5005 (Australia)
[d] *Prof. J. J. De Voss*
School of Chemistry and Molecular Biosciences
University of Queensland
Brisbane, QLD, 4072 (Australia)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/asia.202200986

(Compound II, Cpd II) and a substrate radical.[2b,3–4] The latter then recombines with a hydroxyl radical of Compound II to yield the alcohol product.[4] Under ideal reaction conditions C–H abstraction and C–O bond formation can merge into a dynamically coupled process.[5] When the substrate is not optimally orientated within the active site relative to Cpd I uncoupling pathways can result in unproductive turnover of the catalytic cycle resulting in hydrogen peroxide or water (oxidase activity) formation.[1b,6]

P450-catalyzed desaturation (dehydrogenation) reactions are found in both catabolic and anabolic metabolic processes.[7] Desaturation is associated with the generation of toxic drug metabolites, for example, the desaturation of valproic acid as well as in vitamin and steroid biosynthesis.[8] The desaturation reaction is unusual for one catalyzed by a monooxygenase enzyme, in that there is no insertion of an oxygen atom into the substrate. Overall, the dioxygen molecule is converted into two water molecules with the corresponding generation of the alkene. The P450 catalyzed desaturation of alkyl-substituted compounds is commonly observed alongside the hydroxylation product(s).[8c,9] The first step of the desaturation pathway is proposed to proceed with the same hydrogen abstraction step by Cpd I to yield a carbon radical.[1a] If this step is dynamically coupled to oxygen rebound or the oxygen rebound step is highly efficient (fast), the alcohol hydroxylation product is formed (Scheme 1). Partition between hydroxylation and desaturation reactions would occur at this step. In the case of alkene formation, a single electron transfer from the radical to Cpd II to generate a carbocation and a ferric-hydroxide species could occur (Scheme 1, solid arrows). Subsequent proton transfer then results in alkene formation. An alternative pathway could involve a second hydrogen abstraction by Cpd II to generate the desaturated product (Scheme 1, dashed arrow).[1a] Both DFT and QM/MM calculations agree that desaturation arises from Cpd I rather than another potential oxidant and favor the role of a cationic intermediate.[10]

CYP199A4, from *Rhodopseudomonas palustris* strain HaA2 has high affinity for *para*-substituted benzoic acids. Furthermore, the oxidation activity of the enzyme is usually restricted to the *para*-substituent.[9a,c,11] For example, it is able to demethylate 4-methoxy- and 4-methylamino-benzoic acids.[9c,11a,c,12] *para*-Substituted alkyl benzoic acids, such as 4-ethyl- and 4-isopropyl-benzoic acids, are hydroxylated and desaturated by CYP199A4 to form alcohol and alkene products.[9c] The oxidation activity of CYP199A4 for alkylbenzoic acids is high and the loss of reducing equivalents due to uncoupling side reactions is low.[9c] Crystal structures of several substrate-bound forms of the CYP199A4 and the related CYP199A2 enzymes have been solved.[9b,c,11b,12] The combination of high activity and selectivity of the CYP199A4 enzyme, significant levels of competing alkene and alcohol formation, and structural information from substrate-bound crystal structures make it a suitable P450 enzyme for studies on the partition of the desaturation and aliphatic hydroxylation activities.

CYP199A4 catalyzed oxidation of 4-ethyl- and 4-isopropyl-benzoic acids generates a major hydroxylation product at the α-carbon (Figure 1).[9c] This would be expected from hydrogen abstraction at the benzylic position, followed by radical rebound.[13] There is a significant amount of an alkene by-product, which could have arisen from initial abstraction at either the benzylic α-carbon or the primary methyl β-carbon. Abstraction at the benzylic α-carbon would be energetically favored over that at the less substituted methyl group. However, the crystal structure of CYP199A4 bound with 4-ethylbenzoic acid (PDB: 4EGM) showed that the β-carbon is significantly closer to the heme iron than Cα.[9c] 4-*t*-Butylbenzoic acid oxidation by CYP199A4 occurs at the primary methyl groups of the β-carbon (Figure 1). This demonstrated that for substrates, in which there are no competing reactions due to the absence of C–H bonds on the benzylic α-carbon, that efficient abstraction at a primary methyl was possible by this enzyme system.[14]

To understand the parameters that govern the regioselectivity of the CYP199A4 enzyme, and the partition of hydroxylation versus desaturation in cytochrome P450 enzymes we assayed related alkyl substituted benzoic acids. 4-*n*-Propylbenzoic acid, 4-*n*-butylbenzoic acid and 4-*n*-heptylbenzoic acid were selected to determine the impact of more extended alkyl side chains that contains a methylene group at the β-position (Figure 1). This could undergo C–H bond abstraction more readily than the primary methyl moiety which are present in the substrates studied previously.[15] The oxidation of 4-isobutylbenzoic acid was investigated to assess the effect of increasing the branching of the alkyl substituent at the β-carbon (Figure 1). 4-Cyclopropylbenzoic acid was also employed to assess if CYP199A4 could bind and oxidize strained cyclic alkyl groups and what impact this would have on the product distribution (Figure 1). Cyclopropyl groups have strong C–H

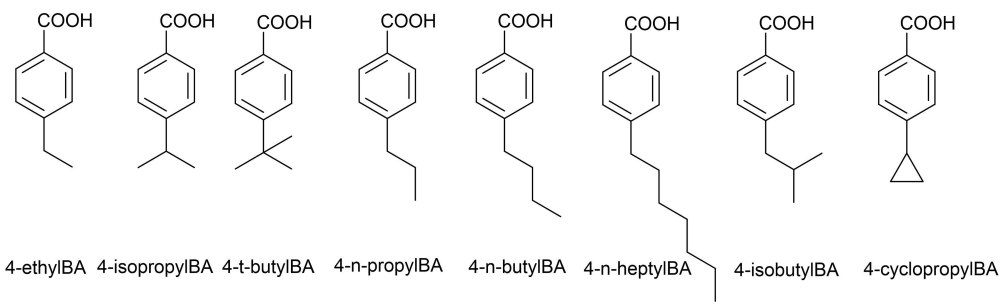


Figure 1. Compounds investigated with CYP199A4 during this and previous related work. BA; benzoic acid.

bonds that would be expected to influence the energetics of hydroxylation.[15–16] We also investigate if the partition between desaturation and hydroxylation of 4-*n*-propylbenzoic can be altered by changing the substrate's binding position in the active site using a mutant of CYP199A4.

Results

Substrate binding

UV-visible spectroscopy has been used to demonstrate that the binding of 4-ethyl-, 4-isopropyl- and 4-*t*-butyl-benzoic acids induce a $\geq$95% shift to the high-spin form of ferric CYP199A4.[9c,14,17] While the first two bind with high affinity ($<$ 0.5 μM, Table 1) 4-*t*-butylbenzoic acid binds with a dissociation constant of 39 $\pm$ 2 μM. 4-Cyclopropyl- and 4-*n*-propyl-benzoic acids generated a similar magnitude of a switch in the spin state upon addition to CYP199A4 (both $\geq$95%, Table 1, Figure S1). The binding affinity of CYP199A4 for both these alkylbenzoic acids was $\leq$0.54 μM (Table 1, Figure S2). The three longer chain alkylbenzoic acids tested here generated similarly high shifts in the spin state upon addition to CYP199A4 (all $\geq$90%, Table 1, Figure S1). The binding affinity of CYP199A4 for all of these alkylbenzoic acids was $<$0.6 μM (Table 1, Figure S2). The substrate with the linear alkyl substituents bound to CYP199A4 with similar affinity to the branched species. Overall, this data highlights that these alkyl substrates were a good fit for the active site of CYP199A4 and displace the distal water ligand. This results in a switch from the 6-coordinate low-spin ferric state to the 5-coordinate high spin form, priming the enzyme for the reduction and oxygen binding steps of the catalytic cycle. That the enzyme is able to bind the larger 4-*n*-heptyl substituent demonstrates that ligands with longer subsituents are good substrates for the CYP199A4 enzyme.

Substrate oxidation and product distributions

First, we demonstrated that 4-ethylbenzoic acid and 4-isopropylbenzoic acid were hydroxylated and desaturated by CYP199A4 with similar rates and product ratios to those reported previously (Table 1, Scheme 2 and Figure S3a and S3b).[9c,12,17] These results highlighted that abstraction of a hydrogen from Cα is strongly favored over that from Cβ. 4-Cyclopropylbenzoic acid and 4-*n*-propylbenzoic acid were chosen as they contain methylene groups at the Cβ positions that would enable further comparisons to be made.

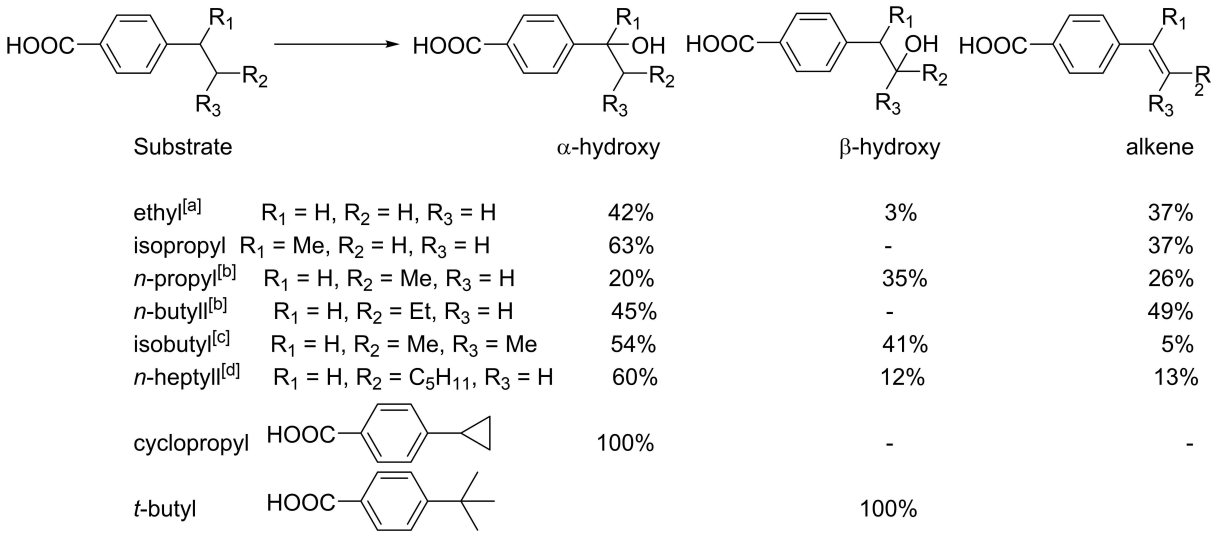
CYP199A4 catalyzed oxidation of 4-*n*-propylbenzoic acid displayed the highest product formation activity, 594 nmol.nmol-CYP^{-1}.min^{-1} (henceforth abbreviated to min^{-1}), of the alkylbenzoic acids tested (Table 1). The NADH oxidation rate was lower than those of 4-ethylbenzoic acid but the coupling efficiency was greater (86%), which resulted in a higher product formation rate (Table 1). Three major oxidation products were observed by HPLC analysis of the reaction mixture, and GC-MS analysis of the extracted and trimethylsilyl (TMS) derivatized reaction samples (Figure S3c). The mass spectra and retention times were consistent with one species arising from desaturation and two hydroxylation products (Figure S4). These were isolated (in certain instances as the methyl ester derivatized species) and characterized by NMR (Figure S5). The products were identified as arising from Cα hydroxylation, 4-(1-hydroxypropyl)benzoic acid (20%); desaturation between Cα and Cβ to yield the *trans* alkene, (*E*)-4-(prop-1-en-1-yl)benzoic acid (26%; See Figure S5) and Cβ hydroxylation, 4-(2-hydroxypropyl)benzoic acid (35%). Two other minor metabolites were detected which had retention times and mass spectra consistent with epoxidation and Cγ hydroxylation of the initial alkene desaturation product (Scheme 2). Compared to 4-ethylbenzoic acid and 4-isopropylbenzoic acid there was a greater proportion of β-hydroxylation product from the methylene carbon of 4-*n*-propylbenzoic acid than the methyl β-carbon of the isopropyl species. The proportion of the metabolites, which arises from desaturation activity, was similar to those measured with 4-isopropyl- and 4-ethyl-benzoic acids (34–37%, Scheme 2).

The hydroxylated metabolites of *n*-propylbenzoic acid could be formed as enantiomers and we analyzed the enzyme-catalyzed oxidations by enantioselective HPLC (Figures S3d and S3e). The enantiomers of 4-(2-hydroxypropyl)benzoic acid were not separated but the enantioselective HPLC analysis demon-

Table 1. Substrate binding and *in vitro* enzyme catalyzed oxidation data for CYP199A4.

Substrate[a]	HS (%)	K_d (μM)	NADH[b]	PFR	Coupling[c] [%]
4-*n*-propylBA	$\geq$95	0.54 $\pm$ 0.02	688 $\pm$ 24	594 $\pm$ 72	86 $\pm$ 8
4-cyclopropylBA	$\geq$95	0.28 $\pm$ 0.02	383 $\pm$ 8	170 $\pm$ 22	45 $\pm$ 5
4-*n*-butylBA	90	0.39 $\pm$ 0.04	167 $\pm$ 6	67 $\pm$ 3	40 $\pm$ 1
4-isobutylBA	$>$95	0.59 $\pm$ 0.03	359 $\pm$ 46	146 $\pm$ 21	41 $\pm$ 5
4-*n*-heptylBA	$\geq$95	0.16 $\pm$ 0.03	278 $\pm$ 15	147 $\pm$ 12	53 $\pm$ 2
4-ethylBA[d]	$\geq$95	0.34 $\pm$ 0.02	812 $\pm$ 7	525 $\pm$ 88	64 $\pm$ 4
4-isopropylBA[d]	$\geq$95	0.29 $\pm$ 0.01	325 $\pm$ 36	221 $\pm$ 27	68 $\pm$ 1
4-*t*-butylBA[d]	$\geq$95	39 $\pm$ 2	227 $\pm$ 4	227 $\pm$ 32	100 $\pm$ 13

[a] Abbreviations: BA, benzoic acid; HS is the change in the heme spin state from low-spin to high-spin upon addition of substrate; NADH is the frequency of NADH oxidation, including product formation and uncoupling reactions. PFR is the product formation rate; coupling is the percentage of NADH utilized for the formation of products. The enzyme-catalyzed oxidations were measured using a HaPuR:HaPux:CYP199A4 concentration ratio of 1:10:1 (0.5 μM CYP enzyme, 50 mM Tris, and pH 7.4). Rates are reported as mean $\pm$S.D. ($n \geq 3$) and given in nmol.nmol-CYP^{-1}.min^{-1}. [b] The NADH leak rate was measured to be ~9.0 min^{-1}. [c] The oxidations of most of these substrates showed low levels of hydrogen peroxide uncoupling ($\leq$2%), as measured by the HRP assay (see experimental). [d] Reported previously.[14]



Scheme 2. Product distributions arising from oxidation of 4-alkylbenzoic acids by CYP199A4. A dash (–) indicates no product formed. Empty cells indicate reactions which cannot occur. [a] small amounts of the epoxide and ketone further oxidation products were also detected (~18% of total products: 17% epoxide, 1% ketone) [b] The remainder of the product distribution was an alcohol arising from oxidation of γ C–H bonds of the alkene 7% and 4%, respectively (for 4-*n*-propylBA and 4-*n*-butylBA). A further oxidation product, consistent with epoxide formation was also detected 12% and 2% respectively. [c] a small amount of an unidentified further oxidation product was detected in these turnovers (<1%). [d] the remainder of the metabolites were: 3% epoxide and 12% α,β-diol.

strated that 4-(1-hydroxypropyl)benzoic acid was not formed with a significant excess of one enantiomer over the other (Figure S3e).

The NADH oxidation and product formation rates for the oxidation of 4-cyclopropylbenzoic acid were comparable to 4-isopropylbenzoic acid (383 and 170 min^{-1}, respectively; Table 1). The coupling efficiency was the lowest among the propyl derivatives (Table 1). HPLC analysis indicated a single product was generated (Figure S3f). The GC-MS of the TMS-derivatized extracted reaction mixture also contained a single product with an MS consistent with an alcohol (m/z = 322.00 versus expected 322.14, Figure S4). This metabolite was purified by semi-prep HPLC and characterized by NMR as 4-(1-hydroxycyclopropyl)benzoic acid (Scheme 2 and Figure S5). There was no evidence for the formation of other minor metabolites via hydroxylation at the methylene β-carbons of the cyclopropyl ring, or an alkene from desaturation.

CYP199A4 oxidized 4-*n*-butyl- and 4-isobutyl-benzoic acid at slower product formation rates (67 and 146 min^{-1}, respectively) than their isomer 4-*t*-butylbenzoic acid or any of the shorter alkyl chain substituted benzoic acids (Table 1). The NADH oxidation rate was slower for the linear *n*-butyl substrate (Table 1). HPLC analysis of the *in vitro* turnovers of 4-*n*-butylbenzoic acid exhibited evidence of three metabolites with one potential additional minor product (Scheme 2 and Figure S3g). Of these the three major products were isolated by scaling up the reaction using a whole-cell turnover. They were separated by semi-prep HPLC, isolated and characterized using NMR (Figure S5). The most abundant product was assigned as the alkene (*E*)-4-(but-1-en-1-yl)benzoic acid (50%). The alkene stereochemistry was assigned by comparison of the NMR data to literature for alkylbenzoic acids. In these studies, the vinylic ^{1}H signals of (*E*)-alkenes tended to overlap. In contrast, (*Z*)-alkene signals were typically able to be resolved.[18] The second most abundant species was 4-(1-hydroxybutyl)benzoic acid (46%, Scheme 2). The HPLC and GC retention times and the MS of the TMS-derivatized forms of both were consistent with these assignments (Figure S4). The third metabolite, which constituted 4% of the total products formed, did not arise from hydroxylation at Cβ but could be assigned from the NMR as the further oxidation product (*E*)-4-(3-hydroxybut-1-en-1-yl)benzoic acid. The alkene stereochemistry was assigned based on the coupling constant observed for the vinylic protons of 4-(3-hydroxybut-1-en-1-yl)benzoic acid (15.9 Hz). The mass spectrum and fragmentation pattern of the TMS-derivatized products also supported these assignments (Figure S4). This double oxidation product could arise from initial formation of the alkene followed by hydroxylation at Cγ. The fourth, unidentified minor product (<2%), was not generated in sufficient yield for further characterization but had a retention time consistent with that of a further oxidation product (epoxide or ketone) rather than a β-hydroxylation metabolite. To confirm this assignment we isolated the alkene metabolite using prep-scale HPLC. This alkene bound to CYP199A4 (35% HS, Figure S6a). The oxidation by CYP199A4 generated two products one of which coeluted with (*E*)-4-(3-hydroxybut-1-en-1-yl)benzoic acid and the other with the unidentified minor metabolite (Figure S6b). This infers that (*E*)-4-(3-hydroxybut-1-en-1-yl)benzoic acid in the oxidation of 4-*n*-butylbenzoic acid by CYP199A4 is formed from further oxidation of the alkene and that the other metabolite is most likely to be the epoxide.

Three metabolites were observed in the HPLC analysis of the CYP199A4 turnovers of 4-isobutylbenzoic acid (Scheme 2, Figure S3h). However, the product distribution was significantly different to that of the linear isomer. Two major products were formed in similar quantities (Scheme 2). HPLC purification and

NMR analysis enabled their identification as 4-(2-hydroxy-2-methylpropyl)benzoic acid (41%) and 4-(1-hydroxy-2-methylpropyl)benzoic acid (54%). The third product had an HPLC retention time and NMR spectrum (Figure S5h) consistent with that of an alkene desaturation product, 4-(2-methylprop-1-en-1-yl)benzoic acid (5%). The masses observed in the GC-MS chromatograms of the TMS-derivatized turnover supported the product assignments of the alkene and the Cα hydroxylation product (Figure S4). The major product, 4-(2-hydroxy-2-methylpropyl)benzoic acid, which arises from hydroxylation at the tertiary Cβ, had an MS fragmentation pattern consistent with a singly derivatized species suggesting that this alcohol may be more difficult to derivatize due to the increased steric bulk at this position.

The long chain alkyl substrate, 4-*n*-heptylbenzoic acid, was also oxidized efficiently by CYP199A4. As a result of a high coupling efficiency (53%), the product formation rate of 147 min^{-1} exceeded that of 4-*n*-butylbenzoic acid. However, the catalytic activity was lower than both those of 4-ethyl- and 4-*n*-propyl-benzoic acids (Table 1). HPLC analysis of the *in vitro* and whole-cell turnovers revealed a complex mixture of product peaks, five of which could be isolated by semi-prep HPLC for characterization (Scheme 2, Figure S3i). The major metabolites could be assigned by NMR (Figure S5) and GC-MS (of the TMS derivatized forms, Figure S4). These arose from Cα hydroxylation, 4-(1-hydroxyheptyl)benzoic acid (60%); Cβ hydroxylation, 4-(2-hydroxyheptyl)benzoic acid (12%) and desaturation across the Cα and Cβ positions, (*E*)-4-(hept-1-en-1-yl)benzoic acid (13%). The two other minor metabolites were characterized by a combination of NMR and GC-MS as double oxidation products: 4-(1,2-dihydroxyheptyl)benzoic acid (12%) and 4-(3-pentyloxiran-2-yl)benzoic acid (3%). These were both formed in greater quantities in the whole-cell oxidation system with an accompanying decrease in the amount of the alkene metabolite. The epoxide most likely arises from the P450 catalyzed epoxidation of the alkene. The diol could originate from a second hydroxylation of either of the Cα or Cβ alcohol products.

Crystal structures of substrate bound forms of CYP199A4

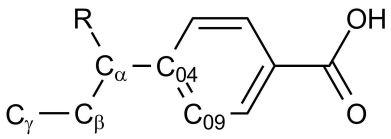
The CYP199A4 catalyzed oxidation of the alkylbenzoic acids all generated products that arise from initial attack at hydrogen atoms on either Cα or Cβ. Despite the different sizes of the *para*-substituents, the binding affinity of most of the benzoic acids tested was high. To help rationalize both observations, we obtained crystal structures of CYP199A4 bound with the benzoic acid substrates. Previously we have determined the crystal structure of CYP199A4 bound with 4-ethylbenzoic acid (PDB: 4EGM). This structure established that the substrate methyl group at Cβ was significantly closer to the heme iron than the benzylic Cα. In this work we have been able to obtain crystals and structures of CYP199A4 in the presence of 4-*n*-propyl- (under two different conditions), 4-cyclopropyl-, 4-*n*-butyl-, 4-isobutyl-, and 4-*n*-heptyl-benzoic acids (Table S1 and S2, PDBs: 5YQA, 7R8S, 5UVB, 8E5J, 6VJX and 6C2D).

In all the other substrate-bound CYP199A4 complexes in this work, the overall protein fold was very similar to the 4-ethylbenzoic acid-bound CYP199A4 structure (Figure S7) and

Table 2. Distances and angles between the substrate and the heme iron and internal angles the substrate molecule in the different alkylbenzoic acid-bound structures of CYP199A4.

Distance (Å)	4-ethyl BA[a]	4-*n*-propyl BA[b]	4-cyclopropyl BA[c]	4-*n*-butyl BA	4-isobutyl BA	4-*n*-heptyl BA
Cα–Fe	4.7	5.1	4.9	4.6	4.9	4.5
Cα–**O**=Fe	3.5	3.7	3.4	3.1	3.4	3.1
Cβ–Fe	3.2	3.8	3.7 (β′ 4.2)	5.3	5.5	4.7
Cβ–**O**=Fe	2.0	2.3	2.1 (β′ 2.3)	4.0	4.2	3.5
Cγ–Fe	–	3.9	–	6.8	5.3	6.2
Cγ–**O**=Fe	–	2.6	–	5.4	4.2	4.9
C_x–F298	X = α, 3.9	X = α, 3.6	X = β′, 3.4	X = δ, 3.6	X = γ, 4.3	X = ω-1, 2.9
Angle (°)						
C_{04}–C_α–C_β	113.7	112.6	126.4 (126.5)	111.0	112.4	113.3
Dihedral [d]	50.2	68.4	54.8 (22.9)	74.1	73.7	64.0
C_α–C_β–C_γ	–	108.8	–	112.2	112.4	110.7
Fe=O–C_α	161.6	156.2	155.5	153.1	151.0	158.7
Fe=O–C_β	162.1	165.6	157.7 (131.0)	135.2	133.6	136.3
Fe=O–C_γ	–	138.8	–	143.0	125.3	139.3

[a] The structure of 4-ethylbenzoic acid was reported previously (PDB: 4EGM).[9c] [b] The distances measured in the two *n*-propylbenzoic acid structures reported here are also in broad agreement (Table S3). [c] For 4-cyclopropylbenzoic acid, which contains two β carbons, these are designated β and β′. [d] The dihedral angle, C_{09}–C_{04}–C_α–C_β, is represented below.



there was electron density in the active site that was readily modelled by the appropriate substrate molecule. The carboxylate and benzene moieties of each substrate were found in comparable positions to that of 4-ethylbenzoic acid. The location of the active site amino acids, the capping chloride anion and active site water molecule were in similar positions, unless otherwise stated (Figure S8 and S9). In the Cpd I intermediate the position of the oxo group would be expected to be 1.62 Å above the heme plane.[19] This oxygen atom was modelled into our structures of substrate-bound P450 enzymes, in order to compare the positions of the carbons of the alkyl moieties relative to the heme (Table 2).

The structure of the CYP199A4 enzyme bound with 4-*n*-propylbenzoic acid was solved using two sets of conditions (PDBs: 5YQA and 7R8S). Both structures resulted in a similar binding orientation of the substrate within the active site of the CYP199A4 enzyme (Figure 3, Figure S8 and Table S3). In both 4-*n*-propylbenzoic acid-bound CYP199A4 crystal structures, the position of the alkyl moiety was comparable to that observed for 4-ethylbenzoic acid (Figure 2). The benzylic α-carbon of 4-*n*-propylbenzoic acid was found slightly further from the heme iron than that of 4-ethylbenzoic acid (5.1 Å versus 4.7 Å), as was the β-carbon (3.8 Å versus 3.2 Å). The γ-carbon points towards the valine 295 residue (Figure 2) and was 3.9 Å from the iron center. If, when the Cpd I intermediate is formed, this substrate remained bound in the same location, the β-carbon would be even closer to the oxo group than Cγ. The angles of the alkyl chain carbons to the plane of the Fe=O bond were 156.2 (Cα), 165.6 (Cβ) and 138.8 (Cγ). The C–H bonds of Cγ are directed away from the Cpd I intermediate which would place them in a

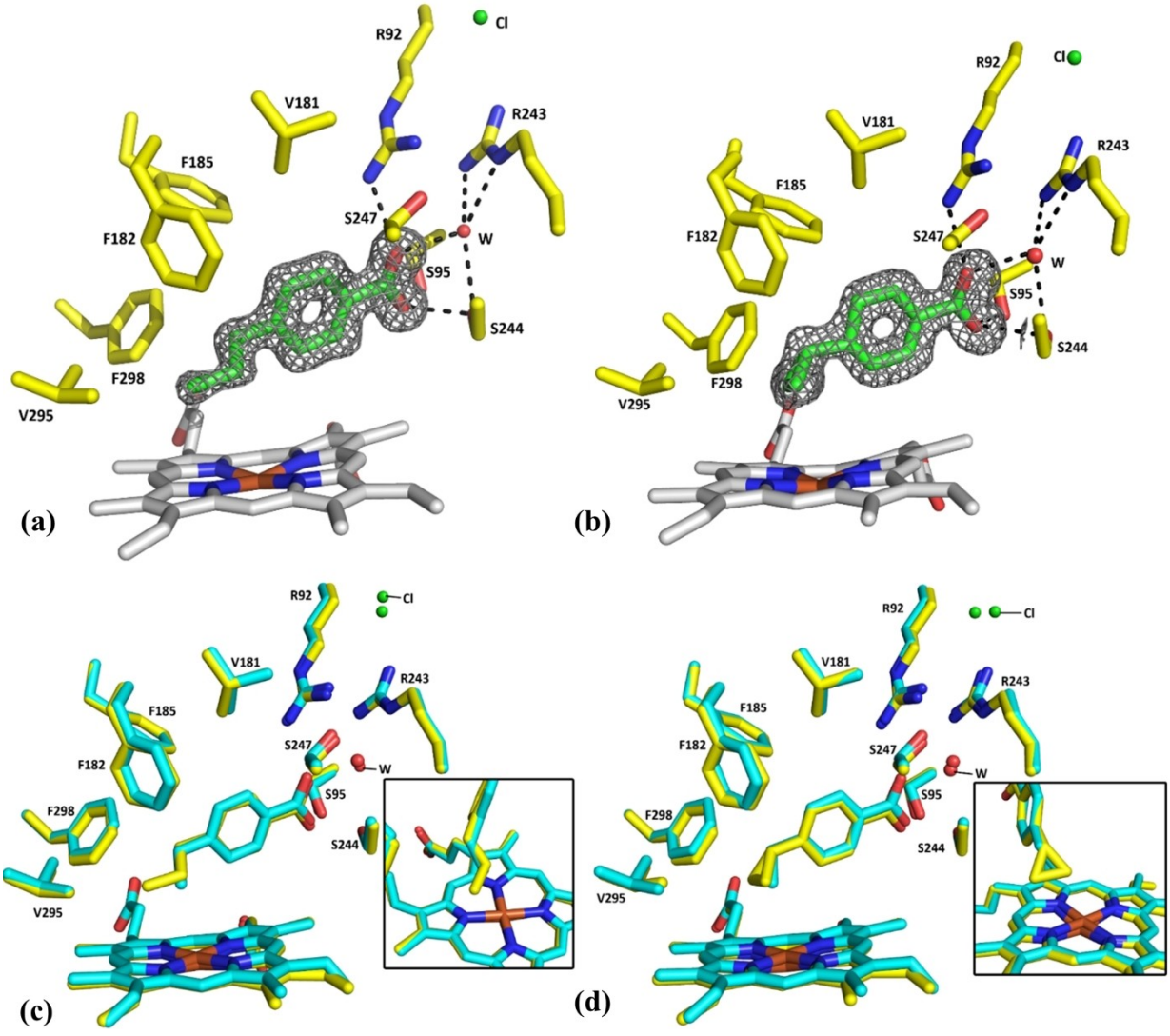


Figure 2. Composite omit maps of (a) the 4-*n*-propylbenzoic acid-CYP199A4 and (b) 4-cyclopropylbenzoic acid-CYP199A4 complexes (PDB: 5YQA and 5UVB).The heme is grey with the iron center in brown, the substrate green, and the red and green spheres are a water molecule and chloride ligand, respectively. The electron density of the substrate is shown contoured as a grey mesh (σ = 1.2) using a reduced bias *2mFo-DFc* composite omit map. (c) Comparison of the active site and substrate binding orientation of 4-*n*-propylbenzoic acid-bound CYP199A4 (yellow) with that of 4-ethylbenzoic acid structure (cyan, PDB: 4EGM). (d) Comparison of the active site and substrate binding orientation of 4-cyclopropylbenzoic acid-bound CYP199A4 (yellow) with that of 4-ethylbenzoic acid structure (magenta, PDB: 5EGM). The labels (W, Cl) point to those located in the 4-*n*-propylbenzoic acid structure. The insets in (c) and (d) show the overlay of the 4-*n*-propyl and the 4-cyclopropylbenzoic acid structures with that of 4-ethylbenzoic from above.

poor position for abstraction.[14] These are important considerations and indicate that the β-carbon is in the best position for C–H bond abstraction by Cpd I (Table 2).

In the crystal structure of CYP199A4 with 4-cyclopropylbenzoic acid the location of the substrate in the active site was comparable with that of 4-ethylbenzoic acid (Figure 2). The α-carbon and the closest β-carbon of the cyclopropyl group are in comparable positions to the equivalent carbons of the ethyl group. The α- and β-carbons of the substrate were 4.9 Å and 3.7 Å, respectively, from the iron of the heme. The other β-carbon of 4-cyclopropylbenzoic acid, which is held further from the heme (designated β′) is at 4.2 Å (Table 2). The resemblant nature of the substrate binding positions in these crystal structures should enable differences in reactivity to be assigned according to the properties of the substrates, such as the strength of C–H bonds of a cyclopropyl ring and other properties such as ring strain.

The *n*-butyl, isobutyl and *n*-heptyl *para*-substituents in their respective structures with CYP199A4 have a significantly different orientation in the active site compared to the ethyl or *n*-propyl moieties of the other alkylbenzoic acids described above (Figures 3 and 4). In these structures the longer alkyl chains are orientated to point away from the heme-iron and towards the F182 and F298 residues (Figures 3 and 4). In addition, the F298 residue changes its orientation relative to the *n*-propylbenzoic acid-bound CYP199A4 crystal structures. The F298 residue undergoes a rotation so that it points down towards the heme instead of into the substrate binding pocket (Figures 3 and 4). A similar shift in the position of the F298 residue has previously been observed in the 4-ethylthio-, 4-cyclohexyl- and 4-phenyl-

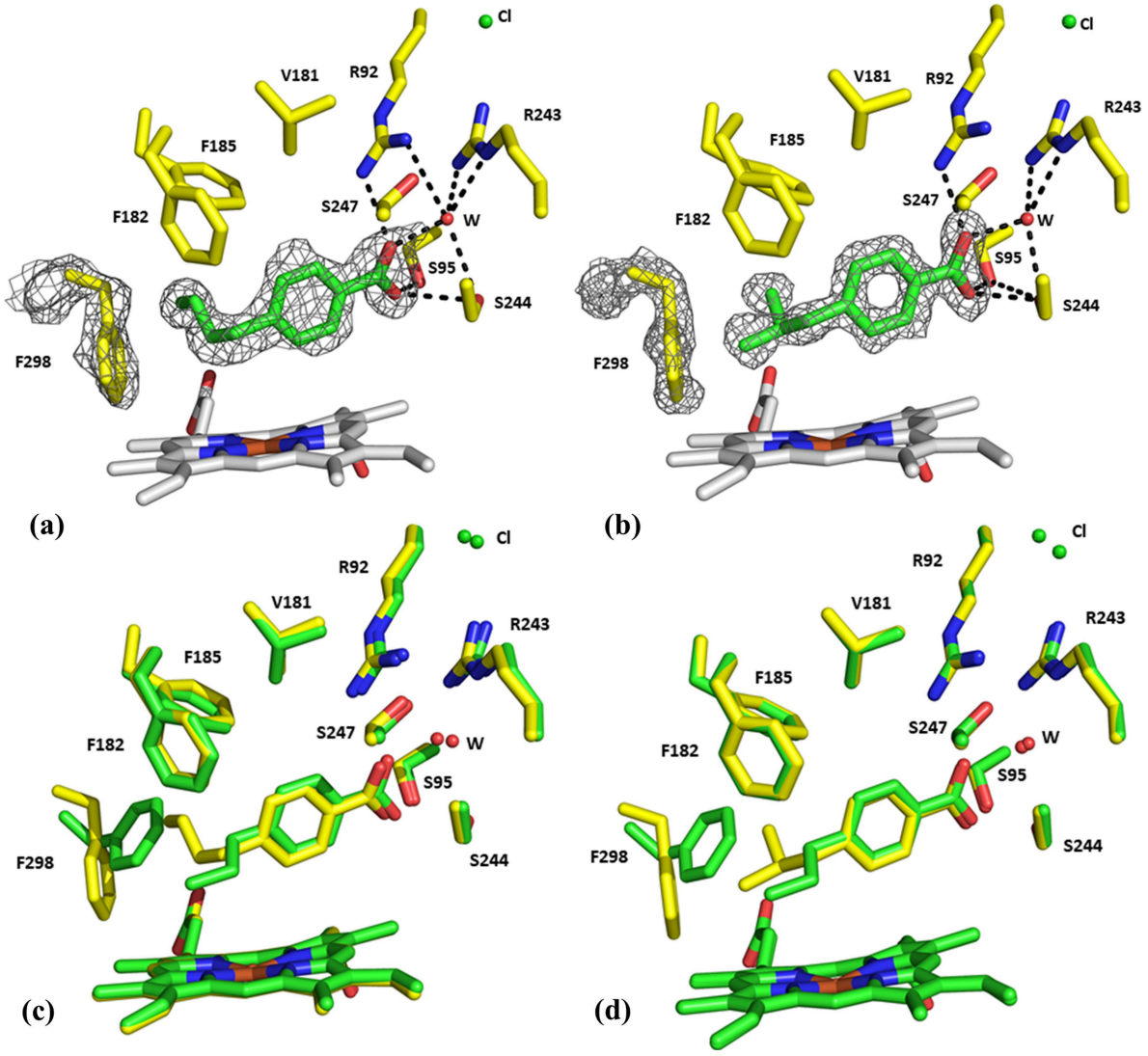


Figure 3. Feature-enhanced maps of the active site of (a) 4-*n*-butyl- and (b) 4-isobutyl-benzoic acid-bound structures of CYP199A4 (PDB 8E5J and 6VJX). The color scheme and electron densities are presented as per Figure 2. (c) Comparison of the active sites of the 4-*n*-propylbenzoic acid and 4-*n*-butylbenzoic acid bound CYP199A4 crystal structures. The 4-*n*-propylbenzoic acid structure (green) is overlaid with the 4-*n*-butylbenzoic acid structure (yellow). (d) Comparison of the active sites of the 4-*n*-propylbenzoic acid and 4-isobutylbenzoic acid bound CYP199A4 crystal structures. The 4-*n*-propylbenzoic acid structure (green) is overlaid with the 4-isobutylbenzoic acid structure (yellow). A view of the substrates and heme centre from above are included in the SI (Figure S9).

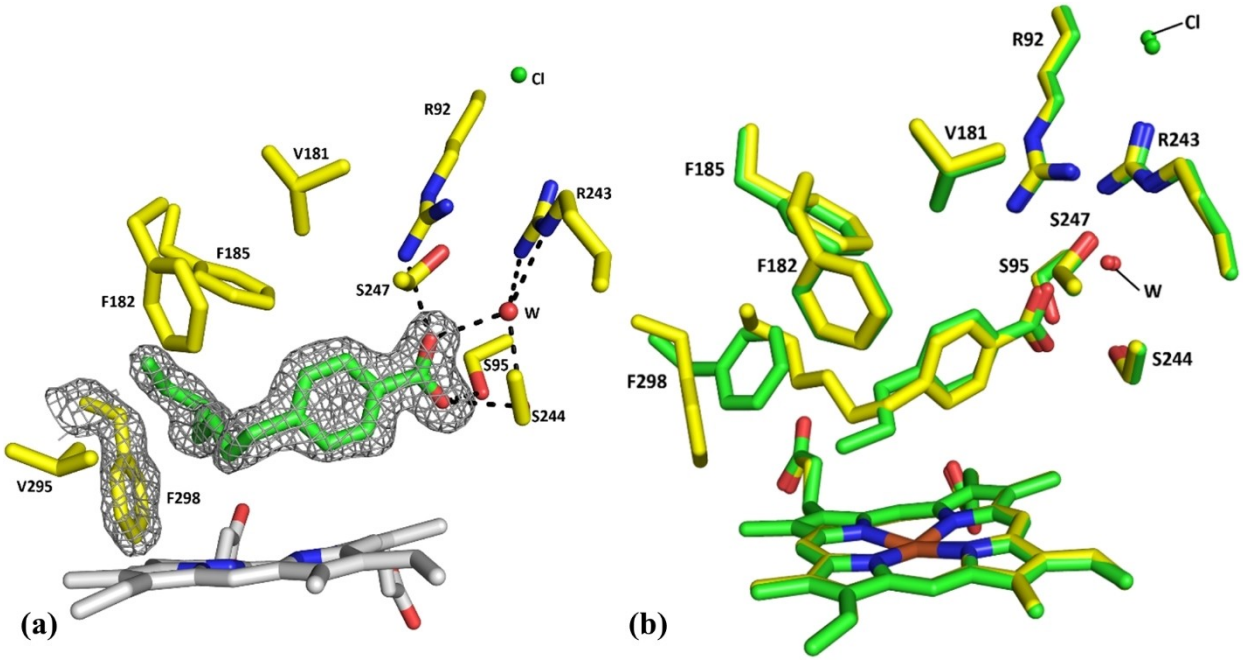


Figure 4. (a) Composite omit maps of the 4-*n*-heptylbenzoic acid-CYP199A4 (PDB: 6C3H). The color scheme and electron density are presented as per Figure 2. (b) Comparison of the active site and substrate binding orientation of 4-*n*-heptylbenzoic acid-bound CYP199A4 (yellow) with the equivalent 4-*n*-propylbenzoic acid structure (green, PDB: 5YQA The superimposition with the 4-*n*-butylbenzoic acid and 4-isobutylbenzoic acid structures are provided in the supporting information (Figure S10).

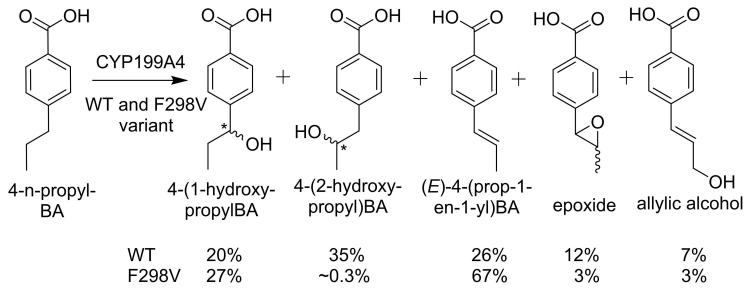
benzoic acid-bound structures of CYP199A4 (PDB: 5U6U, 6C2D and 7JW5, respectively).[13,20]

In all three structures the α-carbon atom was held closer to the heme iron than Cβ (4.5–4.9 Å vs. 4.7–5.5 Å, Table 2). For *n*-heptyl- and *n*-butyl-benzoic acids Cα was closer than the equivalent benzylic carbon of either 4-ethyl-, 4-cyclopropyl and 4-*n*-propyl-benzoic acids (4.5 and 4.6 Å vs 4.7, 4.9 and 5.1 Å, respectively, Table 2). The different orientation of the side-chain results in the γ-carbons being significantly further away from the heme iron compared to 4-*n*-propylbenzoic acid (5.3–6.8 Å versus 3.9 Å, Table 2). These structural features enable the selectivity of the enzyme for C–H bond abstraction at the α- and β-carbons to be rationalized.

Engineering CYP199A4 to enhance desaturation

We hypothesized that it should be possible to alter the partition of desaturation and hydroxylation by favoring a different binding orientation of the alkyl substituent within the active site. In the crystal structures the longer *n*-butyl alkyl substituent occupies a different orientation compared to that of *n*-propylbenzoic acid. This is accompanied by movement of the phenylalanine 298 residue into a different position. Based on the above crystal structures of CYP199A4 with the alkyl substituted benzoic acids we propose that mutating the bulky phenylalanine residue to a smaller valine residue could enable the substituent of *n*-propylbenzoic acid to occupy a similar position to that observed in the *n*-butylbenzoic acid CYP199A4 structure. As CYP199A4 catalyzed oxidation of 4-*n*-butylbenzoic acid demonstrates a strong preference for the metabolites arising from desaturation ($>$50% of products formed) we would expect an increase in the partition towards alkene formation in the F298V mutant with 4-*n*-propylbenzoic acid.

The addition of 4-*n*-propylbenzoic acid to the F298V mutant of CYP199A4[13] resulted in a $\geq$95% shift to the high-spin form of ferric CYP199A4 (Figure S11). The F298V mutant of CYP199A4 catalyzed the oxidation of 4-*n*-propylbenzoic with a lower NADH oxidation rate than the wild-type (WT) enzyme (378 ± 5 min^{-1}). The product distribution of the F298V variant was significantly different to that of the WT enzyme (Scheme 3). The alkene desaturation metabolite was the major metabolite (67%) along with hydroxylation at Cα to generate 4-(1-hydroxypropyl)benzoic acid (27%). In contrast, to the WT enzyme almost none of the 4-(2-hydroxypropyl)benzoic acid metabolite was observed from hydroxylation at Cβ (Figure 5a). Three minor metabolites were also observed in these oxidation reactions. Trace amounts of 4-propionylbenzoic acids arising from further oxidation of the Cα hydroxy metabolite were observed. The other minor metabolites had retention times and mass spectra consistent with epoxidation and Cγ hydroxylation of the alkene desaturation metabolite comparable with what is observed with 4-*n*-butylbenzoic acid (Figure S12 and Figure S13). The levels of product formation were slightly reduced in the oxidation of 4-*n*-propylbenzoic acid by F298V (coupling efficiency of 71% versus 86% in the WT enzyme, Table 1). Overall, the metabolites arising from initial desaturation increased from 45% to 73% in the F298V variant versus WT CYP199A4.



Scheme 3. Product distributions arising from oxidation of 4-*n*-propylbenzoic acid by WT CYP199A4 and the F298V variant.

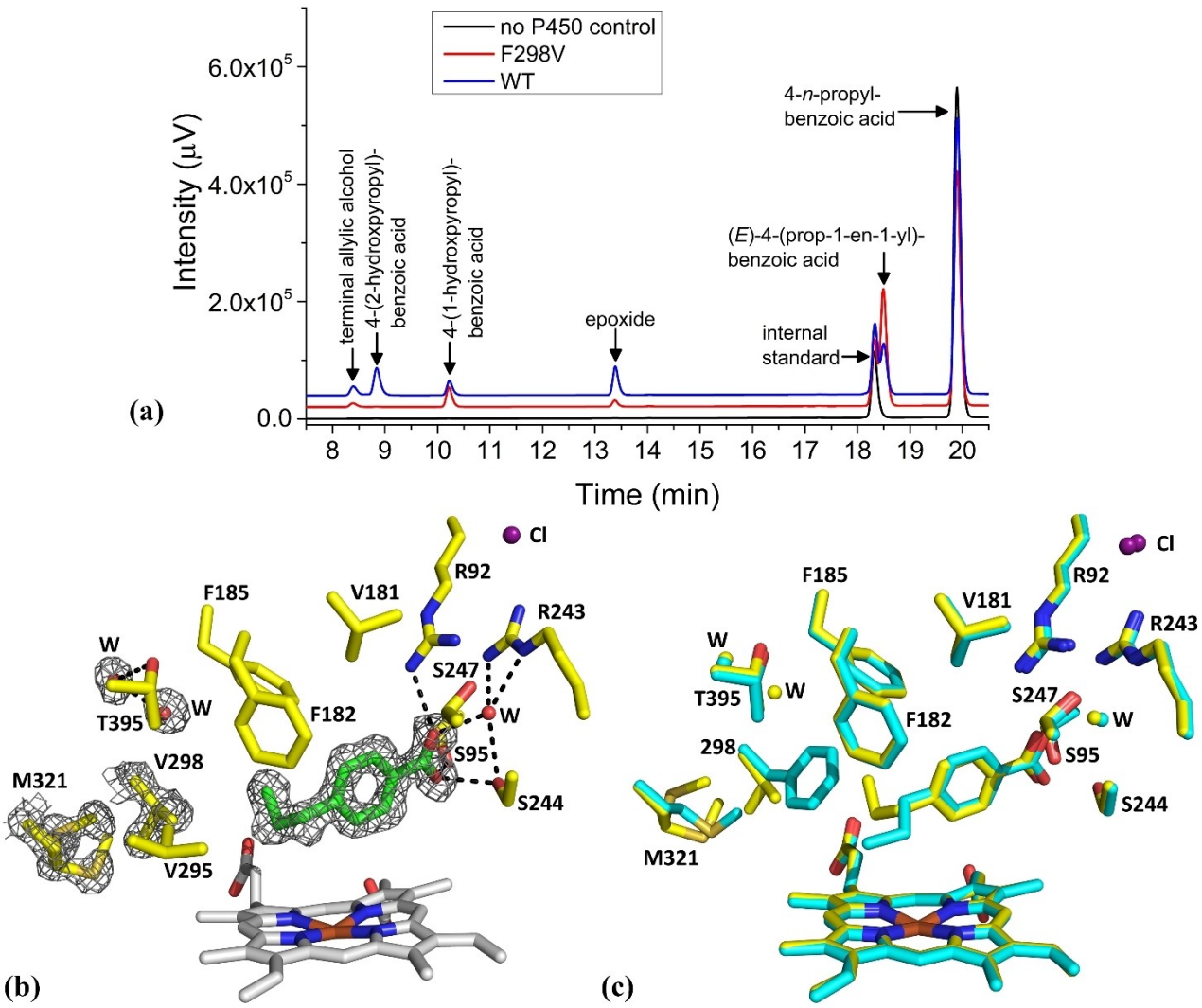


Figure 5. (a) HPLC analysis comparing the *in vitro* oxidation of 4-*n*-propylbenzoic acid by the F298V mutant and WT CYP199A4 enzymes. **(b)** The crystal structure of the F298V mutant of CYP199A4 in complex with 4-*n*-propylbenzoic acid, solved at 1.54-Å resolution (PDB 7UDF). A $2mF_o$-DF_c feature-enhanced map (contoured at 1.0 σ) of the substrate, active-site waters and V298 and M321 residues is shown as grey mesh. M321 has two distinct conformations with occupancies of 67% and 33%. **(c)** Superimposed structures of WT CYP199A4 (cyan, PDB code: 5YQA) and the F298 V mutant (yellow) bound to 4-*n*-propylbenzoic acid.

We were able to crystallize the F298 V variant of CYP199A4 in the presence of 4-*n*-propylbenzoic acid. The crystal structure was determined to 1.54 Å (PDB 7UDF; Table S4). It crystallized in $P2_1$ symmetry with a single subunit and the overall protein fold was very similar to the 4-propylbenzoic acid-bound WT CYP199A4 structure (Figure S14, the RMSD was 0.224 Å). There was electron density in the active site that was readily modelled by 4-*n*-propylbenzoic (Figure 5b). The carboxylate and benzene moieties of each substrate were found in comparable positions to that of 4-*n*-propylbenzoic acid in WT CYP199A4 (Figure 5c).

The electron density clearly showed the mutation of phenylalanine 298 to valine and that the alkyl substituent now occupies a substantially altered position in the active site (Figure 5c and Figure S15). The new conformation of the *n*-propyl chain closely resembles that of 4-*n*-butylbenzoic acid in WT CYP199A4 (Figure S16). The benzylic α-carbon of 4-*n*-propylbenzoic acid in the F298 V variant was found closer to the heme iron than that of the same substrate in the WT enzyme (3.3 Å versus 3.6 Å, from the Cpd I Fe=O oxygen, Table S5). However, the β-carbon (4.0 Å versus 2.2 Å) and γ-carbon (5.4 Å versus 2.6 Å) were significantly further away iron-oxo center (Table S5). The angles of the alkyl chain carbons to the plane of the Fe=O bond are also different and are provided in Table S5. These changes would be expected to result in changes in product distribution and are in agreement with the reduction of the hydroxylation metabolite at Cβ that is observed.

Discussion

The selectivity for the Cα and Cβ positions over others could be rationalized from the crystal structures of the 4-*n*-propylbenzoic acid-bound forms of CYP199A4. These showed that the longer alkyl side chains were oriented so the Cγ C–H bonds would be placed in an unfavorable location for abstraction. The distribution of oxidized metabolites was substrate and C–H bond type dependent. The strength of C–H bonds is known to decrease with the degree of substitution and benzylic bonds are easier to abstract than aliphatic ones (Table 3). The C–H bond strength of a cyclopropyl moiety is higher than those of a straight chain alkyl group (Table 3). The C–H bonds of a tertiary carbon are weaker than those of a secondary carbon which are in turn weaker than these bonds in primary carbons.[15–16]

Altering the type of alkyl side chain in the *para*-position of the benzoic acid had an influence on the product distribution. Previous results for 4-*t*-butylbenzoic acid oxidation established that the enzyme was capable of efficiently oxidizing non-activated primary methyl C–H bonds.[9c,14] However, with other substrates the amount of product generated from abstraction of a primary Cβ methyl group was low when a suitable benzylic C–H bond at Cα was available, e.g., 4-ethyl- and 4-isopropylbenzoic acid. For certain substrates in which Cβ was a methylene group, C–H bond hydroxylation at this position was able to compete with that at the benzylic Cα. For example, the CYP199A4 catalyzed oxidation of 4-*n*-propylbenzoic acid generated substantial quantities of the β-OH metabolite. Previously we have demonstrated that the oxidation of 4-cyclohexylbenzoic acid by CYP199A4, which also has a Cβ methylene group, generated higher levels of the alcohol at this position.[13]

In contrast, CYP199A4 catalyzed oxidation of 4-*n*-butylbenzoic acid did not generate any detectable product arising from β-methylene hydroxylation. Instead, alkene desaturation (Cα–Cβ) was the major product and this was accompanied by comparable levels of hydroxylation at Cα. A further oxidation product, which arose from hydroxylation of the alkene at Cγ, was also detected. In the CYP199A4 structure with this substrate the longer alkyl side chain is held in such a way that abstraction of a hydrogen at Cβ would be disfavoured over one at Cα. With 4-isobutylbenzoic acid formation of the Cβ alcohol was much higher than the linear alkyl substrates and the levels of desaturation were low. The structure of CYP199A4 bound with 4-isobutylbenzoic acid highlighted that hydroxylation at the more substituted Cβ atom of this species occurred despite its distance from the heme. This could be in part related to the nature of the tertiary C–H bond at this position, which would be energetically favoured over the primary and secondary equivalents in the other substrates.

The oxidation of 4-cyclopropylbenzoic acid only generated product arising from C–H abstraction at Cα. The crystal structure with this substrate revealed that the cyclopropyl group was held in a comparable location relative to the heme iron as the ethyl moiety. The strong C–H bonds of the cyclopropyl group are hypothesized to prevent abstraction at the β-position but the effect of the benzoic moiety must make the reaction feasible at the benzylic Cα. H-abstraction of a benzylic C–H bond has been calculated to be approximately 11-fold more favorable compared to H-abstraction from an equivalent secondary carbon center.[21] The coupling efficiency determined for the CYP199A4 catalyzed oxidation of 4-cyclopropylbenzoic acid was lower compared to other propyl-substituted benzoic acids (Table 1). In addition, there was no increase in the hydrogen peroxide uncoupling ($\leq$2%). In this instance it appears that C–H bond abstraction may be suppressed, resulting in an increase in oxidase uncoupling. This could be due to an unfavorable positioning of the Cα C–H bond that the enzyme abstracts, which points away from the heme in the crystal structure.

The proportion of metabolites arising from CYP199A4 catalyzed desaturation varied significantly across the substrates. Desaturation could arise via initial hydrogen abstraction at Cα generating a radical at this position. Radical rebound to yield a hydroxylation product would occur as it does in the majority of P450 catalyzed transformations. Desaturation can also occur if the initial abstraction step is followed by electron transfer to Cpd II or hydrogen abstraction by Cpd II. If electron transfer occurs the final step would be the loss of a proton from an organic cation to produce the alkene. An increase in alkene desaturation product would therefore be indicative of a greater impediment to radical rebound at Cα. Some trends in the proportion of desaturation product were apparent. No desatu-

Table 3. Experimentally derived C–H bond energies of alkanes and cycloalkanes. The bond breaking energies were all determined experimentally in the gas phase at 298 K.[15–16]

C–H bond type	Dissociation energy [kJ mol^{-1}]
Methane (H_3C–**H**)	439
Primary (H_3C–CH_2–**H**)	423
Secondary ($(H_3C)_2CH$–**H**)	413
Tertiary ($(CH_3)_3C$–**H**)	404
Benzene (C_6H_5–**H**)	472
Benzylic ($PhCH_2$-**H**)	377
Allylic (H_2C=CH–CH_2–**H**)	372
Cyclopropane (C_3H_5–**H**)	456

ration was observed with 4-cyclopropylbenzoic acid. A cyclopropyl radical is less readily oxidized to a cation than the ethyl or isopropyl equivalent and the cyclopropenyl moiety is known to be unstable.[22] These may preclude this pathway from occurring in the oxidation of this substrate. The level of alkene formation in the CYP199A4 catalyzed oxidation of 4-*n*-propyl- and 4-*n*-heptyl-benzoic acids were comparable to those found in 4-isopropylbenzoic acid and only marginally lower than that of 4-ethylbenzoic acid. The oxidation of 4-isobutylbenzoic acid resulted in very low level of desaturation but with 4-*n*-butylbenzoic alkene formation was the major metabolite. The structures obtained with WT CYP199A4 and these alkylbenzoic acids provide important parameters to assist with understanding the relative amounts of hydroxylation at Cα and Cβ and desaturation. The F298V mutant of CYP199A4 significantly altered the product distribution of 4-n-propylbenzoic acid. There was a loss of the Cβ hydroxylation metabolite and an increase in the proportion of metabolites arising from desaturation. The similar orientation of the n-propyl substituent in the F298V mutant and the 4-n-butyl moiety in the WT enzyme results in increased levels of desaturation relative to the other substrate/enzyme reactions reported here.

Where alkenes are generated in sufficient quantities and characterized, we observed a strong preference for the *trans*-alkene product over the *cis*. The observation of further oxidation of the (*E*)-4-(but-1-en-1-yl)benzoic acid and the (*E*)-4-(prop-1-en-1-yl)benzoic acid products were the only example of abstraction at Cγ. The difference in spin state shifts induced in CYP199A4 by the *para*-alkenyl versus the alkyl substituted benzoic acids infers that the side chains of these substrates bind in the active site in a significantly different orientation.

The orientation of the C–H bond to be abstracted, as well as its distance from the iron-oxo intermediate, will be important in determining the product distribution. In these enzyme/substrate complexes, the more reactive benzylic Cα atom is always further away from the heme iron than Cβ. This could reduce the efficiency of the rebound step, or the β-substituent could interfere with this step. It is tempting to speculate that, when desaturation is energetically feasible, this slowing down of oxygen rebound results in an increase in the amount of alkene product observed. This work highlights that the alkyl substituents can adopt positions within the active site, relative to the heme, that favor desaturation over hydroxylation. These structural parameters must be important in P450 enzymes which catalyze desaturation as part of their physiological function, such as the CYP61 and CYP710 families of sterol 22-desaturases.[8f,g]

Another significant observation from this work is that in 4-cyclopropyl- and 4-isobutyl-benzoic acid-bound crystal structures of CYP199A4 the C–H bond to be abstracted is not in a favorable position for abstraction. While the benzoic acid moiety is well anchored in the CYP199A4 substrate-binding pocket, the alkyl side chain of some of these substrates will have more flexibility to move. The binding positions of the alkyl side chains of 4-*n*-butyl-, 4-isobutyl- and 4-*n*-heptyl-benzoic acids in their crystal structures demonstrate that alternative binding positions are possible. The alkyl substituents could have multiple orientations within the active site at ambient temperatures. These changes could occur before or after dioxygen binding and the subsequent steps in the catalytic cycle. This would result in a different orientation of the alkyl group relative to the iron-oxo moiety of Cpd I. Recently we demonstrated that there was limited restrictions to the rotation of the alkene moiety of 4-vinylbenzoic acid within the active site of CYP199A4 using MD simulations.[23] In addition, simulations demonstrated that the larger phenyl and cyclohexyl substituents could also rotate within the active site of this enzyme.[13] The alkyl chains investigated in this work may be freer to move or rotate. Movement of the substrate from the position observed in the crystal structures may be required before C–H bond abstraction occurs.

Conclusions

Overall, we have shown that CYP199A4 can bind *para*-substituted alkylbenzoic acids with high affinity. Aliphatic benzylic and methylene C–H bond abstraction proceeds with high efficiency and selectivity. For certain side chains the desaturation pathway can compete effectively with hydroxylation. This work will enable the design of bespoke substituted probe substrates and mutant forms of the CYP199A4 enzyme, which can provide new and unprecedented levels of detail into mechanistic aspects of P450 catalyzed desaturation reactions. Suitably isotopically labelled substrate analogues could elucidate the details of the mechanism of desaturation. Further mutagenesis studies to alter the substrate-binding mode and perturb the selectivity of the oxidation reactions could also be undertaken. The structural data obtained here will also be important for future molecular dynamic simulations and theoretical calculations to understand the selectivity of P450 catalyzed hydroxylations and the partition between hydroxylation and desaturation.

Experimental Section

General: General reagents and organic substrates were from Sigma-Aldrich, TCI, Fluorochem or VWR. Buffer components, NADH and isopropyl β-D-thiogalactopyranoside (IPTG) were from Astral Scientific, Australia. The expression, purification, and quantitation of CYP199A4, HaPux and HaPuR have been described elsewhere.[9c,11a–b] The proteins were stored at −20 °C in 50% glycerol which was removed prior to experiments using a 5 mL gel filtration column (PD-10, Cytiva). UV-visible spectra and spectroscopic activity assays were recorded at 30 ± 0.5 °C on an Agilent CARY-60 spectrophotometer.

Analytical High Performance Liquid Chromatography (HPLC) was performed on samples prepared as previously described.[12] Briefly, 132 μL of the reaction mixture was mixed with 2 μL of an internal standard solution (10 mM 9-hydroxyfluorene in ethanol). For 4-isopropylbenzoic acid, the 9-hydroxyfluorene coeluted with the alkene product so instead diphenylmethane was used as the internal standard (10 mM in ethanol). This was mixed with 66 μL of acetonitrile before analysis by HPLC. Gas Chromatography-Mass Spectrometry (GC-MS) data on derivatized samples were analyzed

on a DB-5 MS fused silica column (30 m × 0.25 mm, 0.25 μm) using on a Shimadzu GC-17A system using the same method as described previously[12] or a Shimadzu GC-2010 gas chromatograph equipped with an autoinjector and a GCMS-QP2010S detector. The interface and injection port temperatures were held at 280 and 250 °C. The column was held at 120 °C for 3 min, and the temperature was then increased to 240 °C at a rate of 7.5 °C min^{-1} and held at 240 °C for 1 min. The temperature was raised to 250 °C at a rate of 2 °C min^{-1} and held at 250 °C for 5 min. The HPLC and GC retention times for the substrates and products are given in the appropriate chromatograms in the supporting information.

Preparative HPLC was performed after whole-cell oxidation reactions as described previously.[12] Those fractions containing a single product (≥95%) were combined and freeze-dried using an Alpha 2-4 LDplus freeze-dryer (Christ, Germany), followed by characterization using NMR and GC-MS. Analytical HPLC was performed on Shimadzu Prominence HPLC system with a LC-20AD pump, SIL-20A autosampler, SPD-20A UV-visible detector and CTO-20A column oven. A Phenomenex Lux 3u Cellulose-1 column was used for enantioselective HPLC. A gradient of 10% acetonitrile in H_2O to 20% acetonitrile in H_2O over 30 min with a further increase to 90% acetonitrile in H_2O over 30 min was used.

Substrate binding: spin state determination and binding titrations: The high-spin heme content was determined using an enzyme concentration of ~2–3 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 approximately ±5%) as described previously.[11c,24]

For dissociation constant determination CYP199A4 was diluted to 0.5–1.5 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. The Soret band peak-to-trough difference (ΔA) in absorbance was recorded between 700 nm and 250 nm. Additional aliquots of substrate were added until ΔA did not change further. The dissociation constants, K_d, were obtained as described previously.[12,17,25]

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

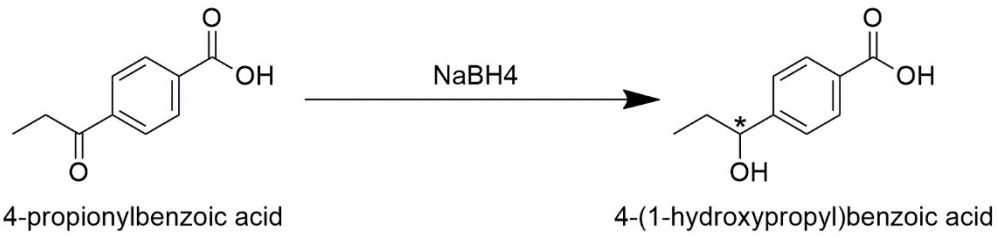
The concentration of hydrogen peroxide was calculated using the horseradish peroxidase (HRP)/phenol/4-aminoantipyrine (4-AP) assay as described previously.[26]

Analysis of metabolites: Calibrations of fixed amounts of product standards were used to quantitate the level of the metabolites in the enzyme-catalyzed oxidations by HPLC or GC-MS. As authentic samples were unavailable, they were synthesized using whole-cell oxidation systems or the coupling was estimated based on the closest available compound. For example, for the product of 4-cyclopropylbenzoic acid for which no authentic standard was available, the product peak was calibrated against a 4-(1-hydroxyethyl)benzoic acid.

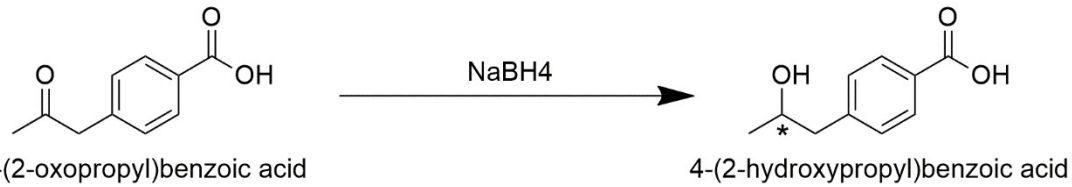
Metabolite synthesis: 4-(1-hydroxypropyl)benzoic acid (Scheme 4). To a stirring solution containing 4-propionylbenzoic acid (5.0 mg, 28 μmol) in MeOH (1 mL) at 0 °C was added sodium borohydride (1.1 mg, 28 μmol) in MeOH (0.5 mL). The reaction was stirred for 1.5 hr at 0 °C. The solvent of the reaction mixture was removed under N_2, and the resulting residue was dissolved in acetonitrile for HPLC analysis.

4-(2-hydroxypropyl)benzoic acid (Scheme 5). To a stirring solution containing 4-(2-oxopropyl)benzoic acid (5.0 mg, 28 μmol) in MeOH (1 mL) at 0 °C was added sodium borohydride (0.6 mg, 14 μmol) in MeOH (0.5 mL). The reaction was stirred for 1.5 hr at 0 °C. The solvent of the reaction mixture was removed under N_2, and the resulting residue was dissolved in acetonitrile for HPLC analysis.

Whole-cell oxidation reactions: When product identification required the synthesis of the metabolite, whole-cell oxidations 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 LB broth, containing the required antibiotics and 1.5 mL of trace elements (20.1 g Na_2EDTA, 16.7 g $FeCl_3 \cdot 6H_2O$, 0.74 g $CaCl_2 \cdot H_2O$, 0.25 g $CoCl_2 \cdot 6H_2O$, 0.18 g $ZnSO_4 \cdot 7H_2O$, 0.132 g $MnSO_4 \cdot 4H_2O$, 0.10 g $CuSO_4 \cdot 5H_2O$), at 37 °C overnight. The incubator temperature was lowered to 18 °C for



Scheme 4. Reduction of 4-propionylbenzoic acid to generate racemic 4-(1-hydroxypropyl)benzoic acid.



Scheme 5. Reduction of 4-(2-oxopropyl)benzoic acid to generate racemic 4-(2-hydroxypropyl)benzoic acid.

1 hour before the addition of 250 μM IPTG (from a stock of 0.5 M in water). The culture was grown for a further 24 hours, and then harvested by centrifugation (5,500 *g*, 10 min). The harvested cells were resuspended in *E. coli* minimal media (EMM; K_2HPO_4 7 g, KH_2PO_4 3 g, $(NH_4)_2SO_4$ 1 g, Na_3citrate 0.5 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 (3×100 mL), washed with brine (100 mL) and dried with $MgSO_4$. The organic extracts were pooled, and the solvent was removed using a rotary evaporator. 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 ^{1}H and 126 MHz for ^{13}C. A combination of ^{1}H 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.

For analysis of the reactions of 4-*n*-propylbenzoic acids, the products and remaining starting material were derivatized into their equivalent methyl esters using excess N,N-dimethylformamide dimethyl acetyl (DMF-DMA). The resulting mixtures were purified using silica column chromatography to isolate the products which were then characterized as above.

Protein crystallization : For crystallization, the CYP199A4 mutants were further purified by size exclusion chromatography (ENrich SEC 650, 10×300 mm or HiPrep Sephacryl S-200 HR size-exclusion column (16 mm×600 mm) at a flow rate of 1 mL min^{-1} using Tris buffer (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 vapor 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 μm by 150 μm by 20 μm were obtained in 24 hours to 1 week using a reservoir solution containing 0.2 M magnesium acetate tetrahydrate, 20–23% w/v PEG-3,350 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 Laboratories) 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.[27]

Crystallization – 4-n–propylbenzoic acid: For crystallization of CYP199A4 the enzyme was purified using a different protocol (see supporting information). For co-crystallization with 4-*n*-propylbenzoic acid, CYP199A4 was concentrated to ~30 mg mL^{-1} in 20 mM KH_2PO_4, pH 7.4, 200 mM KCl, 10 mM β-mercaptoethanol and saturated with substrate. The sitting-drop vapor-diffusion method was used to perform crystal screening with Hampton Research Crystal Screen HT and Index HT. Good quality crystals of the complexes were obtained in about two weeks at 20 °C using 1 μL protein solution mixed with 2 μL reservoir solution and equilibrated with 200 μL reservoir solution. The final conditions were: 0.25 M NaCl, 1.5 M $(NH_4)_2SO_4$, 0.1 M Bis-Tris, pH 8.0. All crystals were cryoprotected by the addition of 20% v/v glycerol to the crystallization solutions. X-ray diffraction data were collected at 100.15 K on beamline 5 A at the High Energy Accelerator Research Organization (KEK), Japan. Complete data collection statistics are summarized in Table S2 (PDB code 5YQA).

X-ray structures : The diffraction data were indexed and integrated using the HKL2000 package,[28] iMosflm.[29] or by automated data processing at the Australian Synchrotron (MX2 beamline).[30] The data was scaled, merged and R-free flags were added using Aimless,[31] both available in the CCP4 suite of programs.[32] The phase problem was solved using the molecular replacement method with Phaser in CCP4,[33] using the CYP199A4 structures (PDB: 4DO1, PDB: 4DNZ or PDB 5UVB with the ligand, waters, chloride and heme removed) as the initial search models. Electron density maps were obtained after initial model building and the model was rebuilt using Coot.[34]

For the other benzoic acids and structural refinements were performed over multiple iterations using Phenix Refine available in the Phenix suite of programs.[35] Composite omit maps were generated using the composite omit or feature enhanced maps program in Phenix.[36] For structure determination of 4-*n*-propylbenzoic acid, the phase problem was solved using molecular replacement with Phaser in the CCP4 suite using. Electron density maps were obtained after initial model building and the model was rebuilt using Coot and refined using Refmac5.[37] The structures of the substrates were manually built and adjusted under the guidance of Fo–Fc difference maps using program Coot and refined with Refmac5. The stereochemical qualities of the refined structures were checked with PROCHECK.[38]

Data collection and structural refinement statistics are summarized in Table S1 and Table S2. The coordinates for the crystal structures have been deposited in the Protein Data Bank (accession codes presented in Table S1 and S2).

Acknowledgements

This work was supported by ARC grant DP140103229 (to J.J.D.V. and S.G.B.). 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., J.H.Z.L, and M.N.P.). We also acknowledge the University of Adelaide for a Constance Fraser Scholarships (J.H.Z.L and M.N.P) and the CSIRO Synthetic Biology Future Science Platform for a Scholarship (J.H.Z.L.). This research was also supported by an AINSE Ltd. Postgraduate Research Award (PGRA) to M.N.P. 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. This research was undertaken in part using the MX1 beamline at the Australian Synchrotron, part of ANSTO. This research was undertaken in part using the MX2 beamline at the Australian Synchrotron, part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector. 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

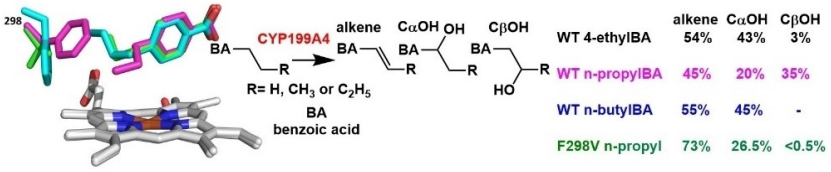
The data that support the findings of this study are available in the supplementary material of this article.

Keywords: cytochrome P450 enzymes · desaturation · enzyme mechanism · metalloenzymes · X-ray crystallography

[1] a) P. R. Ortiz de Montellano, *Chem. Rev.* **2010**, *110*, 932–948; b) T. L. Poulos, *Chem. Rev.* **2014**, *114*, 3919–3962.
[2] a) F. Ogliaro, S. P. de Visser, S. Cohen, P. K. Sharma, S. Shaik, *J. Am. Chem. Soc.* **2002**, *124*, 2806–2817; b) J. Rittle, M. T. Green, *Science* **2010**, *330*, 933–937; c) P. K. Sharma, S. P. De Visser, S. Shaik, *J. Am. Chem. Soc.* **2003**, *125*, 8698–8699; d) B. Meunier, S. P. de Visser, S. Shaik, *Chem. Rev.* **2004**, *104*, 3947–3980; e) S. Shaik, S. Cohen, Y. Wang, H. Chen, D. Kumar, W. Thiel, *Chem. Rev.* **2010**, *110*, 949–1017.
[3] J. T. Groves, G. A. McClusky, *J. Am. Chem. Soc.* **1976**, *98*, 859–861.
[4] X. Huang, J. T. Groves, *J. Biol. Inorg. Chem.* **2017**, *22*, 185–207.
[5] 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..
[6] M. E. Albertolle, F. P. Guengerich, *J. Inorg. Biochem.* **2018**, *186*, 228–234.
[7] H. Wong Siew, G. Bell Stephen, J. De Voss James, *Pure Appl. Chem.* **2017**, *89*, 841.
[8] a) V. M. Kramlinger, L. D. Nagy, R. Fujiwara, K. M. Johnson, T. T. Phan, Y. Xiao, J. M. Enright, M. B. Toomey, J. C. Corbo, F. P. Guengerich, *FEBS Lett.* **2016**, *590*, 1304–1312; b) K. M. Johnson, T. T. N. Phan, M. E. Albertolle, F. P. Guengerich, *J. Biol. Chem.* **2017**, *292*, 13672–13687; c) A. E. Rettie, M. Boberg, A. W. Rettenmeier, T. A. Baillie, *J. Biol. Chem.* **1988**, *263*, 13733–13738; d) C. A. Reilly, G. S. Yost, *Drug Metab. Dispos.* **2005**, *33*, 530–536; e) R. S. Obach, *Drug Metab. Dispos.* **2001**, *29*, 1599–1607; f) S. L. Kelly, D. C. Lamb, B. C. Baldwin, A. J. Corran, D. E. Kelly, *J. Biol. Chem.* **1997**, *272*, 9986–9988; g) T. Morikawa, M. Mizutani, D. Ohta, *Biochem. Soc. Trans.* **2006**, *34*, 1202–1205.
[9] a) S. G. Bell, N. Hoskins, F. Xu, D. Caprotti, Z. Rao, L. L. Wong, *Biochem. Biophys. Res. Commun.* **2006**, *342*, 191–196; b) S. G. Bell, F. Xu, I. Forward, M. Bartlam, Z. Rao, L.-L. Wong, *J. Mol. Biol.* **2008**, *383*, 561–574; 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) C. J. Whitehouse, S. G. Bell, L. L. Wong, *Chemistry* **2008**, *14*, 10905–10908.
[10] a) L. Ji, A. S. Faponle, M. G. Quesne, M. A. Sainna, J. Zhang, A. Franke, D. Kumar, R. van Eldik, W. Liu, S. P. de Visser, *Chemistry* **2015**, *21*, 9083–9092; b) D. Kumar, S. P. De Visser, S. Shaik, *J. Am. Chem. Soc.* **2004**, *126*, 5072–5073; c) W. Lai, H. Chen, S. Cohen, S. Shaik, *J. Am. Chem. Soc.* **2011**, *2*, 2229–2235.
[11] 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) T. Coleman, R. R. Chao, J. B. Bruning, J. De Voss, S. G. Bell, *RSC Adv.* **2015**, *5*, 52007–52018; d) R. R. Chao, J. J. De Voss, S. G. Bell, *RSC Adv.* **2016**, *6*, 55286–55297.
[12] T. Coleman, S. H. Wong, M. N. Podgorski, J. B. Bruning, J. J. De Voss, S. G. Bell, *ACS Catal.* **2018**, *8*, 5915–5927.
[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] M. N. Podgorski, T. Coleman, R. R. Chao, J. J. De Voss, J. B. Bruning, S. G. Bell, *J. Inorg. Biochem.* **2020**, *203*, 110913.
[15] S. J. Blanksby, G. B. Ellison, *Acc. Chem. Res.* **2003**, *36*, 255–263.
[16] Z. Tian, A. Fattahi, L. Lis, S. R. Kass, *J. Am. Chem. Soc.* **2006**, *128*, 17087–17092.
[17] T. Coleman, J. E. Stok, M. N. Podgorski, J. B. Bruning, J. J. De Voss, S. G. Bell, *J. Biol. Inorg. Chem.* **2020**, *25*, 583–596.
[18] a) N. Zhu, J. Zhao, H. Bao, *Chem. Sci.* **2017**, *8*, 2081–2085; b) S. E. Denmark, J. M. Kallemeyn, *J. Am. Chem. Soc.* **2006**, *128*, 15958–15959.
[19] a) A. Gumiero, C. L. Metcalfe, A. R. Pearson, E. L. Raven, P. C. Moody, *J. Biol. Chem.* **2011**, *286*, 1260–1268; b) H. Chen, H. Hirao, E. Derat, I. Schlichting, S. Shaik, *J. Phys. Chem. B* **2008**, *112*, 9490–9500; c) I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, S. G. Sligar, *Science* **2000**, *287*, 1615–1622.
[20] T. Coleman, S. H. Wong, M. N. Podgorski, J. B. Bruning, J. J. De Voss, S. G. Bell, *ACS Catal.* **2018**, *8*, 5915–5927.
[21] S. P. de Visser, D. Kumar, S. Cohen, R. Shacham, S. Shaik, *J. Am. Chem. Soc.* **2004**, *126*, 8362–8363.
[22] R. Woolard, J. Mck, *Summary of Stability of Cyclopropenes, Vol. 5* (Ed. B. Halton), JAI Press, Greenwich, CT **1996**, pp. 245–270.
[23] 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**, **2021**, *11*, 1995–2010.
[24] N. K. Maddigan, S. G. Bell, *Arch. Biochem. Biophys.* **2017**, *615*, 15–21.
[25] J. W. Williams, J. F. Morrison, *Methods Enzymol.* **1979**, *63*, 437–467.
[26] F. Xu, S. G. Bell, Z. Rao, L. L. Wong, *Protein Eng. Des. Sel.* **2007**, *20*, 473–480.
[27] a) 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; b) T. M. McPhillips, S. E. McPhillips, H. J. Chiu, A. E. Cohen, A. M. Deacon, P. J. Ellis, E. Garman, A. Gonzalez, N. K. Sauter, R. P. Phizackerley, S. M. Soltis, P. Kuhn, *J. Synchrotron Radiat.* **2002**, *9*, 401–406.
[28] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307–326.
[29] T. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, A. G. Leslie, *Acta Crystallogr. Sect. D* **2011**, *67*, 271–281.
[30] 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.
[31] P. R. Evans, G. N. Murshudov, *Acta Crystallogr. Sect. D* **2013**, *67*, 1204–1214.
[32] 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.
[33] a) 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; b) C. C. Project, *Acta Crystallogr. Sect. D* **1994**, *50*, 760–763.
[34] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D* **2010**, *66*, 486–501.
[35] 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.
[36] P. V. Afonine, N. W. Moriarty, M. Mustyakimov, O. V. Sobolev, T. C. Terwilliger, D. Turk, A. Urzhumtsev, P. D. Adams, *Acta Crystallogr. Sect. D* **2015**, *71*, 646–666.
[37] a) P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126–2132; b) G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sect. D* **1997**, *53*, 240–255.
[38] R. A. Laskowski, M. W. Macarthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **1993**, *26*, 283–291.

Manuscript received: September 26, 2022
Revised manuscript received: October 18, 2022
Accepted manuscript online: October 21, 2022
Version of record online: ■■, ■■■

RESEARCH ARTICLE



The enzyme catalyzed oxidation of alkyl substituted substrates results in significant changes in the regioselectivity of hydroxylation, and in the relative levels of desaturated alkene product. Using X-ray crystal structures and protein engineering we reveal that the position of the alkyl group relative to the heme and the C–H bond strengths are important in determining these outcomes.

*Dr. T. Coleman, D. Z. Doherty, T. Zhang, M. N. Podgorski, R. Qiao, J. H. Z. Lee, Dr. J. B. Bruning, Prof. J. J. De Voss, Prof. W. Zhou, Dr. S. G. Bell**