A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Substrate recognition by a dual functional P450 monooxygenase GfsF involved in FD-891 biosynthesis

Authors: Tadashi Eguchi, Akimasa Miyanaga, Ryuichi Takayanagi, Takashi Furuya, Ayano Kawamata, Tomohiro Itagaki, Yoshiharu Iwabuchi, Naoki Kanoh, and Fumitaka Kudo

To be cited as: *ChemBioChem* 10.1002/cbic.201700429

Link to VoR: http://dx.doi.org/10.1002/cbic.201700429

A Journal of

ChemPubSoc Europe

WILEY-VCH

www.chembiochem.org

Substrate recognition by a dual functional P450 monooxygenase GfsF involved in FD-891 biosynthesis

Akimasa Miyanaga,[a] Ryuichi Takayanagi,[a] Takashi Furuya,[a] Ayano Kawamata,[b] Tomohiro Itagaki,[b] Yoshiharu Iwabuchi,[b] Naoki Kanoh,[b] Fumitaka Kudo[a] and Tadashi Eguchi*[a]

Abstract: GfsF is a multifunctional P450 monooxygenase that catalyzes the epoxidation and subsequent hydroxylation in the biosynthesis of macrolide polyketide FD-891. Here, we describe the biochemical and structural analysis of GfsF. To obtain the structural basis of a dual functional reaction, we determined the crystal structure of ligand-free GfsF, which revealed GfsF to have a predominantly hydrophobic substrate binding pocket. The docking models in conjunction with the results of the enzymatic assay with substrate analogs as well as site-directed mutagenesis suggested two distinct substrate binding modes for epoxidation and hydroxylation reactions, which explained how GfsF regulates the order of two oxidative reactions. These findings provide new insights into the reaction mechanism of multifunctional P450 monooxygenases.

Introduction

Cytochrome P450 monooxygenases are a large superfamily of heme-containing enzymes. They catalyze various regio- and stereo-specific oxidative reactions such as hydroxylation, epoxidation, dehydrogenation, demethylation, aryl–aryl coupling and C–C bond cleavage in primary and secondary metabolism.[1] P450 enzymes are often involved in the biosynthesis of natural products in bacteria, fungi and plants, and provide the structural diversification of natural products.[2] In general, P450s catalyze a single oxidative reaction, but some P450s catalyze multistep oxidative reactions at separate sites of the substrate.[3] For example, MycG catalyzes sequential hydroxylation and epoxidation reactions in mycinamicin biosynthesis (Scheme S1A).[4] AurH catalyzes the hydroxylation and the tetrahydrofuran ring formation in aureothin biosynthesis (Scheme S1B).[5] TamI catalyzes two hydroxylations and one epoxidation in tirandamycin biosynthesis.[6] These multifunctional P450s seem to have an apparent hierarchy in the order of catalytic steps. Each oxidation step is a prerequisite for the next oxidation step. Although some of the multifunctional P450s have been characterized in terms of function and structure, the underlying mechanisms for their diverse reactivity and ordered reaction sequence are not well understood.

FD-891 (**1a**) is a 16-membered macrolide antibiotic isolated from *Streptomyces graminofaciens* A-8890.[7] FD-891 (**1a**) shows toxicity against human leukemia (HL-60) cells. The FD-891 (**1a**) biosynthetic gene cluster has been identified and found to contain five type I polyketide synthase (PKS) genes (*gfsA*, *gfsB*, *gfsC*, *gfsD* and *gfsE*) and two post-PKS modification enzyme genes (*gfsF* and *gfsG*).[8] The five PKSs are responsible for the biosynthesis of FD-892 (**2c**). Cytochrome P450 GfsF and methyltransferase GfsG catalyze the post-PKS modification in parallel.[9] GfsF catalyzes the epoxidation at the C8–C9 olefin of FD-892 (**2c**) and 25-*O*-methyl-FD-892 (**1c**) and subsequent hydroxylation at the C10 position of the resulting intermediate to produce 25-*O*-demethyl-FD-891 (**2a**) and FD-891 (**1a**), respectively (Scheme 1). GfsG catalyzes the *O*-methylation at the C25 hydroxy group of FD-892 (**2c**) and 25-*O*-demethyl-FD-891 (**2a**) to produce 25-*O*-methyl-FD-892 (**1c**) and FD-891 (**1a**), respectively. Among these post-PKS reactions, GfsF catalyzes epoxidation and hydroxylation in a stepwise manner. The order of the two oxidative reactions is strictly regulated. It is unclear how GfsF regulates the order of dual reactions.

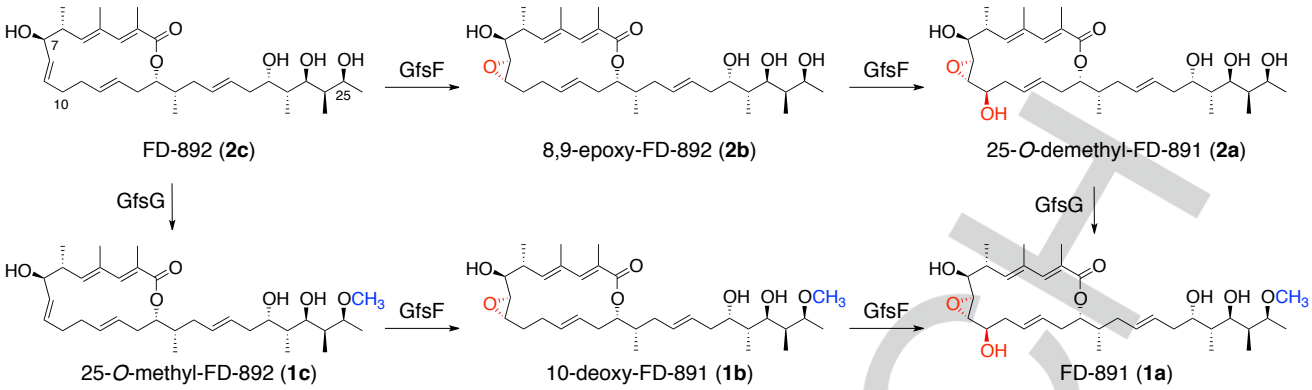
Here, we describe the biochemical characterization and structural determination of GfsF. The structural data combined with a computational docking study and mutational study provides insights into the mechanism of substrate recognition important for tandem oxygenation processes.

Results

Substrate Specificity of GfsF. Previously, we showed that GfsF can accept 25-*O*-methyl-FD-892 (**1c**) and FD-892 (**2c**) as substrates to produce FD-891 (**1a**) and 25-*O*-demethyl-FD-891 (**2a**), respectively (Scheme 1).[9] To further understand the substrate specificity of GfsF, we conducted the reaction of GfsF with **1c** and **2c** to compare the velocity of the GfsF reaction. As a result, GfsF efficiently converted **2c** to **2a**, whereas GfsF converted **1c** to **1a** with lower efficiency (Figure 1A, B). This result suggested that GfsF prefers the hydroxy group at the C25 position rather than the methoxy group. We also found that GfsF catalyzed the epoxidation step faster than the hydroxylation step. After **1c** or **2c** was almost converted to **1b** or **2b** in the GfsF reaction, the formation of **1a** or **2a** started to be observed, respectively. We estimated both

[a] Dr. A. Miyanaga, R. Takayanagi, T. Furuya, Professor Dr. F. Kudo, Professor Dr. T. Eguchi
Department of Chemistry
Tokyo Institute of Technology
2-12-1 O-okayama, Meguro-ku, Tokyo 152-8551 (Japan)
E-mail: eguchi@chem.titech.ac.jp
[b] A. Kawamata, T. Itagaki, Professor Dr. Y. Iwabuchi, Professor Dr. N. Kanoh
Graduate School of Pharmaceutical Sciences
Tohoku University
6-3 Aza-aoba, Aramaki, Aoba-ku, Sendai 980-8578, (Japan)

Supporting information for this article is given via a link at the end of the document.



Scheme 1. Parallel biosynthetic pathway from FD-892 (**2c**) to FD-891 (**1a**)

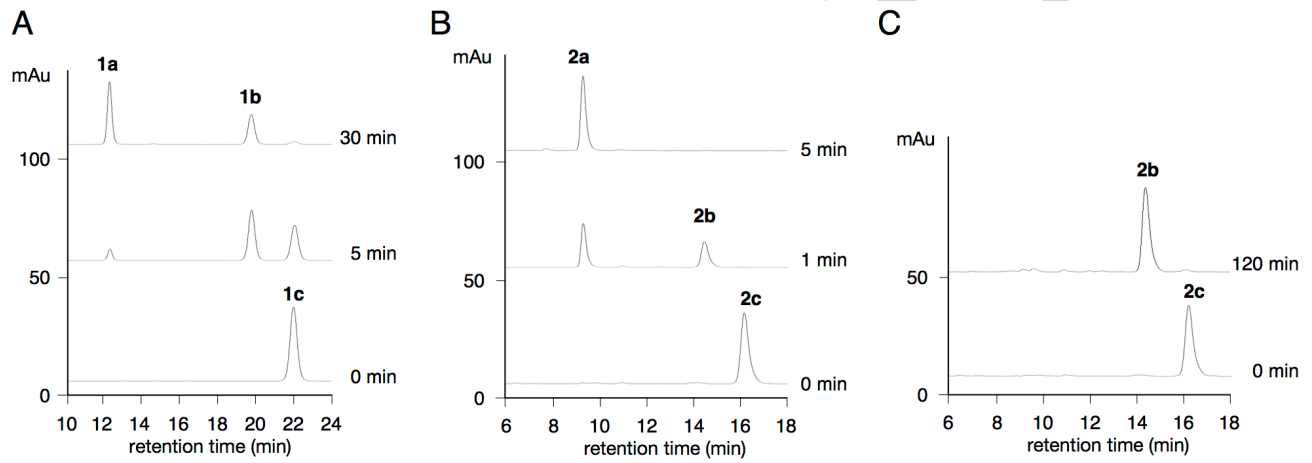
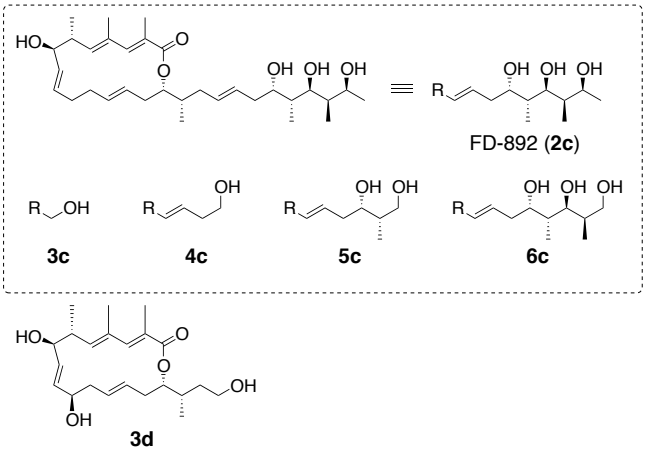


Figure 1. Time-course of the reactions of GfsF wild-type with **1c** (A) and **2c** (B), and GfsF T300V mutant with **2c** (C). HPLC traces (detection at 275 nm) are shown.

initial velocities of epoxidation (0.21 units/mg for **1c** and 1.5 units/mg for **2c**) and hydroxylation (0.045 units/mg for **1b** and 0.23 units/mg for **2b**) reactions (Figure S1).

Next, we investigated the GfsF reaction with substrate analogs (Scheme 2) to obtain an insight into the substrate specificity in each oxidation step of the GfsF reaction. Previously, we synthesized FD-892 analog **3c** that has a shorter C16–C18 alkyl side-chain,[10] so we first carried out the reaction with **3c** to investigate whether the alkyl side-chain moiety is important for the GfsF reaction. We analyzed the reaction product by high-performance liquid chromatography (HPLC) and liquid chromatography-mass spectrometry (LC-MS) and then found that GfsF catalyzed only a single oxidation to produce **3b** (Figure 2A). We isolated **3b** and analyzed its chemical structure by nuclear magnetic resonance (NMR) spectroscopy, which showed that **3b** has an epoxide at the C8–C9 positions and the methylene at the C10 position (Table S1 and Figures S2–S3). This finding suggested that GfsF catalyzed only the epoxidation of the double bond at the C8–C9 position of **3c** (Fig. 2B). Thus, the recognition of a trihydroxyalkyl side-chain moiety does not seem to be critical for the epoxidation step, although the formation rate of the epoxy product was significantly lower (0.96% of relative activity towards **2c**). Conversely, GfsF showed no detectable hydroxylation activity (<0.05% of relative activity towards **2b**), suggesting that GfsF strictly recognizes the presence of a trihydroxyalkyl side-chain in the hydroxylation reaction. Thus, GfsF seems to recognize the alkyl side-chain moiety in a different manner between epoxidation and hydroxylation reactions.



Scheme 2. Synthetic analogs that have a truncated alkyl side-chain

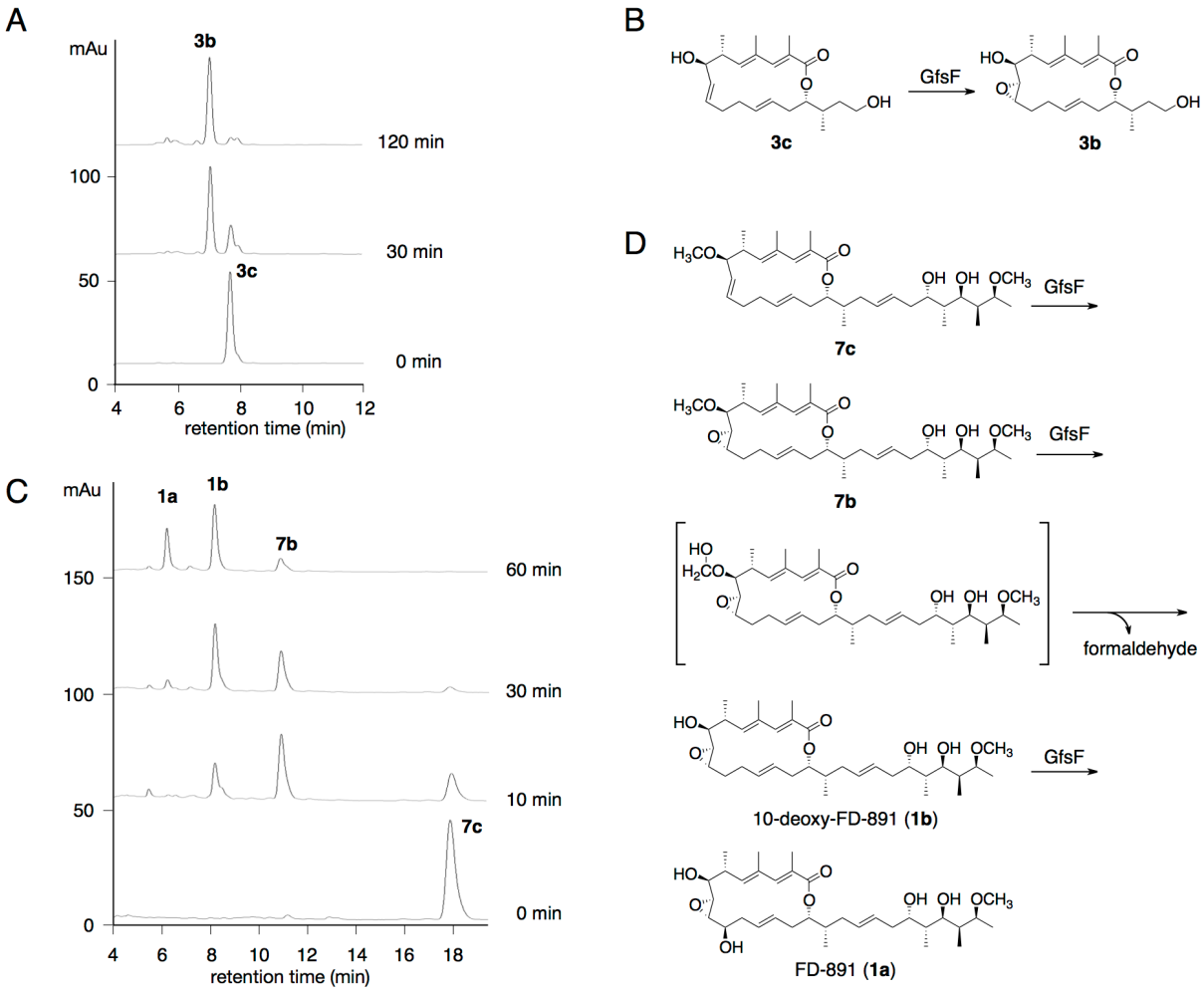


Figure 2. Reactions of GfsF wild-type with **3c** (A and B) and **7c** (C and D). (A) Time-course of the reaction with **3c.** (B) Formation of **3b** from **3c**. (C) Time-course of the reaction with **7c.** (B) Formation of **1a** from **7c**. (A and C) HPLC traces (detection at 275 nm) are shown.

To further investigate the recognition of an alkyl side-chain moiety in the GfsF reaction, we synthesized additional FD-892 analogs possessing a truncated alkyl side-chain using the same strategy as for the synthesis of FD-891 side-chain truncated analogs[11] and examined the reaction (Figures S1 and S4). In the reaction with FD-892 analog **4c** that has a monohydroxyalkyl side-chain (i.e., C16–C21), GfsF showed moderate epoxidation activity (35% relative activity towards **2c**) and significantly low hydroxylation activity (2.7% relative activity towards **2b**). These findings suggested that the C21 hydroxy group of substrate is recognized by GfsF in the epoxidation step because the epoxidation activity against **4c** was 36-fold higher than that against **3c**. In the reaction with FD-892 analog **5c** that has a dihydroxyalkyl side-chain (i.e., C16–C23), GfsF also showed moderate epoxidation activity (25% relative activity towards **2c**) and significantly low hydroxylation activity (1.4% relative activity towards **2b**), both of which were at the same level as those with **4c**. The presence of a C23 hydroxy group did not improve epoxidation or hydroxylation activity, suggesting that GfsF might not recognize the C23 hydroxy group. In the reaction with FD-892 analog **6c** that has a one-carbon truncated alkyl side-chain (i.e., C16–C25), GfsF efficiently catalyzed epoxidation (190% relative activity towards **2c**) and hydroxylation (260% relative activity towards **2b**). GfsF seems to strictly recognize the C24–C25 region in the hydroxylation step because the hydroxylation activity against **6b** was 190-fold higher than that against **5b**. GfsF showed even higher epoxidation and hydroxylation activities against **6c** than those against FD-892 (**2c**). The flexible C25 hydroxy group of **6c** might facilitate the interaction with GfsF.

To investigate whether GfsF recognizes the C7 hydroxy group of the substrate, we synthesized 7,25-*O*-dimethyl-FD-892 (**7c**) from 25-*O*-methyl-FD-892 (**1c**). Unexpectedly, GfsF converted **7c** to FD-891 (**1a**) in the reaction (Figure 2C). This implied that GfsF accepted **7c** as a substrate and catalyzed the demethylation at the C7 hydroxy group during the reaction. To clarify the timing of demethylation, we investigated the time-course of this conversion (Figure 2C). We found that the reaction with **7c** gave two reaction intermediates, one of which was 10-deoxy-FD-891 (**1b**). We isolated another reaction intermediate, **7b**, and structurally identified it as 7-*O*-methyl-10-deoxy-FD-891 by LC-MS and NMR (Table S2 and Figures S5–

S6). We did not detect the formation of 7-*O*-methyl-FD-891 or 25-*O*-methyl-FD-892 (**1c**) from **7c**. We also carried out the reaction with **7b**, which was converted to FD-891 (**1a**) via 10-deoxy-FD-891 (**1b**) (Figure S7). These results clearly suggested that the C7 methoxy group was retained in the epoxidation step and that demethylation occurred before the hydroxylation step (Figure 2D). We detected the formation of formaldehyde (Figure S8), which was concomitant with the conversion of **7c** to **1a**. This observation confirmed that GfsF catalyzed the hydroxylation of the methyl group at the C7 methoxy moiety of **7b** to produce **1b** via a hemiacetal form.

We also carried out the reaction with FD-892 analog **3d** that has a hydroxy group at the C10 position[10] to obtain an insight into the order of dual oxidative reactions. As a result, GfsF did not catalyze the epoxidation of **3d**. Thus, the hydroxylation at the C10 position prior to the epoxidation at the C8–C9 positions completely blocked the further oxidative reaction of GfsF.

Crystal Structure of GfsF. To understand the structural basis of the dual oxidation reactions, we attempted to crystallize GfsF, but we failed. Computational analysis using Protein DisOrder prediction System (PrDOS)[12] suggested that the first 15 amino acid residues at the N-terminus of GfsF are flexible. To minimize the conformational flexibility, we constructed the heterologous expression system of GfsFΔN15 protein that lacks the N-terminal 15 amino acid residues. After we confirmed the activity of GfsFΔN15 protein, we succeeded in the crystallization of GfsFΔN15 and determined the ligand-free structure at 2.0 Å resolution (Table S3).

GfsFΔN15 has a typical P450 fold and contains a heme with invariant Cys363 as an axial thiolate (Figures 3 and S9). A water molecule is coordinated to the heme iron as a distal pocket ligand, as observed in many ligand-free P450 structures.[2a] The overall structure of GfsFΔN15 is most similar to that of P450nor (PDB code 2ROM; Z-score = 52.2, rmsd of 1.3 Å, sequence identity of 38%), which catalyzes the reduction of nitric oxide in *Fusarium oxysporum*.[13] GfsFΔN15 also shows structural similarity with post-PKS modification P450s such as CYP105P1 (PDB code 3E5L; Z-score = 49.0, rmsd of 1.7 Å, sequence identity of 42%) and CYP105D6 (PDB code 3ABB; Z-score = 48.9, rmsd of 1.6 Å, sequence identity of 44%), both of which are involved in filipin biosynthesis in *Streptomyces avermitilis*.[14] In addition, GfsFΔN15 shows similarity with multifunctional P450s such as MycG[4] (PDB code 2Y5N; Z-score = 47.6, rmsd of 1.9 Å, sequence identity of 36%) and AurH[5] (PDB code 3P3X; Z-score = 41.5, rmsd of 2.2 Å, sequence identity of 28%).

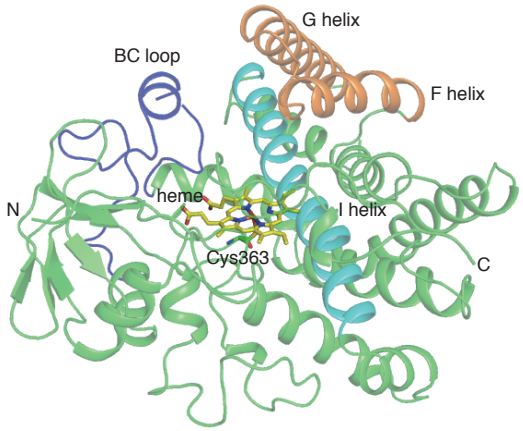


Figure 3. Overall structure of GfsFΔN15. The N- and C-terminals of GfsFΔN15 are denoted as N and C, respectively. The BC loop region (Asp70–Asp104), FG helix region (Tyr171–Leu213), I helix region (Gln235–Glu266) and heme are shown in blue, orange, cyan and yellow, respectively.

The FG helices, which are known to undergo closing motion on ligand binding in P450s,[2,14b] adopt an open conformation in the GfsFΔN15 structure. The BC loop located at the entrance of the substrate binding pocket shows a relatively high B-factor and contains a disordered region (Met94–Pro97), suggesting the flexibility of the BC loop. These structural features are, in general, observed in ligand-free P450 structures.[2,14b] GfsFΔN15 has a predominantly hydrophobic pocket constructed by the BC loop (His75, Tyr82–Ile85, Phe89 and Phe99–Gly101), the I helix (Met244, Gly247, Ala248 and Thr252), the loop after the K helix (Thr300), and the C-terminal loop (Ile403 and Ala404). We observed the electron density for di(hydroxyethyl)ether molecule which likely originated from polyethylene glycol (PEG) in the crystallization solution, in the substrate binding pocket (Figure 4A). A PEG-type molecule was reported to bind in the hydrophobic tunnel or pocket in some enzymes.[15] The di(hydroxyethyl)ether molecule is positioned almost parallel to the heme molecule (3.8–4.5 Å). The di(hydroxyethyl)ether molecule forms a hydrogen bond (3.3 Å) with Thr300 and makes hydrophobic contacts (3.6–4.1 Å) with Ala248 and Thr252. The Thr252 residue of GfsFΔN15 is positioned at the same position as the corresponding Thr residues in many P450s, such as P450cam, CYP105P1 and CYP105D6. This conserved Thr residue is proposed to play an important role in protonation for oxygen activation in the P450 reaction.[16]

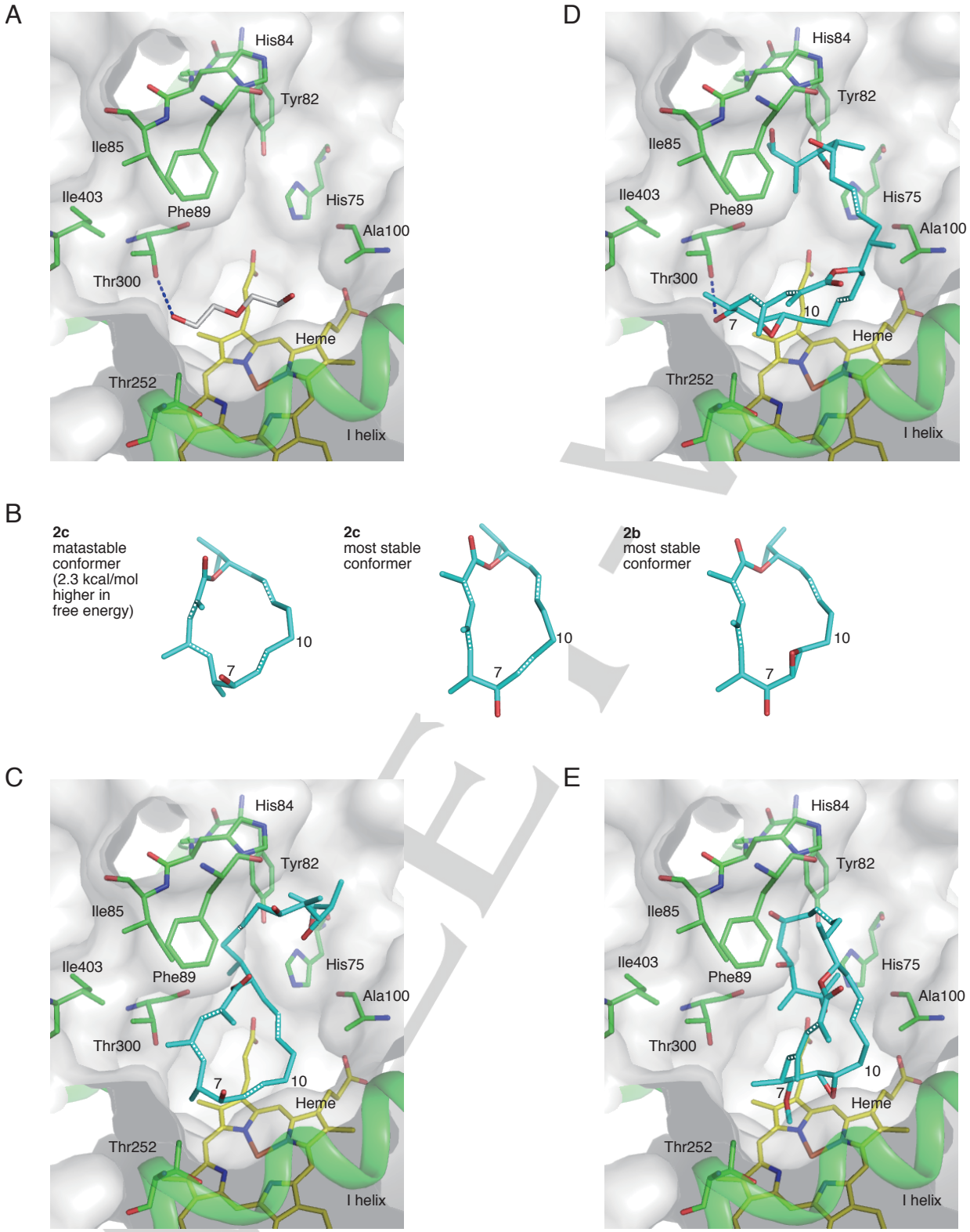


Figure 4. GfsF docking models. The GfsFΔN15 and docked FD-892-related molecules are shown in green and cyan, respectively. The C7 and C10 positions of docked molecules are denoted as 7 and 10, respectively. The hydrogen bond with Thr300 is shown as a blue broken line. The olefin bonds of docked molecules are shown as white dotted lines. (A) The crystal structure of GfsFΔN15. The bound di(hydroxyethyl)ether molecule is shown as gray sticks. (B) A comparison of stable conformers for FD-892 (**2c**) and 8,9-epoxy-FD-892 (**2b**) macrolactones. **2c** and **2b** macrolactone fragments that have a truncated C16–C17 alkyl side-chain were used for MacroModel calculation. (C) The docking model with FD-892 (**2c**). (D) The docking model with 8,9-epoxy-FD-892 (**2b**). (E) The docking model with 7-*O*-methyl-8,9-epoxy-FD-892 (**7b**).

Docking Analysis and Mutational Study. We attempted to cocrystallize GfsFΔN15 with a substrate or substrate analog.

However, these attempts were unsuccessful. Therefore, we conducted computational docking analysis with FD-892 (**2c**) and 8,9-epoxy-FD-892 (**2b**) models whose macrolide rings are energy-minimized by MacroModel® software (Schrödinger, New York, NY, USA). Based on MacroModel calculation, the macrolide ring of **2c** seems to adopt multiple conformations (Figures 4B and S10). The conformation of the most stable conformer closely resembles the X-ray structure of the FD-892 derivative.[10] However, the most stable **2c** conformer did not correctly dock near the heme cofactor. Therefore, we conducted the docking analysis with several other **2c** conformers. We found that the metastable **2c** conformer (2.3 kcal/mol higher in free energy) showing a different conformation in the C1–C9 polyene region was docked near the heme cofactor (Figure 4C). The C6 methyl and C7 hydroxy groups of the most stable conformer are oriented horizontally along the macrolide ring, whereas those of the metastable conformer are oriented almost vertically. The metastable conformation of **2c** might be necessary to fit into the substrate binding pocket of GfsF. In particular, the vertically oriented conformations of C6 methyl and C7 hydroxy groups seem to be necessary for C8–C9 olefin to access near the heme group without steric hindrance. The C8–C9 olefin is located within 4.7 Å of the heme iron and positioned to give an epoxide with the correct stereochemistry in FD-891 (**1a**). In the case of the **2b** docking model, the most stable conformer was well docked (Figure 4D). The orientations of C6 and C7 substituent groups of the **2b** conformer are similar to those of the most stable **2c** conformer (Figures 4B, S10 and S11). The C10 hydroxylation site of **2b** is located near the heme iron (4.0 Å) and positioned to give a hydroxy group with the correct stereochemistry in the docking model.

We found that the position of macrolide ring is significantly different between FD-892 (**2c**) and 8,9-epoxy-FD-892 (**2b**) binding modes (Figure 4C, D). In the **2c** docking model, the macrolide ring is sandwiched between the heme and Phe89, which seems to be important for fixing the position of the C8–C9 olefin region for the epoxidation reaction (Figure 4C). The macrolide ring moiety forms hydrophobic interactions with Ile85, Phe89, Ala100 and Ile403. The alkyl side-chain moiety is positioned around the flexible BC loop region. The C21 and C25 hydroxy groups are located close to the main-chain carbonyl oxygen atoms of Phe89 and His75, respectively. In the docking model with **2b**, the macrolide ring is bound near the I helix, which is on the opposite side from Phe89 (Figure 4D). The introduced C8–C9 epoxide does not interact with the GfsF residue. The region between the C7 hydroxy and C12 methylene groups of the docked **2b** molecule almost overlaps with the di(hydroxyethyl)ether molecule in the crystal structure (Figure 4A, D). The alkyl side-chain moiety is oriented perpendicular to the macrolide ring, and the C21–C25 region reaches into the BC loop region. The C21–C25 region is anchored in the hydrophobic groove constructed by the side-chains of Tyr82, Ile85 and Phe89. The C24 methyl group is located close to the side-chains of Ile85 and Phe89, and the C25 hydroxy group is close to the main-chain nitrogen atom of Ile85. This observation seems to be consistent with the result that the hydroxylation activity against **6b** was 190-fold higher than that against **5b** (Figure S1). The presence of an alkyl side-chain moiety might be important for fixing the position of the macrolide ring for the hydroxylation reaction. As mentioned above, Phe89 seems to be important for fixing the macrolide ring in the epoxidation reaction and for binding the alkyl side-chain in the hydroxylation reaction. To confirm the role of Phe89 in both reactions, we constructed a F89A mutant. The F89A mutant showed significantly low epoxidation activity (2.3% of relative activity of the wild-type) and no detectable hydroxylation activity in the reaction with **2c**.

The position of the C7 hydroxy group is different between FD-892 (**2c**) and 8,9-epoxy-FD-892 (**2b**) in the docking models. The C7 hydroxy group of **2c** is directed toward the space in the substrate binding pocket and shows no direct interaction with any GfsF residue (Figure 4C). Thus, the methoxy group at the C7 position could be accommodated with no steric hindrance in this binding mode, which is consistent with the result that GfsF catalyzed the epoxidation of **7c**. In contrast, the C7 hydroxy group of **2b** forms a hydrogen bond with the side-chain hydroxy group of Thr300 (Figure 4D). This finding might explain why the hydroxylation reaction occurred only after the demethylation at the C7 hydroxy group in the reaction with **7c**. To evaluate the importance of Thr300 in the hydroxylation reaction, we constructed T300V and T300A mutants. The T300V mutant retained the epoxidation activity towards **2c** (250% relative activity of the wild-type), suggesting that Thr300 is not involved in the epoxidation step. The substitution of Thr300 with Val might even be preferable in the epoxidation step because the C3–C5-conjugated olefin region was docked close to Thr300 in the **2c** docking model (Figure 4C). Conversely, the T300V mutant completely lost the hydroxylation activity (Figure 1C). The substitution of Thr300 with Val likely caused repulsion with the C7 hydroxy group of **2b**. Similarly, the T300A mutant retained the epoxidation activity (93% relative activity of the wild-type) and reduced the hydroxylation activity (13% relative activity of the wild-type), probably because of the loss of interaction with the C7 hydroxy group in the hydroxylation step. Thus, the result of the mutational experiment supported the docking models.

To obtain mechanistic insights into the demethylation process, we conducted docking analysis with **7b**. In the docking model with **7b**, the positions of the macrolide ring and alkyl side-chain moieties of **7b** occupies relatively similar positions to those of 8,9-epoxy-FD-892 (**2b**) rather than those of FD-892 (**2c**) (Figures 4D, E). However, the C7 position of **7b** is moved away from Thr300 by ≈3 Å compared with that of **2b**. The methyl group of the C7 methoxy moiety of **7b** is placed close to the heme iron (3.1 Å) in the docking model so that GfsF could catalyze the hydroxylation of the methyl group to produce 10-deoxy-FD-891 (**1b**) via a hemiacetal form. Because the methoxy group is not bulky, GfsF might be able to accommodate **7b** in the substrate binding pocket for the demethylation reaction.

Discussion

Several multifunctional P450s have been structurally analyzed and their substrate recognition mechanisms proposed.

For example, the complex structure of MycG with its substrate was determined by Li and colleagues.[4] The C14 hydroxylation site and the C12–C13 epoxidation site of the substrate are 8.9 Å and 10.0 Å away from the heme iron in the structure of the MycG complex, respectively. Li and colleagues proposed that this substrate binding mode is the initial recognition mode and that the substrate translocates from the initial recognition site to the active site of the enzyme. Translocation of the substrates in two directions might account for the bifunctional activity of MycG. In the case of AurH, docking studies have suggested that Gln91 forms a hydrogen bond with the introduced hydroxy group of the hydroxylated intermediate.[5] The hydroxylated intermediate was proposed to be pushed deeper into the substrate binding pocket so that the C9 methyl group could be placed closer to the heme iron for the second oxygenation–heterocylization step. The relocation also likely happened during the GfsF reaction based on the docking models with FD-892 (**2c**) and 8,9-epoxy-FD-892 (**2b**). The docking models suggested two substrate binding modes for epoxidation and hydroxylation reactions.

In the GfsF reaction, only epoxidized compounds were initially generated from all of the accepted FD-892-related compounds (**1c**–**7c**). The conformation of macrolide ring could influence the order of catalytic steps. GfsF seems to select the metastable conformer of FD-892 (**2c**) rather than the most stable conformer with the hydrophobic substrate binding pocket. The conformation of the metastable conformer of **2c** is important for fixing the C8–9 site near the heme iron in the first epoxidation reaction. GfsF hydrophobic residues such as Phe89 contribute greatly to the correct placement of **2c** in the substrate binding pocket (Figure 4C). The substitution of olefin at C8–C9 with epoxide causes a conformational change in the macrolide ring, which allows 8,9-epoxy-FD-892 (**2b**) to relocate and adopt the orientation in which the C10 site is placed close to the heme iron for the second hydroxylation step. GfsF recognizes the macrolide ring moiety of the substrate mainly through hydrophobic interactions, although the hydrogen bond between Thr300 and the C7 hydroxy group is important for the second hydroxylation step. These predominant hydrophobic interactions might allow two distinct substrate binding modes.

It is known that several P450 enzymes catalyze both double bond epoxidation and allylic hydroxylation of cyclohexene,[17] although the double oxidation of cyclohexene has never been reported. The reaction energy barriers of the double bond epoxidation and the allylic hydroxylation of cyclohexene were calculated to be similar.[18] However, actually, the reactivity of alkenes largely depends on the structures of substrate and enzyme.[17] In the GfsF reaction, the reaction order seems to be completely controlled by the substrate structure, presumably due to the preferable conformation of FD-892 macrolactone as described above. The rate of the second hydroxylation was 15% of that of the first epoxidation in the reaction with **2c**, even though the available substrate in the second hydroxylation was limited in this highly ordered oxidation. Thus, the rate of the second hydroxylation at the sp^3 carbon adjacent to the oxirane ring was not so significantly slower than the first epoxidation. The preferable conformation of the epoxidized compounds could fit well at the active site of GfsF toward the efficient hydroxylation at C10.

A comparison of GfsF reactions with FD-892 substrates (**1c** and **2c**) and analogs (**3c**, **4c**, **5c** and **6c**) suggested that the presence of an alkyl side-chain moiety affects the reaction efficiencies of epoxidation and hydroxylation steps (Figures 1, 2A, S1 and S3). The presence of C21 and C25 hydroxy groups greatly improved the reaction efficiency, suggesting that these hydroxy groups are recognized by GfsF. In addition, the C24 methyl group seems to be important for the hydroxylation step. Previously, we proposed the parallel post-PKS modification by GfsF and GfsG in FD-891 (**1a**) biosynthesis (Scheme 1),[9] but the conversion from **2c** to **1a** via 25-*O*-demethyl-FD-891 (**2a**) might be the preferred pathway in the producer strain, as judged from the substrate preference of GfsF. GfsF catalyzed the epoxidation of **3c** with low efficiency, whereas GfsF showed no hydroxylation activity (Figure 2A). Thus, the macrolide ring itself is not sufficient for the productive binding of the substrate in the hydroxylation step. The presence of an alkyl side-chain moiety seems to be necessary for the position of the macrolide ring of the substrate in the hydroxylation reaction.

Conclusions

We conducted biochemical and structural analyses on the multifunctional P450 monooxygenase GfsF, which catalyzes the epoxidation and hydroxylation reactions in the biosynthesis of the macrolide antibiotic FD-891. We determined the crystal structure of ligand-free GfsF, which enabled the computational docking analysis with FD-892 and 8,9-epoxy-FD-892. The docking models in conjunction with the results of the enzymatic assay with FD-892 analogs and site-directed mutagenesis suggested two distinct substrate binding modes for epoxidation and hydroxylation reactions. The substitution of C8–C9 olefin of FD-892 with epoxide causes a conformational change in the macrolide ring, which allows 8,9-epoxy-FD-892 to relocate for the second hydroxylation reaction. These results provide an insight into how GfsF regulates the order of epoxidation and hydroxylation reactions.

Experimental Section

Synthesis of FD-892 Side-chain Truncated Analogs (4c, 5c and 6c). See details in the Supporting Information.

Synthesis of 7,25-*O*-dimethyl-FD-892 (7c). See details in the Supporting Information.

Preparation of Recombinant GfsF Protein and Related Proteins. *Escherichia coli* BL21(DE3) cells harboring pET28b-*gfsF* plasmids[9] were grown at 28 °C in Luria–Bertani broth containing kanamycin (50 μg mL^{-1}). After the optical density at 600 nm reached 0.6, protein expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (0.05 mM), $FeCl_3$ (0.15 mM) and 5-aminolevulinic acid (40 μg mL^{-1}), and the cells were then cultured for an additional 16 h at 28 °C. The recombinant protein with an N-terminal His-tag was collected from cell-free extracts

prepared by sonication, and was purified on a His60 Ni-Superflow affinity column (Clontech, Mountain View, CA, USA). The protein was then desalted and concentrated using a PD-10 column (GE Healthcare, Buckinghamshire, UK) and an Amicon Ultra centrifugal filter (Merck Millipore, Billerica, MA, USA), respectively. For preparation of GfsFΔN15 protein, the gfsFΔN15 fragment was first amplified from pET28b-*gfsF* with the primers 5'-AAAAAAACATATGGCCCCCGAGTGGCCC-3' and 5'-ATGCTAGTTATTGCTCAGCGG-3' and then inserted between the NdeI and XhoI sites of pET28b. The resulting pET28b-*gfsFΔN15* was transformed into *E. coli* BL21(DE3) cells. The GfsFΔN15 protein was expressed and purified by a His60 Ni-Superflow affinity column as described above. The purified GfsFΔN15 protein was further treated with thrombin to remove the His-tag for crystallization. After subjection to another round of a His60 Ni-Superflow affinity chromatography, the His-tag free GfsFΔN15 protein was purified by Resource Q (GE Healthcare) anion-exchange chromatography with a linear gradient from 0.15 to 0.45 M NaCl in 10 mM HEPES-Na buffer (pH 7.7) containing 10% (*v/v*) glycerol. The purified GfsFΔN15 protein was concentrated to 12.5 mg/mL in 10 mM HEPES-Na (pH 7.7). The sodium dithionite-reduced GfsF solution was bubbled with carbon monoxide and then analyzed with a spectrophotometer UV-2450 (Shimadzu, Tokyo, Japan) to determine the functional P450 concentration using an extinction coefficient of 91,000 $M^{-1}cm^{-1}$.[19] CamA and CamB proteins were prepared as described previously.[9]

GfsF Reaction. The GfsF assay mixture (100 µL each) consisted of 0.1 mM **1c**, **2c** or FD-892 analogs, 1 mM NADH, 30 nM or 3 µM GfsF, 8 µM CamA, and 20 µM CamB in 50 mM HEPES-Na (pH 7.7) containing 10% glycerol. The enzymatic reaction was carried out at 28°C for 1 min–3 h. The product of the enzymatic reaction was extracted twice using 150 µL of EtOAc, and the solvents from the combined organic layers were removed using a centrifugal evaporator. The residue was dissolved in 10 µL of CH_3OH, and a 5-µl aliquot was analyzed by HPLC using an ELITE LaChrom L-2455 DAD Detector and L-2130 Pump (Hitachi, Tokyo, Japan) equipped with a PEGASIL ODS column (100 Å, 250 × 4.6 mm; Senshu, Tokyo, Japan). Except for the reaction with **7c**, CH_3OH (80%) in water was used as an eluent at a flow rate of 0.7 mL min^{-1}. In the reaction with **7c**, CH_3OH (90%) in water was used as an eluent at a flow rate of 0.7 mL min^{-1}. One unit was defined as the amount of enzyme that released 1 µmol of product per minute.

Detection of Formaldehyde. The GfsF assay mixture (50 µL each) consisted of 0.1 mM **7c**, 1 mM NADH, 3 µM GfsF, 8 µM CamA, and 20 µM CamB in 50 mM HEPES-Na (pH 7.7) containing 10% glycerol. The enzymatic reaction was carried out at 28°C for 10 min–2 h. Then, 150 µL of CH_3CN was added to quench the reaction. After the addition of 4 µL of 20% phosphoric acid and 10 µL of 1 mg mL^{-1} 2,4-dinitrophenylhydrazine, the reaction mixture was incubated at 28°C for 20 min. The resulting mixture was analyzed by HPLC using an ELITE LaChrom L-2455 DAD Detector and L-2130 Pump equipped with a PEGASIL ODS column (100 Å, 250 × 4.6 mm). CH_3CN (50%) in water was used as an eluent at a flow rate of 1.0 mL min^{-1}. The derivatized formaldehyde (formaldehyde 2,4-dinitrophenylhydrazone) was detected at 360 nm.

Isolation and Structural Determination of 3b. GfsF-containing reactant was prepared by using *E. coli* BL21(DE3) cells harboring p*gfsF-camAB* plasmid as described previously.[8] **3c** (6.2 mg) were then mixed with the GfsF-containing reactant. The reaction was carried out at 28 °C with agitation (200 rpm) for 36 h. After the reaction, the reaction product was extracted with EtOAc. The organic layers were dried on Na_2SO_4 and the solvent was removed by evaporation. The crude residue was purified by silica gel chromatography with hexane/EtOAc (1:1) to give **3b** (1.9 mg). HR-FAB-MS (positive mode): *m/z* calculated for $C_{22}H_{35}O_5$: 379.2484 ($[M+H]^+$); found: 379.2512. The assignments of the ^1H and ^{13}C NMR spectroscopy signals are shown in the Supporting information (Table S1 and Figures S2–S3).

Isolation and Structural Determination of 7b. 0.5 mM **7c** (20 mg) were mixed with 3 µM GfsF, 10 µM CamA, 20 µM CamB and 1 mM NADH in 20 mM HEPES-Na (pH 7.7) buffer containing 10% glycerol and 1% DMSO in a total volume of 71.5 mL. The reaction was carried out at 28 °C with agitation (600 rpm) for 20 min. After the reaction, the reaction product was extracted with EtOAc. The organic layers were dried on Na_2SO_4 and the solvent was removed by evaporation. The crude residue was purified by HPLC (90% CH_3OH) to give **7b** (3.4 mg). HR-FAB-MS (positive mode): *m/z* calculated for $C_{34}H_{57}O_7$: 577.4104 ($[M+H]^+$); found: 577.4106. The assignments of the ^1H and ^{13}C NMR spectroscopy signals are shown in the Supporting information (Table S2, Figures S5–S6).

Crystallization, Data Collection and Structural Determination. GfsFΔN15 crystals were grown from a 1:1 mixture of a protein solution (12.5 mg mL^{-1} in 10 mM HEPES-Na (pH 7.5)) and a reservoir solution containing 0.2 M KCl, 0.1 M Tris-HCl (pH 8.5) and 27.5% PEG3350 using the sitting-drop vapor diffusion method at 26 °C. Prior to collection of the X-ray data, the crystals were flash-frozen in a stream of liquid nitrogen. The X-ray diffraction data were collected on a beamline BL-5A at the Photon Factory (Tsukuba, Japan) and were subsequently indexed, integrated, and scaled using the HKL2000 program.[20] The initial phase was determined by molecular replacement using the Molrep program[21] with the MoxA structure (PDB code: 2Z36)[22] as a search model. Model building of GfsF was carried out automatically with the ARP/wARP program[23] and subsequently inspected by Coot.[24] Refmac[25] was used for refinement of the structures. The structural representations were prepared with PyMOL (DeLano Scientific, Palo Alto, CA, USA). The geometries of the final structure were evaluated using the program Rampage.[26] The resulting coordinates and structure factors have been deposited in the Protein Data Bank (PDB code: 5Y1I).

Docking Analysis. The docking study was carried out using AutoDock v4.2.[27] **2c**, **2b** and **7b** molecules were generated using MacroModel (Schrödinger) and the PRODRG2 Server.[28] See details for MacroModel calculation in the Supporting Information. Chain A of the crystal structure of GfsF was used for the docking study. Using AutoDockTools, polar hydrogen atoms were added to amino acid residues, and Gasteiger charges were assigned to all atoms of the protein. All rotatable bonds of the alkyl side-chain moiety of the ligand molecule were set to be flexible for the calculation, whereas all of the protein residues and the macrolide ring of the ligand molecule were kept rigid. The cubic energy grid was centered at the substrate binding pocket and had an extension of 50 Å in each direction.

Site-directed Mutagenesis. pET28-*gfsF* was used for construction of GfsF mutants. Site-directed mutagenesis was carried out with the following oligonucleotides and their complementary oligonucleotides: F89A, 5'-CACATCAGCGCGGACGCCAAGTTCCTCAGC-3'; T300A, 5'-GATCGCTCAGGACGCCGTGCGCCGGATCG-3'; T300V, 5'-GATCGCTCAGGACGTCGTGCGCCGGATCG-3'. The mutations were confirmed by determining the nucleotide sequences. The plasmids were transformed into *E. coli* BL21(DE3) cells, and the mutated enzymes were prepared as described above.

Acknowledgements

This work was performed with the approval of the Photon Factory Program Advisory Committee (Proposal No. 2012G508).

This work was supported in part by Grants-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science and Technology (16H06451 to TE, 17H05434 to FK and 23102013 to NK) and JSPS A3 Foresight Program (16822333), The Agricultural Chemical Research Foundation to AM and The Naito Foundation to FK in Japan.

Keywords: Biosynthesis, Crystal structure, Cytochromes, Multifunctional oxygenase, Polyketides

[1] a) D. C. Lamb, M. R. Waterman, S. L. Kelly, F. P. Guengerich, *Curr. Opin. Biotechnol.* **2007**, *18,* 504–512; b) F. P. Guengerich, *J. Biochem. Mol. Toxicol.* **2007**, *21,* 163–168; c) S. C. Moody, E. J. Loveridge, *J. Appl. Microbiol.* **2014,** 117, 1549–1563.

[2] a) L. M. Podust, D. H. Sherman, *Nat. Prod. Rep.* **2012,** *29,* 1251–1266; b) J. D. Rudolf, C. Y. Chang, M. Ma, B. Shen, *Nat. Prod. Rep.* **2017**, doi: 10.1039/c7np00034k.

[3] a) R. V. Cochrane, J. C. Vederas, *Acc. Chem. Res.* **2014**, *47*, 3148–3161; b) L. Hang, N. Liu, Y. Tang, *ACS Catal.* **2016**, *6*, 5935–5945.

[4] S. Li, D. R. Tietz, F. U. Rutaganira, P. M. Kells, Y. Anzai, F. Kato, T. C. Pochapsky, D. H. Sherman, L. M. Podust, *J. Biol. Chem.* **2012**, *287*, 37880–37890.

[5] G. Zocher, M. E. Richter, U. Mueller, C. Hertweck, *J. Am. Chem. Soc.* **2011**, *133*, 2292–2302.

[6] J. C. Carlson, S. Li, S. S. Gunatilleke, Y. Anzai, D. A. Burr, L. M. Podust, D. H. Sherman, *Nat. Chem.* **2011**, *3*, 628–633.

[7] a) M. Seki-Asano, T. Okazaki, M. Yamagishi, N. Sakai, K. Hanada, K. Mizoue, *J. Antibiot.* **1994**, *47*, 1226–1233; b) T. Eguchi, K. Yamamoto, K. Mizoue, K. Kakinuma, *J. Antibiot.* **2004**, *57*, 156–157; c) S. Inaba, T. Eguchi, A. Motegi, K. Mizoue, T. Usui, K. Nagai, T. Kataoka, *J. Antibiot.* **2009**, *62*, 507–512.

[8] F. Kudo, A. Motegi, K. Mizoue, T. Eguchi, *Chembiochem* **2010**, *11*, 1574–1582.

[9] F. Kudo, K. Kawamura, T. Furuya, H. Yamanishi, A. Motegi, A. Komatsubara, M. Numakura, A. Miyanaga, T. Eguchi, *Chembiochem* **2016**, *17*, 233–238.

[10] N. Kanoh, A. Kawamata, T. Itagaki, Y. Miyazaki, K. Yahata, E. Kwon, Y. Iwabuchi, *Org. Lett.* **2014**, *16*, 5216–5219.

[11] T. Itagaki, A. Kawamata, M. Takeuchi, K. Hamada, Y. Iwabuchi, T. Eguchi, F. Kudo, T. Usui, N. Kanoh, *J. Antibiot.* **2016**, *69*, 287–293.

[12] T. Ishida, K. Kinoshita, *Nucleic Acids Res.* **2007**, *35*, W460–W464.

[13] S. Y. Park, H. Shimizu, S. Adachi, A. Nakagawa, I. Tanaka, K. Nakahara, H. Shoun, E. Obayashi, H. Nakamura, T. Iizuka, Y. Shiro, *Nat. Struct. Biol.* **1997**, *4*, 827–832.

[14] a) L. H. Xu, S. Fushinobu, H. Ikeda, T. Wakagi, H. Shoun, *J. Bacteriol.* **2009**, *191*, 1211–1219; b) L. H. Xu, S. Fushinobu, S. Takamatsu, T. Wakagi, H. Ikeda, H. Shoun, *J. Biol. Chem.* **2010**, *285*, 16844–16853.

[15] a) L. H. Xu, H. Ikeda, L. Liu, T. Arakawa, T. Wakagi, H. Shoun, S. Fushinobu, *Appl. Microbiol. Biotechnol.* **2015**, *99*, 3081–3091; b) R. Satou, A. Miyanaga, H. Ozawa, N. Funa, Y. Katsuyama, K. Miyazono, M. Tanokura, Y. Ohnishi, S. Horinouchi, *J. Biol. Chem.* **2013**, *288*, 34146–34157.

[16] a) I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, S. G. Sligar, *Science* **2000**, *287*, 1615–1622; b) Nagano S, Poulos TL. *J. Biol. Chem.* **2005,** *280*, 31659–31663.

[17] a) A. D. Vaz, D. F. McGinnity, M. J. Coon, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 3555–3560; b) E. T. Farinas, M. Alcalde, F. Arnold, *Tetrahedron* **2004**, *60*, 525–528.

[18] R. Gupta, X. X. Li, K. B. Cho, M. Guo, Y. M. Lee, Y. Wang, S. Fukuzumi, W. Nam, *J. Phys. Chem. Lett.* **2017**, *8*, 1557–1561.

[19] T. Omura, R. Sato, *J. Biol. Chem.* **1964**, *239*, 2379–2385.

[20] Z. Otwinowski, W. Minor, *Method Enzymol.* **1997**, *276*, 307–326.

[21] A. Vagin, A. Teplyakov, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 22–25.

[22] Y. Yasutake, N. Imoto, Y. Fujii, T. Fujii, A. Arisawa, T. Tamura T, *Biochem. Biophys. Res. Commun.* **2007**, *361*, 876–882.

[23] R. J. Morris, A. Perrakis, V. S. Lamzin, *Acta Crystallogr. D Biol. Crystallogr.* **2002**, *58*, 968–975.

[24] P. Emsley, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.

[25] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. D Biol. Crystallogr.* **1997**, *53*, 240–255.

[26] S. C. Lovell, I. W. Davis, W. B. Arendall, 3rd, P. I. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, D. C. Richardson, *Proteins* **2003**, *50*, 437–450.

[27] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785–2791.

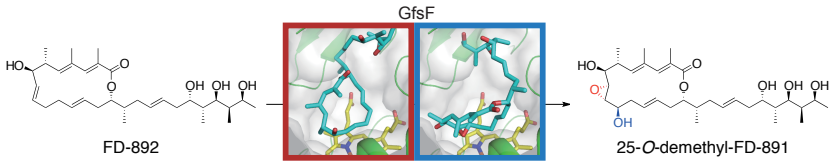
[28] A. W. Schüttelkopf, D. M. F. van Aalten, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 1355–1363.

10.1002/cbic.201700429

WILEY-VCH

Table of Contents

FULL PAPER



GfsF is a multifunctional P450 monooxygenase that catalyzes the epoxidation and subsequent hydroxylation in the biosynthesis of FD-891. X-ray structural and biochemical analyses of GfsF revealed two distinct substrate binding modes for epoxidation and hydroxylation reactions, which explained how GfsF regulates the order of two oxidative reactions.

*Akimasa Miyanaga, Ryuichi Takayanagi, Takashi Furuya, Ayano Kawamata, Tomohiro Itagaki, Yoshiharu Iwabuchi, Naoki Kanoh, Fumitaka Kudo and Tadashi Eguchi**

Page No. – Page No.

Substrate recognition by a dual functional P450 monooxygenase GfsF involved in FD-891 biosynthesis