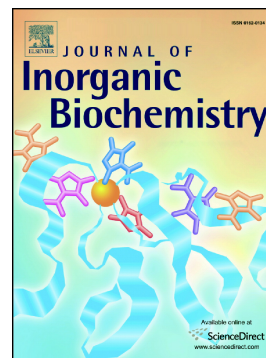


Investigation of the requirements for efficient and selective cytochrome P450 monooxygenase catalysis across different reactions

Matthew N. Podgorski, Tom Coleman, Rebecca R. Chao, James J. De Voss, John B. Bruning, Stephen G. Bell



PII: S0162-0134(19)30642-7

DOI: <https://doi.org/10.1016/j.jinorgbio.2019.110913>

Reference: JIB 110913

To appear in: *Journal of Inorganic Biochemistry*

Received date: 25 September 2019

Revised date: 26 October 2019

Accepted date: 7 November 2019

Please cite this article as: M.N. Podgorski, T. Coleman, R.R. Chao, et al., Investigation of the requirements for efficient and selective cytochrome P450 monooxygenase catalysis across different reactions, *Journal of Inorganic Biochemistry* (2019), <https://doi.org/10.1016/j.jinorgbio.2019.110913>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Investigation of the requirements for efficient and selective cytochrome P450 monooxygenase catalysis across different reactions

Matthew N. Podgorski,[a] Tom Coleman,[a] Rebecca R. Chao,[a] James J. De Voss,[b] John B. Bruning,[c] and Stephen G. Bell*[a]

[a] Department of Chemistry, University of Adelaide, Adelaide, SA 5005, Australia

E-mail: stephen.bell@adelaide.edu.au

[b] School of Chemistry and Molecular Bioscience, University of Queensland, St Lucia, Qld 4072, Australia

[c] School of Biological Sciences, University of Adelaide, SA 5005, Australia

* Corresponding Author; Stephen G. Bell; E-mail: stephen.bell@adelaide.edu.au; Fax, +61 8 8313 4380

Abstract

The cytochrome P450 metalloenzyme CYP199A4 from *Rhodopseudomonas palustris* HaA2 catalyzes the highly efficient oxidation of *para*-substituted benzoic acids. Here we determined crystal structures of CYP199A4, and the binding and turnover parameters, with different *meta*-substituted benzoic acids in order to establish which criteria are important for efficient catalysis. When compared to the *para* isomers, the *meta*-substituted benzoic acids were less efficiently oxidized. For example, 3-formylbenzoic acid was oxidized with lower activity than the equivalent *para* isomer and 3-methoxybenzoic acid did not undergo O-demethylation by CYP199A4. The structural data highlighted that the *meta*-substituted benzoic acids bound in the enzyme active site in a modified position with incomplete loss of the distal water ligand of the heme moiety. However, for both sets of isomers the *meta*- or *para*-substituent pointed towards, and was in close proximity, to the heme iron. The absence of oxidation activity with 3-methoxybenzoic acid was assigned to the observation that the C-H bonds of this molecule point away from the heme iron. In contrast in the *para* isomer they are in an ideal location for abstraction. These findings were confirmed by using the bulkier 3-ethoxybenzoic acid as a substrate which removed the water ligand and reoriented the *meta*-substituent so that the methylene hydrogens pointed towards the heme, enabling more efficient oxidation. Overall we show relatively small changes in substrate structure and position in the active site can have a dramatic effect on the activity.

Keywords Metalloenzyme; C-H bond abstraction; X-Ray crystal structures; Heme monooxygenase; Enzyme catalysis; Molecular docking

1. Introduction

The heme-dependent cytochrome P450 (CYP) monooxygenase family of metalloenzymes catalyzes the oxidation of biological molecules often with high selectivity.[1-3] Moreover, they can catalyze C-H bond abstractions resulting in carbon hydroxylation at ambient temperatures using a reactive ferryl Compound I (Cpd I) intermediate.[1, 4-8] The insertion of the oxygen atom into the C-H bond is hypothesized to occur via a radical rebound mechanism after the Cpd I intermediate abstracts a hydrogen atom from the alkyl substrate.[5, 9] Importantly, members of this enzyme superfamily can catalyze a diverse range of other oxidative transformations including epoxidations, heteroatom dealkylations, sulfoxidations and other more complex reactions.[10-16]

Investigations into the substrate range of the soluble bacterial P450 CYP199A4 revealed that the enzyme is predisposed to bind and rapidly oxidize *para*-substituted benzoic acid substrates.[17-24] The fastest reaction catalyzed by CYP199A4 is *O*-demethylation of 4-methoxybenzoic acid, performed at a rate of 1220 min⁻¹ with a coupling efficiency of 91%.[17, 19] CYP199A4 is also able to bind and oxidize the bulkier substrate 3,4-dimethoxybenzoic acid (veratric acid). In this instance, demethylation occurs only at the *para* position, affording 4-hydroxy-3-methoxybenzoic acid; the *meta*-substituent is not attacked.[21] CYP199A4 does not hydroxylate the aromatic ring of any of these substrates and also fails to oxidize benzoic acid.[19]

Crystal structures of substrate-bound CYP199A4 have been solved to elucidate the orientation of substrates in the binding pocket and examine the tendency of CYP199A4 to attack only the *para*-substituent (Figure 1).[18, 19, 22] These crystal structures revealed that the benzoate moiety, held almost perpendicular to the heme, forms both hydrophilic and

hydrophobic interactions with residues in the binding pocket.[22] The substrate carboxylate hydrogen bonds to the hydroxyl groups of S95 and S244 and forms a salt bridge with the guanidinium of R92.[22] Additionally, it interacts with a molecule of water which is in turn hydrogen-bonded to S244 and R243.[22]

If the substrate carboxylate is replaced with alternate functional groups, it abolishes the ability of the enzyme to tightly bind the substrate.[23, 24] Van der Waals interactions exist between the benzene ring and L98, A248, V181, and F185 and also between the *para*-substituent and F182 and F298.[18] Consistent with the $\geq 95\%$ shift to high-spin observed when 4-methoxybenzoic acid binds, the crystal structure (PDB: 4DO1) revealed that the iron-bound water is completely displaced by this substrate. The *para*-methoxy group is held in close proximity to the heme iron (the carbon is 4.1 Å from the iron) and thus C-H bond abstraction occurs solely at this position, resulting in *O*-demethylation.[22] When veratric acid binds (PDB: 4EGN), its benzoate moiety and *para*-methoxy group occupy equivalent positions to those of 4-methoxybenzoic acid (Figure 1).[18] As a result, the *para*-methoxy group is close to the heme, but the *meta*-methoxy group of veratric acid is oriented away from the heme and is too far (8.1 Å) from the iron to react.[18]

The activity of CYP199A4 towards multiple other substituted benzoic acid substrates has been investigated (Figure S1).[17-21, 23, 24] 2,4-Dimethoxybenzoic acid was also solely oxidized at the *para* position.[19] On the other hand, 2- and 3-methoxybenzoic acid were not oxidized at all.[19] 3-Methoxybenzoic acid binds substantially less tightly than 4-methoxybenzoic acid ($K_d = 69 \mu\text{M}$ vs $0.22 \mu\text{M}$) and only shifts the spin-state to 40% high-spin, implying that the substrate is not bound in a suitable orientation relative to the heme to efficiently displace the iron-bound water.[19] This is consistent with previous studies which

found that other *meta*-substituted benzoic acids bind weakly to CYP199A2, which shares 86% sequence identity with CYP199A4.[22] For example, 3-chlorobenzoic acid was found to induce a substantially smaller spin-state shift than the *para* isomer in CYP199A2 (10% vs 80%).[26] In the absence of a CYP199A4 crystal structure with a substrate containing only a *meta*-arrangement of substituents, it was proposed that 3-methoxybenzoic acid could bind to the active site in a similar orientation to that of veratric acid (Figure 1). This would position the *meta*-substituent to point away from the heme iron, rather than towards it.[19] If this was the case the *meta*-substituent would be held too far from the iron to react. Of note is that 3,5-dimethoxybenzoic acid, which must bind with one *meta*-methoxy group pointing away from the heme and one methoxy group pointing towards it, is oxidized with low efficiency (at a product formation rate of $7 \mu\text{M} (\mu\text{M-P450})^{-1} \text{ min}^{-1}$).[19]

CYP199A4 can catalyze reactions other than *O*-dealkylation such as the sulfoxidation of 4-methylthiobenzoic acid, *N*-dealkylation of 4-methylaminobenzoic acid and the hydroxylation of 4-methylbenzoic (4-toluic) acid.[17, 18] It also oxidizes bulkier substrates such as 4-ethoxy-, and 4-isopropyl-benzoic acid.[17, 18] Crystal structures have been obtained of CYP199A4 in complex with several of these substrates (PDB codes: 4-methylthio-, 5KT1; 4-methylamino-, 5U6W and 4-ethoxy-benzoic acid, 5U6T).

In contrast to 3-methoxybenzoic acid, both 3-methylamino- and 3-methylthio-benzoic acids were oxidized by CYP199A4[17]; however, the activities were significantly lower than those of the *para* isomers. We therefore set out to more fully investigate the ability of CYP199A4 to bind and oxidize benzoic acid substrates with alternative substituents at the *meta*-position. Docking and structural studies were performed to assess whether 3-substituted benzoic acids bind with the *meta*-substituent pointing away from the heme. The substrates were chosen to

enable exploration of a range of different substituents which would alter the steric bulk and reactivity at this position (Figure 2).

The results were compared with those of the *para*-substituted isomers and used to clarify the binding mode of these substrates in order to rationalize the preference of CYP199A4 for *para*- compared to *meta*-substituted benzoic acids. This provided a greater insight into how the coordination environment of the heme iron and the substrate orientation relative to this cofactor affects the catalytic activity in this important and ubiquitous superfamily of heme metalloenzymes.

2. Experimental methods

2.1 General methods

Production and purification of CYP199A4, crystal growth and structure determination, substrate binding analysis, *in vitro* NADH activity assays, chromatographic analysis of metabolites, hydrogen peroxide determination and catalytic activity as well as the use of a whole-cell oxidation system to generate metabolites for characterization were all performed using methods which have been reported previously (full details also provided in the Supplementary information).[17, 19, 21-23]

2.2 Docking of ligands into CYP199A4

To identify the likely binding mode of the substrates in the binding pocket of CYP199A4, the ligands were docked into a crystal structure of CYP199A4 (PDB ID: 5U6W) using ICM-Pro software, version 3.8-6a (Molsoft LLC, San Diego CA).[27] First, hydrogens were added to the crystal structure and optimized. Charges were also automatically assigned to residues which should be protonated/deprotonated at neutral pH and the orientation of histidine, proline, asparagine, glutamine and cysteine residues was also optimized.[27, 28] Waters were deleted from the structure with the exception of W618 which forms a hydrogen bond with the carboxylate of bound benzoic acid substrates, as well as active site residues (Figure 1).[17, 18, 22] The co-crystal ligand (4-methylaminobenzoic acid) was then removed from the active site of the enzyme and this region was used to define the binding site. The 4-methylaminobenzoic acid ligand was modified to 3-methylaminobenzoic acid and this ligand was re-docked into the binding pocket. The top scoring (lowest energy) pose had the 3-methylamino substituent pointing down towards the heme. A positional restraint was imposed on the nitrogen atom in

order to dock the ligand in an alternative configuration with the 3-methylamino moiety pointing away from the heme. The docking 'score' was then used to evaluate the docked poses, with a value of < -32 indicating a good fit.[27, 28] The other ligands were docked into CYP199A4 using the same method.[29]

3. Results

3.1 Substrate binding and catalytic activity of CYP199A4 towards *meta*-substituted benzoic acids

The ability of CYP199A4 to bind and oxidize 3-methyl- and 3-formyl-benzoic acid was explored. Compared to 3-methoxybenzoic acid, which induces a 40% spin-state shift in the heme iron to the high-spin form upon binding to CYP199A4 (Figure S2),[19] both 3-methyl- and 3-formyl-benzoic acid induced smaller type I shifts: 30% and 10%, respectively (Table 1, Figure S2, Table S2). These spin-state shifts are also lower than those reported for the corresponding *para* isomers and similar to those previously reported for 3-methylthio- and 3-methylamino-benzoic acid.[17] 4-Methylamino-, 4-methylthio- and 4-methyl-benzoic acid all induce 70% shifts to high-spin,[17] consistent with the *para*-substituent being held in close proximity to the heme and resulting in loss of the iron-bound water. By way of contrast, 4-formylbenzoic acid induced a smaller spin-state shift (25%), though this is larger than that induced by 3-formylbenzoic acid (Table 1, Figure S2, Table S2).

Based on the weak binding reported for 3-methoxy-, 3-methylamino- and 3-methylthio-benzoic acid, the substrates investigated here were predicted to bind weakly to CYP199A4.[17] 3-Methylbenzoic acid binds more than 100-fold less tightly than 4-methylbenzoic acid, and 3-methylamino- and 3-methylthio-benzoic acids ($K_d = 89 \mu\text{M}$, compared to 31 and 33 μM , respectively, Table 1, Figure S3). In line with its lower spin-state shift, 4-formylbenzoic acid bound with a lower affinity than the other *para*-substituted benzoic acids. 3-Formylbenzoic acid binds even more weakly than the other *meta*-substituted benzoic acids studied ($K_d = 416 \mu\text{M}$, Figure S3) and ~9-fold more less tightly than the corresponding *para* isomer (Table 1). This data indicates that all the *meta*-substituted benzoic acids fit less well in the CYP199A4 active site

compared to the corresponding *para* isomers. It has been hypothesized that this weaker binding is due to lower complementarity between the enzyme active site and the *meta*-substituted benzoic acids.[19]

The smaller spin-state shifts induced by the *meta*-substituted benzoic acids implied that the substrate is not bound close enough to the heme to efficiently displace the iron-bound water. Both 3-formyl- and 3-methylamino-benzoic acid slightly red shifted the Soret maximum (by 0.5 nm) despite giving type I difference spectra, implying that the heteroatoms of these substituents may interact with a heme-bound water (Supporting information, Figure S2, Table S2). The spin state experiments also suggest that NADH oxidation by CYP199A4 in the presence of these substrates will be slow because electron transfer is gated by displacement of the water and the related changes in spin state and redox potential.[17, 19, 30-33] If these substrates were to bind exclusively with the *meta*-substituent on the side of the benzene ring which is further away from the heme, we would predict little or no product to be formed. Therefore, *in vitro* NADH activity assays were performed to evaluate whether these substrates are metabolized by CYP199A4 (Figure S4). The rates of NADH oxidation and product formation by CYP199A4 are substantially lower with the *meta*-substituted benzoic acids than the equivalent *para* isomers (Table 1). Whereas 3-methoxybenzoic acid is not oxidized by CYP199A4, HPLC and GC-MS analysis revealed the presence of oxidation metabolites arising from CYP199A4-catalyzed oxidation of all the other *meta*-substituted substrates. We have previously reported some of these including 3-aminobenzoic acid and 3-methylsulfinylbenzoic acid, which arise from *N*-dealkylation and sulfoxidation of the nitrogen- and sulfur-containing substrates, respectively.[17]

In vitro turnovers with 3-methylbenzoic acid oxidized NADH at a rate of 78 μM ($\mu\text{M-P450}$)⁻¹ min⁻¹ (henceforth abbreviated to min⁻¹), compared to a rate of 444 min⁻¹ with the

para isomer (Table 1). CYP199A4 has previously been reported to convert the *para* isomer exclusively into 4-hydroxymethylbenzoic acid.[18] CYP199A4 similarly converted 3-methylbenzoic acid into a single product (Figure 3). The mass of the TMS-derivatized product ($m/z = 296.10$) is consistent with the calculated mass of doubly derivatized 3-hydroxymethylbenzoic acid ($m/z = 296.1264$). The product could also conceivably be 4-hydroxy-3-methylbenzoic acid, arising from hydroxylation of the benzene ring. However, the retention time of authentic 4-hydroxy-3-methylbenzoic acid was different to that of the P450 product, discounting the possibility that the substrate was hydroxylated at the *para* aromatic C-H bond of the benzoic acid moiety (Figure S5 and S6). To verify that the product was 3-hydroxymethylbenzoic acid, this compound was chemically synthesized via reduction of 3-formylbenzoic acid with NaBH_4 . The CYP199A4 oxidation product and synthesized 3-hydroxymethylbenzoic acid co-eluted via HPLC (Figure S6). The coupling efficiency for this reaction was 48%, and the product formation rate 38 min^{-1} . CYP199A4 therefore oxidizes 3-methylbenzoic acid at a ~ 10 -fold slower rate than it hydroxylates 4-methylbenzoic acid.

The CYP199A4 system oxidized NADH in the presence of 3-formylbenzoic acid at a rate (10.6 min^{-1}) roughly equal to the leak rate (the NADH oxidation rate in the absence of substrate; $\sim 9 \text{ min}^{-1}$). However, low levels of the carboxylic acid product (isophthalic acid) were detected (Figure S7). The coupling efficiency and product formation rate were only 8% and 0.9 min^{-1} , respectively. CYP199A4 converts 4-formylbenzoic acid into terephthalic acid at a rate of 110 min^{-1} with a coupling efficiency of $\sim 83\%$ (Figure 4). This demonstrates that CYP199A4 is capable of catalyzing efficient aldehyde oxidation. Control reactions with these aldehyde-containing substrates and the others showed that no product was detected when the P450 was omitted from the reaction mixture (Figure 4 and S7).

Finally, we assessed the binding and turnover of the *ortho*-substituted substrate 2-methylbenzoic acid. CYP199A4 displayed weak affinity for 2-methylbenzoic acid ($K_d = 611 \mu\text{M}$) and only a small spin-state shift ($\sim 5\%$) was observed (Table 1, Figure S8). In line with the small spin-state shift induced by this substrate, the NADH oxidation rate was only 13 min^{-1} , barely faster than the leak rate (Table 1). No product was detected by GC-MS or HPLC analysis of the turnover mixture (Figure S9 and S10). This result reaffirms that CYP199A4 does not oxidize *ortho*-substituted benzoic acids.[19]

Minimal levels of H_2O_2 were detected in all CYP199A4-catalyzed oxidations. However, it is important to note here that when the rate of NADH oxidation is slow, as it is with the *meta*-benzoic acid isomers, H_2O_2 generated via uncoupling may be lost before the reaction is complete. Indeed, in control experiments we found that when $\sim 300 \mu\text{M}$ H_2O_2 was incubated with a mixture containing HaPux, HaPuR and CYP199A4 in Tris-HCl buffer (but no NADH, substrate or catalase) or with just the HaPux [2Fe-2S] ferredoxin, almost all H_2O_2 was lost over a 60-minute period (Supporting information, Figure S11). Therefore, this method is unlikely to provide an accurate measurement of the amount of H_2O_2 formed at the end of a slow reaction ($>15 \text{ min}$) and the values obtained are likely to be an underestimate of the actual amount generated (Figure S11). This is an important observation when considering the quantitation of the contribution of different uncoupling pathways in P450 reactions.

3.2 Docking of substrates into the active site of CYP199A4

The ability of CYP199A4 to oxidize all the *meta*-substituted benzoic acids with the exception of 3-methoxybenzoic acid led us to speculate that the orientation of these substrates in the binding pocket of CYP199A4 may differ. We therefore decided to dock the *meta*-substituted benzoic acids into the active site of CYP199A4 to predict their likely binding mode. ICM-Pro software

was used to dock all the substrates (except 3-formylbenzoic acid due to the lower binding affinity of the aldehyde isomers) into the CYP199A4 structure (PDB: 5U6W).[27]

In all instances the top-scoring pose had the *meta*-moiety pointing towards the heme (Figure 5, Supporting information and Table S3). For all the substrates, except 3-methylthiobenzoic acid, an alternative (lower-scoring) pose having this substituent pointing away from the heme was found. For 3-methylthiobenzoic acid the next lowest scoring pose also had the *meta*-substituent pointing towards the heme.

In contrast, when veratric acid was docked into the active site, the top-scoring poses had the *meta*-substituent pointing away from the heme, consistent with the crystal structure (Figure S12). Docking of 2-methoxybenzoic acid was conducted as an example of an *ortho*-substituted substrate and revealed that it preferred to bind with the *ortho*-methoxy group pointing away from the heme (Figure S12).

For 3-methylaminobenzoic acid the methyl carbon was only 4.1 Å from the heme iron and the sulfur of 3-methylthiobenzoic acid and methyl group of 3-methylbenzoic acid were at distances of 4.2 Å and 4.0 Å, respectively (Figure 5 and Table S3). This could rationalize why these substrates were oxidized at these positions. However, the top-scoring pose of 3-methoxybenzoic acid, which is not oxidized by CYP199A4, also had the *meta*-substituent pointing towards the heme, with the methyl group 4.0 Å from the iron. Another important observation from these docking experiments was that the position of the benzene moiety of the *meta*-substituted substrates was slightly different to that observed in the X-ray crystal structures of the *para*-substituted benzoic acids (Figure S13). To more satisfactorily answer the question of how these different *meta*-substituted benzoic acids bind, we decided to solve crystal structures of these substrates bound to CYP199A4.

3.3 X-ray crystal structures of CYP199A4 with *meta*-substituted benzoic acids

CYP199A4 was co-crystallized with 3-methoxy-, 3-methylthio-, 3-methylamino-, 3-methyl- and 4-methyl-benzoic acid (Figure S14) and X-ray diffraction data were collected at the MX1 or MX2 beamlines at the Australian Synchrotron at 100 K.[34, 35] Crystal structures of these protein-ligand complexes were solved at resolutions of 1.60-1.89 Å (PDB codes 6PQ6, 6PQD, 6PRR, 6PQW and 6PQS). Data collection and refinement statistics are presented in Table S4. The RMSD between C α atoms of the superimposed structures is ≤ 0.178 Å, indicating that the overall protein fold is virtually identical for all five complexes. These crystal structures also had a similar fold to previously reported structures of CYP199A4 (Figure S15). Clear electron density corresponding to the substrate permitted modelling of each ligand into each structure (Figure 6). These substrates all bind with the *meta*-substituent pointing towards the heme in agreement with the docking studies. Overall this rationalizes why, with the exception of 3-methoxybenzoic acid, oxidation occurred at this position.

The crystal structures of the *meta*-substituted benzoic acids also demonstrated that the water ligand was maintained with a degree of occupancy at the sixth coordination position (21-90%; Table 1, Figure 6). The crystal structures also provide evidence that the water ligand and the N, S, or O atoms of the *meta*-substituent of 3-methylamino-, 3-methylthio- and 3-methoxybenzoic acid, were interacting with this water molecule. The small red shift observed in the Soret maximum with 3-methylaminobenzoic acid may result from such an interaction. 3-Methylbenzoic acid, on the other hand, which has a more hydrophobic substituent, largely displaced the water in the 3-methylbenzoate-CYP199A4 crystal structure (water occupancy = 21%).

The occupancy of the water ligand in the other *meta*-substituted benzoic acid crystal structures closely correlated with the measured fraction of high/low-spin heme (Table 1). For example, the 3-methylaminobenzoic acid-CYP199A4 complex is 90% low-spin and the occupancy of the water ligand in the crystal structure is also 90%. The 3-methylthiobenzoic acid-CYP199A4 complex is 70% low-spin and the refined water occupancy is 77%, while the 3-methoxybenzoic acid-CYP199A4 complex is 60% low-spin and the water ligand occupancy is 50%. 4-Methylbenzoic acid, on the other hand, induces a larger spin-state shift to the high-spin form (70%) and a water ligand is not observed in the crystal structure. The *para*-substituted benzoic acids displace the water molecule and induce larger shifts to the high-spin ferric state.[17] It should be noted that partial spin-state shifts and water occupancy in the crystal structure have also been reported with different substrates and P450_{cam} (CYP101A1).[36, 37]

To ascertain why we observe such an acute difference in the binding and activity of *meta*- versus *para*-substituted benzoic acids, we next compared the structures of the different substrate-bound analogues. When the structures of the *meta*-substituted benzoate-bound CYP199A4 were superimposed with the equivalent *para* substrate-bound structures, only subtle differences in the positions of active-site residues were observed (Figures 7 and 8). However, the benzene ring of the *meta* isomers are all shifted up away from the heme relative to same moiety of the equivalent *para*-substituted benzoic acids. This was in agreement with what was observed in the docking experiments and the positions of 3-methoxy-, 3-methylthio-, 3-methylamino-, and 3-methylbenzoic acid in the crystal structure agreed with the top-scoring docked poses. When the position of the substrates was compared in the X-ray and the modelled structures all the atoms, including those of the substituent, were found to align (Figure S16). This suggests that energy minimization modelling could be a valuable tool in understanding the binding of different

benzoic acids and related substrates in this family of P450 enzymes. This movement of the substrate, when the position of substitution changes, seems to occur to avoid a steric clash between the *meta*-substituent and the heme (Figure S17).

A common feature of all of the structures with *meta*-substituted isomers is that there is more space directly above the heme center as the substituent is more offset from the area directly above the iron (Figures 7 and 8). This is consistent with the partial occupancy of the 6th ligand in these structures compared to the *para*-substituted benzoic acid counterparts, in which this water was not present. Based on the productive catalytic activity and NADH oxidation rates observed (Table 1) it appears that this water molecule can be displaced enabling the enzyme to bind dioxygen and enabling turnover of the catalytic cycle albeit at a lower rate than for the equivalent *para* isomer.

While the heme iron in 4-methylthiobenzoate-bound CYP199A4 is five-coordinate, when 3-methylthiobenzoate is bound a high-occupancy water ligand is present (Fe–O distance 2.3 Å; Table S5). The sulfur of 3-methylthiobenzoic acid is 3.3 Å from this water. The methylthio moiety of 3-methylthiobenzoic acid is rotated out of the plane of the benzene ring by 13° and that of 4-methylthiobenzoic acid is rotated by 34°. In both structures, the methylthio moiety is held close to the heme, though the sulfur of 3-methylthiobenzoic acid is held nearer to the iron than that of 4-methylthiobenzoic acid (4.3 vs 5.0 Å; Table S5). The methyl group of 3- and 4-methylthiobenzoic acid is also held in close proximity to the heme iron (4.4 vs 4.5 Å). When comparing the structures, we wanted to ascertain how close the substrate would be to the oxygen of the reactive Cpd I intermediate (as well as the heme iron). It has been calculated that the Fe=O bond length of Cpd I is 1.62 Å.[38, 39] This is in good agreement with the measured Cpd I Fe=O bond length using EXAFS, which was found to be 1.67 Å.[40] Therefore, to estimate the likely

distances between the ferryl oxygen of Cpd I and the relevant atoms of the bound substrate, an oxygen atom was modelled 1.62 Å above the heme iron directly opposite the Fe-S bond and we assumed that the position of the substrate is not significantly changed after oxygen binding and activation (Table S5). The measured distance between the Cpd I oxygen and the sulfur of 3-methylthiobenzoic acid is only 3.2 Å. This is slightly shorter than the distance between the Cpd I oxygen and the sulfur of 4-methylthiobenzoic acid (3.4 Å). For the *meta*-methylthio isomer the $\text{Fe}^{\text{IV}}=\text{O}-\text{S}$ angle is 126.2°, whereas for the *para* isomer the angle is larger (162.5°).

In the crystal structure of 3-methylaminobenzoate-bound CYP199A4, a water ligand to the heme is present at 90% occupancy (Fe–O distance 2.2 Å; Table S5), and this water is hydrogen-bonded to the substrate nitrogen (3.1 Å). The nitrogen and methyl group of 3-methylaminobenzoic acid are 4.2 and 4.4 Å from the heme iron, respectively. The distance between the Cpd I oxygen and methyl carbon of 3-methylaminobenzoic acid is 3.1 Å, which is similar to that of the methyl carbon of 4-methylaminobenzoic acid (3.0 Å). The $\text{Fe}^{\text{IV}}=\text{O}-\text{C}_{\text{Me}}$ angle was similar for both the *meta* and *para* isomers (129.9 vs 124.6°). The nitrogen heteroatom of the *para* isomer had a larger $\text{Fe}^{\text{IV}}=\text{O}-\text{N}$ angle than 3-methylaminobenzoic acid (148.6 vs 126.8°; Table S5).

In both the structures of 3-methoxybenzoic acid- and 4-methoxybenzoic acid-bound CYP199A4, the methoxy groups are held above the heme. The O and C_{Me} of 3-methoxybenzoic acid are equidistant from the heme iron (4.4 Å), and the methoxy group is rotated 20° out of the plane of the benzene ring. In contrast, the methoxy group of 4-methoxybenzoic acid lies almost in the plane of the benzene ring, and the C_{Me} is closer to the iron (4.1 Å) with the oxygen being further away (5.2 Å). When the ferryl oxygen of Cpd I was considered, the C_{Me} and O of 3-methoxybenzoic acid are both 3.1 Å from the Cpd I oxygen, whereas the C_{Me} and O of 4-

methoxybenzoic acid are 2.7 and 3.6 Å from the ferryl oxygen. The $\text{Fe}^{\text{IV}}=\text{O}-\text{C}_{\text{Me}}$ angle is similar for both 3-methoxy- and 4-methoxy-benzoic acid (136.3 vs 140.7°).

The methyl group of both 3-methyl- and 4-methyl-benzoic acid is held close to the heme iron (4.4 vs 4.5 Å). A low-occupancy water ligand is present in the 3-methylbenzoate-CYP199A4 structure with a Fe–O distance of 2.5 Å, but no water ligand was detected in the 4-methylbenzoic acid structure. The benzene ring of 3-methylbenzoic acid is shifted away from the heme relative to that of 4-methylbenzoic acid, though the carboxylate groups are held in similar positions. The methyl of 3-methylbenzoic acid is held slightly further from the ferryl oxygen than that of 4-methylbenzoic acid (3.2 vs 3.0 Å). The $\text{Fe}^{\text{IV}}=\text{O}-\text{C}_{\text{Me}}$ angle is larger for 4-methylbenzoic acid than for 3-methylbenzoic acid (154.4 vs 130.4°).

3.4 Substrate binding and catalytic activity of bulkier *meta*-substituted benzoic acids

The crystal structures of 3-methoxy-, 3-methylamino-, and 3-methylthio-benzoate-bound CYP199A4 revealed that although the *meta*-substituent was positioned close to the heme, the water ligand was not fully displaced. We hypothesized that substrates with bulkier *meta*-substituents would more efficiently displace the water ligand and that CYP199A4 would display higher oxidation activity towards them. We selected 3-ethoxybenzoic acid and 3-*tert*-butylbenzoic acid and docked these into the active site of CYP199A4 (Figure S18). The top-scoring poses have the *meta*-substituent positioned on the heme side of the substrate (Table S6). However, docking 3-*tert*-butylbenzoic acid into CYP199A4 resulted in a steric clash between the *tert*-butyl group and a heme nitrogen. This implies that this substrate may be too bulky to fit into the active site.

We subsequently measured the spin-state shifts induced by these substrates. The minimal spin-state shift (<5%) induced by 3-*tert*-butylbenzoic acid implies that the bulky *tert*-butyl group hinders binding (Table 2, Figure S19). The *para* isomer, 4-*tert*-butylbenzoic acid, bound to CYP199A4 with a 90% shift to the high spin form, but had lower binding affinity ($39 \pm 2 \mu\text{M}$, Figure S20) than the other *para*-substituted benzoic acid substrates, such as 4-methoxybenzoic acid (Table 1). In contrast to the other *meta*-substituted benzoic acids we tested, 3-ethoxybenzoic acid induced an almost complete (85%) shift of the ferric iron to the HS state, indicating efficient removal of the 6th water ligand (Figure S19). This spin-state shift is larger than that induced by certain *para*-substituted benzoic acids (Table 1). However, 3-ethoxybenzoic acid was found to bind substantially less tightly than the *para* isomer ($K_d = 82 \pm 2$ vs $0.2 \pm 0.1 \mu\text{M}$, Figure S21).

There was no increase in the NADH oxidation rate above the leak rate (as defined in Table 1) in the presence of 3-*tert*-butylbenzoic acid. No product was detected by HPLC analysis of the reaction mixture (Figure S22). These observations are evidence that 3-*tert*-butylbenzoic acid does not bind to CYP199A4. The corresponding *para* isomer is hydroxylated by CYP199A4 at a rate of 227 min^{-1} with 100% coupling efficiency and chromatographic analysis indicated it was oxidized to a single metabolite (Table 2, Figure S23 and S24). The GC-MS of the TMS-derivatized turnover confirmed the formation of a single alcohol product ($m/z = 338.25$ vs expected 338.17, Figure S24). The product was prepared using a whole-cell oxidation system[21] and, after conversion to the methyl ester, purified and characterized by NMR as 4-(1,1-dimethyl-2-hydroxyethyl)benzoic acid (Supporting information, Figure S25). This arose from methyl hydroxylation due to the absence of any benzylic hydrogens within the substrate.

On the other hand, the CYP199A4 system oxidized NADH upon addition of 3-ethoxybenzoic acid at a moderately fast rate (208 min^{-1}). Whereas CYP199A4 is unable to

oxidize 3-methoxybenzoic acid, the enzyme did catalyze the *O*-deethylation of 3-ethoxybenzoic acid. The coupling efficiency and product formation rate were calculated to be 66% and 138 min^{-1} , and only 7% of the NADH consumed was measured as being channeled into H_2O_2 production (Table 2). A single product, 3-hydroxybenzoic acid, was detected, indicating that no hydroxylation occurred at the less reactive terminal CH_3 group (Figure 9). CYP199A4 converts the *para* isomer, 4-ethoxybenzoic acid, into 4-hydroxybenzoic acid at a rate of 527 min^{-1} and with a coupling efficiency of 100%.[17] CYP199A4 therefore still exhibits substantially higher activity towards 4-ethoxybenzoic acid than 3-ethoxybenzoic acid, which again demonstrates the generality of the preference of this enzyme for *para*- over *meta*-substituted benzoic acids.

3.5 Crystal structure of 3-ethoxybenzoate-bound CYP199A4

We obtained a crystal structure of 3-ethoxybenzoate-bound CYP199A4 which was solved at 2.37 Å resolution (PDB code 6PRS, Figure 10, Table S7, Figure S26). It confirms that the *meta*-substituent points down towards the heme and that the 6th iron water ligand is displaced, in agreement with the 85% shift to high spin and the docked structure (Figure S27). There is a change in the ($\text{CH}_2\text{-O-C}_3\text{-C}_4$, Table S8) dihedral angle of the larger ethoxy substituent compared to that of the methoxy substituent which results in the methylene group being held more over the heme iron than the methyl group of the methoxy substituent (Figure 10, Table S8).

Both the methylene (CH_2) and methyl carbons (CH_3) of the *meta*-ethoxy substituent are held in close proximity to the heme iron (3.6 and 4.2 Å, respectively, Table S8). When the distances were measured between the substrate and the hypothetical position of the Cpd I oxygen, the methylene carbon (CH_2) is also held closer than CH_3 (2.1 vs 2.7 Å). The $\text{Fe}^{\text{IV}}=\text{O}-\text{CH}_2$ and $\text{Fe}^{\text{IV}}=\text{O}-\text{CH}_3$ angles are 155.3 and 147.7°, respectively. C-H bond abstraction

occurs solely at CH₂ by CYP199A4, with no hydroxylation being observed at the methyl group (Table S8). This can be rationalized by CH₂ being the closest carbon to the iron and the weaker strength of the secondary methylene C-H bonds (especially those adjacent to oxygen)[1] compared to the primary C-H bonds of the terminal methyl group. Overall this structure highlights that certain *meta*-substituted benzoic acid substrates can efficiently bind to the active site and thus remove the 6th water ligand of CYP199A4 and that this enables more efficient catalytic oxidation.

4. Discussion

The cytochrome P450 family of enzymes catalyze many reactions which are important in xenobiotic catabolism and secondary metabolite formation. The majority of cytochrome P450s are thought to follow the consensus catalytic cycle observed with P450cam, as does CYP199A4. The reactive Cpd I intermediate has been calculated to have similar reactivity across different forms. We hypothesize that the observations reported here, with CYP199A4, would be applicable to many other P450 enzymes. However, it is important to note that an increasing number of P450 enzymes which catalyze varied reactions are being discovered, highlighting the complex roles they play in metabolism and their potential for mechanistic and function diversification.

CYP199A4 displays a strong preference for benzoic acid substrates with a *para*-substituent. Despite having no activity towards 3-methoxybenzoic acid and 3-*tert*-butylbenzoic acid, this enzyme nevertheless is capable of oxidizing certain *meta*-substituted benzoic acids. The binding mode of *meta*-substituted benzoic acids revealed by the crystal structures has the *meta*-substituent positioned close to the heme, not directed away as previously observed in the disubstituted veratric acid structure. All the *meta*-substituted benzoic acid substrates studied here bind with the benzene ring and carboxylate held in virtually identical positions (Figure S28). While the distance of the *meta*-substituent to the heme iron/Cpd I model oxygen is often comparable to that of the equivalent position in the *para*-substituted species, the altered position of the benzene ring and change in substitution pattern results in there being more space directly above the heme iron in the former set of substrates.

Within each pair of *meta* and *para* isomers, the extent of the spin-state shift correlated with the NADH oxidation rate, suggesting that this step is somewhat gated by displacement of the

water. The observation of partial occupancy of the 6th heme iron water ligand in the crystal structures of the *meta*-substituted benzoic acid substrates would result in a lower spin-state shift to the high-spin form and slower turnover of the catalytic cycle which is consistent with the first electron reduction being the rate determining step in the catalytic cycle. This is consistent with previous observations with this enzyme and other related class I bacterial systems such as P450_{cam}. However, the extent of the spin-state shift does not always directly correlate with NADH oxidation rate or product formation of CYP199A4 across all these substrates. Another observation from this work is that hydrogen peroxide is consumed by the P450 system (by the [2Fe-2S] ferredoxin) which can make its quantification in slow reactions difficult or unfeasible.

The binding orientations of 3-methylamino-, 3-methylthio- and 3-methyl-benzoic acid enable one to explain why these compounds can be oxidized by CYP199A4. Despite their lower activities, the coupling efficiency of these oxidation reactions were comparable to the equivalent *para* isomers, in line with their similar distances and orientations relative to the heme iron. The crystal structures and molecular docking experiments with the *meta*-substituted benzoic acids provide excellent insight into why *meta*-substituted and larger substrates are not as good substrates for the CYP199A4 enzyme. These would be beneficial in the future design of variant forms of CYP199A4 to improve the oxidation activity of these and similar substrates. If more space was generated at the top of the active site by mutating the hydrophobic residues in this region (F182, F185 and V181, Figure 7 and 8) the *meta* isomer could bind in an orientation which more efficiently displaces the 6th water ligand and improve the activity. Such changes may also enhance the binding and activity of larger substrates such as substituted naphthoic acids. However, such amino acid substitutions could result in the substrate binding with the *meta*-

substituent pointing into the space generated, which as with veratric acid, would place them further away from the heme.

The results with the methylthio and methylamino analogues are perhaps unsurprising given their higher reactivity. The sulfur of 3-methylthiobenzoic acid is in range for oxidation by the ferryl Cpd I reactive intermediate. For sulfur-containing substrates, Cpd I reacts with the heteroatom via direct oxygen transfer, or via single electron transfer followed by oxygen rebound (we note that other intermediates of the P450 catalytic cycle have been proposed to catalyze sulfoxidations).[8, 12] The observation that 3-methylaminobenzoic acid is oxidized by CYP199A4 but 3-methoxybenzoic acid is not requires more detailed analysis. These substrates have virtually identical orientations in the two crystal structures and the distances of the methyl group and heteroatom of each of the substituents to the heme iron are comparable. The reactivity of methylamines and the potential of *N*-dealkylation to occur via the alternative single electron transfer (SET) pathway - which would involve initial reaction at the closer nitrogen atom- rather than a hydrogen abstraction (HAT) mechanism, are likely to be important here.[12]

The CYP199A4 oxidations of 3-methoxy-, 3-ethoxy- and 3-methyl-benzoic acids (and their *para* isomer equivalents) would all proceed via an initial C-H bond abstraction. Factors that would affect whether Cpd I can abstract a hydrogen are the distance of the C-H bond from the ferryl oxygen, the C-H bond strength and the angle of approach of the substrate (Figure 11).[41] The C-H bond strength should not be problematic in any of these oxidations. 3-Methoxybenzoic acid is not oxidized by CYP199A4 which indicates that it must be inappropriately positioned in the active site for C-H bond activation despite turnover of the catalytic cycle (as judged by the NADH oxidation rate, Table 1). Whereas 3-methoxybenzoic acid is not oxidized, CYP199A4 is able to dealkylate the bulkier 3-ethoxybenzoic acid and hydroxylate the methyl C-H bonds of 3-

methylbenzoic acid. There is a significant change in the conformation of the *meta*-substituent in the ethoxy structure which pushes the methylene group closer to the heme iron and changes its orientation with respect to where the reactive oxygen atom would be located.

It has been proposed that a linear arrangement of the C-H-O atoms is ideal for hydrogen abstraction[42] and that it is optimal for the substrate to approach the $\text{Fe}^{\text{IV}}=\text{O}$ species at an angle of $\sim 130^\circ$ (Fe-O-C angle).[41] In the 4-methoxybenzoic acid crystal structure, the methoxy group and its hydrogens point towards the heme. However, the methoxy group of 3-methoxybenzoic acid is oriented such that the hydrogens could not form a linear C-H-O arrangement with the iron-oxo atom of Cpd I (Figure 11). The change in the conformation in the 3-ethoxybenzoic acid side chain results in the C-H bonds on the methylene pointing towards the iron-oxo atom in contrast to those on the respective methyl groups of 3-methoxy and 3-ethoxybenzoic acid (Figure 11).

It is important to note that while the benzoic acid moiety of the substrates seems to be tightly confined within the active site, we cannot rule out that the orientation of the substituent observed in the crystal structures may not be the same as in solution and its position relative to the heme iron may be altered once oxygen binds or in the presence of the reactive Cpd I intermediate. However, the crystal structure and biochemical data reported are in agreement with theoretical work related to this family of enzymes (Figure 11).[41, 42] The presence of heteroatoms, which can potentially interact with ligands bound to the heme iron, could affect heme reduction or oxygen binding and activation thus altering different steps of the catalytic cycle. In the 4-methoxy- and 3-ethoxy-benzoic acid structures, the methyl or methylene groups are held significantly closer to the heme than the heteroatom. In contrast, in the 3-methoxybenzoic acid structure, the methyl group and oxygen are held at almost equal distances

from the iron. There is a possibility that this close approach of the oxygen heteroatom could interfere with substrate oxidation.

5. Conclusions

Overall the crystal structures of CYP199A4 provide detailed information on the factors which promote efficient catalytic monooxygenase activity for a range of substrates. In summary, this study reveals that CYP199A4 does oxidize certain *meta*-substituted benzoic acids and provides an understanding of why the activity towards these substrates is substantially lower than those of the *para* isomers. The crystal structures of CYP199A4 with these substrates bound in the active site provide important insights into the structural parameters, in this system, which define whether C-H bond abstraction (or other P450-catalyzed activities) can occur efficiently. This provides experimental validation of theoretical studies on other forms of these metalloenzymes and, accordingly these observations could be extended to other P450s. This work highlights that relatively minor changes in the substrate and its position in the enzyme active site relative to the heme can have a dramatic effect on the activity of oxidation. Some P450 activities, for example the C-H abstraction that results in *O*-dealkylation, were more sensitive to these parameters than others which involve more reactive substituents (*N*-dealkylation and *S*-oxidation). These considerations are essential when designing or evolving enzyme systems for efficient biocatalytic reaction or understanding and interpreting the reactivity of computational docking or structural studies of these enzymes. The range of monooxygenase reactions catalyzed by CYP199A4 (Figure S1) infer that this system could be used to investigate other types of cytochrome P450-catalyzed activities, which are important in drug metabolism or natural product synthesis.

Abbreviations

AcCN, acetonitrile; BA, benzoic acid; Cpd I, iron(IV)–oxo porphyrin radical cation intermediate; CYP, cytochrome P450 enzyme; CYP199A4, cytochrome P450 enzyme from *Rhodopseudomonas palustris* strain HaA2; FAD, flavin adenine dinucleotide; FEM, feature enhanced maps; GC-MS, gas chromatography-mass spectrometry; HaPuR, FAD containing flavoprotein from *Rhodopseudomonas palustris* strain HaA2 (Uniprot: Q2ITV9); HaPux; [2Fe-2S] ferredoxin from *Rhodopseudomonas palustris* strain HaA2 (Uniprot: Q2IU01); HAT, hydrogen abstraction; HPLC, high performance liquid chromatography; HS, high spin; NADH, reduced nicotinamide adenine dinucleotide; PFR, product formation rate; SET, single electron transfer; TFA, trifluoroacetic acid; TMS, trimethylsilyl; Tris, tris(hydroxymethyl)aminomethane.

Supplementary Information. Experimental methods, additional substrate binding parameters and catalytic turnover activity data for CYP199A4, details of the distances and angles of pertinent substrate atoms from the heme iron in the crystal structures of CYP199A4, spin state shift assays, dissociation constant analysis, examples of NADH oxidation assays, H₂O₂ assays, HPLC and GC-MS analysis of the *in vitro* turnovers, NMR spectra and details of isolated products, computational docking studies, crystallography details and structures of the C α chains of different CYP199A4 crystal structures.

Acknowledgements

This work was in part supported by ARC grant DP140103229 (to JJDV and SGB). SGB acknowledges the ARC for a Future Fellowship (FT140100355). The authors also acknowledge the award of an Australian Government Research Training Program Scholarships (PhD to TC

and MPhil to MNP and RRC). We would like to thank the scientists at MX1 and MX2 beamline for help with data collection. We acknowledge ANSTO for financial support and in providing the facility used in this work.

Journal Pre-proof

References

- [1] P.R. Ortiz de Montellano, *Chem. Rev.*, 110 (2010) 932-948.
- [2] T.L. Poulos, *Chem. Rev.*, 114 (2014) 3919-3962.
- [3] P.R. Ortiz de Montellano, in: P.R. Ortiz de Montellano (Ed.), *Cytochrome P450: Structure, Mechanism, and Biochemistry*, Springer International Publishing, Cham, 2015, pp. 111-176.
- [4] F. Ogliaro, S.P. de Visser, S. Cohen, P.K. Sharma, S. Shaik, *J. Am. Chem. Soc.*, 124 (2002) 2806-2817.
- [5] J. Rittle, M.T. Green, *Science*, 330 (2010) 933-937.
- [6] P.K. Sharma, S.P. De Visser, S. Shaik, *J. Am. Chem. Soc.*, 125 (2003) 8698-8699.
- [7] B. Meunier, S.P. de Visser, S. Shaik, *Chem. Rev.*, 104 (2004) 3947-3980.
- [8] S. Shaik, S. Cohen, Y. Wang, H. Chen, D. Kumar, W. Thiel, *Chem. Rev.*, 110 (2010) 949-1017.
- [9] X. Huang, J.T. Groves, *J. Biol. Inorg. Chem.*, 22 (2017) 185-207.
- [10] M.J. Cryle, J.E. Stok, J.J. De Voss, *Aust. J. Chem.*, 56 (2003) 749-762.
- [11] P.R. Ortiz de Montellano, J.J. De Voss, *Nat. Prod. Rep.*, 19 (2002) 477-493.
- [12] F.P. Guengerich, *Chem. Res. Toxicol.*, 14 (2001) 611-650.
- [13] E.M. Isin, F.P. Guengerich, *Biochim. Biophys. Acta*, 1770 (2007) 314-329.
- [14] J.B. Behrendorff, W. Huang, E.M. Gillam, *Biochem. J.*, 467 (2015) 1-15.
- [15] M.T. Reetz, *J. Am. Chem. Soc.*, 135 (2013) 12480-12496.
- [16] V.B. Urlacher, M. Girhard, *Trends Biotechnol.*, 30 (2012) 26-36.
- [17] T. Coleman, S. Wong, M. Podgorski, J. Bruning, J. De Voss, S. Bell, *ACS Catal.*, 8 (2018) 5915-5927.

- [18] S.G. Bell, R. Zhou, W. Yang, A.B. Tan, A.S. Gentleman, L.L. Wong, W. Zhou, *Chemistry*, 18 (2012) 16677-16688.
- [19] T. Coleman, R.R. Chao, J.B. Bruning, J. De Voss, S.G. Bell, *RSC Adv.*, 5 (2015) 52007 - 52018.
- [20] R.R. Chao, I.C.K. Lau, J.J. De Voss, S.G. Bell, *ChemCatChem*, 8 (2016) 3626-3635.
- [21] S.G. Bell, A.B. Tan, E.O. Johnson, L.L. Wong, *Mol. Biosyst.*, 6 (2010) 206-214.
- [22] S.G. Bell, W. Yang, A.B. Tan, R. Zhou, E.O. Johnson, A. Zhang, W. Zhou, Z. Rao, L.L. Wong, *Dalton Trans.*, 41 (2012) 8703-8714.
- [23] T. Coleman, R.R. Chao, J. De Voss, S.G. Bell, *Biochim. Biophys. Acta Proteins Proteomics*, 1864 (2016) 667-675.
- [24] R.R. Chao, J.J. De Voss, S.G. Bell, *RSC Adv.*, 6 (2016) 55286-55297.
- [25] S.G. Bell, F. Xu, I. Forward, M. Bartlam, Z. Rao, L.-L. Wong, *J. Mol. Biol.*, 383 (2008) 561-574.
- [26] S.G. Bell, N. Hoskins, F. Xu, D. Caprotti, Z. Rao, L.L. Wong, *Biochem. Biophys. Res. Commun.*, 342 (2006) 191-196.
- [27] R. Abagyan, M. Totrov, D. Kuznetsov, *J. Comput. Chem.*, 15 (1994) 488-506.
- [28] A.J.W. Orry, R. Abagyan, in: A.J.W. Orry, R. Abagyan (Eds.), *Homology Modeling: Methods and Protocols*, Humana Press, Totowa, NJ, 2012, pp. 351-373.
- [29] R. Abagyan, A.J.W. Orry, E. Raush, M. Totrov.
- [30] M.T. Fisher, S.G. Sligar, *J. Am. Chem. Soc.*, 107 (1985) 5018-5019.
- [31] S.G. Sligar, *Biochemistry*, 15 (1976) 5399-5406.
- [32] M.J. Honeychurch, A.O. Hill, L.L. Wong, *FEBS Lett.*, 451 (1999) 351-353.

- [33] S.G. Bell, J.H. McMillan, J.A. Yorke, E. Kavanagh, E.O. Johnson, L.L. Wong, *Chem. Commun.*, 48 (2012) 11692-11694.
- [34] D. Aragao, 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.*, 25 (2018) 885-891.
- [35] 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.*, 22 (2015) 187-190.
- [36] R. Raag, T.L. Poulos, *Biochemistry*, 30 (1991) 2674-2684.
- [37] R. Raag, T.L. Poulos, *Biochemistry*, 28 (1989) 917-922.
- [38] R. Lonsdale, J. Olah, A.J. Mulholland, J.N. Harvey, *J. Am. Chem. Soc.*, 133 (2011) 15464-15474.
- [39] J.C. Schoneboom, H. Lin, N. Reuter, W. Thiel, S. Cohen, F. Ogliaro, S. Shaik, *J. Am. Chem. Soc.*, 124 (2002) 8142-8151.
- [40] C.M. Krest, A. Silakov, J. Rittle, T.H. Yosca, E.L. Onderko, J.C. Calixto, M.T. Green, *Nat. Chem.*, 7 (2015) 696-702.
- [41] R. Lonsdale, K.T. Houghton, J. Žurek, C.M. Bathelt, N. Foloppe, M.J. de Groot, J.N. Harvey, A.J. Mulholland, *J. Am. Chem. Soc.*, 135 (2013) 8001-8015.
- [42] J.C. Schöneboom, S. Cohen, H. Lin, S. Shaik, W. Thiel, *J. Am. Chem. Soc.*, 126 (2004) 4017-4034.

Table 1. Binding data and catalytic activity of CYP199A4 towards *meta*-substituted benzoic acids. Turnover data for *para*-substituted benzoic acid substrates is included for comparison. Values given are the mean $\pm$ SD, with $n \geq 3$. The NADH oxidation rate (N) and product formation rate (PFR) are given in units of $\mu\text{M} (\mu\text{M-P450})^{-1} \text{ min}^{-1}$ and were determined in the presence of 1 mM substrate (see Supporting Information for experimental details). C is the coupling efficiency. The NADH leak rate of the system in the absence of a substrate was $\sim 9 \text{ min}^{-1}$.

Substrate	% HS (% H ₂ O) ^a	K _d (μM)	N (min^{-1})	PFR (min^{-1})	C (%) ^d
3-methoxyBA[17, 19]	40% (50%)	69 $\pm$ 2	498 $\pm$ 5	— ^b	— ^b
4-methoxyBA[19, 22]	$\geq 95\%$ (0%)	0.22 $\pm$ 0.02	1340 $\pm$ 28	1220 $\pm$ 120	91 $\pm$ 2
3-methylaminoBA[17]	10% ^e (90%)	31 $\pm$ 1	255 $\pm$ 2	175 $\pm$ 1	69 $\pm$ 1
4-methylaminoBA[17]	70% (0%)	1.6 $\pm$ 0.07	923 $\pm$ 200	669 $\pm$ 15	64 $\pm$ 2
3-methylthioBA[17]	30% (77%)	33 $\pm$ 0.5	56 $\pm$ 0.2	37 $\pm$ 1	66 $\pm$ 2
4-methylthioBA[17]	70% (0%)	2.3 $\pm$ 0.3	1430 $\pm$ 180	1180 $\pm$ 130	83 $\pm$ 3
3-methylBA	30% (21%)	89 $\pm$ 3	78 $\pm$ 1	37.8 $\pm$ 0.6	48 $\pm$ 1
4-methylBA[18]	70% (0%)	0.66 $\pm$ 0.05	444 $\pm$ 8	397 $\pm$ 22	89 $\pm$ 4
3-formylBA	10% ^e (-)	416 $\pm$ 25	10.6 $\pm$ 0.2 ^c	0.9 $\pm$ 0.1	8 $\pm$ 1
4-formylBA	25% (-)	48 $\pm$ 2	134 $\pm$ 1	110 $\pm$ 2	83 $\pm$ 1

^a The % occupancy of the 6th ligand water molecule in the X-ray crystal structure is given in brackets (where known). ^b No product was formed. ^c The NADH oxidation rate reported is the average rate over the first 15 minutes. ^d The coupling efficiency of the reaction is the proportion of the NADH reducing equivalents which are converted into oxidized metabolites. The level of H₂O₂ formed during each turnover was found to be $\leq 4\%$ in all instances. ^e 3-Methylamino- and 3-formyl-benzoic acid, which induce 10% type I switches to the HS form, both cause the Soret maximum to red shift (0.5 nm, Table S2, Figure S2).

The data for 2-methylbenzoic acid was also measured but no product formation was observed (%HS, $\sim 5\%$; K_d, 611 μM ; and NADH, 13 $\mu\text{M} (\mu\text{M-P450})^{-1} \text{ min}^{-1}$).

Table 2. *In vitro* turnover data for WT CYP199A4 with bulky *meta*-substituted benzoic acid substrates. The details are as defined in Table 1. Turnover data for CYP199A4 with the corresponding *para* isomers, 4-ethoxybenzoic acid which was previously reported,[17] is included for comparison.

Substrate	% HS	K_d (μ M)	N (min^{-1})	PFR (min^{-1})	C (%)
3-ethoxyBA	85	82 ± 2	208 ± 5	138 ± 6	66 ± 1^d
4-ethoxyBA[17]	95	0.2 ± 0.1	527 ± 10	527 ± 10	100 ± 8
3- <i>tert</i> -butylBA	<5%	— ^a	8.3^b	— ^c	— ^c
4- <i>tert</i> -butylBA	90	39 ± 2	227 ± 4	227 ± 32	100 ± 13

a The lack of any appreciable spin-state shift prevented measurement of binding affinity. *b* The NADH oxidation rate reported is the average rate over the first 15 minutes. *c* No product detected. *d* The highest levels of H₂O₂ production were observed with 3-ethoxybenzoic acid and was $7 \pm 0.4\%$ compared to $\leq 1\%$ for the 4-ethoxy isomer.[17]

Figure 1. Binding modes of 4-methoxybenzoic acid (cyan; 4DO1) and veratric acid (yellow; 4EGN) in the binding pocket of CYP199A4. F182 rotates to accommodate the *meta*-methoxy group of veratric acid, which points away from the heme. A chloride ion (purple sphere) is also bound at the entrance to the active site. This chloride is thought to stabilize the closed conformation of the enzyme which prevents excess water entering the active site.[22, 25] Hydrophilic interactions are represented by dashed lines and the distances are all between 2.5-3.1 Å.

Figure 2. The *meta*-substituted benzoic acid substrates investigated with CYP199A4. Several of the *para*-substituted isomers have been studied previously while others were investigated in this work.[17-19]

Figure 3. GC-MS analysis of the CYP199A4 *in vitro* reaction with 3-methylbenzoic acid. In black is the *in vitro* turnover, and in gray is a control reaction omitting the P450. Derivatized 3-methylbenzoic acid appears at 5.1 min and the product, 3-hydroxymethylbenzoic acid, appears at 12.3 min. The chromatograms have been offset along both the x- and y-axes for clarity.

Figure 4. HPLC analysis of the CYP199A4 *in vitro* reaction with 4-formylbenzoic acid. In black is the *in vitro* turnover and in gray is a control reaction performed without P450. Authentic terephthalic acid (dashed chromatogram) co-eluted with the enzyme product at 6.1 min. The substrate appears at 8.9 min. Gradient: 20-95% AcCN in H₂O with 0.1% TFA. Detection wavelength: 254 nm.

Figure 5. Benzoic acids docked into CYP199A4 (PDB ID: 5U6W). The docked substrates are **(a)** 3-methylaminobenzoic acid, **(b)** 3-methylthiobenzoic acid, **(c)** 3-methylbenzoic acid and **(d)** 3-methoxybenzoic acid. The top-scoring pose is shown in **green**, and lower-scoring poses in **orange**.

Figure 6. Crystal structures of **(a)** 3-methoxy-, **(b)** 3-methylamino-, **(c)** 3-methylthio-, **(d)** 3-methyl- and **(e)** 4-methyl-benzoate-bound CYP199A4. The 2mFo-DFc electron density of the substrate and iron-bound water ligand is shown as grey mesh. The electron density map is either a composite omit map or a polder omit map. The details of the contour levels are given in Table S4. Hydrogen bonds are represented by black dashed lines. In grey is the heme, active-site residues are in yellow, the substrate is in green, waters are represented by red spheres, and the chloride ion is in purple.

Figure 7. (top) Overlaid crystal structure of 4-methylthio- (cyan) and 3-methylthio-benzoate-bound CYP199A4 (yellow). (bottom) Overlaid crystal structure of 4-methylamino- (cyan) and 3-methylamino-benzoate-bound CYP199A4 (yellow). The waters and chloride in the *meta*-substituted structures are labelled.

Figure 8. (top) Overlaid crystal structure of 4-methoxy- (cyan) and 3-methoxy-benzoate-bound CYP199A4 (yellow). (bottom) Overlaid crystal structures of 4-methyl- (cyan) and 3-methyl-benzoate-bound CYP199A4 (yellow). The waters and chloride in the *meta*-substituted benzoic acid structure are labelled.

Figure 9. HPLC analysis of the CYP199A4 *in vitro* reaction with 3-ethoxybenzoic acid. In black is the *in vitro* reaction mixture and in gray is a control reaction without P450. Authentic 3-hydroxybenzoic acid (dashed line) co-eluted with the enzyme product at 6.9 min. The substrate, 3-ethoxybenzoic acid, appears at 13.8 min. Gradient: 20-95% AcCN in H₂O with 0.1% TFA. Detection wavelength: 254 nm.

Figure 10. Crystal structure of 3-ethoxybenzoate-bound CYP199A4. A feature-enhanced map of the substrate is shown as grey mesh contoured at 1.0 σ . The insets show the orientation of the 3-

ethoxybenzoic acid substrate relative to the heme and the overlaid structures of 3-methoxy- (cyan) and 3-ethoxy-benzoate-bound CYP199A4 (yellow).

Figure 11. Crystal structures of **(a)** 3-methoxy-, **(b)** 4-methoxy- and **(c)** 3-ethoxy-benzoate-bound CYP199A4 with modelled hydrogens. Hydrogens were added to the crystal structure using ICM-Pro. **(d)** A linear C-H-O angle is calculated to be ideal for hydrogen abstraction[42] and **(e)** the optimal approach angle of the substrate to the $\text{Fe}^{\text{IV}}=\text{O}$ species is postulated to be $\sim 130^\circ$. [41] It has been reported that it is not ideal for the C-H bond of the substrate to approach at angles $>130^\circ$ because this results in larger barriers to abstraction.[41]

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Journal Pre-proof

Graphical Abstract

X-ray crystallography was used to determine the important parameters which lead to highly efficient oxidation of *para*-substituted benzoic acids by the P450 enzyme CYP199A4 compared to their *meta* isomer equivalents. The orientation of the C-H bond relative to the heme iron was found to be critical for efficient O-demethylation.

Journal Pre-proof

Highlights

Six new substrate-bound crystal structures of the P450 monooxygenase CYP199A4.

Related binding modes are observed for *para*- and *meta*-substituted benzoic acids.

These binding modes can be used to rationalize the higher activity of the *para* isomers.

Small positional changes have a significant impact on binding and catalytic activity.

The orientation of C-H bonds relative to heme iron are critical for efficient abstraction.

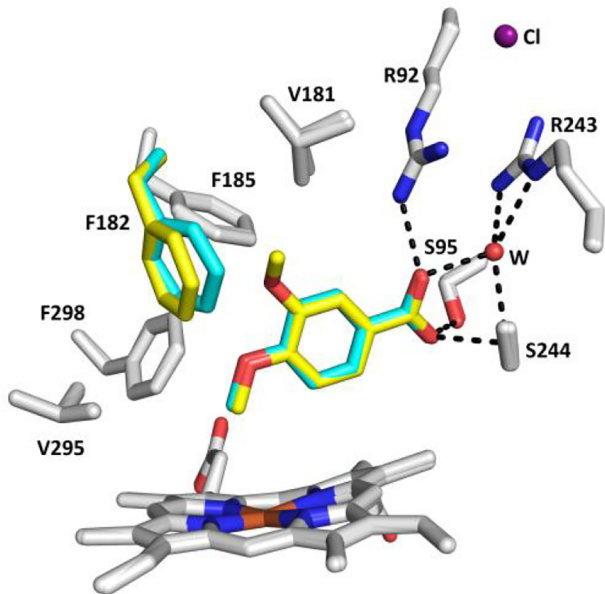
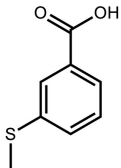
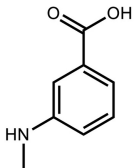


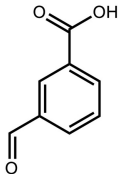
Figure 1



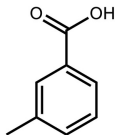
3-methylthioBA



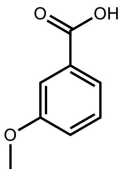
3-methylaminoBA



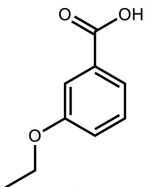
3-formylBA



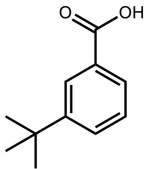
3-methylBA



3-methoxyBA



3-ethoxyBA



3-*tert*-butylBA

Figure 2

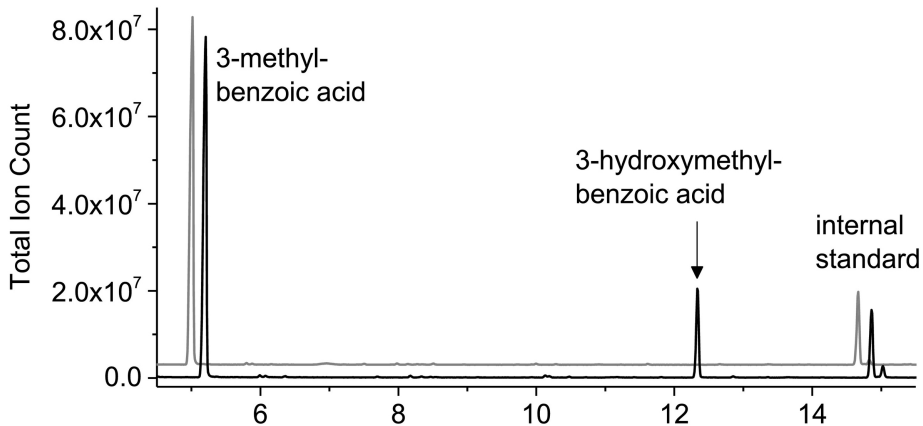


Figure 3

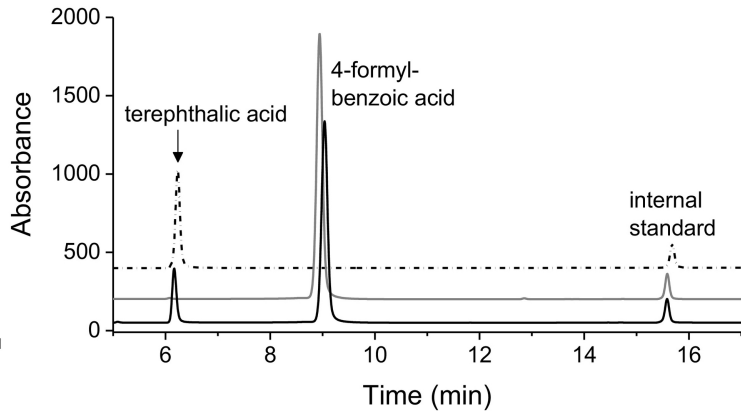
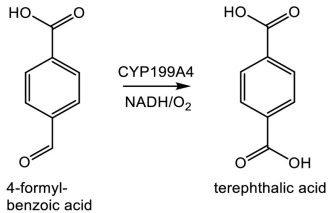
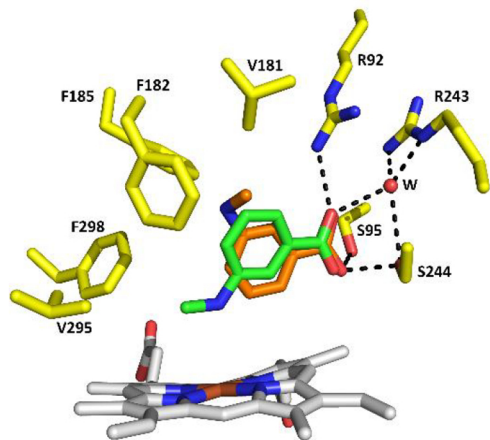
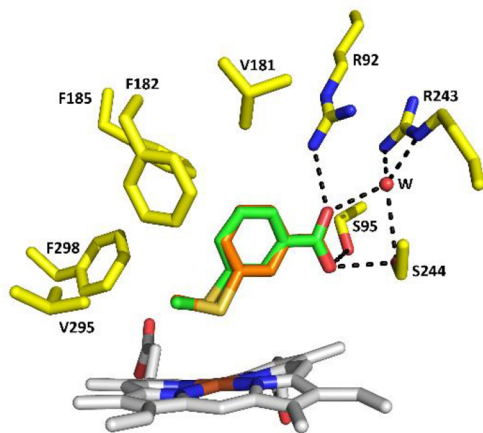


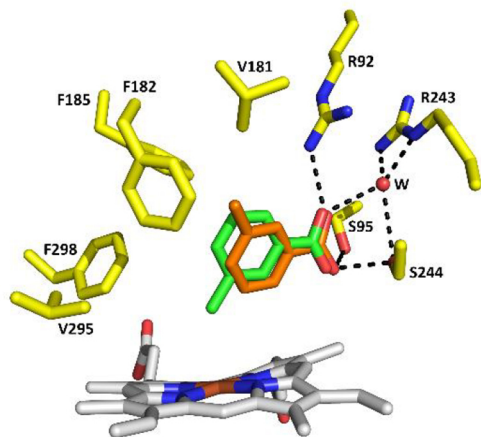
Figure 4



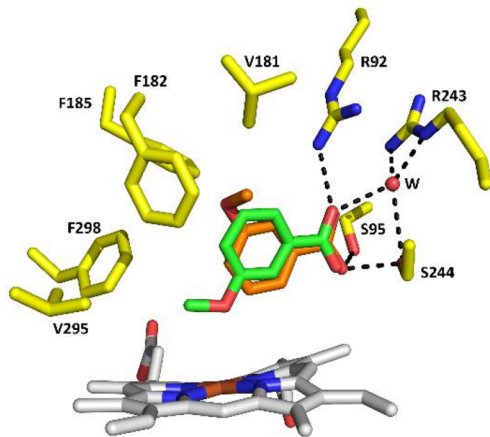
(a)



(b)



(c)



(d)

Figure 5

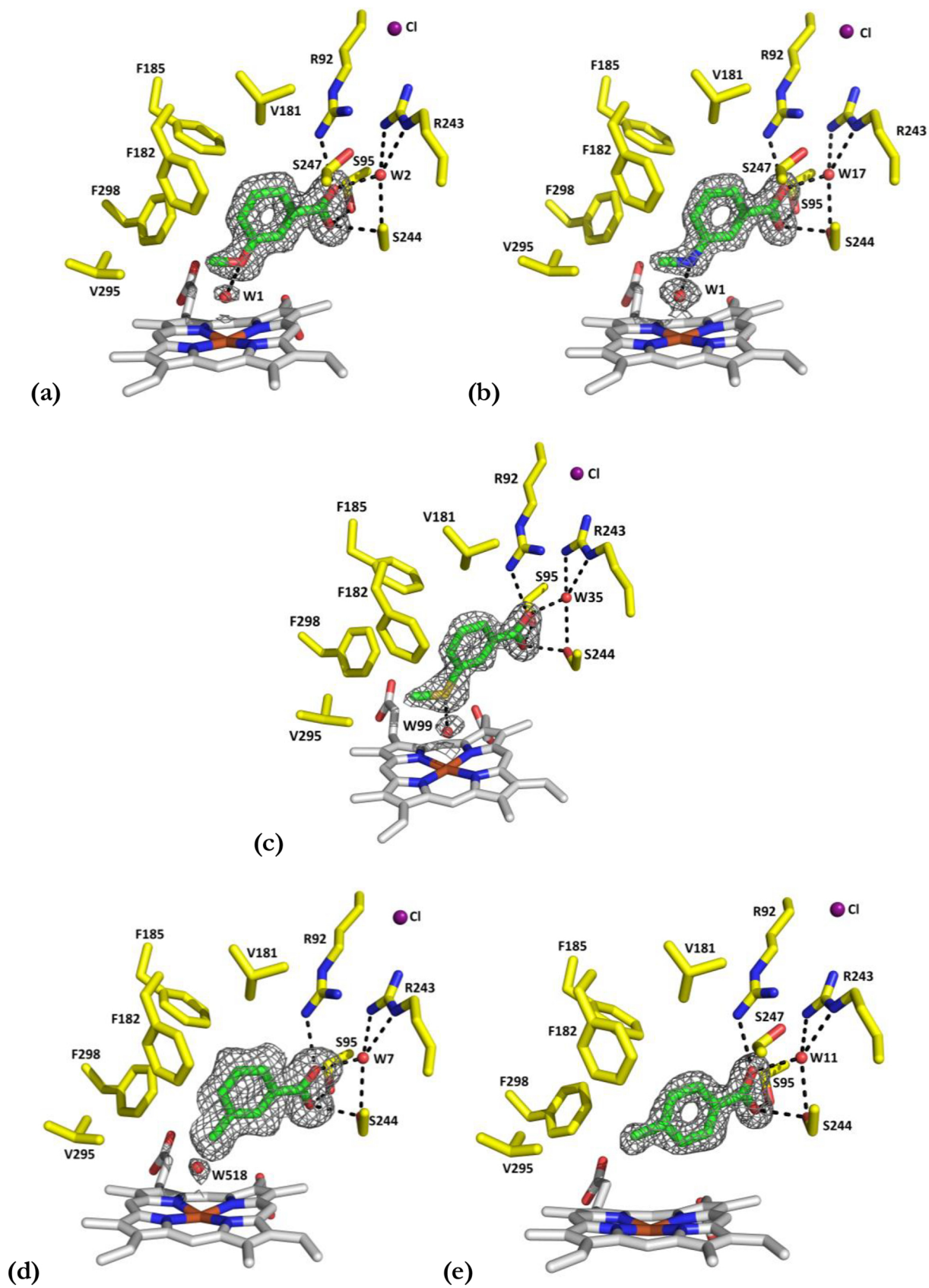


Figure 6

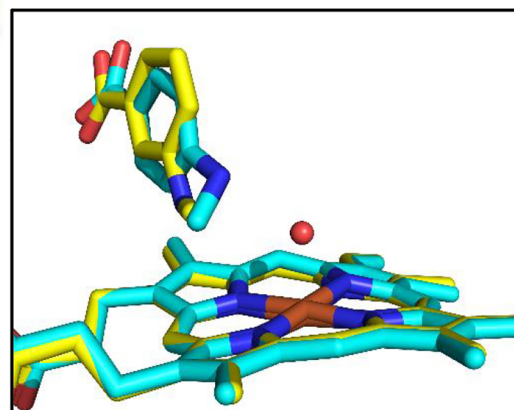
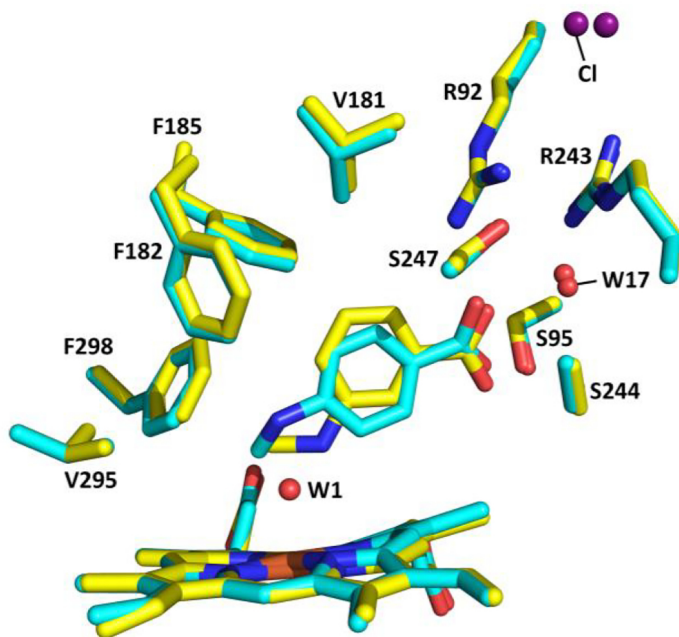
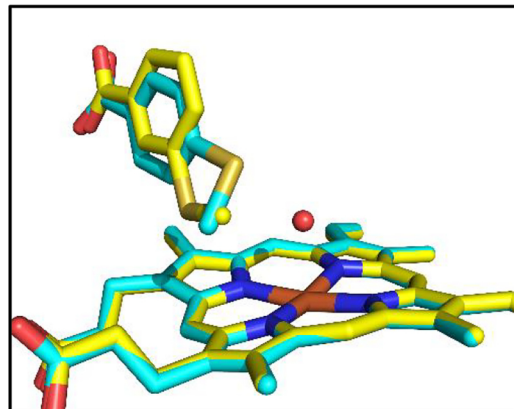
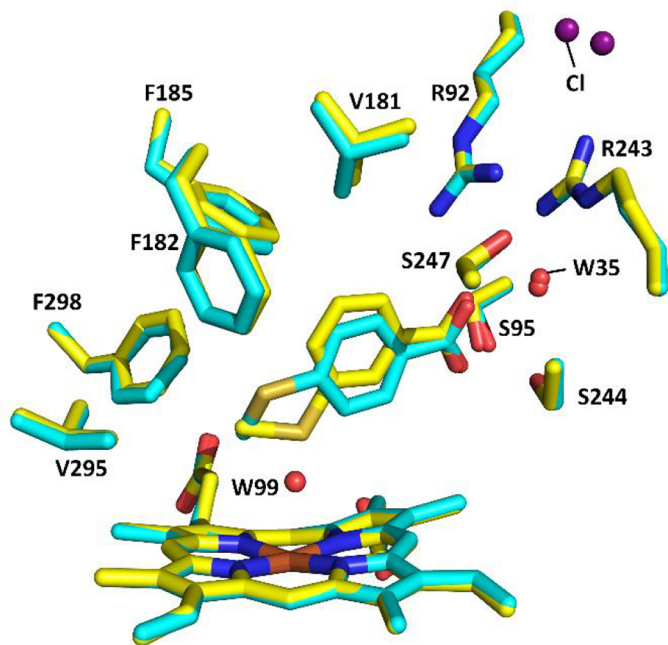


Figure 7

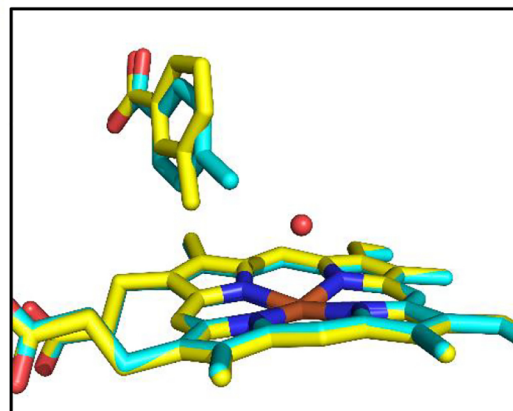
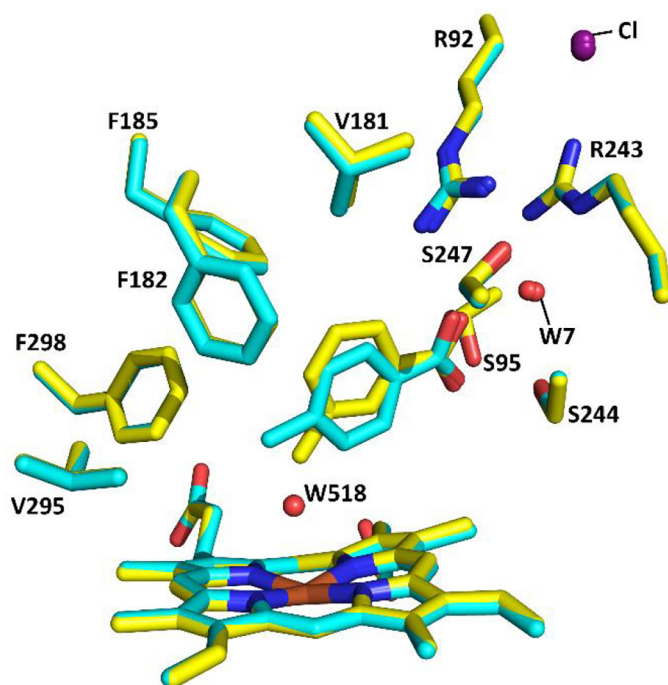
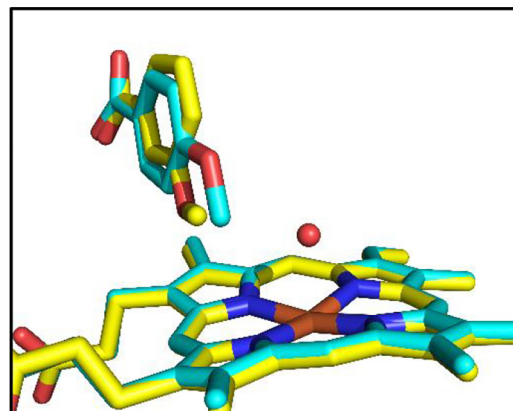
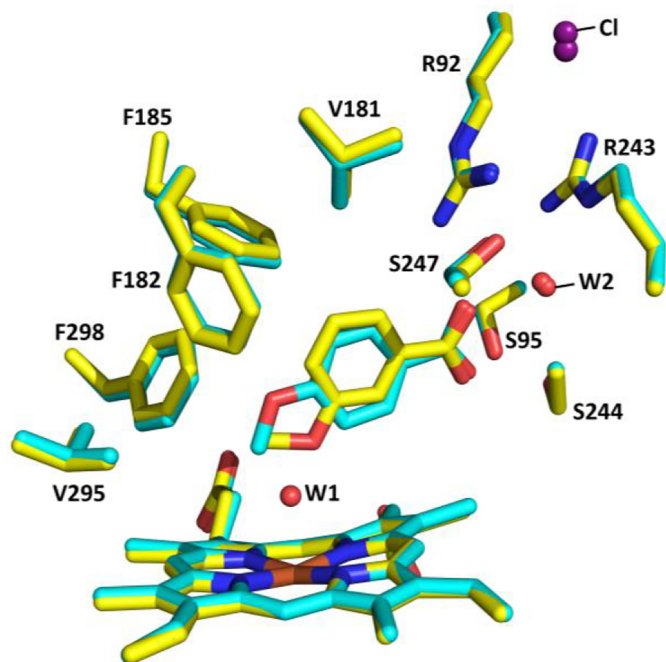


Figure 8

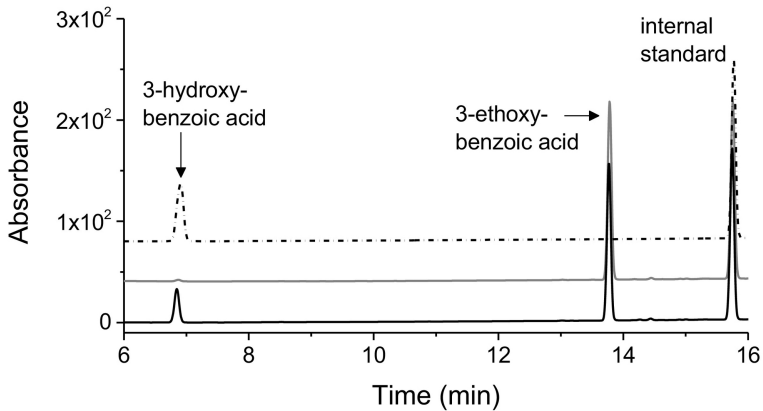
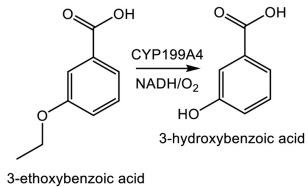


Figure 9

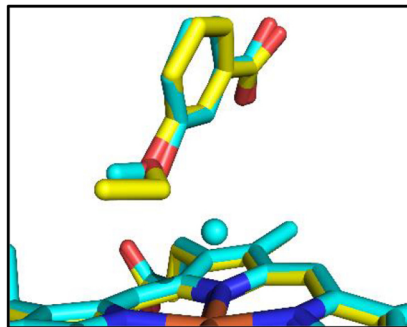
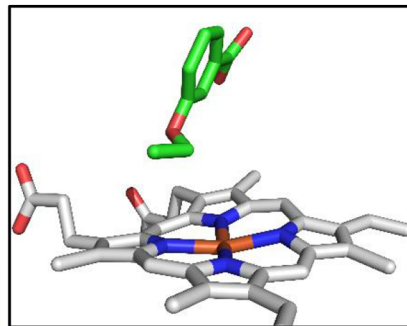
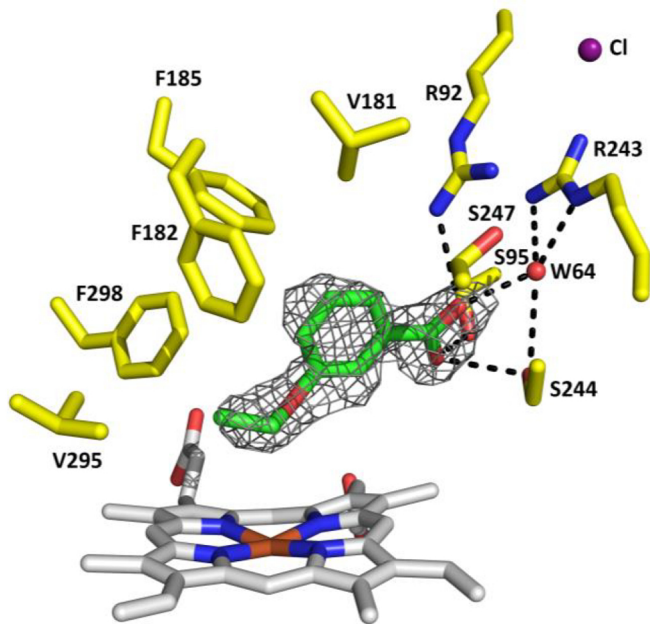


Figure 10

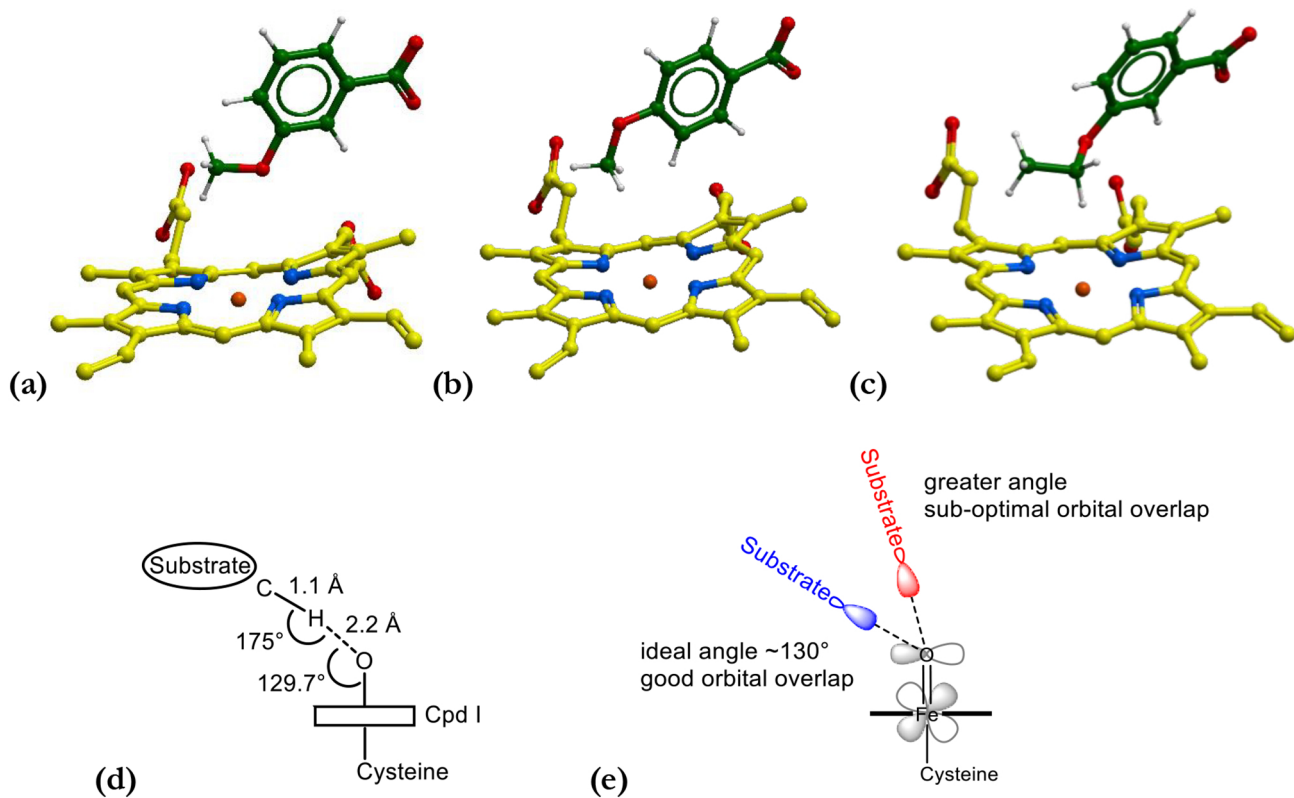


Figure 11