Biochemical Journal (2018) **475** 3875–3886
https://doi.org/10.1042/BCJ20180762

Research Article

Specificity and mechanism of carbohydrate demethylation by cytochrome P450 monooxygenases

Craig S. Robb[1,2], Lukas Reisky[3], Uwe T. Bornscheuer[3] and Jan-Hendrik Hehemann[1,2]

[1]Max Planck Institute for Marine Microbiology, Celsiusstrasse 1, Bremen 28359, Germany; [2]University of Bremen, Center for Marine Environmental Sciences (MARUM), Bremen 28359, Germany; [3]Department of Biotechnology and Enzyme Catalysis, Institute of Biochemistry, University of Greifswald, Greifswald 17487, Germany

Correspondence: Jan-Hendrik Hehemann (jhhehemann@marum.de; jheheman@mpi-bremen.de)

Degradation of carbohydrates by bacteria represents a key step in energy metabolism that can be inhibited by methylated sugars. Removal of methyl groups, which is critical for further processing, poses a biocatalytic challenge because enzymes need to overcome a high energy barrier. Our structural and computational analysis revealed how a member of the cytochrome P450 family evolved to oxidize a carbohydrate ligand. Using structural biology, we ascertained the molecular determinants of substrate specificity and revealed a highly specialized active site complementary to the substrate chemistry. Invariance of the residues involved in substrate recognition across the subfamily suggests that they are critical for enzyme function and when mutated, the enzyme lost substrate recognition. The structure of a carbohydrate-active P450 adds mechanistic insight into monooxygenase action on a methylated monosaccharide and reveals the broad conservation of the active site machinery across the subfamily.

Introduction

Microbial polysaccharide utilization is a key component of the global carbon cycle. Complex carbohydrate degradation requires cognate suites of enzymes to enable utilization by bacteria. Modifications of carbohydrate structures found in nature can shield the sugar against enzymatic degradation in the course of polysaccharide use. Acetyl-, sulfate- and methyl groups need to be enzymatically removed before metabolic enzymes are able to generate energy from monosaccharide conversion. Dealkylation of carbohydrates is particularly difficult to overcome, because methoxy group protections belong to the easy on, hard off category [1,2], i.e., carbohydrates can be readily methylated using methyl-iodide, but it is challenging to remove the methyl groups chemically without damaging the underlying carbohydrate [2]. Plants and algae use this modification to modulate the properties of the polysaccharides and perhaps even prevent or limit their degradation by bacteria [3,4]. For example, in the gut environment, bacteria completely degrade many highly complex carbohydrate structures, but there is evidence that methylated sugars are resistant to degradation [5]. Similarly, methylated carbohydrates accumulate in surface waters of the ocean suggesting that they are relatively recalcitrant in the marine environment as well [6].

A variety of enzyme families have been implicated in demethylation including cytochrome P450 monooxygenases [7], [2Fe–2S]-dependent [8] or folate-dependent [9] enzymes. Cytochrome P450 proteins are ubiquitous oxidizing enzymes that can act as monooxygenases which, in most cases, hydroxylate hydrocarbons leading to the formation of, e.g., ω-hydroxy-fatty acids (P450 BM3) [10], regioselective hydroxylation of steroids like progesterone (CYP109E1) [11] or of camphor ($P450_{cam}$) [12]. The catalytic cycle for these enzymes, typified by $P450_{cam}$ [13], uses molecular oxygen, two electrons from NAD(P)H and two protons at a central heme ligand in the formation of water and a hydroxylated product. Structurally, P450s have a conserved fold and the best understood is $P450_{cam}$

Received: 19 September 2018
Revised: 2 November 2018
Accepted: 6 November 2018

Accepted Manuscript online: 7 November 2018
Version of Record published: 12 December 2018

Biochemical Journal (2018) 475 3875–3886
https://doi.org/10.1042/BCJ20180762

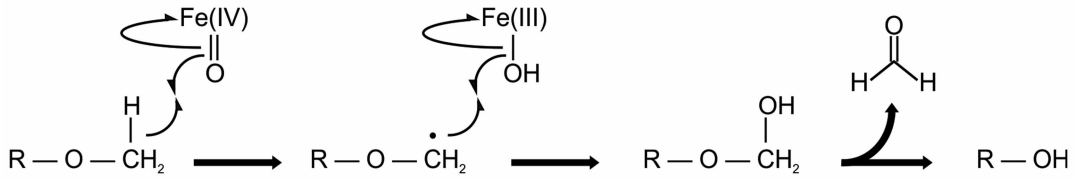

Figure 1. Overall reaction scheme.
The general oxidative demethylation mechanism of cytochrome P450 monooxygenases.

[14]. It has also been trapped *in crysto* in a large variety of intermediate stages of catalysis along the reaction coordinate [13]. Oxidative demethylases, in general, proceed via the insertion of a single oxygen atom into a C–H bond of the methyl group followed by spontaneous decomposition of the hydroxylated intermediate yielding formaldehyde and a demethylated product (Figure 1). A cytochrome P450 *O*-demethylase for lignin degradation has also recently been structurally characterized [15].

We have recently described a novel activity for cytochrome P450s present in marine bacteria that can remove the methyl group from a carbohydrate, 6-*O*-methyl-D-galactose (G6Me), in an oxidative mechanism [16]. This process is dependent on its cognate accessory proteins, ferrodoxin and ferrodoxin reductase, as opposed to many other P450s that accept electrons from generic accessory proteins [17]. These marine bacterial P450s have a tight substrate repertoire as they only accept G6Me out of more than 40 compounds investigated including monosaccharides methylated at the other positions. Furthermore, the carbohydrate-active enzymes did not consume significant amounts of reducing equivalents in the absence of substrate [16]. These enzymes are representative of P450s that act on polar substrates, and represent the first P450 monooxygenases involved in sugar metabolism. A detailed bioinformatic analysis identified that they belong to a new subfamily of P450 monooxygenase and cluster with agarolytic genes in the genomes of marine bacteria. *Zobellia galactanivorans*, the model microbe in the present study, is a marine Flavobacterium that degrades a wide variety of algal polysaccharides including porphyran [18]. In an expression analysis, the target gene of this study that encodes for $P450_{ZoGa}$ (CYP236A2, WP_013995999) was expressed at a higher level on porphyran compared with other polysaccharides just below the 2-fold cutoff suggesting that this enzyme may be used in *Z. galactanivorans* when the organism is growing on this substrate [19] that can contain up to 25% of G6Me [4].

Microorganisms can be engineered to convert algal biomass into valuable chemicals [20]. If G6Me-rich biomass were to be used, the demethylation by $P450_{ZoGa}$ could be used to increase the available amount of fermentable sugars and thereby improve the conversion yield. Furthermore, a molecular understanding of the acceptance of a hydrophilic substrate by a P450 monooxygenase would enable enzyme engineering to adapt the enzyme to different substrates. This could be used, for example, to perform the selective demethylation of various carbohydrates, which is chemically not feasible [21].

Here, we examined the molecular mechanism of substrate recognition and *O*-demethylation of carbohydrates by $P450_{ZoGa}$ using calorimetry, X-ray crystallography and mutagenesis of active site residues. The structure demonstrates positioning for hydroxylation of the distal carbon of the methyl group and a comparison to characterized enzymes illustrates the distinct mechanism of polar substrate recognition in this subfamily of P450s. Mutagenesis of the active site residues demonstrates that the enzyme is acutely sensitive to perturbations in the side chains that mediate the interaction with the substrate. The present study adds mechanistic insight to the ability of P450s to specifically recognize and demethylate carbohydrate ligands.

Experimental procedures

Protein production, purification and crystallization

Recombinant expression of $P450_{ZoGa}$ was carried out using *Escherichia coli* BL21 (DE3) with pET22-$P450_{ZoGa}$ in media containing 1% tryptone, 0.5% yeast extract, 100 mM $MgSO_4$, 0.5% glycerol, 0.05% glucose, 0.2% lactose, 25 mM ammonium sulfate, 100 mM potassium phosphate monobasic, 100 mM sodium phosphate dibasic, 100 μg/ml ampicillin, 0.3 mM $FeSO_4$ and 0.5 mM 5-aminolevulinic acid in unbaffled flasks at 20°C and 180 rpm for three days. Cells were harvested by centrifugation and stored at −20 °C until use. Cell pellets were disrupted by chemical lysis using sucrose solution (25% sucrose, 50 mM Tris pH 8.0) to which lysozyme was added to a concentration of 1 mg/ml and incubated for 10 min. Deoxycholate solution (1% deoxycholate, 1% Triton X-10,

1 mM $MgCl_2$ and 100 mM NaCl) was added and the released DNA was degraded with 1 mg/ml DNase. The resulting lysate was clarified by centrifugation at 16 000×*g* for 45 min at 4°C and applied to a 5 ml prepacked TALON column (GE) equilibrated in Buffer A (20 mM Tris pH 8.0 and 500 mM NaCl). Unbound protein was washed away with Buffer A and then bound protein was eluted with a gradient of imidazole up to 500 mM in Buffer A. Purified protein was concentrated using a stirred cell ultrafiltration device on ice with a 10 kDa molecular mass cutoff and dialyzed into crystallization buffer (20 mM Tris pH 8.0 and 150 mM NaCl). The extinction coefficient $\varepsilon_{450\ nm} = 91\ mM^{-1}\ cm^{-1}$ of the carbon monoxide difference spectrum was used to quantify the protein concentration. Dialyzed protein was screened for crystallization conditions using the Midwest Center for Structural Genomics screen 1 (Molecular Dimensions) and conditions were optimized to obtain large crystals. The final optimized condition crystals were obtained by mixing protein solution 1 : 1 with well solution (20% PEG 3350, 150 mM NaCl). Crystals were soaked in 6-*O*-methyl-D-galactose (Carbosynth) and cryo-protected in glycerol prior to flash freezing in liquid nitrogen.

Mutagenesis

Site-directed mutagenesis was used to change the active site residues V58, E62, F73, L267 and M270 to alanine. PCRs were performed with 0.4 µl *Pfu*Plus! DNA polymerase (5.0 U/µl; roboklon), 2.5 µl 10× *Pfu*-buffer, 1.25 µl forward and reverse primer (10 µM), 0.5 µl dNTPs (10 mM each), ∼40 ng template DNA and water added to a total volume of 25 µl. After an initial 30 s at 95 °C, the cycle of 30 s at 95 °C, 30 s at 63 °C and 7 min at 72 °C was repeated 20 times prior to a final elongation step for 10 min at 72 °C. Afterward, the reactions were stored at 4 °C. To digest template DNA, 1 µl DpnI (20 U/µl, New England Biolabs) was added and incubated for 2 h at 37 °C prior to inactivation for 20 min at 80 °C. Chemically competent *E. coli* Top10 were transformed with 2.5 µl of the mixture and plated on LB agar supplemented with 50 µg/ml kanamycin for selection. Successful generation of the mutants was confirmed by sequencing.

Data collection and structure solving

Diffraction data were collected for both the unbound and 6-*O*-methyl-galactose bound $P450_{ZoGa}$. Data were automatically indexed and integrated using XDS [22] and scaled in Aimless [23]. The structure was solved by molecular replacement using an ensemble of the co-ordinates from four P450 structures (pdb id: 5l90, 4ubs, 1cl6 and 3ejd) [11,24–26] and phaser [27] in the phenix suite [28]. A combination of buccaneer, Refmac5, phenix.refine and coot were used to automatically build and manually complete the structure of $P450_{ZoGa}$ [29–33]. Structure figures were prepared in PyMOL.

Isothermal titration calorimetry

Recombinant protein was desalted using PD-10 columns (GE Healthcare) and concentrated into ITC (isothermal titration calorimetry) buffer (50 mM NaPi, 100 mM NaCl, pH 7.4) using Vivaspin 6 (Sartorius) with a 30 kDa membrane. The ligand was resuspended in the same buffer and the ITC experiment was conducted using a MicroCal PEAQ-ITC instrument (Malvern Panalytical) with analysis and control software version 1.1.0.1262. The protein was at 1.42 mM in the cell and ligand at 50 mM in the syringe. As controls, the buffer was titrated into buffer, the buffer into protein and the ligand into buffer to estimate the heat of dilution.

Binding assay

Mutants of $P450_{ZoGa}$ were purified and dialyzed into binding buffer (25 mM Tris, 100 mM NaCl, pH 8.0). Binding was initially screened for at a substrate concentration of 5 mM to detect binding activity. Mutants that showed binding were further analyzed by generating an endpoint binding curve that was fit to the equation using data from three trials to obtain a dissociation constant, K_d.

Results

The overall structure of $P450_{ZoGa}$ reveals structural rearrangements relative to characterized cytochrome P450 monooxygenases

Generally, eukaryotic cytochrome P450 monooxygenases are generalist enzymes that act on relatively nonpolar substrates lacking strong specificity, as they are often used to detoxify various xenobiotics. But, many bacterial cytochrome P450s tend to have narrower substrate specificities sometimes acting on a single substrate [15]. $P450_{ZoGa}$ has only been shown to be active on G6Me despite testing of other *O*-methylated monosaccharides [16].

Biochemical Journal (2018) **475** 3875–3886
https://doi.org/10.1042/BCJ20180762

In its biological context, this substrate specificity is reasonable as the biological function of $P450_{ZoGa}$ is the removal of the shielding methyl group from the D-galactose sugar present in agar. This is the most abundant methoxy sugar in these complex polysaccharides that can be up to 25% of raw agar. Consequently, $P450_{ZoGa}$ has evolved for this specific function to provide the marine bacteria with a unique advantage over other microbes competing for carbon sources. To better understand the mechanism of specificity and substrate recognition, we used X-ray crystallography to study the structure–function relationship of this enzyme and its carbohydrate substrate. Co-purification of recombinant $P450_{ZoGa}$ with its heme cofactor yielded bright red protein. Crystals of the protein were obtained in the space group P212121 that diffracted to 2.25 Å. The structure was solved using molecular replacement with two molecules in the asymmetric unit (Table 1). Model building led to a nearly complete model between residues K2 and L385 out of 387 residues in the native sequence. Gaps remained between G153 and D167 in chain A and between G153 and A163 in chain B. The structure of $P450_{ZoGa}$ consists of a single domain of mixed alpha–beta topology and an overall triangular globule form (Figure 2A). The heme ligand binds deep within the core of the

Table 1 X-ray data collection and refinement statistics

Dataset	$P450_{ZoGa}$	$P450_{ZoGa}$ G6Me
X-ray source	DESY P11	DESY P11
Wavelength (Å)	1.7389	1.7389
Resolution (Å)	99.87–2.25 (2.37–2.25)	99.84–2.40 (2.49–2.40)
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell constants (Å)	$a = 63.28$, $b = 68.15$, $c = 199.7$	$a = 63.02$, $b = 67.88$, $c = 199.7$
Number of reflections	418 105(57 238)	233 876(24 106)
Number unique	41 766(6000)	34 391(3545)
R_{sym} (%)	0.111 (0.809)	0.127 (0.789)
Completeness (%)	99.5 (99.7)	99.9 (100)
Redundancy	10.0 (9.5)	6.8 (6.8)
$\langle I/\sigma(I)\rangle$	14.4 (3.3)	10.1 (2.4)
Mosaicity	0.15	0.27
Wilson B factor	24.1	24.1
Refinement		
R_{work}/R_{free} (%)	20.35/24.57	16.92/22.78
Number of atoms	6206	6375
Protein	5924	5983
Ligand	86	112
Water	196	272
B factors		
Overall	36.76	38.75
Protein	38.36	39.04
Ligand	21.72	29.30
Water	32.98	36.24
R.m.s. deviations		
Bond lengths (Å)	0.017	0.013
Bond angles (°)	1.659	1.608
Ramachandran statistics (%)		
Favored	97.3	96.39
Allowed	2.7	3.34
Outliers	0.0	0.27
PDB accession code	6G5Q	6G5O

Biochemical Journal (2018) **475** 3875–3886
https://doi.org/10.1042/BCJ20180762

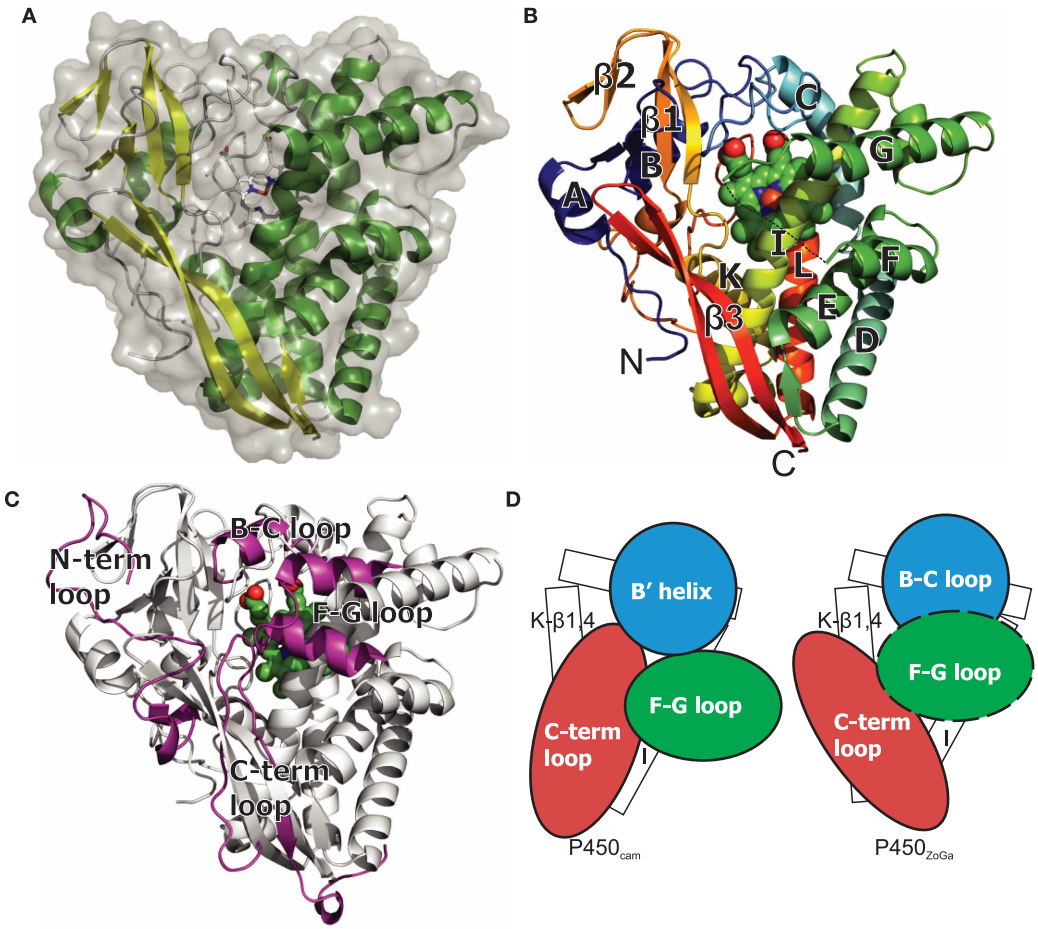

Figure 2. Overall structure of $P450_{ZoGa}$.

(**A**) Overall triangular shape of $P450_{ZoGa}$. The surface of the protein is represented as a transparent surface with the secondary structure elements labeled in yellow and green. (**B**) The structure is shown as a cartoon colored as a rainbow from *N*- to *C*-terminus gradually from blue to red. The heme prosthetic group is shown as calotte model. The molecule is viewed from the substrate access (distal) face and visible structure elements are labeled, as are the *N*- and *C*-termini. The gap in the F-G loop is indicated using a dotted line. (**C**) The four major differences from $P450_{cam}$ (PDB ID code 3CPP) are shown in purple on a background of $P450_{ZoGa}$ in white. The heme ligand is shown as spheres in green. (**D**) The arrangement of the structural features that dominate the opening to the catalytic site in $P450_{cam}$ and $P450_{ZoGa}$. The active site (below) is framed by the I helix, helix K-β1-4 and the B-C loop. Overlain upon this framework are the three loops: the B-C loop (blue), F-G loop (green) and the C-terminal loop (red). The F-G loop is disordered in the structure of $P450_{ZoGa}$ and has been modeled as a dashed line.

fold, forming a covalent bond with the thiolate side chain of C330 and is only accessible to solvent on the distal face. The P450 protein family has a set of standard nomenclature for the conserved secondary structure elements [14]. The helices are named with letters from A to L and the beta-sheets are numbered 1–3 (Figure 2B). β1 is a four-stranded sheet with two strands from the *N*-terminus and two from the center of the protein sequence. β2 consists of a pair of antiparallel strands inserted between β1–3 and β1–4 that stacks against β1. β3 is a long twisted β-sheet with two antiparallel strands that extends along the longest edge of the triangular globule and a small three-residue strand inserted between helices D and E. The orientation and tight packing of the C-terminal loop suggests that this unique structural feature is fixed and cannot participate in closing the active site. The active site is positioned above the heme ligand and is formed by residues from the I helix, the B-C loop, the β1-4 strand and the preceding K-β1-4 loop. This active site is capped by the F-G loop that, in the structure of $P450_{ZoGa}$, is not resolved and this is likely due to intrinsic flexibility of this loop. This region of P450 is typically flexible to allow substrate access to a closed active site and in at least one other case is only visible in the substrate-bound form [34].

Biochemical Journal (2018) 475 3875–3886
https://doi.org/10.1042/BCJ20180762

Exploring the structures similar to $P450_{ZoGa}$ using the Dali server revealed the closest structural homologs are the bacterial P450s CYP124, P450cin and $P450_{cam}$ [35–37]. As such, the structure of $P450_{ZoGa}$ shares many structural features with the archetypal $P450_{cam}$ [37] and the more recently characterized structures. When structurally aligned with $P450_{cam}$, the structures share a sequence identity of 24.6% and a core rmsd of 1.76 Å. In Figure 2C, the shared structural features are hidden and only the major differences are shown. This view highlights four major differences at the structural level that lie in the secondary structure elements involved in capping the conserved active site. In $P450_{ZoGa}$, as previously noted, the active site is composed of the K-β1-4 loop, the B-C loop and the I helix. In $P450_{cam}$, the *C*-terminal loop, the F-G loop and the B′ helix in place of the B-C loop come together to close the active site during catalysis (Figure 2D). However, in $P450_{ZoGa}$ the *C*-terminal loop is in the form of a long twisted β-hairpin that runs along the edge of the structure leaving the active site wide open. As a closed active site is required, the unresolved residues of the F-G loop in $P450_{ZoGa}$ were modeled to test whether they could reach across (Supplementary Figure S1). The structural space occupied by the *C*-terminal loop of $P450_{ZoGa}$ is replaced in $P450_{cam}$ by an extended *N*-terminal loop that extends toward the β-sandwich (β1–β2). Overall, the P450 fold is highly conserved but here we observe a rearrangement relative to the active site architecture found in other P450 homologs.

Complex of $P450_{ZoGa}$ with G6Me reveals the specific residues that mediate substrate recognition

To evaluate the molecular interactions involved in substrate recognition and catalysis, a dataset was collected using a protein crystal soaked in substrate. The resulting co-structure reveals the specific residues that bind and orient the substrate above the active center (Figure 3A). The substrate is found surrounded by hydrophobic residues with a pair of hydrophilic residues orienting the target methyl-group toward the iron center of the heme ligand (Figure 3B). The B-C loop provides Glu62 and Gln70 that coordinate the O2, O3 and O4 hydroxyl groups of the pyranose ring through hydrogen bonding. The adjacent residue Arg66 is close enough to modulate the pKa of Glu62 allowing it to act as a proton acceptor in substrate binding (Figure 3C). Furthermore, the loop provides the hydrophobic residues Val58 and Phe73. The latter is within range to form a C–H···π interaction between C6 of the ligand as donor and center of the aromatic ring of Phe73 as acceptor [38]. The ligand is tightly packed between the I helix and the β1–4 residues as the distance between Cα of Gly220 and Met270 is 8 Å and between G220 and Phe73 is also 8 Å. These residues are close enough to exclude water from the immediate vicinity when the substrate is bound. In summary, the ligand is bound by a hydrophobic environment with additional electrostatic interaction on the distal side of the binding site. These structural data suggest both a favorable entropic and enthalpic contribution to binding and with this hypothesis we turned to calorimetry.

Isothermal titration calorimetry supports the proposed mechanism of substrate recognition

The first step of the catalytic cycle of $P450_{ZoGa}$ is binding of substrate to the active site. This binding event leads to a spectral shift as the iron center shifts from hexa- to penta-coordinated as apical water is displaced [12,37]. To date, the ligands of well-characterized P450s are mostly hydrophobic and as such, these enzymes accommodate the substrate in nonpolar binding pockets and the interaction is often driven by entropy [39]. However, carbohydrates are polar and hydrophilic and the crystal structure revealed that they bind through a disparate mechanism. We used ITC to investigate the interaction of the polar substrate with the P450 (Figure 4). This technique provides direct measurement of changes in heat in the system. Integration of the binding isotherm leads to thermodynamic parameters in addition to the affinity.

Fitting the data with a fixed single site-binding model based on the crystal structure led to a dissociation constant of 0.98 mM. The binding has a favorable enthalpic parameter and a slightly larger favorable entropic parameter (Table 2). This value deviates somewhat from the value obtained from difference spectrum measurements, but this may simply be due to the difference in pH and salinity between the measurements. Protein carbohydrate interactions are typically enthalpy driven through electrostatic interactions with an unfavorable entropy component [40,41]. But, P450 interactions with substrate are typically entropy driven due to the hydrophobic substrate [39]. In $P450_{ZoGa}$, the favorable entropic parameter of the carbohydrate-active P450 may be due to the hydrophobicity of the active site compared with other carbohydrate-active proteins. The relative energetic cost of solvation of the free ligand and hydrophobic-binding site for the additional

Biochemical Journal (2018) **475** 3875–3886
https://doi.org/10.1042/BCJ20180762

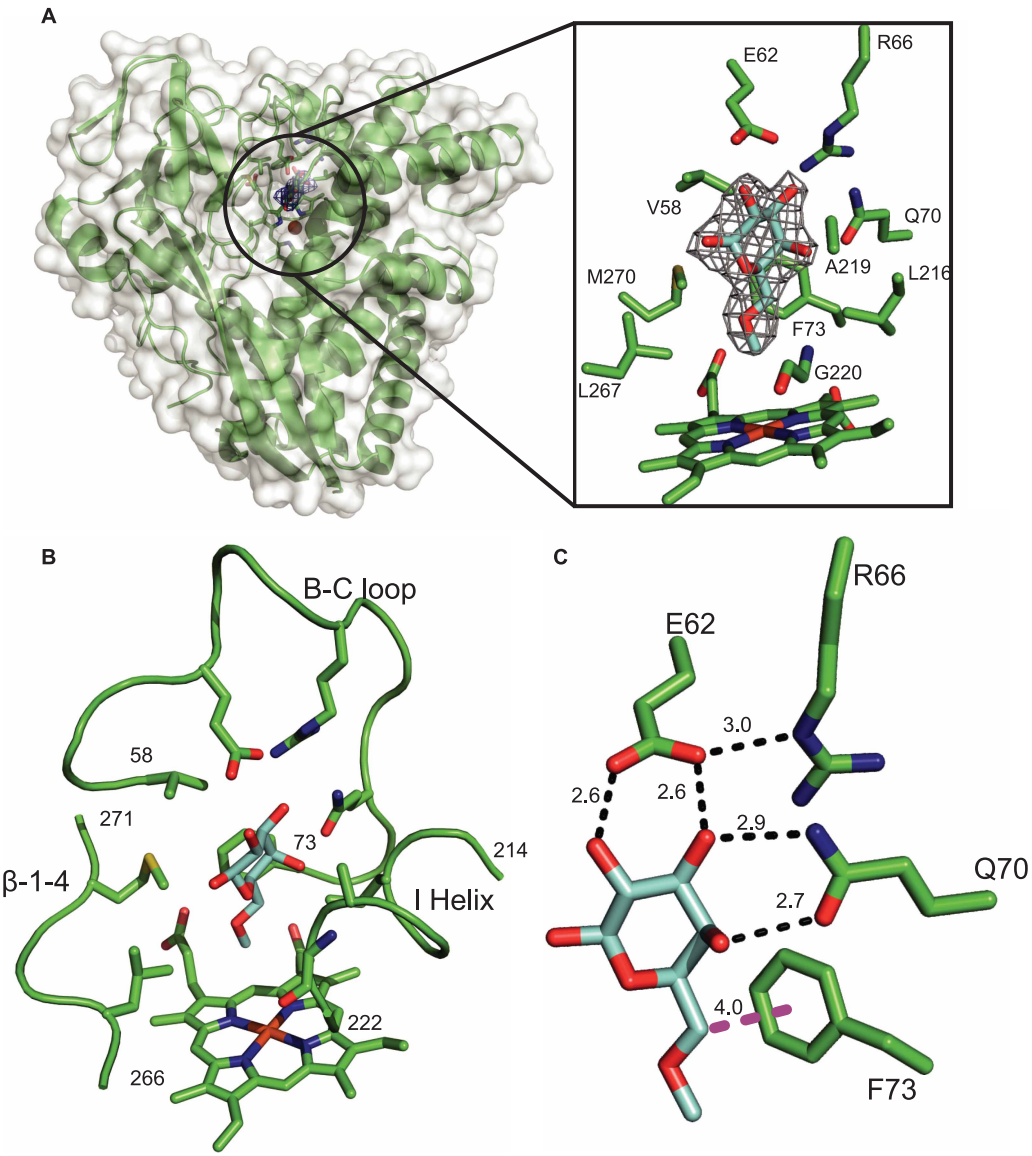

Figure 3. Coordination of G6Me in the active site of $P450_{ZoGa}$.

(**A**) The $2F_o - F_c$ map at 1 sigma is shown in blue along with the adjacent residues of the active site. Ligand coordination in the complex of 6-*O*-methyl-D-galactose and $P450_{ZoGa}$. (**B**) Overall view of catalytic site architecture with the relevant secondary structure elements and the residue ranges labeled. The iron is represented as a brown colored sphere. (**C**) The possible electrostatic interactions found within the catalytic site that coordinate the ligand. Hydrogen bonds are shown in black and the C–H···π interaction is shown in purple with the distances in Å.

methyl-group bound to the monosaccharide ligand could explain this difference. The favorable enthalpic parameter is expected for proteins that rely on hydrogen bonding for substrate recognition and suggests that $P450_{ZoGa}$ employs this mechanism. Overall, the ITC data suggests that both the hydrophobic effect and hydrogen bonding are involved in substrate recognition. To further investigate the contribution of these residues to substrate binding, we used mutagenesis.

Mutants of catalytic site residues abrogate binding to 6-*O*-methyl-D-galactose

Mapping the sequence conservation among the close homologs of $P450_{ZoGa}$ within the CYP236A family onto the structure revealed absolute conservation of the amino acids identified in and around the active site

Biochemical Journal (2018) 475 3875–3886
https://doi.org/10.1042/BCJ20180762

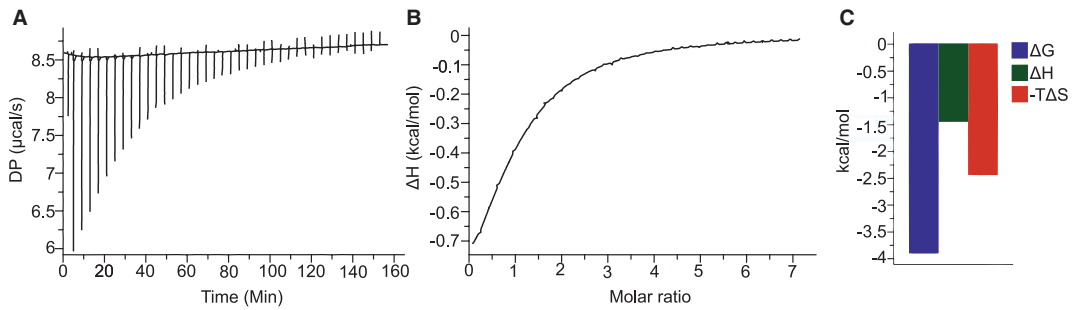

Figure 4. Representative ITC titration of $P450_{ZoGa}$ and 6-*O*-methyl-D-galactose.

(**A**) The raw calorimetric data, (**B**) the integrated heats and (**C**) a bar diagram showing the contributions of Δ*H* and −TΔS to Δ*G*.

(Figure 5A). This conservation underlines the important biological function of these residues and reveals that the subfamily probably exhibits one single activity. We performed site-directed mutagenesis of active site residues to probe their importance. Of six mutants, only the M270A mutant still bound the substrate at an affinity ∼50 times lower as measured in a spectral-difference assay (Table 3). The M270A mutant from the β1–4 strand was the only modification that did not result in a complete loss of substrate recognition. However, the adjacent L267A mutant resulted in a protein that did not bind suggesting that this residue is even more important for the maintenance of a hydrophobic environment. The four mutants from the B-C loop all resulted in a loss of binding to the substrate. These four included the residues Glu62 and Gln70 that mediate hydrogen bonding with the substrate and Val58 and Phe73 that contribute to the hydrophobic environment. The B-C loop, in particular, contains many residues that are absolutely conserved across the protein subfamily and may represent a signature for the carbohydrate-active P450 (Figure 5B). Combined, the strict conservation of many residues shown to be interacting with the ligand and the exquisite sensitivity to perturbations in the active site, reveal a finely tuned enzyme family focused on the removal of 6-*O*-methylation from D-galactose.

Discussion

Monooxygenases act as *O*-demethylases by hydroxylating the substrate followed by spontaneous decomposition of the acetal group [8,42]. To the best of our knowledge, P450 *O*-demethylases have never been examined structurally in complex with their substrate and therefore this line of direct mechanistic evidence is lacking. One key aspect for the mechanism is the distance between the carbon of the target and the iron center. In $P450_{ZoGa}$ the *O*-methyl group in G6Me and the heme iron are 4.28 and 4.74 Å apart in chains A and B, respectively. This is in line with the distance between ligand and heme group observed in $P450_{cam}$ (Figure 6). Furthermore, conservation of Thr224, Asp223 and the kinked I helix which together form the acid/alcohol motif and proton channel [43,44] suggests that the reaction proceeds in an analogous manner to other P450s typified by $P450_{cam}$ [13] up to the point where a hydroxylated intermediate is formed (Figure 6). In the end, this short-lived species rapidly decomposes to form formaldehyde and the demethylated substrate; in this case D-galactose.

Cytochrome P450 monooxygenases catalyze a variety of hydroxylation reactions. The nature of the P450 reaction mechanism and the highly reactive intermediates require a diffusion limited active site with controlled water access [45]. However, many P450s *in vitro* suffer from large rates of uncoupling that results in the production of reactive oxygen species, e.g., hydrogen peroxide [46]. In a physiological context, such reactions would not only waste chemical energy by depleting the NAD(P)H pool but also directly generate harmful, reactive oxygen species [47]. In P450ZoGa, the polar hydroxyl groups of the sugar substrate are bound by a relatively polar active site compared with other P450 monooxygenases. This might render this active site particularly susceptible to disruption of the catalytic cycle of P450s due to excessive hydration near the iron center. However, both previously enzymatically characterized members of CYP236A showed high coupling

Table 2 Thermodynamic parameters of ligand interaction with $P450_{ZoGa}$ derived from ITC at 20°C

	N (sites)	K_d (mM)	Δ*G* (kcal/mol)	Δ*H* (kcal/mol)	*T*Δ*S* (kcal/mol)
$P450_{ZoGa}$	1	0.98 ± 0.12	−4.04 ± 0.08	−1.17 ± 0.07	2.85 ± 0.16

Biochemical Journal (2018) **475** 3875–3886
https://doi.org/10.1042/BCJ20180762

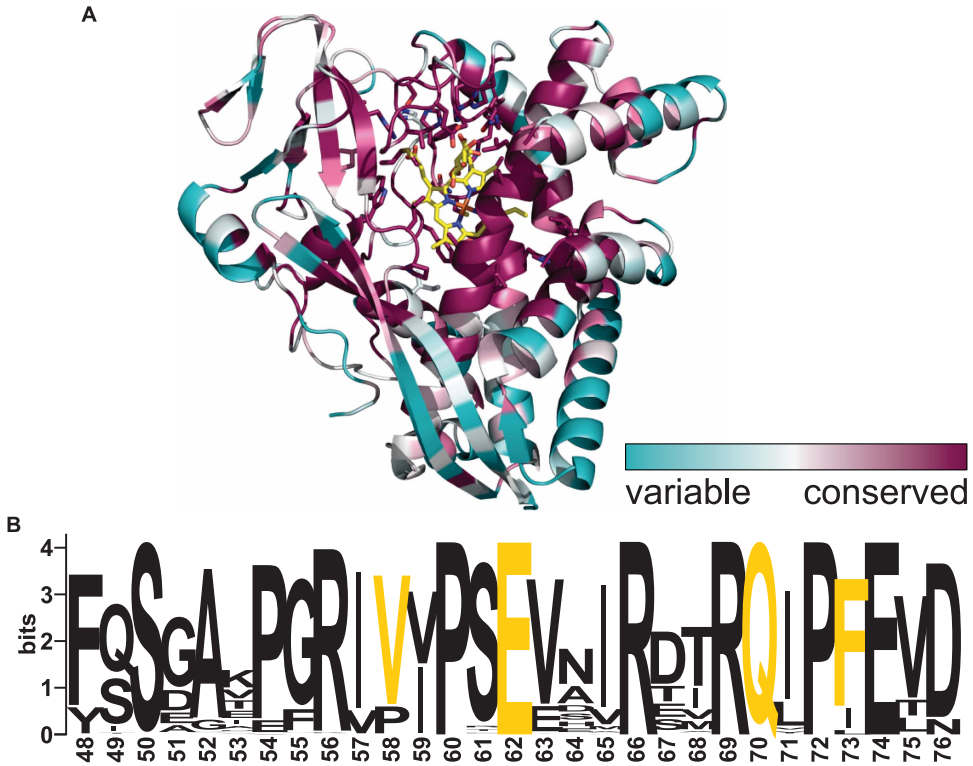

Figure 5. Conservation of P450 CYP236A residues.

(**A**) Sequence conservation was mapped onto the structure of $P450_{ZoGa}$ [50] to illustrate the absolute conservation of the core including the residues implicated in substrate recognition. Conservation is represented by color with dark and light colors representing conserved and divergent residues, respectively. One hundred and fifty bacterial sequences were extracted from NCBI using blast and aligned. (**B**) Sequence conservation of the B-C loop is shown in a Weblogo [51]. The residues highlighted in yellow were tested in mutagenesis experiments.

demonstrating that these enzymes have evolved to minimize uncoupling, while converting a relatively polar substrate. This is underlined by invariance of the residues that provide the fine-tuned recognition of G6Me. This hypothesis is supported by the high sensitivity of the enzyme to perturbations of the side chains in the active site (Table 3).

The relatively low affinity of $P450_{ZoGa}$ for its native substrate that is tightly coordinated by conserved residues suggests that this enzyme is highly specialized for G6Me. Coordination of the substrate in a longitudinal orientation with the hydrogen bonding occurring on the 2, 3 and 4 positions of the carbohydrate and a ring of hydrophobic residues surrounding the 6-*O*-methyl-group takes advantage of the unique chemistry of the substrate. Among the methylated monosaccharides tested using this enzyme, only G6Me has a 6-*O*-methyl group. Monosaccharides methylated at other positions would likely be excluded, barring a significant conformational change, because they cannot bridge the hydrogen-bonding potential of the B-C loop and the iron center at the same time. The high specificity observed could also be necessary to prevent unwanted oxidation reactions of other cellular compounds. If other sugars than G6Me would be nonspecifically hydroxylated, the cell would lose energy and valuable carbon sources. It is also important that the reaction product D-galactose, which differs from the substrate by the absence of the methyl group, is discriminated and not oxidized

Table 3 Effect of mutagenesis of active site residues on binding G6Me

Mutation	Wild type [16]	V58A	E61A	Q70A	F73A	L267A	M270A
K_d (mM)	1.80 ± 0.05	ND	ND	ND	ND	ND	74.6 ± 9.9

ND, no binding detected.

Biochemical Journal (2018) **475** 3875–3886
https://doi.org/10.1042/BCJ20180762

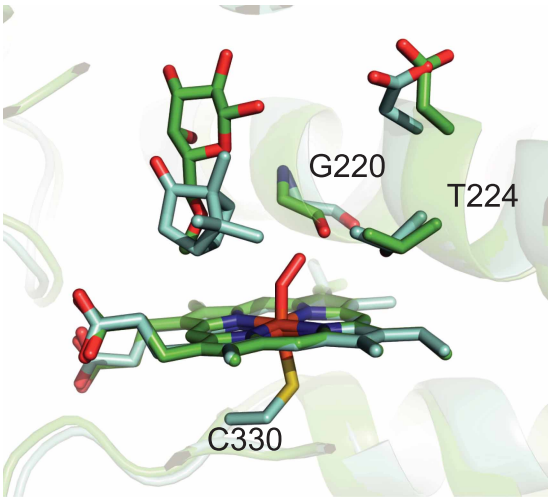

Figure 6. Comparison of the I helix and substrate binding of $P450_{cam}$.

$P450_{cam}$ is in complex with camphor and dioxygen (cyan) and the complex of $P450_{ZoGa}$ with G6Me is shown (green). The numbering comes from $P450_{ZoGa}$.

further while still in the active center. This so-called over oxidation has been reported for several P450 monooxygenases [10,48].

The apparent K_M of $P450_{ZoGa}$ is 0.97 mM [16] which is roughly equal to the dissociation constant of the enzyme. Most P450s have a much lower K_M and dissociation constant compared with these values. A typical K_d or K_M for a lipophilic substrate is closer to 1 μM [15]. This may be a consequence of the polarity of the ligand as a strong entropy-driven affinity is not accessible for enzyme acting on a monosaccharide. Indeed, if other P450s had such weak affinities, quantitative experiments would be much more challenging due to the insolubility of the ligands.

Mechanistically, $P450_{ZoGa}$ acts in an identical manner to most other characterized P450s up to the point of G6Me hydroxylation. This is followed by decomposition of this intermediate into free product and formaldehyde. This mechanism has been proposed based on isotopic effects [49] and is logical based on the products that are formed [16]. The recent structural analysis of a lignin *O*-demethylase examined adjacent hydroxyl groups donating protons in the formation of water in the reaction mechanism [15]. In $P450_{ZoGa}$, the protons could not come from the hydroxyl groups of the carbohydrate because they are too far away given the binding orientation of G6Me in the active site. Therefore, the structure of the complex of $P450_{ZoGa}$ and its substrate supports the conserved mechanism where protons are donated from the acid/alcohol motif and connect to solvent through the conserved water channel.

Our structure–function analysis of the carbohydrate-active P450 has created a model for the specificity of these enzymes. The conservation of specificity determinants in CYP236A suggests that the subfamily may be specific for G6Me. This observation is consistent with their general genomic context within agar utilization loci. Overall, our results bring a molecular understanding to the substrate recognition and activity of a carbohydrate-active cytochrome P450 monooxygenase.

Abbreviations

ITC, isothermal titration calorimetry.

Author Contribution

C.S.R. and J.-H.H. designed research; C.S.R. and L.R. performed research; C.S.R., L.R. and J.-H. H. analyzed data; C.S.R., L.R., U.B. and J.-H.H. wrote the manuscript.

Funding

J.-H.H. acknowledges funding by the Emmy-Noether-Program of the DFG, grant number HE 7217/1-1. We are especially grateful to the German Research Foundation (DFG) for funding through the Research Unit FOR2406.

Acknowledgements

Parts of this research were carried out at PETRAIII at DESY, a member of the Helmholtz Association (HGF). We thank Olga Lorbeer for the assistance in using beamline P11. We thank Prof. Dr Klaus Weisz and Andrea Funke for their support with the ITC measurements.

Competing Interests

The Authors declare that there are no competing interests associated with the manuscript.

References

1 Blanksby, S.J. and Ellison, G.B. (2003) Bond dissociation energies of organic molecules. *Acc. Chem. Res.* **36**, 255–263 https://doi.org/10.1021/ar020230d
2 Kocienski, P.J. (2005) *Protecting Groups (Thieme Publishers Series)*, Georg Thieme Verlag, Stuttgart
3 Urbanowicz, B.R., Pena, M.J., Ratnaparkhe, S., Avci, U., Backe, J., Steet, H.F. et al. (2012) 4-O-methylation of glucuronic acid in *Arabidopsis* glucuronoxylan is catalyzed by a domain of unknown function family 579 protein. *Proc. Natl Acad. Sci. U.S.A.* **109**, 14253–14258 https://doi.org/10.1073/pnas.1208097109
4 Stevenson, T.T. and Furneaux, R.H. (1991) Chemical methods for the analysis of sulphated galactans from red algae. *Carbohydr. Res.* **210**, 277–298 https://doi.org/10.1016/0008-6215(91)80129-B
5 Ndeh, D., Rogowski, A., Cartmell, A., Luis, A.S., Baslé, A., Gray, J. et al. (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* **544**, 65–70 https://doi.org/10.1038/nature21725
6 Panagiotopoulos, C., Repeta, D.J., Mathieu, L., Rontani, J.F. and Sempéré, R. (2013) Molecular level characterization of methyl sugars in marine high molecular weight dissolved organic matter. *Mar. Chem.* **154**, 34–45 https://doi.org/10.1016/j.marchem.2013.04.003
7 Abdel-Rahman, S.M., Marcucci, K., Boge, T., Gotschall, R.R., Kearns, G.L. and Leeder, J.S. (1999) Potent inhibition of cytochrome P-450 2D6-mediated dextromethorphan O-demethylation by terbinafine. *Drug Metab. Dispos.* **27**, 770–775 PMID:10383919
8 Dumitru, R., Jiang, W.Z., Weeks, D.P. and Wilson, M.A. (2009) Crystal structure of dicamba monooxygenase: a Rieske nonheme oxygenase that catalyzes oxidative demethylation. *J. Mol. Biol.* **392**, 498–510 https://doi.org/10.1016/j.jmb.2009.07.021
9 Häggblom, M.M., Berman, M.H., Frazer, A.C. and Young, L.Y. (1993) Anaerobic O-demethylation of chlorinated guaiacols by *Acetobacterium woodii* and *Eubacterium limosum*. *Biodegradation* **4**, 107–114 https://doi.org/10.1007/BF00702327
10 Neufeld, K., Marienhagen, J., Schwaneberg, U. and Pietruszka, J. (2013) Benzylic hydroxylation of aromatic compounds by P450 BM3. *Green Chem.* **15**, 2408 https://doi.org/10.1039/c3gc40838h
11 Jóźwik, I.K., Kiss, F.M., Gricman, Ł., Abdulmughni, A., Brill, E., Zapp, J. et al. (2016) Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*. *FEBS J.* **283**, 4128–4148 https://doi.org/10.1111/febs.13911
12 Sharrock, M., Debrunner, P.G., Schulz, C., Lipscomb, J.D., Marshall, V. and Gunsalus, I.C. (1976) Cytochrome $P450_{cam}$ and its complexes. Mössbauer parameters of the heme iron. *Biochim. Biophys. Acta* **420**, 8–26 https://doi.org/10.1016/0005-2795(76)90340-8
13 Schlichting, I., Berendzen, J., Chu, K., Stock, A.M., Maves, S.A., Benson, D.E. et al. (2000) The catalytic pathway of cytochrome p450cam at atomic resolution. *Science* **287**, 1615–1622 https://doi.org/10.1126/science.287.5458.1615
14 Hasemann, C.A., Kurumbail, R.G., Boddupalli, S.S., Peterson, J.A. and Deisenhofer, J. (1995) Structure and function of cytochromes P450: a comparative analysis of three crystal structures. *Structure* **3**, 41–62 https://doi.org/10.1016/S0969-2126(01)00134-4
15 Mallinson, S.J.B., Machovina, M.M., Silveira, R.L., Garcia-Borràs, M., Gallup, N., Johnson, C.W. et al. (2018) A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioconversion. *Nat. Commun.* **9**, 2487 https://doi.org/10.1038/s41467-018-04878-2
16 Reisky, L., Büchsenschütz, H.C., Engel, J., Song, T., Schweder, T., Hehemann, J.-H. et al. (2018) Oxidative demethylation of algal carbohydrates by cytochrome P450 monooxygenases. *Nat. Chem. Biol.* **14**, 342–344 https://doi.org/10.1038/s41589-018-0005-8
17 Urlacher, V.B. and Girhard, M. (2012) Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends Biotechnol.* **30**, 26–36 https://doi.org/10.1016/j.tibtech.2011.06.012
18 Hehemann, J.-H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M. and Michel, G. (2010) Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* **464**, 908–912 https://doi.org/10.1038/nature08937
19 Thomas, F., Bordron, P., Eveillard, D. and Michel, G. (2017) Gene expression analysis of *Zobellia galactanivorans* during the degradation of algal polysaccharides reveals both substrate-specific and shared transcriptome-wide responses. *Front. Microbiol.* **8**, 1–14 https://doi.org/10.3389/fmicb.2017.01808
20 Wargacki, A.J., Leonard, E., Win, M.N., Regitsky, D.D., Santos, C.N.S., Kim, P.B. et al. (2012) An engineered microbial platform for direct biofuel production from brown macroalgae. *Science* **335**, 308–313 https://doi.org/10.1126/science.1214547
21 Lewis, J.C., Bastian, S., Bennett, C.S., Fu, Y., Mitsuda, Y., Chen, M.M. et al. (2009) Chemoenzymatic elaboration of monosaccharides using engineered cytochrome P450BM3 demethylases. *Proc. Natl Acad. Sci. U.S.A.* **106**, 16550–5 https://doi.org/10.1073/pnas.0908954106
22 Kabsch, W. (2010) *XDS. Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 https://doi.org/10.1107/S0907444909047337
23 Evans, P.R. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 282–292 https://doi.org/10.1107/S090744491003982X
24 Xu, L.-H., Ikeda, H., Liu, L., Arakawa, T., Wakagi, T., Shoun, H. et al. (2015) Structural basis for the 4′-hydroxylation of diclofenac by a microbial cytochrome P450 monooxygenase. *Appl. Microbiol. Biotechnol.* **99**, 3081–3091 https://doi.org/10.1007/s00253-014-6148-y
25 Shimizu, H., Obayashi, E., Gomi, Y., Arakawa, H., Park, S.Y., Nakamura, H. et al. (2000) Proton delivery in NO reduction by fungal nitric-oxide reductase. Cryogenic crystallography, spectroscopy, and kinetics of ferric-NO complexes of wild-type and mutant enzymes. *J. Biol. Chem.* **275**, 4816–4826 https://doi.org/10.1074/jbc.275.7.4816
26 Cryle, M.J. and Schlichting, I. (2008) Structural insights from a P450 carrier protein complex reveal how specificity is achieved in the P450(Biol) ACP complex. *Proc. Natl Acad. Sci. U.S.A.* **105**, 15696–15701 https://doi.org/10.1073/pnas.0805983105
27 McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) *Phaser* crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 https://doi.org/10.1107/S0021889807021206

Biochemical Journal (2018) **475** 3875–3886
https://doi.org/10.1042/BCJ20180762

28 Adams, P.D., Afonine P, V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N. et al. (2010) *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 https://doi.org/10.1107/S0907444909052925

29 Afonine P, V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M. et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 https://doi.org/10.1107/S0907444912001308

30 Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 https://doi.org/10.1107/S0907444910007493

31 Cowtan, K. (2006) The *Buccaneer* software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002–1011 https://doi.org/10.1107/S0907444906022116

32 Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A. et al. (2011) *REFMAC* 5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 https://doi.org/10.1107/S0907444911001314

33 Collaborative Computational Project, Number 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 https://doi.org/10.1107/S0907444994003112

34 Wester, M.R., Johnson, E.F., Marques-Soares, C., Dijols, S., Dansette, P.M., Mansuy, D. et al. (2003) Structure of mammalian cytochrome P450 2C5 complexed with diclofenac at 2.1 Å resolution: evidence for an induced fit model of substrate binding. *Biochemistry* **42**, 9335–9345 https://doi.org/10.1021/bi034556l

35 Johnston, J.B., Kells, P.M., Podust, L.M. and de Montellano PR, O. (2009) Biochemical and structural characterization of CYP124: a methyl-branched lipid omega-hydroxylase from *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. U.S.A.* **106**, 20687–20692 https://doi.org/10.1073/pnas.0907398106

36 Meharenna, Y.T., Li, H., Hawkes, D.B., Pearson, A.G., De Voss, J. and Poulos, T.L. (2004) Crystal structure of P450cin in a complex with its substrate, 1,8-cineole, a close structural homologue to D-camphor, the substrate for $P450_{cam}$. *Biochemistry* **43**, 9487–9494 https://doi.org/10.1021/bi049293p

37 Poulos, T.L., Finzel, B.C., Gunsalus, I.C., Wagner, G.C. and Kraut, J. (1985) The 2.6-Å crystal structure of *Pseudomonas putida* cytochrome P-450. *J. Biol. Chem.* **260**, 16122–16130 PMID:4066706

38 Tsuzuki, S. and Fujii, A. (2008) Nature and physical origin of CH/π interaction: significant difference from conventional hydrogen bonds. *Phys. Chem. Chem. Phys.* **10**, 2584–2594 https://doi.org/10.1039/b718656h

39 Wilderman, P.R., Shah, M.B., Jang, H.-H., Stout, C.D. and Halpert, J.R. (2013) Structural and thermodynamic basis of (+)-α-pinene binding to human cytochrome P450 2B6. *J. Am. Chem. Soc.* **135**, 10433–10440 https://doi.org/10.1021/ja403042k

40 Dam, T.K. and Brewer, C.F. (2002) Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. *Chem. Rev.* **102**, 387–429 https://doi.org/10.1021/cr000401x

41 Ficko-Blean, E. and Boraston, A.B. (2006) The interaction of a carbohydrate-binding module from a *Clostridium perfringens N*-acetyl-β-hexosaminidase with its carbohydrate receptor. *J. Biol. Chem.* **281**, 37748–37757 https://doi.org/10.1074/jbc.M606126200

42 Hagel, J.M. and Facchini, P.J. (2010) Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy. *Nat. Chem. Biol.* **6**, 273–275 https://doi.org/10.1038/nchembio.317

43 Gotoh, O. (1992) Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J. Biol. Chem.* **267**, 83–90 PMID:1730627

44 Imai, M., Shimada, H., Watanabe, Y., Matsushima-Hibiya, Y., Makino, R., Koga, H. et al. (1989) Uncoupling of the cytochrome P-450cam monooxygenase reaction by a single mutation, threonine-252 to alanine or valine: possible role of the hydroxy amino acid in oxygen activation. *Proc. Natl Acad. Sci. U.S.A.* **86**, 7823–7827 https://doi.org/10.1073/pnas.86.20.7823

45 Miao, Y. and Baudry, J. (2011) Active-site hydration and water diffusion in cytochrome $P450_{cam}$: a highly dynamic process. *Biophys. J.* **101**, 1493–1503 https://doi.org/10.1016/j.bpj.2011.08.020

46 Degregorio, D., Sadeghi, S.J., Di Nardo, G., Gilardi, G. and Solinas, S.P. (2011) Understanding uncoupling in the multiredox centre P450 3A4-BMR model system. *J. Biol. Inorg. Chem.* **16**, 109–116 https://doi.org/10.1007/s00775-010-0708-0

47 Wienkers, L.C. and Heath, T.G. (2005) Predicting in vivo drug interactions from in vitro drug discovery data. *Nat. Rev. Drug. Discov.* **4**, 825–833 https://doi.org/10.1038/nrd1851

48 Bell, S.G., Chen, X., Sowden, R.J., Xu, F., Williams, J.N., Wong, L.L. et al. (2003) Molecular recognition in (+)-α-pinene oxidation by cytochrome $P450_{cam}$. *J. Am. Chem. Soc.* **125**, 705–714 https://doi.org/10.1021/ja028460a

49 Hales, D.B., Ho, B. and Thompson, J.A. (1987) Inter- and intramolecular deuterium isotope effects on the cytochrome P-450-catalyzed oxidative dehalogenation of 1,1,2,2-tetrachloroethane. *Biochem. Biophys. Res. Commun.* **149**, 319–325 https://doi.org/10.1016/0006-291X(87)90369-X

50 Ashkenazy, H., Erez, E., Martz, E., Pupko, T. and Ben-Tal, N. (2010) Consurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, W529–W533 https://doi.org/10.1093/nar/gkq399

51 Crooks, G.E., Hon, G., Chandonia, J.-M. and Brenner, S.E. (2004) Weblogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 https://doi.org/10.1101/gr.849004