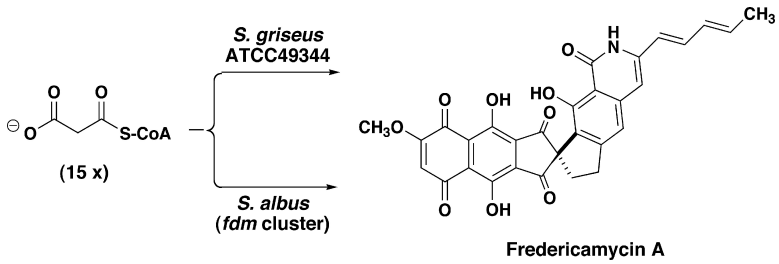
Article

Cloning, Sequencing, Analysis, and Heterologous Expression of the Fredericamycin Biosynthetic Gene Cluster from *Streptomyces griseus*

Evelyn Wendt-Pienkowski, Yong Huang, Jian Zhang, Bensheng Li, Hao Jiang, Hyungjin Kwon, C. Richard Hutchinson, and Ben Shen

J. Am. Chem. Soc., **2005**, 127 (47), 16442-16452• DOI: 10.1021/ja054376u • Publication Date (Web): 05 November 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 8 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 11/05/2005

Cloning, Sequencing, Analysis, and Heterologous Expression of the Fredericamycin Biosynthetic Gene Cluster from *Streptomyces griseus*

Evelyn Wendt-Pienkowski,[†] Yong Huang,[†] Jian Zhang,[‡] Bensheng Li,[†] Hao Jiang,[†] Hyungjin Kwon,[†] C. Richard Hutchinson,[†,§] and Ben Shen*[,†,‡]

Contribution from the Division of Pharmaceutical Sciences and Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53705, and Kosan Biosciences Inc., 3832 Bay Center Place, Hayward, California 94545

Received July 1, 2005; E-mail: bshen@pharmacy.wisc.edu

Abstract: Fredericamycin (FDM) A, a pentadecaketide featuring two sets of *peri*-hydroxy tricyclic aromatic moieties connected through a unique chiral spiro carbon center, exhibits potent cytotoxicity and has been studied as a new type of anticancer drug lead because of its novel molecular architecture. The *fdm* gene cluster was localized to 33-kb DNA segment of *Streptomyces griseus* ATCC 49344, and its involvement in FDM A biosynthesis was proven by gene inactivation, complementation, and heterologous expression experiments. The *fdm* cluster consists of 28 open reading frames (ORFs), encoding a type II polyketide synthase (PKS) and tailoring enzymes as well as several regulatory and resistance proteins. The FDM PKS features a KS_α subunit with heretofore unseen tandem cysteines at its active site, a KS_β subunit that is distinct phylogenetically from KS_β of hexa-, octa-, or decaketide PKSs, and a dedicated phosphopantetheinyl transferase. Further study of the FDM PKS could provide new insight into how a type II PKS controls chain length in aromatic polyketide biosynthesis. The availability of the *fdm* genes, in vivo characterization of the *fdm* cluster in *S. griseus*, and heterologous expression of the *fdm* cluster in *Streptomyces albus* set the stage to investigate FDM A biosynthesis and engineer the FDM biosynthetic machinery for the production of novel FDM A analogues.

Introduction

Several clinically used drugs are aromatic polyketides, such as the antitumor agents daunorubicin/doxorubicin and the antibacterial antibiotic oxytetracycline. They are biosynthesized by bacterial aromatic polyketide synthases (PKSs), also known as type II PKS.[1,2] Type II PKSs minimally consist of four subunits: ketoacyl synthase α (KS_α), ketoacyl synthase β (KS_β, also known as chain length factor), acyl carrier protein (ACP), and malonyl CoA:ACP acyltransferase (MAT). A PKS catalyzes the biosynthesis of the polyketide backbone from acyl-CoA precursors, which usually are activated by a MAT for loading to the holo-ACP, and then condensed iteratively with a unique chain starter unit, via a series of decarboxylative Claisen condensation reactions until the specified chain length is reached. Aromatases and cyclases then catalyze the regiospecific folding and cyclization of the nascent polyketide chain to form the linear, angular, or discoid aromatic ring systems.[3] The latter are then further customized by a host of tailoring enzymes that perform modifications such as oxidation, reduction, methylation, and glycosylation to yield the final polyketides. These modifications introduce a large amount of structural diversity, and many of the resulting products have interesting biological activities.[4]

Combinatorial biosynthesis is an emerging technology that has enabled the production of numerous "unnatural" natural aromatic polyketides with various pharmacological properties. For example, through gene inactivation of *mtmW*, a ketoreductase involved in mithramycin biosynthesis, the major product mithramycin-SK showed an improved therapeutic index compared to the parent drug mithramycin.[5] Using bimodular polyketide synthases consisting of tetracenomycin (TCM) or actinorhodin (ACT) minimal PKSs and the R1128 priming PKS module that specifies for various acyl-CoA starters, several novel anthracyclines with anticancer activities or inhibitory activities toward glucose-6-phosphate translocase have been produced.[6] More recently, through mutasynthetic or chemoenzymatic routes, the production of an aminocoumarin library facilitated the study of structure-activity relationships for this family of natural products.[7,8] However, obstacles remain when attempting to alter the chain length of the nascent polyketide intermediate giving

[†] Division of Pharmaceutical Sciences.
[‡] Department of Chemistry.
[§] Kosan Biosciences Inc..

(1) Hopwood, D. A. *Chem. Rev.* **1997**, *97*, 2465−2497.
(2) Shen, B. *Top. Curr. Chem.* **2000**, *209*, 1−51.
(3) Jakobi, K.; Hertweck, C. *J. Am. Chem. Soc.* **2004**, *126*, 2298−2299.
(4) Rix, U.; Fischer, C.; Remsing, L. L.; Rohr, J. *Nat. Prod. Rep.* **2002**, *19*, 542−580.
(5) Remsing, L. L.; Gonzalez, A. M.; Nur-e-Alam, M.; Fernandez-Lozano, M. J.; Brana, A. F.; Rix, U.; Oliveira, M. A.; Mendez, C.; Salas, J. A.; Rohr, J. *J. Am. Chem. Soc.* **2003**, *125*, 5745−5753.
(6) Tang, Y.; Lee, T. S.; Khosla, C. *PLoS Biol.* **2004**, *2*, E31.
(7) Galm, U.; Dessoy, M. A.; Schmidt, J.; Wessjohann, L. A.; Heide, L. *Chem. Biol.* **2004**, *11*, 173−183.
(8) Xu, H.; Heide, L.; Li, S. M. *Chem. Biol.* **2004**, *11*, 655−662.

10.1021/ja054376u CCC: $30.25

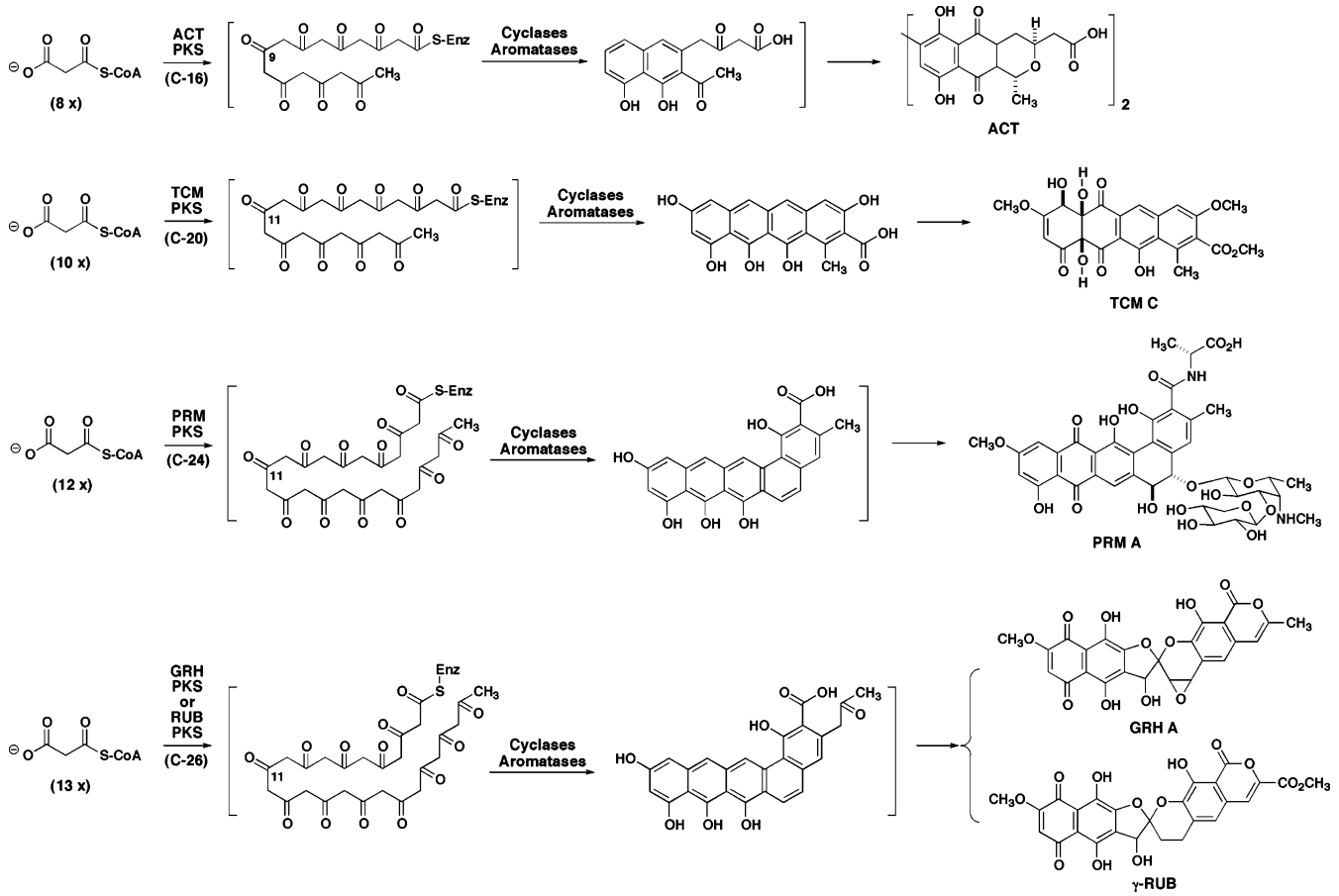


Figure 1. Examples of type II PKS-catalyzed biosynthesis of aromatic polyketides with variable chain length. ACT, octaketide actinorhodin (C-16); TCM C, decaketide tetracenomycin C (C-20), PRM A, dodecaketide pradimicin A (C-24); and GRH A and γ-RUB, tridecaketides griseorhodin A (C-26) and γ-rubromycin (C-26).

rise to the aromatic core structures due to the inherent high chemical reactivity of the poly-β-keto thioester intermediate and rudimentary understanding of how a type II PKS controls the chain length in polyketide biosynthesis.

Studies of ACT (an octaketide) and TCM (a decaketide) biosynthesis have revealed that they are produced from 8 or 10 malonyl-CoA condensations, respectively, and that the chain lengths were most likely controlled by the KS_α and KS_β heterodimer of the type II PKS.[9,10] Khosla and co-workers recently reported the site-directed mutagenesis study of ACT and TCM KS_β.[11] By changing two amino acids in the KS_α/KS_β heterodimer interface, they showed that the F109A/F116A ACT KS_β mutant produced 66% decaketide in vivo, while the G116T TCM KS_β mutant produced >65% octaketide. The identified amino acids, called "gatekeepers", were further confirmed by the X-ray crystallographic structure of ACT KS_α/KS_β.[12] It was therefore proposed that the seven amino acids at the KS_α/KS_β heterodimer interface constitute a tunnel, the depth of which controls the polyketide chain length. Most of the current studies on how a type II PKS controls polyketide chain length have been focused on the ACT and TCM PKSs, however, because the majority of the type II PKSs sequenced to date catalyze the formation of polyektides with chain length ranging mainly between 16 and 20 carbons with a few up to 26 carbons (Figure 1). Characterization of type II PKSs for polyketides with longer chains, such as fredericamycin (FDM) A (a pentadecaketide, Figure 2), will contribute to the investigation of how to control the chain length in aromatic polyketide biosynthesis and thus will facilitate combinatorial biosynthesis for aromatic polyketide structural diversity.

FDM A was first isolated from *Streptomyces griseus* ATCC 49344 in 1981,[13] and its structure was established by X-ray crystallographic analysis in 1982 after extensive spectroscopic studies failed to resolve the tautomeric structures.[14,15] Two minor metabolites, FDM B and C, were also reported in the original study[13] but their structures were not determined until 2004, together with the identification of an additional shunt metabolite, FDM C_1 (Figure 2).[16] Structurally, FDM A features two sets of *peri*-hydroxy tricyclic aromatic moieties connected through a chiral spiro carbon center, a molecular architecture that has not been found to date in any other natural product. The biosynthetic origin of FDM A was established by isotope labeling experiments, revealing that FDM A is a pentadecaketide, most likely

(9) MacDaniel, R.; Ebert-Kbosla, S.; Hopwood, D. A.; Khosla, C. *J. Am. Chem. Soc.* **1993**, *115*, 11671−11675.
(10) Shen, B.; Summers, R. G.; Wendt-Pienkowski, E.; Hutchinson, C. R. *J. Am. Chem. Soc.* **1995**, *117*, 6811−6821.
(11) Tang, Y.; Tsai, S. C.; Khosla, C. *J. Am. Chem. Soc.* **2003**, *125*, 12708−12709.
(12) Keatinge-Clay, A. T.; Maltby, D. A.; Medzihradszky, K. F.; Khosla, C.; Stroud, R. M. *Nat. Struct. Mol. Biol.* **2004**, *11*, 888−893.
(13) Pandey, R. C.; Toussaint, M. W.; Stroshane, R. M.; Kalita, C. C.; Aszalos, A. A.; Garretson, A. L.; Wei, T. T.; Byrne, K. M.; Geoghegan, R. F., Jr; White, R. J. *J. Antibiot.* **1981**, *34*, 1389−1401.
(14) Misra, R.; Pandey, R. C.; Hilton, B. D.; Roller, P. P.; Silverton, J. V. *J. Antibiot.* **1987**, *40*, 786−802.
(15) Misra, R.; Pandey, R. C. *J. Am. Chem. Soc.* **1982**, *104*, 4478−4479.
(16) Stontag, B.; Muller, J. G.; Hansske, F. G. *J. Antibiot.* **2004**, *57*, 823−828.

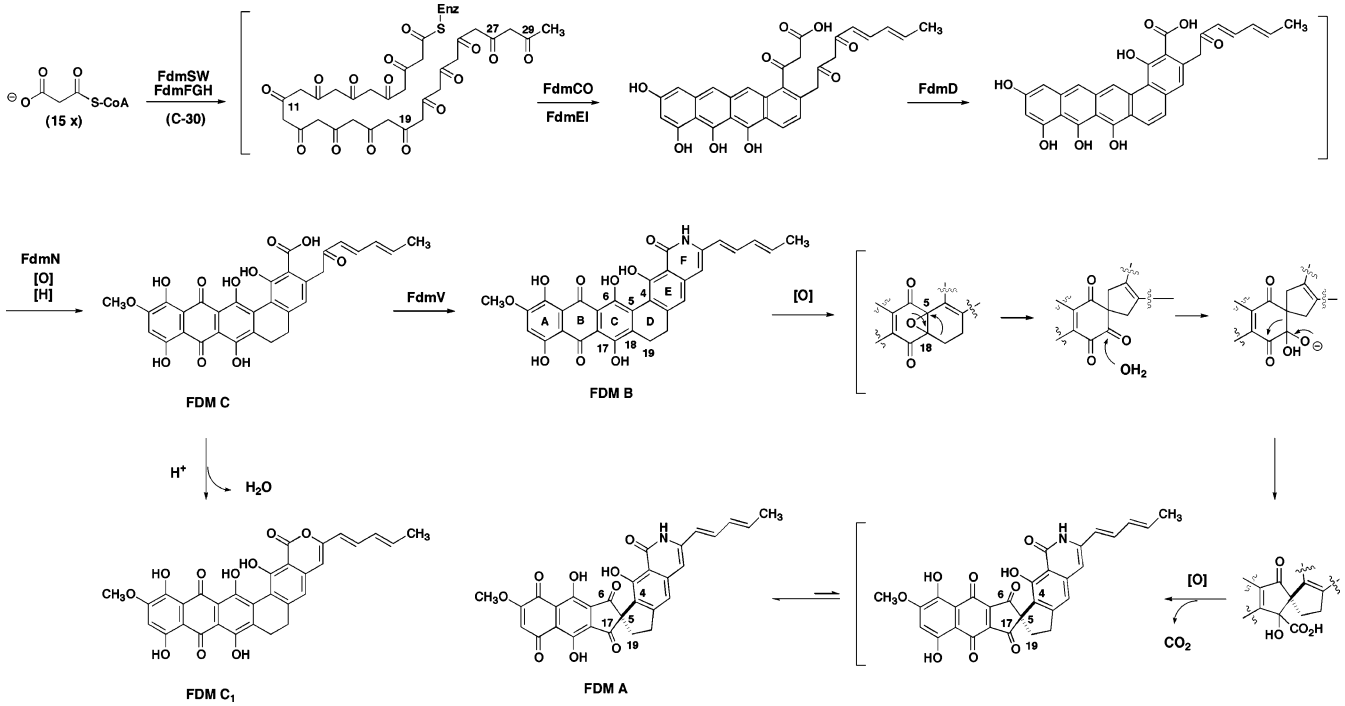


Figure 2. Proposed pathway for the fredericamycin (FDM) A biosynthesis featuring a type II PKS specifying a pentadecaketide (C-30) intermediate and novel oxygenases for the formation of a chiral spiro carbon center. Structures in brackets are hypothetical. [H], reduction; [O], oxidation.

derived from sequential condensations of 15 malonyl-CoAs with $C_{29}-C_{30}$ representing the starter acetate unit of the pentadecaketide chain (Figure 2).[17]

FDM A was reported to have moderately potent antibacterial and antifungal activities and potent in vivo anticancer activities against P388 mouse leukemia, CD8F mammary tumor, and B_{16} melanoma.[18] The unprecedented molecular architecture of FDM A clearly distinguishes it from all other antitumor agents, and it is quite possible that its mode of action involves unusual features that could be exploited for anticancer drug discovery. For example, FDM A inhibits both topoisomerase I and II at biologically relevant concentrations and additional DNA processing enzymes at higher concentrations.[19] FDM A is also known to form a stable oxidized free radical upon exposure to oxygen, which has been suggested to be another mechanism for its cytotoxicity.[20,21] More recently, FDM A was found to be an irreversible peptidyl-prolyl *cis*−*trans* isomerase (PPI) inhibitor.[22] This finding is exciting because Pin1, the human orthologue of the parvulin family of PPIs, has been shown to play an important role in mitotic regulation and oncogenesis and also to be a critical regulator of the tumor suppressor p53 during DNA damage response.[23−25] Pin1 could therefore represent a new anticancer target,[22] and FDM A could serve as an attractive lead for a novel type of anticancer drug.[26−28]

Here we report the cloning and sequencing of the *fdm* biosynthetic gene cluster from *S. griseus* ATCC 49344, the determination of the *fdm* cluster boundaries, and predictions for the functions of many of the gene products. The involvement of the cloned gene cluster in FDM A biosynthesis was confirmed by gene inactivation and complementation experiments in *S. griseus* and by heterologous expression of the *fdm* cluster to produce FDM A in *Streptomyces albus*. The *fdm* cluster consists of 28 open reading frames (ORFs), which encode the components of a type II PKS, tailoring enzymes, and several regulatory and resistance proteins. The FDM PKS features a KS_α subunit with tandem cysteines at its active site, a KS_β subunit that is distinct phylogenetically from the KS_β subunits of hexa-, octa-, or decaketide PKSs, and a dedicated phosphopantetheinyl transferase (PPTase).

Experimental Procedures

Bacterial Strains and Plasmids. *E. coli* DH5α, JM109, GM2929 were used for routine subcloning and plasmid preparations. *E. coli* XL1-Blue MR (Stratagene, La Jolla, CA) was used for cosmid library preparation. The pGEM series of vectors (Promega, Madison, WI), pSPORT1 (Invitrogen, Carisbad, CA), and Litmus 28, 38 (New England Biolabs, Beverly, MA) were from commercial sources, pKC1139,[29] pKC1218,[29] pOJ260,[29] and pSET151,[29] pIJ2926,[30] pWHM3[31] and pWHM1250,[31] the *aac(3)IV* apramycin resistance cassette,[30] and *S. albus* J1074[32] were described previously. A modified version of SuperCos 1 (Stratagene) that lacks the kanamycin resistance marker was used for cosmid library preparation. *E. coli* strains were grown in

(17) Byrne, K. M.; Hilton, B. D.; White, R. J.; Misra, R.; Pandey, R. C. *Biochemistry* **1985**, *24*, 478−486.
(18) Warnick-Pickle, D. J.; Byrne, K. M.; Pandey, R. C.; White, R. J. *J. Antibiot.* **1981**, *34*, 1402−1407.
(19) Latham, M. D.; King, C. K.; Gorycki, P.; Macdonald, T. L.; Ross, W. E. *Cancer Chemother. Pharmacol.* **1989**, *24*, 167−171.
(20) Hilton, B. D.; Misra, R.; Zweier, J. L. *Biochemistry* **1986**, *25*, 5533−5539.
(21) Dalal, N. S.; Shi, X. L. *Biochemistry* **1989**, *28*, 748−750.
(22) Lu, K. P.; Fischer, G. PCT Int. Appl. WO 2004002429, 2004.
(23) Shaw, P. E. *EMBO Rep.* **2002**, *3*, 521−526.
(24) Ryo, A.; Liou, Y.-C.; Lu, K. P.; Wulf, G. *J. Cell Sci.* **2003**, *116*, 773−783.
(25) Miyashita, H.; Uchida, T.; Mori, S.; Echigo, S.; Motegi, K. *Oncol. Rep.* **2003**, *10*, 1045−1048.
(26) Simon, W.; Abel, U. PCT Int. Appl. WO 2004004713, 2004.
(27) Simon, W.; Abel, U. PCT Int. Appl. WO 03087060, 2003.
(28) Abel, U.; Simon, W. PCT Int. Appl. WO 03080582, 2003.
(29) Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Rao, R. N.; Schoner, B. E. *Gene* **1992**, *116*, 43−49.
(30) Kieser, T.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Steptomyces Genetics*; The John Innes Foundation: Norwich, 2000.
(31) Lomovskaya, N.; Otten, S. L.; Doi-Katayama, Y.; Fonstein, L.; Liu, X.-C.; Takatsu, T.; Inventi-Solari, A.; Filippini, S.; Torti, F.; Colombo, A. L.; Hutchinson, C. R. *J. Bacteriol.* **1999**, *181*, 305−318.
(32) Sanchez, C.; Zhu, L.; Brana, A. F.; Salas, A. P.; Rohr, J.; Mendez, C.; Salas, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 461−466.

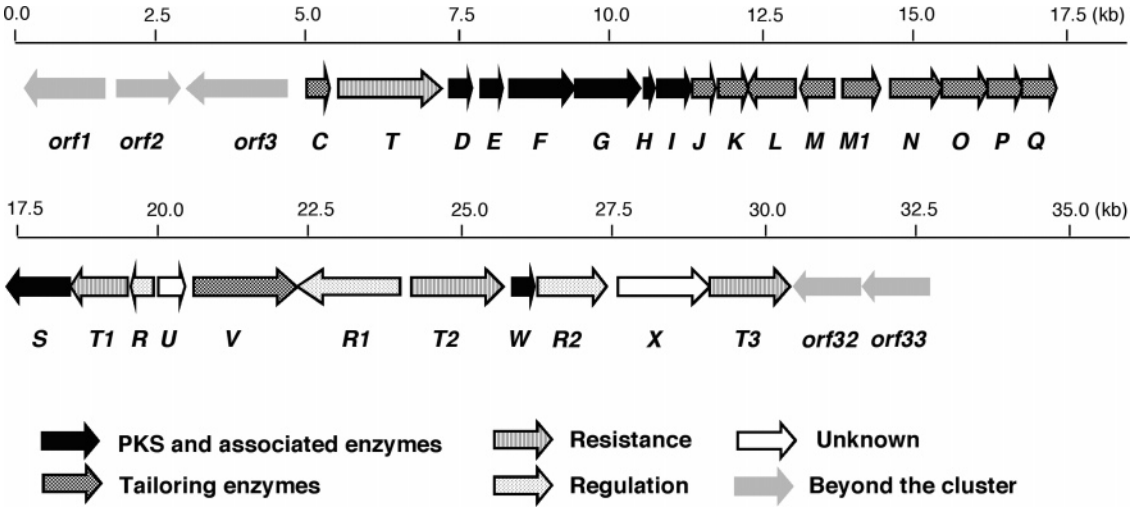


Figure 3. Genetic organization of the *fdm* biosynthetic gene cluster from *S. griseus* ATCC 49344. Proposed functions for individual ORFs are coded with various patterns and summarized in Table 1.

Luria−Bertani (LB) medium with appropriate antibiotics where necessary. *S. griseus* ATCC 49344 was purchased from ATCC (American type Culture Collection, Rockville, MD).

Biochemicals, Chemicals, and Media. Common biochemicals and chemicals were from standard commercial sources unless specified otherwise. Soft nutrient agar (SNA) and R2YE were prepared according to standard protocol.[30] R2YE was used for routine *S. griseus* culture growth, transformation, sporulation, and seed culture. Modified antibiotic production medium (APM) for FDM A production by *S. griseus* consisted of 8 g of yeast extract, 20 g of malt extract, 2 g of NaCl, 10 g of MOPS sodium salt in 500 mL of tap water and 380 mL of Milli-Q water (Millipore Corp, Bedford, MA). After autoclaving, 1 mL of 10% $MgSO_4$, 1 mL of 1% $FeSO_4$, 0.1 mL of 10% $ZnSO_4$ and 120 mL of 50% glucose were added.

Production, Isolation and Analysis of FDM A. *S. griseus* spores were inoculated into 5 mL of R2YE and grown overnight. The resultant seed culture (1 mL) was transferred to the modified APM (50 mL) in a 250-mL baffled Erlenmeyer flask for FDM A production. Flasks were incubated in a rotary incubator at 300 rpm for 6 days at 28 °C. To isolate FDM A, the culture broth was first adjusted to pH 2 with 2 N HCl, and the mycelia were harvested by centrifugation. The mycelia were extracted with acetone (3 × 10 mL), and the combined extract was used directly for HPLC analysis. The HPLC analysis was carried out on an Alltima C-18 column (5 μ, 250 × 4.6 mm, Alltech, Deerfield, IL). The column was equilibrated with 50% solvent A (1% AcOH in H_2O) and 50% solvent B (1% AcOH in CH_3CN) and developed with the following program (0−5 min, 50% A/50% B; 5−15 min, a linear gradient from 50% A/50% B to 5% A/95% B; 15−25 min, 5% A/95% B) at flow rate of 0.8 mL/min and UV detection at 375 nm using a Varian Prostar 330 PDA detector (Varian, Palo Alto, CA). The identity of FDM A was confirmed by ESI−MS analysis on an Agilent 1000 HPLC−MSD SL instrument (Agilent Technologies, Palo Alto, CA), MALDI−FT-ICR MS analysis on an IonSpec HiRes MALDI-FT-Mass spectrometer (IonSpec Corp., Lake Forest, CA), and NMR analysis on a Varian Unity Inova 500 MHz instrument operating at 500 MHz for ^1H and 125 MHz for ^{13}C nuclei in $CDCl_3$, in comparison with a FDM A authentic standard.[17] The isolated yield for FDM A from the wild-type *S. griseus* strain was ∼170 mg/L.

Cloning, Isolation and Sequencing of the *fdm* Biosynthetic Gene Cluster. PCR primers for KS_α (KS−FP: 5′-AGC TCC ATC AAG TCC/G ATG A/GTC GG-3′) and KS_β (KS−RP: 5′-CC GGT GTT C/GAC C/GGC GTA GAA CCA GGC G-3′) were designed and used to amplify putative type II PKS genes from *S. griseus* genomic DNA. A distinct 600-bp DNA fragment was obtained by PCR, cloned and sequenced, and labeled using the digoxigenin (DIG) DNA labeling kit (Roche Diagnostics Corp., Indianapolis, IN). This was used as a probe in a mini chromosomal library screening of the *S. griseus* DNA. A 7.2-kb *Bam*HI-*Xho*I fragment was isolated and cloned into the same sites of pSPORT1 to afford pBS4012, which was subsequently sequenced. Chromosomal walking was done to extend this region in both directions, and a cosmid library was concurrently constructed in a modified version of SuperCos1 in order to clone the entire *fdm* cluster. To construct the cosmid library, chromosomal DNA was partially digested with *Sau*3AI and dephosphorylated using calf intestine alkaline phosphatase. The vector was prepared according to manufacturer's instructions, and ligation was performed using T4 DNA ligase (Promega) at 16 °C overnight. The ligation mixture was packaged with Gigapack III XL packaging extract according to manufacturer's instructions (Stratagene) and used to transfect *E. coli* XL1 Blue MR cells. DIG-labeled probes were prepared (Roche) and colony and Southern hybridizations were performed on Hybond N (Amersham, Little Chalfont, Bucks, UK) as recommended by the manufacturer.

DNA Sequencing and Analysis. Both DNA strands of selected cosmids and subclones were sequenced either by primer walking or using M13 universal primers or primers designed for SuperCos1. Sequencing reactions were run using Big Dye Terminator mix (Applied Biosystems, Foster City, CA) and cleaned using CleanSeq magnetic beads (Agencourt Biosciences, Beverly, MA). DNA sequencing was carried out on an Applied Biosynthesis 3700 automated DNA sequencer at the University of Wisconsin Biotechnology Center. Data were analyzed using PE-Biosystems version 3.7 of Sequencing Analysis. ORF assignments were made with the assistance of Codon Preference (GCG package, Madison, WI) and BLAST analyses. The nucleotide sequence reported in this paper has been deposited in GenBank database under accession number AF525490.

Inactivation by in-frame Deletion and Complementation. To inactivate the FDM PKS by in-frame deletion, a 7.2-kb *Bam*HI-*Eco*RI fragment was cut from pBS4012 and further digested with *Xmn*I. The resultant 1.7-kb *Bam*HI-*Xmn*I and 1.4-kb *Xmn*I-*Eco*RI fragments were isolated and co-ligated into the *Bam*HI and *Eco*RI site of pKC1139 to yield pBS4013, yielding an in-frame deletion of a 4.1-kb *Xmn*I fragment that contains the C-terminal portion of *fdmE*, all of *fdmFGHIJ*, and the N-terminal portion of *fdmK* (Figure 4A). pBS4013 was transferred into wild-type *S. griseus* by protoplast-mediated transformation.[30] Transformants were selected by overlaying plates with 3 mL of SNA containing thiostrepton (25 μg/mL) at 28 °C. Once transformants were visible on the plates, the incubation temperature was shifted to 39 °C to force integration of pBS4013 into the chromosome via homologous recombination. Single crossover mutants were picked and grown nonselectively for several rounds before screening for a thiostrepton

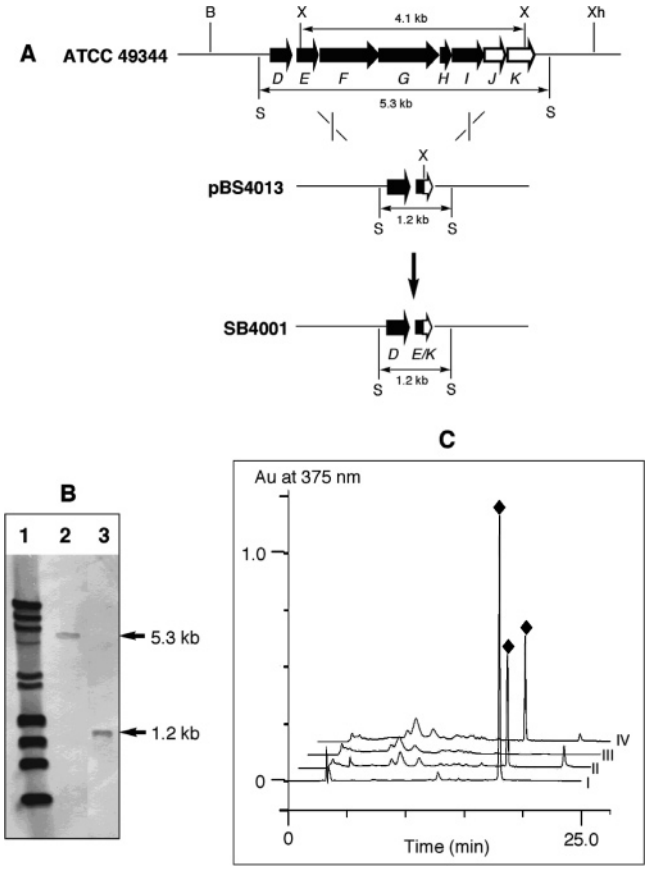


Figure 4. Confirmation of the cloned cluster to encode FDM A biosynthesis by inactivation of the FDM PKS via in-frame deletion. (A) Isolation of the FDM PKS in-frame deletion mutant SB4001 and restriction maps of the *S. griseus* ATCC49344 and SB4001 mutant strains showing predicted fragment sizes upon *Srf*I digestion. S, *Srf*I. (B) Southern analysis of wild-type (lane 2) and SB4001 (lane 3) genomic DNAs digested with *Srf*I using a 458-bp *Xmn*I-*Srf*I fragment of *fdmK/L* as a probe. Lane 1, molecular weight markers. (C) HPLC analysis of FDM A (◆) authentic standard (I) and production isolated from wild-type (II), SB4001 (III), and SB4005 [i.e., SB4001 (pBS4015)] (IV).

sensitive double crossover mutant to afford the SB4001 mutant strain, the genotype of which was confirmed by Southern analysis (Figure 4B). To make the FDM PKS expression construct, a 7.1-kb *Bam*HI-*Kpn*I fragment from pBS4012 and a 300-bp *Eco*RI-*Bam*HI fragment that harbored the *ErmE** promoter from pWHM1250 were cloned into the *Eco*R1 and *Kpn*I site of pKC1218 to afford pBS4015. The latter was introduced into *S. griseus* SB4001 by protoplast-mediated transformation to complement the Δ*fdmEFGHIJK* deletion, yielding strain SB4005. Recombinant strains were cultured and analyzed for FDM A production by HPLC and ESI-MS analyses with the wild-type *S. griseus* strain as a control (Figure 4C). The isolated yield for FDM A from the *S. griseus* SB4005 recombinant strain was ∼170 mg/L.

Inactivation by Gene Replacement and Determination of the *fdm* Cluster Boundaries. To inactivate *orf1*, a 2.4-kb *Kpn*I-*Sph*I DNA fragment containing *orf1* was cloned into the same sites of pIJ2926 to yield pBS4016. A blunt ended 1.5-kb *aac(3)IV* apramycin resistance gene was then cloned into the *Rsr*II site (Klenow treated) within *orf1* of pBS4016 to afford pBS4017. The inactivated *orf1* was moved as a 3.9-kb *Bgl*II fragment into the *Bam*HI site of pSET151 to finally afford pBS4018, in which *orf1* was insertionally inactivated by *aac(3)IV* (Figure 5A).

To inactivate *orf3*, a 3.2-kb *Mlu*I-*Xho*I fragment containing *orf3* was first cloned into the *Mlu*I-*Sal*I sites of Litmus 38, yielding pBS4020. A 1.5-kb *Kpn*I-*Nsi*I fragment of the *aac(3)IV* apramycin resistance gene was then cloned into the same sites of pBS4020, replacing an 850-bp *Kpn*I-*Nsi*I internal fragment of *orf3*, to afford pBS4021. The inactivated *orf3* was moved as 3.8-kb *Hpa*I-*Eco*RI fragment into the *Eco*RV and *Eco*RI sites of pOJ260 to yield pBS4022 and finally as a 3.8-kb *Eco*RI-*Xba*I fragment into the same sites of pSET151 to afford pBS4023, in which *orf3* was partially replaced by *aac(3)IV* (Figure 5B).

To inactivate *orf33*, a 3.2-kb *Pst*I-*BamH*I DNA fragment containing *orf33* was cloned into the same sites of Litmus 38 to afford pBS2025. A blunt ended 1.5-kb *aac(3)IV* apramycin resistance gene was then cloned into the *Sma*I site within *orf33* of pBS4025 to afford pBS4026. The inactivated *orf33* was moved as a 4.5-kb *Sph*I fragment into the same site of pSET151 to finally yield pBS4027, in which *orf33* was insertionally inactivated by *aac(3)IV* (Figure 5C).

pBS4018, pBS4023, or pBS4027 was introduced into the wild-type *S. griseus* strain by protoplast-mediated transformation. The desired double-crossover homologous recombination mutant strains were selected by using the apramycin-resistant and thiostrepton sensitive phenotype. Thus, pBS4018, pBS4023, or pBS4027 resulted in the isolation of mutant strains SB4002 (*orf1::aac(3)IV*) (Figure 5A), SB4003 (*orf3::aac(3)IV*) (Figure 5B), or SB4004 (*orf33::aac(3)IV*) (Figure 5C), respectively, whose genotypes were further confirmed by Southern analysis (Figure 5D). These mutant strains were cultured and analyzed for FDM A production by HPLC and ESI−MS analyses with the wild-type *S. griseus* strain as a control (Figure 5E). The isolated yields for FDM A from the *S. griseus* SB4002, SB4003, and SB4004 mutant strains were ∼170, ∼90, and ∼90 mg/L, respectively.

Heterologous Expression of the *fdm* Cluster in *S. albus*. To make the expression construct for the entire *fdm* cluster, a 25-kb fragment harboring *fdmC* through *fdmT3*, from the *Csp*45I site (approximately 70 bp upstream of *fdmC*) to the *Bgl*II site (approximately 250 bp downstream of *fdmT3*), was assembled in multiple steps and finally inserted into the multiple cloning site of pWHM3 to yield pBS4028. This plasmid was then introduced into *S. albus* J1074 via standard protoplast transformation. Transformants were selected by overlaying the plates with SNA containing thiostrepton to afford recombinant *S. albus* SB4006, which carries pBS4028, and the negative control *S. albus* SB4007, which carries the pWHM3 vector. Both *S. albus* SB4006 and SB4007 were cultured under the identical conditions as the wild-type *S. griseus* strain and analyzed for FDM A production by HPLC analysis with the wild-type *S. griseus* as a positive control (Figure 5E). The isolated yield for FDM A from the *S. albus* SB4006 recombinant strain is ∼120 mg/L.

Results

Cloning and Sequencing of the *fdm* Gene Cluster. Since the KS_α and KS_β subunits share high sequence similarity and the KS_α and KS_β genes always reside next to each other,[1,2] two highly conserved regions, SSIKSM(V/I)G in the KS_α and TFAWFYAVN (S/T) in the KS_β, were used to design primers, KS−FP and KS−RP, to clone the FDM PKS genes by PCR. A distinct product of about 600 bp was readily amplified from the *S. griseus* chromosomal DNA, cloned and sequenced, to reveal part of a putative type II PKS gene set. This DNA fragment was then used as a probe to screen an *S. griseus* mini-chromosomal library, from which a 7.2 kb *BamH*I-*Xho*I DNA fragment was isolated and cloned as pBS4012. Sequence analysis of the latter showed that it contained genes encoding typical type II PKS enzymes such as KS_α, KS_β, and ACP. After this locus was confirmed to encode genes essential for FDM A biosynthesis by gene inactivation and complementation experiments, a combination of chromosomal walking and cosmid library screening was done to obtain the full gene cluster. DNA sequencing of the cloned 33-kb DNA region revealed 33 ORFs (Figure 3), 28 of which were assigned to the *fdm* cluster with the proposed functions summarized in Table 1.

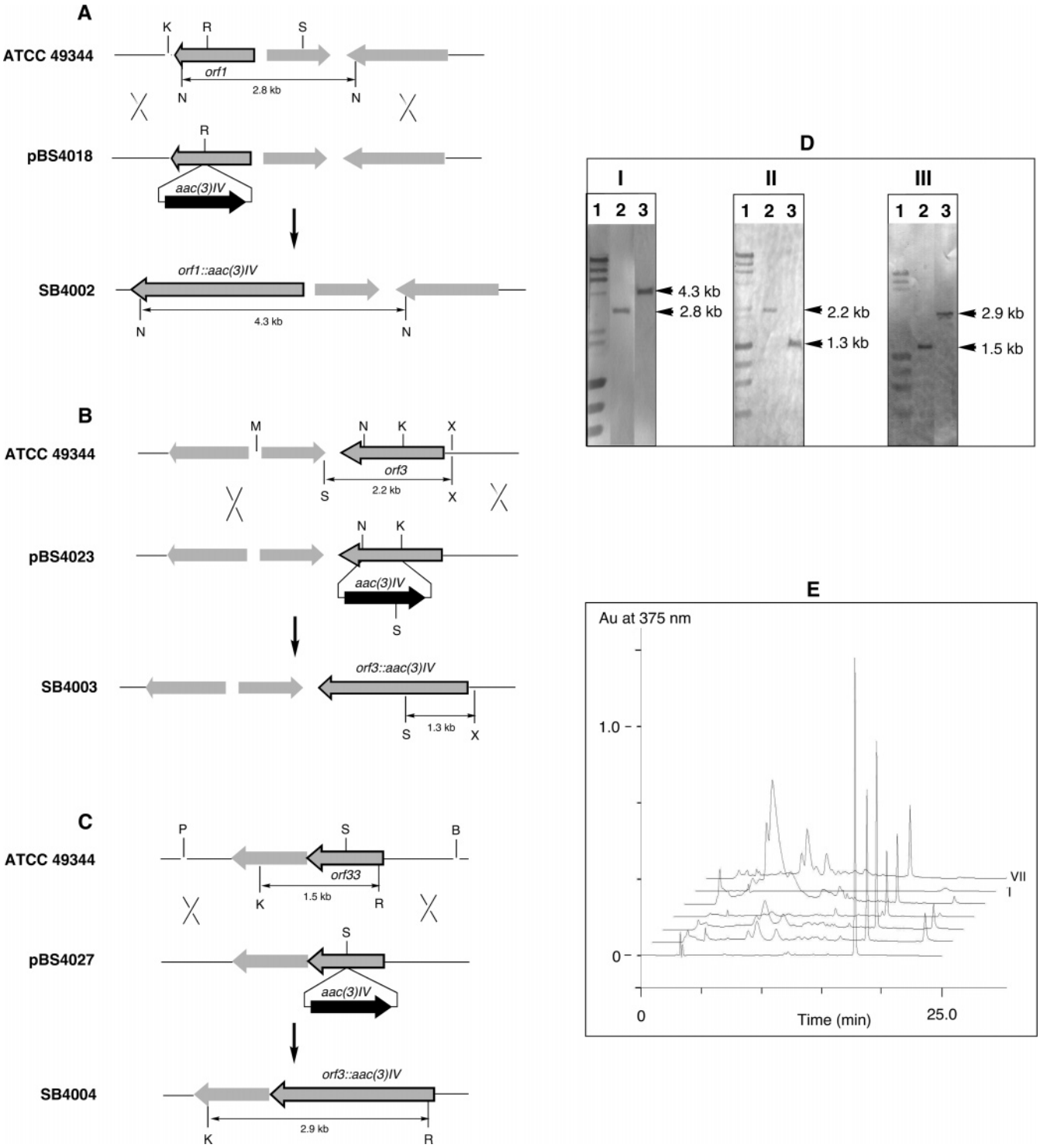


Figure 5. Determination of the *fdm* biosynthetic gene cluster boundaries by inactivation of *orf1*, *orf3*, and *orf33* and by heterologous expression of the *fdm* cluster (from *fdmC* to *fdmT3*) in *S. albus*. (A) The *orf1* gene replacement construct pBS4018 and restriction maps of the *S. griseus* ATCC wild-type and SB4002 mutant strains showing predicted fragment sizes upon *Nsi*I digestion. K, *Kpn*I; N, *Nsi*I; R, *Rsr*II; S, *Sph*I. (B) The *orf3* gene replacement construct pBS4023 and restriction maps of the *S. griseus* ATCC wild-type and SB4003 mutant strains showing predicted fragment sizes upon *Sst*I/*Xho*I digestion. K, *Kpn*I; M, *Mlu*I; N, *Nsi*I, S, *Sst*I; X, *Xho*I. (C) The *orf33* gene replacement construct pBS4027 and restriction maps of the *S. griseus* ATCC wild-type and SB4004 mutant strains showing predicted fragment sizes upon *Kpn*I/*Sma*I digestion. B, *Bam*HI; K, *Kpn*I; P, *Pst*I; R, *Rsr*II; S, *Sma*I. (D) Southern analysis, lane 1, molecular weight markers. (I) Wild-type (lane 2) and SB4002 (lane 3) genomic DNAs digested with *Nsi*I with the 600-bp *Nis*I-*Rsr*II fragment of *orf1* as a probe. (II) Wild-type (lane 2) and SB4003 (lane 3) genomic DNAs digested with *Sst*I/*Xho*I with the 1.1-kb *Kpn*I-*Xho*I fragment of *orf3* as a probe. (III) Wild-type (lane 2) and SB4004 (lane 3) genomic DNAs digested with *Kpn*I/*Rsr*II with the 850 bp fragment of *Kpn*I-*Sma*I fragment of *orf32-orf33* as a probe. (E) HPLC analysis of FDM A (◆) authentic standard (I) and production isolated from wild-type (II), SB4003 [III, *orf1::aac(3)IV*], SB4004 [IV, *orf3::aac(3)IV*], SB4005 [V, *orf33::aac(3)IV*], SB4007 [VI, *S. albus* (pWHM3)], and SB4006 [VII, *S. albus* (pBS4028)].

Confirmation of the Cloned Locus for FDM A Biosynthesis by in-frame Deletion and Complementation. To confirm that the cloned genes were essential for FDM A biosynthesis, an in-frame deletion mutant of *S. griseus* was constructed. The pKC1139 based plasmid pBS4013 was created by removing an internal 4130 bp *Xmn*I-*Xmn*I fragment from

Table 1. Deduced Function of ORFs in the *fdm* Biosynthetic Gene Cluster

gene	size[a]	protein homologues and the accession numbers	proposed function
orf1−*orf3*			ORFs beyond the upstream boundary
fdmC	266	SimJ2 (AAK06809), SAV3653 (BAC71365)	3-Ketoacyl ACP reductase
fdmT	493	SCE25.05 (CAB89436), RUB−ORF3 (AAM97376)	peptide transporter
fdmD	106	TcmI (P39890), GrhQ (AAM33678)	polyketide cyclase
fdmE	141	TcmJ (AAA67514), GrhS (AAM33682)	polyketide cyclase
fdmF	421	TcmK (AAA67515), GrhA (AAM33653)	KS_α
fdmG	403	TcmL (AAA67516), GrhB (AAM33654)	KS_β
fdmH	84	TcmM (P12884), GrhC (AAM33655)	ACP
fdmI	155	TcmN (AAA67518), GrhT (AAM33685)	polyketide cyclase
fdmJ	107	GrhU (AAM33683), RubH (AAG03072)	oxygenase
fdmK	144		oxygenase
fdmL	246		oxygenase
fdmM	149	GrhM (AAM33665), RubQ (AAM97373)	oxygenase
fdmM1	154	GrhM (AAM33665), PRM-ORF5 (BAA23148)	oxygenase
fdmN	353	TcmN (AAA67518), PRM-ORF6 (BAA23149)	O-methyltransferase
fdmO	249	GrhO10 (AAM33668), PRM-ORF7 (BAA23150)	3-Ketoacyl ACP reductase
fdmP	103	GrhU (AAM33683), PRM-ORF8 (BAA23151)	oxygenase
fdmQ	105	PRM-ORF9 (BAA23152), GrhV (AAM33684)	oxygenase
fdmS	333	ZhuH (AAG30195), HedS (AAA65208)	3-Ketoacyl ACP synthase
FdmT1	343	RUB-ORF3(AAM97376), SCM2.02c (CAB65630)	transporter
fdmR	143	GrhR3 (AAM33681), RubM (AAM97363)	regulator
fdmU	141	−	unknown
fdmV	622	RubR (AAM97368), TcsG (BAB12569)	asparagine synthetase
fdmR1	612	GrhR2 (AAM33680), RubS (AAM97369)	regulator
fdmT2	531	GrhK (AAM33663), TcmA (A41901)	transporter
fdmW	159	SC6G4.22C (CAA20400), *E. coli* ACPS (B42294)	holo-ACP synthase
fdmR2	170	SCG11A.22c (CAB61603), GrhR3 (AAM33681)	regulator
fdmX	405	SC3C8.10 (NP 625301), S0055 (CAD33753)	transposase
fdmT3	525	TcmA (A41901), GrhK (AAM33663)	transporter
orf32−*orf33*			ORFs beyond the upstream boundary

[a] Numbers are in amino acids.

the original 7.2-kb *BamH*I-*Xho*I fragment encoding the KS_α, KS_β, and ACP subunits (Figure 4A). pBS4013 was transferred into wild-type *S. griseus* by protoplast-mediated transformation. A double crossover mutant was selected and its genotype was confirmed by Southern analysis (Figure 4B). The resulting mutant, SB4001, completely lost its ability to produce FDM A when grown under the conditions used for FDM A production by the wild-type strain, as evidenced by HPLC analysis (Figure 4C). A complementation construct, pBS4015, was generated in pKC1218 by placing the *ErmE** promoter in front of a 7.1-kb *BamH*I-*Kpn*I fragment from pBS4012 to ensure efficient expression of the KS_α, KS_β, and ACP subunits. This plasmid was then introduced by transformation into SB4001, and restoration of FDM A production was confirmed by HPLC and UV−vis analyses with authentic FDM A as a standard (Figure 4C) and by electron spray ionization-mass spectrometry (ESI−MS) analysis, yielding the characteristic molecular ions (*m*/*e* for $[M + H]^+$) of 540, and MALDI−FT-ICR MS analysis, yielding (*m*/*e* for $[M + H]^+$) of 540.1136, consistent with the molecular formula of FDM A ($C_{30}H_{21}NO_9$, calcd. 539.1216). These results confirmed that the cloned gene cluster is required for FDM A biosynthesis.

Determination of the *fdm* Gene Cluster Boundaries by Gene Inactivation and Heterologous Expression in *Streptomyces albus.* The *fdm* gene cluster boundaries were defined through inactivation of selected genes (*orf1*, *orf3*, and *orf33*) residing at the distal ends of the sequenced region (Figure 3). Thus, for the upstream boundary, *orf1* encodes a putative acyl-CoA decarboxylase, which shows high sequence similarity to GhrG (74% identity)[33] and JadN (69% identity).[34] *Orf2* encodes an acyltransferase with high sequence similarity to known acyltransferases found in several aromatic PKS gene clusters, such as DspD (42% identity)[35] and ZhuC (43% identity).[6] The gene product of *orf3* shows marginal sequence similarity to CumA (26% identity), a known Mn(II)-oxidation-associated multicopper oxidase that exists in diverse *Pseudomonas* strains[36] and DHGO (41% identity), a dihydrogeodin oxidase that is responsible for the conversion of dihydrogeodin to geodin.[37] For the downstream boundary, both *orf32* and *orf33* encode proteins of unknown function. *Orf1*, *orf3*, and *orf33* were inactivated by gene replacement (Figure 5A−C), and the resultant mutant strains SB4002 (*orf1::aac(3)IV*), SB4003 (*orf3::aac(3)IV*), and SB4004 (*orf33::aac(3)IV*) were confirmed by Southern analysis (Figure 5D) and fermented for FDM A production with the wild-type strain as a control (Figure 5E). HPLC analysis of the fermentation culture showed that inactivation of *orf1*, *orf3*, or *orf33* had little or insignificant effect on FDM A production (Figure 5E, panels II to V), indicating that these genes are either outside of the *fdm* gene cluster or dispensable under the growth conditions used.

To ascertain that the FDM biosynthetic machinery does not recruit genes residing somewhere else in the *S. griseus* chromosome and that the genes within the defined cluster boundaries are sufficient for FDM A biosynthesis, pBS4028, which carries a 25-kb fragment of the *fdm* cluster harboring *fdmC* to *fdmT3*,

(33) Li, A.; Piel, J. *Chem. Biol*. **2002**, *9*, 1017−1026.
(34) Han, L.; Yang, K.; Kulowski, K.; Wendt-Pienkowski, E.; Hutchinson, C. R.; Vining, L. C. *Microbiology* **2000**, *146*, 903−910.
(35) Bao, W.; Sheldon, P. J.; Wendt-Pienkowski, E.; Hutchinson, C. R. *J. Bacteriol*. **1999**, *181*, 4690−4695.
(36) Francis, C. A.; Tebo, B. M. *Appl. Environ. Microbiol*. **2001**, *67*, 4272−4278.
(37) Huang, K. X.; Fujii, I.; Ebizuka, Y.; Gomi, K.; Sankawa, U. *J. Biol. Chem*. **1995**, *270*, 21495−21502.

was introduced into *S. albus* to demonstrate FDM A production in a heterologous host. HPLC analysis of the fermentation culture showed that the resultant *S. albus* SB4006 strain, which carried the pBS4028 expression construct, produced FDM A in an excellent yield, while FDM A was completely absent from *S. albus* SB4007 negative control strain (Figure 5E, panels VI and VII). The identity of FDM A produced was verified by HPLC and UV−vis analyses with authentic FDM A as a standard.

Given the fact that genes encoding the production of secondary metabolites are often clustered in one region of the microbial chromosome,[1] these results, together with the deduced functions of the sequenced region (Table 1), allowed the assignment of the *fdm* gene cluster to span 29-kb, including 28 ORFs from *fdmC* to *fdmT3* (Figure 3). Among them, 8 are biosynthetic genes encoding PKS and associated enzymes (*fdmD*, *fdmE*, *fdmF*, *fdmG*, *fdmH*, *fdmI*, *fdmS*, *fdmW*), 11 genes encoding tailoring enzymes (*fdmC*, *fdmN*, *fdmJ*, *fdmK*, *fdmL*, *fdmM*, *fdmM1*, *fdmO*, *fdmP*, *fdmQ*, *fdmV*), 3 regulatory genes (*fdmR*, *fdmR1*, and *fdmR2*), 4 resistance genes (*fdmT*, *fdmT1*, *fdmT2*, and *fdmT3*), and 2 additional genes (*fdmU*, *fdmX*) whose deduced functions cannot be predicted or assigned for FDM A production on the basis of sequence comparison alone (Table 1).

Genes Encoding PKS and Associated Enzymes. The minimal FDM PKS consists of FdmF (KS_α), FdmG (KS_β), and FdmH (ACP), which share high sequence similarity to known type II PKSs, such as ACT PKS (KS_α, 58% identity and KS_β, 52% identity), TCM PKS (KS_α, 58% identity and KS_β, 52% identity), pradimicin (PRM) PKS (KS_α, 65% identity and KS_β, 55% identity), rubromycin (RUB) PKS (KS_α, 60% identity and KS_β, 55% identity), and griseorhodin (GRH) PKS (KS_α, 58% identity and KS_β, 55% identity). However, unlike the others, FdmF contains two cysteines, Cys170 and Cys171, at the presumed KS_α active sites. All other KS_αs have the highly conserved sequence STGCTSGLD around the active site amino acid Cys in contrast to SSGCCAGID for FdmF. The role of the tandem cysteines and the reason for the differences in the conserved sequence around the FdmF active site remain unclear. FdmG contains the characteristic Gln162 for KS_β and is more closely related to KS_β found for the sub-group of PKSs that synthesize polyketides with longer chains, such as the dodecaketide PRM A and the tridecaketides GRH A and γ-RUB (Figures 1 and 6).[33,38,39] The *fdmH* gene is 64 bases downstream of *fdmG*, encoding an ACP with the characteristic conserved G(x)DS motif for attachment of the phosphopantetheinyl prosthetic group by a PPTase.[40]

Three putative polyketide cyclases, FdmD, FdmI, and FdmE, are found within the *fdm* gene cluster. FdmI shares high sequence similarity to the N-terminal part of the bifunctinal protein TcmN (62% identity) and might be responsible for initial folding of the nascent linear polyketide intermediate at the C_{11} position and subsequent cyclizations involved in formation of the aromatic rings A−D of FDM A.[41] FdmD shares high sequence similarity to TcmI (60% identity), supporting its role in the cyclization of the E ring.[42] FdmE shares high sequence similarity to TcmJ (53% identity). Although the function of TcmJ is not completely understood, it has been shown that the presence of TcmJ seems to enhance TCM production, possibly through facilitating the formation of the TCM PKS complex.[43] A similar role could be proposed for FdmE to facilitate the formation of the FDM PKS complex. Interestingly, the organization of the genes encoding the FDM minimal PKS and its associated cyclases is identical to that for TCM, with *fdmE* and *fdmIJ* flanking the *fdmFGH* genes encoding the minimal KS_α, KS_β, and ACP subunits (Figures 2 and 3).

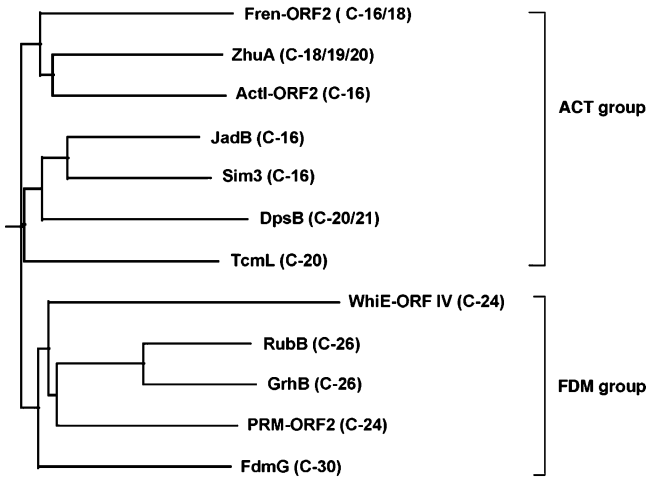


Figure 6. Phylogenetic analysis of the FdmE KS_β and its homologues . PDB ID codes: ActI−ORF2, Q02062; JadB, AAB36563; Sim3, AAL15581; Fren-ORF2, AAA19617; ZhuA, AAG30188; TcmL, AAA67516; DpsB, AAA65207; WhiE-ORF IV, CAA39409; PRM-ORF2, BAA23145; RubB, AAG03068; GrhB, AAM33654.

Two putative accessory PKS genes were identified from *fdm* gene cluster. FdmS is a FabH homologue and shares high sequence similarity to the ketoacyl synthases from several known aromatic polyketide pathways, such as the ZhuH (53% identity) from R1128 gene cluster. ZhuH is involved in biosynthesis of the alkylacyl intermediate for priming the R1128 PKS that utilizes starter units other than acetate.[6] A similar mechanism involving a FabH ortholog and an acyltransferase was also proposed for hedamycin biosynthesis.[44] FdmW shares high sequence similarity to ACPS-type PPTases involved in fatty acid biosynthesis pathways in *Streptomyces coelicolor* (53% identity)[45] and other bacterial species such as *Escherichia coli* (26% identity).[46] FdmW also shares high sequence similarity to RemD (45% identity), a putative ACPS-type PPTase from the resistomycin biosynthetic gene cluster.[3] The co-localization of *fdmW* and *remD* with aromatic polyketide gene clusters suggests that they presumably encode dedicated PPTases for the posttranslational modification of the apo-ACP FdmH or RemC, respectively.

(38) Dairi, T.; Hamano, Y.; Igarashi, Y.; Furumai, T.; Oki, T. *Biosci. Biotechnol. Biochem.* **1997**, *61*, 1445−1453.
(39) Martin, R.; Sterner, O.; Alvarez, M. A.; de Clercq, E.; Bailey, J. E.; Minas, W. *J. Antibiot.* **2001**, *54*, 239−249.
(40) Lambalot, R. H.; Gehring, A. M.; Flugel, R. S.; Zuber, P.; LaCelle, M.; Marahiel, M. A.; Reid, R.; Khosla, C.; Walsh, C. T. *Chem. Biol.* **1996**, *3*, 923−936.
(41) Shen, B.; Hutchinson, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6600−6604.
(42) Shen, B.; Hutchinson, C. R. *Biochemistry* **1993**, *32*, 11149−11154.
(43) Bao, W.; Wendt-Pienkowski, E.; Hutchinson C. R. *Biochemistry* **1998**, *37*, 8132−8138.
(44) Bililign, T.; Hyun, C. G.; Williams, J. S.; Czisny, A. M.; Thorson, J. S. *Chem. Biol.* **2004**, *11*, 959−969.
(45) Cox, R. J.; Crosby, J.; Daltrop, O.; Glod, F.; Jarzabek, M. E.; Nicholson, T. P.; Reed, M.; Simpson, T. J.; Smith, L. H.; Soulas, F.; Szafranska, A. E.; Westcott, J. *J. Chem. Soc., Perkin Trans. 1* **2002**, *14*, 1644−1649.
(46) Lambalot, R. H.; Walsh, C. T. *J. Biol. Chem.* **1995**, *270*, 24658−24661.

Genes Encoding Tailoring Enzymes or Proteins of Unknown Function. Similar to GRH or RUB, which are considered some of the most heavily oxidized bacterial polyketides known to date,[33] the *fdm* gene cluster contains genes for two known reductases and numerous putative novel oxygenases. The two ketoacyl ACP reductases (KR), FdmC and FdmO, might be responsible for reduction of the nascent polyketide intermediate (Figure 2). FdmO shares high sequence similarity to KRs from clusters encoding longer polyketide chains such as the dodecaketide PRM A (PRM-ORF7, 51% identity), the tridecaketides γ-RUB (RubG, 58% identity and RubF, 48% identity) and GRH A (GrhO10, 59% identity and the KR part of a bifuntional protein GrhT, 47% identity). It shares relatively low sequence similarity to KRs from aromatic polyketide clusters encoding shorter polyketide chains such as the octaketide ACT (ActIII, 37% identity).[47] In contrast, FdmC is more similar to KR SAV3653 (54% identity), identified within the putative *pks*8 polyketide biosynthetic gene cluster from *Streptomyces avermitilis*,[48] and SimJ2 (48% identity) from *Streptomyces antibioticus,* which is involved in simocyclinone biosynthesis,[49] as compared to GrhO10 (39% identity) or RubG (40% identity). These KRs, FdmC, SAV3653, and SimJ2, are only distantly related to KRs found in other aromatic polyketide biosynthetic pathways and have been proposed to reduce a polyketide intermediate into a polyunsaturated chain. This supports the role of FdmC in the biosynthesis of the pentadienyl side chain of FDM A (Figure 2).

FdmN shares high sequence similarity to the C-terminal region of TcmN (47% identity), which functions as an *O*-methyltransferase to catalyze the *O*-methylation of the phenolic OH group of TCM D3.[41] By analogy, FdmN presumably is the methyltransferase responsible for *O*-methylating the OH group of ring D of FDM A. This agrees well with the fact that homologues of FdmN could be found in the PRM A (PRM-ORF6, 46% identity), GRH A (GrhL, 26% identity), and γ-RUB (RubO, 35% identity) biosynthetic pathways, all of which share the common methoxyquinone moiety (ring A) (Figures 1 and 2).

FdmV shares high sequence similarity to a family of asparagine synthetases, such as TcsG (43% identity), as well as RubR (41% identity) and GrhP (39% identity). TcsG has been proposed to be involved in chlorotetracycline biosynthesis;[50] FdmV is thus possibly responsible for transferring an amino group from an amino acid into a polyketide intermediate to afford the 2-pyridone moiety (ring F) of FDM A (Figure 2).

Nine ORFs within the *fdm* cluster encode proteins whose function cannot be predicted by sequence comparison alone. FdmJ, FdmL, FdmM, FdmM1, FdmP, and FdmQ all contain the so-called "antibiotic monooxygenase domain" identified from several *Streptomyces* species.[51] The latter was recently characterized from the ActAV ORF6 monooxygenase, which catalyzes the oxidation of phenols to quinones without any of the typical prosthetic groups for oxygen activation.[52] All of these genes, except *fdmL*, have a close homologue in the *prm* cluster, and some have homologues in the *grh* or *rub* cluster as well. Given the structural and biosynthetic resemblance among PRM A, GRH A, γ-RUB, and FDM A and the fact that the former metabolites are considered to be the most heavily oxidized bacterial polyketides known to date,[33] the finding that these clusters share many of these genes encoding unknown function supports the hypothesis that such genes play roles in the common post-PKS oxidative tailoring steps.

FdmK, with no homologue identified in the *prm*, *grh*, or *rub* cluster, shares significant sequence similarity to the C-terminal region (127 amino acids) of Fox2 (30% identity), which is involved in peroxisomal β-oxidation of long branched fatty acids in the mycorrhizal fungus *Glomus mosseae*.[53] Since the formation of the spiro carbon center is unique in FDM A in contrast to PRM A, GRH A, or γ-RUB, FdmK (as well as FdmL that is also unique to the *fdm* cluster) might be specifically involved in the formation of the spiro carbon center of FDM A. Finally, FdmU does not have any homologue in the database, and FdmX resembles tranposases such as SC3C8.10 from *S. coelicolor* (90% identity); their role in FDM A biosynthesis cannot be speculated.

Genes Encoding Regulation and Resistance Proteins. Three genes, the deduced products of which show sequence homology to known regulatory proteins, were identified within the *fdm* cluster. FdmR and FdmR2 belong to MarR-family of transcriptional regulators,[54] while FdmR1 is an AsfR-type regulatory protein, presumably responsible for globally controlling secondary metabolism in *Streptomyces*.[55] All three regulators have homologues in the *fdm*, *grh*, or *rub* clusters, such as GrhR3 (41% identity) for FdmR, RubS (38% identity) or GrhR2 (42% identity) for FdmR1, and RubM (43% identity) for FdmR2, suggesting a common regulatory mechanism for all these pathways.

Four putative transporter genes were identified within the *fdm* cluster. FdmT shows significant sequence similarity to RUB−ORF3 (33% identity), a putative peptide transporter from *S. collinus*, and FdmT1, FdmT2, and FdmT3 are all putative membrane transporter proteins with homologues in the *fdm*, *grh*, or *rub* clusters, such as PRM-ORF3 (34% identity) for FdmT1, GrhK (34% identity) for FdmT2, and GrhK (38% identity) for FdmT3. These proteins serve as candidates to confer FDM A resistance presumably via metabolite export pumps powered by transmembrane electrochemical gradients.

Discussion

Cloning and Confirmation of the FDM Biosynthetic Gene Cluster. Here we report the cloning and sequence analysis of the gene cluster for the biosynthesis of FDM A, an antitumor antibiotic with unique structural features, from *S. griesus* ATCC49344. Feeding experiments with [1-^{13}C], [2-^{13}C], and [1,2-^{13}C]acetate previously established the pentadecaketide biosynthetic origin for FDM A.[17] Since the biosynthesis of aromatic polyketides in bacteria is typically catalyzed by type

(47) Meurer, G.; Gerlitz, M.; Wendt-Pienkowski, E.; Vining, L. C.; Rohr, J.; Hutchinson, C. R. *Chem. Biol.* **1997**, *4*, 433−443.
(48) Omura, S.; Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Takahashi, C.; Shinose, M.; Takahashi, Y.; Horikawa, H.; Nakazawa, H.; Osonoe, T.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12215−12220.
(49) Galm, U.; Schimana, J.; Fiedler, H. P.; Schmidt, J.; Li, S. M.; Heide, L. *Arch. Microbiol.* **2002**, *178*, 102−114.
(50) Nakano, T.; Miyake, K.; Endo, H.; Dairi, T.; Mizukami, T.; Katsumata, R. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 1345−1352.
(51) Marchler-Bauer, A.; et al. *Nucleic Acids Res.* **2003**, *31*, 383−387.
(52) Sciara, G.; Kendrew, S. G.; Miele, A. E.; Marsh, N. G.; Federici, L.; Malatesta, F.; Schimperna, G.; Savino, C.; Vallone, B. *EMBO J.* **2003**, *22*, 205−215.
(53) Requena, N.; Fuller, P.; Franken, P. *Mol. Plant Microbe Interact.* **1999**, *12*, 934−942.
(54) Seoane, A. S.; Levy, S. B. *J. Bacteriol.* **1995**, *177*, 3414−3419.
(55) Floriano, B.; Bibb, M. *Mol. Microbiol.* **1996**, *21*, 385−396.

II PKSs,[1,2] we adopted a PCR strategy to clone the *fdm* cluster by amplifying putative PKS fragments from the FDM A producer *S. griseus*. Using the PCR-amplified PKS fragment as a probe, we screened the *S. griseus* genomic library to isolate the FDM PKS locus, followed by chromosome walking to clone the entire *fdm* gene cluster in a 33,058-bp DNA region. Sequence analysis of this DNA revealed 33 ORFs (Figure 3 and Table 1). An in-frame deletion of the putative *fdm* minimal PKS genes abolished FDM A production, and subsequent complementation of this mutation by expressing the *fdm* minimal PKS genes *in trans* restored FDM A production to the wild-type level, confirming that the cloned gene cluster is essential for FDM A biosynthesis (Figure 4). In contrast, inactivation of *orf1*, *orf3*, or *orf33* had little or insignificant effect on FDM A production, suggesting that these genes are not involved in FDM A biosynthesis (Figure 5). Finally, introduction of the 25-kb fragment harboring the *fdmC* to *fdmT3* genes into *S. albus* resulted in the production of FDM A with an excellent yield in the heterologous host, unambiguously establishing the precise boundaries of the *fdm* cluster (Figure 5). From these results, the *fdm* gene cluster was defined to consist of 28 ORFs, featuring a unique type II PKS with a dedicated PPTase for the biosynthesis of a pentadecaketide intermediate and numerous novel enzymes for the subsequent tailoring steps to afford FDM A (Figures 2 and 3).

Chain Length Determination. Dozens of aromatic polyketide gene clusters have been sequenced to date, coding for the biosynthesis of polyketides with chain lengths ranging from 16 to 26 carbons (among the ones whose metabolites have been unambiguously identified) as exemplified in Figure 1. While most of the current studies focused primarily on the ACT PKS (for an octaketide) and TCM PKS (for a decaketide), the FDM PKS catalyzes the biosynthesis of a pentadecaketide, a polyketide with the longest chain length characterized to date. The recently solved ACT PKS structure has revealed that seven amino acids, two from the KS_α subunit (L143 and F140) and five from the KS_β subunit (T112, F116, W194, F109, G195), at the KS_α/KS_β heterodimer interface form part of a tunnel, the depth of which has been proposed to control the length of the polyketide chain intermediate in type II PKSs.[12] The identification of these residues, called "gatekeepers", is a major step forward in understanding the chain length control for the type II PKSs. KS_α/KS_β subunit pairs from aromatic polyketides with different chain lengths, ranging from the ACT (C-16) to FDM (C-30), need to be analyzed for similar amino acid gatekeepers in order to improve the understanding of how chain length can be controlled and altered for the purpose of combinatorial biosynthesis.

The primary sequences of the FDM KS_α/KS_β pair fit the polyketide tunnel hypothesis well, with the amino acid gatekeepers GLYVGVY (G113, L117, Y195, V110, G196 from the FdmF KS_β, V145, and Y142 from the FdmG KS_α) similar to the gatekeepers for the GRH PKS, as suggested by Khosla and co-workers,[12] as well as for the PRM and RUB PKSs. These gatekeeper amino acids are indeed more similar among the KS_α and KS_β subunits of type II PKSs for aromatic polyketides with long chains (C-24 and longer) than that for aromatic polyketides with short chains (C-20 and shorter). To further probe how a type II PKS controls the chain length, we performed a phylogenetic analysis of KS_βs for type II PKSs specifying various chain lengths, since KS_β is the main chain length determinant for aromatic polyketide biosynthesis. While the overall primary amino acid sequences are highly similar, the KS_β subunits specifying for aromatic polyketides with longer chains, such as WhiE-ORFIV (C-24), PRM-ORF2 (C-24), GrhB (C-26), RubB (C-26), and FdmG (C-30), named as the FDM group, and the KS_β subunits specifying for aromatic polyketides with shorter chains, such as Act-Orf2 (C-16), JadB (C-16), Sim3 (C-16), Fren-ORF2 (C-16/18), ZhuA (C-18/19/20), and DpsB (C-20/21), named the ACT group, fall into two distinct phylogenetic groups (Figure 6). The latter finding should inspire future efforts in the identification of additional gatekeepers that may be responsible for accommodating the biosynthesis of aromatic polyketides with chain lengths longer than those currently studied.

Novel Tailoring Enzymes for Post-PKS Steps in FDM A Biosynthesis. After FDM PKS-catalyzed assembly of the nascent linear pentaketide intermediate from 15 malonyl-CoA precursors, minimally five steps are needed to finish the biosynthesis of FDM A (Figure 2). They are as follows: (i) the regiospecific reduction of the carbonyl groups at C_{19} by FdmO and at C_{27} and C_{29} by FdmC followed by dehydration to afford the pentadienyl side chain; (ii) the cyclization of the linear acyl-S-ACP species into the pentacyclic intermediate by the FdmDEI cyclases/aromatases; (iii) the methylation of phenolic OH of ring A by the FdmN methyltransferase; (iv) the generation of the 2-pyridone ring by asparagine synthetase homologue FdmV; (v) the oxidation of rings A, B, and C to form the characteristic quinone/hydroquinone functionalities and subsequent oxidative rearrangement of rings C and D to afford the unique spiro carbon center of FDM A by novel oxygenases such as FdmJKLMM1PQ. While the timing of these steps remains to be determined, this proposal agrees well with the recently solved structures of FDM B and C as key biosynthetic intermediates and FDM C_1 as a shunt metabolite for FDM A biosynthesis.[16] The current study sets the stage to further test this proposal experimentally.

The proposal of the pentacyclic intermediate for FDM A biosynthesis is based on the presence of common angular aromatic ring systems as well as the highly similar cyclases shared by the FDM A, PRM A, GRH A, and γ-RUB pathways and agrees well with the recent isolation of FDM C as the earliest intermediate for FDM A biosynthesis.[16] The finding that the cyclases FdmI and FdmD also share high similarity to TcmN and TcmI is not unexpected because of the similar C_{11} folding pattern and initial formation of the first aromatic ring (ring A) catalyzed by TcmN or FdmI and the subsequent formation of the ring D by TcmI or FdmD for TCM or FDM A, respectively (Figures 1 and 2). Angucycline-specific cyclases, such as JadI, have been characterized previously.[56] However, no JadI homologue could be identified within the *fdm*, *prm*, *grh*, or *rub* clusters, which raises the question if there are additional unknown cyclases specifically for this type of structure formation. Only one extra cyclase, GrhE, has been identified from the GRH A pathway, which is an ActVII-type cyclase that directs the first ring closure at the C_9-reduced position for aromatic polyketide biosynthesis; GrhE was proposed to be inactive, though, since no ActVII-type cyclization product could be detected.[33]

(56) Kulowski, K.; Wendt-Pienkowski, E.; Han, L.; Yang, K.; Vining, L. C.; Hutchinson, C. R. *J. Am. Chem. Soc.* **1999**, *121*, 1786−1794.

FdmV homologues were also identified in the amide containing chlortetracycline aromatic polyketide gene cluster.[50] For the formation of the 2-pyridone ring, the asparagine synthetase homologue FdmV could catalyze the formation of the amide from glutamine, ATP, and an FDM intermediate such as FDM C. Following the attack of the closest carbonyl carbon by the amine, dehydration and aromatization could afford the 2-pyridone ring (ring F), and the isolation of both FDM C_1 as a shunt metabolite and FDM B as an intermediate for FDM A biosynthesis support this proposal (Figure 2).[16]

The fact that the FDM A biosynthetic machinery shares similarity with many of the novel oxygenases seen in the PRM A, GRH A, or γ-RUB biosynthetic machinery supports the idea of a common mechanism for the formation of the A/B/C ring moiety for this family of metabolites. However, the spiro carbon feature of FDM A is unique. An oxidative ring contraction mechanism has been proposed to account for the loss of C_{18} as suggested by feeding experiments with isotope-labeled precursors.[17] Taking the oxidative nature of this carbon rearrangement into account, the two FDM A pathway-specific novel oxygenases, FdmL and FdmK, serve as probable candidates for catalyzing this reaction. While various mechanisms for oxidative ring fission or contraction have been proposed for the biosynthesis of polyketides such as gilvocarcin,[57] jadomycin,[58] GRH A and γ-RUB,[33] and kinamycin,[59] little is known about a spiro carbon center biosynthesis. FDM A biosynthesis provides an excellent opportunity to investigate the mechanism of these novel oxygenases that catalyze an unprecedented biotransformation.

Significance. FDM A is an unusual aromatic polyketide from several viewpoints (atypical processing of the first three chain extension units, longest length of the poly-β-ketone intermediate, a novel chiral spiro carbon suggestive of unusual oxidative biochemistry among the enzymes involved). Its discovery elicited considerable interest because of its unusual molecular shape[15] and its potential as a novel anticancer drug lead.[22−27] Later, attempts to define its mechanism of action as a cytotoxin with possible antitumor activity led to contrasting conclusions,[18−22] which have left this question unresolved. Consequently, the availability of the biosynthetic genes described here opens the way to studies of the unusual biochemistry and to attempts to use the information about the FDM PKS uncovered here in combinatorial biosynthesis to refine the understanding of how the chain length of the incipient poly-β-ketone intermediate(s) is controlled and how their cyclization is affected. The results of such work may also provide new FDM A analogues for further investigation of its mechanism of action and potential use as an antitumor agent.

Acknowledgment. We thank Dr. Ramesh C. Pandey, NCI-Frederick Cancer Research Facility, Frederick, MD, for an authentic sample of fredericamycin A, Dr. Jose A. Salas, University of Oviedo, Spain, for the *S. albus* J1074 strain, and the Analytical Instrumentation Center of the School of Pharmacy, UW-Madison, for support in obtaining MS and NMR data. This work is supported in part by the School of Pharmacy, UW-Madison, and National Institutes of Health Grant No. CA35381. B.S. is the recipient of National Science Foundation CAREER Award MCB9733938 and National Institutes of Health Independent Scientist Award AI51687.

Supporting Information Available: Complete ref 51. This material is available free of charge via the Internet at http://pubs.acs.org.

JA054376U

(57) Liu, T.; Fischer, C.; Beninga, C.; Rohr, J. *J. Am. Chem. Soc.* **2004**, *126*, 12262−12263.
(58) Rix, U.; Zheng, J.; Remsing, L. L.; Greenwell, L.; Yang, K.; Rohr J. *J. Am. Chem. Soc.* **2004**, *126*, 4496−4497.
(59) Gould, S. J.; Melville, C. R.; Cone, M. C.; Chen, J.; Carney, J. R. *J. Org. Chem.* **1997**, *62*, 320−324.