Article

Molecular Basis for the Biosynthesis of an Unusual Chain-Fused Polyketide Gregatin A

Wei-Guang Wang, Hang Wang, Lian-Qiong Du, Min Li, Lin CHEN, Jian Yu, Gui-Guang Cheng, Meng-Tao Zhan, Qiu-Fen Hu, Lihan Zhang, Min Yao, and Yudai Matsuda

J. Am. Chem. Soc., **Just Accepted Manuscript** • DOI: 10.1021/jacs.0c02337 • Publication Date (Web): 10 Apr 2020

Molecular Basis for the Biosynthesis of an Unusual Chain-Fused Polyketide Gregatin A

Wei-Guang Wang,*[,†,‡] Hang Wang,[§,‡] Lian-Qiong Du,[†,‡] Min Li,[†] Lin Chen,[#] Jian Yu,[§] Gui-Guang Cheng,[∥] Meng-Tao Zhan,[†] Qiu-Fen Hu,[†] Lihan Zhang,[⊥] Min Yao,[§] and Yudai Matsuda*[,#]

[†]Key Laboratory of Chemistry in Ethnic Medicinal Resources, State Ethnic Affairs Commission and Ministry of Education, Yunnan Minzu University, Kunming 650031, Yunnan, China.

[§]Graduate School of Life Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan.

[∥]Yunnan Institute of Food Safety, Kunming University of Science and Technology, Kunming 650500, Yunnan, China.

[⊥]School of Science, Westlake University, 18 Shilongshan Road, Hangzhou 310024, Zhejiang, China.

[#]Department of Chemistry, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong SAR, China.

ABSTRACT: Gregatin A (**1**) is a fungal polyketide featuring an alkylated furanone core, but the biosynthetic mechanism to furnish the intriguing molecular skeleton has yet to be elucidated. Herein, we have identified the biosynthetic gene cluster of gregatin A (**1**) in *Penicillium* sp. sh18, and investigated the mechanism that produces the intriguing structure of **1** by in vivo and in vitro reconstitution of its biosynthesis. Our study established the biosynthetic route leading to **1,** and illuminated that **1** is generated by the fusion of two different polyketide chains, which are, amazingly, synthesized by a single PKS GrgA with the aid of a *trans*-acting enoylreductase GrgB. Chain fusion, as well as chain hydrolysis, is catalyzed by an α/β hydrolase GrgF, hybridizing the C_{11} and C_4 carbon chains by Claisen condensation. Finally, structural analysis and mutational experiments using GrgF provided insight into how the enzyme facilitates the unusual chain-fusing reaction. In unraveling a new biosynthetic strategy involving a bifunctional PKS and a polyketide fusing enzyme, our study expands our knowledge concerning fungal polyketide biosynthesis.

INTRODUCTION

Polyketides are one of the major groups of natural products with a great diversity of structure and of biological functions. The backbone structures of polyketides are assembled by polyketide synthases (PKSs) through repeated elongation of C_2-carbon units in a manner similar to that of the fatty acid synthase.[1] However, PKSs generate an enormous number of molecules due to several factors, including the utilization of non-standard starter/extender units, and variation in the number of chain elongation cycles, and the extent of chain reduction at each elongation step. Furthermore, recent studies have discovered unusual enzymology within the polyketide chain biogenesis,[2] which further contributes to the structural diversification of this class of natural products. Fusion of two polyketide chains to provide a hybrid molecule is one such example (Figures 1A and 1B). Several different mechanisms for the chain fusion are known; for example, MxnB, involved in the biosynthesis of myxopyronins, combines two carbon chains synthesized by two individual PKSs using Claisen condensation.[3] Nevertheless, given the significant diversity in polyketide natural products, it is highly likely there are many more as yet unidentified reactions to afford polyketide skeletons, and biosynthetic studies on structurally unique polyketides could further expand our understanding of the logic behind the polyketide biosynthesis.

Gregatin A (**1**; Figure 1C) is a fungal metabolite originally isolated from the fungus *Cephalosposium gregatum* in 1975 and found to be produced by several different fungal species.[4] Structural determination of gregatin A was finally achieved in 2012 after two rounds of revisions.[5] Gregatin A features an alkylated furanone core, and the molecular architecture suggests a polyketide origin of **1**. Previous studies suggested that the polyketide **2** undergoes oxidative cleavage of the aromatic ring to yield the core skeleton of **1** (Figure 1C);[6] however, no experimental evidence to support this hypothesis has been provided. Furthermore, it might also be possible that the carbon chain of **1** is generated by the fusion of two distinct polyketide chains as seen in the myxopyronin biosynthesis (Figure 1D), although only one type-III PKS is known to catalyze such a chain fusion reaction in fungal natural product pathways (Figure 1B).[7] Collectively, it was expected that the biosynthetic study on gregatin A might result in identification of a novel enzymatic strategy employed by fungal PKSs, and would thereby facilitate future efforts to rationally engineer the PKS machinery.

In this study, we identified the biosynthetic gene cluster of gregatin A (**1**) in *Penicillium* sp. sh18, and successfully reconstituted its biosynthesis in a heterologous fungus, *Aspergillus oryzae*. Isotope-incorporation experiments then indicated that **1** is biosynthesized by fusion of two carbon chains synthesized by a single PKS, GrgA. In vitro enzymatic reactions further confirmed that a predicted hydrolase GrgF is responsible for the fusion of the C_{11} and C_4 carbon chains to produce the backbone skeleton of **1**. Finally, mutational experiments and examination of the X-ray crystal structure of GrgF provided a plausible mechanism for the chain-fusion reaction.

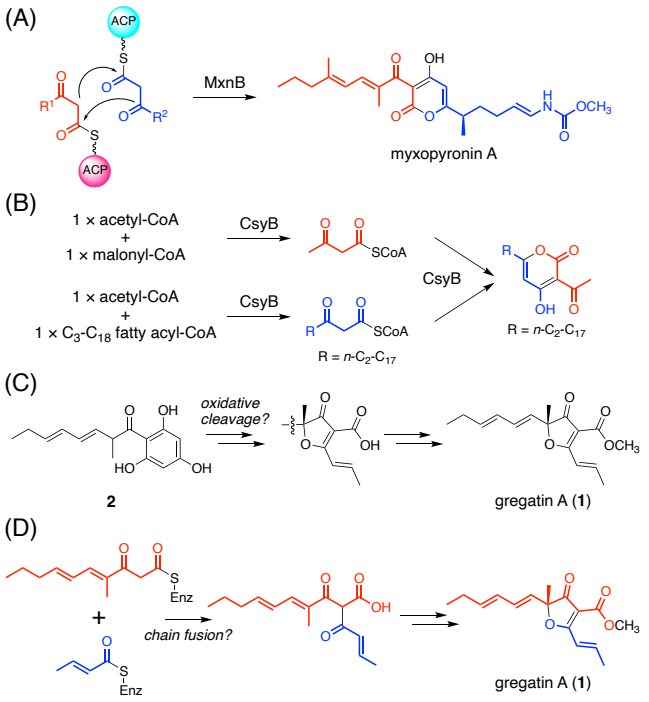


Figure 1. (A) (B) Representative polyketide chain-fusion reactions: Reactions catalyzed by (A) MxnB and (B) CsyB. (C) Structure of gregatin A (**1**), and the previously suggested mechanism to generate the furanone core of **1**. (D) An alternative possible mechanism to produce the backbone skeleton of **1**.

RESULTS

Discovery of the Gregatin A Biosynthetic Gene Cluster and Heterologous Reconstitution of the Biosynthesis. To identify the biosynthetic gene cluster for gregatin A (**1**), we initially obtained the genome sequence of a fungus *Penicillium* sp. sh18, from which **1** was isolated.[4b] On the basis of the structure, the gene cluster for **1** should at least encode a highly reducing (HR)-PKS with ketoreductase (KR), dehydratase (DH), enoylreductase (ER), and methyltransferase (MT) domains, an *O*-methyltransferase, and an oxidative enzyme for the furanone ring formation. Examination of the genome sequence identified one candidate gene cluster that meets the criteria, which was designated as the *grg* cluster (Figure 2A; DDBJ/EMBL/GenBank accession number: LC522971). The *grg* cluster encodes the HR-PKS GrgA, the *trans*-acting ER (*trans*-ER) GrgB, the *O*-methyltransferase GrgD, the α/β hydrolase GrgF, and the cytochrome P450 monooxygenase GrgG. Thus, the biosynthetic pathway of **1** was roughly predicted as follows; GrgA and GrgB collaboratively synthesize the backbone carbon chain, GrgF serves as a thioesterase (TE) and releases the polyketide chain, GrgG forms the furanone ring, and GrgD methylates the carboxylate.

To confirm the involvement of the *grg* cluster in the gregatin A biosynthesis, we then expressed the five enzyme-coding genes in *Aspergillus oryzae*,[8] which is a robust host for heterologous expression of fungal biosynthetic genes.[9] Subsequent HPLC analysis revealed that the transformant produced several new metabolites (Figure 2B, trace vii), which were not detected in the host strain (Figure 2B, trace i). Unfortunately, isolation and structural characterization of most peaks were hampered due to their instability. Nevertheless, we have successfully isolated one of the new products **1**, which was determined to be gregatin A by NMR analysis. Thus, it was confirmed that the *grg* cluster is indeed responsible for the biosynthesis of **1**. We also sought to purify another new peak **3** and found that **3** is a mixture of approximately equal amounts of two gregatin A analogues, namely gregatins C (**3a**) and D (**3b**),[5] in which one of the olefins in **1** is hydrated. Given the diastereomeric nature of **3**, it was suggested that **3a** and **3b** are nonenzymatic products. In accordance, **1** is spontaneously converted to **3a** and **3b** in water (Figure S1), indicating these compounds are off-pathway products.

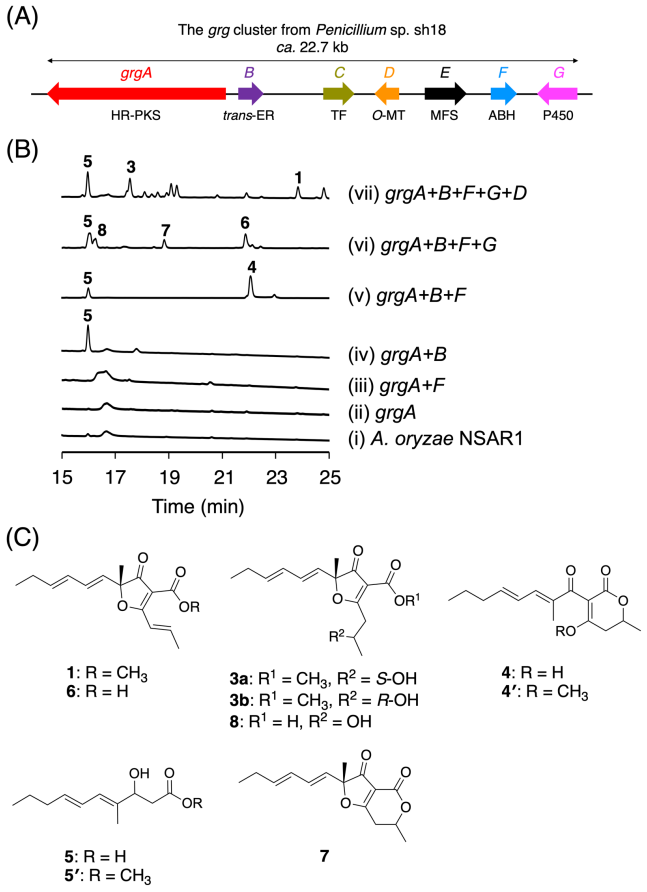


Figure 2. (A) Schematic representations of the *grg* cluster. HR-PKS: highly reducing polyketide synthase; *trans*-ER: *trans*-acting enoylreductase; TF: transcription factor; *O*-MT: *O*-methyltransferase; MFS: major facilitator superfamily transporter; ABH: α/β hydrolase; P450: cytochrome P450 monooxygenase. (B) HPLC profiles of culture supernatant extracts from *A. oryzae* NSAR1 transformants. The chromatograms were monitored at 230 nm. (C) Structures of the metabolites isolated or detected in this study.

Investigation of the Biosynthetic Pathway Leading to Gregatin A. With the successful reconstitution of the gregatin A biosynthesis in *A. oryzae*, we next investigated the biosynthetic origin of **1** using [1-^{13}C]- and [2-^{13}C]-sodium acetate (Figures 3 and S2). When the *A. oryzae* transformant was cultivated in the presence of [1-^{13}C]-sodium acetate, **1** was labeled at C-2, C-4, C-6, C-8, C-10, C-12, and C-15 positions. Meanwhile, [2-^{13}C]-sodium acetate labeled **1** at C-3, C-5, C-7, C-9, C-11, C-14, and C-16 positions. Collectively, only C-13 and C-17 were not labeled either by [1-^{13}C]- and [2-^{13}C]-sodium acetate, and therefore, C-13 and C-17 methyl groups are introduced by the activity of the MT domain of the PKS and by an *O*-methyltransferase, respectively. Notably, the labeling

experiments indicated that C-12 is a part of the polyketide chain, implying that two polyketide chains are fused to afford **1** and that a single PKS GrgA is responsible for the production of both chains.

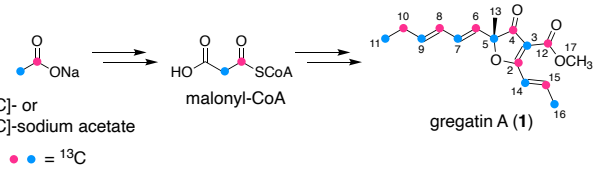


Figure 3. Isotope-labeling experiment of **1** using [1-^{13}C] and [2-^{13}C]sodium acetate.

To further obtain insight into the biosynthetic route leading to **1**, several different *A. oryzae* transformants were created. Initially, we evaluated the transformant with the PKS GrgA, the *trans*-ER GrgB, and the hydrolase GrgF, since GrgB and GrgF were predicted to work together with the PKS. The strain with the three genes produced one major product **4** (Figure 2B, trace v). Unfortunately, structural characterization of **4** was unsuccessful, as we failed to completely purify the compound due to the unstable nature of the compound. Therefore, we methylated **4** using TMS-diazomethane, which yielded relatively stable compound **4'**. NMR analysis of **4'** revealed that **4'** possesses the same carbon chain backbone as that of **1** (Figure 2C), but lacks the characteristic furanone ring; instead, **4'** features a six-membered lactone ring, which might be formed spontaneously after chain release. Thus, we concluded that **4** is the demethylated form of **4'**. Although we could not determine the absolute configuration of **4'**, based on the specific rotation of the compound ($[\alpha]^{25}_{D}$ -50.0), **4'** could be optically pure, or at least one of the enantiomers should be dominant, indicating the presence of a chiral center in the precyclized intermediate. We also confirmed that **4** can be synthesized only in the presence of all the three enzymes (Figure 2B, traces ii to iv), establishing the minimal components required for chain-fused polyketide synthesis. Interestingly, co-expression of the PKS and *trans*-ER genes gave a specific metabolite **5** (Figure 2B, trace iv). Since it was difficult to completely purify **5**, the compound was methylated for further purification and structural analysis. The methylated product **5'** was found to be a linear polyketide with 11 carbon atoms (Figure 2C), and the carbon skeleton of **5** can be found in **1** as well as **4**. Taken together, the hydrolase GrgF appears to be the chain-fusion enzyme.

We next investigated the late-stage biosynthesis of gregatin A (**1**). Considering that the desmethyl analogue of **1** can be found in nature,[6a] we reasoned that the methyltransferase GrgD is responsible for the very last step of the biosynthesis. Thus, we analyzed the metabolites from the four gene-expressing strain harboring the P450 gene *grgG* as well as the aforementioned three genes. The transformant yielded three new products **6**, **7**, and **8**; compound **6** was characterized to be the demethylated analogue of **1** and was named desmethylgregatin A, whereas **7** was identified as a known natural product, cyclogregatin.[10] During the purification of the other product **8**, we found that the compound is spontaneously converted to **6** and **7** (Figure S3). To characterize the structure, **8** was methylated, which provided a mixture of gregatins C (**3a**) and D (**3b**) (Figures S48 and S49). Based on this observation as well as the molecular formula of **8** established by HR-MS analyses, **8** was determined to be the demethylated analogue of **3a** and **3b**. Collectively, the P450 GrgG is responsible for the oxidative heterocyclization to furnish the gregatin scaffold with the furanone ring.

In vitro Reconstitution of the Chain-Fusing Reaction. To obtain in-depth insight into the chain-fusion reaction, we sought to reconstitute the biosynthesis of the fused product **4** in vitro. To this end, the PKS GrgA was obtained as a C-terminal His-tagged protein using *A. oryzae* expression system according to a previously reported method,[11] whereas the *trans*-ER GrgB and hydrolase GrgF were purified from an *Escherichia coli* expression system as a N-terminal His-tagged protein (Figure S4). The three enzymes were then incubated with malonyl-CoA, NADPH, and *S*-adenosylmethionine (SAM), and the enzymatic products were analyzed by LC-MS. The reaction successfully yielded a product, which is identical to the chain fused product **4** obtained from the *A. oryzae* transformant (Figure 4A, traces i and v), confirming that the PKS GrgA is solely responsible for the synthesis of whole carbon chain of **4**. In accordance with the in vivo experiment, **4** was only produced under the existence of the three enzymes (Figure 4A, traces ii to iv). Meanwhile, compound **5** was not detected in the reaction with GrgA and GrgB, suggesting the involvement of an endogenous enzyme(s) of *A. oryzae* in the production of **5**.

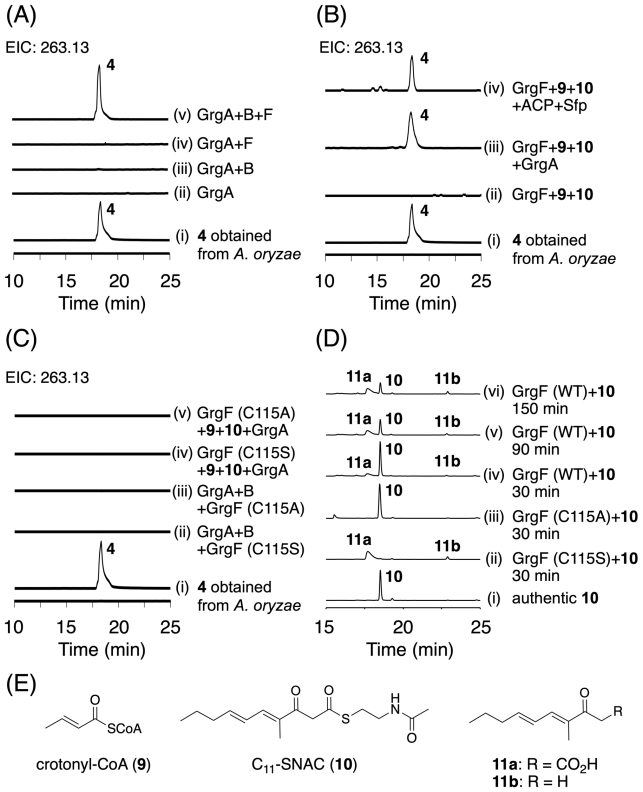


Figure 4. In vitro reconstitution of the chain-fusion reaction. LC-MS chromatograms of products from (A) reactions performed with malonyl-CoA, NADPH, and SAM; (B) reactions of GrgF with the substrate analogues **9** and **10**; and (C) reactions of GrgF mutants. The chromatograms were extracted at *m*/*z* 263.13. Compound **4** partially purified from the *A. oryzae* transformant was used as a standard. (D) HPLC chromatograms of the products from the reactions of GrgF and its mutants with **10**. The chromatograms were monitored at 280 nm. WT: wild type enzyme. (E) Structures of substrate analogues used in this study (**9** and **10**) and the predicted structures of **11a** and **11b**.

Subsequently, we further investigated the function of the hydrolase GrgF, which serves as the chain-fusing enzyme. Given the production of the C_{11}-polyketide **5** in the *A. oryzae* transformant with

the PKS and *trans*-ER genes, we expected that the gregatin A biosynthesis involves the fusion of C_{11} and C_4 carbon chains. To investigate this hypothesis, potential substrate analogues shown in Figure 4E were purchased or synthesized; these molecules have either CoA- or *N*-acetylcysteamine (NAC) substructures, which are known to serve as surrogates of ACP-bound reaction intermediates and can be accepted by a variety of PKS-associated domains/enzymes.[12] The reaction of the hydrolase GrgF with the two substrate analogues, crotonyl-CoA (**9**) and C_{11}-SNAC (**10**), however, did not provide the expected product **4** at detectable levels (Figure 4B, trace ii). Intriguingly, **4** was efficiently produced by the addition of the PKS GrgA to the reaction mixture (Figure 4B, trace iii). It is unclear how exactly the PKS assists the reaction, but the substrate analogues might directly acylate the ACP domain[13] and/or be first accepted by the AT domain to be transferred to the ACP perhaps because of the promiscuity of the AT,[14] which are then utilized by GrgF for the subsequent chain fusion. To investigate the effect of the substrate loading onto the ACP, we purified the ACP domain of GrgA from *E. coli* expression system. We next sought to load **9** onto the ACP using the promiscuous phosphopantetheinyl transferase Sfp derived from *Bacillus subtilis*.[15] Consequently, the reaction with the hydrolase GrgF, the ACP, Sfp, **9**, and **10** successfully generated **4** at levels comparable to that obtained with the whole PKS, judging from the peak areas in the LC-MS analysis (Figure 4B, trace iv). Since our attempt to synthesize the C_{11}-CoA was unfortunately unsuccessful, we could not examine the effect of loading the C_{11} chain to the ACP. Nevertheless, the current experimental results indicate that the crotonyl moiety must be attached to the ACP to allow for chain-fusion and that GrgF is indeed engaged in the reaction.

Mutational Experiments of GrgF. To better understand the function of the chain-fusion enzyme GrgF, we performed mutational experiments of GrgF. It is well known that the majority of α/β hydrolase fold enzymes adopt a Ser-His-Asp (Glu) catalytic triad;[16] however, sequence alignment of GrgF with homologous enzymes revealed that GrgF harbors a cysteine residue (Cys115) instead of a serine at the catalytic site, which is relatively rare within this family of enzymes (Figure S5). To examine the importance of Cys115 in the chain fusion reaction, we mutated the cysteine to serine or alanine, and performed in vitro reactions using the PKS GrgA, *trans*-ER GrgB, and malonyl-CoA as a substrate. Both mutants failed to produce the chain-fused product **4** (Figure 4C, traces ii and iii). We also performed reactions using the substrate analogues **9** and **10** as well as the PKS GrgA, but production of **4** was again not observed (Figure 4C, traces iv and v), indicating that Cys115 is critical for the fusion reaction.

Next, we examined the properties of GrgF and its mutants using the substrate analogues **9** and **10**. Crotonyl-CoA (**9**) was not used by any of the enzymes (Figure S6), which is consistent with the observation that loading of the crotonyl moiety onto the ACP is essential for the chain-fusion reaction. On the other hand, the C_{11}-SNAC **10** was very efficiently consumed by the C115S mutant to give new products **11a** and **11b** (Figure 4D, trace ii), whose molecular formulae were determined to be $C_{11}H_{16}O_3$ and $C_{10}H_{16}O$, respectively, by HR-MS analysis (Figure S7). The molecular formulae respectively correspond to the hydrolyzed form of **10** and its decarboxylated form, and therefore, the C115S mutant functions as a thioesterase, while losing the ability to perform chain fusion. Importantly, the hydrolysis occurs much slower in the wild type enzyme-catalyzed reaction (Figure 4D, traces iv to vi), and no hydrolyzed product was detected from the C115A mutant (Figure 4D, trace iii). Based on these observations, we reasoned that GrgF initially accepts the C_{11} chain and that Cys115 is important to retain the carbon chain until the C_4 chain approaches the active site of GrgF.

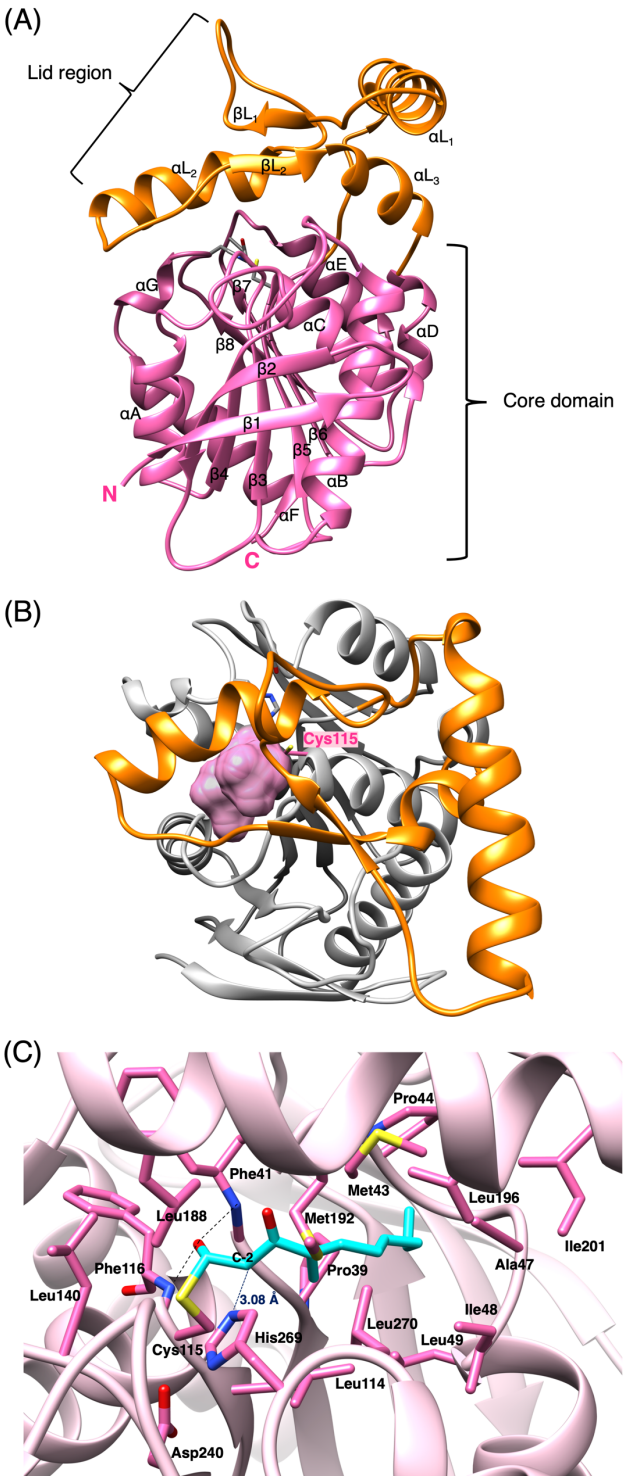


Figure 5. (A) Overall structure of GrgF. The core domain and lid region are shown in *pink* and *orange*, respectively. One of the monomers in the asymmetric unit is shown. (B) Probable substrate binding site of GrgF. The lid region (*orange*) seals the hydrophobic chamber (*pink*). The binding site was predicted using CASTp.[17] (C) Docked model of the C_{11} chain (*cyan*). C-1 is covalently bound to Cys115.

Structural Characterization of GrgF and Docking Simulation. To further clarify the mechanism by which GrgF performs the chain fusion reaction, we solved the X-ray crystal structure of GrgF at 1.9 Å resolution (PDB ID: 6LZH; Figure 5A and Table S6). The crystal structure revealed the presence of a dimer in the asymmetric unit. GrgF possesses an α/β hydrolase fold with an inserted region between β6 and αD, which is a commonly found but variable region in TE domains of PKSs/NRPSs known as the "lid region".[18] Unlike typical TE domains with a lid region consisting of two helices,[19] the corresponding region of GrgF harbors three helices αL_1, αL_2 and αL_3 as well as two sheets βL_1 and βL_2, and it seals the active site from solvents, leading to the closed state of GrgF. Although some other studies successfully obtained open state structures of TEs,[18b,20] the structure of open form GrgF was not obtained in our study, suggesting that the approach of the substrate is required for the conformational change. Structural comparison of GrgF with other enzymes using the Dali program[21] indicated that GrgF shares structural similarities with hydrolase family enzymes PA2218 (PDB ID: 2HDW) and YcjY (PDB ID: 5XB6), whose functions have yet to be clarified, with rms deviation of 2.2 and 2.1 Å, respectively (Figure S8).

As predicted from the protein sequence, GrgF forms a Cys-His-Asp catalytic triad with contributions from Cys115, His269, and Asp240. The carboxylate residue at the catalytic triad of TEs fused with a PKS or NRPS is normally found at a noncanonical position, C-terminus of β6,[18] whereas Asp240 of GrgF is located in the typical position, the loop between β7-αE. We also found that the backbone amide groups of Phe41 and Phe116 form a probable oxyanion hole, which is commonly seen in α/β hydrolase fold enzymes, and is important for the stabilization of the tetrahedral intermediate.[16] Subsequently, we investigated the potential substrate binding site, and found a highly hydrophobic chamber in the vicinity of the active site residues (Figure 5B), which is isolated from solvents because of the lid region.

To investigate whether the hydrophobic chamber could accommodate the predicted substrate, we sought to dock the C_{11} chain, which should be accepted by GrgF first, to the probable substrate binding site of the enzyme. The docking simulation, in which the C_{11} chain is covalently bound to Cys115, revealed that the long carbon chain could indeed be accepted at this site (Figure 5C). In the identified binding mode, the C_{11} chain is located in a pocket consisting of Pro39, Gly40, Phe41, Met43, Pro44, Ala47, Ile48, Leu49, Leu114, Leu140, Leu188, Met192, Leu196, Ile201, and Leu270. Additionally, the carbonyl oxygen of the thioester linkage between the substrate and GrgF was found in the oxyanion hole formed by Phe41 and Phe116. The C-2 of the substrate, to which the second C_4 chain fuses, resides proximate to the catalytic His269, which might deprotonate C-2 for the chain-fusion reaction. In order to further confirm whether the docking simulation is reasonable, a molecular dynamics (MD) simulation based on the obtained docking model was performed (Figures S9 to S11). The MD simulation showed that across 99% and 92% of the simulation time, hydrogen bonds are maintained between the substrate thioester carbonyl oxygen and Phe41 and Phe116, respectively (Figure S10).

DISCUSSION

In this study, we have discovered the biosynthetic gene cluster of gregatin A (**1**) and successfully reconstituted the biosynthesis in a heterologous fungus *A. oryzae*. We have also performed a series of in vitro enzymatic reactions to unveil the mechanism to afford the backbone skeleton of **1**. Based on all the experimental data, we now propose the biosynthetic pathway leading to **1** as follows (Figure 6). The PKS GrgA synthesizes C_{11} and C_4 polyketide chains in the presence and absence of the *trans*-ER GrgB, respectively, which are then hybridized by the chain-fusing enzyme GrgF. GrgF would first undergo a conformational change to the open form, and the active site Cys115 is then acylated by the C_{11} chain. The carbonyl oxygen of the thioester would be placed at the oxyanion hole formed by Phe41 and Phe116, which facilitates the transthioesterification reaction. After the elimination of the phosphopantetheinyl chain, the second polyketide chain of four carbon long is delivered adjacent to the enzyme-bound C_{11} chain. The catalytic histidine, His269, deprotonates a proton from C-2 of the long chain, and the resultant carbanion attacks the C-1 carbonyl of the crotonyl group to perform Claisen condensation, by which the phosphopantetheinyl chain is released. Eventually, hydrolysis of the thioester linkage probably by a His269-activated water molecule completes the reaction to afford the reaction product **12**. Compound **12** would spontaneously be transformed to **4** by the attack of the conjugated olefin at C-3' by the carboxy group followed by keto-enol tautomerization (Figure 6A). Next, the P450 GrgG accepts **12** as a substrate and performs the oxidative cyclization to furnish the gregatin scaffold. In this transformation, GrgG would initially abstract a hydrogen atom from C-8 to generate the substrate radical **13**, from which one electron is transferred to the iron-heme center to yield the carbocationic species **14** (Figure 6B, path **a**). Heterocyclization along with double bond isomerizations provide the product **6** with the furanone ring. Alternatively, GrgG might provide the hydroxylated product **15**, which is followed by dehydration to give the cyclized product **6** (Figure 6B, path **b**). Finally, the *O*-methyltransferase GrgD methylates the carboxyl group to provide gregatin A (**1**). The gregatin A analogue, cyclogregatin (**7**), would originate from **4** in a similar manner to that for **1**, whereas gregatins C (**3a**) and D (**3b**) are nonenzymatic products derived from **1** (Figure S12).

The most intriguing point of the gregatin A biosynthetic pathway would be the chain-fusion reaction catalyzed by GrgF. Actually, such reactions can be found in other polyketide biosynthetic pathways,[3,7,22] but, so far, all characterized enzymes for polyketide chain fusion by Claisen condensation are thiolase-like enzymes. Thus, GrgF provides the first example in which an α/β hydrolase fold enzyme performs a polyketide chain-fusion reaction. Importantly, GrgF adopts cysteine residue (Cys115), as the catalytic residue and as the site for the polyketide chain attachment, instead of most commonly found serine residue. Mutational experiment illuminated the importance of Cys115 to prevent the first polyketide chain from being hydrolyzed before the fusion with the second chain. It is still unclear on what role the cysteine residue plays in retaining the C_{11} chain, but it is generally thought that a cysteine residue generates more stable acyl-enzyme complex in this family of enzymes. For example, the serine to cysteine mutation in the catalytic triad of cytosolic type I acyl-CoA thioesterase (CTE-I) resulted in a significantly reduced activity, in which the enzyme was efficiently acylated by the substrate.[23] Furthermore, the naturally-occurring Cys-His-Asp triad can be found in some TE domains associated with polyketide/nonribosomal peptide biosynthesis, which possess an activity distinct from simple hydrolysis.[24] As observed in our case, the cysteine to serine mutation of these TE domains resulted in the loss of the original activity, but the mutants still produced the chain-hydrolyzed product, and therefore, it would be reasonable that GrgF possesses the cysteine residue for its catalysis.

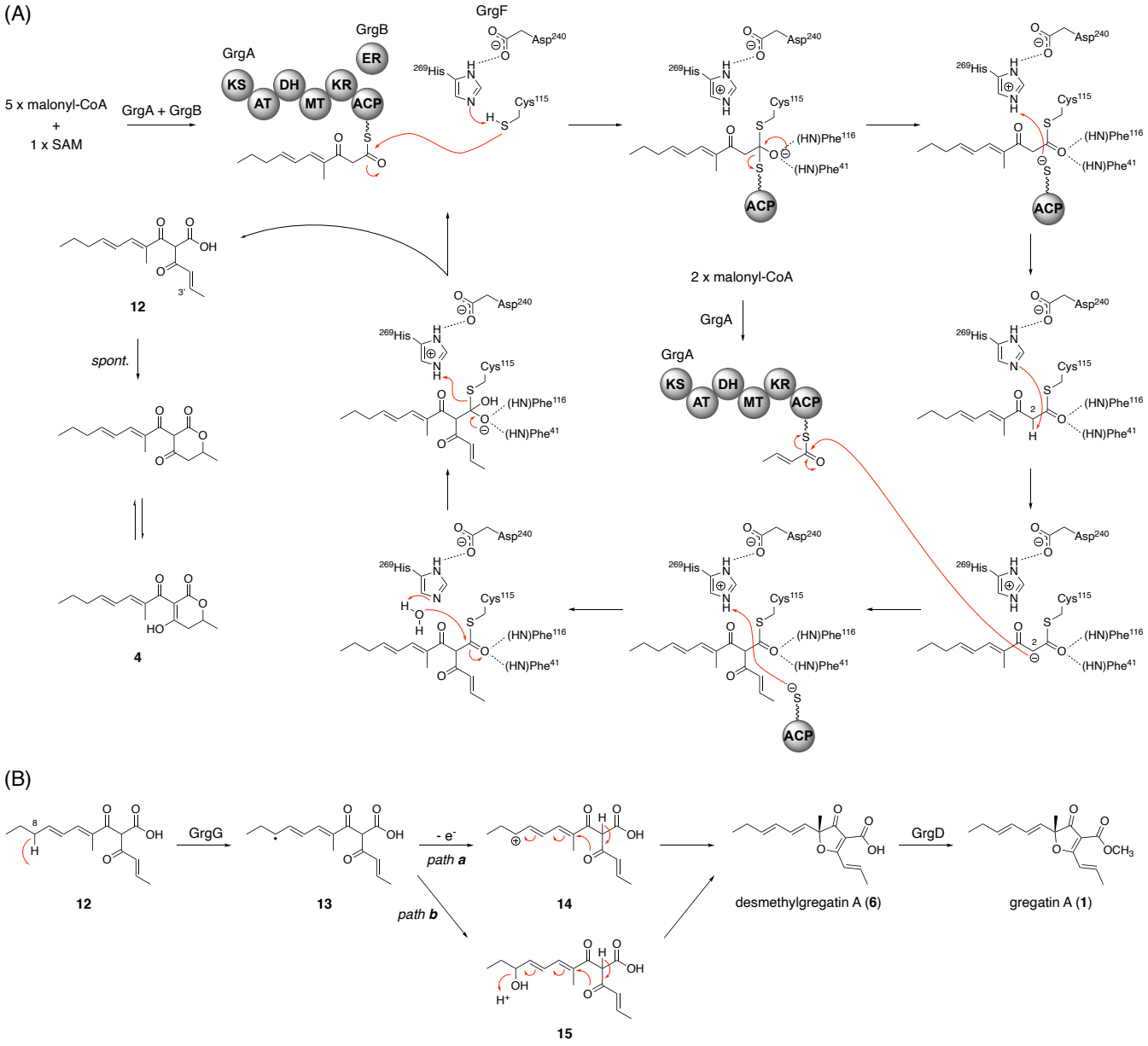


Figure 6. (A) Proposed biosynthetic pathway of the chain-fused product **4** and catalytic mechanism of GrgF. (B) Proposed late-stage biosynthesis of gregatin A (**1**).

Another important feature of the gregatin A pathway is that the biosynthesis involves the fusion of two distinct polyketide chains synthesized by a single PKS. At the moment, only a few examples of PKSs synthesize two different carbon chains to eventually generate a fused product,[7,25] and to the best of our knowledge, this is the first characterization of a fungal type I PKS that naturally synthesizes two distinct chain products for further hybridization. Interestingly, bioinformatic searches revealed that the gene cluster for biosynthesis of gregatin A (**1**) is highly similar to that of other fungal natural products, fujikurins,[26] which are apparently biosynthesized by the fusion of two polyketide chains. The gene cluster, found in the genome of *Fusarium fujikuroi*, encodes proteins that are highly homologous to the PKS GrgA, the *trans*-ER GrgB, and the chain-fusing enzyme GrgF, each exhibiting ~70% protein sequence identity with the corresponding enzyme (Figure S13). It should be noted that the GrgF homologue in the fujikurin pathway, FFUJ_12241, also harbors a cysteine residue in site corresponding to Cys115. Although the biosynthetic pathway of fujikurins has yet to be elucidated, given the high similarity of its biosynthetic enzymes with the Grg enzymes and their structures, fujikurins should also be biosynthesized in a manner similar to that described for **1** (Figure S13). Thus, the PKS (PKS19), together with the *trans*-ER, would produce two different polyketide chains with seven and five carbon atoms, which are then fused together to give fujikurin D. Although the biosyntheses of gregatins and fujikurins differ in the chain lengths and degree of *C*-methylation, we note that the enoylreductions by the *trans*-ERs are only required in the first-round elongation of the longer chains. Therefore, the occurrence of the enoylreduction in the first-round elongation step might be the key determinant of the chain length in the biosynthesis of these molecules. Moreover, the database search identified many proteins homologous to GrgF and FFUJ_12241 containing the Cys-His-Asp triad (Figure S14), and the genes encoding these proteins are often clustered with a PKS gene. Indeed, some other

fungal natural products appear to be biosynthesized in a similar manner, including alternaric acid[27] and ophiofuranone A[28] (Figure S15). Thus, it is suggested that chain-fused polyketides are widespread in fungi and that genome mining targeting these homologues provides valuable leads to the discovery of natural products with a novel scaffold.

CONCLUSION

In this study, we have provided, for the first time, the complete biosynthetic pathway of the fungal polyketide gregatin A (**1**), more than 40 years after its first isolation. Notably, the single PKS GrgA produces two distinct carbon chains, which are further hybridized by the α/β hydrolase fold enzyme GrgF. Future structural biological characterizations of GrgF as well as the ACP domain will further clarify the mechanism behind the unusual chain-fusion reaction, and would also contribute to the synthesis of unnatural fused polyketides. The discovery and characterization of the unusual chain-fusion enzyme GrgF indicates that there must still be a number of overlooked but intriguing chemical reactions in polyketide pathways, despite the decades of intensive studies on their biosynthetic machineries. Collectively, our study has unraveled a novel biosynthetic strategy employed by fungi, which should facilitate biosynthetic studies on analogous natural products, and mining of unexploited metabolites with unique molecular architectures.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.
Experimental details and supplementary tables and figures. (PDF)

AUTHOR INFORMATION

Corresponding Author

*wwg@live.cn
*ymatsuda@cityu.edu.hk

Author Contributions

‡These authors contributed equally.

Notes

The authors declare no competing interest.

ACKNOWLEDGMENT

We are grateful to Prof. Katsuya Gomi (Tohoku University) and Prof. Jun-ichi Maruyama (The University of Tokyo) for kindly providing the expression vectors and the fungal strain. We thank Prof. Li-Dong Shao (Yunnan University of Chinese Medicine) and Baomin Fan (Yunnan Minzu University) for their valuable suggestions in the synthesis of the substrate analogue. We appreciate the Kunming Institute of Botany, Chinese Academic Sciences for using the Schrodinger Suite software and the valuable suggestions from Prof. Zhi-Li Zuo in the institute. This project was in part supported financially by the National Natural Science Foundation of China (Nos. 31760096 and 31960095 to W.-G.W.), Yunnan Ten-thousand Talents Program to W.-G.W., the National Natural Science Foundation of Yunnan provinces (No. 2019FB027 to W.-G.W.), Yunnan Innovative Research Team for Discovery and Comprehensive Utilization of Functional Small Molecules in Medicinal Plants, and the Science Foundation of Yunnan to W.-G.W., Grants-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science and Technology of Japan (JSPS KAKENHI Grant No. JP19H04633 to M.Y.), and grants from City University of Hong Kong (Project No. 7200579 and 9610412 to Y.M.).

REFERENCES

(1) Fischbach, M. A.; Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem. Rev.* **2006**, *106*, 3468-3496.

(2) Hertweck, C. Decoding and reprogramming complex polyketide assembly lines: Prospects for synthetic biology. *Trends Biochem. Sci.* **2015**, *40*, 189-199.

(3) (a) Sucipto, H.; Wenzel, S. C.; Müller, R. Exploring chemical diversity of α-pyrone antibiotics: Molecular basis of myxopyronin biosynthesis. *ChemBioChem* **2013**, *14*, 1581-1589. (b) Sucipto, H.; Sahner, J. H.; Prusov, E.; Wenzel, S. C.; Hartmann, R. W.; Koehnke, J.; Müller, R. In vitro reconstitution of α-pyrone ring formation in myxopyronin biosynthesis. *Chem. Sci.* **2015**, *6*, 5076-5085.

(4) (a) Kotayashi, K.; Ui, T. Isolation of phytotoxic substances produced by *Cephalosposium gregatum* Allington & Chamberlain. *Tetrahedron Lett.* **1975**, *16*, 4119-4122. (b) Wang, W.-G.; Li, A.; Yan, B.-C.; Niu, S.-B.; Tang, J.-W.; Li, X.-N.; Du, X.; Challis, G. L.; Che, Y.; Sun, H.-D.; Pu, J.-X. LC-MS-guided isolation of penicilfuranone A: A new antifibrotic furancarboxylic acid from the plant endophytic fungus *Penicillium* sp. sh18. *J. Nat. Prod.* **2016**, *79*, 149-155.

(5) Burghart-Stoll, H.; Brückner, R. Total syntheses of the gregatins A–D and aspertetronin A: Structure revisions of these compounds and of aspertetronin B, together with plausible structure revisions of gregatin E, cyclogregatin, graminin A, the penicilliols A and B, and the huaspenones A and B. *Eur. J. Org. Chem.* **2012**, *2012*, 3978-4017.

(6) (a) Yin, G.-P.; Wu, Y.-R.; Yang, M.-H.; Li, T.-X.; Wang, X.-B.; Zhou, M.-M.; Lei, J.-L.; Kong, L.-Y. Citrifurans A–D, four dimeric aromatic polyketides with new carbon skeletons from the fungus *Aspergillus* sp. *Org. Lett.* **2017**, *19*, 4058-4061. (b) Yin, G.-P.; Wu, Y.-R.; Han, C.; Wang, X.-B.; Gao, H.-L.; Yin, Y.; Kong, L.-Y.; Yang, M.-H. Asperones A–E, five dimeric polyketides with new carbon skeletons from the fungus *Aspergillus* sp. AWG 1–15. *Org. Chem. Front.* **2018**, *5*, 2432-2436.

(7) (a) Seshime, Y.; Juvvadi, P. R.; Kitamoto, K.; Ebizuka, Y.; Fujii, I. Identification of csypyrone B1 as the novel product of *Aspergillus oryzae* type III polyketide synthase CsyB. *Bioorg. Med. Chem.* **2010**, *18*, 4542-4546. (b) Mori, T.; Yang, D.; Matsui, T.; Hashimoto, M.; Morita, H.; Fujii, I.; Abe, I. Structural basis for the formation of acylalkylpyrones from two β-ketoacyl units by the fungal type III polyketide synthase CsyB. *J. Biol. Chem.* **2015**, *290*, 5214-5225.

(8) Jin, F. J.; Maruyama, J.; Juvvadi, P. R.; Arioka, M.; Kitamoto, K. Development of a novel quadruple auxotrophic host transformation system by *argB* gene disruption using *adeA* gene and exploiting adenine auxotrophy in *Aspergillus oryzae*. *FEMS Microbiol. Lett.* **2004**, *239*, 79-85.

(9) (a) 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.* **2014**, *136*, 15326-15336. (b) Matsuda, Y.; Bai, T.; Phippen, C. B. W.; Nødvig, C. S.; Kjærbølling, I.; Vesth, T. C.; Andersen, M. R.; Mortensen, U. H.; Gotfredsen, C. H.; Abe, I.; Larsen, T. O. Novofumigatonin biosynthesis involves a non-heme iron-dependent endoperoxide isomerase for orthoester formation. *Nat. Commun.* **2018**, *9*, 2587. (c) Liu, C.; Tagami, K.; Minami, A.; Matsumoto, T.; Frisvad, J. C.; Suzuki, H.; Ishikawa, J.; Gomi, K.; Oikawa, H. Reconstitution of biosynthetic machinery for the synthesis of the highly elaborated indole diterpene penitrem. *Angew. Chem. Int. Ed.* **2015**, *54*, 5748-5752.

(10) Anke, H.; Casser, I.; Schrage, M.; Steglich, W. Cyclogregatin, a new metabolite from *Aspergillus panamensis*. *J. Antibiot.* **1988**, *41*, 1681-1684.

(11) Awakawa, T.; Yokota, K.; Funa, N.; Doi, F.; Mori, N.; Watanabe, H.; Horinouchi, S. Physically discrete β-lactamase-type thioesterase catalyzes product release in atrochrysone synthesis by iterative type I polyketide synthase. *Chem. Biol.* **2009**, *16*, 613-623.

(12) Pohl, N. L.; Gokhale, R. S.; Cane, D. E.; Khosla, C. Synthesis and incorporation of an *N*-acetylcysteamine analogue of methylmalonyl-CoA by a modular polyketide synthase. *J. Am. Chem. Soc.* **1998**, *120*, 11206-11207.

(13) (a) Misra, A.; Sharma, S. K.; Surolia, N.; Surolia, A. Self-acylation properties of type II fatty acid biosynthesis acyl carrier protein. *Chem. Biol.* **2007**, *14*, 775-783. (b) Walker, M. C.; Thuronyi, B. W.; Charkoudian, L. K.; Lowry, B.; Khosla, C.; Chang, M. C. Y. Expanding the fluorine chemistry of living systems using engineered polyketide synthase pathways. *Science* **2013**, *341*, 1089-1094.

(14) Rittner, A.; Paithankar, K. S.; Huu, K. V.; Grininger, M. Characterization of the polyspecific transferase of murine type I fatty acid synthase (FAS) and implications for polyketide synthase (PKS) engineering. *ACS Chem. Biol.* **2018**, *13*, 723-732.

(15) (a) Quadri, L. E. N.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **1998**, *37*, 1585-1595. (b) Yin, J.; Lin, A. J.; Golan, D. E.; Walsh, C. T. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nat. Protoc.* **2006**, *1*, 280-285.

(16) Rauwerdink, A.; Kazlauskas, R. J. How the same core catalytic machinery catalyzes 17 different reactions: The serine-histidine-aspartate catalytic triad of α/β-hydrolase fold enzymes. *ACS Catal.* **2015**, *5*, 6153-6176.

(17) Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0: Computed atlas of surface topography of proteins. *Nucleic Acids Res.* **2018**, *46*, W363-W367.

(18) (a) Tseng, C. C.; Bruner, S. D.; Kohli, R. M.; Marahiel, M. A.; Walsh, C. T.; Sieber, S. A. Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase. *Biochemistry* **2002**, *41*, 13350-13359. (b) Samel, S. A.; Wagner, B.; Marahiel, M. A.; Essen, L.-O. The thioesterase domain of the fengycin biosynthesis cluster: A structural base for the macrocyclization of a non-ribosomal lipopeptide. *J. Mol. Biol.* **2006**, *359*, 876-889. (c) Korman, T. P.; Crawford, J. M.; Labonte, J. W.; Newman, A. G.; Wong, J.; Townsend, C. A.; Tsai, S.-C. Structure and function of an iterative polyketide synthase thioesterase domain catalyzing Claisen cyclization in aflatoxin biosynthesis. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 6246-6251.

(19) Horsman, M. E.; Hari, T. P. A.; Boddy, C. N. Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: Logic gate or a victim of fate? *Nat. Prod. Rep.* **2016**, *33*, 183-202.

(20) (a) Chakravarty, B.; Gu, Z.; Chirala, S. S.; Wakil, S. J.; Quiocho, F. A. Human fatty acid synthase: Structure and substrate selectivity of the thioesterase domain. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15567-15572. (b) Whicher, J. R.; Florova, G.; Sydor, P. K.; Singh, R.; Alhamadsheh, M.; Challis, G. L.; Reynolds, K. A.; Smith, J. L. Structure and function of the RedJ protein, a thioesterase from the prodiginine biosynthetic pathway in *Streptomyces coelicolor*. *J. Biol. Chem.* **2011**, *286*, 22558-22569.

(21) Holm, L. Benchmarking fold detection by DaliLite v.5. *Bioinformatics* **2019**, *35*, 5326-5327.

(22) (a) Song, L.; Barona-Gomez, F.; Corre, C.; Xiang, L.; Udwary, D. W.; Austin, M. B.; Noel, J. P.; Moore, B. S.; Challis, G. L. Type III polyketide synthase β-ketoacyl-ACP starter unit and ethylmalonyl-CoA extender unit selectivity discovered by *Streptomyces coelicolor* genome mining. *J. Am. Chem. Soc.* **2006**, *128*, 14754-14755. (b) Zocher, G.; Vilstrup, J.; Heine, D.; Hallab, A.; Goralski, E.; Hertweck, C.; Stahl, M.; Schäberle, T. F.; Stehle, T. Structural basis of head to head polyketide fusion by CorB. *Chem. Sci.* **2015**, *6*, 6525-6536. (c) Mori, T.; Awakawa, T.; Shimomura, K.; Saito, Y.; Yang, D.; Morita, H.; Abe, I. Structural insight into the enzymatic formation of bacterial stilbene. *Cell Chem. Biol.* **2016**, *23*, 1468-1479. (d) Kresovic, D.; Schempp, F.; Cheikh-Ali, Z.; Bode, H. B. A novel and widespread class of ketosynthase is responsible for the head-to-head condensation of two acyl moieties in bacterial pyrone biosynthesis. *Beilstein J. Org. Chem.* **2015**, *11*, 1412-1417. (e) Zhang, D.; Zhang, F.; Liu, W. A KAS-III heterodimer in lipstatin biosynthesis nondecarboxylatively condenses C8 and C14 fatty acyl-CoA substrates by a variable mechanism during the establishment of a C22 aliphatic skeleton. *J. Am. Chem. Soc.* **2019**, *141*, 3993-4001.

(23) Huhtinen, K.; O'Byrne, J.; Lindquist, P. J. G.; Contreras, J. A.; Alexson, S. E. H. The peroxisome proliferator-induced cytosolic type I acyl-CoA thioesterase (CTE-I) is a serine-histidine-aspartic acid α/β hydrolase. *J. Biol. Chem.* **2002**, *277*, 3424-3432.

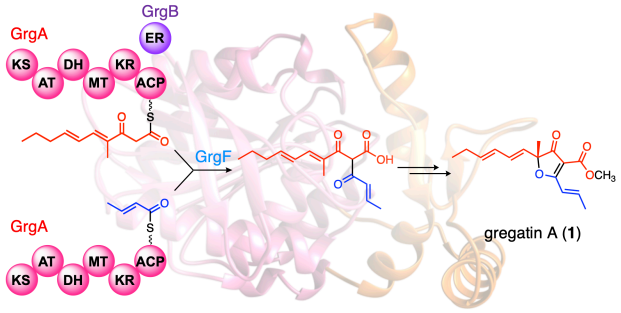
(24) (a) Schaffer, J. E.; Reck, M. R.; Prasad, N. K.; Wencewicz, T. A. β-Lactone formation during product release from a nonribosomal peptide synthetase. *Nat. Chem. Biol.* **2017**, *13*, 737-744. (b) Tang, M.-C.; Fischer, C. R.; Chari, J. V.; Tan, D.; Suresh, S.; Chu, A.; Miranda, M.; Smith, J.; Zhang, Z.; Garg, N. K.; St. Onge, R. P.; Tang, Y. Thioesterase-catalyzed aminoacylation and thiolation of polyketides in fungi. *J. Am. Chem. Soc.* **2019**, *141*, 8198-8206.

(25) Cohen, D. R.; Townsend, C. A. A dual role for a polyketide synthase in dynemicin enediyne and anthraquinone biosynthesis. *Nature Chem.* **2018**, *10*, 231-236.

(26) von Bargen, K. W.; Niehaus, E.-M.; Krug, I.; Bergander, K.; Würthwein, E.-U.; Tudzynski, B.; Humpf, H.-U. Isolation and structure elucidation of fujikurins A–D: Products of the PKS19 gene cluster in *Fusarium fujikuroi*. *J. Nat. Prod.* **2015**, *78*, 1809-1815.

(27) Tabuchi, H.; Oikawa, H.; Ichihara, A. Biosynthetic study of alternaric acid: Isolation of plausible biosynthetic intermediates and origins of the hydrogen and oxygen atoms. *J. Chem. Soc., Perkin Trans. 1* **1994**, 2833-2839.

(28) Li, Y.-L.; Zhu, R.-X.; Li, G.; Wang, N.-N.; Liu, C.-Y.; Zhao, Z.-T.; Lou, H.-X. Secondary metabolites from the endolichenic fungus *Ophiosphaerella korrae*. *RSC Adv.* **2019**, *9*, 4140-4149.


GrgB
ER
GrgA
KS
DH
KR
AT
MT
ACP
GrgF
GrgA
KS
AT
MT
ACP
DH
KR
gregatin A (1)