Binding of Two Flaviolin Substrate Molecules, Oxidative Coupling, and Crystal Structure of *Streptomyces coelicolor A3(2) Cytochrome P450 158A2*

Bin Zhao‡*, F. Peter Guengerich‡, Aouatef Bellamine‡, David C. Lamb¶, Miho Izumikawa║, Li Lei‡, Larissa M. Podust‡, Munirathinam Sundaramoorthy‡§, John A. Kalaitzis‖, L. Manmohan Reddy£, Steven L. Kelly§, Bradley S. Moore‖, Donald Stec¥, Markus Voehler¥, John R. Falck£, Tsutomu Shimada‡, and Michael R. Waterman‡

*From the Departments of ‡Biochemistry and §Medicine and the Center for Structural Biology, and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, TN 37232-0146, ‖Swansea Clinical School, University of Wales Swansea, UK, ‖College of Pharmacy, University of Arizona, Tucson, AZ 85721-0207, £Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75390, and ¥Department of Chemistry, Vanderbilt University, Nashville, TN 37235

*To whom correspondence should be addressed:
Department of Biochemistry, Vanderbilt University School of Medicine
864 Robinson Research Building
23rd and Pierce Avenues
Nashville, Tennessee 37232-0146
Tel: 615-343-4644
Fax: 615-343-0704
SUMMARY

Cytochrome P450 158A2 (CYP158A2) is encoded within a three-gene operon (*sco1206–sco1208*) in the prototypic soil bacterium *Streptomyces coelicolor* A3(2). This operon is widely conserved among streptomycetes. CYP158A2 has been suggested to produce polymers of flaviolin, a pigment that may protect microbes from UV radiation, in combination with the adjacent *rppA* gene, which encodes the type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene synthase. Following cloning, expression, and purification of this cytochrome P450 we have shown that it can produce dimer and trimer products from the substrate flaviolin, and the structures of two of the dimeric products were established using mass spectrometry and multiple NMR methods. Comparison of the X-ray structures of ligand-free (1.75 Å) and flaviolin-bound (1.62 Å) forms of CYP158A2 demonstrates a major conformational change upon ligand binding which closes the entry into the active site, in part due to repositioning of the F and G helices. Particularly interesting is the presence of two molecules of flaviolin in the closed active site. The flaviolin molecules form a quasi-planar three-molecule stack including the heme of CYP158A2, suggesting that oxidative C-C coupling of these phenolic molecules leads to the production of flaviolin dimers.
INTRODUCTION

Cytochrome P450 (P450, CYP)\(^1\) monooxygenases constitute a complex superfamily of proteins found in all biological kingdoms from bacteria to humans and participate in the biosynthesis of physiologically important compounds as well as in detoxification of a wide variety of foreign compounds from the environment (1-3). Largely as a result of genome sequencing projects, more than 4000 genes within the P450 superfamily have been identified (http://drnelson.utmem.edu/CytochromeP450.html). While there has been considerable progress in the expression and characterization of recombinant P450s within the last decade, the endogenous substrates of most remain unknown. For example, there are 18 CYP genes in the model actinomycete *Streptomyces coelicolor* A3(2) (4, 5), whose endogenous functions and biological roles are largely unclear. *S. coelicolor* produces more than 20 secondary metabolites including antibiotics, siderophores, lipids, and pigments (5). Many of these compounds are oxidized and contain hydroxyl groups, which could arise from the function of P450s associated with the biosynthetic pathways. Most P450s catalyze monooxygenase reactions, in which one atom of molecular oxygen is used to hydroxylate the substrate and the other to produce water (6). P450s additionally catalyze a myriad of other reactions, including epoxidations, oxidative rearrangements, and oxidative coupling reactions (7-11).

Chloroeremomycin and balhimycin, which are members of the vancomycin family of glycopeptide antibiotics, possess a rigidified peptidyl backbone in which the aromatic residues are linked via P450-dependent biphenyl and biphenyl ether connections (12, 13). Knockout experiments with the P450 genes *oxyA*, *oxyB*, and *oxyC* in *Amycolatopsis orientalis* demonstrated that P450s catalyze the cross-linking steps in balhimycin biosynthesis (14, 15). Thus, OxyA and OxyB catalyze C-O-C coupling reactions and OxyC is involved in C-C...
coupling. Schlichting and her associates (16, 17) have determined the high-resolution X-ray structures of OxyB and OxyC, which provide templates with which to characterize these oxidative coupling reactions.

CYP158A2 from *S. coelicolor* A3(2) has also been implicated in phenol oxidative coupling (18), generating red-brown pigments in actinomycetes. The gene encoding this P450 (*sco1207*) is located in the three-gene operon *sco1206-sco1208* that also contains the type III polyketide synthase (PKS) 1,3,6,8-tetrahydroxynaphthalene synthase (THNS) (19, 20) and a gene of unknown function designated open reading frame 3. This three-gene operon is highly conserved and present in many *Streptomyces* species (18). Homodimeric type III PKSs are involved in biosynthetic pathways in bacteria for the assembly of small aromatic metabolites (19, 21, 22). Bacterial aromatic polyketides represent a large group of biologically active natural products, whose potency is often dictated by tailoring enzymes such as oxygenases, methyltransferases and glycosyltransferases. P450s are often associated with polyketide biosynthetic gene clusters where they catalyze late-stage stereo- and regiospecific oxidations. In the *S. coelicolor* operon containing CYP158A2, the 5′-gene *rppA* (*sco1206*), whose stop codon overlaps the start codon of *CYP158A2* (*sco1207*), encodes THNS. This type III PKS catalyzes the sequential conversion of five molecules of malonyl-CoA to 1,3,6,8-tetrahydroxynaphthalene (THN), which then undergoes spontaneous oxidation to flaviolin (19, 20).

In the present study, we focused on the catalytic activity of CYP158A2 and demonstrated that CYP158A2 catalyzes the oxidative coupling of two or three THN-based flaviolin molecules, involving C-C coupling. To explore the molecular basis of this catalytic reaction, we have determined the crystal structures of CYP158A2 in the absence and presence of the substrate flaviolin, which provides the molecular basis for P450-catalyzed oxidative coupling reactions.
EXPERIMENTAL PROCEDURES

Expression and Purification of CYP158A2—The gene encoding *S. coelicolor* A3(2) CYP158A2, engineered with four histidine codons at the C-terminus, was subcloned into the *Escherichia coli* expression vector pET17b (Novagen, Madison, WI) using the *Nde*I and *Hind*III sites (5). Recombinant proteins were expressed in *E. coli* HMS174 (DE3). Transformed *E. coli* was grown at 37 °C in 3 liters of Terrific Broth containing 100 µg/ml ampicillin until the optical density at 600 nm reached 0.8. After induction with 0.5 mM isopropyl-β-D-thiogalactopyranoside and addition of δ-aminolevulinic acid to a final concentration of 1 mM for heme synthesis, growth was allowed to continue for an additional 24 h at 24 °C and the cells were harvested by centrifugation for 10 min at 4,000 rpm. The pellet was resuspended in 100 ml of lysis buffer (50 mM Tris-HCl, pH 7.5, containing 500 mM NaCl, 0.5 mM EDTA, and 10% (v/v) glycerol). The soluble CYP158A2 were purified by Ni²⁺-NTA (Qiagen, Chatsworth, CA) and S-Sepharose/Q-Sepharose (Amersham Pharmacia Biotech, Piscataway, NJ) chromatography. The Ni²⁺-NTA column was washed with 10 bed volumes of lysis buffer and 10 bed volumes of lysis buffer with 3 mM imidazole (pH 7.5). The protein was eluted with elution buffer (20 mM Tris-HCl, pH 7.5, containing 10% (v/v) glycerol and 80 mM imidazole). The eluted CYP158A2 was passed over a S-Sepharose/Q-Sepharose column, washed with 5 bed volumes of buffer (20 mM Tris-HCl, pH 7.5, containing 0.5 mM EDTA and 10% (v/v) glycerol) and then eluted with a linear gradient of 0-500 mM NaCl in 20 mM Tris-HCl, pH 7.5. Individual fractions were analyzed by SDS-PAGE, pooled, and concentrated using a Microcon-30 filtration device (Millipore, New York, NY). *E. coli* flavodoxin and flavodoxin reductase were expressed in *E. coli* strain HMS174 (DE3) and purified by DEAE column chromatography as described previously (23).
Spectral Substrate Binding Assay—Absorbance spectra were recorded using a double beam Shimadzu UV-2401PC spectrophotometer. The interaction of flaviolin with CYP158A2 was examined by perturbation of the heme Soret spectrum. CYP158A2 (2.5 μM) in 20 mM Tris-HCl (2.0 ml, pH 7.5) was divided between two tandem cuvettes. After thermal equilibration at 25 °C a baseline was established between 350 nm 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, pH 7.5) were added to the sample cuvette to give a final ligand concentration in the range of 0-100 μM. An equal volume of 5% (v/v) methanol in 20 mM Tris-HCl (pH 7.5) was added to the reference cuvette, and the difference spectrum recorded after each titration. The Hill coefficient was obtained from plot of \( \log_{10}(\Delta A_{387-418}/[\Delta A_{\text{max}}-\Delta A_{387-418}]) \) vs \( \log_{10}[\text{flaviolin}] \). \( K_d \) values were estimated by fitting plots of \( \Delta A_{387-418} \) vs [flaviolin].

Enzyme Activity Assay—CYP158A2 (1 nmol), flavodoxin (20 nmol), and flavodoxin reductase (10 nmol) were reconstituted in 400 μl of 20 mM Tris-HCl buffer (pH 7.5) containing 10% (v/v) glycerol and flaviolin (0.26 μmol). Flaviolin was obtained from E. coli BL21 (DE3)pLysS/pHIS8-THNS as previously described (19) or organic synthesis from 3,5-dimethoxybenzoic acid following established protocols (24, 25). Following incubation of this mixture for 5 min on ice, the reconstituted enzyme solution was placed in a shaking bath at 37 °C. The reaction was started 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 quenched with 4 μl of concentrated HCl and extracted three times with 400 μl of ethyl acetate. The extracts were dried under a stream of N₂, the residues were dissolved in 50 μl of methanol, and 20 μl of extract was injected onto an HPLC/MS system for analysis. UV detection was at 254 nm. Negative control incubations (minus CYP158A2) were carried out as above.
**Large-scale Incubations**—For the structural identification of those products converted from flaviolin, multiple large-scale incubations of flaviolin with CYP158A2 were carried out. The reaction mixture consisted of CYP158A2 (3 nmol of CYP/ml), flavodoxin (35 nmol/ml), and flavodoxin reductase (15 nmol/ml), flaviolin (42 µmol), 10 mM glucose 6-phosphate, 1 mM NADP⁺, and 2 U/ml of glucose-6-phosphate dehydrogenase in a final volume of 72 ml of 20 mM Tris-HCl buffer (pH 7.5) containing 10% (v/v) glycerol. The incubation proceeded for 2 h, and then the reaction mixture was acidified by concentrated HCl and extracted three times into an equal volume of ethyl acetate. The solvent was removed from the combined organic extracts in *vacuo*. The residue was dissolved in acetonitrile, and purified by HPLC on a 5 µm HyPURITY C18 column (10 mm × 150 mm) eluted with a linear solvent gradient from 65% of 0.5% formic acid/0.01% trifluoroacetic acid in water (solvent A) to 60% of 0.5% formic acid/0.01% trifluoroacetic acid in acetonitrile (solvent B) over 25 min at a flow rate of 1.5 ml/min, monitored with UV detection at 254 nm. The desired fractions were collected and dried under a stream of N₂.

**Identification of Flaviolin Metabolites**—For HPLC/MS analysis, electrospray ionization (ESI) mass spectra were obtained using a Finnigan TSQ Quantum mass spectrometer in the positive ion mode with electrospray voltage 4.0 kV, 300°C capillary temperature, N₂ sheath gas (43 p.s.i.), and N₂ auxiliary gas (25 p.s.i.). The ESI interface and mass spectrometer parameters were optimized to obtain maximum sensitivity. HPLC separations were achieved using a YMC ODS-AQ 2.0 mm × 250 mm reversed-phase octadecylsilane column with a linear solvent gradient from 90% of 0.5% formic acid/0.01% trifluoroacetic acid in water (solvent A) to 100% of 0.5% formic acid/0.01% trifluoroacetic acid in acetonitrile (solvent B) over 20 min, followed by isocratic elution with solvent B for 10 min at a flow rate of 200 µl/min. NMR
spectra were recorded on a Bruker DPX 500 AVANCE spectrometer operating at a proton frequency of 500.13 MHz and carbon frequency of 125.76 MHz. Samples were prepared in either DMSO-$d_6$ or acetone-$d_6$. $^1$H and $^{13}$C chemical shifts are reported as $\delta$ values with reference to DMSO-$d_6$ at 2.49 and 39.5 ppm, respectively, or acetone-$d_6$ at 2.04 and 206 ppm, respectively. All measurements were run at 25 °C using standard pulse sequences. Two-dimensional (2D) techniques such as $^1$H-$^1$H correlated spectroscopy (COSY), and $^1$H-$^{13}$C heteronuclear multiple quantum coherence (HMQC), and $^1$H-$^{13}$C heteronuclear multiple bond coherence (HMBC) were also used for structure elucidations.

**Crystallization and Structure Determination**—Crystals of CYP158A2 were obtained using the hanging-drop vapor diffusion method, in which 2 $\mu$l of a 16 mg/ml protein solution was mixed with an equal volume of 0.1 M 2-[N-morpholino]ethanesulfonic acid (pH 6.0), 5 mM spermine-HCl, and 20-25% (w/v) of PEG 2000 monomethyl ether. At 20 °C, plate crystals appeared within three weeks. The solution used as cryoprotectant for flash cooling these crystals at liquid nitrogen temperatures was the original reservoir solution, supplemented with 10% (v/v) glycerol. The crystals belong to the monoclinic space group P2$_1$ with unit cell parameters: $a = 59.55$ Å, $b = 79.2$ Å, $c = 87.43$ Å, $\beta = 92.26^\circ$. Full diffraction data were collected at 100 K at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID 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 (26). The substrate-free structure was solved by molecular replacement using the program CNS1.1 (27) and the inhibitor-bound CYP158A2 structure as a search model. The initial model was built in O (28) and refinement was performed using CNS1.1 (27). There were two molecules of CYP158A2 in the asymmetric unit, resulting in a solvent content of 47.5%.
Co-crystals of CYP158A2 and flaviolin were grown using a well solution containing 0.1 M bis-Tris (pH 6.5) and 25% (w/v) of PEG 3350 supplemented with 1 mM flaviolin. At 20 °C, rod-like crystals were obtained within 1-4 days. The crystals belong to the orthorhombic space group P2₁2₁2₁ with unit cell parameters: a = 54.8 Å, b = 70.74 Å, c = 102.57 Å. The flaviolin complex structure was also solved by molecular replacement using the program CNS1.1 (27) and the inhibitor-bound CYP158A2 structure as a search model. Refinement was performed using CNS1.1 (27) and CNS parameter and topology files were generated by PRODRG (29). Final refinement statistics for both crystal structures are given in Table I. The coordinates and associated structure factors have been deposited with the Protein Data Bank (accession code: substrate-free: 1SE6; substrate-bound: 1T93). The inhibitor-bound CYP158A2 structure used as search model for both structures presented here was produced by crystallization in the presence of 4-phenylimidazole that binds to the heme iron. This structure is refined to 1.5 Å (accession code: 1S1F) and since it is not directly relevant to the CYP158A2 function of oxidative coupling of flaviolin is not presented herein, but will be published elsewhere (B. Zhao, unpublished data).
RESULTS

An Endogenous Function and Probable Biological Role of CYP158A2—

Titration shows that CYP158A2 binds flaviolin ($K_d = 7.3$ µM), producing a typical type I P450 binding spectrum (Fig. 1) and demonstrating that this potential substrate can enter the active site. The Hill coefficient was $1.6 \pm 0.2$, suggesting the possible presence of more than one flaviolin molecule bound to the protein. Analysis of the reaction of CYP158A2 by electrospray mass spectrometry indicated four different products from flaviolin (Fig. 2), three isomers with mass 410 (MH$^+$ 411, $P_1 t_R$ 15.6 min; $P_2 t_R$ 16.0 min; $P_3 t_R$ 17.2 min) and one product with mass 614 (MH$^+$ 615, $P_4 t_R$ 17.4 min). The molecular masses of 410 and 614 (Fig. 2) are those of the dimer and trimer of flaviolin, respectively. The turnover number for flaviolin oxidation was estimated to be $1.4 \text{ min}^{-1}$ based on the rate of substrate consumption under the in vitro condition used here.

The structures of the flaviolin dimers $P_1$ and $P_2$ were obtained from multiple NMR analyses, which included one-dimensional (1D) $^1$H and two-dimensional (2D) $^1$H-$^1$H COSY and $^1$H-$^1$H HMQC, and $^1$H-$^1$C HMBC experiments. $^1$H and $^{13}$C NMR signals of flaviolin were completely assigned in DMSO-$d_6$ and used as a reference in the characterization of $P_1$ and $P_2$. The singlet at $\delta 6.03$ was assigned to H-3 and the two 2.4 Hz doublets at $\delta 6.94$ and 6.53 were assigned to H-8 and H-6, respectively. The proton NMR spectrum$^2$ of $P_1$ showed two doublets and two singlets at $\delta 7.2$ (d, $J = 2.4$ Hz, 1H, for H-8), 6.81 (s, 1H, for H6’), 6.69 (d, $J = 2.4$ Hz, 1H, for H6), and 6.16 (s, 1H, for H-3’) accounting for four carbon-bound hydrogen atoms, thereby suggesting that $P_1$ is a non-symmetrical flaviolin C–C dimer (Fig. 3A, 3C). Analysis of the $^1$H-$^{13}$C HMQC and $^1$H-$^{13}$C HMBC spectra ($H-3' \rightarrow C-1'$, $C-2'$, $C-10'$; $H-6 \rightarrow C-5$, $C-7$, $C-8$, $C-10$; $H-6' \rightarrow C-5'$, $C-8'$, $C-10'$; $H-8 \rightarrow C-1$, $C-6$, $C-10$) established $P_1$ as 3,8'-biflaviolin (Fig. 3E). On the other hand, $P_2$ was clearly a symmetric structure as judged by its $^1$H NMR spectra (Fig.
3B, 3D), which showed two doublets at δ 6.99 (d, J=2.0 Hz, 2H, for H-8/8’) and 6.55 (d, J=2.0 Hz, 2H, for H-6/6’). This structure was assigned as 3,3’-biflaviolin (Fig. 3F) (30), which was further verified through analysis of the $^1$H-$^1$H COSY, $^1$H-$^{13}$C HMQC, and $^1$H-$^{13}$C HMBC spectra (see Supplemental Data).

**CYP158A2/Substrate-Free Crystal Structure**—The global features of the substrate-free structure (Fig. 4A), determined by X-ray crystallography, show a more “open” conformation with respect to many other P450 structures (32, 33). This “open state” conformation should allow the substrate access to the active site of the enzyme (34). In the case of the substrate-free structure (1.75 Å), the F/G helices are rotated out of the active site so that the cleft reflects a gap between the F helix and β-sheet 4. Additionally, owing to the absence of substrate, the substrate access pocket is filled with water molecules. The water molecule coordinating the heme iron at the sixth position is present with an Fe-H$_2$O bond distance of 2.38 Å.

**CYP158A2/Flaviolin Complex Crystal Structure**—Examination of the general protein fold for the structure of the substrate-complex indicates that the substrate access channel closes with respect to the substrate-free structure (Fig. 4B). The root mean square difference of superposition for the $C_a$ atoms upon binding of flaviolin to CYP158A2 as compared to the substrate-free structure is 0.93 Å (Fig. 5), which implies significant conformational change. Overall, the largest conformational changes occur in the regions R14-P22, L44-W50, G70-P98 (SRS$^1$-1) (35), P99-S107, G162-D213 (SRS-2 & SRS-3), E231-T247 (SRS-4), H287-R295 (SRS-5), and F388-G396 (SRS-6), which represent approximately 35% of the total amino acid sequence of CYP158A2. Among these regions, R14-P22, L44-W50, G70-P98, G162-D213, and
F388-G396 are located on the distal surface of the protein. The most significant change occurs in the F/G region (G162-D213), and the largest movement of F/G loop residues is at S182 and H183, where the Cα atoms are displaced 15-16 Å when substrate is bound (Fig. 5). In the flaviolin complex, the F helix slides toward the loop between F/G helices and rotates into the active site, and the adjacent G helix is disrupted into two helices, an alpha helix (G) and a 3_10 helix (G’), accompanied by the complementary movement of this G helix region. The 3_10 G’ helix is not present in the substrate-free structure. The residues in the BC loop dip into the active site to contact the substrate. In addition, there are 1-2 Å changes in residues R14-P22, the loop between strands β5-1 and β1-3 as well as in the β-4 loop with respect to the substrate-free structure. The conformational changes, which are largely confined to the distal surface loops of the protein, contribute to the closure of the substrate access channel and appear to be responsible for substrate orientation. In addition, one significant structural change of the C helix on the proximal surface is observed. This region around the C helix appears to be moved toward the proximal surface, which in several P450s is reported to have an important role in docking the redox partner and in electron transfer (36). The other two regions undergoing significant conformational changes, E231-T247 and H287-R295, are in the structural core of CYP158A2.

An intriguing result from this structure is that there are two molecules of flaviolin present in the active site, according to electron density, exhibiting planar shaped density for the π-electron system of the naphthaquinone ring (Fig. 6). The positions of the three hydroxyls of flaviolin are clearly recognizable due to the high resolution of this complex (1.62 Å). The average crystallographic temperature factors of these two flaviolin molecules are 16.2 Å² and 16.5 Å², respectively, indicating complete occupancy at both positions (37). The π-π electron system of flaviolin results in the planar surface of the aromatic ring of the first flaviolin being
almost parallel with the heme (Fig. 6). The five-coordinated iron is pulled slightly toward the
cysteine, away from the plane of the heme (toward the proximal surface). The location of the
first flaviolin places the carbon and oxygen atoms (C6, 5-OH, 7-OH) closest to the heme iron
between 4.6-5.3 Å, which falls in the normal range of substrate positions for activation to occur
in P450s (38, 39). The C6’, 5’-OH, 7’-OH, and C8’ in the second flaviolin molecule are
positioned over the sixth coordination position between 6.4-8.9 Å, significantly farther from the
heme iron.

The binding of the flaviolin leads to displacement of the sixth coordinated water
molecule away from the heme iron and the closing of the substrate access channel. The two
flaviolin molecules are surrounded by nine ordered water molecules (Fig. 7). These ordered
water molecules directly hydrogen bond with the two flaviolin molecules, for example, 2-OH
with WAT501 and WAT506, carbonyl C4 with WAT507, 5-OH with WAT505, 7-OH with
WAT508 and WAT600, 2’-OH with WAT555, carbonyl C4’ with WAT537, 5’-OH with WAT508,
and 7’-OH with WAT600 and WAT640. Interestingly, WAT508 and WAT600 are involved in
interactions with both flaviolin molecules, which must contribute to the stability of the geometry
of substrate orientation. The nine ordered water molecules are also able to form an extensive
network of hydrogen bonds with adjacent amino acid residues and seven additional ordered
water molecules in the active site which do not contact the substrate molecules. Besides ordered
waters, the side chains of R71, and R288 and the amide of L293 also form hydrogen bonds
directly with the two flaviolin molecules (Fig. 7). The side chains of R71 and R288 are oriented
toward the active site pocket of the enzyme, and the guanidium group of R288 forms hydrogen
bonds with the carbonyl of C4 in the first flaviolin but also forms hydrogen bonds with the
carbonyl of C1’ and with the 2’-OH in the second flaviolin molecule, indicating that R288 is a
key residue for CYP158A2 substrate binding. The guanidium group of R71 forms hydrogen bonds with the carbonyl of C1 and the propionate of the heme. Hydrogen bonding interaction is also observed between 2-OH in the first flavolin molecule and the backbone amide nitrogen atom of L293. These hydrogen bonding interactions provide a polar environment to accommodate the polar substrate in the active site, also seen in CYP2C5-diclofenac complex (40). Conversely, polar groups of flavolin molecules also play an important role in ordering the water molecules in the active site pocket of the flavolin-complex.
DISCUSSION

S. coelicolor A3(2) is a model actinomycete that produces more than 20 different secondary metabolites, including pigments, antibiotics, siderophores, hopanoids and other lipids (4). A majority of these compounds are oxidized, in some cases by the function of P450 monooxygenases. The functions of the 18 S. coelicolor CYPs, however, are not at all clear from genome analysis. CYP105D5 is predicted to metabolize toxic molecules from the environment based on known activities of CYP105D1 from S. griseus (41). CYP154C1 can participate in macrolide antibiotic biosynthesis in vitro (42), but such antibiotics are not produced by S. coelicolor A3(2). Based on studies in other Streptomyces strains on the biosynthesis of THN-based red-brown pigments (18, 21), we tested flaviolin as a substrate for CYP158A2 and established the first endogenous function of a S. coelicolor A3(2) CYP in secondary metabolite biosynthesis. Not only have we shown a function of CYP158A2 to dimerize and even trimerize flaviolin in vitro, we have found that two flaviolin molecules occupy the active site in the substrate-bound structure.

In the CYP158A2 structure with two flaviolin molecules, the positions of amino acid side chains and ordered waters in the active site are optimized. To accommodate two flaviolin molecules, the amide nitrogen of A245 is rotated toward the heme pocket providing hydrogen bonds to WAT503 that is also hydrogen bonded to the carbonyl oxygen of I241 and WAT640 which hydrogen bonds with 7’-OH and WAT505, which extends a hydrogen bond to 5-OH. WAT600 hydrogen bonds with 7’-OH and the carbonyl oxygen of L238. This hydrogen bonding network leads to bending of the N-terminal portion of the I-helix toward the active site, which seems to stabilize water molecules. The flaviolin-bound structure exhibits conformational
changes with respect to the substrate-free form, which reflect both the hydration and the size of the substrate.

Hydrogen bonding interactions between 2-OH (2'-OH) and carbonyls in two flavilolin molecules with R71, L293 and R288, and the 5-OH (5'-OH) and 7-OH (7'-OH) in flavilolin molecules appear to play an important role in formation of the water cluster seen in the substrate binding cavity of the flavilolin-complex (Fig. 7). Based on these observations, the critical role of the 2-OH (2'-OH) might be for anchoring to the enzyme and the 5-OH (5'-OH) and/or 7-OH (7'-OH) may be responsible for proton delivery, as has been suggested for P450<sub>eryF</sub> (43), or stabilization of water molecules needed for cleavage of dioxygen. Absence of one or both of these two hydroxyl groups may prevent metabolism or decrease the turnover rate, which suggests enzyme selectivity for the specific substrate. The role of ordered networks of water molecules in the binding of hydroxyl and carbonyl moieties of flavilolin suggests the importance of such networks in the binding of polar substrates and in stabilizing the significant conformational changes after they have occurred.

In addition, the crystal structure indicates that CYP158A2 may also have the capacity to bind a larger substrate, because CYP158A2 has a relatively large active site (495 Å<sup>3</sup>) and there is a large movement of the F/G region. One of the four products (Fig. 2) is a flavilolin trimer. This result might imply that flavilolin dimer and single flavilolin can bind to the active site pocket at the same time, which presumably is the case for trimer formation.

The general P450 catalytic cycle (44) is a multistep pathway and can be described as follows: a) substrate binding, b) first electron transfer, c) dioxygen binding, d) second electron transfer, e) protonation and splitting of the oxygen-oxygen bond with generation of a perferryl iron complex (FeO<sup>3+</sup>), f) reaction of FeO<sup>3+</sup> with substrate to form product, and g) product
dissociation. The chemical mechanism of formation of the flaviolin dimers has not been established at this point. Examples of coupling by P450s have been reported (45), and the coupling of resonance-stabilized diradicals has been proposed as a mechanism. The flaviolin atom closest to the P450 iron atom in the crystal structure is C6 (Fig. 6), but the products we characterized are 3,3′- and 3,8′-dimers (Fig. 3).

A mechanism can be proposed that involves initial oxidation of the proximal flaviolin molecule by hydrogen atom abstraction near the iron in the crystal structure (Fig. 8). The initial abstraction is postulated to occur from the (phenolic) OH, which is located 5.3 Å from the iron in the structure. The C6 atom is located 4.6 Å away, slightly closer, but the localization of a radical or a cationic charge there (i.e., with an Fe-O-C6 complex) leads to localization of the positive charge or radical density at C7, C6, C10, and C2 but not a C3 or C8. The mechanism shown has radical localization at C3 and C8 (and also C6 and C1), which may also explain that the multiple products were formed during the reaction.

An alternate mechanism can be proposed with an initial ipso attack at the C5 atom (46), which yields a covalent intermediate having a positive charge at C3 (Fig. 9). In this mechanism, this electronegative center would react with the double bond of the adjacent flaviolin molecule.

A requirement of the diradical mechanism proposed in Fig. 8 is that the proximal flaviolin radical either exchanges position with the distal flaviolin molecule in the active site, so that both flaviolins form radicals, or undergo an abstraction process within the active site. Alternatively, radicals could leave the P450 and react to form dimers in solution outside of the P450. The products would reflect the stability and reactivity of individual radicals rather than steric influences of the P450 active site. At this point we cannot exclude any of the possible mechanisms discussed here.
In conclusion, we report the first endogenous substrate firmly identified in the eighteen P450s from *S. coelicolor*. CYP158A2 can polymerize flaviolin to red-brown pigments, which may afford physical protection to the organism, possibly against the deleterious effects of UV radiation to which this soil bacterium is exposed (18). In combination with biochemical data, the structures reported here provide valuable insights into the general mechanism of the oxidative coupling reactions of phenols in P450s.
ACKNOWLEDGEMENT

We thank Dr Nobutaka Funa for generous gift of highly purified flaviolin; Drs. Norio Kagawa, and Galina I. Lepesheva and Prof. David Hopwood for helpful discussions and advice; personnel at the SER-CAT 22-ID beamline at APS Argonne National Laboratory for expert technical assistance; Drs. David Hachey, M. Lisa Manier, and Dawn Overstreet for assistance with mass spectrometry; Drs. Jorge H. Capdevila and Zhongliu Wu for advice on HPLC analysis; Dr. Joel Harp for x-ray technical assistance; and Dr. Youngchang Kim for substrate-free form data collection. Support was provided by National Institutes of Health Grants GM37942 (M. R. W.), GM69970 (M.R.W., F. P. G., D.C.L. and S.L.K.), ES00267 (F. P. G. and M.R.W.), DK62524 (M. S.), AI52443 (B. S. M.), and GM31278 (J.R.F.) and the Wellcome Trust (D.C.L.) and Biological and Biotechnological Sciences Research Council (S.L.K.).
REFERENCES


FOOTNOTES

1 The abbreviations used are: P450 or CYP, cytochrome P450 monooxygenase; THN, 1,3,6,8-tetrahydroxynaphthalene; THNS, THN synthase; F_o, observed structure factor; F_c, calculated structure factor; SRS, predicted substrate recognition sequences originally suggested in P450 sequences by Gotoh. COSY, correlated spectroscopy, and HMQC, heteronuclear multiple quantum coherence, and HMBC, heteronuclear multiple bond coherence.

2 Because only very small amounts of P1 and P2 were available, a few additional peaks coming from impurities and solvents also appeared in the NMR spectra. However, these additional peaks could be rationally excluded by analyzing their two dimensional NMR spectra.
FIGURE LEGENDS

FIG. 1. **Flaviolin binding by CYP158A2.** A, Type I binding spectra resulting from CYP158A2 (2.5 µM) titration (details presented in Experimental Procedures) with increasing concentration of flaviolin. B, Data are fitted to form a plot of ΔA(387-418) vs [flaviolin]. The $K_d$ value (7.3 µM) was estimated. The flaviolin structure and carbon numbering are shown.

FIG. 2. **Catalytic activity of CYP158A2 supported by flavodoxin and flavodoxin reductase.** Oxidation reactions were carried out as described under Experimental Procedures. A, Product profile obtained following the flavodoxin and flavodoxin reductase supported reaction using CYP158A2 (1 nmol) and flaviolin (0.26 µmol) at 37 °C for 2 h. Four major products are noted P1, P2, P3, and P4, respectively. B, Negative control incubation, as in A but without CYP158A2.

FIG. 3. **Chemical structure identification of P1 and P2 by $^1$H NMR analyses.** NMR spectra were recorded on a Bruker DPX 500 AVANCE spectrometer operating at a proton frequency of 500 MHz. Chemical shifts are reported as δ values with reference to acetone-$d_6$ for P1 and to DMSO-$d_6$ for P2 as internal standards, respectively. A, $^1$H NMR spectrum of P1; B, $^1$H NMR spectrum of P2; C, $^1$H NMR spectrum (aromatic region) of P1 and peak assignments from the assigned structure; D, $^1$H NMR spectrum (aromatic region) of P2 and peak assignments from the assigned structure; E, chemical structure of P1; F, chemical structure of P2.

FIG. 4. **Two views of the ribbon diagram of CYP158A2;** A, Substrate-free structure in orange. B, Substrate-bound in cyan; heme is colored red; two flaviolin molecules are colored purple. The
left hand structures show the traditional P450 orientation. The right hand structures have been rotated 90° into the plane of the figure. The figure was produced using Setor (31).

FIG. 5. **Two views of the superposition of the substrate-free (orange) and substrate-bound structures (cyan).** Two flaviolins are colored purple. CYP158A2 undergoes a significant conformational change in the BC loop and F/G region upon flaviolin binding. The positions are identical to those shown in Fig. 4.

FIG. 6. **Electron density for flaviolin in the structure of substrate-bound CYP158A2.** The electron density map was calculated using σA-weighted 2|Fo| - |Fc| coefficients and is contoured at 1.5σ. The two electron densities above the heme were from two flaviolin molecules. The labeled distances are in Angstroms. The first flaviolin and second flaviolin molecules are denoted “F1” and “F2”, respectively. Note that the two flaviolin molecules have their C6 (C6’) ends in close proximity but are rotated 180° from each other.

FIG. 7. **Stereoview of the hydrogen bonding network between R71, R288, L293 and ordered waters and the two flaviolin molecules in the CYP158A2 active site;** The side chain atoms and flaviolin atoms are rendered as stick figures. Water molecules are displayed as red spheres. Potential hydrogen bonds are dotted green lines.
FIG. 8. Proposed diradical mechanism for formation of the flaviolin dimers by CYP158A2.

FIG. 9. A cationic mechanism for formation of flaviolin dimers, proposed as a possible alternative to the mechanism in Fig 8.
### Table I Data Collection and Refinement Statistics

<table>
<thead>
<tr>
<th></th>
<th>Substrate-free</th>
<th>Substrate-bound</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection statistics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
<td>P₂₁2₁2₁</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Total observations</td>
<td>460826</td>
<td>280222</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>73147</td>
<td>54055</td>
</tr>
<tr>
<td>Redundancy</td>
<td>6.3(5.8)</td>
<td>5.2(4.9)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>89.2(92.4)</td>
<td>97.8(99)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>21.6(6.5)</td>
<td>30.1(3.2)</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>5.8(25)</td>
<td>7.8(37)</td>
</tr>
<tr>
<td><strong>Refinement statistics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>36-1.75</td>
<td>38-1.62</td>
</tr>
<tr>
<td>No. of reflections used in refinement</td>
<td>71548</td>
<td>46733</td>
</tr>
<tr>
<td>No. of water molecules</td>
<td>539</td>
<td>242</td>
</tr>
<tr>
<td>Protein atoms</td>
<td>6045</td>
<td>3095</td>
</tr>
<tr>
<td>Heme atoms</td>
<td>86</td>
<td>43</td>
</tr>
<tr>
<td>Ligand atoms</td>
<td>0</td>
<td>30ᵄ</td>
</tr>
<tr>
<td>R_work %</td>
<td>23.4</td>
<td>24.5</td>
</tr>
<tr>
<td>R_free %</td>
<td>26.8</td>
<td>27.8</td>
</tr>
<tr>
<td>Rmsd in bond lengths (Å)</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td>Rmsd in bond angles (°)</td>
<td>1.6</td>
<td>1.4</td>
</tr>
</tbody>
</table>

* Values for the highest resolution shell in parentheses

* Residues 5-175/194-406(A) and 5-406(B) of substrate-free; residues 4-406 of substrate-bound

* Two flavolin molecules
FIG. 2

(A) Chromatogram of flavinol. The peaks correspond to different mass-to-charge ratios (m/z): m/z 411 and m/z 615. Peaks labeled P1, P2, P3, and P4 denote specific compounds.

(B) Second chromatogram showing a different compound profile. The chromatogram is labeled with time (t_g) in minutes from 12 to 24.
FIG. 3
FIG. 7
FIG. 9
Binding of two flaviolin substrate molecules, oxidative coupling, and crystal structure of streptomyces coelicolor A3(2) cytochrome P450 158A2

Bin Zhao, F. Peter Guengerich, Aouatef Bellamine, David C. Lamb, Miho Izumikawa, Li Lei, Larissa M. Podust, Munirathinam Sundaramoorthy, John A. Kalaitzis, L. Manmohan Reddy, Steven L. Kelly, Bradley S. Moore, Donald Stec, Markus Voehler, John R. Falck, Tsutomu Shimada and Michael R. Waterman

J. Biol. Chem. published online January 19, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M410933200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2005/01/25/M410933200.DC1

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2005/01/19/jbc.M410933200.citation.full.html#ref-list-1