Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*

Luisa Laureti[a,b,1,2], Lijiang Song[c,1], Sheng Huang[a,b,3], Christophe Corre[c], Pierre Leblond[a,b], Gregory L. Challis[c,4], and Bertrand Aigle[a,b,4]

[a]Unité Mixte de Recherche1128 Génétique et Microbiologie, Université Henri Poincaré, Institut Fédératif de Recherche 110 EFABA, F-54506 Vandœuvre-lès-Nancy, France; [b]Unité Mixte de Recherche 1128 Génétique et Microbiologie, Institut National de la Recherche Agronomique, F-54506 Vandœuvre-lès-Nancy, France; and [c]Department of Chemistry, University of Warwick, CV4 7AL Coventry, United Kingdom

Edited by Chaitan Khosla, Stanford University, Stanford, CA, and accepted by the Editorial Board February 25, 2011 (received for review December 20, 2010)

There is a constant need for new and improved drugs to combat infectious diseases, cancer, and other major life-threatening conditions. The recent development of genomics-guided approaches for novel natural product discovery has stimulated renewed interest in the search for natural product-based drugs. Genome sequence analysis of *Streptomyces ambofaciens* ATCC23877 has revealed numerous secondary metabolite biosynthetic gene clusters, including a giant type I modular polyketide synthase (PKS) gene cluster, which is composed of 25 genes (nine of which encode PKSs) and spans almost 150 kb, making it one of the largest polyketide biosynthetic gene clusters described to date. The metabolic product(s) of this gene cluster are unknown, and transcriptional analyses showed that it is not expressed under laboratory growth conditions. The constitutive expression of a regulatory gene within the cluster, encoding a protein that is similar to Large ATP binding of the LuxR (LAL) family proteins, triggered the expression of the biosynthetic genes. This led to the identification of four 51-membered glycosylated macrolides, named stambomycins A–D as metabolic products of the gene cluster. The structures of these compounds imply several interesting biosynthetic features, including incorporation of unusual extender units into the polyketide chain and in *trans* hydroxylation of the growing polyketide chain to provide the hydroxyl group for macrolide formation. Interestingly, the stambomycins possess promising antiproliferative activity against human cancer cell lines. Database searches identify genes encoding LAL regulators within numerous cryptic biosynthetic gene clusters in actinomycete genomes, suggesting that constitutive expression of such pathway-specific activators represents a powerful approach for novel bioactive natural product discovery.

antibiotic | silent gene cluster | genome mining | structure elucidation | anticancer agent

Natural products continue to be of utmost importance because of their broad spectrum of applications in medicine and agriculture (1, 2). In the past decade, microbial genomics has uncovered numerous cryptic secondary metabolite biosynthetic pathways not associated with the production of known metabolites (3), suggesting that the reservoir of bioactive natural products is far from exhausted, even in actinomycetes, which are responsible for the production of nearly 70% of known microbial natural products. Such cryptic biosynthetic pathways promise to deliver a multitude of novel bioactive compounds with the potential to rejuvenate stalling drug discovery pipelines. Several strategies have been developed to identify the metabolic products of cryptic biosynthetic pathways (4). However, many cryptic biosynthetic gene clusters appear to be expressed poorly, or not at all, under laboratory growth conditions (i.e., they are silent). This presents a major potential bottleneck for this exciting approach to novel natural product discovery, and the activation of these pathways is thus a major challenge of current interest (4).

Streptomyces ambofaciens is well known to produce two antibiotics: the macrolide spiramycin, used for a long time in human therapy as an antibacterial agent and for the treatment of toxoplasmosis; and the pyrrole-amide congocidine (5). Bioinformatics analyses of the *S. ambofaciens* partially sequenced chromosome led to the identification of the congocidine biosynthetic gene cluster (6) and provided the complete sequence of the gene cluster that directs spiramycin biosynthesis (7). More importantly, these analyses also revealed numerous cryptic putative secondary metabolite biosynthetic gene clusters, mostly located in the chromosome arms (8). The metabolic products of three of these gene clusters have been identified as the known siderophores coelichelin and desferrioxamines E/B and the known kinamycin angucyclinone antibiotics, none of which had been previously reported as metabolites of *S. ambofaciens* (9–11). The metabolic products of the other cryptic biosynthetic gene clusters remain unknown.

We have focused our attention on a large cryptic type I modular polyketide synthase (PKS) gene cluster located in the right arm of the *S. ambofaciens* chromosome. Polyketides are of particular interest because they encompass several different chemical classes, including macrolides, polyenes, aromatics, and polyethers, which have found therapeutic applications as antibiotics, antitumor agents, immunosuppressants, and cholesterol-lowering drugs (12). Here we report the identification of stambomycins A–D, four polyketides that are metabolic products of the cryptic type I PKS gene cluster. The stambomycins constitute a unique family of 51-membered glycosylated macrolides with promising antiproliferative activity against human cell lines. Constitutive expression of a gene within the cryptic gene cluster encoding a putative pathway-specific activator triggered the expression of the PKS genes (which are not expressed under laboratory growth conditions), providing the key to the identification of the stambomycins.

Results and Discussion

Identification and in Silico Analysis of a Unique Type I PKS Gene Cluster in *S. ambofaciens* ATCC23877. Among the cryptic biosynthetic gene clusters identified by analysis of the partial *S. ambofaciens* genome

Author contributions: P.L., G.L.C., and B.A. designed research; L.L., L.S., S.H., and C.C. performed research; L.L., L.S., G.L.C., and B.A. analyzed data; and L.L., L.S., G.L.C., and B.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. C.K. is a guest editor invited by the Editorial Board.

[1]L.L. and L.S. contributed equally to this work.

[2]Present address: Faculté de Médecine, Université Paris Descartes, INSERM U1001, 75730 Paris Cedex 15, France.

[3]Present address: Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA 50011.

[4]To whom correspondence may be addressed. E-mail: bertrand.aigle@scbiol.uhp-nancy.fr or g.l.challis@warwick.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1019077108/-/DCSupplemental.

sequence, one cluster located at ~500 kb from the right end of the linear chromosome contains nine genes that code for putative type I modular PKSs. Sixteen further genes, encoding proteins predicted to be involved in PKS substrate supply, post- and on-PKS tailoring reactions, deoxysugar biosynthesis, regulation, and resistance flank, and are interspersed with the PKS genes (Fig. 1*A* and *SI Appendix*, Table S1). The limits of the cluster were proposed as the *samR0465* and *samR0487* genes on the basis of sequence comparisons (*SI Appendix*, Table S1). The proposed functions of the flanking genes *samR0464* (predicted to encode an endoribonuclease) and *samR0488* (predicted to encode a "cold-shock" DNA binding protein) are unrelated to the biosynthesis of secondary metabolites. Thus, this putative gene cluster contains 25 genes that span 150 kb, making it one of the largest polyketide biosynthetic gene clusters identified to date. Sequence analyses of the PKSs encoded by the gene cluster using SEARCHPKS (13) revealed that they contain 112 enzymatic domains organized into 25 modules (Fig. 1*B*). All of the enzymatic domains, except the ketoreductase (KR) domain in the last module, are predicted to harbor a functional active site. The presence of a ketosynthase^{Q} (KS^{Q}) domain at the N terminus of SamR0467 and a thioesterase (TE) domain at the C terminus of SamR0474 indicates that SamR0467 and SamR0474 contain the modules that initiate and terminate polyketide chain assembly, respectively (14, 15). Following the precedent of other type I modular PKS systems (16), the order in which the remaining seven PKSs act in chain assembly was assumed to follow the order of the PKS genes in the cluster (Fig. 1*B*). Sequence analyses of the acyltransferase (AT) domain within each module suggested that the product of the PKS is assembled from 16 molecules of malonyl-CoA, eight molecules of methylmalonyl-CoA (one of which is used to create the propionyl starter unit), and one molecule of an unknown extender unit, loaded by the AT13 domain (13, 17) (Fig. 1*C* and *SI Appendix*, Fig. S1). These analyses, combined with the putative β-carbon–processing domains in each module (Fig. 1*B*), predicted the planar structure of the fully processed acyl thioester intermediate attached to the acyl carrier protein (ACP) domain in the last module of the PKS, with a single ambiguity—the nature of the C-26

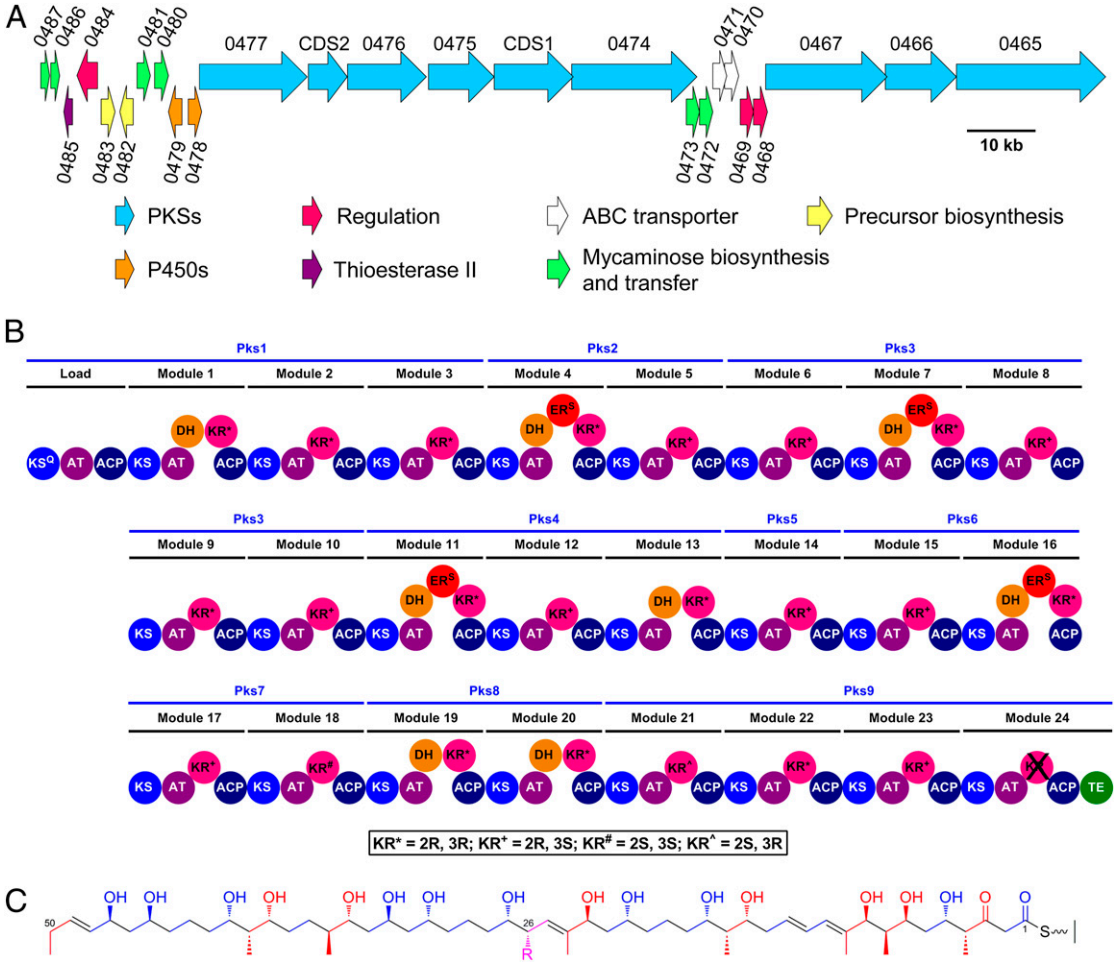

Fig. 1. Stambomycin biosynthetic gene cluster. (*A*) Organization of the stambomycin biosynthetic gene cluster. Limits of the cluster have been assigned using sequence comparisons. (*B*) Module and domain organization of the PKS encoded by the cluster, assigned using the SEARCHPKS program (13). The ketoreductase domain in module 24 is predicted to be nonfunctional because the tyrosine and asparagine residues of the catalytic triad are absent (42). ACP, acyl carrier protein; AT, acyltransferase; DH, dehydratase; ER^{R}, enoyl reductase predicted to generate a *2R* acyl thioester intermediate; ER^{S}, enoyl reductase predicted to generate a *2S* acyl thioester intermediate; KR*, ketoreductase predicted to generate a *2R*, *3R*-acyl thioester intermediate; KR^{+}, ketoreductase predicted to generate a *2R*, *3S*-acyl thioester intermediate; KR$^{\#}$, ketoreductase predicted to generate a *2S*, *3S*-acyl thioester intermediate; KR$^{\wedge}$, ketoreductase predicted to generate a *2S*, *3R*-acyl thioester intermediate; KS, β-ketoacyl synthase; KS^{Q}, β-ketoacyl synthase in which the active site cysteine residue is replaced by glutamine; TE, thioesterase. (*C*) Predicted structure of the fully assembled polyketide chain attached to the ACP domain in the last module of the PKS. The structure of the side chain at C-26 could not be predicted because the nature of the substrate of the AT domain in module 12 could not be inferred from sequence comparisons. Atoms in the polyketide chain are colored according to the precursors from which they derive: blue, malonyl-CoA; red, methylmalonyl-CoA; purple, malonyl-CoA derivative of unknown structure.

substituent loaded by the AT13 domain (Fig. 1*C*). The stereochemistry of each stereocenter in this intermediate was predicted by sequence analyses of the KR domains, which control the stereochemistry of the α and β carbons, and the enoylreductase (ER) domains, which control only the α carbon stereochemistry in the intermediates attached to each module of the PKS (Fig. 1 *B* and *C* and *SI Appendix*, Figs. S2 and S3) (18, 19). Substructure and similarity searches of the Chemical Abstracts database using the predicted structure of the final PKS-bound intermediate indicated that the product of the cryptic gene cluster was likely to be novel.

Five genes in the cluster were found to be homologous to genes within the spiramycin biosynthetic gene cluster that are responsible for the biosynthesis of mycaminose, one of three deoxysugars within the spiramycin structure (7) (*SI Appendix*, Fig. S4). Thus, these five genes are predicted to encode enzymes catalyzing the conversion of α-D-glucose-1-phosphate to nucleotide diphosphate (NDP)-D-mycaminose (*SI Appendix*, Table S1 and Fig. S4). In addition, analysis with SEARCHGTr (20) indicated that the *samR0481* gene encodes a putative glycosyl transferase (GT) capable of transferring a mycaminosyl residue from NDP to one or more of the hydroxyl groups in the product of the PKS. Taken together, these data strongly suggested that the final product of the cluster is a glycosylated lactone. It has been observed that GTs often require an auxiliary protein to glycosylate their acceptors, especially when the donor substrate is an amino sugar like mycaminose (21). No gene encoding such a protein has been identified within the cryptic biosynthetic gene cluster. Nguyen et al. (22) recently characterized an auxiliary protein, encoded by the *srm6* gene within the spiramycin biosynthetic gene cluster, which interacts with the mycaminosyltransferase Srm5, responsible for catalyzing transfer of the mycaminose residue from NDP to the spiramycin aglycone. They also showed that another auxiliary protein, Srm28, is flexible, and that in addition to the forosaminyltransferase Srm29, Srm28 can partner Srm5. Thus, it is possible that SamR0481 requires either Srm6 or Srm28, or both, for activity.

Three other genes are postulated to be involved in the biosynthesis of the predicted unique glycosylated polyketide. The *samR0478* and *samR0479* genes encode putative cytochromes P450, which usually catalyze post-PKS hydroxylation reactions (23). The predicted product of *samR0485* is a type II thioesterase. Such enzymes typically catalyze hydrolysis of aberrant acyl groups attached to the ACP domains of the PKS that block chain elongation (24).

Activation of the Silent Cryptic Gene Cluster. Poor or no expression of the gene cluster could explain why its metabolic product has never been identified. To test this hypothesis, we examined the expression of selected PKS genes within the cluster under two growth conditions known to elicit production of the macrolide spiramycin (25). RT-PCR analyses indicated that the PKS genes were expressed very poorly or not at all throughout growth (Fig. 2), implying that the product of the gene cluster either is not produced or is produced in quantities likely to be below the detection limit of most analytical tools.

The in silico analysis of the gene cluster identified three genes coding for potential pathway-specific regulators. The *samR0468* and *samR0469* genes encode a putative two-component system, and *samR0484* encodes a putative large ATP-binding regulator of the LuxR family (LAL) protein (*SI Appendix*, Table S1). LAL regulators contain an ATP-binding domain at the N terminus (Walker A and B motifs) and a helix-turn-helix domain at the C terminus, most likely responsible for DNA binding (26). Members of the LAL family have been previously described as activators of polyketide biosynthesis, such as PikD and RapH, which activate pikromycin and rapamycin production, respectively (27, 28). Thus, we decided to examine whether constitutive overexpression of the *samR0484* gene triggers expression of the PKS genes within the cryptic cluster. The *samR0484* gene was cloned into the conjugative and integrative vector pIB139 (29), placing it under the control of the strong constitutive *ermE** promoter. The resulting plasmid was integrated into the *att*B site of the chromosome of *S. ambofaciens* to create the strain ATCC/OE484, and the empty pIB139 vector was separately integrated into the chromosome of *S. ambofaciens* as a control to create the strain ATCC/pIB139. Comparative transcriptional analyses by RT-PCR showed that the PKS genes within the cryptic cluster were expressed in the ATCC/OE484 strain throughout growth, thus suggesting that the LAL regulator is a pathway-specific activator of the gene cluster (Fig. 2). In contrast, little or no expression of the PKS genes was detected in the control strain. Constitutive expression of the genes encoding the two-component system resulted in no induction of expression of the PKS genes relative to the control (*SI Appendix*, *SI Materials and Methods* and Fig. S5).

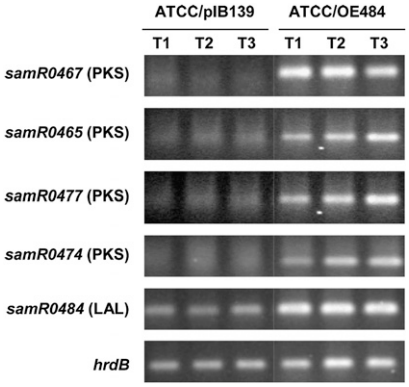

Fig. 2. Induction of PKS gene expression. Transcriptional analysis by RT-PCR of the control ATCC/pIB139 and the mutant strain ATCC/OE484, grown in MP5 medium. Expression of four biosynthetic genes (*samR0467*, *samR0465*, *samR0477*, *samR0474*), together with expression of the regulatory gene *samR0484*, were analyzed by RT-PCR using 4 μg total RNA. The constitutively expressed *hrdB* gene, coding for the major sigma factor, was used as positive control. Experiments carried out on three separate occasions gave the same results. T1, exponential phase; T2, transition phase; T3, stationary phase.

Isolation and Structure Elucidation of the Metabolic Products of the Cryptic PKS Gene Cluster. To identify the product(s) of the gene cluster, the supernatant and the mycelium extracts of the strains ATCC/pIB139 and ATCC/OE484 were analyzed by comparative metabolic profiling using liquid chromatography-electrospray ionization-time of flight-mass spectrometry (LC-ESI-TOF-MS). Thus, two major species with molecular formulas $C_{73}H_{133}NO_{22}$ (calculated for $C_{73}H_{134}NO_{22}^{+}$: 1376.9392; found: 1376.9391) and $C_{72}H_{131}NO_{22}$ (calculated for $C_{72}H_{132}NO_{22}^{+}$: 1362.9236; found: 1362.9232) were identified in the methanolic mycelial extracts of the ATCC/OE484 strain that were absent in extracts of the ATCC/pIB139 strain (Fig. 3 and *SI Appendix*, Figs. S6 and S7). To confirm that these compounds were the products of the cryptic gene cluster, we replaced the *samR0467* gene (coding for the first PKS in the assembly line) in the ATCC/OE484 strain with a kanamycin resistance cassette. LC-ESI-TOF-MS analysis of mycelial extracts of this strain confirmed that production of the compounds was abrogated (Fig. 3). Similarly, inactivation of the *samR0481* gene, encoding the glycosyl transferase, also abolished production of the compounds. LC-ESI-TOF-MS analysis indicated that the corresponding aglycones lacking the predicted β-D-mycaminose residue were produced instead (calculated for $C_{65}H_{119}O_{19}^{+}$: 1203.8340, found: 1203.8344; calculated for $C_{64}H_{117}O_{19}^{+}$: 1189.8184, found: 1189.8183) (*SI Appendix*, Figs. S8 and S9).

A search of the Chemical Abstracts database did not retrieve any compounds with molecular formulas corresponding to the

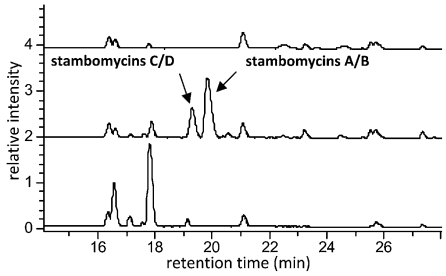

Fig. 3. Detection of the stambomycins. Base peak chromatograms from LC-ESI-TOF-MS analyses of methanolic mycelial extracts of ATCC/pIB139 (*Top*), ATCC/OE484 (*Middle*), and ATCC/OE484/Δ467 (*Bottom*) strains. Peaks corresponding to stambomycins A/B and C/D in the ATCC/OE484 strain are labeled.

identified metabolic products of the cryptic gene cluster. Thus, we purified these unique compounds from mycelial extracts of the ATCC/OE484 strain using semipreparative HPLC. ^{1}H NMR analysis of the higher-molecular-weight species indicated that it consists of two almost identical compounds, which we named stambomycins A and B, in a ratio of ~5:1. COSY, TOCSY, HMBC, HSQC, NOESY, and PENDANT NMR experiments showed that stambomycins A and B consist of a common 51-membered glycosylated macrolide core containing a tetrahydropyran resulting from the addition of the C-7 hydroxyl group to the C-3 keto group (Fig. 4 and *SI Appendix*, Figs. S10–S22 and Table S2). The NMR experiments confirmed the close relationship between the planar structure of the macrolide aglycone and the predicted structure of the final intermediate attached to the ACP domain in the last module of the PKS (Figs. 1*C* and 4). $^{3}J_{H\text{-}H}$ coupling constants of ~15 Hz for H-12/H-13 and H-48/H-49 and NOESY correlations between the C-10 methyl group and C-12, and between the C-24 methyl group and C-26, indicate that the C-12/C-13, C-48/C-49, C-10/C-11, and C-24/C-25 double bonds all have *E* configuration (Fig. 4). COSY and HMBC correlations combined with comparisons of ^{1}H NMR data with the literature (30, 31) established that β-mycaminose is attached to the aglycone (Fig. 4 and *SI Appendix*, Fig. S14 and Table S2), as predicted by sequence analysis of the deoxysugar biosynthesis genes within the cluster. NMR spectroscopic analysis of the aglycones isolated from the *samR0481* mutant of the ATCC/OE484 strain confirmed that they lacked the β-mycaminose residue at C-5 (*SI Appendix*, *SI Materials and Methods*, Figs. S31–S40, and Tables S4 and S5).

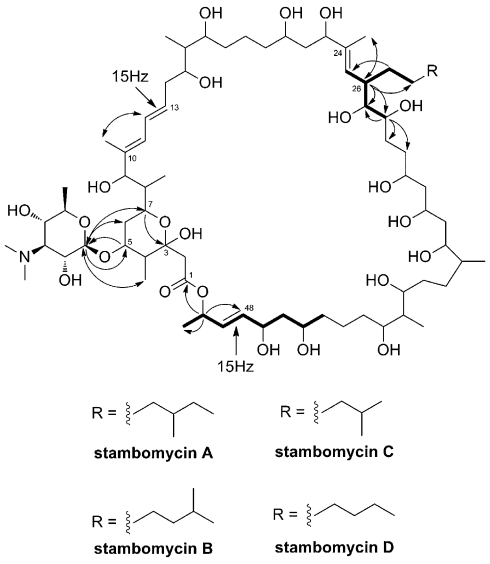

Fig. 4. Structure elucidation of the stambomycins. Structures of stambomycins A–D proposed on the basis of HRMS and NMR spectroscopic analyses. Key COSY, HMBC, and NOESY correlations for structural elements that were not predicted by sequence analysis of the biosynthetic enzymes are indicated by bold lines, single-headed arrows, and double-headed arrows, respectively. ^{3}J coupling constants observed in ^{1}H NMR spectra for H-12/H-13 and H-49/H-50 are indicated.

Several structural features of the stambomycins were not predictable from sequence analyses of the biosynthetic enzymes, including the nature of the C-26 side chain and the sites of glycosylation, hydroxylation, and macrocycle formation. A 4-methyl-n-hexyl chain is proposed to be attached to C-26 in stambomycin A on the basis of COSY and HMBC correlations (Fig. 4 and *SI Appendix*, Figs. S10, S12, S13, S17, and S20). Similar analyses suggested that stambomycin B contains a 5-methyl-n-hexyl side chain at C-26. Several HMBC and NOESY correlations indicate that the deoxysugar residue is attached to the C-5 hydroxyl group of the aglycone (Fig. 4). COSY and HMBC correlations suggest that non–PKS-derived hydroxyl groups are attached to C-28 and C-50, and that the macrolactone results from formation of an ester bond between C-1 and the C-50 hydroxyl group (Fig. 4).

^{1}H, COSY, HMBC, HSQC, NOESY, and PENDANT NMR experiments indicated that the lower-molecular-weight species consists of two compounds, which we named stambomycins C and D, in a ratio of ~3:1, that differ from stambomycins A and B only in the structure of the alkyl chain attached to C-26 (*SI Appendix*, Table S3 and Figs. S23–S30). Stambomycin C is proposed to have a 4-methyl-n-pentyl group at this position, whereas stambomycin D is proposed to bear an n-hexyl group (Fig. 4 and *SI Appendix*, Figs. S10, S12, S13, S26, and S29).

The experimentally elucidated structures of the stambomycins suggest the assembly of these remarkable macrolides involves unprecedented biosynthetic chemistry. The various hexyl/heptyl substituents at C-26 indicate that the AT domain of PKS module 12 selectively recognizes and loads hexyl/heptylmalonyl-CoA extender units onto the ACP domain of this module (*SI Appendix*, Fig. S41). Malonyl-CoA, methylmalonyl-CoA, ethylmalonyl-CoA, (2-propenyl)malonyl-CoA, and (2-chloroethyl)malonyl-CoA are all known or hypothesized to be used as extender units by type I modular PKSs (32). Crotonyl-CoA reductase/carboxylase-like enzymes catalyze reduction and carboxylation of α, β-unsaturated thioester precursors in the biosynthesis of ethylmalonyl-CoA, (2-propenyl)malonyl-CoA and (2-chloroethyl)malonyl-CoA (32–34). No enzymes with sequence similarity to crotonyl-CoA reductase/carboxylases are encoded by genes within the stambomycin biosynthetic gene cluster, suggesting that a different mechanism is used to biosynthesize the hexyl/heptylmalonyl-CoA extender units in *S. ambofaciens*.

The C-1/O-50 ester linkage in the stambomycins suggests that hydroxylation of C-50, probably catalyzed by one of the two cytochrome P450s encoded within the cluster, occurs during polyketide chain assembly on one or more ACP-bound intermediates attached to the PKS (*SI Appendix*, Fig. S41). The C-50 hydroxyl group is not present in the (*2S*)-methylmalonyl-CoA starter unit predicted to be recognized and decarboxylated by the loading module of the PKS, and is required to offload the fully assembled polyketide chain from the active site Ser residue of the TE domain by macrocyclization. Further experiments will be required to test this hypothesis.

Bioactivity of the Stambomycins. The broad range of biological activities exhibited by macrolides, including antibacterial (35), antifungal (36), and antiproliferative (37) activities, prompted us to investigate the biological activity of the stambomycins. They showed moderate activity against Gram-positive bacteria (*SI Appendix*, Table S6) but no activity against Gram-negative bac-

MICROBIOLOGY

teria, or fungi. More interestingly, they inhibited proliferation of human adenocarcinoma (HT29) cell lines with a similar IC_{50} (1.77 ± 0.04 μM for stambomycins A/B and 1.74 ± 0.04 μM for stambomycins C/D) to the clinical anticancer agent doxorubicin (1.32 ± 0.08 μM), but were fourfold less toxic toward adult Chinese hamster ovary sane cells (CHO-K1; 8.47 ± 0.67 μM for stambomycins A/B, 8.46 ± 0.52 μM for stambomycins C/D) than doxorubicin (1.99 ± 0.25 μM). Significant antiproliferative activities were also observed against the human breast (MCF7), lung (H460), and prostate (PC3) cancer cell lines (*SI Appendix*, Table S6). These data indicate that stambomycins may represent promising new leads for anticancer drug discovery.

Conclusions

In summary, we have used constitutive overexpression of a putative pathway-specific LAL regulator to induce expression of a silent type I modular PKS gene cluster in *S. ambofaciens*. This enabled us to identify a unique structural class of polyketides with promising antitumor activity that represents the largest macrolides ever to be isolated from an actinomycete. In addition, structure elucidation of these molecules disclosed unique and interesting chemical features; in particular, the incorporation of hexyl- and heptylmalonyl-CoA extender units and a mechanism of macrolactonization involving in *trans* hydroxylation of PKS-bound intermediate(s) to release the fully assembled polyketide chain from the PKS. Manipulation of a putative pathway-specific regulatory gene has been used to uncover a unique natural product by activating expression of a silent hybrid PKS-NRPS gene cluster in a filamentous fungus (38). A similar tactic has hitherto not been used for the identification of novel metabolic products of cryptic gene clusters in actinomycetes, which are one of the main producers of microbial antibiotics and other bioactive natural products. Database searches indicate that LAL regulators are present in numerous gene clusters that direct the production of known secondary metabolites (*SI Appendix*, Table S7), as well as several cryptic biosynthetic gene clusters for which the metabolic product(s) remain to be discovered (*SI Appendix*, Table S8). The findings reported here suggest that our approach is likely to represent a powerful general strategy for realizing the drug-discovery potential offered by the numerous cryptic biosynthetic pathways identified in actinomycete genome sequences.

Materials and Methods

Bacterial strains, plasmids and BACs are described in *SI Appendix*, Table S9, and their manipulation is described in *SI Appendix*, *SI Materials and Methods*.

Bioinformatics Analysis. The programs SEARCHPKS (13) and SEARCHGTr (20) were used to analyze the PKSs and the glycosyltransferase encoded by genes within the cluster. Functional analysis of each gene product and sequence comparisons were carried out using BlastP and Blast bl2seq (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Overexpression of the LAL Regulator. The SamR0484-encoding sequence was amplified by PCR from *S. ambofaciens* ATCC23877 genomic DNA with the primers OE484-F and OE484-R (sequences in *SI Appendix*, Table S10). The PCR conditions, using Takara polymerase (Fermentas), were as follows: 2 min at 94 °C, 30 cycles of 10 s at 98 °C, and 15 min at 68 °C, followed by a final extension of 10 min at 72 °C. The PCR product was cloned into pGEMT-easy (Promega) and the integrity of the insert was confirmed by sequencing. After restriction digestion with *Nde*I and *Xba*I, the insert was cloned into the pIB139 vector under the control of the *ermEp** promoter modified to have a typical *Streptomyces* ribosome binding site (39). The resulting recombinant plasmid (pOE-0484) was introduced in *S. ambofaciens* ATCC23877 by conjugation from *Escherichia coli* and integrated into the chromosomal $att B_{\phi C31}$ site by site-specific recombination. The mutant strain ATCC/OE484 was checked by PCR, Southern blotting, and pulsed-field gel electrophoresis.

Transcriptional Analysis. Total RNA of *Streptomyces* WT and mutant strains was isolated from MP5 liquid cultures, and cDNAs were prepared as described elsewhere (39). The primers used for the analysis are listed in *SI Appendix*, Table S10. The PCR conditions for the transcriptional analysis were as follows: 4 min at 94 °C; 28 cycles of 30 s at 94 °C, 30 s at 60 °C, and 30 s at 72 °C; and a final extension of 5 min at 72 °C. To check possible contamination of genomic DNA, the same PCR program (35 cycles instead of 28) was applied to RNA samples, after DNase treatment, using as a control primers designed to amplify the *hrdB* gene, which encodes the major sigma factor and is considered to be constitutively expressed.

Construction of *S. ambofaciens* Mutant Strains. The REDIRECT system (40) was used to make an in-frame deletion of *samR0467* in *S. ambofaciens* ATCC/OE484 as described previously (41). The *neo* + *oriT* mutagenesis cassette, derived from the plasmid pIJ776 (40), was used as template in the PCR with the primers D467-F and D467-R. *E. coli* BW25113/pKD20 was transformed, first with the BAC of interest (BBB), and then with the PCR product (~1400 bp) to replace the biosynthetic gene *samR0467* by homologous recombination. The chloramphenicol resistance gene of the vector pBeloBAC11 was replaced by a spectinomycin resistance gene, using the same strategy. *E. coli* ET12567/pUZ8002 was transformed with the mutated BAC (BBBspec/Δ467::*neo*+*oriT*) for conjugation with the ATCC/OE484 strain. The kanamycin-resistant and spectinomycin-sensitive clones were picked and analyzed by PCR, Southern blotting, and pulsed-field gel electrophoresis.

Identification, Isolation, and Structural Elucidation of Stambomycins A/B and C/D. The PKS products were identified by LC-MS comparative metabolic profiling of *S. ambofaciens* ATCC23877, ATCC/OE484, and ATCC/OE484/Δ467 strains grown in liquid MP5 medium (details in *SI Appendix*). To purify stambomycins, solid MP5 medium was used. A 50-mL quantity of MP5 molten solid medium was poured into a 10 × 10-cm square Petri dish. A sterile cellophane membrane was placed on the solidified plate and 20 μL *S. ambofaciens* ATCC/OE484 spores from the stock were spread on top of it. The plate was incubated for 4 d at 30 °C, and the cellophane membrane was lifted off the plate. The mycelia were scraped off the cellophane membrane and combined with mycelia from three other cultures grown in the same way. The combined mycelia were transferred to a 50-mL plastic tube, and 40 mL methanol was added. The mixture was sonicated for 10 min, and the mycelia were pelleted by centrifugation for 10 min at 4 °C. The extract was decanted, and the mycelia were extracted twice more with 40 mL methanol. The combined methanolic fractions were concentrated under reduced pressure at 20–25 °C. Stambomycins were partially purified from the concentrated methanol extract using semipreparative HPLC on a reverse-phase column (C18, 100 × 21 mm, fitted with C18 precolumn 10 × 21 mm). The mobile phases used were water (A) and acetonitrile (B), both with 0.1% formic acid added. The gradient used was as follows: 0 min, 80% A/20% B; 10 min, 80% A/20% B; 20 min, 100% B; 35 min 100% B. Absorbance was monitored at a wavelength of 240 nm. Fractions containing stambomycins were identified using ESI-MS and combined. The combined fractions were evaporated under reduced pressure and resuspended in a small volume of 50% aqueous methanol. Stambomycins were purified from the combined fractions by semipreparative HPLC on the same column using the following gradient: 0 min, 60% A/40% B; 15 min, 5% A/95% B; 20 min, 100% B; 25 min, 100% B. Fractions containing a mixture of stambomycins A and B were collected, combined, and lyophilised, as were separate fractions containing a mixture of stambomycins C and D. An 7.8-mg quantity of stambomycins A/B and 4.3 mg stambomycins C/D were obtained from 10 plates. NMR spectra were recorded in d_4-MeOH and, in the case of HMBC spectra, for stambomycins A/B also d_6-DMSO, on Bruker Avance 700 MHz or 500 MHz spectrometers equipped with cryoprobes.

ACKNOWLEDGMENTS. We thank Peter Leadlay (University of Cambridge, Cambridge, United Kingdom) for the kind gift of pIB139. We likewise thank Dr. Peter Grice (University of Cambridge, Cambridge, United Kingdom) with assistance in the recording of NMR spectra. The Bruker MaXis mass spectromter used in this research was obtained, through Birmingham Science City: Innovative Uses for Advanced Materials in the Modern World (West Midlands Centre for Advanced Materials Project 2), with support from Advantage West Midlands and was partly funded by the European Regional Development Fund. This research was supported by the European Commission under the 6th framework program (integrated project actinogen; FP6-5224).

1. Bérdy J (2005) Bioactive microbial metabolites. *J Antibiot (Tokyo)* 58:1–26.

2. Newman DJ, Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. *J Nat Prod* 70:461–477.

3. Challis GL (2008) Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology* 154:1555–1569.
4. Zerikly M, Challis GL (2009) Strategies for the discovery of new natural products by genome mining. *ChemBioChem* 10:625–633.
5. Pinnert-Sindico S (1954) Une nouvelle espèce de *Streptomyces* productrice d'antibiotiques: *Streptomyces ambofaciens* n. sp. caractères culturaux. *Ann Inst Pasteur (Paris)* 87:702–707.
6. Juguet M, et al. (2009) An iterative nonribosomal peptide synthetase assembles the pyrrole-amide antibiotic congocidine in *Streptomyces ambofaciens*. *Chem Biol* 16: 421–431.
7. Karray F, et al. (2007) Organization of the biosynthetic gene cluster for the macrolide antibiotic spiramycin in *Streptomyces ambofaciens*. *Microbiology* 153:4111–4122.
8. Choulet F, et al. (2006) Evolution of the terminal regions of the *Streptomyces* linear chromosome. *Mol Biol Evol* 23:2361–2369.
9. Barona-Gómez F, et al. (2006) Multiple biosynthetic and uptake systems mediate siderophore-dependent iron acquisition in *Streptomyces coelicolor* A3(2) and *Streptomyces ambofaciens* ATCC 23877. *Microbiology* 152:3355–3366.
10. Pang X, et al. (2004) Functional angucycline-like antibiotic gene cluster in the terminal inverted repeats of the *Streptomyces ambofaciens* linear chromosome. *Antimicrob Agents Chemother* 48:575–588.
11. Bunet R, et al. (2011) Characterization and manipulation of the pathway-specific late regulator AlpW reveals *Streptomyces ambofaciens* as a new producer of kinamycins. *J Bacteriol* 193:1142–1153.
12. Hertweck C (2009) The biosynthetic logic of polyketide diversity. *Angew Chem Int Ed Engl* 48:4688–4716.
13. Yadav G, Gokhale RS, Mohanty D (2003) SEARCHPKS: A program for detection and analysis of polyketide synthase domains. *Nucleic Acids Res* 31:3654–3658.
14. Bisang C, et al. (1999) A chain initiation factor common to both modular and aromatic polyketide synthases. *Nature* 401:502–505.
15. Gokhale RS, Hunziker D, Cane DE, Khosla C (1999) Mechanism and specificity of the terminal thioesterase domain from the erythromycin polyketide synthase. *Chem Biol* 6:117–125.
16. Callahan B, Thattai M, Shraiman BI (2009) Emergent gene order in a model of modular polyketide synthases. *Proc Natl Acad Sci USA* 106:19410–19415.
17. Yadav G, Gokhale RS, Mohanty D (2003) Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *J Mol Biol* 328:335–363.
18. Keatinge-Clay AT (2007) A tylosin ketoreductase reveals how chirality is determined in polyketides. *Chem Biol* 14:898–908.
19. Kwan DH, et al. (2008) Prediction and manipulation of the stereochemistry of enoylreduction in modular polyketide synthases. *Chem Biol* 15:1231–1240.
20. Kamra P, Gokhale RS, Mohanty D (2005) SEARCHGTr: A program for analysis of glycosyltransferases involved in glycosylation of secondary metabolites. *Nucleic Acids Res* 33(Web Server issue):W220–225.
21. Hong JS, et al. (2007) Functional analysis of desVIII homologues involved in glycosylation of macrolide antibiotics by interspecies complementation. *Gene* 386: 123–130.
22. Nguyen HC, et al. (2010) Glycosylation steps during spiramycin biosynthesis in *Streptomyces ambofaciens*: Involvement of three glycosyltransferases and their interplay with two auxiliary proteins. *Antimicrob Agents Chemother* 54:2830–2839.
23. Parajuli N, Basnet DB, Chan Lee H, Sohng JK, Liou K (2004) Genome analyses of *Streptomyces peucetius* ATCC 27952 for the identification and comparison of cytochrome P450 complement with other *Streptomyces*. *Arch Biochem Biophys* 425: 233–241.
24. Kim BS, Cropp TA, Beck BJ, Sherman DH, Reynolds KA (2002) Biochemical evidence for an editing role of thioesterase II in the biosynthesis of the polyketide pikromycin. *J Biol Chem* 277:48028–48034.
25. Pernodet JL, Alegre MT, Blondelet-Rouault MH, Guérineau M (1993) Resistance to spiramycin in *Streptomyces ambofaciens*, the producer organism, involves at least two different mechanisms. *J Gen Microbiol* 139:1003–1011.
26. De Schrijver A, De Mot R (1999) A subfamily of MalT-related ATP-dependent regulators in the LuxR family. *Microbiology* 145:1287–1288.
27. Wilson DJ, Xue Y, Reynolds KA, Sherman DH (2001) Characterization and analysis of the PikD regulatory factor in the pikromycin biosynthetic pathway of *Streptomyces venezuelae*. *J Bacteriol* 183:3468–3475.
28. Kuscer E, et al. (2007) Roles of *rapH* and *rapG* in positive regulation of rapamycin biosynthesis in *Streptomyces hygroscopicus*. *J Bacteriol* 189:4756–4763.
29. Wilkinson CJ, et al. (2002) Increasing the efficiency of heterologous promoters in actinomycetes. *J Mol Microbiol Biotechnol* 4:417–426.
30. Ramu K, Shringarpure S, Cooperwood S, Beale JM, Williams JS (1994) 1H-NMR and 13C-NMR spectral assignments of spiramycins I and III. *Pharm Res* 11:458–465.
31. Melançon CE, 3rd, Liu HW (2007) Engineered biosynthesis of macrolide derivatives bearing the non-natural deoxysugars 4-epi-D-mycaminose and 3-n-monomethylamino-3-deoxy-D-fucose. *J Am Chem Soc* 129:4896–4897.
32. Chan YA, Podevels AM, Kevany BM, Thomas MG (2009) Biosynthesis of polyketide synthase extender units. *Nat Prod Rep* 26:90–114.
33. Eustáquio AS, et al. (2009) Biosynthesis of the salinosporamide A polyketide synthase substrate chloroethylmalonyl-coenzyme A from S-adenosyl-L-methionine. *Proc Natl Acad Sci USA* 106:12295–12300.
34. Mo S, et al. (2011) Biosynthesis of the allylmalonyl-CoA extender unit for the FK506 polyketide synthase proceeds through a dedicated polyketide synthase and facilitates the mutasynthesis of analogues. *J Am Chem Soc* 133:976–985.
35. Brisson-Noel A, Trieu-Cuot P, Courvalin P (1988) Mechanism of action of spiramycin and other macrolides. *J Antimicrob Chemother* 22(Suppl B):13–23.
36. Fjaervik E, Zotchev SB (2005) Biosynthesis of the polyene macrolide antibiotic nystatin in *Streptomyces noursei*. *Appl Microbiol Biotechnol* 67:436–443.
37. Olano C, Méndez C, Salas JA (2009) Antitumor compounds from marine actinomycetes. *Mar Drugs* 7:210–248.
38. Bergmann S, et al. (2007) Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat Chem Biol* 3:213–217.
39. Bunet R, et al. (2008) Regulation of the synthesis of the angucyclinone antibiotic alpomycin in *Streptomyces ambofaciens* by the autoregulator receptor AlpZ and its specific ligand. *J Bacteriol* 190:3293–3305.
40. Gust B, Challis GL, Fowler K, Kieser T, Chater KF (2003) PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci USA* 100:1541–1546.
41. Aigle B, Pang X, Decaris B, Leblond P (2005) Involvement of AlpV, a new member of the *Streptomyces* antibiotic regulatory protein family, in regulation of the duplicated type II polyketide synthase *alp* gene cluster in *Streptomyces ambofaciens*. *J Bacteriol* 187:2491–2500.
42. Reid R, et al. (2003) A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. *Biochemistry* 42:72–79.