Analysis and Refactoring of the A-74528 Biosynthetic Pathway

Jay T. Fitzgerald,[†,‖] Louise K. Charkoudian,[†,‖] Katharine R. Watts,[†] and Chaitan Khosla*[,†,‡,§]

[†]Department of Chemistry, [‡]Department of Chemical Engineering, and [§]Department of Biochemistry, Stanford University, Stanford, California 94305, United States

ⓈSupporting Information

ABSTRACT: A-74528 is a C_{30} polyketide natural product that functions as an inhibitor of 2′,5′-oligoadenylate phosphodiesterase (2′-PDE), a key regulatory enzyme of the interferon pathway. Modulation of 2′-PDE represents a unique therapeutic approach for regulating viral infections. The gene cluster responsible for biosynthesis of A-74528 yields minute amounts of this natural product together with considerably larger quantities of a structurally dissimilar C_{30} cytotoxic agent, fredericamycin. Through construction and analysis of a series of knockout mutants, we identified the genes necessary for A-74528 biosynthesis. Remarkably, the formation of six stereocenters and the regiospecific formation of six rings in A-74528 appear to be catalyzed by only two tailoring enzymes, a cyclase and an oxygenase, in addition to the core polyketide synthase. The inferred pathway was genetically refactored in a heterologous host, *Streptomyces coelicolor* CH999, to produce 3 mg/L A-74528 in the absence of fredericamycin.

The development of new antiviral agents is an important goal in contemporary medicinal chemistry. Because viruses typically rely on native cell machinery for their replication, a primary challenge in engineering effective therapeutics is the ability to differentiate between viral and host targets. In this respect, the polyketide A-74528 (**1**) (Figure 1) is a promising antiviral lead. Its unique mode of action
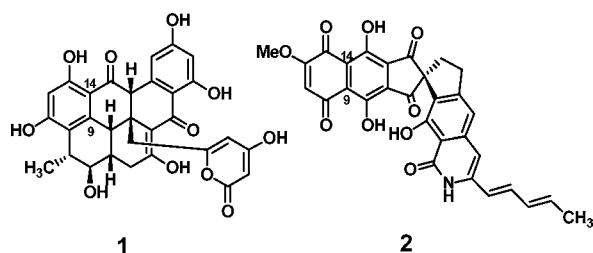

Figure 1. The products of the *san* gene cluster, A-74528 (**1**) and FDM A (**2**). The C9 and C14 positions on the two molecules are labeled to highlight the extent of the similarity in their cyclization patterns.

entails boosting the native 2-5A antiviral pathway through inhibition of 2′,5′-oligoadenylate phosphodiesterase (2′-PDE).[1]

We recently cloned and sequenced the *san* gene cluster responsible for A-74528 production in *Streptomyces* sp. SANK 61196.[2] We also reported heterologous production of A-74528 in *Streptomyces lividans* K4-114 using cosmid pKZ11, which

harbors the entire A-74528 gene cluster except for the putative phosphopantetheinyl transferase *sanW*.[2] In both the native and heterologous systems, A-74528 was produced in impractically low quantities (≪1 mg/L), a problem that was exacerbated by the copious production (∼100 mg/L) of cytotoxic fredericamycin A (FDM A, **2**)[3] by the same *san* cluster [Figure S1 in the Supporting Information (SI)].

In this work, we sought to identify the minimal set of proteins necessary for A-74528 production and to engineer a system capable of producing A-74528 to the exclusion of FDM A. We began by constructing a set of mutants of cosmid pKZ11 in which candidate biosynthetic genes were individually deleted using polymerase chain reaction (PCR) targeting.[4] Mutant cosmids were introduced into *S. lividans* K4-114, and the metabolite profiles of the resulting transformants were analyzed. Using this information, we then sought to refactor the minimal A-74528 pathway in the heterologous host *Streptomyces coelicolor* CH999.[5] For these studies, the shuttle vector pRM5 was used,[5] as it allows for a greater degree of control over enzyme expression than pKZ11.

The *san* gene cluster encodes a putative ketosynthase/chain length factor (KS/CLF) heterodimer (*sanFG*), an acyl carrier protein (ACP) (*sanH*), and a phosphopantetheinyl transferase (*sanW*).[6] Aside from these crucial genes, the *san* cluster also harbors a putative C9−C14 cyclase (*sanI*) and an initiation module (*sanS2C*). All of these proteins were presumed to be required for both A-74528 and FDM A production. [See the SI for a complete list of genes and putative functions (Table S2), as well as proposed pathways to A-74528 (Scheme S2) and FDM A (Scheme S1).] A number of additional genes encode putative regulatory proteins and transporters (*sanTT1T2RR1*). Lastly, on the basis of the work of Shen and co-workers,[7,8] some genes were thought to be specific to FDM A production (*sanMM1VED*) and were thus presumed to lie downstream of the branch point between the FDM A and A-74528 pathways. We therefore focused on the remaining genes, including five putative oxygenases (*sanJKLPQ3*), a ketoreductase (*sanO*), and two genes of unknown function (*sanUX*).

In-frame deletion mutants of pKZ11 were constructed using the previously described PCR targeting system and flippase (FLP) recombination (for details, see the SI).[4] The metabolite profile of each mutant plasmid was analyzed by extracting at least 10 plates of the corresponding *S. lividans* K4-114 transformant with organic solvent (for details, see the SI).

We had previously hypothesized that the protein product of *sanX* may play a role in A-74528 production.[2] This gene is the
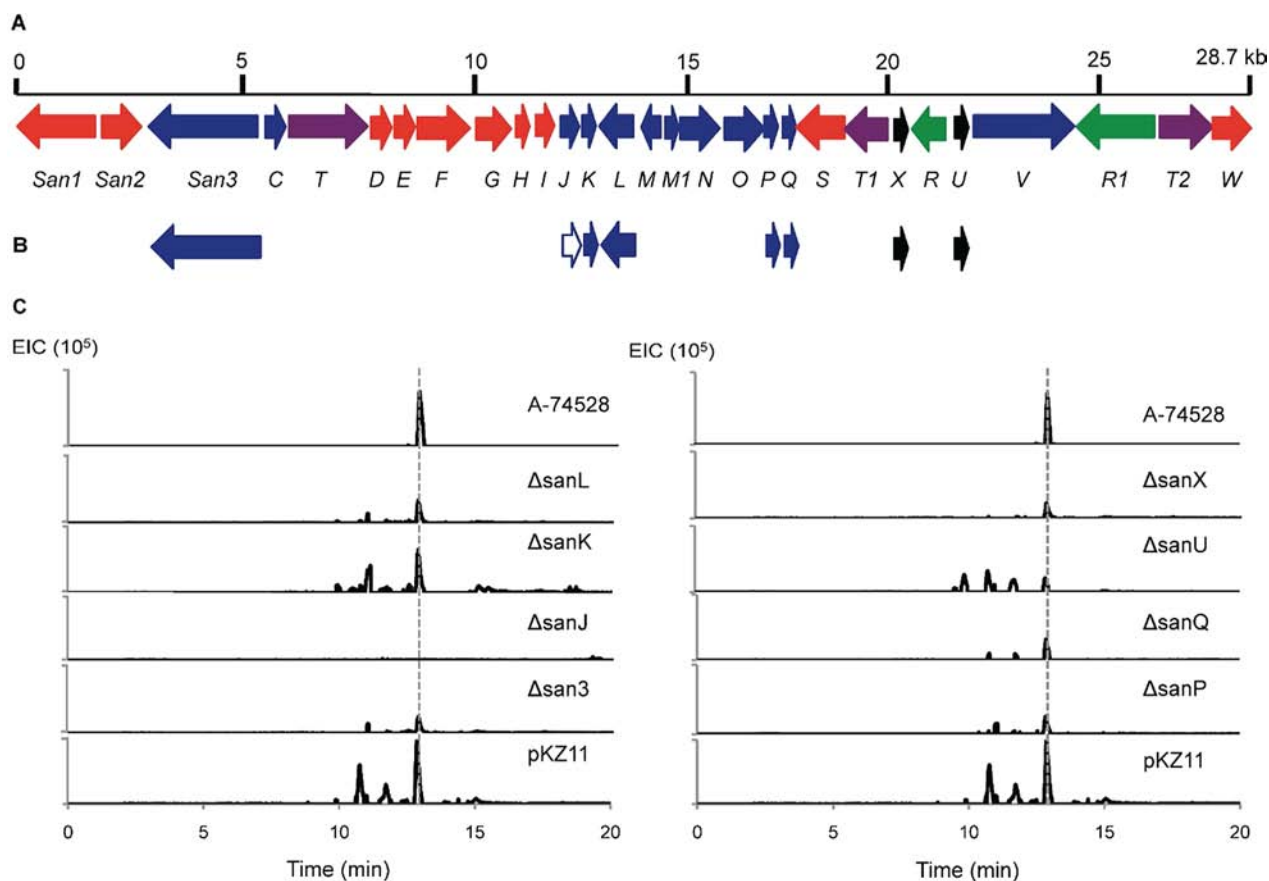
Figure 2. PCR targeting of pKZ11, which harbors the biosynthetic genes necessary for A-74528 production. (A) The 28.7 kilobase *san* gene cluster encodes for PKS and associated genes (red), tailoring genes (blue), resistance genes (purple), regulatory genes (green), and genes of unknown function (black). (B) Genes encoding oxygenases and enzymes of unknown function were individually disrupted by PCR targeting. Only *sanJ* was found to be essential for A-74528 biosynthesis. (C) LC−MS analysis of extracts isolated from the mutant strains revealed the presence of A-74528 in most of the secondary metabolite extracts but not in pKZ11/Δ*sanJ*. The extracted ion chromatogram (EIC) for the positive ion of A-74528 ([M + H]$^+$, m/z 561.1−561.15) is shown on the Y axis. The traces for pKZ11, Δ*sanJ*, Δ*sanU*, and Δ*sanQ* were obtained from extracts derived from liquid cultures followed by ammonium sulfate precipitation that were analyzed at 50 mg/mL. The traces for Δ*san3*, Δ*sanK*, Δ*sanL*, Δ*sanP*, and Δ*sanX* were obtained using samples derived from agar plate cultures that were analyzed at 25 mg/mL. See the SI for experimental details.

only qualitative difference between the *fdm* cluster in *Streptomyces griseus*, which exclusively produces FDM A, and the *san* cluster in *Streptomyces sp.* SANK 61196, which produces both FDM A and A-74528. Surprisingly, the *sanX* knockout retained the ability to produce A-74528 (Figure 2C). This led us to speculate that a more subtle difference controls the branch point between the FDM A and A-74528 pathways. (All other genes in the *san* cluster bear >80% sequence identity to their homologues in the *fdm* cluster.)

As shown in Figure 2, the putative oxygenase, SanJ, is the only protein among those tested that is necessary for A-74528 biosynthesis. None of the other oxygenases (SanK, SanL, SanP, SanQ, or San3) is essential, nor is the ketoreductase (SanO) or the protein of unknown function (SanU). Because biosynthesis of FDM A (but not A-74528) requires reduction of the C19 ketone of the putative polyketide backbone, we correctly anticipated that the Δ*sanO* mutant would not produce FDM A (Figure S5).

The above finding is consistent with our hypothesis (Scheme S2) that only one oxygenation is required at the C28 position for A-74528 biosynthesis, whereas several other oxygenation reactions occur in the pathway leading to FDM A.[8] Oddly, the Δ*sanK*, Δ*sanL*, Δ*sanP*, Δ*sanQ*, and Δ*san3* mutants retained some ability to produce FDM A (Figure S4). This may be due

to partially redundant function of the oxygenases, not unlike what has previously been observed in the oxytetracycline system.[9]

To test our hypothesis that A-74528 biosynthesis requires *sanJ*, we sought to refactor the minimal pathway for this antiviral agent in *S. coelicolor* CH999. Genes for the minimal polyketide synthase (PKS) (*sanFGH*), the C9−C14 cyclase (*sanI*), and the oxygenase (*sanJ*) as well as the initiation module (*sanS2C*) were cloned from pKZ11 using primers described in Table S1. Separately, the gene for the phosphopantetheinyl transferase (*sanW*) was amplified from pKZ2.[2] These genes were then assembled into small segments using PCR splicing by overlap extension (PCR SOEing)[10] and ligated into intermediate vectors. These segments were then transferred into pRM5 (see the SI for details). Three constructs were obtained: the minimal PKS, *sanFGHW* (pJF77); the minimal PKS, the initiation module, and the C9−C14 cyclase, *sanFGHWIXS2C* (pJF76); and a construct that also included *sanJ*, *sanFGH-WIS2CJ* (pJF111).[11]

Each plasmid was introduced into *S. coelicolor* CH999/pBOOST*,[12] yielding strains CH999/pBOOST*/pJF77, CH999/pBOOST*/pJF76, and CH999/pBOOST*/pJF111. Each strain was grown on semisolid agar medium (500 mL for analytical purposes and 3 L for compound isolation and
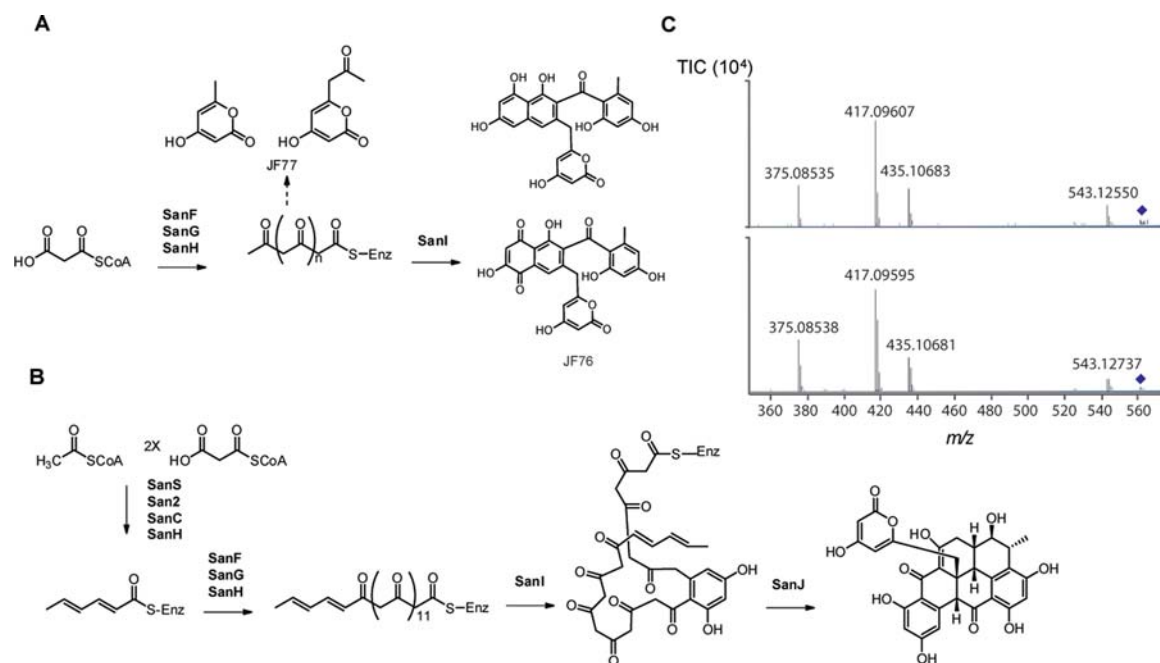
Figure 3. Refactoring of A-74528 biosynthesis in *S. coelicolor* CH999. (A) Strains harboring the minimal PKS enzymes (CH999/pBOOST*/pJF77) produced tri- and tetraketide shunt products. Upon the addition of the initiation module and first ring cyclase enzymes (CH999/pBOOST*/pJF76), the dodecaketides TW95a and TW95b were observed. In this case, the initiation module appears to have failed to provide the appropriate hexadienyl starting unit to this major polyketide product. (B) Proper function of the initiation module along with addition of the oxygenase SanJ (CH999/pBOOST*/pJF111) afforded a strain capable of producing A-74528. As in previous work,[16] the origins of the hexadienyl starting unit remain unclear but here are shown as acetyl-CoA and malonyl-CoA for simplicity. (C) LC−MS/MS confirmed the identity of A-74528 produced by CH999/pBOOST*/pJF111 (top panel) through comparison to an authentic standard (bottom panel). Diamonds represent the A-74528 parent ion.

characterization). The medium was then extracted with equal volumes of a 1% AcOH/10% MeOH/89% EtOAc solution overnight. Small-scale extracts were evaporated to dryness and resuspended in methanol for analysis by Q-TOF LC−MS. For characterization of major metabolites, C_{18} solid-phase extraction cartridges and preparative HPLC were used to purify individual molecules (see the SI).

The pigmentation pattern of *S. coelicolor* CH999/pBOOST*/pJF77 (yellow/orange) was distinct from the brown/black pigment produced by the other two strains. Unsurprisingly, it also did not produce A-74528, as gauged by LC−MS analysis (Figure S6). It lacked any major products with absorbance at both 280 and 350 nm [wavelengths at which many conjugated, oxygenated, aromatic compounds appear (Figure S2)]. The only isolable compounds from this strain were previously reported tri- and tetraketides (Figure 3A).[13] Although the absence of longer-chain polyketides cannot be definitively interpreted, it appears that the minimal PKS lacks auxiliary subunits that allow the KS/CLF to fashion a binding site for the growth of longer polyketide chains.

In addition to the minimal PKS, *S. coelicolor* CH999/pBOOST*/pJF76 also harbors the initiation module and the C9−C14 cyclase. It produced a more instructive polyketide mixture than the minimal construct alone, although it too was unable to produce A-74528 (Figure S6). In addition to the tri- and tetraketide products produced by CH999/pBOOST*/pJF77, the strain harboring pJF76 produced two new peaks that absorbed at 280 and 350 nm. These peaks matched the retention time and UV signature of TW95a and TW95b (Figure S2), two known dodecaketides originally isolated from strains harboring the *whiE* minimal PKS and the *whiE* C9−C14 cyclase.[14] The production of these dodecaketides has three

significant implications for the *san* KS/CLF. First, as has been seen with other type-II PKSs that have the capacity to synthesize very long chain products, auxiliary enzyme components can be presumed to modulate the chain-length specificity of the *san* KS/CLF.[13,15] Second, although TW95a and TW95b are produced in small quantities by this strain, the product profile remains similar to that of JF77 (Figure S2), indicating that long-chain polyketide formation is inefficient. Third, the lack of any identifiable hexadienyl-primed products indicates that priming of the minimal PKS by the initiation module is either inefficient or does not occur at all in this construct.

S. coelicolor CH999/pBOOST*/pJF111 differs from its counterpart harboring pJF76 by only the putative *sanJ* oxygenase gene. It produced A-74528 (Figures 3C, S6, and S7 and Table S5), confirming our hypothesis that *sanJ* is indeed required for production of this metabolite. The strain produced isolable quantities of A-74528 from a standard agar extraction (titer estimated at 3 mg/L in comparison with an authentic reference standard). This quantity of A-74528 was previously inaccessible from both the native *Streptomyces* sp. SANK 61196 and the heterologous *S. lividans* K4114/pKZ11 expression strain, both of which produced quantities too small to be isolated easily. Unfortunately, A-74528 remains a relatively minor product of *S. coelicolor* CH999/pBOOST*/pJF111; the strain also yields larger quantities of other shunt metabolites, including TW95a and TW95b (Figure S2). Thus, efficient priming of the minimal PKS remains a problem.

Our analysis of the *san* cluster has provided unexpected and interesting new insights into the biosynthesis and tailoring of one of the longest and most intricately fashioned polyketides made by a type-II PKS. Remarkably, the formation of six

stereocenters and the regiospecific formation of six rings in A-74528 appear to occur through the action of only two tailoring enzymes, the C9—C14 cyclase SanI and the putative oxygenase SanJ. It is of course possible that the minimal PKS also aids in fashioning the natural product. Thus, the paradigm in which the chain length, primer unit selection, and cyclization regiospecificity are modularly controlled appears to need modification when very large chains and complex priming units are involved. Further studies to understand this interplay between enzyme—substrate and protein—protein interactions are warranted.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

Experimental procedures and compound characterizations. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

khosla@stanford.edu

Author Contributions

‖J.T.F. and L.K.C. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Silverman, R. H. *Biochemistry* **2003**, *42*, 1805−1812.
(2) Zaleta-Rivera, K.; Charkoudian, L. K.; Ridley, C. P.; Khosla, C. *J. Am. Chem. Soc.* **2010**, *132*, 9122−9128.
(3) Wendt-Pienkowski, E.; Huang, Y.; Zhang, J.; Li, B.; Jiang, H.; Kwon, H.; Hutchinson, C. R.; Shen, B. *J. Am. Chem. Soc.* **2005**, *127*, 16442−16452.
(4) Gust, B.; Kieser, T.; Chater, K. F. *REDIRECT Technology: PCR Targeting Systems in Streptomyces coelicolor*; John Innes Centre: Norwich, U.K., 2002.
(5) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D.; Khosla, C. *Science* **1993**, *262*, 1546−1550.
(6) Das, A.; Khosla, C. *Acc. Chem. Res.* **2009**, *42*, 631−639.
(7) Chen, Y.; Wendt-Pienkowski, E.; Ju, J.; Lin, S.; Rajski, S. R.; Shen, B. *J. Biol. Chem.* **2010**, *285*, 38853−38860.
(8) Chen, Y.; Wendt-Pienkoski, E.; Rajski, S. R.; Shen, B. *J. Biol. Chem.* **2009**, *284*, 24735−24743.
(9) Wang, P.; Zhang, W.; Zhan, J.; Tang, Y. *ChemBioChem* **2009**, *10*, 1544−1550.
(10) Horton, R.; Hunt, H.; Ho, S.; Pullen, J.; Pease, L. *Gene* **1989**, *77*, 61−68.
(11) The gene *sanX* was not included in pJF111, though it was present in pJF76, and it was found to have no effect on A-74528 production.
(12) Hu, Z.; Hopwood, D. A.; Hutchinson, C. R. *J. Ind. Microbiol. Biotechnol.* **2003**, *30*, 516−522.
(13) Das, A.; Khosla, C. *Chem. Biol.* **2009**, *16*, 1197−1207.
(14) Yu, T.-W.; Shen, Y.; McDaniel, R.; Floss, H. G.; Khosla, C.; Hopwood, D. A.; Moore, B. S. *J. Am. Chem. Soc.* **1998**, *120*, 7749−7759.
(15) Szu, P.-H.; Govindarajan, S.; Meehan, M. J.; Das, A.; Nguyen, D. D.; Dorrestein, P. C.; Minshull, J.; Khosla, C. *Chem. Biol.* **2011**, *18*, 1021−1031.
(16) Das, A.; Szu, P.-H.; Fitzgerald, J. T.; Khosla, C. *J. Am. Chem. Soc.* **2010**, *132*, 8831−8833.