ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

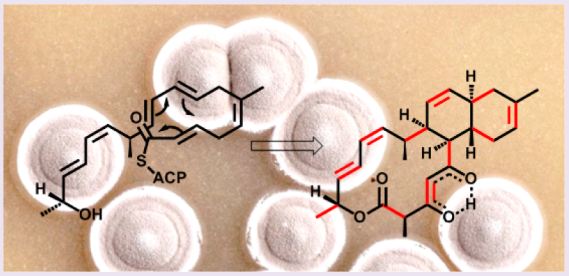
Biosynthesis of the Novel Macrolide Antibiotic Anthracimycin

Silke Alt and Barrie Wilkinson*

Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich, Norfolk NR4 7UH, United Kingdom

Supporting Information

ABSTRACT: We report the identification of the biosynthetic gene cluster for the unusual antibiotic anthracimycin (*atc*) from the marine derived producer strain *Streptomyces* sp. T676 isolated off St. John's Island, Singapore. The 53 253 bps *atc* locus includes a *trans*-acyltransferase (*trans*-AT) polyketide synthase (PKS), and heterologous expression in *Streptomyces coelicolor* resulted in anthracimycin production. Analysis of the *atc* cluster revealed that anthracimycin is likely generated by four PKS gene products AtcC−AtcF without involvement of *post*-PKS tailoring enzymes, and a biosynthetic pathway is proposed. The availability of the *atc* cluster provides a basis for investigating the biosynthesis of anthracimycin and its subsequent bioengineering to provide novel analogues with improved pharmacological properties.



Anthracimycin (**1**) (Figure 1) is a structurally unique 14-membered macrolide antibiotic that was first reported in

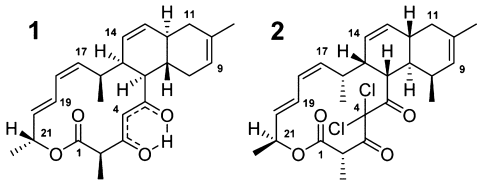


Figure 1. Chemical structures of anthracimycin (**1**) produced by *Streptomyces* sp. strain T676 and the related structure of chlorotonil A (**2**) produced by *Sorangium cellulosum* 1525.

2013 from *Streptomyces* sp. CNH365 isolated off the Santa Barbara (CA) coast.[1] It had previously been isolated from *Streptomyces* sp. T676 by MerLion Pharma (Singapore) in 1995 following high-throughput screening for new antibacterials (originally named N1097; unpublished). *Streptomyces* sp. T676 was isolated from a sediment sample taken off St. John's Island (Singapore) and culture extracts exhibited potent antibacterial activity (≤0.06 μg/mL) toward Gram-positive bacteria including methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant enterococci (VRE). The active component was identified as N1097 which is identical to **1**. **1** also displays excellent activity against *Bacillus anthracis* (0.03 μg/mL) and minimal toxicity toward eukaryotic cells with an IC_{50} = 70 μg/mL against human carcinoma cells.[1,2] Most macrolide antibiotics operate through the inhibition of protein biosynthesis by binding to the 50S ribosomal subunit.[3] However, macromolecular synthesis studies suggest that **1** inhibits DNA and/or RNA synthesis, with this occurring in the absence of DNA intercalation.[2]

1 is structurally distinct from typical macrolide antibiotics such as erythromycin A by virtue of a tricyclic system, which includes a decalin moiety, and an absence of any deoxyhexose moieties. Only chlorotonil A (**2**) (Figure 1), produced by the myxobacterium *S. cellulosum So ce* 1525, has been reported to possess a similar carbon skeleton.[4] Chlorotonil A differs from **1** in that all sp^3 stereocenters are inverted plus it has an additional methyl group at C8, and a *gem*-dichloro moiety at C4.

The structure of **1** suggests a polyketide origin with chain assembly on a type I polyketide synthase (PKS). The decalin moiety is likely to be the product of a [4 + 2] cycloaddition and three of the double bonds are located in noncanonical positions for classical modular PKS biosynthesis. PKSs catalyze the formation of specialized metabolites by successive condensation of carboxylic acids, and three major classes of PKSs are known: type I PKSs, iterative type II PKSs, and acylcarrier protein (ACP)-independent type III PKSs, each of which consists of additional variants. *cis*-Type I modular PKSs are characterized by (generally) noniteratively acting modules, each of which minimally contains a β-ketoacyl-ACP synthase (KS) domain, an acyltransferase (AT) domain, and an acyl carrier protein (ACP) domain. These are responsible for selecting carboxylic acid precursors from the cellular pool (AT domain) and loading them onto the PKS assembly line (onto the ACP domains), before catalyzing chain elongation *via* a decarboxylative Claisen-like condensation (KS domain), to generate a β-ketoacyl-S-ACP intermediate. Optional domains such as β-ketoacyl-ACP reductases (KR), dehydratases (DH), enoyl reductases (ER) or methyltransferases (MT) are found between the AT and ACP domains within a module and carry out further transformation of the newly formed β-keto group prior to transfer of the elongated chain to the KS domain of the following module. Once the processing polyketide chain reaches the final module of the PKS, it is released by a C-

Received: July 8, 2015
Accepted: August 23, 2015
Published: September 8, 2015

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479

Table 1. Deduced Function of Gene Products for the *atc* Cluster

protein/gene	size (Da/bp)	proposed function, PKS modules, and protein domains	position in *atc* cluster (bps)	database ID of protein homologue (similarity/identity %)[a]
AtcA/*atcA*	52342/1548	Multidrug MFS (Major Facilitator Superfamily) transporter	140−1687	I0L732_9ACTN (70/54)
AtcB/*atcB*	22382/615	Transcriptional regulator TetR family	1648−2298	V6L528_9ACTN (64/52)
AtcC/*atcC*	120961/3327	*trans* acyltransferase and oxidoreductase: AT1, AT2, ER	2782−6108	H5XCT6_9PSEU (70/57)
AtcD/*atcD*	691113/19686	PKS	6105−25790	H5XCT5_9PSEU (61/52)
		Module 1: KS1, KR, ACP1		
		Module 2: KS2 [clade IX], D, DH, KR, ACP2		
		Module 3: KS3 [clade IV], D, DH, KR, MT, ACP3		
		Module 4: AD, KS4 [clade I], DH		
AtcE/*atcE*	467966/13407	PKS	25780−39186	D9WSJ9_9ACTN (50/39)
		ACP4		
		Module 5: KS5 [clade IX], DH, KR, ACP5		
		Module 6: KS6 [clade IX], ACP6a, D, ACP6b		
		Module 7: KS7 [clade I], DH, KR, MT, ACP7		
AtcF/*atcF*	409749/11610	PKS	39183−50792	A0A017T6E6_9DELT (49/38)
		Module 8: KS8 [clade I], DH, KR, ACP8		
		Module 9: KS9 [clade IX], (ACP9a), ACP9b		
		Module 10: KS10 [clade IV], D, ACP10, TE, MT		
AtcG/*atcG*	31768/852	ABC transporter	50888−51739	A0A0B7CII6_9PSEU (72/59)
AtcH/ *atcH*	26524/750	Putative membrane protein	51736−52485	M0IKY4_9EURY (62/44)
AtcI/*atcI*	24415/708	Putative membrane protein	52482−53189	A0A0B7CLF4_9PSEU (49/31)

[a]NCBI BLAST+ at EBI was used to query UniProt Knowledgebase protein database (ID given corresponds to UniProt accession).

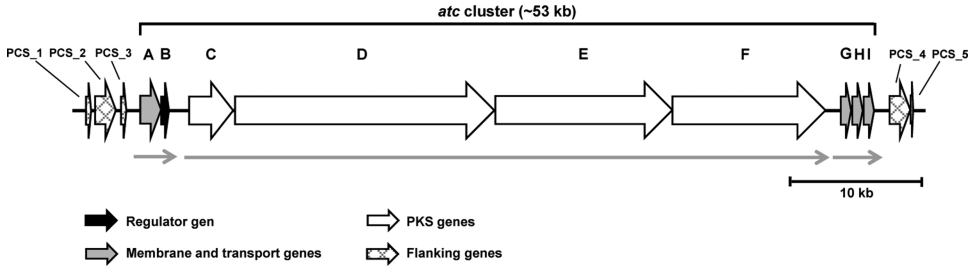


Figure 2. Organization of the 53 253 bp **1** (*atc*) biosynthetic gene cluster in *Streptomyces* sp. T676. The function of each gene is listed in Table 1. The transcriptional regulator gene is shown in black, genes encoding polyketide synthases are white, and transporter genes are gray. The gray arrows illustrate the three mRNA molecules likely to be generated by transcription of the cluster. Five protein coding sequences (PCSs) are illustrated upstream and downstream of the *atc* cluster.

terminal domain, most commonly a thioesterase (TE) domain, to yield the first enzyme free intermediate of the biosynthetic process. These can then undergo further "*post*-PKS" modifications to generate the final, mature polyketide natural product.[5]

One alternative variant of the modular type I PKSs is the *trans*-AT PKS in which each module lacks an integral AT domain. The essential AT activity is instead provided by a discrete AT protein that acts *in trans*. In addition to the unique acyl transfer mechanism, these PKS enzymes are noteworthy for exhibiting highly aberrant architectures in contrast to the *cis*-AT PKSs where the domain and modular architecture has a generally high level of correspondence to the molecule produced: this is known as co-linearity. Additionally, there are modules with unusual domain orders, novel types of domains or repeated domain sets, modules split between two proteins, as well as module architectures that do not seem to correspond with the polyketide structure due to missing or apparently superfluous domains.[6] The numerous contradictions of *trans*-AT PKSs to the co-linearity rules often make it difficult to correlate the PKS architecture with function. However, unlike KS domains from *cis*-AT PKSs, KSs from *trans*-AT PKSs can be grouped into distinct clades according to the substrates they process. With the use of biosynthetic and phylogenetic methods, the preferred substrates of KS domains from *trans*-AT PKSs can be predicted.[7] The functional disparities between *cis*-AT and *trans*-AT synthases appear to reside in their different evolutionary paths, with *cis*-AT PKS evolution directed by module duplication and domain diversification and *trans*-AT evolution guided by horizontal gene transfer between bacteria.[8,7] The first member of the *trans*-AT PKS family, the bacillaene PKS, was identified from the sequenced genome of *Bacillus subtillis* 168.[9] The acylation of entire PKS systems by *trans*-acting ATs was first confirmed *in vitro* for the leinamycin biosynthesis in *Streptomyces atroolivaceus* during genetic and biochemical studies.[10,11]

In this paper, we report the identification, cloning, and heterologous expression of the anthracimycin biosynthetic gene cluster from *Streptomyces* sp. T676 which provides an example of a polyketide megasynthase with some highly unusual features. Chief among these is the lack of any discrete genes encoding *post*-PKS acting enzymes that might be responsible

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479

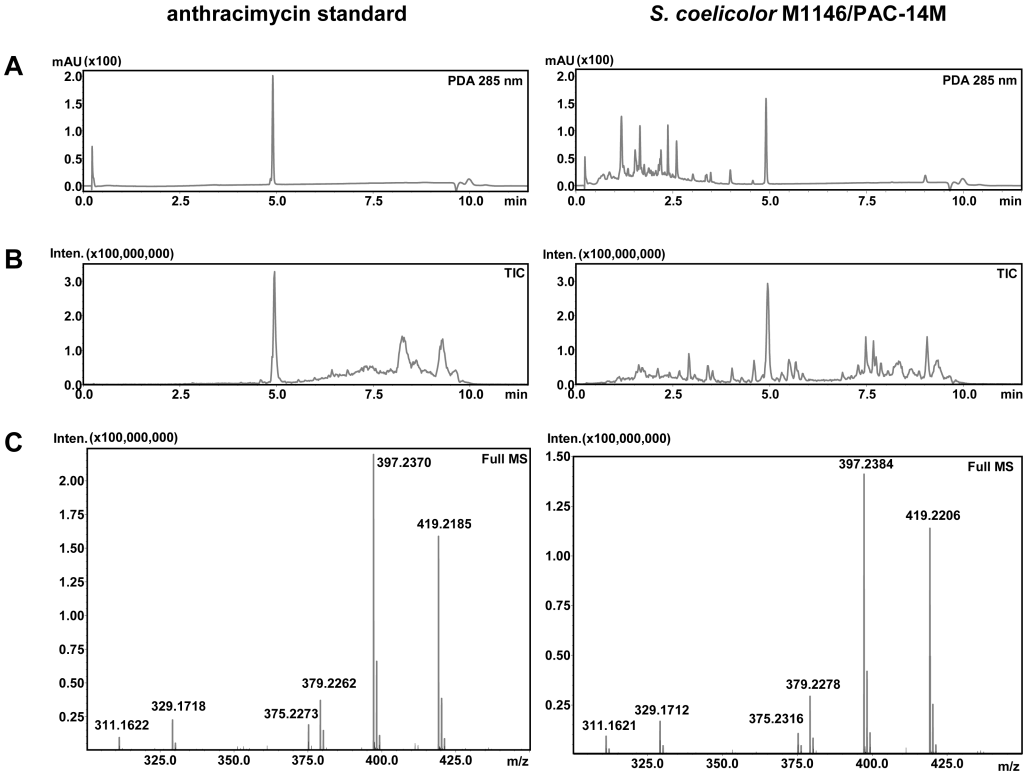


Figure 3. Production of **1** in the heterologous host *Streptomyces coelicolor* M1146/PAC-14M compared with a **1** standard (*m*/*z* 397.23). (A) UV chromatograms at 285 nm; (B) total ion chromatograms; (C) mass spectra.

for the cyclization events required to form the decalin moiety. This strongly suggests that all of the transformations required to produce **1** occur while the polyketide chain is attached to the *trans*-AT PKS enzymes. On the basis of our analysis, a model for the biosynthetic pathway to **1** is proposed.

RESULTS AND DISCUSSION

Identification of the Anthracimycin (*atc*) Biosynthetic Gene Cluster. The genome sequence of *Streptomyces* sp. T676 was obtained using the Pacific Biosciences (PacBio) RS II sequencing platform and yielded approximately 6.9 Mb of contiguous sequence plus one large (*ca.* 57 kb) plasmid.[12] Since we hypothesized that **1** is synthesized by a modular type I PKS, we focused on the two type I modular PKS gene clusters identified by the genome mining platform antiSMASH 3.0.[13] One of these was a type I *cis*-AT PKS, but the predicted domain and module architecture were not consistent with the polyketide backbone of **1**. The second gene cluster encoded a putative *trans*-AT type I PKS consisting of 10 modules with integrated methyltransferase domains and a discrete *trans*-AT gene encoding a protein with two AT-like domains. Functional assignments were made for each gene by comparing the deduced amino acid sequence with proteins of known functions from the databases. A summary of the *atc* cluster organization with the functional description of the gene products of each protein coding sequence (PCS) is shown in Table 1.

Organization of the *atc* Biosynthetic Gene Cluster. The **1** gene cluster spans 53 253 base pairs and bioinformatic analysis identified nine genes, *atcA* to *atcI* (Figure 2). The first two genes, *atcA* and *atcB*, encode, respectively, a putative membrane transport protein of the major facilitator family (Pfam family MFS_1 (PF07690), complete HMM alignment with score 168.6 and e-value of 1.5×10^{-49}), and a putative transcriptional regulator of the TetR family (Pfam family TetR_N (PF00440), complete HMM alignment with score 56.5 and e-value of 1.5×10^{-15}). The following four genes, *atcCDEF*, encode the modules of a putative *trans*-AT type I PKS, with the AT functionality provided by the *atcC* gene product. Of the last three genes, *atcG* encodes a putative transmembrane ABC transporter, with an amino-terminal ATP-binding domain of ABC transporters (Pfam family ABC_tran (PF00005) complete HMM alignment with score 101.5 and e-value of 4.3×10^{-29}); a search with NCBI Conserved Domains Database (CDD) server narrowed the domain down to subfamily A (ABC_DR_subfamily_A (cd03230) complete alignment with score 221.89 and e-value of 3.29×10^{-72}).[14] NCBI Blast, Pfam and CDD searches with the proteins encoded by *atcH* and *atcI* did not reveal any putative function or close homologues;[15,14] however, both AtcH and AtcI were predicted (TMHMM server) to form six transmembrane helices each and therefore most likely are membrane associated proteins.[16]

All nine genes are transcribed in the same direction and likely organized in three transcription units: *atcAB*, *atcCDEF*, and *atcGHI*. It appears that all of the genes within each operon are translationally coupled as in all cases the start codon of the gene overlaps with the stop codon of the gene immediately upstream. Similar operon-type organizations in PKS clusters are well-known and have been postulated to facilitate co-translation of the genes within the operon to yield equimolar amounts of proteins for optimal interactions.[17]

The extent of the *atc* gene cluster was deduced on the basis of sequence synteny with other *Streptomyces* genomes. The protein coding sequences (PCSs) located immediately adjacent to *atcA-I* show striking synteny with the genome of *Streptomyces coelicolor* A3(2) (SCO).[18] The PCSs upstream of *atcA* showed

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479

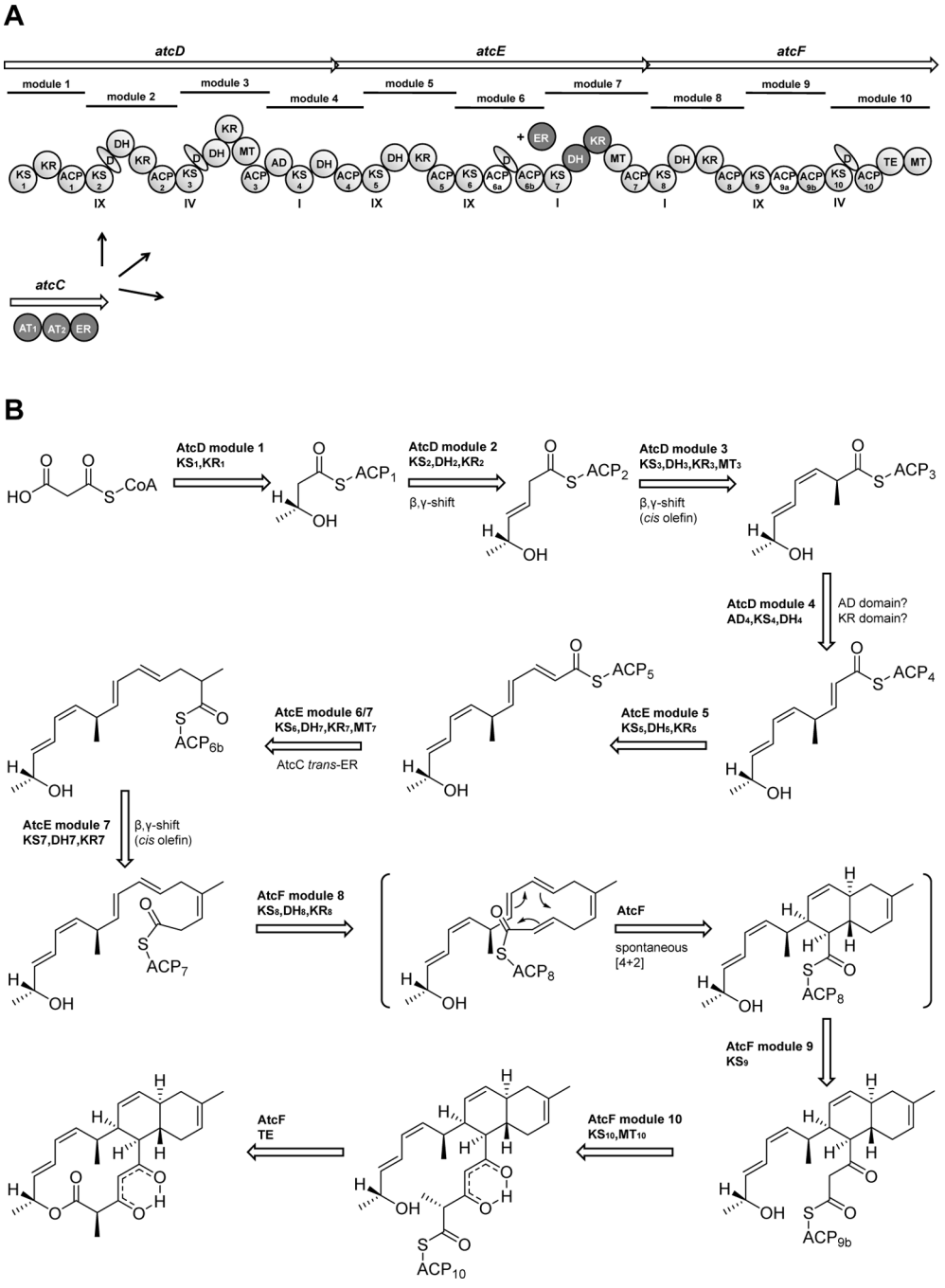


Figure 4. (A) Architecture of the *atc* PKS AtcC-AtcF. Dark shading indicates iteratively used domains: ACP, acyl-carrier protein; AD, alcohol dehydrogenase; AH, acylhydrolase; AT, acyltransferase; D, docking domain; DH, dehydratase; KS, β-ketoacyl synthase; KR, β-ketoreductase; MT, methyltransferase; ER, enoylreductase; TE, thioesterase. KS clades are indicated in roman numerals below each domain. (B) Proposed **1** biosynthetic pathway; all reactions likely occur while the growing polyketide chain is covalently attached to the *atc* PKS.

very high amino acid identity with SCO1013 (Mut-like protein, 85% identity), SCO1014 (apolipoprotein acyltransferase involved in cell envelope biosynthesis, 82% identity), SCO1017 (lipoprotein, 86% identity), and continuing beyond SCO1020 with 80−88% amino acid identity; furthermore, a blast search showed well over 20 kb of sequence upstream *atcA* sharing more than 80% nucleotide identity with that region of *S. coelicolor* A3(2). The PCSs located downstream of *atcI* showed homology with SCO5535 and SCO5536, which encode the highly conserved α and ε chains, respectively, of an acetyl-CoA carboxylase β-subunit essential for the primary and secondary metabolism of *S. coelicolor*.[19] Therefore, while the boundaries of the *atc* biosynthetic gene cluster are yet to be determined experimentally, it is reasonable to propose that *atcA* and *atcI* represent the left- and right-hand ends of the *atc* cluster, respectively.

Concomitant with this work, the biosynthetic gene cluster for **1** from *Streptomyces* sp. CNH365 was identified; the genetic organization of the biosynthetic genes for **1** is identical in both strains, while the genes outside the proposed gene cluster completely differ between both strains, providing further support for the proposed boundaries of the *atc* cluster

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479

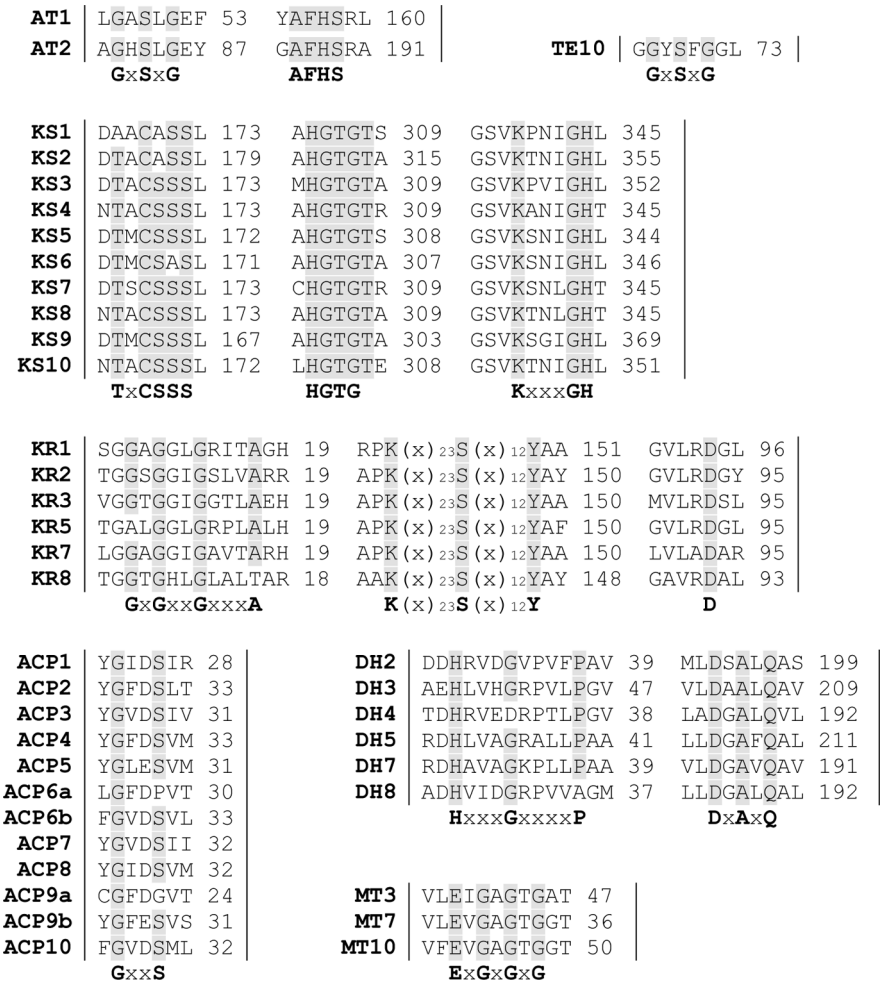


Figure 5. Clustal W alignments of the core regions of AT domains, DH domains, KS domains, KR domains, MT domains, and TE domains. Conserved amino acid residues are highlighted in gray. The numbers indicate amino acid positions within sequences and (x)*n* indicates the number of amino acids that separate the active-site residues. The conserved consensus motif is given in bold letters.

(Jungmann *et al.*[20]). This work also describes the biosynthetic gene cluster of chlorotonil A from *S. cellulosum* 1525. The chlorotonil A gene cluster is similar to that for **1** with the same number of genes but a different PKS domain organization. In addition, there are two extra genes encoding the expected halogenase and a discrete methyltransferase. Again, no apparent *post*-PKS genes are present to account for the modifications seen in the polyketide backbone. These combined observations are consistent with the notion that, in both cases, formation of the decalin system and noncanonical arrangement of double bonds occurs on the PKS.

Heterologous Expression of the Anthracimycin Biosynthetic Gene Cluster in *S. coelicolor*. A P1-derived artificial chromosome (PAC) library of genomic DNA of *Streptomyces* sp. T676 was custom-made by Bio S&T Co. (Montreal, Canada). The PAC library was screened with the three primer pairs listed in Table S2, amplifying fragments flanking either side and in the center of the *atc* cluster. Two PAC clones showed amplification with all three primer pairs and the smaller PAC (designated PAC-14M) containing a 100 kb insert was chosen for heterologous expression studies. On the basis of the insert size and genome sequence, it was ascertained that no other secondary metabolite cluster is encoded on PAC-14M.

PAC-14M, as well as the corresponding empty pESAC13 vector control, were first moved into the nonmethylating *Escherichia coli* ET12567 strain, containing the driver plasmid pR9604, *via* triparental mating,[21,22] and introduced into the improved *S. coelicolor* host strains M1146, M1152, and M1154 by conjugation.[21,23] After cultivation in production media, production of **1** was detected in all strains containing PAC-14 M using HPLC-UV, LCMS and an isolated standard of **1** (Figure 3 and Figure S1). Average production levels of ex-conjugants from strains M1146, M1152, and M1154 were 13.8 ± 2.4, 9.7 ± 2.6, and 8.6 ± 0.5 μg/mL, respectively.

Anthracimycin Is Assembled from Malonyl-CoA. To verify that **1** is biosynthesized from malonyl-CoA, consistent with a *trans*-AT PKS product, we performed stable isotope labeling studies with [1,2-$^{13}C_2$] sodium acetate.[24] To aid this, the production profile of **1** in GPP production medium was determined: production of **1** starts after 1 day and the maximum titer is achieved between days six and nine (*ca.* 100 μg/mL). Labeled substrate was fed to growing cultures between days three and seven and the cultures extracted after day eight. Following purification, ^{13}C NMR analysis indicated the incorporation of 11 intact acetate units as expected (Figure S2 and Table S3).

Bioinformatic Analysis of the Anthracimycin Gene Cluster. During the course of this study, the genome sequence

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479

of *Streptomyces* sp. NRRL F-5065 was reported (accession no. JOHV00000000). This organism is closely related to *Streptomyces* sp. T676 and contains a putative *atc* cluster (cluster 1 from 1 to 53 432 bps). The F-5065 sequence was obtained with Illumina technology and therefore may comprise a less reliable assembly than the PacBio sequence, but is likely to have a more accurate sequence at the nucleotide level. A 186 bp insertion and a 15 bp omission were identified in the PKS genes *atcD* and *atcE*, respectively. These regions were PCR-amplified and Sanger-sequenced, and the PacBio sequence was confirmed to be correct for strain T676 in both cases. The differences are therefore either errors in F-5065 sequence or real differences between the strains (see Methods).

Potential Self-Resistance and Regulatory Mechanisms. Resistance to **1** is most likely associated with AtcA, a major facilitator superfamily (MFS) transporter that is predicted to have 12 transmembrane helices according to analysis with TMHMM.[16] Multidrug transporter gene expression is often regulated by multidrug-binding transcriptional regulators.[25] The transcriptional regulator AtcB of the *atc* cluster belongs to the TetR family of regulators (Pfam entry PF00440) which are known to act as homodimeric transcriptional repressors of the multidrug resistance pump, and to control the expression of genes that direct the biosynthesis of natural products.[26,27] Numerous TetR family regulators have been identified in the biosynthetic clusters for antibiotics and other secondary metabolites in species of *Streptomyces* and related Actinobacteria. Some TetR family regulators (*i.e.*, ActR, SimR) have been shown to bind the products or intermediates of the biosynthetic pathways in which they are encoded.[28,29] These TetR family regulators primarily regulate the expression of efflux pumps required for antibiotic export but may also regulate the expression of late-stage biosynthetic genes. On the basis of the orientation and proximity of *atcB* relative to adjacent genes on the chromosome, it can be classified into the type II family of TetR regulators which are predicted to be co-transcribed with one or more adjacent genes.[30]

The Anthracimycin Megasynthase Is a Type I *trans*-AT PKS. The *atc* cluster contains four multifunctional genes *atcCDEF* with predicted PKS type I functions. The domain organization of *atcDEF* identified 10 PKS modules without internal AT domains (Figure 4A, Table 1). Those domains typical of polyketide production identified within the *atc* cluster are summarized below.

In trans Activities (AtcC). *atcC* encodes a protein with three distinct domains. The N-terminal and central domains contain the conserved core motif GxSxG.[31,32] Also, an AFHS motif is present, including the most significant phenylalanine residue, suggesting malonyl-CoA as the substrate (Figure 5).[33,32] Blastp and clustalW alignments revealed that the two AT domains share only 45% similarity with each other, suggesting different roles. While the N-terminal AT domain (AT1) showed highest similarity (62%) with the AT1 domain of OzmM involved in oxazolomycin biosynthesis,[34] AT2 showed highest similarity (67%) with the AT and enoyl reductase (ER) CorA from corallopyronin A biosynthesis (Figure S3A).[35] The presence of multiple *trans*-AT homologues, either as discrete proteins or fused with additional domains, including a second AT-like domain, is a common feature of *trans*-AT PKS gene clusters. Combined *in vivo* and biochemical studies have shown distinct functions for these different AT-like domains, with one encoding the malonyltranferase activity required to load the PKS carrier proteins, while the other acts as an acyl hydrolase (AH) domain for the removal of aberrant or stalled acyl units from blocked *trans*-AT PKS modules.[36] Phylogenetic analysis of AtcC suggests that the N-terminal domain is an acyl hydrolase (AH), while the central domain comprises the AT activity (Figure S3A).

The C-terminal domain of AtcC shows similarity to proteins of the oxidoreductase superfamily (Pfam 03060) and most likely serves as a *trans*-acting enoyl reductase.[37] Such *trans*-acting ERs are involved in polyunsaturated fatty acid biosynthesis and their involvement in *trans*-AT PKS pathways has been proven biochemically.[37] ER domains acting *in trans* are also found associated with the iterative modular PKSs of fungi.[38]

KS Domains. The PKS encodes 10 KS domains, all of which contain the conserved active-site cysteine (TxCxxS) and two histidine residues (HGTGT; GSVKxxxGH) (Figure 5).[39] The KS domains of *trans*-AT PKSs generally group according to their substrate types, suggesting a close link between KS evolution and substrate specificity.[7] Phylogenetic analysis revealed that three KSs (4, 7, and 8) of the *atc* cluster belong to clade I which recognize carbon branched substrates.[7] Four KSs (2, 5, 6, and 9) group into clade IX, indicating an olefinic thioester substrate. KS3 and KS10 fit into clade IV which might be specific for β-hydroxyl substrates, and is a paraphyletic group basal clade V which indicates a saturated or olefinic substrate.[7] KS1 shares highest identity with KS1 of the disorazol (Dis) PKS (60%) and KS1 of the macrolactin (Mln) PKS (52%).[40,41] DisKS1 could not be grouped into any clade,[7] but interestingly, all three PKSs responsible for anthracimycin, disorazol, and macrolactin biosynthesis lack a loading module and we suggest that they may be considered as a new clade of KSs associated with chain elongation initiation (Figure S3B).

KR Domains. All KR domains contain the Rossmann fold (GxGxxGxxxA) which is required for NADP(H) binding, although KR5 and KR8 have a slight variation thereof (Figure 5).[42,43] They all contain the conserved lysine-serine-tyrosine catalytic triad that is necessary for ketoreduction,[44] and have the diagnostic aspartic acid residue which in most cases allows their assignment as B-type leading to D-configured β-hydroxy products.[45]

ACP Domains. All ACPs contain the signature GxDS motif around the invariable 4′-phosphopantetheine attachment site serine residue, other than the tandem ACP6a and ACP9a domains.[46] The presence of tandem ACP domains within modules of *trans*-AT PKSs is often associated with non-canonical functionality.[37] Consistent with this, we suggest AtcE module 6 may represent a "skipped" extension module with module 7 operating iteratively, or, alternatively, that an extension reaction competent module 6 might utilize the reductive domains of module 7, in concert with the *in trans* activity of the C-terminal AtcC ER domain, in order to generate the required structural modification which includes a non-canonically located double bond after the seventh round of chain elongation.

Methyltransferase (MT) Domains. All three MTs contain the conserved *S*-adenosylmethionine binding motif ExGxGxG found in *C*-methyltransferases and belong to Pfam family Methyltransf_12 (PF08242).[47] The incorporation of methyl groups from *S*-adenosylmethionine catalyzed by *cis*-MTs has been described for other *trans*-AT systems, *e.g.*, chivosazol biosynthesis.[48] Methylation of the α-carbon is predicted to occur after the third, sixth and final chain extension reactions.

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479

DH Domains. Modules 2−5, 7, and 8 contain regions similar to DH domains which contain the active site HxxxGxxxxP signature motif with a conserved histidine residue and D(A/V)(V/A)(A/L)(Q/H) with a conserved aspartic acid residue.[49] These conserved histidine and aspartic acid residues comprise the catalytic dyad of dehydratases, and all DH domains of the *atc* cluster are proposed to be functional.[50] In conjunction with the all B-type KR domains, the DH domains are all proposed to generate *trans* double bonds.

Thioesterase (TE) Domain. Module 10 of AtcF contains a domain which shows significant similarity with PKS thioesterase (TE) domains and contains the GxSxG motif found in other TE domains but does not contain a second conserved GxH motif.[51]

Docking (D) Domains. The remnants of four embedded ATs are found within the PKS and have been proposed to function as "AT docking domains" or recognition elements between the *trans*-AT (AtcC) and PKS modules in AtcD, AtcE, and AtcF. Docking domains lack critical catalytic motifs (GxSxG and AFHS) and are therefore catalytically inactive.[52]

Proposed Pathway for Anthracimycin Biosynthesis. Analysis of the *atc* PKS architecture allowed us to propose a model for **1** biosynthesis (Figure 4) which accounts for the structural features observed. As no acetate-specific AT was identified, we suggest that the biosynthetic route may be initiated by loading of a malonyl residue catalyzed by the AtcC acyltransferase protein followed by decarboxylation providing the appropriately acylated KS1 domain of AtcD. The first two rounds of chain extension would proceed to generate an ACP2 bound triketide intermediate with a directly introduced β,γ-double bond rather than the canonical α,β-double bond typically produced by PKSs. This is in line with experimental observations for bacillaene (Bae) and rhizoxin (Rhi) biosynthesis on *trans*-AT PKSs.[53,54] In both cases, double bonds are introduced by an apparently canonical module, and sequence and phylogenetic analysis of the associated DH domains (DHs of Rhi-PKS modules 7 and 8, DHs of Bae-PKS modules 11−13) fails to identify any obviously difference from typical PKS DH domains. Careful analysis of the *atc* DH2 domain was consistent with a typically functional DH domain (Figure 5).

The third chain elongation step again most likely occurs with introduction of a β,γ-double bond (but with *cis* geometry), to yield a conjugated diene. β,γ-shifted conjugated dienes are rarely found in polyketides, but one is present in rhizoxin (all *trans*). During rhizoxin biosynthesis, introduction of the second β,γ-shifted double bond involves a dedicated "DB shift" module containing an unusual DH* domain which lacks critical catalytic residues.[54] However, sequence analysis of the *atc* DH3 domain does not suggest any unusual functionality that would be associated with a β,γ-shift or the observed *cis* stereochemistry. The striking anomaly for this module is the presence of a Zn-dependent alcohol dehydrogenase-like domain immediately after the ACP, and it is unclear whether this plays a role in the β,γ-double bond shift, or is required to fulfill a noncanonical KR-like function for the fourth extension module which lacks a KR domain. β,γ-Shifting of the diene moiety intermediate prior to further chain elongation means the α-carbon of the triketide intermediate becomes sp^3 hybridized, and thus after chain elongation by modules 4 and 5, the introduction of two further double bond occurs without formation of an extended tetraene system. This may be critical for the putative [4 + 2] cycloaddition proposed to occur later in the pathway.

Chain extension by modules 4 and 5 leads to a hexaketide intermediate with the introduction of a canonical α,β-unsaturated diene moiety. Module 4 is split between two proteins (AtcD and AtcE), a structural modification which may be important in allowing the pentaketide intermediate to access either the KR5 domain or the AD domain which sits between modules 4 and 5. Split module organization is not unusual in *trans*-AT PKSs, *e.g.*, in the chivosazol, difficidin or myxovirescin PKSs.[48,55,56]

A significant deviation from colinearity then occurs for module 6 which houses no β-keto processing or MT domains. An additional nonfunctional ACP6a domain is present, and we hypothesize that α-methylation and β-keto reduction after the sixth extension event are catalyzed by the domains present in module 7 with *in trans* reduction of the double bond by the AtcC ER domain. The aberrant, cross module functioning of MT domains has also been observed for oxazolomycin biosynthesis.[34] This is followed by chain extension on module 7 with the direct introduction of a β,γ-shifted double bond with *cis* geometry: the DH7 domain does not appear to deviate from typical DHs according to sequence analysis (Figure 5).

Chain elongation by module 8 appears to occur in a canonical manner to give an α,β-unsaturated nonaketide attached to ACP8, an intermediate which appears ideal for the postulated [4 + 2] cycloaddition. The putative *trans*-substituted dienophile is electron deficient and an *endo* transition state arising from the conformation shown in Figure 4B should lead to *trans*-decalin formation with the appropriate geometry observed in **1**. This may occur while attached to ACP8 or possibly on module 9 which contains a tandem ACP9a/9b pair. The final two rounds of chain elongation occur with no reductive β-keto modification and the final module 10 contains an MT domain, albeit, unusually, at the C-terminal end of a TE domain; this is consistent with the C2 methyl group of **1**. The TE domain will catalyze chain release with concomitant macrolactone formation.

Discussion. The screening of microbial fermentation extracts remains an effective method for the discovery of antibacterial molecules with new structural architectures and novel modes of action as exemplified by the recent discovery of teixobactin.[57] Anthracimycin **1** and the structurally related chlorotonil A **2** are exciting molecules with excellent activity against Gram-positive pathogens including *B. anthracis* and methicillin-resistant and vancomycin-resistant *S. aureus* (MRSA and VRSA).[4,1] Chlorotonil A exhibits additional activity against laboratory and clinical isolates of *Plasmodium falciparum* with no cross resistance to existing treatments (Jungmann and Müller, personal communication).[58] Both compounds display good *in vivo* efficacy, but their biophysical properties are not optimal for development as therapeutic agents.

On the basis of these observations, and intrigued by the novel structure of **1**, we decided to investigate its mode of biosynthesis and uncover the mechanism responsible for formation of the *trans*-decalin system which we hypothesized could involve an intramolecular [4 + 2] cycloaddition. Identification and validation of the biosynthetic gene cluster for **1** was greatly facilitated by recent technological developments. First, single molecule sequencing approaches using the PacBio SMRT RSII platform enabled the rapid whole genome sequencing of *Streptomyces* sp. T676 to yield a single high-quality genomic contig that includes complete assemblies for the loci encoding repetitive modular biosynthetic gene clusters. This approach was recently validated with reporting of the

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479

Streptomyces leeuwenhoekii genome sequence.[59] Analysis of this data using the biosynthetic gene cluster prediction software antiSMASH 3.0 allowed straightforward identification of the *atc* gene cluster which surprisingly encoded a *trans*-AT PKS.[13] We then used heterologous expression of the *atc* cluster in enhanced *S. coelicolor* hosts to verify its functionality for the biosynthesis of **1**.[23] To achieve this, the *atc* cluster was cloned into a PAC (P1-derived artificial chromosome) vector which was conjugatively transferred into *S. coelicolor* hosts where it stably and irreversibly integrated into the genomes.[60]

Analysis of the *atc* cluster suggests that the *trans*-AT PKS gene products AtcC to AtcF are sufficient for the biosynthesis of **1** as we can find no genes encoding discrete enzymes that might act upon a final PKS-free product, or even a PKS bound product, to generate the *trans*-decalin moiety *via* an intramolecular [4 + 2] cycloaddition. Enzymes capable of catalyzing [4 + 2] cycloadditions have been identified for a small number of natural product biosynthetic pathways, including for the generation of decalin moieties.[61−67] Very recently, Liu and coworkers showed that an FAD dependent enzyme PyrE3 from the pyrroindomycin pathway is essential for formation of the *trans*-decalin moiety of this molecule and acts upon a linear polyene produced by the *pyr cis*-AT PKS.[68] PyrE3 shows similarity to a characterized hydroxylase OxyS from the oxytetracycline pathway, but no homologues to *pyrE3* or *oxyS2* were identified in the *atc* cluster. In fact, after careful analysis, we could not identify genes that might encode products related to any of the small number of thus far validated enzymes capable of catalyzing [4 + 2] cycloadditions.[68]

To the best of our knowledge, the only other enzyme implicated in a [4 + 2] cycloaddition on an ACP-tethered intermediate is the LovB PKS required for biosynthesis of the fungal metabolite lovastatin, which also yields a *trans*-decalin product.[65,66,69] For several of the characterized [4 + 2] cycloaddition enzymes, it is possible that the cyclization reaction can occur spontaneously. As we highlight above, the nonaketide intermediate attached to AtcF (linked to either ACP8 or ACP9) has an electron deficient dieneophile with *trans* geometry which is ideal for the preferred *endo* selective reaction leading to the stereochemcial outcome observed (consistent with the Woodward−Hoffmann rules). By virtue of being covalently attached to the PKS, and with the rigidity and orientation imparted by the two β,γ-shifted double bond systems, this intermediate has reduced degrees of freedom compared with an equivalent enzyme free species meaning that the putative intramolecular [4 + 2] cycloaddition may be aided by an "entropy trap" mechanism rather than through any particular transition state stabilization.

Modifying PKSs through targeted synthase re-engineering is a promising approach for accessing functionally optimized polyketides.[70] To date, most examples of successful modular PKS bioengineering have been for *cis*-AT PKSs. This is in part due to their more common occurrence but is also because of their close adherence to the co-linearity principle which allows "rational" modification. However, the proposed evolution of these systems, by gene duplication and subsequent mutational optimization,[8] might suggest that components of *cis*-PKSs have been optimized *"in situ"* to perform specific tasks within an assembly line context and are hence less obviously "trans-plantable". Conversely, while *trans*-AT PKSs are inherently less easy to assign functionally, due to often significant deviations from colinearity and the common occurrence of apparently superfluous enzymatic domains, their highly mosaic architecture and quite different evolutionary origins[6,7] may be inherently more amenable to bioengineering. Success in this endeavor will, however, require advances in our mechanistic and structural understanding of these less well investigated systems.

The studies we report here, in conjunction with those reported in parallel by Jungmann *et al.*,[20] now set the scene for more detailed experiments to determine the functionality of the *atc* PKS proteins which may aid us in their bioengineering to produce modified variants of **1** that may have value as leads for further optimization as antibacterial agents.

METHODS

Chemicals and Molecular Biology Materials. Antibiotics, buffer components, and media compounds were purchased from Sigma-Aldrich, Roquette Pharma, and ADM (Archer Daniels Midland Company). Anthracimycin standard was kindly provided by MerLion Pharma (Singapore). [1,2-$^{13}C_2$] sodium acetate was purchased from Cambridge Isotope Laboratories, Inc. Molecular biology enzymes, reagents and kits were used according to manufacturer's instructions. PCR amplifications were performed with GoTaq Green Master Mix (Promega).

Bacterial Strains, Plasmids, and Culture Conditions. *E. coli* and *Streptomyces* strains, plasmids, and oligonucleotides used or constructed in this study are listed in Tables S1and S2. *E. coli* DH5α was used as general cloning host and was cultivated in liquid or on solid LB medium at 37 °C.[71] *E. coli* ET12567 was used as the host for conjugal transfer of plasmids to *S. coelicolor*.[72] The wild type producer strain of anthracimycin, *Streptomyces* sp. T676, was kindly provided by MerLion Pharma (Singapore).

Streptomyces strains were routinely cultivated in liquid TSB medium,[21] SV2 seed medium (15 g/L glucose, 15 g/L glycerol, 15 g/L soy peptone, 1 g/L $CaCO_3$; pH 7.0),[73] GPP production medium (5 g/L glucose, 80 g/L glucidex-12 malto dextrine (Roquette), 60 g/L Pharmamedium (ADM), 8 g/L $CaCO_3$; pH 7.0) (Merlion Pharma, Singapore) (optimized for the wild type producer strain T676), and anthracimycin production medium (10 g/L starch, 4 g/L yeast extract, 2 g/L peptone, 1 g/L $CaCO_3$, 40 mg/L $Fe_2(SO_4)_3{\cdot}4\ H_2O$, 100 mg/L KBr, 40 g/L sea salts) (used for the cultivation of the heterologous host strains).[1] Antibiotic concentrations were used as follows: 50 μg/mL apramycin, 100 μg/mL carbenicillin, 25 μg/mL chloramphenicol, 50 μg/mL fosfomycin, 50 μg/mL kanamycin, 25 μg/mL nalidixic acid, and 25 μg/mL thiostrepton. Standard methods for cultivation, DNA isolation and manipulation were performed according to standard procedures for *E. coli* and *Streptomyces*.[74,21]

Isolation of Genomic DNA from *Streptomyces* T676. High molecular weight genomic DNA was extracted from *Streptomyces* sp. T676 according to the salting out procedure of the *Streptomyces* manual with some modifications:[21] 0.5 mL of wet mycelium from a 30 h SV2 culture was washed with 10 mL 10% sucrose before resuspension in 5 mL SET buffer (75 mM NaCl, 25 mM EDTA, 20 mM TrisHCl pH 8.0) containing 200 μL of lysozyme (50 mg mL^{-1}) and 15 μL of ribonuclease A (10 mg mL^{-1}). The cells were incubated at 37 °C overnight, and 300 μL of fresh lysozyme was added after *ca.* 17 h followed by additional 2 h incubation. The following steps were performed according to Kieser *et al.*.[21]

Genome Sequencing, Cluster Annotation, and Bioinformatic analysis. Briefly, genomic DNA was sequenced with Pacific Biosciences (PacBio) RSII SMRT technology (commissioned to The Genome Analysis Centre (TGAC), Norwich, U.K.); the assembly obtained five contigs of approximately 11.7, 5919.7, 15.7, and 968.1 kb eventually merged in a 6.9 MB contig that contains most of the chromosome, and a 56.5 kb contig that represents a plasmid. The 6.9 MB contig was submitted to Prodigal server for identification of protein coding sequences (PCSs);[75] putative gene function was assigned with BASys server.[76] Putative gene clusters for the

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479

biosynthesis of specialized metabolites were identified with antiSMASH 3.0 server.[13]

The anthracimycin biosynthetic gene cluster was analyzed in detail using Artemis for visualization and annotation;[77] gene function was assigned based on a combination of searches with BLAST at NCBI server,[78] Pfam at Sanger server,[15] and Conserved Domain Database (CDD) at NCBI server;[14] modules and domains of the polyketide synthase were assigned on the basis of Pfam, CDD and antiSMASH search output. The frame Plot function of Artemis was used to confirm the extent of PCSs and for identification of putative frame-shifts due to errors in the sequence.[79] One obvious frame-shift was identified causing the split of a PKS gene in two; PCR amplification and sequencing of a segment covering the region identified a missing G causing the frame-shift. Furthermore, in comparison with *Streptomyces* sp. NRRL F-5065 (accession no. JOHV00000000), a 186 bp insertion and a 15 bp omission were identified in *atcD* and *atcE*, respectively. These regions were PCR-amplified and Sanger-sequenced. PCR conditions were as follows: 98 °C, 2 min; 30 cycles of 98 °C, 30 s; 60 °C, 30 s; 72 °C, 90 s and a final extension cycle at 72 °C, 7 min. GoTaq Green Master Mix (Promega) and 2.5% dimethyl sulfoxide (DMSO) were used for all amplifications; primers are listed in Table S2. The sequence of the anthracimycin biosynthetic gene cluster has been deposited in the European Nucleotide Archive under the accession number LN871452.

Production, Purification, and Analysis of Anthracimycin. *Streptomyces* T676 was precultivated in TSB medium for 2 days at 30 °C and 250 rpm. A 10% of the TSB culture was used to inoculate SV2 seed medium. After 2 days cultivation at 30 °C and 250 rpm, 10% of the seed culture was used to inoculate the GPP or anthracimycin production medium, respectively. The cultivation was continued for 8 days at 28 °C (GPP medium) or 30 °C (anthracimycin medium), and 250 rpm. For analytic purposes, 1 mL bacterial culture was acidified with HCl to pH 2 and extracted with an equal volume of ethyl acetate. After evaporation of the solvent, the residue was redissolved in 300 μL of methanol. Ten microliters was analyzed by HPLC over a Gemini-NX C_{18} 110A column (150 × 4.6 mm, 3 μm; Phenomenex) at a temperature of 25 °C and a flow rate of 1 mL/min. A linear gradient from 40% to 100% solvent B (solvent A, H_2O/HCOOH 99:1; solvent B, methanol 100%) over 17 min was used. UV detection was carried out at 285 nm. An amount of 1 mM authentic anthracimycin was used as standard. Samples were analyzed on a Shimadzu LC−MS system equipped with a NexeraX2 liquid chromatograph (LC30AD) fitted with a Prominence photo diode array detector (SPD-M20A) and an LCMS-IT-TOF mass spectrometer. Culture extracts were prepared as described for the HPLC analysis. Samples (typically 5 μL) were analyzed on a Kinetex XB C_{18} 100A (50 × 2.1 mm, 2.6 μm; Phenomenex) at a temperature of 40 °C and a flow rate of 0.6 mL/min. A linear gradient from 2% to 100% solvent B in solvent A over 7 min was used (solvent A, H_2O/HCOOH 99:1; solvent B, methanol 100%). UV detection was carried out at 285 nm. The mass spectrometer was operated in the electrospray mode with positive ionization. Identification of **1** was made by comparison of retention times and fragmentation patterns with those from authentic samples.

Preparation and Screening of Phage P1 Artificial Chromosome (PAC) Library. A P1-derived artificial chromosome (PAC) library of genomic DNA of *Streptomyces* sp. T676 was custom-made and screened by Bio S&T Co. (Montreal, Canada). Mycelium from a 2 days SV2 seed culture was prepared for PAC library construction. The mycelium was washed twice in 2 vol of buffer (200 mM NaCl, 10 mM Tris-HCl pH 8.0, 100 mM EDTA), resuspended in 20% glycerol, and flash frozen in liquid nitrogen before being stored at −80 °C. Three PCR primer pairs (Table S2) were designed for PAC library screening to isolate positive clones containing the entire **1** biosynthetic gene cluster, and were tested beforehand using *Streptomyces* sp. T676 genomic DNA (20 ng/μL) as PCR template. PCR conditions were: 95 °C, 3 min; 25 cycles of 95 °C, 30 s; 63 °C, 30 s; 72 °C, 45 s and a final extension cycle at 72 °C, 7 min. GoTaq Green Master Mix (Promega) and 2.5% dimethyl sulfoxide (DMSO) were used for all amplifications. Screening of 1536 PAC clones (= 4 × 384 well plates) resulted in two clones (PAC-14M and PAC-18J) showing amplification with all three primer pairs. Gel electrophoresis suggested the inserts to be 100 and 155 kb in size, respectively. The smaller PAC, designated PAC-14M, containing the 100 kb insert, was chosen for heterologous expression studies.

Heterologous Expression of PAC-14M in *S. coelicolor* Host Strains. The *E. coli*−*Streptomyces* Artificial Chromosome vector (pESAC13) contains an *ori*T, a ΦC31 integrase gene and a ΦC31 *attP* site that permit conjugation into *Streptomyces* strains and integration at the *attB* recombination locus. The *E. coli* DH10B strain containing PAC-14M was used in a triparental mating with an *E. coli* Top 10 strain that contained the driver plasmid pR9406, and the nonmethylating *E. coli* strain ET12567.[80,72] With the use of overnight starter cultures, these three strains were grown to an OD_{600} of 0.4 and mixed together on LB plates without antibiotic selection. After 24 h, the resulting *E. coli* patch was streaked on selective LB plates. The resulting *E. coli* ET12567 ex-conjugants were tested by PCR to confirm that they contained the entire PAC-14M clone insert. The *E. coli* ET12567 strain containing pR9406 and PAC-14M were conjugated with spores of *S. coelicolor* M1146, M1152, and M1154 according to standard protocols.[22,21,23] The SFM plates containing 10 mM $MgCl_2$ were overlaid after 18 h with thiostrepton and nalidixic acid. Ex-conjugants were streaked on new SFM plates containing $MgCl_2$, thiostrepton, and nalidixic acid at the same concentrations and allowed to grow for 9 days prior to harvesting spores for inoculating cultures. To create control strains, pESAC13 as well as *BamH*I digested and religated pESAC13 (to remove the pUC backbone what is lost during the PAC clone construction)[81] were also transferred to *E. coli* ET12567 by triparental mating for later conjugation into *S. coelicolor* strains. The strains were cultivated in anthracimycin production medium and analyzed using HPLC and LCMS in comparison to wild type samples and authentic anthracimycin standard as described in production, purification, and analysis of **1**.

Labeled Sodium Acetate Feeding and NMR Analysis. The GPP production media was inoculated with a 10% inoculum of the SV2 seed culture (see Culture Conditions). Aqueous solutions of [1,2-$^{13}C_2$] sodium acetate were passed through a 0.2 μm filter and added to the culture at concentrations of 2 mg per 100 mL medium from day three up to day seven. After a total incubation of 8 days at 28 °C and 250 rpm, the culture was harvested, acidified from pH 9 to pH 2 with concentrated hydrochloric acid, and extracted with an equal volume of ethyl acetate. The organic phase was taken to dryness, and the residue was dissolved in methanol and purified by solid phase extraction (Discovery reversed-phase columns DSC-Ph; Supelco) according to the manufacturers instruction and semipreparative reversed-phase HPLC with a Gemini-NX C_{18} 110A column (150 × 10 mm, 5 μm; Phenomenex) at a flow rate of 2 mL/min. A linear gradient from 85% to 100% solvent B in solvent A over 12 min was used (solvent A H_2O/HCOOH 99:1; solvent B MeOH 100%). The ^{13}C-labeled anthracimycin (2 mg) was dissolved in $CDCl_3$ (500 μL) and analyzed in a high precision NMR tube. NMR spectra were recorded at 500 MHz on a Bruker Avance III instrument equipped with a DUL cryoprobe at 25 °C. Chemical shifts in $CDCl_3$ are reported relative to the solvent peak (7.25 ppm).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acschembio.5b00525.

Table of strains, plasmids, and PCR primers; HPLC chromatograms of **1** production in heterologous host strains; ^{13}C NMR data and spectra of [1,2-$^{13}C_2$] sodium acetate labeled **1**; domain sequences and phylogenetic trees of AT and KS domains (PDF)

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479

AUTHOR INFORMATION

Corresponding Author

*E-mail: barrie.wilkinson@jic.ac.uk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We are grateful to Dr. Tony Buss (MerLion Pharma) for providing *Streptomyces* sp. T676. We thank Dr. Lijiang Song and Professor Gregory L. Challis (University of Warwick) for NMR support. This work was funded by the Biotechnology and Biological Sciences Research Council, U.K. *via* Institute Strategic Programme Grant BB/J004561/1 to the John Innes Centre.

REFERENCES

(1) Jang, K. H., Nam, S.-J., Locke, J. B., Kauffman, C. A., Beatty, D. S., Paul, L. A., and Fenical, W. (2013) Anthracimycin, a potent anthrax antibiotic from a marine-derived actinomycete. *Angew. Chem., Int. Ed. 52*, 7822−7824.

(2) Hensler, M. E., Jang, K. H., Thienphrapa, W., Vuong, L., Tran, D. N., Soubih, E., Lin, L., Haste, N. M., Cunningham, M. L., Kwan, B. P., Shaw, K. J., Fenical, W., and Nizet, V. (2014) Anthracimycin activity against contemporary methicillin-resistant Staphylococcus aureus. *J. Antibiot. 67*, 549−553.

(3) Corcoran, J. W. (1984) Mode of action and resistance mechanisms of macrolides. In *Macrolide Antibiotics* (Omura, S., Ed.), pp 231−259, Academic Press, Orlando, FL.

(4) Gerth, K., Steinmetz, H., Höfle, G., and Jansen, R. (2008) Chlorotonil A, a macrolide with a unique gem-dichloro-1,3-dione functionality from Sorangium cellulosum, So ce1525. *Angew. Chem., Int. Ed. 47*, 600−602.

(5) Staunton, J., and Weissman, K. J. (2001) Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep. 18*, 380−416.

(6) Piel, J. (2010) Biosynthesis of polyketides by trans-AT polyketide synthases. *Nat. Prod. Rep. 27*, 996−1047.

(7) Nguyen, T., Ishida, K., Jenke-Kodama, H., Dittmann, E., Gurgui, C., Hochmuth, T., Taudien, S., Platzer, M., Hertweck, C., and Piel, J. (2008) Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nat. Biotechnol. 26*, 225−233.

(8) Jenke-Kodama, H., and Dittmann, E. (2009) Evolution of metabolic diversity: insights from microbial polyketide synthases. *Phytochemistry 70*, 1858−1866.

(9) Scotti, C., Piatti, M., Cuzzoni, A., Perani, P., Tognoni, A., Grandi, G., Galizzi, A., and Albertini, A. M. (1993) A Bacillus subtilis large ORF coding for a polypeptide highly similar to polyketide synthases. *Gene 130*, 65−71.

(10) Cheng, Y.-Q., Tang, G.-L., and Shen, B. (2002) Identification and localization of the gene cluster encoding biosynthesis of the antitumor macrolactam leinamycin in Streptomyces atroolivaceus S-140. *J. Bacteriol. 184*, 7013−7024.

(11) Cheng, Y.-Q., Tang, G.-L., and Shen, B. (2003) Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc. Natl. Acad. Sci. U. S. A. 100*, 3149−3154.

(12) Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J., and Turner, S. (2009) Real-time DNA sequencing from single polymerase molecules. *Science 323*, 133−138.

(13) Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., Lee, S. Y., Fischbach, M. A., Müller, R., Wohlleben, W., Breitling, R., Takano, E., and Medema, M. H. (2015) antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res. 43*, W237.

(14) Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., and Bryant, S. H. (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res. 43*, 222−226.

(15) Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L. L., Tate, J., and Punta, M. (2014) Pfam: the protein families database. *Nucleic Acids Res. 42*, 222−230.

(16) Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol. 305*, 567−580.

(17) Cane, D. E. (1997) Introduction: Polyketide and Nonribosomal Polypeptide Biosynthesis. From Collie to Coli. *Chem. Rev. 97*, 2463−2464.

(18) Bentley, S. D., Chater, K. F., Cerdeño-Tárraga, A.-M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.-H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M.-A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J., and Hopwood, D. A. (2002) Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature 417*, 141−147.

(19) Rodriguez, E., Banchio, C., Diacovich, L., Bibb, M. J., and Gramajo, H. (2001) Role of an essential acyl coenzyme A carboxylase in the primary and secondary metabolism of Streptomyces coelicolor A3(2). *Appl. Environ. Microbiol. 67*, 4166−4176.

(20) Jungmann, K., Jansen, R., Gerth, K., Huch, V., Krug, D., Fenical, W., and Müller, R. (*ACS Chem. Biol.*2015) Two of a kind: The biosynthetic pathways of chlorotonil and anthracimycin. DOI: 10.1021/acschembio.5b00523.

(21) Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000) *Practical Streptomyces Genetics*, John Innes Foundation, Norwich, U.K..

(22) Jones, A. C., Gust, B., Kulik, A., Heide, L., Buttner, M. J., and Bibb, M. J. (2013) Phage p1-derived artificial chromosomes facilitate heterologous expression of the FK506 gene cluster. *PLoS One 8* (7), e69319.

(23) Gomez-Escribano, J. P., and Bibb, M. J. (2011) Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters. *Microb. Biotechnol. 4*, 207−215.

(24) Simpson, T. J. (1998) Application of Isotopic Methods to Secondary Metabolic Pathways. *Top. Curr. Chem. 195*, 1−48.

(25) Schumacher, M. A., Miller, M. C., Grkovic, S., Brown, M. H., Skurray, R. A., and Brennan, R. G. (2002) Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR. *EMBO J. 21*, 1210−1218.

(26) Ramos, J. L., Martnez-Bueno, M., Molina-Henares, A. J., Terán, W., Watanabe, K., Zhang, X., Gallegos, M. T., Brennan, R., and Tobes, R. (2005) The TetR family of transcriptional repressors. *Microbiol. Mol. Biol. Rev. 69*, 326−356.

(27) Cuthbertson, L., and Nodwell, J. R. (2013) The TetR family of regulators. *Microbiol. Mol. Biol. Rev. 77*, 440−475.

(28) Tahlan, K., Ahn, S. K., Sing, A., Bodnaruk, T. D., Willems, A. R., Davidson, A. R., and Nodwell, J. R. (2007) Initiation of actinorhodin export in Streptomyces coelicolor. *Mol. Microbiol. 63*, 951−961.

(29) Le, T. B. K., Schumacher, M. A., Lawson, D. M., Brennan, R. G., and Buttner, M. J. (2011) The crystal structure of the TetR family transcriptional repressor SimR bound to DNA and the role of a flexible

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479

N-terminal extension in minor groove binding. *Nucleic Acids Res. 39*, 9433−9447.
(30) Ahn, S. K., Cuthbertson, L., and Nodwell, J. R. (2012) Genome context as a predictive tool for identifying regulatory targets of the TetR family transcriptional regulators. *PLoS One 7* (11), e50562.
(31) Haydock, S. F., Aparicio, J. F., Molnár, I., Schwecke, T., Khaw, L. E., König, A., Marsden, A. F., Galloway, I. S., Staunton, J., and Leadlay, P. F. (1995) Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases. *FEBS Lett. 374*, 246−248.
(32) Yadav, G., Gokhale, R. S., and Mohanty, D. (2003) SEARCHPKS: A program for detection and analysis of polyketide synthase domains. *Nucleic Acids Res. 31*, 3654−3658.
(33) Reeves, C. D., Murli, S., Ashley, G. W., Piagentini, M., Hutchinson, C. R., and McDaniel, R. (2001) Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry 40*, 15464−15470.
(34) Zhao, C., Coughlin, J. M., Ju, J., Zhu, D., Wendt-Pienkowski, E., Zhou, X., Wang, Z., Shen, B., and Deng, Z. (2010) Oxazolomycin biosynthesis in Streptomyces albus JA3453 featuring an ″acyltransferase-less″ type I polyketide synthase that incorporates two distinct extender units. *J. Biol. Chem. 285*, 20097−20108.
(35) Erol, O., Schäberle, T. F., Schmitz, A., Rachid, S., Gurgui, C., El Omari, M., Lohr, F., Kehraus, S., Piel, J., Müller, R., and König, G. M. (2010) Biosynthesis of the myxobacterial antibiotic corallopyronin A. *ChemBioChem 11*, 1253−1265.
(36) Jensen, K., Niederkrüger, H., Zimmermann, K., Vagstad, A. L., Moldenhauer, J., Brendel, N., Frank, S., Pöplau, P., Kohlhaas, C., Townsend, C. A., Oldiges, M., Hertweck, C., and Piel, J. (2012) Polyketide proofreading by an acyltransferase-like enzyme. *Chem. Biol. 19*, 329−339.
(37) Bumpus, S. B., Magarvey, N. A., Kelleher, N. L., Walsh, C. T., and Calderone, C. T. (2008) Polyunsaturated fatty-acid-like trans-enoyl reductases utilized in polyketide biosynthesis. *J. Am. Chem. Soc. 130*, 11614−11616.
(38) Kennedy, J., Auclair, K., Kendrew, S. G., Park, C., Vederas, J. C., and Hutchinson, C. R. (1999) Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science 284*, 1368−1372.
(39) White, S. W., Zheng, J., Zhang, Y.-M., and Rock, C. O. (2005) The structural biology of type II fatty acid biosynthesis. *Annu. Rev. Biochem. 74*, 791−831.
(40) Kopp, M., Irschik, H., Pradella, S., and Müller, R. (2005) Production of the tubulin destabilizer disorazol in Sorangium cellulosum: biosynthetic machinery and regulatory genes. *ChemBioChem 6*, 1277−1286.
(41) Schneider, K., Chen, X.-H., Vater, J., Franke, P., Nicholson, G., Borriss, R., and Süssmuth, R. D. (2007) Macrolactin is the polyketide biosynthesis product of the pks2 cluster of Bacillus amyloliquefaciens FZB42. *J. Nat. Prod. 70*, 1417−1423.
(42) Scrutton, N. S., Berry, A., and Perham, R. N. (1990) Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature 343*, 38−43.
(43) Aparicio, J. F., Caffrey, P., Marsden, A. F., Staunton, J., and Leadlay, P. F. (1994) Limited proteolysis and active-site studies of the first multienzyme component of the erythromycin-producing polyketide synthase. *J. Biol. Chem. 269*, 8524−8528.
(44) Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D. V., Hutchinson, C. R., and McDaniel, R. (2003) A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. *Biochemistry 42*, 72−79.
(45) Caffrey, P. (2003) Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases. *ChemBioChem 4*, 654−657.
(46) Joshi, V. C., and Wakil, S. J. (1971) Studies on the mechanism of fatty acid synthesis. XXVI. Purification and properties of malonyl-coenzyme A–acyl carrier protein transacylase of Escherichia coli. *Arch. Biochem. Biophys. 143*, 493−505.
(47) Kagan, R. M., and Clarke, S. (1994) Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependant methyltransferases suggests a common structure for these enzymes. *Arch. Biochem. Biophys. 310*, 417−427.
(48) Perlova, O., Gerth, K., Kaiser, O., Hans, A., and Müller, R. (2006) Identification and analysis of the chivosazol biosynthetic gene cluster from the myxobacterial model strain Sorangium cellulosum So ce56. *J. Biotechnol. 121*, 174−191.
(49) Bevitt, D. J., Cortes, J., Haydock, S. F., and Leadlay, P. F. (1992) 6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme. *Eur. J. Biochem. 204*, 39−49.
(50) Valenzano, C. R., You, Y.-O., Garg, A., Keatinge-Clay, A., Khosla, C., and Cane, D. E. (2010) Stereospecificity of the dehydratase domain of the erythromycin polyketide synthase. *J. Am. Chem. Soc. 132*, 14697−14699.
(51) Tsai, S. C., Miercke, L. J., Krucinski, J., Gokhale, R., Chen, J. C., Foster, P. G., Cane, D. E., Khosla, C., and Stroud, R. M. (2001) Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: versatility from a unique substrate channel. *Proc. Natl. Acad. Sci. U. S. A. 98*, 14808−14813.
(52) Tang, G.-L., Cheng, Y.-Q., and Shen, B. (2004) Leinamycin biosynthesis revealing unprecedented architectural complexity for a hybrid polyketide synthase and nonribosomal peptide synthetase. *Chem. Biol. 11*, 33−45.
(53) Calderone, C. T., Kowtoniuk, W. E., Kelleher, N. L., Walsh, C. T., and Dorrestein, P. C. (2006) Convergence of isoprene and polyketide biosynthetic machinery: isoprenyl-S-carrier proteins in the pksX pathway of Bacillus subtilis. *Proc. Natl. Acad. Sci. U. S. A. 103*, 8977−8982.
(54) Bretschneider, T., Heim, J. B., Heine, D., Winkler, R., Busch, B., Kusebauch, B., Stehle, T., Zocher, G., and Hertweck, C. (2013) Vinylogous chain branching catalysed by a dedicated polyketide synthase module. *Nature 502*, 124−128.
(55) Chen, X.-H., Vater, J., Piel, Jö., Franke, P., Scholz, R., Schneider, K., Koumoutsi, A., Hitzeroth, G., Grammel, N., Strittmatter, A. W., Gottschalk, G., Süssmuth, R. D., and Borriss, R. (2006) Structural and functional characterization of three polyketide synthase gene clusters in Bacillus amyloliquefaciens FZB 42. *J. Bacteriol. 188*, 4024−4036.
(56) Simunovic, V., Zapp, J., Rachid, S., Krug, D., Meiser, P., and Müller, R. (2006) Myxovirescin A biosynthesis is directed by hybrid polyketide synthases/nonribosomal peptide synthetase, 3-hydroxy-3-methylglutaryl-CoA synthases, and trans-acting acyltransferases. *ChemBioChem 7*, 1206−1220.
(57) Ling, L. L., Schneider, T., Peoples, A. J., Spoering, A. L., Engels, I., Conlon, B. P., Mueller, A., Schäberle, T. F., Hughes, D. E., Epstein, S., Jones, M., Lazarides, L., Steadman, V. A., Cohen, D. R., Felix, C. R., Fetterman, K. A., Millett, W. P., Nitti, A. G., Zullo, A. M., Chen, C., and Lewis, K. (2015) A new antibiotic kills pathogens without detectable resistance. *Nature 517*, 455−459.
(58) Held, J., Gebru, T., Kalesse, M., Jansen, R., Gerth, K., Müller, R., and Mordmüller, B. (2014) Antimalarial activity of the myxobacterial macrolide chlorotonil a. *Antimicrob. Agents Chemother. 58*, 6378−6384.
(59) Gomez-Escribano, J. P., Castro, J. F., Razmilic, V., Chandra, G., Andrews, B., Asenjo, J. A., and Bibb, M. J. (2015) The Streptomyces leeuwenhoekii genome: de novo sequencing and assembly in single contigs of the chromosome, circular plasmid pSLE1 and linear plasmid pSLE2. *BMC Genomics 16*, 485.
(60) Sosio, M., Giusino, F., Cappellano, C., Bossi, E., Puglia, A. M., and Donadio, S. (2000) Artificial chromosomes for antibiotic-producing actinomycetes. *Nat. Biotechnol. 18*, 343−345.
(61) Oikawa, H., Yagi, K., Watanabe, K., Honma, M., and Ichihara, A. (1997) Biosynthesis of macrophomic acid: plausible involvement ofintermolecular Diels−Alder reaction. *Chem. Commun. 1*, 97−98.
(62) Ose, T., Watanabe, K., Mie, T., Honma, M., Watanabe, H., Yao, M., Oikawa, H., and Tanaka, I. (2003) Insight into a natural Diels-Alder reaction from the structure of macrophomate synthase. *Nature 422*, 185−189.

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479

(63) Oikawa, H., Katayama, K., Suzuki, Y., and Ichihara, A. (1995) Enzymatic activity catalysing *exo*-selective Diels−Alder reaction in solanapyrone biosynthesis. *J. Chem. Soc., Chem. Commun.*, 1321−1322.
(64) Kasahara, K., Miyamoto, T., Fujimoto, T., Oguri, H., Tokiwano, T., Oikawa, H., Ebizuka, Y., and Fujii, I. (2010) Solanapyrone synthase, a possible Diels-Alderase and iterative type I polyketide synthase encoded in a biosynthetic gene cluster from Alternaria solani. *ChemBioChem 11*, 1245−1252.
(65) Auclair, K., Sutherland, A., Kennedy, J., Witter, D. J., Van den Heever, J. P., Hutchinson, C. R., and Vederas, J. C. (2000) Lovastatin nonaketide synthase catalyzes an intramolecular Diels-Alder reaction of a substrate analogue. *J. Am. Chem. Soc. 122*, 11519−11520.
(66) Ma, S. M., Li, J. W.-H., Choi, J. W., Zhou, H., Lee, K. K. M., Moorthie, V. A., Xie, X., Kealey, J. T., Da Silva, N. A., Vederas, J. C., and Tang, Y. (2009) Complete reconstitution of a highly reducing iterative polyketide synthase. *Science 326*, 589−592.
(67) Kim, H. J., Ruszczycky, M. W., Choi, S.-H., Liu, Y.-N., and Liu, H.-W. (2011) Enzyme-catalysed [4 + 2] cycloaddition is a key step in the biosynthesis of spinosyn A. *Nature 473*, 109−112.
(68) Wu, Q., Wu, Z., Qu, X., and Liu, W. (2012) Insights into pyrroindomycin biosynthesis reveal a uniform paradigm for tetramate/tetronate formation. *J. Am. Chem. Soc. 134*, 17342−17345.
(69) Campbell, C. D., and Vederas, J. C. (2010) Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers 93*, 755−763.
(70) Hertweck, C. (2015) Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. *Trends Biochem. Sci. 40*, 189−199.
(71) Bertani, G. (1951) Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. *J. Bacteriol. 62*, 293−300.
(72) MacNeil, D. J., Gewain, K. M., Ruby, C. L., Dezeny, G., Gibbons, P. H., and MacNeil, T. (1992) Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene 111*, 61−68.
(73) Wilkinson, B., Foster, G., Rudd, B. A., Taylor, N. L., Blackaby, A. P., Sidebottom, P. J., Cooper, D. J., Dawson, M. J., Buss, A. D., Gaisser, S., Böhm, I. U., Rowe, C. J., Cortés, J., Leadlay, P. F., and Staunton, J. (2000) Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polyketide synthase. *Chem. Biol. 7*, 111−117.
(74) Sambrook, J. and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, New York.
(75) Hyatt, D., Chen, G.-L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinf. 11*, 119.
(76) Van Domselaar, G. H., Stothard, P., Shrivastava, S., Cruz, J. A., Guo, A., Dong, X., Lu, P., Szafron, D., Greiner, R., and Wishart, D. S. (2005) BASys: a web server for automated bacterial genome annotation. *Nucleic Acids Res. 33*, 455−459.
(77) Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., and Barrell, B. (2000) Artemis: sequence visualization and annotation. *Bioinformatics 16*, 944−945.
(78) Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res. 25*, 3389−3402.
(79) Bibb, M. J., Bibb, M. J., Ward, J. M., and Cohen, S. N. (1985) Nucleotide sequences encoding and promoting expression of three antibiotic resistance genes indigenous to Streptomyces. *Mol. Gen. Genet. 199*, 26−36.
(80) Grant, S. G., Jessee, J., Bloom, F. R., and Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. *Proc. Natl. Acad. Sci. U. S. A. 87*, 4645−4649.
(81) Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene 33*, 103−119.

DOI: 10.1021/acschembio.5b00525
ACS Chem. Biol. 2015, 10, 2468−2479