ChemPubSoc Europe

DOI: 10.1002/cbic.201300147

Bioactivity-Guided Genome Mining Reveals the Lomaiviticin Biosynthetic Gene Cluster in *Salinispora tropica*

Roland D. Kersten,[a] Amy L. Lane,[b] Markus Nett,[c] Taylor K. S. Richter,[a] Brendan M. Duggan,[d] Pieter C. Dorrestein,[a, d, e] and Bradley S. Moore*[a, d]

The use of genome sequences has become routine in guiding the discovery and identification of microbial natural products and their biosynthetic pathways. In silico prediction of molecular features, such as metabolic building blocks, physico-chemical properties or biological functions, from orphan gene clusters has opened up the characterization of many new chemo- and genotypes in genome mining approaches. Here, we guided our genome mining of two predicted enediyne pathways in *Salinispora tropica* CNB-440 by a DNA interference bioassay to isolate DNA-targeting enediyne polyketides. An organic extract of *S. tropica* showed DNA-interference activity that surprisingly was not abolished in genetic mutants of the targeted enediyne pathways, *ST_pks1* and *spo*. Instead we showed that the product of the orphan type II polyketide synthase pathway, *ST_pks2*, is solely responsible for the DNA-interfering activity of the parent strain. Subsequent comparative metabolic profiling revealed the lomaiviticins, glycosylated diazofluorene polyketides, as the *ST_pks2* products. This study marks the first report of the 59 open reading frame lomaiviticin gene cluster (*lom*) and supports the biochemical logic of their dimeric construction through a pathway related to the kinamycin monomer.

Introduction

Microbial genome sequencing has increased dramatically over the past decade, providing access to a tremendous amount of information about the genetic capability of microorganisms to produce natural products.[1] Bioinformatics analyses have revealed that the number of putative natural product biosynthetic gene clusters often greatly exceeds the number of reported metabolites,[2] suggesting that a large portion of nature's chemical potential remains unexplored. Realization of the chemistry encoded by orphan gene clusters may afford access to the next generations of drugs as well as insights into evolutionary benefits of the plethora of biosynthetic genes maintained within microbial genomes.

Researchers have developed a variety of approaches to discover natural products associated with orphan gene clusters.[3,4] Such genome mining strategies include inactivation of selected gene clusters followed by comparative metabolic profiling of mutant and wild-type strains,[5] prediction of physico-chemical properties from biosynthetic gene sequences,[6] manipulation of genes regulating biosynthetic pathway expression,[7] isotope labeling of predicted biosynthetic precursors in combination with isotope-guided fractionation,[8] heterologous expression of orphan biosynthetic genes,[9] and mass spectrometry-guided genome mining.[10]

The discovery of novel natural products that impede DNA replication or function is of interest because DNA-interfering molecules have potential cancer chemotherapeutic properties,[11,12] especially as drug conjugates linked to monoclonal antibodies.[13] Hallmark examples of DNA-targeting natural products include the enediyne antibiotics.[14] Their biosynthesis involves an iteratively acting type I polyketide synthase, which assembles the distinctive core structure common to all members of this natural product family.[15] Enediynes feature two acetylenic groups that are linked through a conjugated double bond. They are differentiated into two subclasses possessing either nine-membered rings, as exemplified by neocarzinostatin, or ten-membered rings, such as calicheamicin and dynemicin (Scheme 1 A).[15] These cytotoxic agents intercalate into chromosomal DNA in a sequence-specific manner and cause strand scission through a radical mechanism.[16]

[a] R. D. Kersten, T. K. S. Richter, Prof. Dr. P. C. Dorrestein, Prof. Dr. B. S. Moore
Center for Marine Biotechnology and Biomedicine
Scripps Institution of Oceanography, University of California at San Diego
9500 Gilman Drive, La Jolla, CA 92093-0204 (USA)
E-mail: bsmoore@ucsd.edu

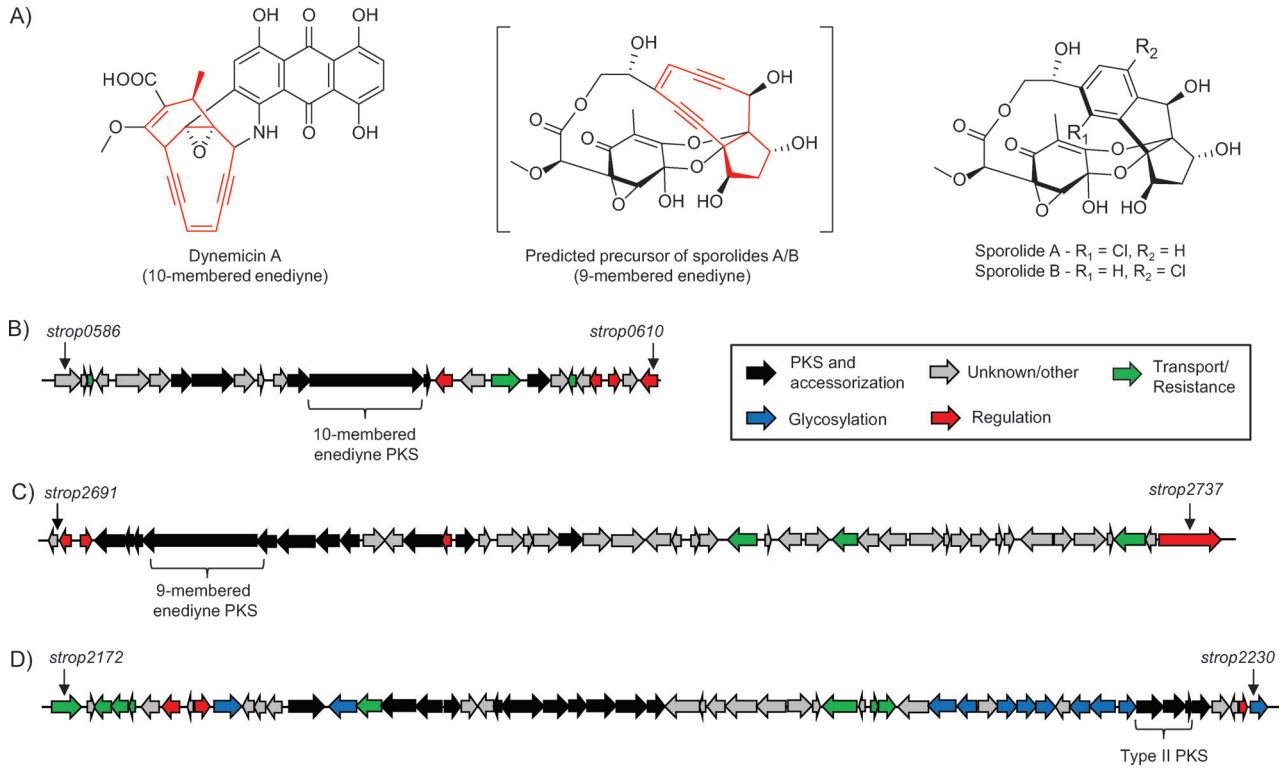
[b] Prof. Dr. A. L. Lane
University of North Florida, Chemistry Department
1 UNF Dr., Jacksonville, FL 32224 (USA)

[c] Dr. M. Nett
Leibniz Institute for Natural Product Research and Infection Biology
Hans-Knöll-Institute
Adolf-Reichwein-Strasse 23, 07745 Jena (Germany)

[d] Dr. B. M. Duggan, Prof. Dr. P. C. Dorrestein, Prof. Dr. B. S. Moore
Skaggs School of Pharmacy and Pharmaceutical Sciences
University of California at San Diego
9500 Gilman Drive, La Jolla, CA 92093 (USA)

[e] Prof. Dr. P. C. Dorrestein
Department of Pharmacology and
Department of Chemistry and Biochemistry
University of California at San Diego
9500 Gilman Drive, La Jolla, CA 92093 (USA)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201300147.



Scheme 1. Candidate DNA-targeting genotypes based on genome mining of *S. tropica* CNB-440. A) Representative ten-membered enediyne, dynemicin A, and predicted enediyne precursor of sporolides A/B. Enediyne scaffold indicated in red. B) *ST_pks1* gene cluster with predicted ten-membered enediyne polyketide biosynthetic machinery. C) *spo* (sporolide A/B) gene cluster with a predicted nine-membered enediyne biosynthetic machinery. D) *ST_pks2* gene cluster with a predicted glycosylated anthracycline product.

Actinomycetes are prevalent sources of DNA-interfering natural products, including enediynes.[12,14] Genome-sequenced *Salinispora* species devote approximately 10% of their genetic material to natural product biosynthesis, which includes several orphan enediyne polyketide synthase (ePKS) gene clusters.[17,18] Although the 5.2-Mbp genome of the *Salinispora tropica* strain CNB-440 harbors at least 19 putative biosynthetic gene clusters,[1] only five of these loci are linked with characterized natural products.[18–22] Its untapped metabolic potential thus makes *S. tropica* a promising candidate for genomics-guided drug discovery. The *S. tropica* CNB-440 genome contains two ePKS gene clusters, the *ST_pks1* and *spo* loci. While the product of the former is not known, the *spo* gene set is associated with the biosynthesis of two chlorinated cyclopenta[*a*]indenes. Sporolides A and B were proposed to be degradation products of a reactive nine-membered enediyne precursor molecule (Scheme 1A, Figure S1A in the Supporting Information).[22–24] While the potent proteasome inhibitor salinosporamide A was isolated from this strain and is currently in clinical trials for treatment of multiple myeloma,[25] no DNA-interfering anticancer natural products have been reported from *S. tropica*.

Herein, we report the identification of DNA-targeting natural products from *S. tropica* CNB-440 through bioactivity-guided genome mining. Allelic-exchange mutagenesis of select biosynthesis genes, DNA interference assays and comparative metabolite profiling tracked the observed biological activity not to one of the anticipated enediyne pathways but exclusively to *ST_pks2*, an orphan biosynthetic gene cluster predicted to yield an anthracycline-type antibiotic. The metabolites associated with the *ST_pks2* locus were subsequently identified as the lomaiviticins, glycosylated diazofluorene polyketides with potent anticancer activity originally discovered from *Salinispora pacifica* strain DPJ-0019 (formerly *Micromonospora lomaivitiensis*).[26–29] The discovery of the lomaiviticin pathway provides the molecular logic for the construction of these complex organic molecules and, in the long term, may help generate new bioactive derivatives by genetic engineering. Furthermore, the described genome mining approach may offer an effective strategy to mine for new DNA-interference natural products and their pathways with potential applicability as antitumor chemotherapies.

Results and Discussion

Bioinformatic prediction and biochemical detection of DNA-damaging chemotypes in *S. tropica* CNB-440

The first step in the search for DNA-interfering chemotypes from the *S. tropica* CNB-440 genome was the bioinformatic analysis of candidate gene clusters. Initially, the two ePKS pathways in the *S. tropica* CNB-440 genome, *spo* and *ST_pks1*, were targeted. The *ST_pks1* locus was uncharacterized and not related to known enediyne molecules from *S. tropica*.

Proteins encoded by the *ST_pks1* genes exhibit strongest homology to enzymes required for the biosynthesis of dynemicin, a ten-membered enediyne.[29] Interestingly, *ST_pks1* (Scheme 1B, Table S1) lacks open reading frames (ORFs) for the synthesis of anthraquinone- or sugar-based groups that are typically clustered with *epks* genes and account for the extensive functionalization of the enediyne core. In dynemicin biosynthesis, genes encoding tailoring enzymes are chromosomally distinct from enediyne-associated genes,[15] and an analogous scenario is plausible for the proposed *ST_pks1*-encoded enediyne as well.

The *spo* cluster (Scheme 1C) includes all genes required for enediyne biosynthesis and is expected to yield a nine-membered enediyne product (Scheme 1A). Most nine-membered enediynes are highly unstable relative to their ten-membered counterparts,[23] and the *spo* enediyne product was proposed to readily undergo nonenzymatic Bergman-type cyclization in the presence of chloride anions, yielding sporolides A and B (Figure S1A).[22,23,30] Even though the enediyne origin of the sporolides has not been experimentally confirmed, previous in vitro studies supported an involvement of *spo* genes in the biosynthesis of the cyclohexenone residue.[22]

To determine whether *S. tropica* produces DNA-damaging antibiotics, as predicted by bioinformatic analysis, the strain was fermented in a commonly employed seawater-based medium, and chemical extracts from multiple time points were evaluated for DNA-interfering natural products using the well-established biochemical prophage induction assay (BIA).[31] DNA-damaging activity was most pronounced for methanol/ethyl acetate extracts of *S. tropica* cultures harvested six to ten days after inoculation (Figure S2A), suggesting that at least one of the biosynthetic clusters predicted to yield DNA-interfering enediyne compounds was expressed.

S. tropica enediyne genotypes are not responsible for biosynthesis of DNA-interfering chemotypes

To explore whether the candidate gene clusters *ST_pks1* and *spo* are associated with the observed DNA-damaging activity, a key biosynthetic gene from each cluster was inactivated and replaced with a gene for apramycin resistance (*aprR*) through PCR-targeted gene replacement.[32] BIA activities were then compared between chemical extracts from individual mutants and the wild-type strain (Figure S2).

Elimination of the ePKS gene *spoE* from the *spo* cluster yielded a *spoE::apr*^{R} mutant incapable of sporolide production (Figure S1B), confirming that the sporolides are indeed ePKS-derived products as originally suspected. However, chemical extracts from the *spoE::apr*^{R} mutant exhibited the same activity profile as the wild-type strain in the BIA (Figure S2C), indicating that neither sporolides A and B nor their proposed nine-membered enediyne precursor (Figure S1A) were responsible for the observed DNA interference in the BIA. We suspect that this enediyne may decompose too rapidly for BIA detection due to the lack of a protective carrier. Most characterized gene clusters for nine-membered enediynes encode an apoprotein that sequesters the enediyne to enhance its stability and thus facilitate biological activity.[15,23] However, no proteins with homology to characterized enediyne-stabilizing proteins were found within the *S. tropica* genome, suggesting that this organism lacks mechanisms for nine-membered enediyne stabilization and offering an explanation for the apparent instability of the proposed sporolide precursor. A similar transformation was hypothesized for the cyanosporasides from *S. pacifica* CNS-143, which also have a cyclopenta[*a*]indene core structure derived from a putative nine-membered enediyne precursor.[33]

From the *ST_pks1* gene cluster, PCR-targeted gene elimination of the putative ten-membered ePKS encoded by *strop0598* yielded a mutant with BIA activity equivalent to that of the wild type (Figure S2B), thus eliminating *ST_pks1* as a candidate for biosynthesis of the observed DNA-interfering active metabolite. Furthermore, LC-MS profiling of the mutant and wild-type extracts revealed no evidence for enediyne natural products (data not shown), suggesting that the orphan *ST_pks1* gene cluster is either inactive under the selected fermentation conditions or expressed at low levels.

Type II PKS pathway *ST_pks2* produces DNA-interfering compounds in *S. tropica* CNB-440

As no enediyne was responsible for causing DNA-interference in *S. tropica* extracts as originally suspected, we evaluated the other orphan biosynthetic pathways for the responsible agent. Another prominent class of DNA-targeting natural products from actinomycetes are aromatic polyketides such as the anthracycline anticancer agents.[34] Aromatic polyketides are commonly biosynthesized by type II PKSs and are often glycosylated.[35] Glycosylated anthracycline anticancer agents, such as daunomycin, can intercalate into chromosomal DNA with their polyketide moiety, while the glycosyl moiety binds the phosphoribose DNA backbone.[34] The *S. tropica* CNB-440 genome contains two orphan type II PKS pathways, *ST_pks2* and *ST_pks3*. The *ST_pks3* locus (Figure S3A, Table S2) encodes the synthesis of a putative spore pigment polyketide related to the *whiE* dodecaketide product of unknown structure from *Streptomyces coelicolor* A3(2).[36] We explored its function by gene elimination of the predicted *ST_pks3* β-ketosynthase (*strop2500*) and sporulation phenotyping on A1 agar. The *ST_pks3* mutant showed loss of spore pigmentation compared to wild-type *S. tropica* confirming a predicted spore pigment product (Figure S3B). The *ST_pks2* gene cluster, on the other hand, is predicted to encode production of a glycosylated aromatic polyketide, thereby making it a more promising alternative target in the search for a DNA-interfering molecule than the *ST_pks3* spore pigment.

To examine the function of the *ST_pks2* gene cluster (Scheme 1D), we eliminated the putative β-ketosynthase gene *strop2223* to yield a *strop2223::apr*^{R} mutant. The organic extract of this mutant was inactive in the BIA (Figure S2D), thereby correlating the DNA-interference activity observed from wild-type *S. tropica* exclusively to the *ST_pks2* gene cluster and not to either of the enediyne-encoding gene clusters (*ST_pks1* and *spo*). In light of prior research correlating the observation of ePKS-encoding genes with BIA activity,[37] we found this result

surprising. This earlier work, however, did not confirm their observations with gene deletions or compound isolation, thus leaving open the possibility that non-enediyne polyketides were responsible for the observed biological activity.

Characterization of *ST_pks2* as the lomaiviticin biosynthetic gene cluster (*lom*)

To identify the DNA-interfering glycosylated polyketide predicted by bioinformatics analysis of the orphan *ST_pks2* gene cluster, we employed an LC-MS^n-based comparative metabolite profiling strategy. In LC-MS^n traces of the wild-type extract, several masses were detected that were absent in the corresponding mutant sample (Figure 1 A and B, Figure S4). One compound with an exact mass of 1338.5319 Da showed candidate deoxysugar fragments (*m/z* 158.12—putative *N,N*-dimethyldideoxysugar; *m/z* 145.09—putative *O*-methyldideoxysugar) in the MS/MS spectrum (Figure 1 C), indicating a putative glycosylated natural product. The compound was isolated as a burgundy-red solid by MS-guided fractionation through reversed-phase flash column chromatography from a methanol/ethyl acetate extract of a culture of *S. tropica* CNB-440. NMR analysis enabled the characterization of the molecule as lomaiviticin C (Figures 1 D and S5, Table S3), a glycosylated aromatic polyketide recently described from *S. pacifica* strain DPJ-0019.[26] In addition, other known derivatives, lomaiviticins A, D and E, were identified by LC-MS^n (Figure S6).[27,28] In the β-ketosynthase mutant, no nucleoside diphosphate-deoxysugar accumulation was detected (data not shown) as these NDP-species should have a generally short lifetime in cell metabolism due to their high energy content.

Lomaiviticin A was first isolated by Carter et al.[27] and is the derivative with most potent anticancer activity.[26] Its C–C dimeric diazofluorene core is also found in its monomeric form in kinamycin natural products from *Streptomyces murayamaensis*.[28,38] While the kinamycin gene cluster (*kin*) was previously reported by Gould et al. in 1998,[39] the lomaiviticin biosynthetic gene cluster (*lom*) remained unknown until now. The *lom* locus is predicted to comprise 59 ORFs, which includes all biosynthetic genes putatively involved in the construction of the diazofluorene core in kinamycin (Table 1, Figure S7).[28,39] The *lom* cluster contains three additional PKS genes adjacent to the minimal type II PKS genes for the incorporation of a propionate starter unit in contrast to a starter acetate unit in kinamycin (Figure S7). Putative biosynthetic gene products include two glycosyltransferases and several deoxysugar biosynthetic enzymes that nicely correlate with the glycosylation pattern of the lomaiviticins (Figure S8). In addition, the gene cluster contains a FAD-dependent monooxygenase gene, *strop2191*, that is homologous to ActVA-Orf4 from the actinorhodin biosynthetic gene cluster in *S. coelicolor* A3(2). ActVA-Orf4 catalyzes the C–C dimerization reaction of two benzoisochromanequi-

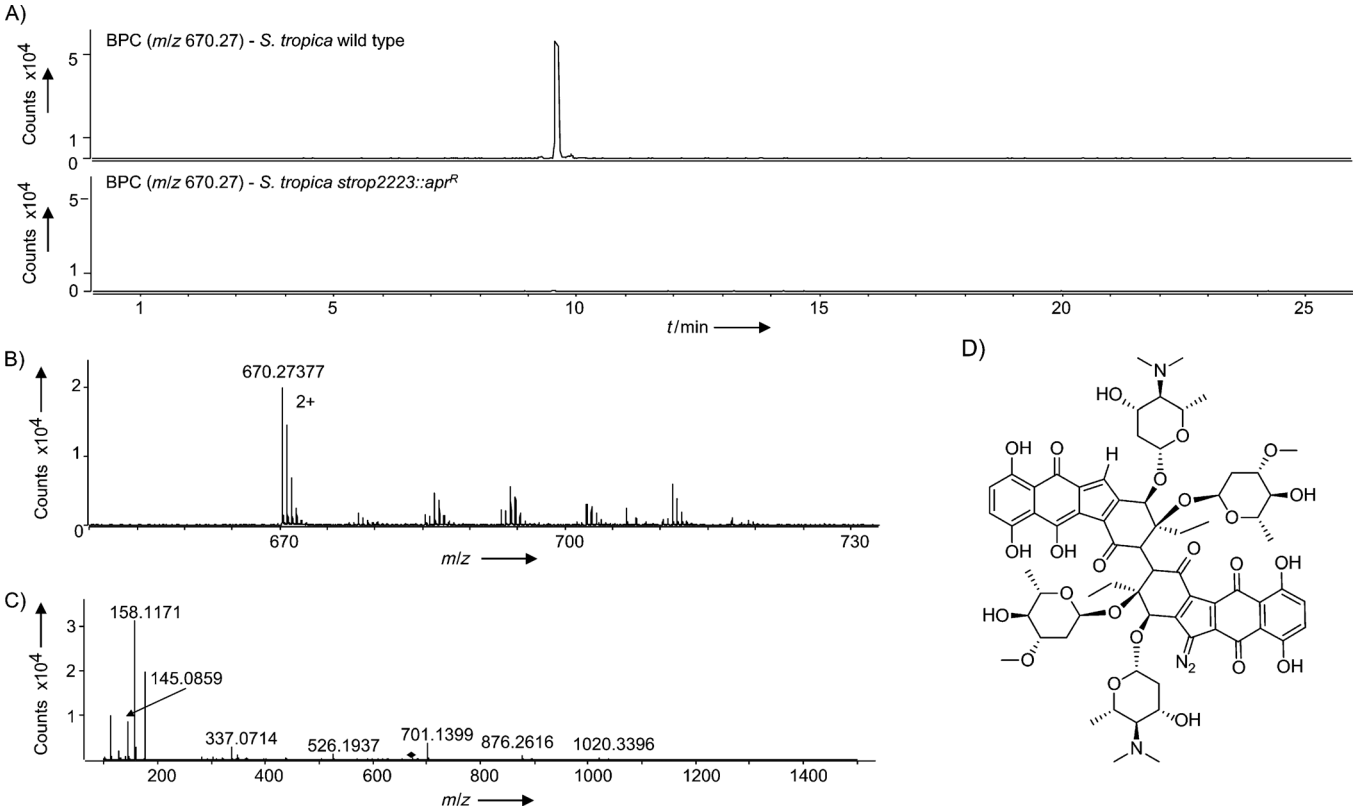


Figure 1. Identification of the *ST_pks2 (lom)* gene cluster product, lomaiviticin C, from extracts of wild-type *S. tropica* by LC-MS-based comparative metabolic profiling. A) Comparative metabolic profiling identified a compound at *m/z* 670.27 with abolished production in the *ST_pks2* mutant (shown by corresponding base peak chromatogram (BPC)). B) MS spectrum of the *ST_pks2* product. C) MS/MS spectrum of the *ST_pks2* product showed deoxysugar-specific fragments indicating a glycosylated natural product. D) The identified *ST_pks2* product lomaiviticin C.

Table 1. Organization of the lomaiviticin biosynthetic gene cluster (*lom*) from *S. tropica* CNB-440.

Gene	Size [aa]	Predicted function	Closest homologue [similarity/identity, (%/%)]	Ref.	Kinamycin homologue
strop2172	524	drug resistance transporter	drug resistance transporter [*Salinispora arenicola* CNS-205] (84/90)	ABV98201	
strop2173	161	hypothetical protein	hypothetical protein [*Salinispora arenicola* CNS-205] (75/86)	ABV98202	
strop2174	276	ABC transporter	ABC transporter [*Catenulispora acidiphila* DSM 44928] (56/39)	ACU70166	
strop2175	316	ABC transporter	ABC transporter [*Streptomyces sviceus* ATCC 29083] (60/74)	EDY54089	
strop2176	171	bleomycin resistance protein	bleomycin resistance protein [*Streptomyces pristinaespiralis* ATCC 25486] (63/79)	EDY62258	
strop2177	258	hypothetical protein	alpha/beta hydrolase [*Streptomyces* sp. Mg1] (49/62)	EDX26142	
strop2178	278	transcriptional activator	transcriptional activator [*Streptomyces coelicoflavus* ZG0656] (68/52)	EHN80233	
strop2179	113	hypothetical protein	hypothetical protein [*Lodderomyces elongisporus* NRRL YB-4239] (63/40)	XP_001524904	
strop2180	281	AraC-family transcriptional regulator	AraC-family transcriptional regulator [*Micromonospora* sp. ATCC 39149] (61/45)	ZP_04603952	
strop2181	477	NDP-hexose 2,3-dehydratase	3-dehydratase [*Streptomyces coelicoflavus* ZG0656] (81/72)	EHN77727	
strop2182	221	hypothetical protein	hypothetical protein [*Streptomyces coelicoflavus* ZG0656] (79/63)	EHN77694	
strop2183	148	carboxymuconolactone decarboxylase	carboxymuconolactone decarboxylase [*Streptomyces coelicoflavus* ZG0656] (73/58)	EHN77723	
strop2184	275	glutamine amidotransferases	peptidase C26 [*Streptomyces coelicoflavus* ZG0656] (77/67)	EHN77724	
strop2185	590	aromatic ring hydroxylase	XiaK [*Streptomyces* sp. SCSIO 02999] (61/47)	AFK78077	
strop2186	488	protoporphyrinogen oxidase	protoporphyrinogen oxidase [*Streptomyces coelicoflavus* ZG0656] (73/62)	EHN77710	
strop2187	368	ParB domain protein nuclease	nuclease, partial [*Streptomyces coelicoflavus* ZG0656] (79/67)	HN77721	
strop2188	503	anthrone hydroxylase	FAD-binding monooxygenase [*Streptomyces coelicoflavus* ZG0656] (78/71)	EHN77709	KinO2 (60/47)
strop2189	490	anthrone hydroxylase	FAD-binding monooxygenase [*Streptomyces coelicoflavus* ZG0656] (75/66)	EHN77708	
strop2190	263	3-oxoacyl-ACP reductase	short-chain dehydrogenase/reductase [*Streptomyces coelicoflavus* ZG0656] (82/71)	EHN77698	
strop2191	290	putative dimerase	NmrA family protein [*Streptomyces coelicoflavus* ZG0656] (76/67)	EHN77697	
strop2192	205	DSBA oxidoreductase	DSBA oxidoreductase [*Streptomyces hygroscopicus* ATCC 53653] (71/57)	ZP_07297742	
strop2193	109	polyketide synthesis cyclase	cyclase [*Streptomyces antibioticus*] (85/75)	CAG14964	KinI (82/70)
strop2194	261	polyketide ketoreductase	short-chain dehydrogenase/reductase [*Streptomyces bingchenggensis* BCW-1] (85/77)	YP_004965103	KinE (80/71)
strop2195	345	polyketide O-methyltransferase	O-methyltransferase [*Streptomyces coelicoflavus*] (78/63)	EHN77696	
strop2196	503	anthrone hydroxylase	SaqE [*Micromonospora* sp. Tu 6368] (70/59)	ACP19351	KinOR (60/47)
strop2197	319	polyketide aromatase	SaqL [*Micromonospora* sp. Tu 6368] (71/61)	ACP19357	KinD (67/55)
strop2198	479	FAD-dependent monooxygenase	putative FAD-dependent monooxygenase [*Streptomyces albaduncus*] (61/52), JagF	CBH32087	
strop2199	492	anthrone hydroxylase	oxygenase-like protein [*Streptomyces murayamaensis*] (73/63)	AAO65343	KinO1 (73/63)
strop2200	247	anthrone oxidase	JadG [*Streptomyces venezuelae*] (62/46)	AAV52247	KinG (68/52)
strop2201	627	hypothetical protein	hypothetical protein [*Streptomyces ambofaciens*] (71/61)	CAK51011	
strop2202	118	4Fe-4S ferredoxin	putative ferredoxin [*Streptomyces ambofaciens* ATCC 23877] (83/75)	CAI78079	
strop2203	264	hypothetical protein	hypothetical protein [*Salinispora tropica* CNB-440] (78/65)	YP_001159312	
strop2204	500	glutamine synthetase	glutamine synthetase [uncultured bacterium BAC AB649/1850] (75/65)	AEE65491	
strop2205	484	amidase	putative amidase [*Streptomyces coelicoflavus* ZG0656] (73/63)	EHN77687	
strop2206	428	adenylosuccinate lyase	putative adenylosuccinate lyase [*Streptomyces coelicoflavus* ZG0656] (80/70)	EHN77655	
strop2207	137	N-acetyltransferase GCN5	putative acetyltransferase [*Streptomyces ambofaciens* ATCC 23877] (79/72)	CAI78074	
strop2208	523	ABC transporter-like protein	NovA [*Streptomyces coelicoflavus* ZG0656] (82/69)	EHN77703	
strop2209	106	dioxygenase	dioxygenase [*Thermobifida fusca* YX] (76/73)	YP_290041	
strop2210	136	bleomycin resistance protein	bleomycin resistance protein [*Salinispora arenicola* CNS-205] (97/92)	YP_001537194	
strop2211	263	bleomycin resistance protein	bleomycin resistance protein [*Nesterenkonia* sp. F] (61/42)	ZP_09541316	
strop2212	520	secreted peptidase	TAP domain-containing protein [*Streptomyces coelicoflavus* ZG0656] (72/61)	EHN77722	
strop2213	380	glycosyltransferase	glycosyltransferase [uncultured bacterium] (57/41)	AEM44235	
strop2214	371	polyketide O-methyltransferase	O-methyltransferase [*Ktedonobacter racemifer* DSM 44963]	ZP_06965007	

Table 1. (Continued)

Gene	Size [aa]	Predicted function	Closest homologue [similarity/identity, (%/%)]	Ref.	Kinamycin homologue
			(66/56)		
strop2215	313	O-methyltransferase	NanM [*Streptomyces nanchangensis*] (77/60)	AAP42862	
strop2216	328	NDP-hexose 4-ketoreductase	4-ketoreductase [*Streptomyces* sp. TP-A0274] (68/56)	BAC55215	
strop2217	199	NDP-hexose 3,5-epimerase	3,5-epimerase [*Streptomyces* sp. TP-A0274] (77/67)	BAC55217	
strop2218	335	NDP-hexose 3-ketoreductase	putative 3-ketoreductase [*Streptomyces galilaeus*] (68/55)	AAF73453	
strop2219	244	glycosyltransferase (auxiliary)	hypothetical protein [*Salinispora arenicola* CNS-205] (88/80)	YP_001537195	
strop2220	344	glycosyltransferase	glycosyltransferase [*Streptomyces cyanogenus*] (55/42)	AAD13553	
strop2221	401	NDP-hexose 3,4-dehydratase/ aminotransferase	3,4-dehydratase-like protein [*Streptomyces* sp. KCTC 0041BP] (74/61)	ABB52533	
strop2222	333	dTDP-glucose-4,6-dehydratase	putative dTDP-glucose-4,6-dehydratase [*Actinoplanes missouriensis* 431] (79/69)	YP_005462170	
strop2223	423	minimal type II PKS, KS	beta-ketoacyl synthase [*Streptomyces acidiscabies* 84-104] (82/71)	ZP_10452176	KinA (79/66)
strop2224	423	minimal type II PKS, CLF	putative chain length factor [*Streptomyces ravidus*] (76/62)	CBH32808	KinB (73/59)
strop2225	84	minimal type II PKS, ACP	ACP [*Thermomonospora curvata* DSM 43183] (73/48)	YP_003300377	KinC (62/47)
strop2226	327	propionate starter unit, AT	AknF [*Streptomyces galilaeus*] (64/53)	BAB72049	
strop2227	322	propionate starter unit, KS	putative modular polyketide synthase [*Kitasatospora setae* KM-6054] (45/30)	BAJ32815	
strop2228	107	propionate starter unit, ACP	actinorhodin polyketide dimerase, ACP [*Thermomonospora fusca* YX] (71/51)	YP_289281	
strop2229	160	AraC family transcriptional regulator	AraC family transcriptional regulator [*Salinispora arenicola* CNS-205] (91/87)	YP_001537196	
strop2230	290	glucose-1-phosphate thymidylyltransferase	G1P thymidylyltransferase [*Thermomonospora curvata* DSM 43183] (87/72)	YP_003300347	

none monomers in actinorhodin biosynthesis.[40] We thus suspect that its homologue, Strop2191, catalyzes the dimerization of kinamycin-like diazofluorene monomers in lomaiviticin biosynthesis. The dimerization could occur from C2 after a dihydroxylation at C3 and C4 and a hydroxylation at C10 as enzymatic hydroxylations might be sterically hindered in the dimer (Figure S7). A homologue of Strop2191 is not present in the homologous *kin* cluster. These diazo-forming and dimerizing enzymes await characterization.

Conclusions

In this study, we identified the lomaiviticin biosynthetic gene cluster (*ST_pks2*=*lom*) in *S. tropica* CNB-440 through bioactivity-guided genome mining. Motivated by *ST_pks1* and *spo*, two *S. tropica* enediyne gene clusters without characterized enediyne products, we analyzed organic extracts with the BIA, an assay for rapid screening of DNA-interfering chemotypes, to detect a *S. tropica* DNA-targeting chemotype. Although we did indeed measure a BIA-based activity, gene elimination experiments showed that this observed bioactivity was not caused by an enediyne pathway product. Instead we correlated the BIA bioactivity to the products of the type II PKS pathway, *ST_pks2*. Comparative metabolic profiling yielded the lomaiviticins as *ST_pks2* products and, thus, as the observed BIA-positive compounds. Our approach of using the BIA in combination with genome mining and gene elimination might prove an effective strategy to identify new DNA-targeting chemo- and genotypes beyond enediyne scaffolds. The identified biosynthetic gene cluster for a highly complex and stereospecific natural product such as lomaiviticin A can open up semisynthetic and heterologous expression routes as alternatives for their challenging production through total synthesis.

Experimental Section

General: All chemicals were acquired from Fisher Scientific, Sigma–Aldrich and Honeywell Burdick & Jackson, and solvents were of HPLC grade or higher. HPLC analyses were conducted with an Agilent 1200 HPLC system with diode array detection. LC-MS analyses were conducted with an Agilent 6530 Accurate-Mass Q-TOF MS (MassHunter software, Agilent) equipped with a Dual electrospray ionization source and with an Agilent 1260 LC system (ChemStation software, Agilent) with diode array detector. NMR data were acquired at the UCSD Skaggs School of Pharmacy and Pharmaceutical Sciences NMR Facility on a 600 MHz Varian NMR spectrometer (Topspin 2.1.6 software, Bruker) with a 1.7 mm cryoprobe.

Gene inactivation: Targeted biosynthetic genes within *S. tropica* gene clusters (*ST_pks1, ST_pks3, lom*, and *spo*) predicted to yield DNA-interfering natural products were inactivated by PCR-based mutagenesis following previously established methods.[32] In brief, the apramycin resistance (*aac(3)IV*) cassette from pIJ773 was PCR amplified and extended using primer sequences (Table S4). Extended antibiotic resistance cassettes were introduced by electroporation into *E. coli* BW25113/pKD20 carrying appropriate pCCFOS-based fosmids BHXS2039 (for *ST_pks1* genes), BHXS0939 (for *ST_pks3* genes), BHXS5407 (for *lom* genes), and BHXS4676 (for *spo* genes). The mutated fosmid was then transferred into *S. tropica* CNB-440 by conjugation from *E. coli* S17-1, and gene replacement was confirmed by colony PCR and sequencing of PCR products. Gene replacement experiments were minimally carried out in duplicate.

***S. tropica* fermentation and BIA:** Wild-type *S. tropica* CNB-440 as well as *spoE::apr*^{R}, *strop2223::apr*^{R}, *strop2500::apr*^{R}, and *stro-*

p0598::apr^R mutants were cultured at 28 °C in Fernbach flasks containing A1 medium (1 L, 1% starch, 0.4% yeast extract, 0.2% peptone, 0.1% calcium carbonate, 3% InstantOcean sea salt). Inoculation occurred with A1 medium starter cultures (10 mL) in falcon tubes (50 mL). Flasks were shaken at 225 rpm, and time point samples (50 mL) were drawn every two days starting on day 4. Each sample (50 mL) was divided into two equal portions. One portion was clarified by centrifugation at 5000 rpm for 10 min to provide a supernatant sample. The remaining mycelial pellet from this portion was suspended in methanol (MeOH, 25 mL), stirred for 30 min at room temperature, and cell debris subsequently removed by centrifugation (7000 *g*, 10 min, 4 °C). The other portion of *S. tropica* fermentation (25 mL) was extracted three times with equal volumes of ethyl acetate. Supernatant samples, as well as MeOH and EtOAc extracts, were concentrated to dryness in vacuo. Samples were resuspended at 1 mg mL^{-1} in their respective extraction solvents. The BIA was performed as previously described.[31] In brief, extract (10 μL) was applied to agar plates seeded with *E. coli* ATCC 33312 and incubated for 5 h at 37 °C. Soft agar containing Fast Blue RR salt (1.2 mg mL^{-1}) and 6-bromo-2-naphthyl-β-D-galactopyranoside (0.4 mg mL^{-1}) was added onto the plate and color development was observed within 15 min. Cisplatin was included as a positive reference (test concentration: 5 mg mL^{-1} and twofold serial dilutions thereof). The fermentation medium and solvents were also assayed alone and negative for BIA activity. Every assay was run at least in duplicate. For sporulation phenotyping, wild-type *S. tropica* CNB-440 and *strop2500::apr*^{R} mutant were inoculated on an A1 agar plate (18 g agar per liter A1 medium) and incubated for 14 days at 28 °C.

Comparative metabolite profiling of wild-type and mutant *S. tropica* chemical extracts by LC-MS: Crude MeOH/EtOAc extracts of wild-type and mutant *S. tropica* strains were filtered through Acrodisc MS PTFE Syringe filters (Pall Inc., Ann Arbor, MI, USA) and adjusted in concentration (0.2 mg mL^{-1}). For HPLC analysis, 25 μL was injected on a Luna reversed-phase C18 (5 μm, 4.6× 100 mm) HPLC column and subjected to a 10–100% MeCN (0.1% TFA)/0.1% aq. TFA gradient (15 min, 0.7 mL min^{-1}). For LC-MS^n analysis, crude extract (2 μg) was injected onto a Phenomenex Luna C18 reversed-phase HPLC column (5 μm, 150×4.6 mm) and was analyzed with an Agilent 6530 Accurate-Mass LC-MS with an Agilent 1260 LC system under the following LC conditions (1–3 min—10% MeCN (0.1% TFA), 90% aq. 0.1% TFA, 3–23 min—10–100% MeCN (0.1% TFA), 23–25 min—100% MeCN (0.1% TFA), 0.7 mL min^{-1}). Q-TOF MS settings during the LC gradient were as follows: Acquisition—mass range *m/z* 300–1500, MS scan rate 1 s^{-1}, MS/MS scan rate 2 s^{-1}, fixed collision energy 20 eV; Source—gas temperature 300 °C, gas flow 11 L min^{-1}; Nebulizer 45 psig, ion polarity positive; Scan source parameters—VCap 3000, Fragmentor 100, Skimmer1 65, OctopoleRFPeak 750. The MS was autotuned using Agilent tuning solution in positive mode before each measurement. LC(DAD) data were analyzed with ChemStation software (Agilent), and MS data were analyzed with MassHunter software (Agilent).

Isolation and characterization of *ST_pks2* product, lomaiviticin C: Lomaiviticin C was isolated by MS-guided fractionation from an A1 medium culture (1 L) of wild-type *S. tropica* CNB-440 after ten days incubation. The cells of 500 mL of the culture were extracted with MeOH for 30 min with stirring, and the remaining supernatant and culture were extracted twice with EtOAc. The crude organic extract was concentrated in vacuo, resuspended in MeOH (2 mL), and loaded on a reversed-phase C18 silica gel column for flash-column chromatography. Lomaiviticin C was further purified as previously described[26] using LC-MS for identification of lomaiviticin C-containing fractions. Purified lomaiviticin C was dissolved in CD_3OD and subjected to NMR structure elucidation (^{1}H, DQF-COSY, ^{1}H-^{13}C HMBC, NOESY). NMR data were analyzed with Topspin 2.1.6 software (Bruker).

Acknowledgements

We thank Paul R. Jensen and William Fenical for providing S. tropica *CNB-440. This work was supported by a postdoctoral fellowship from the German Academic Exchange Service (DAAD) to M.N. and an NIH IRACDA postdoctoral fellowship (GM068524) to A.L.L. This work was supported by grants from the NIH (GM085770 to B.S.M., GM097509 to P.C.D. and instrument grant 1-S10-RR031562-01A1).*

Keywords: biosynthesis · genome mining · lomaiviticins · polyketides · *Salinispora tropica*

[1] M. Nett, H. Ikeda, B. S. Moore, *Nat. Prod. Rep.* **2009**, *26*, 1362 – 1384.
[2] G. L. Challis, *J. Med. Chem.* **2008**, *51*, 2618 – 2628.
[3] C. Corre, G. L. Challis, *Nat. Prod. Rep.* **2009**, *26*, 977 – 986.
[4] J. M. Winter, S. Behnken, C. Hertweck, *Curr. Opin. Chem. Biol.* **2011**, *15*, 22 – 31.
[5] L. J. Song, F. Barona-Gomez, C. Corre, L. K. Xiang, D. W. Udwary, M. B. Austin, J. P. Noel, B. S. Moore, G. L. Challis, *J. Am. Chem. Soc.* **2006**, *128*, 14754 – 14755.
[6] A. S. Eustáquio, S.-J. Nam, K. Penn, A. Lechner, M. C. Wilson, W. Fenical, P. R. Jensen, B. S. Moore, *ChemBioChem* **2011**, *12*, 61 – 64.
[7] a) S. Lautru, R. J. Deeth, L. M. Bailey, G. L. Challis, *Nat. Chem. Biol.* **2005**, *1*, 265 – 269; b) K. Ishida, T. Lincke, S. Behnken, C. Hertweck, *J. Am. Chem. Soc.* **2010**, *132*, 13966 – 13968; c) L. Laureti, L. Song, S. Huang, C. Corre, P. Leblond, G. L. Challis, B. Aigle, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6258 – 6263.
[8] H. Gross, V. O. Stockwell, M. D. Henkels, B. Nowak-Thompson, J. E. Loper, W. H. Gerwick, *Chem. Biol.* **2007**, *14*, 53 – 63.
[9] L. C. Blasiak, J. Clardy, *J. Am. Chem. Soc.* **2010**, *132*, 926 – 927.
[10] R. D. Kersten, Y. L. Yang, Y. Xu, P. Cimermancic, S. J. Nam, W. Fenical, M. A. Fischbach, B. S. Moore, P. C. Dorrestein, *Nat. Chem. Biol.* **2011**, *7*, 794 – 802.
[11] B. La Ferla, C. Airoldi, C. Zona, A. Orsato, F. Cardona, S. Merlo, E. Sironi, G. D'Orazio, F. Nicotra, *Nat. Prod. Rep.* **2011**, *28*, 630 – 648.
[12] a) D. J. Newman, G. M. Cragg, *Curr. Opin. Invest. Drugs* **2009**, *10*, 1280 – 1296; b) C. Olano, C. Mendez, J. A. Salas, *Nat. Prod. Rep.* **2009**, *26*, 628 – 660.
[13] a) P. R. Hamann, L. M. Hinman, I. Hollander, C. F. Beyer, D. Lindh, R. Holcomb, W. Hallett, H. R. Tsou, J. Upeslacis, D. Shochat, A. Mountain, D. A. Flowers, I. Bernstein, *Bioconjugate Chem.* **2002**, *13*, 47 – 58; b) L. Ducry, B. Stump, *Bioconjugate Chem.* **2010**, *21*, 5 – 13.
[14] U. Galm, M. H. Hager, S. G. Van Lanen, J. H. Ju, J. S. Thorson, B. Shen, *Chem. Rev.* **2005**, *105*, 739 – 758.
[15] Z. X. Liang, *Nat. Prod. Rep.* **2010**, *27*, 499 – 528.
[16] A. L. Smith, K. C. Nicolaou, *J. Med. Chem.* **1996**, *39*, 2103 – 2117.
[17] K. Penn, C. Jenkins, M. Nett, D. W. Udwary, E. A. Gontang, R. P. McGlinchey, B. Foster, A. Lapidus, S. Podell, E. E. Allen, B. S. Moore, P. R. Jensen, *ISME J.* **2009**, *3*, 1193 – 1203.
[18] D. W. Udwary, L. Zeigler, R. N. Asolkar, V. Singan, A. Lapidus, W. Fenical, P. R. Jensen, B. S. Moore, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10376 – 10381.
[19] A. A. Roberts, A. W. Schultz, R. D. Kersten, P. C. Dorrestein, B. S. Moore, *FEMS Microbiol. Lett.* **2012**, *335*, 95 – 103.
[20] A. S. Eustáquio, R. P. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Florova, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. A. Reynolds, B. S. Moore, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12295 – 12300.

[21] A. Miyanaga, J. E. Janso, L. McDonald, M. He, H. Liu, L. Barbieri, A. S. Eustáquio, E. N. Fielding, G. T. Carter, P. R. Jensen, X. Feng, M. Leighton, F. E. Koehn, B. S. Moore, *J. Am. Chem. Soc.* **2011**, *133*, 13311 – 13313.
[22] R. P. McGlinchey, M. Nett, B. S. Moore, *J. Am. Chem. Soc.* **2008**, *130*, 2406 – 2407.
[23] M. Jean, S. Tomasi, P. Van de Weghe, *Org. Biomol. Chem.* **2012**, *10*, 7453 – 7456.
[24] G. O. Buchanan, P. G. Williams, R. H. Feling, C. A. Kauffman, P. R. Jensen, W. Fenical, *Org. Lett.* **2005**, *7*, 2731 – 2734.
[25] a) T. A. Gulder, B. S. Moore, *Angew. Chem.* **2010**, *122*, 9534 – 9556; *Angew. Chem. Int. Ed.* **2010**, *49*, 9346 – 9367; b) R. H. Feling, G. O. Buchanan, T. J. Mincer, C. A. Kauffman, P. R. Jensen, W. Fenical, *Angew. Chem.* **2003**, *115*, 369 – 371; *Angew. Chem. Int. Ed.* **2003**, *42*, 355 – 357.
[26] C. M. Woo, N. E. Beizer, J. E. Janso, S. B. Herzon, *J. Am. Chem. Soc.* **2012**, *134*, 15285 – 15288.
[27] H. He, W. D. Ding, V. S. Bernan, A. D. Richardson, C. M. Ireland, M. Greenstein, G. A. Ellestad, G. T. Carter, *J. Am. Chem. Soc.* **2001**, *123*, 5362 – 5363.
[28] S. B. Herzon, C. M. Woo, *Nat. Prod. Rep.* **2012**, *29*, 87 – 118.
[29] a) M. Konishi, H. Ohkuma, T. Tsuno, T. Oki, G. D. VanDuyne, J. Clardy, *J. Am. Chem. Soc.* **1990**, *112*, 3715 – 3716; b) Q. Gao, J. S. Thorson, *FEMS Microbiol. Lett.* **2008**, *282*, 105 – 114.
[30] C. L. Perrin, B. L. Rodgers, J. M. O'Connor, *J. Am. Chem. Soc.* **2007**, *129*, 4795 – 4799.
[31] R. K. Elespuru, R. J. White, *Cancer Res.* **1983**, *43*, 2819 – 2830.
[32] a) B. Gust, G. L. Challis, K. Fowler, T. Kieser, K. F. Chater, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1541 – 1546; b) A. S. Eustáquio, F. Pojer, J. P. Noel, B. S. Moore, *Nat. Chem. Biol.* **2008**, *4*, 69 – 74.
[33] A. L. Lane, S. J. Nam, T. Fukuda, K. Yamanaka, C. A. Kauffman, P. R. Jensen, W. Fenical, B. S. Moore, *J. Am. Chem. Soc.* **2013**, *135*, 4171 – 4174.
[34] A. Rabbani, R. M. Finn, J. Ausio, *Bioessays* **2005**, *27*, 50 – 56.
[35] I. Fujii, Y. Ebizuka, *Chem. Rev.* **1997**, *97*, 2511 – 2523.
[36] T. W. Yu, Y. Shen, R. McDaniel, H. G. Floss, C. Khosla, D. A. Hopwood, B. S. Moore, *J. Am. Chem. Soc.* **1998**, *120*, 7749 – 7759.
[37] E. Zazopoulos, K. X. Huang, A. Staffa, W. Liu, B. O. Bachmann, K. Nonaka, J. Ahlert, J. S. Thorson, B. Shen, C. M. Farnet, *Nat. Biotechnol.* **2003**, *21*, 187 – 190.
[38] a) S. Ito, T. Matsuya, S. Omura, M. Otani, A. Nakagawa, *J. Antibiot.* **1970**, *23*, 315 – 317; b) S. J. Gould, N. Tamayo, C. R. Melville, M. C. Cone, *J. Am. Chem. Soc.* **1994**, *116*, 2207 – 2208.
[39] S. J. Gould, S. T. Hong, J. R. Carney, *J. Antibiot.* **1998**, *51*, 50 – 57.
[40] T. Taguchi, T. Ebihara, A. Furukawa, Y. Hidaka, R. Ariga, S. Okamoto, K. Ichinose, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 5041 – 5045.

Received: March 14, 2013
Published online on May 3, 2013