Functional analysis of environmental DNA-derived type II polyketide synthases reveals structurally diverse secondary metabolites

Zhiyang Feng[a], Dimitris Kallifidas[a], and Sean F. Brady[a,b,1]

[a]Laboratory of Genetically Encoded Small Molecules, and [b]Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, NY 10065

Edited by Jerrold Meinwald, Cornell University, Ithaca, NY, and approved June 14, 2011 (received for review March 14, 2011)

A single gram of soil is predicted to contain thousands of unique bacterial species. The majority of these species remain recalcitrant to standard culture methods, prohibiting their use as sources of unique bioactive small molecules. The cloning and analysis of DNA extracted directly from environmental samples (environmental DNA, eDNA) provides a means of exploring the biosynthetic capacity of natural bacterial populations. Environmental DNA libraries contain large reservoirs of bacterial genetic diversity from which new secondary metabolite gene clusters can be systematically recovered and studied. The identification and heterologous expression of type II polyketide synthase-containing eDNA clones is reported here. Functional analysis of three soil DNA-derived polyketide synthase systems in *Streptomyces albus* revealed diverse metabolites belonging to well-known, rare, and previously uncharacterized structural families. The first of these systems is predicted to encode the production of the known antibiotic landomycin E. The second was found to encode the production of a metabolite with a previously uncharacterized pentacyclic ring system. The third was found to encode the production of unique KB-3346-5 derivatives, which show activity against methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant *Enterococcus faecalis*. These results, together with those of other small-molecule-directed metagenomic studies, suggest that culture-independent approaches are capable of accessing biosynthetic diversity that has not yet been extensively explored using culture-based methods. The large-scale functional screening of eDNA clones should be a productive strategy for generating structurally previously uncharacterized chemical entities for use in future drug development efforts.

Despite the historical success of bacterial natural products as lead structures for the development of small molecule therapeutics and the continued need for new antimicrobials and chemotherapeutics, large screening programs have deemphasized the use of microbial extracts over the past two decades. The reason most frequently cited for this decline is the persistent rediscovery of known metabolites (1–3). Most environmental bacteria remain recalcitrant to standard culture methods (4–6), and the difficulties associated with growing these organisms prohibit their use as new sources of bioactive small molecules. Although it is not yet possible to easily culture the majority of environmental bacteria, it is possible to extract microbial DNA directly from environmental samples (environmental DNA, eDNA) and to clone this DNA into cultured bacteria where it can be functionally characterized. This general approach has been termed metagenomics (7). The application of metagenomics to the study of bacterial secondary metabolism is particularly appealing in light of the fact that the genes required for the biosynthesis of a natural product are typically clustered on a bacterial chromosome. The heterologous expression of natural product gene clusters captured on individual clones or on small numbers of overlapping clones should provide a means of obtaining previously unidentified bioactive small molecules.

A structurally diverse collection of aromatic metabolites, including many important antimicrobials and anticancer agents, arise from iterative (type II, aromatic) polyketide synthases (PKSs) (8). Although the gene clusters that encode the biosynthesis of these diverse metabolites can differ substantially in gene content, they all encode a conserved minimal PKS composed of three proteins: ketosynthase alpha, KS_α; ketosynthase beta/chain length factor, KS_β; and acyl carrier protein, ACP. The minimal PKS is responsible for the iterative condensation of malonyl-CoAs into a nascent polyketide chain that is then cyclized, aromatized, reduced, oxidized, rearranged, and functionalized in pathway-specific ways to generate the extraordinary structural diversity that is known to arise from these systems (8, 9). PCR-based studies as well as shotgun-sequencing efforts indicate that eDNA samples are rich in unique minimal PKS genes (10–13). Through the functional characterization of eDNA-derived type II PKS containing clones, we have identified PKS systems that encode structurally diverse metabolites including compounds with unique and rare carbon skeletons. Among the metabolites we identified are compounds that show activity against both methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant (*vanA*) *Enterococcus faecalis* (VRE). These studies suggest that large-scale heterologous expression of eDNA clones containing diverse KS_β genes will be a productive strategy for producing previously unidentified bioactive metabolites that can be used in future drug discovery efforts.

Results and Discussion

Two previously archived eDNA cosmid-based libraries, constructed using DNA isolated from arid soils collected in Utah (UT) and Arizona (AZ), were screened for clones containing type II PKS systems. To identify type II PKS gene clusters in these libraries, cosmid DNA isolated from each library was used as the template in PCR reactions with degenerate primers designed to amplify full-length KS_β genes (10, 14–16). Amplicons of the correct predicted size (1.5 kb) were gel purified, sequenced, and compared to deposited KS_β genes from cultured bacteria (Fig. 1). Unique KS_β genes were used as probes to recover type II PKS containing clones from the archived libraries. Recovered cosmids were then retrofitted with the genetic elements required for conjugation and site-specific integration into *Streptomyces* spp. and

Author contributions: Z.F., D.K., and S.F.B. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The X-ray data and structure tables for compound **2** have been deposited in the Cambridge Structural Database (CSD), Cambridge Crystallographic Data Centre, Cambridge CB2 1EZ, United Kingdom (CSD reference no. 805477). The sequences reported in this paper have been deposited in the GenBank database [accession nos. HQ828985 (cosAZ154), HQ828986 (BAC:UT-X26/F129), and HQ828984 (cosAZ97)].

[1]To whom correspondence should be addressed. E-mail: sbrady@rockefeller.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1103921108/-/DCSupplemental.

conjugated from *Escherichia coli* S17.1 into *Streptomyces albus* for heterologous expression studies.

Fig. 1 shows a ClustalW-based phylogenetic tree containing eDNA-derived KS_β genes as well as KS_β genes from functionally characterized type II gene clusters that encode the biosynthesis of structurally diverse metabolites. It has been shown previously that KS_α proteins group into clades that correlate well with the chain length and initial cyclization pattern of the polyketide precursor produced by a gene cluster (17). Our analysis indicates that the above grouping is well retained for both KS_α and KS_β genes (Fig. 1). KS_βs, which are primarily responsible for determining the length of the polyketide chain generated by a minimal PKS (8), can therefore be used as bioinformatics markers for predicting both the length of the polyketide chain produced by a minimal PKS and the initial polyketide cyclization pattern of the polyketide chain. For this study, three cosmids (cosAZ97, cosUT-X26, and cosAZ154) that (*i*) contain KS_β genes from distinct clades and (*ii*) confer the production of clone-specific metabolites to *S. albus* were selected for extensive characterization. Small-scale (50 mL) test cultures of *S. albus* transformed with each cosmid were optimized for clone-specific metabolite production, and then extracts obtained from large-scale fermentations (>3 L) grown under optimized heterologous expression conditions were used to obtain sufficient material for structural characterization studies.

Full sequencing of cosAZ97 revealed that, in both sequence and gene organization, the eDNA insert closely resembled the landomycin biosynthetic gene cluster from *Streptomyces cyanogenus* S136 (18, 19) (*Supporting Information*). Culture broth extracts of *S. albus* transformed with cosAZ97 were therefore examined for the presence of landomycin derivatives. HPLC-MS revealed one major clone-specific metabolite in these extracts. Both MS [electrospray ionization (ESI) $[M\text{-}H]^-$ $m/z = 711.24$]

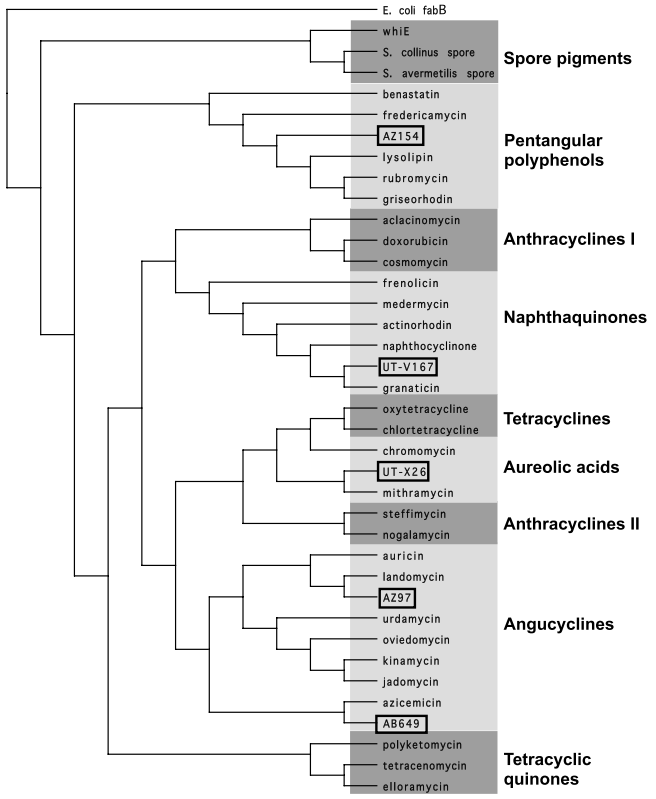

Fig. 1. A ClustalW-based phylogenetic tree containing eDNA-derived KS_β genes (boxes) and KS_β genes from sequenced PKS gene clusters that encode structurally diverse metabolites. The tree was constructed using the Neighbor-Joining method. The *fabB* gene from *E. coli* was used as an outgroup.

and UV data correspond to data reported for landomycin E (compound **1**), indicating that cosAZ97 encodes either the production of landomycin E itself or a close structural analog of landomycin E (18, 19) (Fig. 2 *A* and *C*). Landomycin E has previously been reported as a natural product of *Streptomyces globisporus 1912*.

A number of minor clone-specific metabolites were observed in culture broth extracts derived from *S. albus* transformed with cosUT-X26. Exhaustive attempts to isolate these minor metabolites were unsuccessful. We have observed that the appearance of large numbers of unstable minor metabolites in crude culture broth extracts is often an indication that the polyketide biosynthetic machinery for a particular biosynthetic pathway has not been captured in its entirety on an eDNA clone. To identify overlapping clones that could be assembled into a larger continuous fragment of DNA for use in heterologous expression experiments, the UT library was rescreened with primer pairs designed to recognize the distal ends of cosUT-X26. A single overlapping cosmid clone, cosUT-F129, was identified and recovered from the library in this screen. Transformation-associated recombination (TAR) in *Saccharomyces cerevisiae* was used to reassemble cosUT-X26 and cosUT-F129 into a continuous 60 kb eDNA fragment contained within a bacterial artificial chromosome (BAC: UT-X26/F129) (20–23) (*Supporting Information*). Although *S. albus* transformed with cosUT-X26 alone produced multiple minor clone-specific metabolites and *S. albus* transformed with cosUT-F129 alone did not produce any detectable clone-specific metabolites, *S. albus* transformed with BAC:UT-X26/F129 produced a single major clone-specific metabolite. Silica gel flash chromatography followed by reversed-phase HPLC was used to purify this metabolite (compound **2**) from ethyl acetate extracts of cultures grown for 7 d (30 °C, 200 rpm) in modified R5 medium. The structure of compound **2** was determined by X-ray crystallography and confirmed by 1- and 2D NMR (Fig. 2*B* and *Supporting Information*). To the best of our knowledge, the pentacyclic ring system found in **2** (Fig. 2*C*) does not appear in any previously reported natural or synthetic compounds.

Extracts from cultures of *S. albus* transformed with the third cosmid, cosAZ154, were found to contain two major clone-specific metabolites. Compounds **3** and **4** were purified by silica gel flash chromatography and reversed-phase HPLC from methanol washes of HP-20 resin (Supelo-Diaion™) added to ISP Medium 4 (Difco) grown cultures (12–14 d, 30 °C, 200 rpm). The structures of both metabolites were established by high-resolution MS (HRMS) and extensive 1- and 2D NMR experiments (Fig. 2*B* and *Supporting Information*). Compounds **3** and **4** (Fig. 2*C*) are previously uncharacterized natural products and represent members of the recently reported *Streptomyces* (strain FERM BP-109834) derived KB-3346-5 family of compounds (24).

Metagenomics severs the link between DNA and its source, thus the metabolite(s) produced by an eDNA-derived gene cluster in the context of its natural genetic background is not known. Compounds arising from metagenomic studies will undoubtedly represent a mixture of native metabolites, intermediates from the biosynthesis of native metabolites, as well as metabolites arising from unique interactions of eDNA genes with host biosynthetic machineries. In the case of cosAZ97, where the cloned eDNA closely resembles a previously sequenced biosynthetic gene cluster, we believe it is reasonable to assume that the heterologously produced landomycin derivative likely represents a natively produced metabolite. For cosAZ154 (*Supporting Information*) and BAC:UT-X26/F129 (*Supporting Information*), we were unable to identify closely related functionally characterized gene clusters that would permit analogous simple comparisons. The KB-3346-5 gene cluster from strain FERM BP-109834 likely resembles cosAZ154; however, because this gene cluster is not publicly available, no direct comparison was possi-

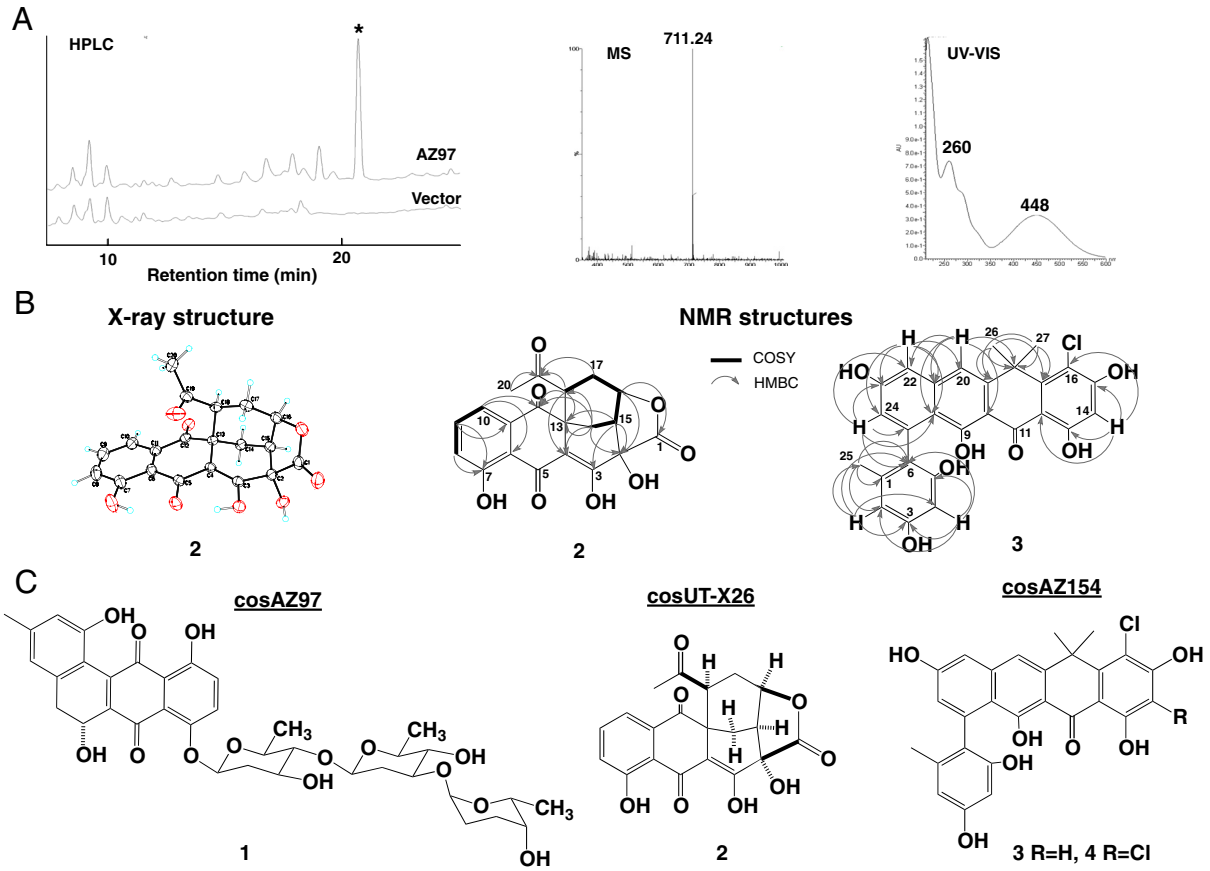

Fig. 2. (*A*) UV-visible chromatogram from the HPLC-MS analysis of organic extracts of *S. albus* transformed with cosAZ97. *S. albus* transformed with cosAZ97 produces one major clone-specific metabolite (*) that is not seen in vector control cultures. MS and UV spectra for this peak match previously reported data for landomycin E (**1**) (18, 19). (*B*) X-ray structure of compound **2**. Extensive 2D NMR data [COSY, heteronuclear multiple bond correlation (HMBC)] was used to define the structures of compounds **2**–**4**. (*C*) Compounds **1**–**4** were isolated from cultures of *S. albus* transformed with eDNA constructs cosAZ97, BAC:UT-X26/F129, and cosAZ154.

ble. Thirteen of the genes found on BAC:UT-X26/F129 have close relatives in a cryptic type II PKS gene cluster from *Streptomyces bingchenggensis* (25). BAC:UT-X26/F129 contains a number of genes predicted to encode polyketide modifying enzymes that are not encoded by the cryptic *S. bingchenggensis* gene cluster, including additional thioesterase, cyclase, ketoreductase, and oxidoreductase genes. It is therefore unlikely that these two gene clusters encode the biosynthesis of identical metabolites or even metabolites based on the same polyketide skeletons. Although comparisons to gene clusters functionally characterized in culture-based studies provided little insight into the biosynthetic origins of compounds **2**–**4**, their biosyntheses can be rationalized based on standard polyketide cyclization paradigms and information gleaned from the clade in which the KS_β gene found on each eDNA cosmid resides (Fig. 1).

Functionally characterized gene clusters containing KS_β genes from the same clade into which KS_β UT-X26 falls generate linear tetracyclic intermediates from decaketide precursors (26) (Fig. 3*A*). This linear tetracyclic intermediate is seen in both tetracycline and aureolic acid biosynthesis. It is therefore not surprising that many genes found on UT-X26/F129 have close homologs in tetracycline and aureolic-acid-type gene clusters (*Supporting Information*). Three predicted cyclases from UT-X26/F129 (ORFs 6, 34, and 29) are closely related to the three cyclases, MtmQ, Y, and X, which produce a linear tetracyclic intermediate in the biosynthesis of the aureolic acid mithramycin. In our biogenic proposal, ORF6, 34, and 29 are predicted to generate a tetracyclic intermediate as well. The reduction of the C-9 ketone by ORF28, a predicted C-9 ketoreductase, would explain the absence of the C-9 hydroxyl in **2**. In the biosynthesis of the aureolic acid family of antitumor compounds, the linear tetracyclic intermediate is converted to a tricyclic structure with a pentyl side chain through the oxidative cleavage of the fourth ring (26) (Fig. 3*A*). This cleavage reaction is accompanied by the loss of a carboxylate. In the biosynthetic proposal for compound **2** (Fig. 3*A*), the tetracyclic intermediate undergoes a similar oxidative cleavage reaction, however, the carboxylate is retained. No close homolog of the oxygenase responsible for the decarboxylation-coupled cleavage of the fourth ring in sequenced aureolic acid gene clusters appears on BAC:UT-X26/F129. A cleavage reaction that retains the carboxylate is therefore likely carried out by one of the nonhomologous oxygenases (ORF20 or 23) that could not be assigned to another role based on simple homology arguments (ORF15 and 25). Lactonization of the carboxylate and the pentyl side chain, followed by reduction and cyclization of the now butyl side chain with the new tetracyclic core would generate the pentacyclic ring system seen in **2**. The formation of both the fourth and fifth rings from the pentyl side chain appears to be unprecedented in the biosynthesis of polyketides characterized in culture-based studies. Although it is possible to rationalize the biosynthesis of **2** from the genes found on BAC:UT-X26/F129 (Fig. 3*A*), as with any heterologous expression experiment, it is also possible that the host proteome plays a role in the biosynthesis of **2**. BAC:UT-X26/F129 contains two predicted glycosyltransferases as well as a number of predicted sugar biosynthesis genes. Compound **2** may therefore represent an aglycone structure. It is not uncommon for heterologous expression experiments to yield aglycone structures as a result of either the

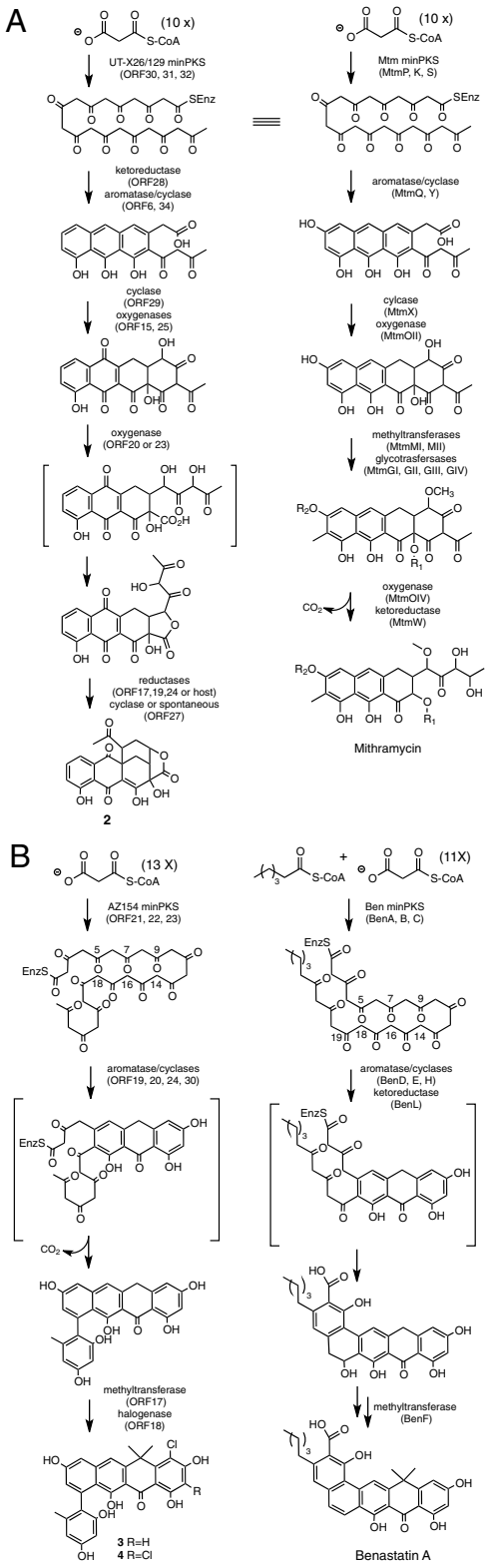

Fig. 3. (*A*) Compound **2** is predicted to arise from a decaketide precursor. Early steps in our proposed biosynthetic scheme match those seen in the biosynthesis of mithramycin-like metabolites. (*B*) The proposed biosynthetic scheme for compound **3** and **4** involves a 13-acetate precursor. Early steps in this cyclization scheme resemble those used in many pentangular biosynthetic pathways including the biosynthesis of the benestatins (29).

failure to heterologously produce required specialized sugars or the absence of required thymidine diphosphate sugars in the host primary metabolome (27, 28).

Fig. 3*B* shows a biogenic proposal for the formation of compounds **3** and **4** from the cyclization and functionalization of a tridecaketide precursor. KS_β AZ154, which is used in the biosynthesis of compounds **3** and **4**, falls in a clade with KS_β genes from functionally characterized minimal PKSs that produce long chain polyketides (>10 acetates) and pathways that proceed through related pentacyclic intermediates (29) (Fig. 3*B*). Our biosynthetic proposal for compounds **3** and **4** shares the C-9/C-14, C-7/C-16, and C-5/C-18 cyclization reactions found in the biosynthesis of metabolites using closely related KS_β genes (e.g., γ-rubromycin, fredericamycin, griseorhodin, benestatin). After the formation of the first three rings, the biosynthesis of compounds **3** and **4** appears to diverge from that seen in γ-rubromycin, fredericamycin, griseorhodin, and benestatin, which undergo keto reduction at C-19 and two subsequent cyclizations to generate what has been termed a "pentangular" polyphenolic intermediate (Fig. 3*B*). CosAZ154 does not contain a predicted C-19 ketoreductase homolog. Instead of the formation of a pentangular intermediate, we propose that a non-C-19 reduced intermediate undergoes C-2/C-19 and C-20/C-25 cyclization reactions to yield a different pentacyclic intermediate. Methylation (ORF17) and halogenation (ORF18) of this pentacyclic intermediate would give rise to compounds **3** and **4**. The biosynthesis of the pentangular derived metabolite benestatin undergoes the same rare geminal bis-methylation reaction that is predicted to be required for the formation of **3** and **4**.

At the highest concentration tested (50 μg/mL), the metabolites **2**–**4** showed no activity against yeast or gram-negative bacteria. Compounds **3** and **4** do, however, exhibit potent activity against gram-positive bacteria (Table 1). Both metabolites are active against MRSA (USA300) and VRE (EF16) (30). The specific halogenation pattern seen in compound **3** appears to be important for potent antimicrobial activity, as the monohalogenated derivative (**3**) is almost an order of magnitude more active than the dihalogenated derivative (**4**) against VRE. No previously described KB-3346-5 derivatives contain this halogenation pattern. Compound **3** is one of the first metabolites with potent bioactivity to be identified through the heterologous expression of soil-derived eDNA (31).

Conclusions

In previous studies, eDNA clones associated with two different KS_β genes, UT-V167 and AB649 (Fig. 4), were found to encode the biosynthesis of the pentacyclic metabolite erdacin (**5**) and previously uncharacterized fluostatin (**6**) derivatives, respectively (14, 15). Of the five eDNA-derived type II PKS systems that have now been functionally characterized in heterologous hosts (Fig. 4), two encode the biosynthesis of metabolites with ring systems that had not been previously reported from culture-based studies [BAC:UT-X26/F129 (**2**) and cosUT-V167 (**5**)], two encode the biosynthesis of structures based on ring systems that have only rarely been observed in culture-based studies [cosAB649 (**6**) and cosAZ154 (**3**, **4**)], and only one was found to encode the biosynth-

Table 1. Minimum inhibitor concentrations (μg/mL) for compounds 3 and 4 against yeast and bacteria; compound 2 did not show activity at the highest concentration tested (50 μg/mL)

Strain	3	4
Saccharomyces cerevisiae w303	>50	>50
Escherichia coli EC100	>50	>50
Ralstonia metallidurans CH34	>50	>50
Burkholderia thailandensis E264	>50	>50
Bacillus subtilis BR151	3.1	6.25
Staphylococcus aureus 6538P	3.1	6.25
Staphylococcus aureus USA300 (MRSA)	3.1	6.25
Enterococcus faecalis OG1RF	0.8	6.25
Enterococcus faecalis EF16 (VRE)	0.8	6.25

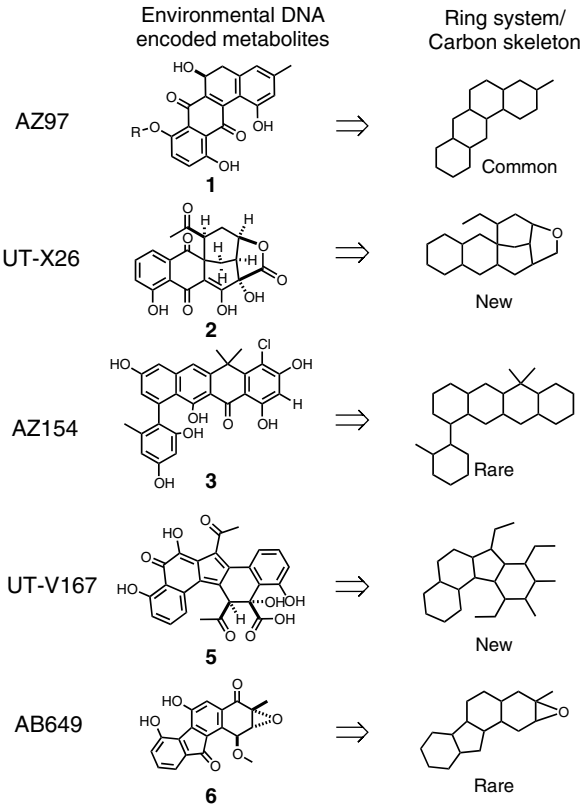

Fig. 4. Molecules characterized from the heterologous expression of eDNA clones containing type II polyketide synthases are shown. Compounds **1**–**4** were identified in this study. Aromatic polyketides characterized from heterologous expression studies using eDNA clones have often been found to contain either rare and previously uncharacterized carbon skeletons.

esis of a carbon skeleton that is well-represented among metabolites characterized in culture-based studies [cosAZ97 (**1**)]. The high frequency with which unique and rare structural families have been found in these studies indicates that metagenomics is likely providing access to biosynthetic diversity that has not yet been extensively explored using culture-based methods.

Methods

See *Supporting Information* for additional methods and analytical data.

Library Construction and Arraying. The two permanently archived eDNA cosmid libraries used in this study were previously constructed and archived using published protocols (16). Briefly, soil was collected in Utah and Arizona, sifted to remove large particulates, and then heated (70 °C) in lysis buffer [100 mM Tris·HCl, 100 mM EDTA, 1.5 M NaCl, 1% (wt/vol) cetyltrimethylammonium bromide, 2% (wt/vol) SDS (pH 8.0)] for 2 h. Soil particulates were removed from the crude lysate by centrifugation, and eDNA was precipitated from the resulting supernatant with the addition of 0.7 vol isopropanol. Crude eDNA was collected by centrifugation, washed with 70% ethanol, and resuspended in TE [10 mM Tris·HCl, 1mM EDTA (pH 8.0)]. Gel purified (1% agarose) high-molecular-weight eDNA was blunt ended (Epicentre, End-It), ligated into the SmaI site of pWEB or pWEB-TNC (Epicentre), packaged into lambda phage, and transfected into *E. coli* EC100. Unique approximately 5,000-membered sublibrary pools were archived as matching pairs of glycerol stocks and minipreps. Each library contains at least 10 million unique cosmid clones. Genetic manipulation of *E. coli* and *Streptomyces* were performed according to standard protocols (32, 33).

KS_β Screening, Clone Recovery, Retrofitting, and Mating. Cosmid DNA isolated from each library served as the template to amplify KS_β sequences using previously described conditions and degenerate primers (dp:KS-TTCGGSGGITTCCAGWSIGCSATG and dp:ACP-TCSAKSAGSGCSAISGASTCGTAICC) (10). Twenty-five microliter PCR reactions contained 50 ng of cosmid DNA, 2.5 μM of each primer, 2 mM dNTPs, 1× ThermoPol reaction buffer (New England Biolabs), 0.5 unit *Taq* DNA polymerase (New England Biolabs), and 5% DMSO. PCR was conducted using the following touchdown protocol: denaturation (95 °C, 2 min), 8 touchdown cycles [95 °C, 45 s; 65 °C (–1 °C per cycle), 1 min; 72 °C, 2 min], 35 standard cycles (95 °C, 45 s; 58 °C, 1 min; 72 °C, 2 min), and a final extension step (72 °C, 2 min). Sublibrary pools from which unique KS_β genes were amplified were used as the starting point for the recovery of unique KS_β containing eDNA clones. Overnight cultures of sublibraries containing KS_β genes of interest were plated into 96-well microtiter plates at a dilution of 10^{-5} or 10^{-6}. After 18 h at 37 °C, the diluted cultures were screened by whole-cell PCR using KS_β specific primers. PCR positive wells were then plated directly on solid media to yield distinct colonies that were screened in a second round of whole-cell PCR. Recovered cosmids were digested with PsiI, which cuts in the cosmid vector but rarely cuts in eDNA inserts. Each linearized cosmid was then ligated with the 6.81 kb DraI fragment from pOJ436. This fragment contains an origin of transfer (oriT), an apramycin resistance marker, and elements of the phiC31 attp-attB integration system needed for integration into diverse *Streptomyces* spp. (34). Retrofitted cosmids were transformed into *E. coli* S17.1 and conjugated into *Streptomyces* using published protocols. Fifty-milliliter test cultures were grown (30 °C, 7–21 d) in modified R5 or ISP Medium 4 media both with and without 5% HP-20 resin. Ethyl acetate extracts of crude culture broths as well as methanol washes of the HP-20 resin were examined for clone-specific metabolites.

TAR Reassembly of Overlapping Clones. PCR primers designed to recognize the sequence at each end of an eDNA clone were used to identify and recover overlapping cosmid clones. In the UT library, cosUT-F129 was the only overlapping clone identified for cosUT-X26. The AZ library yielded overlapping clones for both ends of cosAZ154 (cosAZ809 and cosAZ1678). TAR was used to reassemble these overlapping cosmid clones into larger insert BAC clones (14, 21) (*Supporting Information*).

Compound Fermentation, Analysis, Isolation, and Purification. *S. albus* recombinants containing either cosAZ154 or cosAZ97 were grown in ISP Medium 4 containing Diaion™ 5% HP-20 resin at 30 °C (200 rpm) for 12 d. The HP-20 resin was collected from the fermentation broth, washed with water, and then flushed with methanol. The methanol eluent was fractionated by silica gel flash chromatography using a $CHCl_3$:MeOH step gradient. Compounds **3** and **4** eluted from this column with the 95:5 $CHCl_3$:MeOH fraction. Each metabolite was then purified by reversed-phase (XBridge™ C18, 10 × 150 mm, 5 μm) HPLC (7 mL/min) using a linear gradient from 10:90 H_2O:MeOH (containing 0.1% formic acid) to 100% MeOH (containing 0.1% formic acid) over 30 min (compound **3**, 22.2 min; compound **4**, 23.1 min). Compounds **3** and **4** are each produced at approximately 0.5 mg/L. *S. albus* transformed with the larger TAR construct, BAC:AZ1678/154/809, produced the same two clone-specific metabolites, and therefore, cultures of *S. albus* transformed with cosAZ154 alone were used for all molecule isolation studies. Compound **3** HRMS-ESI (*m/z*): $[M\text{-}H]^-$: calculated for $C_{27}H_{20}O_7Cl$ 491.0898; found 491.0913. Compound **4** HRMS-ESI (*m/z*): $[M\text{-}H]^-$: calculated for $C_{27}H_{19}O_7Cl_2$ 525.0508; found 525.0508.

S. albus recombinants containing UT-X26-associated eDNA constructs were fermented in modified R5 (for 1 L: 100 g sucrose, 0.25 g K_2SO_4, 10.12 g $MgCl_2 \bullet 6H_2O$, 10 g glucose, 0.1 g casamino acids, 5 g yeast extract, 21 g MOPS, 2 g NaOH, 2 mL R2YE trace elements) at 30 °C and 200 rpm for 7 d. Cultures were then extracted twice with an equal volume of ethyl acetate. Crude ethyl acetate extracts were initially fractionated by silica gel flash chromatography using a $CHCl_3$:MeOH step gradient. Compound **2** eluted with 80:20 $CHCl_3$:MeOH. Compound **2** was purified by reversed-phase (XBridge™ C18, 10 × 250 mm, 5 μm) HPLC (3.6 mL/min) using a linear gradient from 60:40 H_2O:MeOH (with 0.1% acetic acid) to 100% MeOH (with 0.1% acetic acid) over 30 min (retention time of 23 min). Compound **2** is produced at approximately 15 mg/L. Compound **2** was crystallized by slow evaporation from H_2O:MeOH. Compound **2**, HRMS-ESI (*m/z*): $[M + Na]^+$: calculated for $C_{20}H_{16}O_8Na$ 407.0743; found 407.0753.

Bioactivity Assays. Overnight cultures grown in either LB or yeast extract-peptone-glucose medium were diluted 10^6-fold. Aliquots (150 μL) of the dilute cultures were added to individual wells of a 96-well plate. Compounds were resuspended in methanol at 10 mg/mL and then diluted 100-fold into culture media. This solution (150 μL) was added to the first well of the microtiter plate and then serially diluted twofold per well across the plate. The plates were incubated at 30 °C for 18–24 h. Minimum inhibitor concentrations are reported as the lowest concentration at which no bacterial growth was observed.

ACKNOWLEDGMENTS. We thank Emil Lobkovsky from the Cornell University X-ray crystallography facility. We thank Dr. John D. Bauer for his assistance with library construction and clone recovery. We thank Ken, Beth, and Mikaela Morris for their assistance with sample collection. S.F.B. is a Howard Hughes Medical Institute Early Career Scientist. This work was supported by National Institutes of Health Grants GM077516 and U54-AI057158.

1. Projan SJ (2003) Why is big Pharma getting out of antibacterial drug discovery. *Curr Opin Microbiol* 6:427–430.
2. Baltz RH (2006) Marcel Faber roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration. *J Ind Microbiol Biotechnol* 33:507–513.
3. von Nussbaum F, Brands M, Hinzen B, Weigand S, Habich D (2006) Antibacterial natural products in medicinal chemistry—exodus or revival. *Angew Chem Int Ed Engl* 45:5072–5129.
4. Rappe MS, Giovannoni SJ (2003) The uncultured microbial majority. *Annu Rev Microbiol* 57:369–394.
5. Torsvik V, Ovreas L, Thingstad TF (2002) Prokaryotic diversity—magnitude, dynamics, and controlling factors. *Science* 296:1064–1066.
6. Torsvik V, Goksoyr J, Daae FL (1990) High diversity in DNA of soil bacteria. *Appl Environ Microbiol* 56:782–787.
7. Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. *Chem Biol* 5:R245–R249.
8. Hertweck C, Luzhetskyy A, Rebets Y, Bechthold A (2007) Type II polyketide synthases: Gaining a deeper insight into enzymatic teamwork. *Nat Prod Rep* 24:162–190.
9. Shen B (2000) Biosynthesis of aromatic polyketides. *Top Curr Chem* 209:1–51.
10. Seow KT, et al. (1997) A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: A means to access and use genes from uncultured microorganisms. *J Bacteriol* 179:7360–7368.
11. Wawrik B, et al. (2007) Biogeography of actinomycete communities and type II polyketide synthase genes in soils collected in New Jersey and Central Asia. *Appl Environ Microbiol* 73:2982–2989.
12. Wawrik B, Kerkhof L, Zylstra GJ, Kukor JJ (2005) Identification of unique type II polyketide synthase genes in soil. *Appl Environ Microbiol* 71:2232–2238.
13. Pang M, Tan GA, Abdullah N, Lee C, Ng C (2008) Phylogenetic analysis of type I and type II polyketide synthase from tropical forest soil. *Biotechnology* 7:660–668.
14. Feng Z, Kim JH, Brady SF (2010) Fluostatins produced by the heterologous expression of a TAR reassembled environmental DNA derived type II PKS gene cluster. *J Am Chem Soc* 132:11902–11903.
15. King RW, Bauer JD, Brady SF (2009) An environmental DNA-derived type II polyketide biosynthetic pathway encodes the biosynthesis of the pentacyclic polyketide erdacin. *Angew Chem Int Ed Engl* 48:6257–6261.
16. Brady SF (2007) Construction of soil environmental DNA cosmid libraries and screening for clones that produce biologically active small molecules. *Nat Protoc* 2:1297–1305.
17. Metsa-Ketela M, et al. (2002) Molecular evolution of aromatic polyketides and comparative sequence analysis of polyketide ketosynthase and 16S ribosomal DNA genes from various streptomyces species. *Appl Environ Microbiol* 68:4472–4479.
18. Erb A, Krauth C, Luzhetskyy A, Bechthold A (2009) Differences in the substrate specificity of glycosyltransferases involved in landomycins A and E biosynthesis. *Appl Microbiol Biotechnol* 83:1067–1076.
19. Ostash B, Korynevska A, Stoika R, Fedorenko V (2009) Chemistry and biology of landomycins, an expanding family of polyketide natural products. *Mini Rev Med Chem* 9:1040–1051.
20. Kouprina N, Larionov V (2008) Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast Saccharomyces cerevisiae. *Nat Protoc* 3:371–377.
21. Kim JH, et al. (2010) Cloning large natural product gene clusters from the environment: Piecing environmental DNA gene clusters back together with TAR. *Biopolymers* 93:833–844.
22. Larionov V, et al. (1994) Transformation-associated recombination between diverged and homologous DNA repeats is induced by strand breaks. *Yeast* 10:93–104.
23. Larionov V, et al. (1996) Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc Natl Acad Sci USA* 93:491–496.
24. Satoshi O, et al. (2009) New KB-3346-5 substance and method for producing the same. Japanese Patent JP2009046404.
25. Wang XJ, et al. (2010) Genome sequence of the milbemycin-producing bacterium Streptomyces bingchenggensis. *J Bacteriol* 192:4526–4527.
26. Lombo F, Menendez N, Salas JA, Mendez C (2006) The aureolic acid family of antitumor compounds: Structure, mode of action, biosynthesis, and novel derivatives. *Appl Microbiol Biotechnol* 73:1–14.
27. Rodriguez E, Menzella HG, Gramajo H (2009) Heterologous production of polyketides in bacteria. *Methods Enzymol* 459:339–365.
28. Luzhetskyy A, et al. (2007) Cloning and heterologous expression of the aranciamycin biosynthetic gene cluster revealed a new flexible glycosyltransferase. *Chembiochem* 8:599–602.
29. Lackner G, et al. (2007) Biosynthesis of pentangular polyphenols: Deductions from the benastatin and griseorhodin pathways. *J Am Chem Soc* 129:9306–9312.
30. Diep BA, et al. (2006) Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant Staphylococcus aureus. *Lancet* 367:731–739.
31. Brady SF, Simmons L, Kim JH, Schmidt EW (2009) Metagenomic approaches to natural products from free-living and symbiotic organisms. *Nat Prod Rep* 26:1488–1503.
32. Sambrook J, Russell DW (2001) Molecular cloning: A laboratory manual. (Cold Spring Harbor Laboratory Press, Plainview, NY).
33. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical *Streptomyces* genetics. (The John Innes Foundation, Norwich, UK).
34. Bierman M, et al. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to Streptomyces spp. *Gene* 116:43–49.