Molecular insights into the unusually promiscuous and catalytically versatile Fe(II)/ α -ketoglutarate-dependent oxygenase SptF

Hui Tao

University of Tokyo

Takahiro Mori

The University of Tokyo <https://orcid.org/0000-0002-2754-5858>

Heping Chen

University of Tokyo

Shuang Lyu

University of Tokyo

Akihito Nonoyama

Sumitomo Dainippon Pharma Co., Ltd.

Shoukou Lee

Sumitomo Dainippon Pharma Co., Ltd.

Huiping Zhang

RIKEN SPring-8 Center

Fumiaki Hayashi

RIKEN Center for Biosystems Dynamics Research <https://orcid.org/0000-0003-2132-8745>

Ikuro Abe (✉ abei@mol.f.u-tokyo.ac.jp)

University of Tokyo <https://orcid.org/0000-0002-3640-888X>

Article

Keywords: Fe/ α KG oxygenases, SptF, biocatalysts

Posted Date: August 3rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-736558/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Communications on January 10th, 2022. See the published version at <https://doi.org/10.1038/s41467-021-27636-3>.

Molecular insights into the unusually promiscuous and catalytically versatile Fe(II)/ α -ketoglutarate-dependent oxygenase SptF

Hui Tao^{1,6}, Takahiro Mori^{1,2,3,6,*}, Heping Chen¹, Shuang Lyu¹, Akihito Nonoyama⁴, Shoukou Lee⁴, Huiping Zhang⁵, Fumiaki Hayashi⁵, and Ikuro Abe^{1,2,*}

¹ Graduate School of Pharmaceutical Sciences, The University of Tokyo, Tokyo, Japan

² Collaborative Research Institute for Innovative Microbiology, The University of Tokyo, Tokyo, Japan

³ PRESTO, Japan Science and Technology Agency, Saitama, Japan

⁴ Sumitomo Dainippon Pharma Co., Ltd., Osaka, Japan

⁵ RIKEN SPring-8 Center, Yokohama, Japan

⁶ These authors contributed equally

*Address correspondence to tmori@mol.f.u-tokyo.ac.jp or abei@mol.f.u-tokyo.ac.jp

Abstract

Non-heme iron and α -ketoglutarate-dependent (Fe/ α KG) oxygenases catalyze various oxidative biotransformations. Due to their catalytic flexibility and high efficiency, Fe/ α KG oxygenases have attracted keen attention for their application as biocatalysts. Here, we report the biochemical and structural characterizations of the unusually promiscuous and catalytically versatile Fe/ α KG oxygenase SptF, involved in the biosynthesis of fungal meroterpenoid emervaridones. The *in vitro* analysis reveals that SptF catalyzes four continuous oxidation reactions, including hydroxylation, desaturation, epoxidation, and skeletal rearrangement. SptF exhibits extremely broad substrate specificity toward various meroterpenoids, and efficiently produced novel cyclopropane-ring-fused 5/3/5/5/6/6 and 5/3/6/6/6 scaffolds from terretonins. Moreover, SptF also hydroxylates steroids, including androsterone, testosterone, and progesterone, with different regiospecificities. Crystallographic and structure-based mutagenesis studies of SptF reveal the molecular basis of the enzyme reactions, and suggest that the malleability of the loop region contributes to the remarkable substrate promiscuity. SptF exhibits great potential as a promising biocatalyst for oxidation reactions.

Introduction

Oxygenases are widely distributed in nature and catalyze the incorporation of oxygen into various natural products. One of the largest oxygenase groups is the non-heme iron and α -ketoglutarate-dependent (Fe/ α KG) oxygenase superfamily, which is involved in both primary and secondary metabolism.¹ In animals, Fe/ α KG oxygenases play important roles in physiological processes *via* hydroxylation or *N*-demethylation reactions, with substrates including proteins, nucleic acids, and lipids.¹ In plants and microorganisms, in addition to the typical hydroxylation reactions, Fe/ α KG oxygenases catalyze a wide range of chemical transformations in the biosynthesis of natural products,²⁻²⁷ including complex skeletal rearrangement,^{11,12} ring-expansion,¹⁵ and C-C bond formation.²¹⁻²³ Given their intriguing chemical reactions, detailed biochemical, structural, and calculation studies of Fe/ α KG oxygenases have been conducted over the past few decades.^{5,6,8,28-32}

Fe/ α KG oxygenases utilize α KG as a co-substrate and Fe(II) as a cofactor.³³⁻³⁵ Through the oxidative decarboxylation of α KG, a reactive Fe(IV)-oxo intermediate is generated to activate a selective aliphatic C–H bond, which is normally difficult to cleave due to high bond-dissociation energy, and then oxidative products are generated following radical recombination. Unlike other oxidases, including P450 monooxygenases, Fe/ α KG oxygenases do not require any reductase partners or expensive cofactors such as NAD(P)H and FAD/FMN.³⁶ Furthermore, these enzymes exhibit high catalytic efficiencies, and typically have large turnover numbers.^{11,14,37,38} These features of Fe/ α KG oxygenases, along with their catalytic versatility, have attracted considerable attention as a source of biocatalysts.^{36,39}

SptF is an Fe/ α KG oxygenase first discovered in the biosynthesis of the fungal

meroterpenoid emervaridones.⁴⁰ It catalyzes two oxidation steps to form emervaridone C (**3**) from andiconin D (**1**) (Fig. 1, Supplementary Fig. 1). The proposed reaction scheme involves the initiation by a C-11 hydroxylation reaction, followed by water elimination, carbon skeletal rearrangement, and deprotonation reactions to yield **2**. Compound **2** is then accepted by SptF again and undergoes an epoxidation reaction to generate **3**. Thus, SptF accepts two skeletally distinct molecules as substrates and plays a key role in the structural complexification of andiconin-derived meroterpenoids. Although the function of SptF has been investigated by *in vitro* enzymatic reactions, the detailed mechanisms of the dynamic skeletal rearrangement and the accommodation of structurally distinct substrates remain to be elucidated.

In this study, we investigated SptF further by *in vitro* enzyme reactions, crystallization, and structure-based mutagenesis. Surprisingly, SptF is an unusually promiscuous and catalytically versatile enzyme that performs up to four rounds of oxidative reactions with its natural substrate **1**, including hydroxylation, desaturation, epoxidation, and skeletal rearrangements. SptF also catalyzed the formation of ~~unprecedented~~ cyclopropane-ring-fused, highly congested 5/3/5/5/6/6 and 5/3/6/6/6 cyclic skeletons from structurally distinct meroterpenoid terretonins. Moreover, SptF hydroxylated steroids, including androsterone, testosterone, and progesterone, with different regiospecificities. Our crystallographic analyses of SptF in complex with (un)natural substrates, along with the structure-based mutagenesis studies, have revealed the molecular basis for the substrate recognition and suggested that the malleability of the loop region contributes to the remarkable substrate promiscuity.

Results

Substrate promiscuity and catalytic versatility

Emervaridone C (**3**) was previously reported as the final product of SptF from andiconin D (**1**).⁴⁰ However, our careful reexamination of the enzyme reaction revealed that SptF actually produced two more products **4** and **5**, in addition to **2** and **3** (Fig. 1a). Structural analyses established that **4** contains an additional C9 hydroxyl group and C1-C2 epoxide (Fig. 1c, Supplementary Figs. 2, 17-22 and Supplementary Table 7). In contrast, **5** has a new ether bridge between C4' and C2', along with a C5' hydroxyl group (Fig. 1c, Supplementary Figs. 23-28 and Supplementary Table 8). Notably, SptF, in the emervaridone biosynthetic pathway, shares 94% amino acid sequence identity with AndF, which is responsible for the generation of anditomin (**7**) from andilesin C (**6**), in the closely related anditomin pathway (Fig. 1c).³⁷ As expected, SptF also accepted **6** as a substrate to produce **8** and **9** in addition to **7** (Fig. 1b-c, Supplementary Figs. 29-40 and Supplementary Tables 9, 10). Interestingly, **8** and **9** share the same skeleton, but have different stereochemistry at the C3'-C9' epoxide ring. It is remarkable that SptF, a single enzyme, accepts structurally distinct meroterpenoids as substrates and catalyzes sequential oxidation reactions, including desaturation, epoxidation, hydroxylation, ring-opening, and carbon skeletal rearrangements, to generate skeletally diverse products.

Encouraged by the promiscuity of SptF, we further explored its substrate scope by testing six fungal meroterpenoids with different core scaffolds, preandiloid B (**10**), preandiloid C (**12**), terretonin J (**13**), terretonin C (**15**), terretonin A (**17**), and terretonin (**19**) (Supplementary Fig. 3). Interestingly, SptF accepted all of them as substrates, except for **12**. SptF thus installed a hydroxyl group at C1 of **10** to form **11** (Fig. 2a and d, Supplementary Figs. 41-46, and Supplementary Table 11). Notably, SptF oxidized all four terretonins to yield products with 2

mass units less than the original (Fig. 2b-d), and structure determinations by spectroscopic analyses established that **14/18** and **16/20** possess highly congested, cyclopropane-ring-fused 5/3/6/6/6 and 5/3/5/5/6/6 scaffolds, respectively (Supplementary Figs. 2-4, 47-58, and Supplementary Tables 12, 13). The conversion rate of the enzyme reactions ranged from 6% to 51% over 24 hours (Supplementary Table 1).

To further evaluate the utility of SptF as a biocatalyst, we also tested various steroids, including androsterone (**21**), testosterone (**24**), and progesterone (**28**), as substrates. To our surprise, SptF accepted all of them with high conversion rates of 34%-53% over 24 hours, and the structures of the products were determined by NMR and the crystalline-sponge-based X-ray diffraction method (Fig. 3, Supplementary Figs. 5, 59-80, and Supplementary Tables 14-17). SptF thus accepted **21** and installed one hydroxyl group at C5 to yield **22**, and the second hydroxyl group at C9 to form the novel steroid **23**. In contrast, **24** underwent hydroxylation at C9 to yield **25**. Moreover, the hydroxylation and following oxidation at C6 of **24** generated **26**, which is interchangeable with **27** according to the NMR data. Although the structures of products **29-31** from **28** have not been determined yet, due to separation and stability problems, they are also predicted to be hydroxylated and/or oxidized products as judged from their molecular weights (Supplementary Fig. 2). The substrate promiscuity and catalytic versatility of SptF are thus quite exceptional.

Substrate binding mode in the active site

To elucidate the molecular basis of the promiscuous SptF enzyme reactions, we solved the structures of SptF-Fe/ α KG/**1**, SptF-Fe/*N*-oxalylglycine (NOG; a catalytically inactive analogue

of α KG⁴¹)/**6**, SptF S114A-Fe/NOG/**3**, and SptF S114A-Fe/NOG/**15** (Fig. 4a-d, Supplementary Figs. 6 and 7a, and Supplementary Table 2). The structure of SptF shares the double stranded β -helix fold and conserved 2-His-1-Asp facial triad found in other fungal meroterpenoid Fe/ α KG oxygenases, including AndA (PDB: 5ZM3)¹² with RMSD values of 1.4 Å for C α -atoms and 38% amino acid sequence identities .

In the active site, **1** or **6** is located on the opposite side of the facial triad in an almost identical fashion (Fig. 4a, b), indicating that the size of the A-ring does not affect the binding mode. The distance between the iron and the initial reaction site C11 of **1** or **6** is 4.3 Å, which is close enough for hydrogen atom abstraction by the C-H bond activation. The active site hydrophobic residues, Ile63, Phe133, and Ile231, are located within van der Waals distances of the B/C-ring, suggesting that a hydrophobic surface facilitates the substrate binding. Similar to the previously reported AndA,¹² the loop between Trp53 and Asn75 of SptF (Supplementary Fig. 7b) was only observed in complex structures with **1** or **6** (Fig. 4a, b). This lid-like loop region interacts with the A-ring of **1** or **6** *via* only one hydrogen bond with Asn65. In contrast, the D/E-rings of **1** and **6** are tightly fixed by several hydrophilic residues located on the bottom surface of the active site cavity: the carbonyl oxygen of the E-ring interacts with Ser114, while the carbonyl oxygen of the D-ring interacts with the main chains of Leu199 and Thr148 *via* water molecules (Fig. 4a, b). Notably, the complex structure of S114A with intermediate **3** revealed a different ligand binding mode (Fig. 4c): the A-ring of **3** interacts with Lys265' from the adjacent monomer *via* two water molecules, whereas the E-ring interacts with Thr148 *via* one water molecule and forms a hydrogen bond directly with Asn150. The putative reaction positions C9 and C7 are located at distances of 5.3 Å and 4.2 Å from the iron center, respectively.

Interestingly, the unnatural substrate **15** also binds to the active site in a similar manner to that of **3**: The D-ring of **15** forms a hydrogen bond network with Thr148 and Leu199 *via* water molecules, but also forms a hydrogen bond directly with Asn150 (Fig. 4d). Notably, the lid-like loop region was not observed upon the binding of **15**, while the conformations of the other active site residues were almost identical. In this structure, Phe133 and Ile231 are located close to the methyl groups of C8, C10, and C13, and potentially form hydrophobic interactions. The C-6 enol group of **15** is close to the iron with a distance of 3.7 Å, suggesting that the enolic hydrogen atom is abstracted to initiate the cyclopropane ring formation reaction *via* the proposed path c in Supplementary Fig. 8.

Structure-based mutagenesis

To investigate the importance of the active site residues, we performed structure-based mutagenesis. Firstly, the hydrophobic residues Ile63, Phe133, and Ile231, lining the active site cavity, were substituted with Ala (Fig. 5a-c). As a result, the I63A variant abolished the formation of **4** and **5**, and instead generated the new product **32** (Fig. 5d) in addition to **2**. The structure of **32** was determined to be an E-ring opened dicarboxylic acid, which is possibly formed by two rounds of oxidation at the C1' of **1** (Fig. 5b and e, Supplementary Figs. 9, 81-86, and Supplementary Table 18). On the other hand, I63A oxidized **2** to **3**, and converted **3** to small amounts of **4** and **5** (Supplementary Fig. 10), whereas F133A abolished the production of **4** and **5**, but accumulated the early-stage intermediates **2** and **3** (Fig. 5b). The $k_{\text{cat}}/K_{\text{M}}$ value of F133A for **1** decreased 900-fold as compared with the wild-type enzyme (Supplementary Table 3). F133A also converted **1** into the C11-hydroxylated product **33** (Fig. 5d) Supplementary Figs.

87-92 and Supplementary Table 19). In contrast, the substitution of Phe133 with Tyr retained the generation of both **4** and **5** (Fig. 5b). Moreover, the I231A variant showed a similar LC-MS profile to that of F133Y, while affording **32** as the major product. The loss of the *sec*-butyl group increases the active site space around the D-ring of **1**, which could lead to an alternative substrate binding mode. Notably, similar results were observed in the enzyme reactions with the unnatural substrate **15**. The activity for the generation of the cyclopropane ring was abolished in F133A, but retained in F133Y. Moreover, I231A significantly reduced the activity to produce **16** (Fig. 5c), suggesting the similar role of the hydrophobic residues in the recognition of the unnatural substrate **15**.

Secondly, the hydrophilic residues that interact with substrates, including Asn65, Ser114, Thr148, and Asn150, were substituted with alanine (Fig. 5a-c). Due to the insolubility of T148A, the T148S variant was used for the enzyme reaction. As a result, N65A significantly reduced the **4** and **5** formation activity to 20% and accumulated **3** (Fig. 5b, e), suggesting that the hydrogen-bond interaction with Asn65 is moderately important for the binding of **3**. Interestingly, as compared with wild-type SptF, the activities for the formation of **4** and **5** were maintained at 83%, 93%, and 52% in the S114A, T148S, and N150A variants, respectively. The ratio of **4** and **5** was altered to 7:3 in S114A and T148S, as compared with 4:6 in the wild type, and decreased to 1:9 in N150A (Fig. 5e). Our observations suggested that these three residues determine the position of the hydrogen atom abstraction (C9 or C7) by anchoring **3** within the binding pocket, to generate **4** and **5** *via* a hydrogen bond network.

Asn65 in SptF is uniquely substituted with threonine in the closely related AndF (Supplementary Fig. 11). To clarify the function of Asn65, we constructed the SptF N65T

variant, which converted **1** into **5** as the major product, with a slight amount of **4** (Fig. 5b, e). Notably, in the crystal structure of SptF N65T-Fe/NOG/**1**, a hydrogen bond was observed between the C3 ketone at the A-ring and the newly introduced Thr65 (Fig. 5f). These results indicated that a subtle change of the hydrogen bond interactions affects the position of the hydrogen atom for abstraction and determines the product specificity. Interestingly, only the N150A variant lacked the activity for the conversion of **15** into **16** (Fig. 5b, c), suggesting that the direct hydrogen-bonding between Asn150 and **15** is crucial for the substrate binding.

Lid-like loop region and substrate recognition

The mutagenesis studies revealed that the hydrogen bond and hydrophobic interactions in the active site cavity are important for the substrate recognition. In contrast, while Ile63 and Asn65 on the lid-like loop region are important for the formation of **4** and **5**, these residues are not essential for the reaction with the unnatural substrate **15**, suggesting that **15** is mainly recognized by active site residues but not *via* residues on the lid-like loop.

To further investigate the importance of the flexible lid-like loop region, loop-truncated variants of 6 residues (His61-Val66), 9 residues (Lys58-Val66), 16 residues (Asn54-Lys69), and 19 residues (Trp53-Lys71) were constructed (Supplementary Figs. 12 and 13) based on the sequence comparison as well as the structures of SptF in complex with **1**. Since the $\Delta 19$ variant was expressed in inclusion bodies, the $\Delta 6$, $\Delta 9$, and $\Delta 16$ variants were tested for their enzyme reaction activities with **1**, **2**, and **3**. No significant differences in overall protein folding of the truncated variants was confirmed by CD-spectra. Further, crystal structure of SptF $\Delta 9$ variant revealed that it maintains almost identical overall structure, catalytic triad, and NOG binding

mode (Supplementary Fig. 12e-i). Interestingly, although the activities of the truncated variants were dramatically decreased as compared to the wild type, all the loop-truncated variants still consumed 60~94% and 17~77% of **1** and **2** over 24 hours, respectively, and generated many minor products, including the E-ring-opened product **32**, but not **3-5** (Fig. 5b, Supplementary Fig. 14 and Supplementary Table 5). The differences in the activities among $\Delta 6$, $\Delta 9$, and $\Delta 16$ should be due to the subtle folding differences. On the other hand, these variants also showed 26~56% activities toward **3** and generated small amounts of **4** and **5**, as well as other minor compounds (Supplementary Fig. 14 and Supplementary Table 5). Although we could not determine the structures of the minor products due to low yields, these results suggested that the lid-like region is not essential for the reactions with **1-3**, but important for the reaction efficiency and product selectivity.

We also tested the activities of the truncated variants with unnatural substrates, including the fungal meroterpenoids **10** and **13**, and steroids **21**, **24**, and **28**. As a result, all the variants still maintained 24~40% and 56~84% activities toward **10** and **13**, but dramatically decreased activities toward **21**, **24**, and **28** (<20%) (Fig. 5c, Supplementary Fig. 15, and Supplementary Table 5). These observations indicated that the lid-like loop region is crucial for the steroid substrate recognition, but not essential for the meroterpenoid substrates, which is consistent with the absence of the electron density of the loop in the complex structure with **15**. These results strongly suggested that the malleability of the lid-like loop region contributes to the remarkable substrate promiscuity and catalytic versatility of SptF.

Discussions

Substrate promiscuity plays a key role in both biomedical and industrial applications, as it provides a starting point for protein engineering and drug design, enzyme evolution with different reaction/substrate specificities, and chemoenzymatic total synthesis.^{36,39,42-45} Therefore, the identification and investigation of promiscuous enzymes enrich our understanding of how the enzymes obtain such a wide range of diverse functions, and will further guide the rational engineering of these enzymes as biocatalysts for the production of useful molecules with improved biological activities.

The Fe/ α KG oxygenase SptF exhibits unusual promiscuity and catalytic versatility, and catalyzes four sequential oxidation reactions in the biosynthesis of fungal meroterpenoid emervaridones. SptF also catalyzes the formation of cyclopropane-ring-fused, highly congested 5/3/5/5/6/6 and 5/3/6/6/6/6 molecular scaffolds from the structurally distinct meroterpenoid terretonins. Moreover, SptF also hydroxylates steroids, including androsterone, testosterone, and progesterone. It is particularly remarkable that SptF accepts both 3-keto- Δ^4 - and 3-hydroxy-steroids with similar efficiencies to catalyze the hydroxylation at the α -face of different positions and generate a series of hydroxylated steroids. Hydroxylation of steroids has been previously reported for cytochrome P450 enzymes,⁴⁶ including the human CYP3A subfamily and the engineered bacterial P450 BM3, which catalyzes the hydroxylation of testosterone, a 3-keto- Δ^4 -steroid, at the β -face of the C3 and C15 positions.⁴⁷ However, in spite of intensive engineering efforts, P450 BM3 does not accept 3-hydroxy-steroids.⁴⁸ Thus, SptF exhibits great potential as a promising biocatalyst for the oxidation of various natural products with important pharmacological activities.

Based on our structure-function analyses, we propose enzyme reaction mechanisms for

the consecutive oxidation reactions in the emervaridone biosynthetic pathways, as follows (Fig. 6). Hydrogen abstraction at C11 of **1** by the ferryl species first generates a radical, which undergoes recombination to cleave the C-C bond between C4' and C3' and bridge C4' and C11. Subsequently, a double bond is formed to yield **2**, which undergoes epoxidation to generate **3**. Compound **3** then serves as the branch point intermediate to produce **4** and **5**. In path a, hydrogen atom abstraction at C9 and subsequent hydroxyl group rebound result in emeridone F, which then undergoes A-ring epoxidation to generate **4**. In path b, hydrogen abstraction at C7 induces the radical recombination, which cleaves the C8-C2' bond to construct a double bond between C7 and C8 on the B-ring. The hydroxyl rebound at C2' forms **IM1**, which is followed by lactone formation and C4'-C5' bond cleavage to yield **IM2**, in a similar manner to the formation of emeridone A.⁴⁰ In the last step, hydrogen abstraction at C7' and relocation of the radical to C5' complete the formation of the final product **5**.

The multifunctional catalytic activities largely stem from the unusual substrate promiscuity of SptF, which can accommodate a series of structurally different molecules in the active site for further oxidation reactions. The crystallographic and mutagenesis investigations have provided the structural basis for the promiscuity of SptF. In particular, SptF is unique since the lid-like loop region interacts with the natural substrate **1** *via* only one hydrogen bond with Asn65, which is clearly distinct from other fungal meroterpenoid Fe/ α KG oxygenases, including AndA and PrhA.^{11,12} In AndA and PrhA, the residues on the lid-like loop tightly recognize the D/E rings of substrates *via* a hydrogen bond network, while the A/B-rings loosely interact with the active site (Supplementary Fig. 16). These relaxed interactions and high malleability of the lid-like loop region in SptF are thought to contribute to its remarkable

substrate promiscuity.

Truncations of the lid-like loop region demonstrated that the loop is not essential for the activity toward the natural and unnatural meroterpenoid substrates, but important for binding the steroid substrates and the product selectivity from natural substrates. The substrate is thought to be mainly recognized by the active site residues deep inside the cavity, which is then covered by the lid-like loop to further support the substrate loading and ensure the precise and efficient positioning for the C-H bond activation by the ferryl species. This whole process would be rigidly controlled for the natural substrate. In contrast, the flexible loop may interact loosely with the unnatural substrates, but still retains the efficiency and product specificity of the enzyme reactions. In particular, it would be difficult to tightly fix smaller steroids with a few functional groups in the active site. These molecules would probably loosely bind to the enzyme in several different conformations, and the enzymatic C-H bond activation could only take place when the substrate is properly positioned for the interactions with the lid-like loop region. This would be the reason why we could not obtain the steroid-bound complex structure, despite our soaking experiments. Therefore, understanding how the lid-like loop region interacts with different substrates would provide key information to manipulate the SptF enzyme reaction.

In conclusion, we characterized the exceptionally promiscuous and catalytically versatile Fe/ α KG oxygenase SptF, which catalyzes a remarkable repertoire of oxidation reactions of meroterpenoid and steroid molecules. We believe these unique features of SptF will serve as a platform for future efforts to control the reactivity and selectivity of the α KG oxygenases. SptF thus exhibits great potential as a promising biocatalyst for oxidation reactions.

Methods

General.

Oligonucleotide primers, listed in Supplementary Table 6, and DNA sequencing were ordered from Eurofins Genomics. The restriction enzymes *Nde*I and *Hind*III, and PrimeSTAR GXL DNA polymerase were purchased from Takara Bio Inc. Solvents and chemicals were purchased from Wako Chemicals, Ltd. (Tokyo, Japan), Merck KGaA Ltd. (Darmstadt, Germany), and Hampton Research (CA, USA), unless noted otherwise. PCR was performed using a TaKaRa PCR Thermal Cycler Dice® Gradient (Takara Bio Inc.). The NMR spectra of compounds were recorded on Bruker AVANCE III HD 800 MHz, AVANCE III HD 900 MHz (Bruker MA, USA), ECX-500 MHz, and EXZ-500 MHz (JEOL Ltd., Tokyo, Japan) spectrometers.

Construction of SptF expression plasmid.

To obtain stable and pure proteins for *in vitro* enzyme reactions and crystallization experiments, expression plasmids for SptF lacking three residues, predicted to be flexible by XtalPred server,⁴⁹ were prepared as follows. The primer pair SptF-N3-F/SptF-N3-R was used for amplification of the SptF DNA fragment from the previously reported pET-28a(+)-sptF⁴⁰ plasmid (see Supplementary Table 6 for primer sequence information). SptF PCR products were purified and ligated into the *Nde*I/*Hind*III-digested pET-22b (+) vector, using an In-Fusion® HD Cloning Kit (TaKaRa Clontech). The identities of the resulting vectors were confirmed by DNA sequencing. *Escherichia coli* DH5α was used for plasmid construction, and *E. coli* Rosetta™2(DE3)pLysS (Novagen) was used for the preparation of recombinant proteins.

Construction of plasmids for expression of SptF variants.

The primers used for the construction of plasmids for site-directed mutagenesis studies are listed in Supplementary Table 6. The plasmid for the expression of wild-type SptF was used as the template for PCR-based site-directed mutagenesis, which was performed with a QuikChange Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's protocol. For the construction of loop-truncation mutants, the primer pairs listed in Supplementary Table 6 were used for the amplification of SptF DNA fragments from pET-28a(+)-sptF⁴⁰ plasmids. SptF PCR products were purified and ligated into the *NdeI/HindIII*-digested pET-22b (+) vector, using an In-Fusion® HD Cloning Kit (TaKaRa Clontech). The identities of the resulting plasmids were confirmed by DNA sequencing.

Production and purification of SptF and its variants.

In this study, the recombinant proteins used for *in vitro* assays were prepared as follows: Plasmids expressing SptF or its variants were transformed into *E. coli* RosettaTM2(DE3)pLysS (Novagen). The resulting strains were cultured in LB medium supplemented with 34 mg/L chloramphenicol and 50 mg/L ampicillin sodium, with shaking at 37 °C, 160 rpm. When the OD₆₀₀ reached 0.6, the cell cultures were cooled down on ice for 30 min, and then IPTG (0.2 mM) was added to induce the target protein expression at 16 °C, 160 rpm. After 18h of post-induction incubation, cells were harvested by centrifugation at 6,000 rpm for 15 min and suspended in lysis buffer, containing 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 10 mM imidazole, and 5% glycerol. The cell suspension was sonicated for 4 minutes on ice. After

removal of the cell debris by centrifugation at 15,000 rpm for 30 min, the supernatant was mixed with 1 mL Ni-NTA agarose resin and loaded onto a gravity flow column. Unbound proteins were removed with 50 mL lysis buffer containing 25 mM imidazole, and then the His-tagged protein was eluted with lysis buffer containing 300 mM imidazole. For the *in vitro* assay, the protein eluate was concentrated to 10 mg/mL after the removal of imidazole, using a 30 kDa Amicon[®] Ultra-15 filtration unit (Millipore). For crystallization, the eluate from the Ni-NTA agarose was applied to a HiLoad 16/60 Superdex 200 preppacked gel filtration column (4 °C, GE Healthcare), and eluted with a solution containing 20 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 1 mM dithiothreitol. The resulting eluate was concentrated to 10 mg/mL, using an Amicon Ultra-4 (MWCO: 30 kDa) filter at 4 °C. The purity of the purified proteins was monitored by SDS-PAGE (Supplementary Fig. 13), and the protein concentrations were determined with a SimpliNano microvolume spectrophotometer.

***In vitro* assay of SptF and its variants.**

Substrates used for *in vitro* enzyme reactions, including **1, 6, 10, 12, 13, 15, 17** and **19**, were prepared according to previous studies,^{12,37,40,50} and the proteins were prepared as described above. All reactions in this study were performed following the previous studies unless noted otherwise.⁴⁰ Reactions were performed on a 50-μL scale with 50 mM PIPES (pH 7.5), 0.2 mM FeSO₄, 5 mM α-ketoglutarate, 4 mM ascorbate, 100 μM of substrate, and 15 μM of wild-type or mutant SptF. The reaction mixture without protein was used as the negative control. Enzyme reactions, except for steady-state enzyme kinetics, were incubated at 30 °C for 1 h or overnight and quenched by adding 50-μL methanol. After centrifugation and filtration, the reaction

mixtures were analyzed by LC-MS. For comparisons of the substrate consumption between truncation variants and wild type, the reactions were performed on a 50- μ L scale with 50 mM PIPES (pH 7.5), 0.2 mM FeSO₄, 5 mM α -ketoglutarate, 4 mM ascorbate, 100 μ M of substrate, and 40 μ M of enzyme.

LC-MS methods used for analyzing *in vitro* reactions.

All samples from *in vitro* assays were analyzed by a Bruker Compact QqTOF mass spectrometer (Bruker) coupled with an LC-20AD ultra-high-performance liquid chromatography (UHPLC) system (Shimadzu). MS data were recorded in the positive electrospray ionization mode. Compounds were separated on a COSMOSIL 5C18-MS-II Packed Column (2.0 mm $\times$ 100 mm, 5 μ m). All reaction mixtures were separated using H₂O supplemented with 0.1% v/v formic acid as solvent A and acetonitrile as solvent B. The gradient for the separation of compounds in reaction mixtures with **1**, **10**, **12**, **17**, or **19** was 20% to 45% B over 20 min, then increase to 100% B in 1 min, and hold for 4 min at 40 °C with a flow rate of 0.2 mL/min. Reaction mixtures with **6** are shown here: 20% to 60% B over 17 min, then increase to 100% B in 0.1 min, and hold for 2.9 min at 40 °C with a flow rate of 0.15 mL/min. Reaction mixtures with **13** or **15** are shown here: 20% to 100% B over 20 min, and hold for 5 min at 40 °C with a flow rate of 0.15 mL/min. Reaction mixtures with **21**, **24**, or **28** are as follows: 10% to 100% B over 20 min, and hold for 5 min at 35 °C with a flow rate of 0.15 mL/min.

Crystallization and structure determination.

Crystals of SptF lacking three residues, as well as the N65T and S114A variants, were obtained after 1 day at 10 °C, except for $\Delta 9$ variant which was obtained at 20 °C, by using the sitting-drop vapor-diffusion method with the following reservoir solutions: SptF-apo: 0.27 M magnesium chloride, 21% w/v PEG 3350, 1.5% trimethylamine N-oxide dihydrate. SptF N65T: 0.2 M magnesium formate dehydrate, 20% w/v PEG 3350. SptF S114A: 0.27 M magnesium chloride, 20% w/v PEG 3350, 0.2 M sodium chloride, 1.5% trimethylamine N-oxide dihydrate. SptF $\Delta 9$ variant: 0.2 M sodium formate (pH 7.0), 20% w/v PEG 3350. The protein concentrations used for the crystallizations of SptF, N65T, S114A, and $\Delta 9$ variants were 7.5 mg/L, 10 mg/L, 7.5 mg/L, and 5 mg/L, respectively. Since S114A can produce **4** and **5**, and has better diffraction as compared with wild-type SptF, we employed it for the soaking experiments of **3** and **15**. The complex structures were prepared by incubating SptF, N65T, or S114A crystals at 10 °C for 12 h with 5 mM of compound, 5 mM of α KG/NOG, and 100 μ M of Fe in the crystallization drop. When α KG was used, the soaking experiment was performed in an anaerobic chamber.

The crystals were transferred into the cryoprotectant solution (reservoir solution with 25% (v/v) glycerol), and then flash cooled at -173 °C in a nitrogen-gas stream. The X-ray diffraction data sets were collected at BL-1A (Photon Factory, Tsukuba, Japan), using a beam wavelength of 1.1 Å. The diffraction data sets for SptF were processed and scaled using the XDS program package⁵¹ and Aimless.⁵² The initial phase of the SptF structure was determined by molecular replacement, using AndA (PDB ID: 5ZM4) as the search model. Molecular replacement was performed with Phaser in PHENIX.^{53,54} The initial phase was further calculated with AutoBuild in PHENIX.⁵⁵ The SptF structures were modified manually with Coot⁵⁶ and refined with

PHENIX.refine.⁵⁷ The cif parameters of the ligands for the energy minimization calculations were obtained by using the PRODRG server.⁵⁸ The final crystal data and intensity statistics are summarized in Supplementary Table 2.

The Ramachandran statistics are as follows: 98.3% favored, 1.7% allowed for SptF apo, 97.7% favored, 2.3 % allowed for SptF complexed with **1** and α KG, 98.5% favored, 1.5% allowed for SptF complexed with **6** and NOG, 98.5% favored, 1.5% allowed for SptF Δ 9 variant complexed with NOG, 97.7% favored, 2.3% allowed for SptF S114A complexed with **3** and NOG, and 98.3% favored, 2.7 % allowed for SptF S114A complexed with **15** and NOG. All crystallographic figures were prepared with PyMOL (DeLano Scientific, <http://www.pymol.org>).

Steady-state enzyme kinetics.

Each assay was performed in a total volume of 50 μ L, containing 50 mM PIPES buffer (pH 7.5), 4 mM ascorbate, 200 μ M FeSO₄, and 1 mM α KG, at 30 °C for 5 min, and quenched by the addition of 50 μ l methanol. Compounds **1** (0.5, 1.0, 2.5, 5.0, 10.0, 15.0, 20.0, and 30.0 μ M) and **6** (0.5, 1.0, 2.5, 5.0, 10.0, 15.0, 20.0, and 50.0 μ M) were used to determine the kinetic parameters of SptF. To obtain the kinetic parameters of F133A, **1** (5.0, 20.0, 50.0, 100.0, 150.0, 200.0, and 300.0 μ M) and **6** (5.0, 20.0, 50.0, 75.0, 100.0, 150.0, and 200.0 μ M) were used. For substrate **1**, 0.05 and 5.0 μ M purified SptF or F133A were used, respectively. For substrate **6**, 0.01 and 2.5 μ M purified SptF or F133A were used, respectively. Substrate consumption was analyzed with LC-MS, by calculating the total peak areas of the substrate. Each reaction was performed in triplicate. GraphPad Prism (GraphPad Prism Software Inc., San Diego, CA) was

used for statistics data analysis (Supplementary Table 3).

Enzyme reaction products from 1, 6, 10, 13, and 15.

To isolate the products from **1**, **6**, **10**, **13**, and **15** for structure determination, *in vitro* enzymatic reactions were performed as described above in a total volume of 100 mL, divided into 0.5 mL portions in 2 mL-tubes. After an overnight incubation at 30 °C, the reactions were quenched and extracted twice with an equal volume of ethyl acetate. After the ethyl acetate removal, 1 mL methanol was added to dissolve the products for separation by semi-preparative HPLC, which was performed on a Shimadzu LC20-AD HPLC system.

Compounds **4**, **5** and **33** were purified from large scale enzymatic reactions as mentioned above, using wild-type SptF and the F133A variant with **1**. Compounds **8** and **9** were prepared using wild-type SptF from **6**. These products were purified on an XSelect HSS T3 OBD Prep column (10.0 mm × 250 mm, 5 µm) using 50% methanol supplemented with 0.1% v/v formic acid at 40 °C. The injection volume was 0.1 mL and the flow rate was 2.5 mL/min. Finally, 0.3-0.4 mg of compounds **4**, **5**, **8**, **9**, and **33** were obtained.

Compound **32** was purified from large scale enzymatic reactions using I231A, on an XSelect HSS T3 OBD Prep column (10.0 mm × 250 mm, 5 µm) with H₂O supplemented with 0.1% v/v formic acid as solvent A and methanol as solvent B, at 40 °C. The gradient for separation was 46% B for 40 min, and 56% B for 60 min. The injection volume was 0.1 mL and the flow rate was 2.5 mL/min. The purification yielded 0.3 mg of compound **32**.

Compound **11** was purified with an XSelect HSS T3 OBD Prep column (10.0 mm × 250 mm, 5 µm), using 45% acetonitrile supplemented with 0.1% v/v formic acid, at 40 °C. The injection

volume was 0.15 mL and the flow rate was 3.0 mL/min, and 0.9 mg of compound **11** was obtained.

Compound **14** and **15** were purified with an XSelect HSS T3 OBD Prep column (10.0 mm × 250 mm, 5 μm) using 47% acetonitrile supplemented with 0.1% v/v formic acid at 40 °C, respectively. Inject volume is 0.15 mL and flow rate is 3.0 mL/min. 0.3 and 0.4 mg compound **14** and **15** was obtained, respectively.

Compound **16** was purified with a COSMOSIL πNAP Packed column (10.0 mm × 250 mm, 5 μm), using 50% acetonitrile supplemented with 0.1% v/v formic acid, at 40 °C. The injection volume was 0.15 mL and the flow rate was 3.0 mL/min, and 0.9 mg compound **16** was obtained.

Enzyme reaction products from 21 and 24.

Reactions were performed in a total volume of 100 mL, with 100 mM HEPES (pH 7.5), 0.2 mM FeSO₄, 4 mM α-ketoglutarate, 2 mM ascorbate, 1.6 mM of substrate, 10% glycerol, and 50 μM of SptF, at 4 °C for 56 h. Products were separated by normal-phase silica gel chromatography. The elution conditions were as follows: CHCl₃ : MeOH = 100 : 0.5 for 540 mL; then CHCl₃ : MeOH = 100 : 1 for 90 mL, and CHCl₃ : MeOH = 100 : 1.5 for 90 mL. Finally, 2.0 mg and 1.3 mg of compounds **22** and **23** were obtained from **21**, respectively. From **24**, 2.2 mg and 1.3 mg of **25** and **26/27** were obtained, respectively.

Structure determination of 25 by crystalline sponge method.

A single crystal of [(ZnCl₂)₃(tpt)₂(cyclohexane)_x] (tpt = 2,4,6-tris-(4-pyridyl)-1,3,5-triazine)⁵⁹ was used as the crystalline sponge in this study. The crystalline sponge (typically approx. 200

$\times 150 \times 50 \mu\text{m}^3$) was placed in a vial and immersed in a small amount of cyclohexane (50 μL). A 1,2-dichloroethane solution (5 μL) of the compound (1 $\mu\text{g}/\mu\text{L}$, 5 μg) was added to the vial. The vials were equipped with a syringe needle for slow solvent evaporation. Guest soaking was carried out at 50 °C for 1 d. After confirming solvent evaporation, the X-ray diffraction data sets were collected at BL02B1 (SPring-8, Hyogo, Japan), using a beam wavelength of 0.4115 Å. The sample crystal was cooled to 100 K, using a cold nitrogen stream. The collected data were integrated, corrected, and scaled using the program CrysAlisPro, ver. 1.171.38.46 (Rigaku Oxford Diffraction, 2015). All crystal structures were solved using the program SHELXT ver. 2014/5⁶⁰ and refined with the program SHELXL ver. 2018/1.⁶¹ The crystal structure of the **25**/crystalline sponge complex was solved and three independent molecules of **25** per asymmetric unit were clearly observed. All non-hydrogen atoms were refined anisotropically. All hydrogen atoms were generated using the proper HFIX command and refined isotropically using the riding model. All non-hydrogen atoms in the framework were refined without restraints and constraints. Guest molecules were refined with restraints on the atomic displacement parameters (SIMU and RIGU). Three independent guest molecules were characterized as geometrically equal molecules by restraint (SAME). The populations of the three independent guest molecules per asymmetric unit were estimated to be 69%, 56% and 31%, according to the structure refinement. The solvent cyclohexane molecules suggested in the electron density map were refined using some restraints (DFIX, DANG, SIMU, and ISOR), because the corresponding electron densities were obscured due to their thermal motion and disorder. The absolute configuration of **25** was thus established as 8*S*, 9*R*, 10*S*, 13*S*, 14*S*, 17*S*, based on the Flack [0.122 (10)] parameters. Crystallographic data are as follows. Refined

formula: $C_{118.33}H_{125.04}Cl_{12}N_{24}O_{4.68}Zn_6$, formula weight (M_r): 2775.94, crystal system: monoclinic, space group: $C2$, $Z = 4$, $R_{int} = 0.0634$. Lattice parameters, R -factor on $F2 > 2\sigma(F2)$, and weighted R -factor are as follows: $a = 32.8346(8) \text{ \AA}$, $b = 14.4445(2) \text{ \AA}$, $c = 30.7569(8) \text{ \AA}$, $\beta = 100.966(2)^\circ$, $V = 14321.0(6) \text{ \AA}^3$, $R = 0.0977$, $wR = 0.3103$, and $S = 1.197$. The calculated density is 1.287. The linear absorption coefficient (μ) is 0.288. The Flack parameter, determined using 12,777 intensity quotients (Parsons' method), is 0.122(10).

Data availability.

The data supporting the findings of this study are available within the article and its Supplementary Information Files or from the corresponding authors on reasonable request. Protein Data Bank (PDB): The coordinates and the structure factor amplitudes for the apo structures of SptF wild type apo, SptF wild type in complex with **1** and α -KG, SptF wild type in complex with **6** and NOG, SptF $\Delta 9$ variant in complex with NOG, SptF S114A variant in complex with **3** and NOG, SptF S114A variant in complex with **15** and NOG were deposited under accession code 7EYR, 7EYS, 7EYT, 7FCB, 7EYV, and 7EYW, respectively. The crystallographic information file (CIF) for crystal structure of **25** was submitted to The Cambridge Crystallographic Data Centre (CCDC), under reference number 2090716.

References

- 1 Aik, W., McDonough, M. A., Thalhammer, A., Chowdhury, R. & Schofield, C. J. Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. *Curr. Opin. Struct. Biol.* **22**, 691-700 (2012).
- 2 Clifton, I. J. *et al.* Structural studies on 2-oxoglutarate oxygenases and related double-stranded β -helix fold proteins. *J. Inorg. Biochem.* **100**, 644-669 (2006).
- 3 Hausinger, R. P. Fe (II)/ α -ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol. Biol.* **39**, 21-68 (2004).
- 4 Schofield, C. & Hausinger, R. 2-Oxoglutarate-dependent oxygenases. (Royal Society of Chemistry, 2015).
- 5 Bräuer, A., Beck, P., Hintermann, L. & Groll, M. Structure of the Dioxygenase AsqJ: Mechanistic Insights into a One - Pot Multistep Quinolone Antibiotic Biosynthesis. *Angew. Chem. Int. Ed.* **55**, 422-426 (2016).
- 6 Mader, S. L., Bräuer, A., Groll, M. & Kaila, V. R. Catalytic mechanism and molecular engineering of quinolone biosynthesis in dioxygenase AsqJ. *Nat. Commun.* **9**, 1168 (2018).
- 7 Liao, H. J. *et al.* Insights into the desaturation of cyclopeptin and its C3 epimer catalyzed by a non - heme iron enzyme: Structural characterization and mechanism Elucidation. *Angew. Chem. Int. Ed.* **57**, 1831-1835 (2018).
- 8 Li, J. *et al.* Epoxidation catalyzed by the nonheme iron(II)- and 2-oxoglutarate-dependent oxygenase, AsqJ: Mechanistic elucidation of oxygen atom transfer by a ferryl intermediate. *J. Am. Chem. Soc.* **142**, 6268-6284 (2020).

- 9 Li, J., van Belkum, M. J. & Vederas, J. C. Functional characterization of recombinant hyoscyamine 6 β -hydroxylase from *Atropa belladonna*. *Bioorg. Med. Chem.* **20**, 4356-4363 (2012).
- 10 Chang, W.-c., Li, J., Lee, J. L., Cronican, A. A. & Guo, Y. Mechanistic investigation of a non-heme iron enzyme catalyzed epoxidation in (–)-4'-methoxycyclopenin biosynthesis. *J. Am. Chem. Soc.* **138**, 10390-10393 (2016).
- 11 Nakashima, Y. *et al.* Structure function and engineering of multifunctional non-heme iron dependent oxygenases in fungal meroterpenoid biosynthesis. *Nat. Commun.* **9**, 104 (2018).
- 12 Nakashima, Y. *et al.* Structural and computational bases for dramatic skeletal rearrangement in anditomin biosynthesis. *J. Am. Chem. Soc.* **140**, 9743-9750 (2018).
- 13 Steffan, N., Grundmann, A., Afiyatullo, S., Ruan, H. & Li, S.-M. FtmOx1, a non-heme Fe (II) and α -ketoglutarate-dependent dioxygenase, catalyses the endoperoxide formation of verruculogen in *Aspergillus fumigatus*. *Org. Biomol. Chem.* **7**, 4082-4087 (2009).
- 14 Matsuda, Y. *et al.* Novofumigatonin biosynthesis involves a non-heme iron-dependent endoperoxide isomerase for orthoester formation. *Nat. Commun.* **9**, 2587 (2018).
- 15 Doyon, T. J. *et al.* Radical Tropolone Biosynthesis. *ChemRxiv* 10.26434/chemrxiv.12780044.v1 (2020).
- 16 McCulloch, K. M. *et al.* Oxidative cyclizations in orthosomycin biosynthesis expand the known chemistry of an oxygenase superfamily. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11547-11552 (2015).

- 17 Li, S. *et al.* Direct genetic and enzymatic evidence for oxidative cyclization in hygromycin B biosynthesis. *ACS Chem. Biol.* **13**, 2203-2210 (2018).
- 18 Harris, N. C. *et al.* Isonitrile Formation by a Non-heme iron (II)-dependent oxidase/decarboxylase. *Angew. Chem. Int. Ed.* **57**, 9707-9710 (2018).
- 19 Chen, T. Y. *et al.* Pathway from N-alkylglycine to alkylisonitrile catalyzed by iron (II) and 2-oxoglutarate-dependent oxygenases. *Angew. Chem. Int. Ed.* **132**, 7437-7441 (2020).
- 20 Li, H. & Liu, Y. Mechanistic investigation of isonitrile formation catalyzed by the nonheme iron/ α -KG-dependent decarboxylase (ScoE). *ACS Catalysis* **10**, 2942-2957 (2020).
- 21 Siitonen, V. *et al.* Divergent non-heme iron enzymes in the nogalamycin biosynthetic pathway. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 5251-5256 (2016).
- 22 Havemann, J., Vogel, D., Loll, B. & Keller, U. Cyclolization of D-lysergic acid alkaloid peptides. *Chem. Biol.* **21**, 146-155 (2014).
- 23 Jakubczyk, D., Caputi, L., Stevenson, C. E., Lawson, D. M. & O'Connor, S. E. Structural characterization of EasH (*Aspergillus japonicus*)—an oxidase involved in cycloclavine biosynthesis. *Chem. Commun.* **52**, 14306-14309 (2016).
- 24 Chang, W.-c. *et al.* Mechanism of the C5 stereoinversion reaction in the biosynthesis of carbapenem antibiotics. *Science* **343**, 1140-1144 (2014).
- 25 Gao, S.-S., Naowarojna, N., Cheng, R., Liu, X. & Liu, P. Recent examples of α -ketoglutarate-dependent mononuclear non-haem iron enzymes in natural product biosyntheses. *Nat. Prod. Rep.* **35**, 792-837 (2018).

- 26 Weichold, V., Milbredt, D. & van Pée, K. H. Specific enzymatic halogenation—from the discovery of halogenated enzymes to their applications in vitro and in vivo. *Angew. Chem. Int. Ed.* **55**, 6374-6389 (2016).
- 27 Abe, I., Bunno, R., Awakawa, T. & Mori, T. Aziridine formation by a Fe (II)/ - ketoglutarate dependent oxygenase and 2-aminoisobutyrate biosynthesis in fungi. *Angew. Chem. Int. Ed.* (2021). doi.org/10.1002/anie.202104644
- 28 Roach, P. L. *et al.* Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. *Nature* **375**, 700-704 (1995).
- 29 Roach, P. L. *et al.* Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. *Nature* **387**, 827-830 (1997).
- 30 Hamed, R. B. *et al.* The enzymes of β -lactam biosynthesis. *Nat. Prod. Rep.* **30**, 21-107 (2013).
- 31 Markolovic, S. *et al.* Structure–function relationships of human JmjC oxygenases—demethylases versus hydroxylases. *Curr. Opin. Struct. Biol.* **41**, 62-72 (2016).
- 32 Einsiedler, M., Jamieson, C. S., Maskeri, M., Houk, K. & Gulder, T. A. M. Fungal dioxygenase AsqJ is promiscuous and bimodal: substrate - directed formation of quinolones versus quinazolinones. *Angew. Chem. Int. Ed.* **60**, 8297-8302 (2021)
- 33 McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J. & Schofield, C. J. Structural studies on human 2-oxoglutarate dependent oxygenases. *Curr. Opin. Struct. Biol.* **20**, 659-672 (2010).
- 34 Loenarz, C. & Schofield, C. J. Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends Biochem.*

- Sci* **36**, 7-18 (2011).
- 35 Schofield, C. J. & Zhang, Z. Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Curr. Opin. Struct. Biol.* **9**, 722-731 (1999).
 - 36 Zwick, C. R. & Renata, H. Harnessing the biocatalytic potential of iron- and α -ketoglutarate-dependent dioxygenases in natural product total synthesis. *Nat. Prod. Rep.* **37**, 1065-1079 (2020).
 - 37 Matsuda, Y., Wakimoto, T., Mori, T., Awakawa, T. & Abe, I. Complete biosynthetic pathway of anditomin: Nature's sophisticated synthetic route to a complex fungal meroterpenoid. *J. Am. Chem. Soc.* **136**, 15326-15336 (2014).
 - 38 Nakamura, H., Matsuda, Y. & Abe, I. Unique chemistry of non-heme iron enzymes in fungal biosynthetic pathways. *Nat. Prod. Rep.* **35**, 633-645 (2018).
 - 39 Li, J., Li, F., King-Smith, E. & Renata, H. Merging chemoenzymatic and radical-based retrosynthetic logic for rapid and modular synthesis of oxidized meroterpenoids. *Nat. Chem.* **12**, 173-179 (2020).
 - 40 Bai, T. *et al.* Structural diversification of andiconin-derived natural products by α -ketoglutarate-dependent dioxygenases. *Org. Lett.* **22**, 4311-4315 (2020).
 - 41 Kalliri, E., Grzyska, P. K. & Hausinger, R. P. Kinetic and spectroscopic investigation of CoII, NiII, and N-oxalylglycine inhibition of the FeII/ α -ketoglutarate dioxygenase, TauD. *Biochem. Biophys. Res. Commun.* **338**, 191-197 (2005).
 - 42 Nobeli, I., Favia, A. D. & Thornton, J. M. Protein promiscuity and its implications for biotechnology. *Nat. Biotechnol.* **27**, 157-167 (2009).
 - 43 Tawfik, O. K. & S, D. Enzyme promiscuity: a mechanistic and evolutionary perspective.

- Annu. Rev. Biochem.* **79**, 471-505 (2010).
- 44 Zhang, X. *et al.* Divergent synthesis of complex diterpenes through a hybrid oxidative approach. *Science* **369**, 799-806 (2020).
- 45 Renata, H. Exploration of iron-and α -ketoglutarate-dependent dioxygenases as practical biocatalysts in natural product synthesis. *Synlett* **32**, 775-784 (2021).
- 46 Agematu, H. *et al.* Hydroxylation of testosterone by bacterial cytochromes P450 using the *Escherichia coli* expression system. *Biosci. Biotechnol. Biochem.* **70**, 307-311 (2006).
- 47 Kille, S., Zilly, F. E., Acevedo, J. P. & Reetz, M. T. Regio-and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution. *Nat. Chem.* **3**, 738-743 (2011).
- 48 Liu, X., Wang, Z.-B., Wang, Y.-N. & Kong, J.-Q. Probing Steroidal Substrate Specificity of Cytochrome P450 BM3 Variants. *Molecules* **21**, 760 (2016).
- 49 Slabinski, L. *et al.* XtalPred: a web server for prediction of protein crystallizability. *Bioinformatics* **23**, 3403-3405, (2007).
- 50 Matsuda, Y., Iwabuchi, T., Wakimoto, T., Awakawa, T. & Abe, I. Uncovering the unusual D-ring construction in terretinin biosynthesis by collaboration of a multifunctional cytochrome P450 and a unique isomerase. *J. Am. Chem. Soc.* **137**, 3393-3401 (2015).
- 51 Kabsch, W. Xds. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **66**, 125-132 (2010).
- 52 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **69**, 1204-1214 (2013).

- 53 Ness, S. R., de Graaff, R. A., Abrahams, J. P. & Pannu, N. S. CRANK: new methods for automated macromolecular crystal structure solution. *Structure* **12**, 1753-1761 (2004).
- 54 McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674 (2007).
- 55 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **66**, 213-221 (2010).
- 56 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **60**, 2126-2132 (2004).
- 57 Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix. refine. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **68**, 352-367 (2012).
- 58 Schüttelkopf, A. W. & Van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **60**, 1355-1363 (2004).
59. Sakurai, F., Khutia, A., Kikuchi, T. & Fujita, M. X-ray Structure Analysis of N-Containing Nucleophilic Compounds by the Crystalline Sponge Method. *Chem. Eur. J.* **23**, 15035–15040 (2017).
60. G. M. Sheldrick, *Acta Crystallogr. Sect. A. Biol. Crystallogr.* **71**, 3-8 (2015).
61. G. M. Sheldrick, *Acta Crystallogr. Sect. C. Biol. Crystallogr.* **71**, 3-8 (2015).

Acknowledgements

We thank Drs. Takayoshi Awakawa and Richiro Ushimaru for critical reading of the manuscript. We also thank Dr. Kuniyoshi Sugimoto for synchrotron radiation experiments performed at the BL02B1 of the SPring-8. The synchrotron radiation experiments were performed at the BL-1A of the Photon Factory. NMR data were measured at RIKEN Yokohama NMR Facility. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan (JSPS KAKENHI Grant Number JP16H06443, JP19K15703, JP20H00490, JP20KK0173, and JP20K22700), the New Energy and Industrial Technology Development Organization (NEDO, Grant Number JPNP20011), the PRESTO program from Japan Science and Technology Agency, Astellas Foundation for Research on Metabolic Disorders, Takeda Science Foundation, Noda Institute for Scientific Research, and Partnership to Realize Innovative Seeds and Medicines from Sumitomo Dainippon Pharma Co., Ltd. H.T. is a recipient of the JSPS Postdoctoral Fellowship for Foreign Researchers (ID No. P18404).

Author Contributions

T.M. and I.A. designed the experiments. H.T., T.M., H.C., and S.Lyu. performed *in vitro* analysis and crystallization experiments. A.N., and S.Lee. determined the structure of product using crystalline sponge method. H.Z., and F.H., measured NMR. H.T., T.M., H.Z., and I.A. analyzed the data. H.T., T.M., and I.A. wrote the paper.

Competing financial interests

The authors declare no competing financial interests.

Figures

Figure 1. Characterization of SptF function *in vitro* using andiconin D (1) and andilesin C (6) as substrates. a, LC-MS EICs of products from the *in vitro* assay using 1 as the substrate, and b, using 6 as the substrate (the + m/z value for each compound is shown). c, Biosynthesis and structures of SptF and AndF products, using 1 and 6 as substrates *in vitro*. Note that the products 4, 5, 8, and 9 were isolated from large scale enzymatic reactions, and their structures were determined by NMR.

Figure 2. Substrate scope investigation of SptF. a-c, LC-MS EICs for *in vitro* assays using preandiloid B (10), preandiloid C (12), terretonin J (13), terretonin C (15), terretonin A (17), and terretonin (19). The + m/z value used for each compound is shown. d, Structures of substrates and corresponding products. Products 11, 14, and 16 were isolated from large scale enzymatic reactions, and their structures were determined by NMR. Note that the structures of 18 and 20 were deduced according to the HRMS, UV pattern, and reaction type (Supplementary Fig. 4).

Figure 3. Oxidation of human steroid hormones by SptF. **a**, LC-MS EICs for *in vitro* assays using androsterone (21), testosterone (24), and progesterone (28). The + m/z value used for each compound is shown. **b**, Structures of substrates and corresponding products. Products 22, 23, and 25-27 were isolated from large scale enzymatic reactions, and the structures of 22, 23, and 26/27 were determined by NMR, while the structure of 25 was solved by the crystal sponge method. Note that 26 and 27 are interchangeable with each other according to the NMR data. Products 29-31 were not determined due to separation problems, but their structures were deduced to be hydroxylated and/or oxidized compounds, considering their molecular weights.

Figure 4. Active site view of SptF in complex with natural substrate, intermediate, or unnatural substrate. **a**, Active site residues interact with andiconin D (**1**), including hydrogen bonding with Asn65 and Ser114, indirect interactions with Thr148 and Leu199 via waters in SptF (shown in salmon). **b**, Active site residues interact with andilesin C (**6**), including hydrogen bonding with Asn65 and Ser114, indirect interactions with Ler199 via waters in SptF. **c**, Active site residues interact with emervaridone C (**3**), including hydrogen bonding with Asn150, and indirect interactions with Thr148, Leu199, and Lys265' (chain B) via waters in S114A (shown in yellow). **d**, Active site residues interact with terretonin C (**15**), including hydrogen bonding with Asn150, and indirect interactions with Thr148 and Leu199 via waters in S114A (shown in yellow). Note that the lid-like loop region, shown as blue cartoon, was only observed in the complex structure with **1** or **6**. Water molecules are shown as red spheres, and Fe are shown as orange spheres. α KG/NOG, **1**, **6**, **3**, and **15** are shown as green, cyan, purple, magenta, and blue sticks, respectively. Hydrogen networks are shown as yellow dash lines, and distances from Fe are shown with green dash lines. Key hydrogen bond interactions are shown in green boxes.

Figure 5. Mutagenesis studies. **a**, Active site cavity of SptF with **1** (shown as cyan sticks). Hydrophobic residues (shown as orange sticks), Ile63, Phe133, Ile231, and V264', form a hydrophobic surface to interact with the hydrophobic part of the substrate, whereas hydrophilic residues (shown as salmon sticks), Asn65, Ser114, Thr148, and Asn150, interact with the hydrophilic part of the substrate or intermediate. **b**, **c**, LC-MS EICs of products from *in vitro* assays with mutants using **1** or **15** as the substrate. The + m/z value used for each compound is shown. **d**, Structures of new products generated by mutants. Compounds **32** and **33** were isolated from large scale enzymatic reactions using the F133A and I231A mutants, respectively. Structures were determined by NMR. **e**, Ratios of **4** and **5** from mutants compared with those from wild-type SptF. **f**, Active site residues interacting with **1** (shown as cyan sticks), including hydrogen bonding directly with Ser114 and the newly introduced Thr65, and indirectly with Thr148 and Leu199 via waters in N65T (shown in purple). Note that the coloring scheme for ligands is the same as in Figure 4.

Figure 6. Proposed mechanism of SptF reactions. The mechanism for the generation of emervaridone B (2) is revised in this study. Paths a (red) and b (magenta) branch from the intermediate 3, to produce 4 and 5, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SptFNOGandC.pdf](#)
- [SptFNOGterrC.pdf](#)
- [SptFS114ANOGP3.pdf](#)
- [SptFaKGandD.pdf](#)
- [SptFdelta9.pdf](#)
- [SptFN63TNOGandD.pdf](#)
- [SptFapo.pdf](#)
- [SuplInfo.pdf](#)
- [checkcifcompound25210617.pdf](#)
- [compound25.cif](#)