Different Binding Modes of Two Flaviolin Substrate Molecules in Cytochrome P450 158A1 (CYP158A1) Compared to CYP158A2[†,‡]

Bin Zhao,[*,§] David C. Lamb,[#] Li Lei,[§] Steven L. Kelly,[#] Hang Yuan,[§] David L. Hachey,[§,||,⊥] and Michael R. Waterman[§]

Departments of Biochemistry and Pharmacology, and the Mass Spectrometry Center and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0146, and Institute of Life Science and School of Medicine, University of Wales Swansea, United Kingdom

Received April 13, 2007; Revised Manuscript Received May 23, 2007

ABSTRACT: Cytochrome P450 158A2 (CYP158A2) has been shown to catalyze an unusual oxidative C−C coupling reaction to polymerize flaviolin and form highly conjugated pigments (three isomers of biflaviolin and one triflaviolin) in *Streptomyces coelicolor* A3(2) which protect the soil bacterium from deleterious effects of UV irradiation (Zhao B. et al. (2005) *J. Biol. Chem.* 280, 11599−11607). The present studies demonstrate that the subfamily partner CYP158A1, sharing 61% amino acid identity with CYP158A2, can also catalyze the same flaviolin dimerization reactions, but it generates just two of the three isomers of biflaviolin that CYP158A2 produces. Furthermore, the two CYP158A1 products have very different molar ratios compared with the corresponding CYP158A2 products, indicating that each enzyme maintains its own stereo- and regiospecificity. To find an explanation for these differences, three CYP158A1 structures have been solved by X-ray crystallography and have been compared with those for CYP158A2. The structures reveal surprising differences. Particularly, only one flaviolin molecule is present close to the heme iron in CYP158A1, and the second flaviolin molecule binds at the entrance of the putative substrate access channel on the protein distal surface 9 Å away. Our work describes two members of the same P450 subfamily, which produce the same products by oxidative C−C coupling yet show very different structural orientations of substrate molecules in the active site.

It is a daunting problem to explore and confirm the biological function and endogenous activity of the multitude of orphan enzymes identified from the large number of genomic sequencing studies within the past decade. A particularly good example of this is the more than 6500 genes within the cytochrome P450 (P450 or CYP[1]) superfamily (http://drnelson.utmem.edu/CytochromeP450.html), about which we know little concerning the biological function and endogenous substrates. For example, there are 18 *CYP* genes in the model actinomycete *Streptomyces coelicolor* A3(2) (*1*, *2*), whose biological roles are largely unknown. Nevertheless, the increasing availability of various approaches (*3*−*5*), such as evaluation of genomic context, gene knockin/knockout techniques associated with measurement of metabolic response, protein structural data, and sequence relatedness, have provided useful clues for deciphering the biological function and activity of such orphan enzymes. In one example, we have shown that cytochrome P450 158A2 (CYP158A2) from *S. coelicolor* catalyzes oxidative C−C coupling reactions of the phenol-like molecule flaviolin to biflaviolin and triflaviolin *in vitro*, which was originally predicted from analysis of the highly conserved operon containing this P450 gene (*6*, *7*). These secondary metabolites are thought to afford physical protection to this soil bacterium against the deleterious effects of UV irradiation on genetic integrity (*8*, *9*). Funa et al. have observed a lack of phenotype for CYP158A2 knockout, but for CYP158B1 (*rppB*) in *S. griseus*, they have shown that knockout of the gene compromises the organism upon UV exposure (*9*). We have confirmed the CYP158A2 knockout result by observing that the metabolic profiles between the CYP158A2 knockout and the wild type strains using different culture media (minimal media, YEME) showed no effects on pigmentation, metabolic profile measured by mass spectrometry, or any other discernible difference in phenotype. One possible explanation for these results is that the appropriate laboratory conditions generating a phenotype have not been used. Nevertheless, these results have led us to speculate that another unknown enzyme can assume the role of CYP158A2 during the life cycle of the *S. coelicolor* knockout. The logical candidate would seem to be CYP158A1, which, while being a member of the same subfamily as that of CYP158A2, is not a member

[†] This work was supported by National Institutes of Health Grants GM69970 (to M.R.W.) and ES00267 (to M.R.W.).

[‡] The coordinates and associated structure factors have been deposited with the Protein Data Bank with the following accession codes: CYP158A1 ligand-free form structure, 2NZA; flaviolin-bound complex, 2NZ5; imidazole complex, 2DKK.

* To whom to correspondence should be addressed. Tel: 615-343-4644. Fax: 615-343-0704. E-mail: bin.zhao@vanderbilt.edu.

[§] Department of Biochemistry, Vanderbilt University School of Medicine.

[||] Department of Pharmacology, Vanderbilt University School of Medicine.

[⊥] Mass Spectrometry Center and Center in Molecular Toxicology, Vanderbilt University School of Medicine.

[#] University of Wales Swansea.

[1] Abbreviations: P450 or CYP, cytochrome P450 monooxygenase; F_o, observed structure factor; F_c, calculated structure factor.

10.1021/bi7006959 CCC: $37.00
Published on Web 07/06/2007

of any operon being unlinked in the chromosome (*1*). CYP158A1 shares 61% amino acid sequence identity with CYP158A2, and we have examined whether CYP158A1 can catalyze the C−C coupling of flaviolin in a manner analogous to that of CYP158A2 and hence might compensate the CYP158A2 knockout, affording a second level of protection to the organism.

Herein, we show that CYP158A1 can catalyze *in vitro* oxidative C−C coupling reactions, which are very similar to those of CYP158A2 using flaviolin as substrate. However, when we solved the crystal structures of both the ligand-free and ligand-bound forms of CYP158A1 (to the inhibitor imidazole and the substrate flaviolin), comparison of the CYP158A1 and CYP158A2 structural architecture reveals that these two monooxygenases from the same subfamily accommodate substrate flaviolin molecules spatially differently, even though they generate closely related or identical products. Our results indicate that these two members of the CYP158A subfamily seemingly use different molecular orientations for oxidative C−C dimerization and present an interesting challenge in understanding how the structural differences of the active site lead to closely related enzymatic stereo- and regiospecificity.

EXPERIMENTAL PROCEDURES

Gene Isolation, Expression, and Purification of CYP158A1. The gene encoding *S. coelicolor* A3(2) CYP158A1, genetically engineered to incorporate six histidine codons at the 3′-terminus, was isolated by PCR and subcloned into the *Escherichia coli* expression vector pET17b (Novagen, Madison, WI) using the NdeI and HindIII sites (*2*). Recombinant proteins were produced in *E. coli* BL21 (DE3) pLysS competent cells. Briefly, the cells were cultured overnight in LB broth containing 50 μg/mL ampicillin. The transformed *E. coli* inoculated (1:100) in 3 L of Terrific Broth containing 100 μg/mL ampicillin were grown at 37 °C and 240 rpm until the optical density at 600 nm reached around 1.0. After induction with 1 mM isopropyl-β-D-thiogalactopyranoside and the addition of 1 mM δ-aminolevulinic acid for heme synthesis, growth was continued for an additional 20 h at 27 °C and 190 rpm. The cells were harvested by centrifugation and resuspended in 100 mL of lysis buffer (50 mM Tris-HCl at pH 8.2 containing 0.5 mM EDTA, 20% (v/v) glycerol, 3 mM imidazole, and 100 mM NaCl). Cells were broken by freeze-thawing and the CYP158A1 cytosol isolated following centrifugation at 100,000*g*. The soluble CYP158A1 was purified by metal (Ni^{2+}) affinity chromatography (Qiagene) and the fraction was further purified by S-sepherose/Q-sepherose (Amersham Bioscience) chromatography as described (*7*).

Spectral Substrate Binding and Activity Assays. Spectral and catalytic activity assays were carried out as reported previously for CYP158A2 using flaviolin as a substrate (*7*). Absorbance difference spectra were recorded using a double beam Shimadzu UV-2401PC spectrophotometer. The interaction of flaviolin with CYP158A1 was examined by perturbation of the heme Soret spectrum. CYP158A1 (2.5 μM) in 20 mM Tris-HCl (2.0 mL, pH 8.2) was divided between two tandem cuvettes. After thermal equilibration at 25 °C, a baseline was established between 350 and 450 nm, and sequential additions (1−5 μL) of a concentrated aqueous solution of flaviolin (1 mM dissolved in 5% (v/v) methanol in 20 mM Tris-HCl at pH 8.2) were added to the sample cuvette to give a final ligand concentration in the range of 1−50 μM. An equal volume of 5% (v/v) methanol in 20 mM Tris-HCl (pH 8.2) was added to the reference cuvette, and the difference spectrum recorded after each titration. The Hill coefficient was obtained from a plot of $\log(\Delta A_{(382-418)}/[\Delta A_{max} - \Delta A_{(382-418)}])$ versus log[flaviolin]. K_d values were estimated by fitting plots of $\Delta A_{382-418}$ versus [flaviolin]. For activity assays, CYP158A1 (1 nmol), flavodoxin (10 nmol), and flavodoxin reductase (2 nmol) were reconstituted in 400 μL of 20 mM Tris-HCl buffer (pH 8.2) and flaviolin (0.26 μmol). Flaviolin was generated as previously described (*10*, *11*). Following incubation of the reaction mixture for 5 min on ice, the reconstituted enzyme solution was placed in a shaking water bath at 37 °C. The reaction was initiated by the addition of NADPH to a final concentration of 5 mM and was conducted for 2 h in a 1.5-mL tube, at which time the reaction was terminated by the addition of 4 μL of concentrated HCl. Subsequently, the mixtures were extracted three times with 400 μL of ethyl acetate. Product formation was analyzed by LC/MSMS (liquid chromatography−mass spectrometry) as described previously (*7*). UV detection was at 254 nm. Positive control incubations utilizing CYP158A2 were carried out as described above.

Crystallization and Data Collection. Crystals of CYP158A1 were obtained using the hanging-drop vapor diffusion method, in which 2 μL of a 20 mg/mL protein solution was mixed with an equal volume of 0.1 M bis-Tris (pH 6.5) and 0.8−1.2 M ammonium dihydrogen phosphate. At 20 °C, the ligand-free crystals appeared within a few days; for the generation of the substrate flaviolin-bound crystals, 1 mM flaviolin was added to the protein solution; for inhibitor-bound imidazole crystals, bis-Tris (pH 6.5) was replaced by 0.1 M imidazole (pH 8.0). All three crystals belong to the monoclinic space group *C*2 with different unit cell parameters (Table 1). Full diffraction data were collected at 100 K at the Southeast Regional Collaborative Access Team (SER-CAT) 22-BM beamline at the Advanced Photon Source, Argonne National Laboratory, Argonne, IL. The X-ray data were processed and scaled with the HKL package programs HKL2000 and MOSFILM (*12*, *13*).

Structure Determination. The CYP158A1 structure was solved by molecular replacement using the program PHASER (*14*) and the substrate-free CYP158A2 structure (protein data bank code: 1SE6) as a search model. The amino acid sequence identity between CYP158A1 and CYP158A2 sequences is 61% in a protein sequence alignment. The initial model was built in O (*15*), and refinement was performed using CNS1.1 (*16*). For the ligand-free structure and the flaviolin-bound complex, there were two molecules of CYP158A1 in the asymmetric unit. However, only one molecule of CYP158A1 is in the asymmetric unit of the imidazole complex. Final refinement statistics of all three structures are given in Table 1. The coordinates and associated structure factors have been deposited with the Protein Data Bank (accession codes: CYP158A1 ligand-free form structure, 2NZA; flaviolin-bound complex, 2NZ5; imidazole complex, 2DKK). Figures were generated using SETOR (*17*).

Table 1: Data Collection and Refinement Statistics

	substrate-free	imidazole bound complex	flaviolin bound complex
	data collection statistics		
space group	*C*2	*C*2	*C*2
unit cell (Å)	186.61	103.97	192.27
	44.34	44.35	44.86
	130.62	102.11	105.05
beta value (°)	98.23	114.44	97.76
molecules/asymmetric unit	2	1	2
data resolution (Å)	2.9	1.97	2.35
redundancy[a]	3.3 (2.9)	4.0 (3.6)	3.2 (2.5)
completeness %[a]	88.3 (73.7)	99 (96.3)	97 (97)
$I/\sigma(I)$[a]	9.7 (3.9)	15.1 (3.6)	17.2 (4.8)
*R*merge %[a]	10.3 (20)	6.5 (22.3)	7.2 (24.5)
	refinement statistics		
no. of reflections used in refinement	23648	30268	50519
no. of water molecules	228	287	265
protein atoms[b]	6058	3094	6218
heme atoms	86	86	86
ligand atoms	0	10[c]	30[d]
R_{work} %	19.7	22.8	23.5
R_{free} %	26.8	28.5	29.3
rmsd in bond lengths (Å)	0.008	0.006	0.008
rmsd in bond angles (°)	1.4	1.3	1.42

[a] The values for the highest resolution shell are in parentheses. [b] Residues 13−407(A) and 13−81/91−409(B) of substrate-free; residues 13−410 of inhibitor imidazole complexes; residues 13−410 (A and B) of substrate-bound. [c] Two imidazole molecules. [d] Two flaviolin molecules.

RESULTS

Substrate Binding and Catalytic Activity of CYP158A1. On the basis of the CYP158A1 sequence relatedness with CYP158A2, we predicted that flaviolin, the established endogenous substrate of CYP158A2, might be a substrate for CYP158A1. When examining flaviolin binding to CYP158A1, a typical type I P450 binding spectrum, which shifts the low-spin heme Soret peak at 418 nm to a high-spin peak at 382 nm, was observed (Figure S1, Supporting Information). The binding affinity constant (*K*d) was estimated to be 10.5 μM, which is comparable to that for flaviolin binding to CYP158A2 (7.3 μM). Interestingly, however, the Hill coefficient in CYP158A1 was 1.1 ± 0.1, indicating one flaviolin molecule binding to the enzyme active site compared to 1.7 in CYP158A2, which binds two flaviolin molecules in the active site as clearly demonstrated by X-ray crystallography (*7*). Analysis of flaviolin product formation by CYP158A1 using electrospray mass spectrometry, as described previously in detail (*7*), indicated two different reaction products (Figure 1), which are both isomers of biflaviolin with a mass of 410 (MH^+ 411, P1 *tR* 14.8 min; P2 *tR* 15.2 min). The products were initially identified from their retention time and mass spectrometry data with respect to corresponding product amounts characterized from CYP158A2 activity (Figures S2 and S3, Supporting Information) and were confirmed by MS/MS data and assigned as 3,8-biflaviolin (P1) and 3,3-biflaviolin (P2). These results indicate that CYP158A1 can catalyze the same oxidative C−C coupling reaction as that described previously for

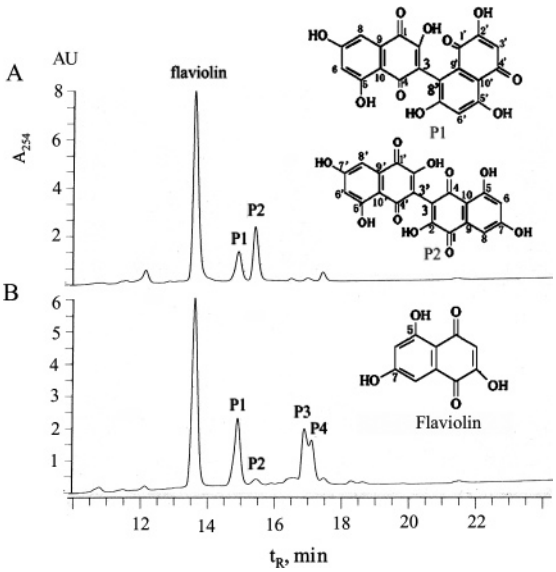

FIGURE 1: Catalytic activity of CYP158A1. Oxidation reactions were carried out as described under Experimental Procedures. (A) Product profile obtained following flavodoxin and flavodoxin reductase supported reactions using CYP158A1 (1 nmol) and flaviolin (0.26 μmol) at 37 °C for 2 h. (B) The same reaction using CYP158A2 as control. Here, the four major products are noted as P1, P2, P3, and P4, and the known chemical structures of P1 and P2 are shown. The products in panel A are found to be the same as P1 and P2 by both HPLC and MS analysis.

CYP158A2 (*7*). It is worth noting that the relative ratio of one biflaviolin product (P1) was reduced to 40%, and P2 was increased about 10-fold compared to the corresponding products generated in CYP158A2 assays. Furthermore, the third biflaviolin isomer P3 and triflaviolin (P4) generated in CYP158A2 catalysis were not observed in reconstituted CYP158A1 experiments. The turnover number for flaviolin dimerization in CYP158A1 catalysis was estimated to be 0.9 min^{-1} on the basis of the rate of substrate consumption under the *in vitro* conditions used here (1.4 min^{-1} for CYP158A2). Our results indicate that CYP158A1/A2 can accommodate the same substrate and generate the same products but with very different molar ratios. Each enzyme has its own catalytic specificity, but there is little difference in the overall efficiency of turnover of flaviolin between them.

CYP158A1 Crystal Structures. All three CYP158A1 crystal structures (ligand-free, bound to the inhibitor imidazole, and bound to the substrate flaviolin) exhibited the traditional P450 fold seen in CYP158A2 (*7*) and other known P450 structures (Figure 2). The overall ligand-free structure shows an open conformation also observed for ligand-free CYP158A2 (*7*). Surprisingly, the inhibitor-bound imidazole and substrate-bound flaviolin complexes do not show substantial conformational changes in the region of the BC loop, β sheet 4, and F/G helices, which are associated with substrate recognition and binding in most other P450 structures including CYP158A2 (*7*, *18*, *19*) (Figure 2). This is a significant structural difference from CYP158A2 structures, in which ligands induce typical P450 conformational movements of secondary structural elements (Figure 3).

Although product formation by CYP158A1 (Figure 1) led us to expect to find two flaviolin molecules bound in the CYP158A1 active site pocket as in the CYP158A2−flaviolin-bound complex (*7*), a Hill coefficient of about 1.0 made this uncertain. As shown in Figure 4, the electron

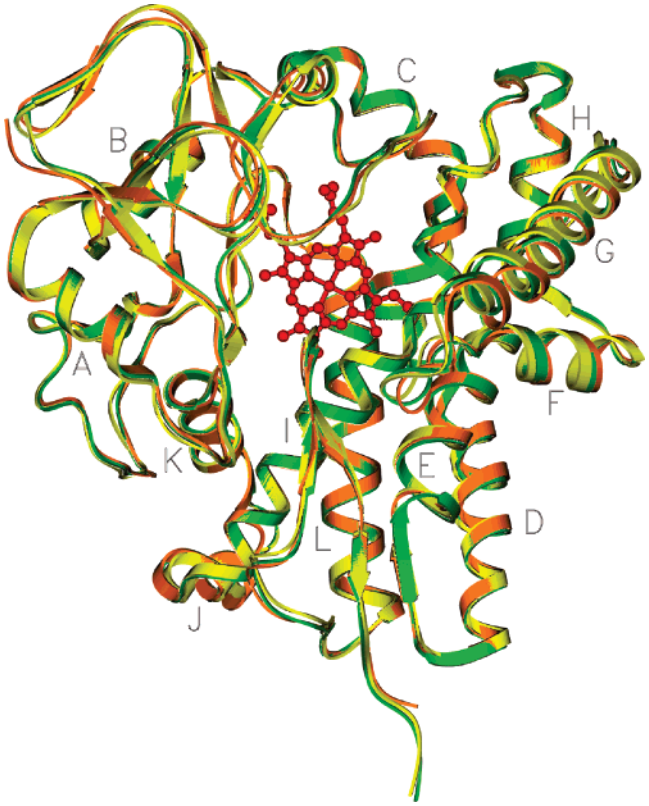

FIGURE 2: Overlay ribbon diagrams of CYP158A1 ligand-free and ligand-bound (to inhibitor imidazole and to substrate flaviolin structures). All three X-ray crystal structures show a typical cytochrome P450 fold. The ligand-free structure is represented in orange, the flaviolin-bound structure in green, and the imidazole-bound structure in yellow. Heme is the red ball and stick model. The ligands flaviolin and imidazole are omitted in order to more clearly show the overall structures. Note the global structural conformations of CYP158A1 are only slightly changed upon ligand binding.

density map of the CYP158A1−flaviolin-bound complex does reveal two flaviolin molecules binding to CYP158A1. However, only one flaviolin molecule is situated in the proximal portion of the active site over the heme, adopting a position similar to that of the proximal flaviolin molecule in the CYP158A2−flaviolin complex structure. The key conserved residue Arg291 (Arg288 in CYP158A2) for flaviolin anchoring as demonstrated in CYP158A2 plays the same role in CYP158A1, being oriented toward the active site pocket and directly forming hydrogen bonds with the carbonyl C1 and 2-OH groups of flaviolin (Figure 4). In CYP158A1, His290 provides a hydrogen bond to carbonyl C4 in the proximal flaviolin molecule to assist flaviolin binding into the active site instead of that mediated by a water molecule in the CYP158A2−flaviolin complex (*7*).

Surprisingly, the distal flaviolin (F2) in CYP158A1 is bound at the substrate entrance, which is located between the BC- and FG-loops in a cleft on the distal surface of the protein (Figure 4). It is partially exposed to the bulk solvent and is approximately 9 Å away from the proximal flaviolin (F1). These flaviolin structural features may explain the Hill coefficient close to 1 because the distal flaviolin molecule is far away from heme and the proximal flaviolin molecule so that it probably does not affect the Soret band spectra during titration experiments, and therefore, it is very possible that the distal flaviolin binds to the enzyme silently. In CYP158A2, the two flaviolin molecules form a quasi-planar stack including the heme molecule and are packed very tightly, approximately 3.5 Å apart from each other (Figure 5). The average crystallographic temperature factors of the two flaviolin molecules in CYP158A1 are 54.8 $Å^3$, compared with 16.4 $Å^3$ for those in the CYP158A2 structure, indicating that the substrate flaviolin molecules in the CYP158A1 structure may either be partially occupied, or the flaviolin molecules occupy multiple positions in their binding sites. Spectral titration data do indicate that flaviolin binds about as effectively to CYP158A1 as to CYP158A2. This supports a conclusion that they occupy multiple positions in the active site. In addition, there are two complex structures within a CYP158A1 asymmetric unit, and the F1 molecules accommodate sharply different orientations over the heme in the active site, and one of the F2 molecules is flipped 180° from the other (Figure 6). These results further confirm that flaviolin molecules occupy more than one position in different CYP158A1 populations.

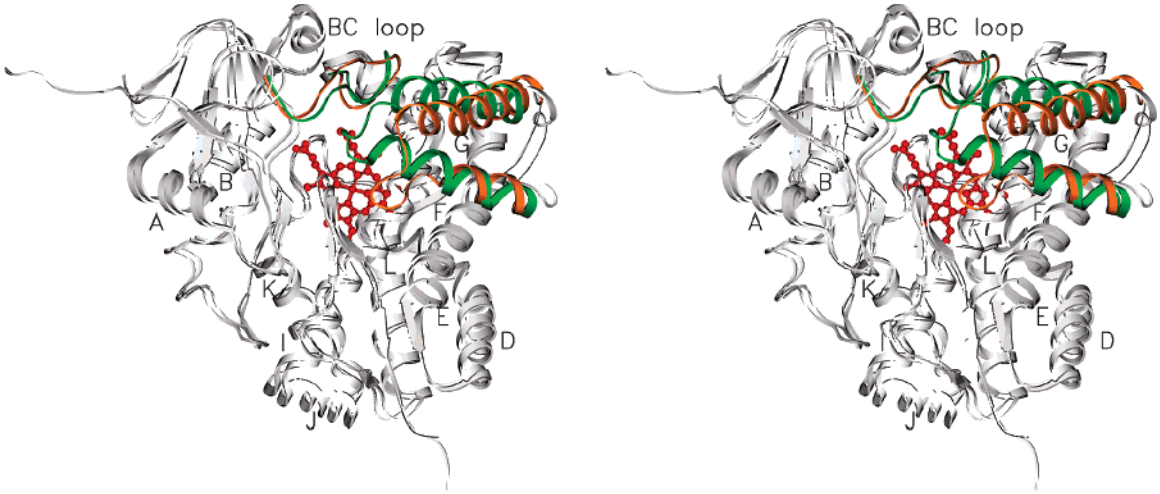

FIGURE 3: Stereoview of the superimposition of CYP158A1 and CYP158A2 flaviolin-bound structures. The BC loop and F and G helices are in orange (CYP158A1) and in green (CYP158A2) showing structural differences between these two P450s. The movement of the BC loop and the F and G helices in CYP158A2 close the active site compared to that in CYP158A1. The flaviolin molecules are omitted for clarity.

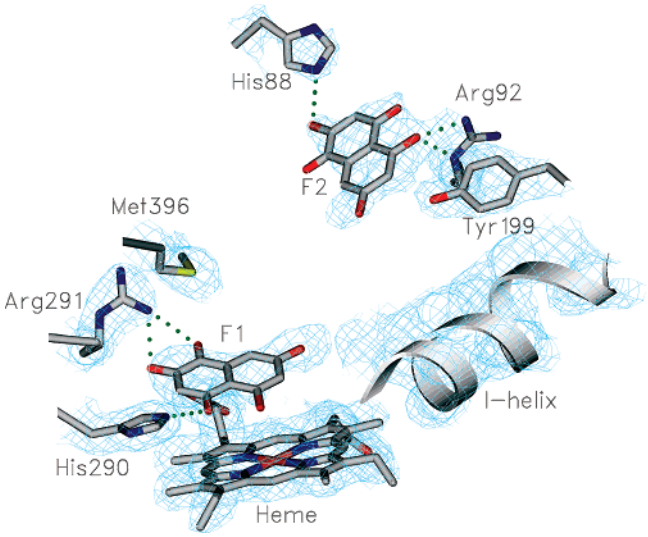

FIGURE 4: Electron density map of flaviolin in the structure of substrate-bound CYP158A1. The electron density map was calculated using σA-weighted $2|F_o| - |F_c|$ coefficients and is contoured at 1.0 σ shown in cyan. The proximal flaviolin and distal flaviolin molecules are denoted F1 and F2, respectively. The side chain atoms, flaviolin, and heme are rendered as stick figures (gray, carbon; blue, nitrogen; yellow, sulfur; red, oxygen and iron). Potential hydrogen bonds are shown as dotted green lines.

Detailed Comparison with CYP158A2. Although CYP158A1 and CYP158A2 belong to the same CYP subfamily sharing 77% sequence similarity (61% identity), there are a number of features that differ between their structures because of the lengths and orientations of several secondary structural elements. The most notable structural differences involve the BC loop and F helix, which are located on the distal face of the protein. As shown in Figure 5, residues 86−95 in the BC loop region show a different conformation between these two P450s, folding into the active site in CYP158A2 (residues 82−90) but oriented away from the active site and pointing to the bulk solvent in CYP158A1. One reason for this is that the residue corresponding to Ile87 in CYP158A2 is substituted by the long side chain in Lys90 of CYP158A1. The side chain of Ile87 points to the active site and contacts F2 in CYP158A2. However, the bulkier side chain of Lys90 in CYP158A1 is oriented away from the BC loop and toward the distal surface of the protein. There would be steric collision with other active site residues if Lys90 served a role in CYP158A1 similar to that of Ile87 in CYP158A2. Lys90 could be the main reason that the BC-loop conformation does not dip into the active site in CYP158A1. In addition, Arg92 in CYP158A1 occupies approximately the same position in three dimensions as Arg90 in CYP158A2. The guanidium group of Arg90 hydrogen bonds with Met75, which is also located in the BC loop, locking the orientation of the side chain of Arg90 away from the G helix in CYP158A2. However, the residue corresponding to Met75 in CYP158A2 is Thr78 in CYP158A1. This side chain is too short to hydrogen bond with Arg92 so that the bulkier side chain of Arg92 is oriented toward and contacts the G helical region in CYP158A1 (Figure S4, Supporting Information). This may explain why the G helix is pushed farther away from the active site compared to CYP158A2, which is consistent with a more open conformation even when flaviolin or imidazole molecules bind to the enzyme.

Another difference is that the F helix of CYP158A2 is significantly longer than that of CYP158A1 (Figure 5). There is one less helical turn at the C-terminal end of the F helix, which results in a longer loop structural motif between the F and G helices in CYP158A1. This longer loop is located above the I helix and substantially increases the interactions between the F/G region and β sheet 4 and forms a different active site shape. The side chains of Arg401 hydrogen bond with the carbonyl oxygen of Thr184 in CYP158A1. Ile182 also lies in this extended loop region, but its side chain points away from the active site. In contrast, the side chain of Leu179 is shown to interact with substrate F2 in CYP158A2 and orients into the active site. Furthermore, there is a significant shift along the helical axis in CYP158A1 so that the E/F loop is pushed away from the structural core in comparison with that in CYP158A2. The F helix shift is due to its short length. The longer FG loop releases the stretch of the observed shift of the F helix so that the G helix adopts approximately the same length and number of helical turns as those in CYP158A2.

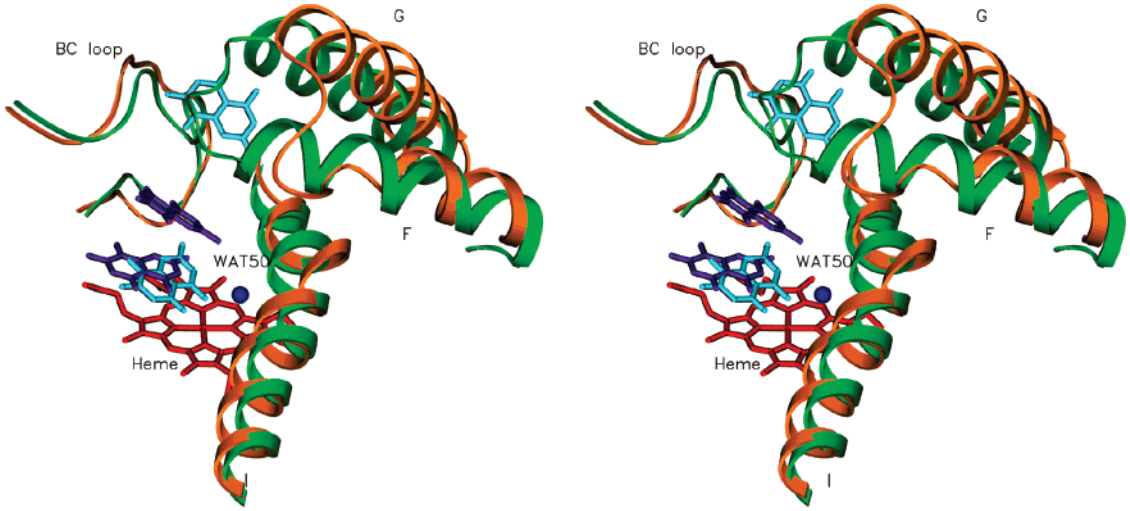

FIGURE 5: Stereoview of the superimposition of two flaviolin molecules in CYP158A1 (orange) and CYP158A2 (green). The two flaviolin molecules are presented as stick models and are colored cyan in CYP158A1 and blue in CYP158A2. Heme is colored red, and water molecule 50 in CYP158A1 is blue.

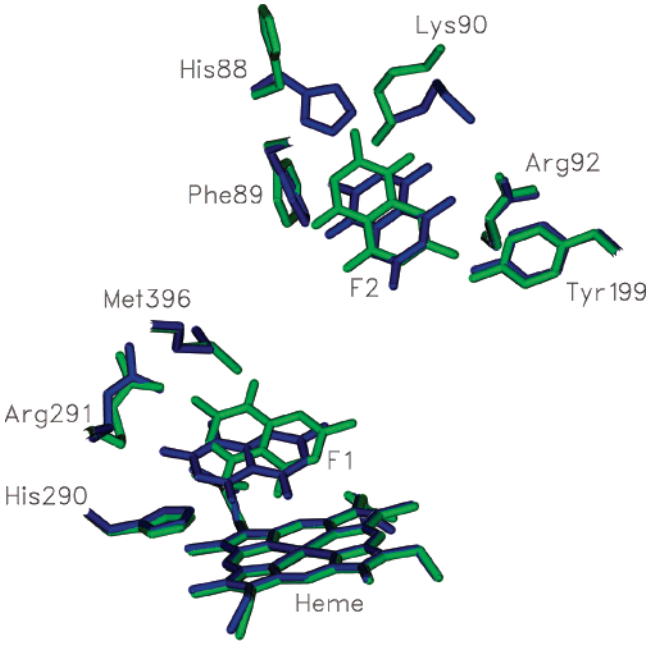

FIGURE 6: Overall view of two flaviolin molecules and key residues showing a comparison between the two CYP158A1−flaviolin complex structures in the asymmetric unit. The flaviolin molecules and residues from one CYP158A1 molecule are colored blue, and those from the other are colored green. Both F1 and F2 accommodate different positions in the active site.

Active Site of CYP158A1. The active site of CYP158A1 is substantially different from that of CYP158A2 owing to the presence of residues Lys90 and Met396 and the orientation of Arg92, resulting in substantial changes in the volume space of the active site. The overall volume of the CYP158A active site pockets calculated with VOIDOO (*20*) is quite different, 1500 Å for CYP158A1, which is three times larger than the 495 Å for CYP158A2. This is because the flaviolin-bound CYP158A1 structure is a much more open conformation compared to the closed conformation in the flaviolin-bound CYP158A2 structure. The CYP158A1 active site volume is apparently plastic and dynamic during the catalytic cycle. Nevertheless, the volume occupied by F1 is about 150 Å, not leaving sufficient space for F2 to bind within the proximal active site as occurs in CYP158A2 because the long side chain of Met396 limits the size of the F1 binding site. This is consistent with the relatively high B factor of the flaviolin molecules, which appears not to be fixed as tightly as those in the closed active site of CYP158A2.

It is surprising to find that only the proximal flaviolin molecule is over the heme in the proximal portion of the active site pocket. This proximal flaviolin molecule slides toward the I helix compared to the proximal flaviolin in CYP158A2. This movement leads to an opening for an ordered water molecule (WAT50), which creates a helical kink in the I helix. The WAT50 donates a hydrogen bond to the carbonyl oxygen of Gly248. This kink in the I helix is not present in CYP158A2 (Figure 5). This proximal substrate binding site contains several important residues that have been implicated in substrate recognition and binding: Met396, Ile397, Leu296, Arg291, His289, and Ala248 creating the F1 binding site in CYP158A1, corresponding to Leu393, Ile394, Leu293, Arg288, His287, and Ala245, respectively, in CYP158A2. A major difference is the presence of the long side-chain residue Met396 in CYP158A1 in place of Leu393 in CYP158A2. The orientation of the Met396 side chain in CYP158A1 is significantly different depending on the presence of a ligand (Figure 7). In the ligand-free form, the side chain of Met396 is rotated out of the active site to open the substrate channel, but this side chain will point into the active site and close the channel when substrate binds to the active site. In this position, the side chain of Met396 appears to present a steric barrier to the movement of substrate in or out of the active site pocket. We propose the Met396 may serve as a gating residue to control the opening or closing of the substrate access channel, which is triggered by ligand binding because the overall structural conformation does not change very much in CYP158A1 upon ligand binding. The repositioning of the Met396 side chain is one of the major conformational changes in CYP158A1 upon ligand binding.

The distal flaviolin binding region in CYP158A1 comprises Lys90 and His88 in the BC loop as well as Lys195 and Tyr199 located in the G helix, all of which contact with the flaviolin molecule within 4 Å. These residues form a small cavity on the distal surface, which binds F2 and

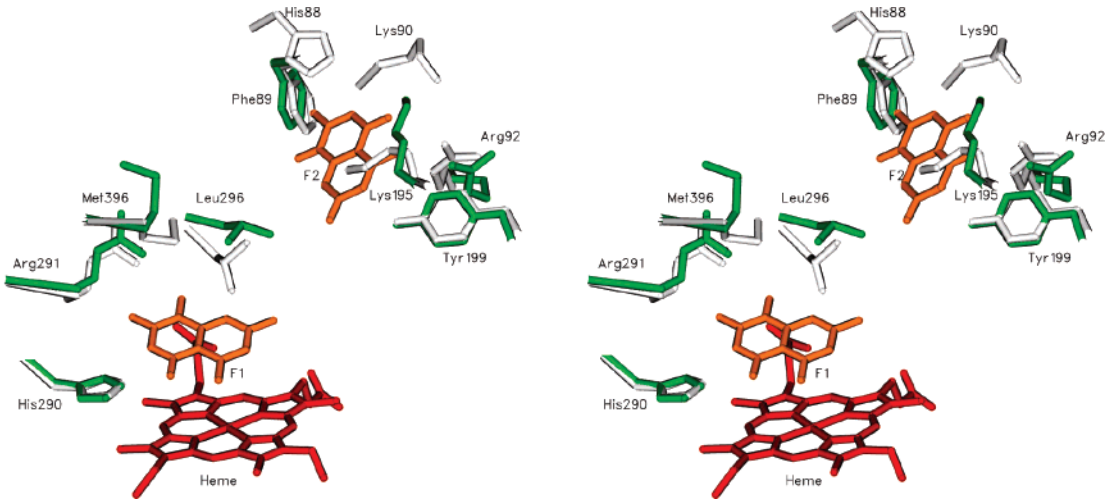

FIGURE 7: Stereoview of superimposed structures of ligand-free and flaviolin-bound CYP158A1. The green and gray stick models show key residues in ligand-free and flaviolin-bound CYP158A1, respectively. Flaviolin molecules are shown in orange. The orientation of side chains (Met396, Leu296, His88, Arg291, Arg92, and Lys295) is different upon the substrate flaviolin binding to the protein. His88 and Lys90 in the ligand-free structure are not shown because of the absence of electron density.

surprisingly also binds a second molecule of imidazole (Figure S5, Supporting Information). The proximal imidazole shows normal azole binding directly to the heme iron, while the distal imidazole binds in the same cleft that binds F2 with His88 and Lys90 on one side and Lys195 and Try199 on the other side (Figure S6, Supporting Information). The imidazole nitrogen atom provides a hydrogen bond with the carbonyl oxygen atom of Lys90. It is worth noting that in the ligand-free structure, this small cavity exists as well but is empty, indicating that the conformational differences between CYP158A1 versus CYP158A2 in the BC loop and FG helical region create a cavity to trap the ligand on the distal surface of the protein. No overall structural change in the loop is triggered upon ligand binding, which confirms that the distal binding site is a unique structural feature for CYP158A1. If these secondary elements moved as in CYP158A2, the distal flaviolin or imidazole molecules would be pushed out of the distal binding site. In addition, His88 has two different conformations in the two structures in one asymmetric unit. This could support the distal flaviolin molecules accommodating flipped orientation positions in the two complex molecules, although poor electron density makes this hard to define (Figure 6), which may also explain why flaviolin molecules show relatively high crystallographic temperature factors. In addition, this weak contact feature could be of benefit for distal flaviolin binding and release. The functional relevance of this distal binding site is still unclear. Perhaps, it is bifunctional in ensuring that substrate bound to this distal binding site subsequently dissociates and moves into the active site as well as orienting the neutral flaviolin molecule for interaction with the F1 oxidation intermediate leading to C−C coupling in the products.

DISCUSSION

The genus *Streptomyces* produces more than two-thirds of the naturally occurring microbial antibiotics and a large array of other secondary metabolites that are important to human and veterinary medicine (*1*). For the 18 P450s in the most studied *Steptomyces* spp., *S. coelicolor*, the biological function of only CYP158A2 has been clearly established. We have undertaken the study of the other CYP158A P450 in this organism after finding that the CYP158A2 knockout shows no phenotype and no secondary metabolite differences when compared to wild type strains. Perhaps this is because CYP158A1 can also dimerize flaviolin, creating redundancy of reaction in *S. coelicolor*. We find that each enzyme can produce 3,8-biflaviolin and 3,3-biflaviolin in *in vitro* assays, although at very different ratios, indicating that each enzyme has their own catalytic regiospecificity. It is believed that these products are involved in protecting *S. coelicolor* from UV irradiation, and different patterns of products do not influence the bacterial phenotype under laboratory conditions.

These two enzymes have relatively high sequence similarity (77%), and it is not particularly surprising that they both share the same substrate (flaviolin) in carrying out an unusual oxidative C−C coupling reaction *in vitro*, even though the product patterns are somewhat different. Phylogenetic analysis of CYP158A1 and CYP158A2 reveal their relatedness to other CYPs. For CYP158A2, the closest homologues are CYPs genetically organized in a similar fashion to this CYP gene, that is, the three-gene operon consisting of a type III polyketide synthase, the cytochrome P450, and a quinone-forming monooxygenase (*21*). This operon and hence the P450 is conserved in *S. avermitilis* (CYP158A3) (*22*), *S. griseus* (CYP158A4) (*21*), and *Saccharopolyspora erythraea* (CYP158B1) (*8*). The closest homologue for CYP158A2 is CYP158A1 also from *S. coelicolor*. CYP158A1 is not genetically organized in an operon within the *S. coelicolor* genome and is predicted to be transcribed as a single mRNA. Furthermore, no strong CYP158A1 orthologue is found in any other bacterial genome analyzed so far. BLAST analysis using CYP158A1 shows that its closest homologues are CYP158A2 and other CYPs linked to the three gene operons (CYP158A3, CYP158A4, and CYP158B1). Continued analysis of other related CYPs show that below the CYP158A families the closest matched CYPs for both CYP158A1/2 belong to the CYP107 family, a common and diverse CYP family within all gram positive bacteria. Given the abundance of CYP107s in streptomycetes (and actinomycetes in general), it is possible that a CYP107 P450 evolved in function to a CYP158A, which then became genetically linked into the three-gene operon because of the biosynthetic nature of the pathway where the gene products are involved in producing a final chemical molecule of benefit to the organism. Why *S. coelicolor* has two members of the CYP158A family and therefore redundancy in oxidative C−C coupling of flaviolin is open to conjecture at present.

Although it is not yet absolutely proven that C−C oxidative coupling of flaviolins occurs in the CYP158A2 active site, the presence of the two flaviolin molecules in the reduced dioxygen CYP158A2 complex strongly supports this notion (*23*). If not in the active site, this coupling is likely to occur somewhere in the substrate access channel and not randomly as a nonenzymatic reaction. At this stage, we would rule out the possibility that the radical or cationic intermediates could leave the P450 and react to form dimers in solution outside of the P450. These results might further indicate that the oxidative coupling reactions in CYP158A1/A2 do not require diradical intermediate for the reaction to occur. Therefore, the regiospecificity of the dimerization products is probably actively driven by the enzyme secondary structural environment and by key substrate contact residues. Probably the proximal flaviolin is activated to a radical or cationic intermediate (*7*) by a very similar reaction in both CYP158A1 and CYP158A2. Then, this intermediate must contact the distal neutral flaviolin molecule somewhere within in the active site or substrate access channel. The orientation of each flaviolin molecule or intermediate must be established by key contact residues. How flaviolin molecules form dimerization products and the control of the regiospecificity is not clear for either of the *S. coelicolor* CYP158A monooxygenases. Further studies regarding transfer of the first and second electrons during the catalytic steps will perhaps provide evidence for the structural changes and the movement of the bound substrate or intermediate necessary to form the C−C connection in both CYP158 monooxygenases.

In CYP158A2, it is much easier to imagine in a positional sense how this coupling reaction occurs rather than in CYP158A1 because in the former, the flaviolin molecules are 3−4 Å apart, whereas in the latter they are 9 Å apart. In CYP158A1, the position occupied by the highly conserved threonine in the I helix of a majority of P450s (>95%) is replaced by alanine, just as in CYP158A2. There are three

water molecules located just over the heme that form hydrogen bonds, WAT50 with 5-OH of the proximal flaviolin and WAT80, and WAT120 with the carbonyl C-4 of the proximal flaviolin. So the hydroxyl groups of the proximal flaviolin in CYP158A1 appear to serve the same structural role in positioning active site water molecules as observed in CYP158A2 (*23*). In addition, these catalytic water molecules are 4.3−6.5 Å from the heme iron, leaving sufficient room for dioxygen atom binding to the heme. The closest water WAT50 also might provide a hydrogen bond to the distal oxygen atom, thereby transferring protons and stabilizing oxygen binding.

Examination of many of the important bacterial antibiotics used in medicine such as vancomycin, the last line of treatment against methicillin resistant *Staphlococcus aureus* (MRSA), reveal conserved chemical syntheses. These include side chain cross-linking of separate chemical moieties involving oxidative C−C coupling reactions catalyzed by cytochromes P450. Although such C−C coupling reactions are widespread in secondary metabolite biosynthesis, the intricate molecular details regarding how a P450 catalyzes such a reaction is limited. Including CYP158A1, there are now four cytochromes P450 whose structural architectures have been revealed by X-ray crystallography and have been shown to catalyze oxidative C−C coupling: CYP158A1/A2 involved in cross-linking flaviolin monomers, CYP165B1 (OxyB) from the vancomycin producer *Amycolatopsis orientalis*, which has been proposed to catalyze the first oxidative coupling reaction in vancomycin biosynthesis (*24*), and CYP165C1 (OxyC), also from *A. orientalis*, which has been proposed to catalyze the last oxidative coupling reaction step in vancomycin biosynthesis (*25*). To date, it has not been possible to generate substrate-bound crystal forms of OxyB and OxyC for a detailed understanding of C−C catalysis at the molecular level (*26*, *27*). Hence, CYP158A1 and A2 provide the first substrate complexes allowing us to speculate on chemical reaction mechanisms for C−C coupling involving the activation of a proximal substrate, which then attacks the distal bound substrate to form the linked product with the expulsion of a water molecule. Analysis of the structures of OxyB and OxyC indicate that there is room for only one substrate molecule within the active sites, suggesting that while a similar mechanism for the generation of an intermediate molecule might exist as predicted for CYP158A1/A2, C−C bond formation is probably quite different from that in the CYP158A subfamily. It has been suggested that OxyB and OxyC interact with polyketide synthase so that the second P450 substrate molecule is produced by this bound synthase, and the activated intermediate can cross-link directly to it in the protein−protein complex (*26*, *27*).

CYP158A1 contains the same amino acid sequence identity (61%) to CYP158A2 in the active site as that for the complete protein sequence. Amino acid differences in the active site compared to CYP158A2 result in altered substrate selectivity and product specificity. The biochemical results reveal that CYP158A1/A2 can catalyze oxidative C−C coupling reactions for polymerizing flaviolin. Producing highly conjugated dimers to protect the soil bacterium from UV radiation, an important protective mechanism that is duplicated in *S. coelicolor* ensure survival. Work is now underway to prepare the double knockout (CYP158A1/A2) strain, which should provide additional information on the function of flaviolin secondary metabolites. However, it is not at all clear why these two protective enzymes show such different bindings of the distal flaviolin molecule. One thought is that the open flaviolin-bound form of CYP158A1 represents an earlier conformation in the catalytic cycle than the closed form in CYP158A2. There are a few other examples of ligand-bound P450s that do not exist in a closed form as determined by X-ray crystallography (*28*). We have wondered whether crystal contacts might be responsible for holding substrate-bound CYP158A1 in an open position. However, three structures of CYP158A1 (ligand-free structure, imidazole- and flaviolin-bound structures) in different unit cells but with the same open conformation and with imidazole bound in the F1 and F2 sites as well, suggest that the lattice effects are not the cause. Furthermore, careful analysis of crystal packing shows that this is not the case (Figure S7, Supporting Information). Perhaps interaction with the ferredoxin that transfers electrons to the CYP158A1 heme iron is necessary in addition to flaviolin binding for the closed conformation to occur. Even if this were true, the positioning of the two flaviolins would be different in CYP158A1 than in CYP158A2. Presumably, however, a closed flaviolin-bound conformation in CYP158A1 would position the flaviolins closer to one another compared to that in the open conformations presented here. Future studies involving mutations of key residues based on the X-ray structures of CYP158A1 and its complex with flaviolin will lead to a better understanding of the very surprising result that very different structures of flaviolin bound to CYP158A1 and CYP158A2 lead to the same products produced via oxidative C−C coupling.

ACKNOWLEDGMENT

We thank Dr. Bradley S. Moore (University of California, San Diego) for the generous gift of purified flaviolin, Dr. Claus Schneider (Vanderbilt University) for helpful discussions and advice, and personnel at the SER-CAT 22-BM beamline at APS Argonne National Laboratory for expert technical assistance.

SUPPORTING INFORMATION AVAILABLE

Figure S1 shows the typical type I P450 titration spectra and binding constant. Figures S2 and S3 provide the LC-MS/MS information to characterize the chemical structures of biflaviolin products as 3,8-biflaviolin (P1) and 3,3-biflaviolin (P2), respectively. Figures S4 to S7 show CYP158A1 structures in detail regarding the orientation of Arg92, the electron density of imidazole, and crystal packing. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Bentey, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbiowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill,

J., and Hopwood, D. A. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2), *Nature 417*, 141−147.
2. Lamb, D. C., Skaug, T., Song, H-L., Jackson, C. J., Podust, L. M., Waterman, M. R., Kell, D. B., Kelly, D. E., and Kelly, S. L. (2002) The cytochrome P450 complement (CYPome) of *Streptomyces coelicolor* A3(2), *J. Biol. Chem. 277*, 24000−24005.
3. Huynen, M., Snel, B., Lathe, III, W., Bork, P. (2000) Predicting protein function by genomic context: quantitative evaluation and qualitative inferences, *Genome Res., 10*, 1204−1210.
4. Eisenberg, D., Marcotte, E. M., Xenarios, I., and Yeates, T. O. (2000) Protein function in the post-genomic era, *Nature 405*, 823−826.
5. Kinoshita, K., and Nakamura., H. (2003) Protein informatics towards function identification, *Curr. Opin. Struct. Biol. 13*, 396−400.
6. Austin, M. A., Izumikawa, M., Bowman, M. E., Udwary, D. W., Ferrer, J.-L., Moore, B. S., and Noel, J. P. (2004) Crystal structure of a bacterial type III polyketide synthase and enzymatic control of reactive polyketide intermediates, *J. Biol. Chem., 279*, 45162−45174.
7. Zhao, B., Guengerich, F. P., Bellamine, A., Lamb, D. C., Izumikawa, M., Lei, L., Podust, L. M., Sundaramoorthy, M., Reddy, L. M., Kelly, S. L., Kalaitzis, J. A., Stec, D., Voehler, M., Falck, J. R., Moore, B. S., Shimada, T., Waterman, M. R. (2005) Binding of two flaviolin substrate molecules, oxidative coupling, and crystal structure of *Streptomyces coelicolor* A3(2) cytochrome P450 158A2, *J. Biol. Chem. 280*, 11599−11607.
8. Cortes, J., Velasco, J., Foster, G., Blackaby, A. P., Rudd, B. A., Wilkinson, B. (2002) Identification and cloning of a type III polyketide synthase required for diffusible pigment biosynthesis in *Saccharopolyspora erythraea, Mol. Microbiol. 44*, 1213−1224.
9. Funa, N., Funabashi, M., Yoshimura, E., and Horinouchi, S. (2005) Biosynthesis of hexahydroxyperylenequinone melanin via oxidative aryl coupling by cytochrome P-450 in *Streptomyces griseus*, *J. Bacteriol. 187*, 8149−8155.
10. Izumikawa, M., Shipley, P. R., Hopke, J. N., O'Hare, T., Xiang, L., Noel, J. P., and Moore, B. S. (2003) Expression and characterization of the type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene synthase from *Streptomyces coelicolor* A3(2), *J. Ind. Microbiol. Biotechnol. 30*, 510−515.
11. Pang, Y. P., and Kozikowski, A. P. (1991) Synthesis of methoxy and hydroxy analogues of 1,2,3,4,4a,9a-hexahydro-4a-fluorenamine: rigid phencyclidine analogues as probes of PCP receptor topography, *J. Org. Chem. 56*, 4499−4508.
12. Owtinowski, Z., and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode, in *Macromolecular Crystallography* (Carter, C. W., and Swwet, R. M., Eds.) Part A, Vol. 276, pp 307−326, Academic Press, New York.
13. Powell, H. R. (1999) The Rossmann Fourier autoindexing algorithm in MOSFLM, *Acta Crystallogr., Sect. D 55*, 1690−1695.
14. Storoni, L. C., McCoy, A. J., and Read, R. J. (2004) Likelihood-enhanced fast rotation functions, *Acta Crystallogr., Sect. D 60*, 432−438.
15. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr., Sect. A 47*, 110−119.
16. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., and Pannu, N. S. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D 54*, 905−921.
17. Evans, S. V. (1993) SETOR: hardware-lighted three-dimensional solid model representations of macromolecules, *J. Mol. Graphics 11*, 134−138.
18. Ravichandran, K. G., Boddupalli, S. S., Hasemann, C. A., Peterson, J. A., and Deisenhofer, J. (1993) Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's, *Science 261*, 731−736.
19. Scott, E. E., He, Y. A., Wester, M. R., White, M. A., Chin, C. C., Halpert, J. R., Johnson, E. F., and Stout, C. D. (2003) An open conformation of mammalian cytochrome P450 2B4 at 1.6-A resolution, *Proc. Natl. Acad. Sci. U.S.A. 100*, 13121−13122.
20. Kleywegt, G. J., and Jones, T. A. (1994) Detection, delineation, measurement and display of cavities in macromolecular structures, *Acta Crystallogr., Sect. D 50*, 178−185.
21. Funa, N., Funabashi, M., Yoshimura, E., and Horinouchi, S. (2005) A novel quinone-forming monooxygenase family involved in modification of aromatic polyketides, *J. Biol. Chem. 280*, 14514−14523.
22. Lamb, D. C., Ikeda, H., Nelson, D. R., Ishikawa, J., Skaug. T., Jackson, C., Omura, S., Waterman, M. R., and Kelly, S. L. (2003) Cytochrome p450 complement (CYPome) of the avermectin-producer *Streptomyces avermitilis* and comparison to that of *Streptomyces coelicolor* A3(2), *Biochem. Biophys. Res. Commun. 307*, 609−610.
23. Zhao, B., Guengerich, F. P., Voehler, M., Waterman, M. R. (2005) Role of active site water molecules and substrate hydroxyl groups in oxygen activation by cytochrome P450 158A2: a new mechanism of proton transfer, *J. Biol. Chem. 280*, 42188−42197.
24. Bischoff, D., Pelzer, S., Holtzel, A., Nicholson, G. J., Stockert, S., Wohlleben, W., Jung, G., and Sussmuth, R. D. (2001) The biosynthesis of vancomycin-type glycopeptide antibiotics: new insights into the cyclization steps, *Angew. Chem., Int. Ed. 40*, 1693−1696.
25. Bischoff, D., Pelzer, S., Bister, B., Nicholson, G. J., Stockert, S., Schirle, M., Wohlleben, W., Jung, G., and Sussmuth, R. D. (2001) The biosynthesis of vancomycin-type glycopeptide antibiotics: the order of the cyclization steps, *Angew. Chem., Int. Ed. 40*, 4688−4691.
26. Pylypenko, O., Vitali, F., Zerbe, K., Robinson, J. A., and Schlichting, I. (2003) Crystal structure of OxyC, a cytochrome P450 implicated in an oxidative C−C coupling reaction during vancomycin biosynthesis, *J. Biol. Chem. 278*, 46727−46733.
27. Zerbe, K., Pylypenko, O., Vitali, F., Zhang, W., Rouset, S., Heck, M., Vrijbloed, J. W., Bischoff, D., Bister, B., Sussmuth, R. D., Pelzer, S., Wohlleben, W., Robinson, J. A., and Schlichting, I. (2002) Crystal structure of OxyB, a cytochrome P450 implicated in an oxidative phenol coupling reaction during vancomycin biosynthesis, *J. Biol. Chem. 277*, 47476−47485.
28. Zhao, Y., White, M. A., Muralidhara, B. K., Sun, L., Halpert, J. R., and Stout, C. D. (2006) Structure of microsomal cytochrome P450 2B4 complexed with the antifungal drug bifonazole: insight into P450 conformational plasticity and membrane interaction, *J. Biol. Chem. 281*, 5973−5981.

BI7006959