CHEMBIOCHEM

DOI: 10.1002/cbic.200800822

Analysis of the Indanomycin Biosynthetic Gene Cluster from *Streptomyces antibioticus* NRRL 8167

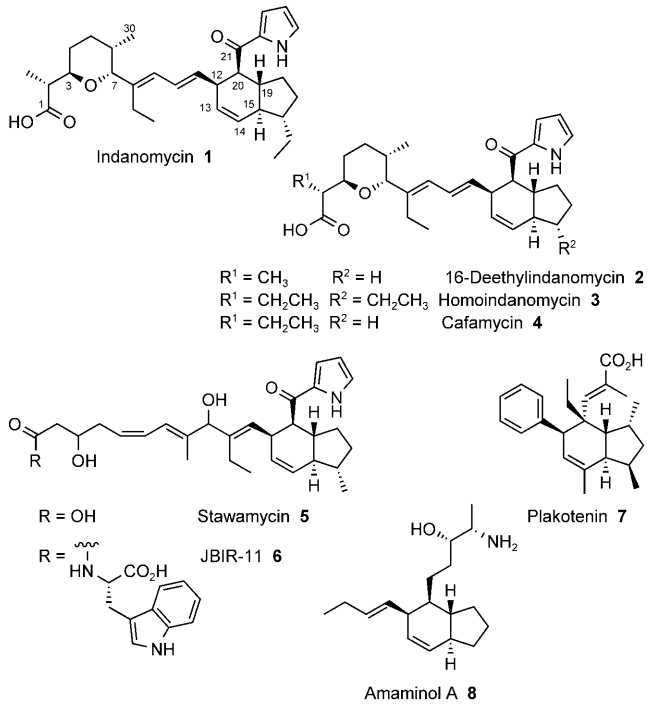
Chaoxuan Li, Kathryn E. Roege, and Wendy L. Kelly*[a]

Metabolites that harbor a core indane scaffold are found to have diverse biological properties. Indanomycin and related pyrroloketoindanes are ionophores and have demonstrated antiparasitic, insecticidal, and antibacterial activities. To understand the biochemical mechanisms guiding formation of the central indane ring, the biosynthetic gene cluster for indanomycin was identified from *Streptomyces antibioticus* NRRL 8167 and sequenced to ~80 kb; this revealed five genes encoding subunits of a polyketide synthase (PKS) and 18 other open reading frames. The involvement of this cluster in indanomycin biosynthesis was confirmed by deletion mutagenesis. The indanomycin PKS lacks the expected thioesterase at the carboxy terminus of the final module, and instead appears to house an incomplete module containing an unusual cyclase domain. These findings now enable additional detailed genetic and biochemical studies of the mechanisms guiding the generation of pyrroloketoindanes.

Introduction

The pyrroloketoindanes are hybrid nonribosomal peptide-polyketide metabolites and are produced by various *Streptomyces* species (Scheme 1). This small family of antibiotics includes indanomycin (**1**), 16-deethylindanomycin (**2**), homoindanomycin (**3**), cafamycin (**4**), stawamycin (**5**), and JBIR-11 (**6**).[1–6] The indanomycins and cafamycin possess a tetrahydropyran ring and are effective ionophoric agents. These ionophores are active against Gram-positive bacteria, including *Staphylococcus aureus*,[7] and demonstrate insecticidal and antiprotozoal activities.[8,9] Other ionophores, such as the polyether-containing monensin and lasalocid are utilized in veterinary medicine and animal husbandry and have exhibited antimalarial properties.[8,10] Cation-coordination of the indanomycin antibiotics is conferred by complexation of the terminal carboxylic acid and the ether oxygen from two molecules of indanomycin to form a jaw-like dimer surrounding the monovalent or divalent metal.[7]

Indane-containing antibiotics lacking the tetrahydropyran moiety demonstrate an array of biological properties. Stawamycin (**5**) is modestly active against the Epstein–Barr virus, inhibiting binding of the viral transcription factor BZLF1, to its DNA target.[4] Antitumor activity was reported for JBIR-11 (**6**), a recently isolated analogue of stawamycin in which the free carboxylate is replaced by an amide-linked tryptophan residue.[1] Other indane-containing metabolites such as the plakotenins (**7**), and the sphingolipid derivative amaminol A (**8**) harbor cytotoxic properties.[11,12] The unifying feature among all of these metabolites is the indane ring system, and it is notable that secondary metabolites of distinct biosynthetic origins have converged to this common scaffold. The ability to utilize the indanomycin biosynthetic machinery for the production of designer indane metabolites with directed biological properties



Scheme 1. Structures of indanomycin and other representative indane metabolites.

[a] *Dr. C. Li, K. E. Roege, Prof. W. L. Kelly*
School of Chemistry and Biochemistry and the Parker H. Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology
Atlanta, Georgia 30332 (USA)
Fax: (+1) 404-894-2295
E-mail: wendy.kelly@chemistry.gatech.edu

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200800822.

would prove a valuable addition to the biosynthetic engineer's toolkit.

Indanomycin is produced by *Streptomyces antibioticus* NRRL 8167 (*S. antibioticus*).[3] Examination of indanomycin and the results of previous feeding studies reveal that the backbone may be assembled by a hybrid nonribosomal peptide synthetase-polyketide synthase (NRPS-PKS) complex from the building blocks of L-proline, malonyl-CoA, methylmalonyl-CoA and ethylmalonyl-CoA.[13] To date, no biochemical or genetic studies have been reported for indanomycin or any other pyrroloketoindane. We describe here the sequence and analysis of the 80 kb indanomycin biosynthetic gene cluster. Requirement of this locus for indanomycin production is confirmed by inactivation of a gene required for the generation of the pyrrole starter unit, and the boundaries confirmed by deletion mutagenesis. The functional assignments of the proteins encoded in this region are in agreement with the enzymes needed for construction of indanomycin. Taken together, these results establish that the identified genetic locus is responsible for indanomycin biosynthesis in *S. antibioticus* and set the stage for further investigation of the enzymatic transformations leading to the core structural features of pyrroloketoindane antibiotics.

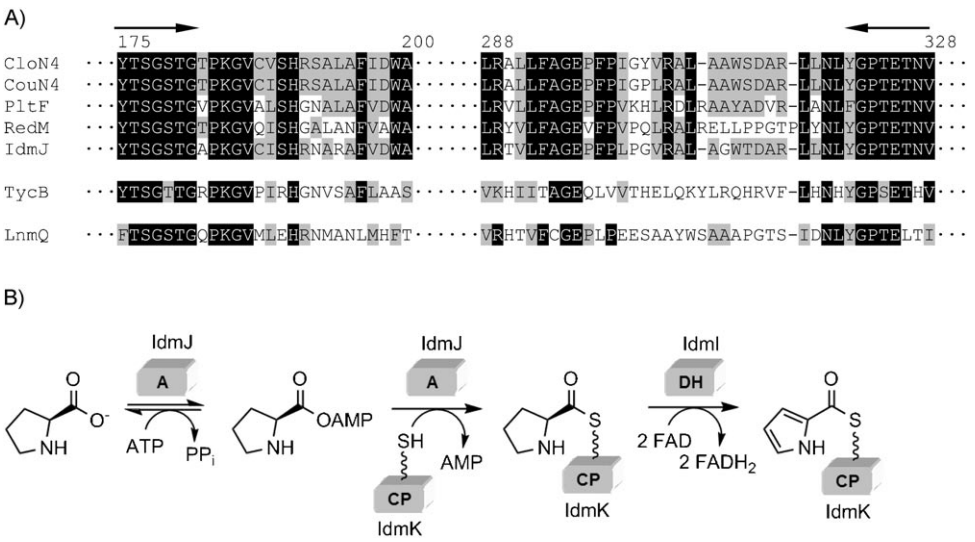


Figure 1. A) Alignment of the partial amino acid sequence of IdmJ with other amino acid activation domains. CloN4, CouN4, PltF, and RedM are four discrete proline-activation domains involved in biosynthesis of clorobiocin, coumermycin, pyoluteorin, and undecylprodigiosin, respectively. TycB is an internal proline-activation domain from the tyrocidine synthetase subunit TycB. LnmQ is a discrete alanine-activation enzyme from the leinamycin synthase. The arrows above the alignment indicate the sequences targeted for PCR amplification. B) Proposed biosynthesis of the indanomycin pyrrole.

Results and Discussion

Identification of the indanomycin biosynthetic gene cluster in *S. antibioticus* NRRL 8167

The strategy adopted to locate the indanomycin biosynthetic gene cluster was to screen for a gene encoding a free-standing proline adenylyltransferase. The proline-derived pyrrole is consistent with a carrier protein-mediated strategy observed for several other pyrrole-containing antibiotics, such as clorobiocin, coumermycin, undecylprodigiosin, and pyoluteorin.[14–16] In these NRPS-based systems, L-proline is covalently sequestered on a prolyl carrier protein (CP) by a free-standing proline activation domain and is then oxidized to the pyrrole by a L-prolyl-*S*-CP dehydrogenase.[15] A pair of degenerate PCR primers was designed according to the A3 and A5 motifs of NRPS amino acid activation (A) domains (Figure 1 A) to produce a product of approximately 460 bp.[17] The A5 motif primer was tailored to target the highly conserved amino acid sequence surrounding this region unique to the discrete proline adenylyltransferases of pyrrole biosynthetic systems (Figure 1 A). Indeed, analysis of the translated 458 bp PCR product from *S. antibioticus* genomic DNA revealed the fragment of a protein, later designated as IdmJ, with significant similarity (~90%) to the sequences of corresponding fragments from adenylation enzymes dedicated to the purpose of pyrrole generation (Figure 1 A). The encoded *S. antibioticus* protein fragment and the free-standing prolyl adenylyltransferases have lower sequence similarity to other A domains, as shown here when compared to a discrete alanine-activation domain and an internal proline activation domain involved in leinamycin and tyrocidine biosynthesis, respectively.[18,19] Comparison of the eight amino acid residues lining the substrate-binding pocket of NRPS A domains enables a prediction of substrate specificity.[20,21] Several discrete proline adenylyltransferases have a nearly identical specificity-conferring code: DLLYLALV, and this is also present in the putative proline-activation protein identified from *S. antibioticus* (Figure S1 in the Supporting Information). These observations support that the identified gene encodes a discrete proline adenylyltransferase and contributes to formation of a pyrrolyl moiety.

A genomic fosmid library of *S. antibioticus* was constructed and screened for the indanomycin biosynthetic locus using the cloned fragment of *idmJ*. Sequencing of the cloned insert ends of one identified fosmid, KR1H6, revealed genes encoding a polyketide synthase and a putative 3-hydroxybutyryl-CoA dehydrogenase. Subsequent screening yielded two overlapping fosmids, KR11A1 and KR19A9 (Figure 2). Sequence analysis of the region surrounding *idmJ* revealed an 80 kb region containing 23 contiguous open reading frames (ORFs; Figure 2) and an overall G+C content of 71.9%. Sequence analysis permitted identification of homologues and assignment of a putative function for each of the ORFs (Table 1), including five subunits of a type I modular polyketide synthase. The functional assign-

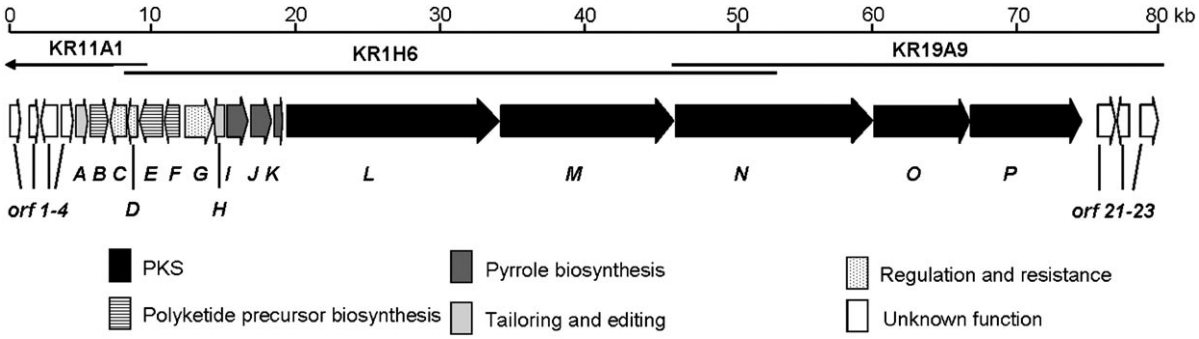


Figure 2. Organization of the indanomycin biosynthetic gene cluster.

Table 1. Deduced functions of open reading frames in the indanomycin biosynthetic gene cluster.

Gene	Size (aa)	Protein homologue,[a] size (aa)	Identity/similarity [%/%]	Proposed function
orf1	254	AAur_0366 (YP_946182), 269	42/65	metallo-β-lactamase superfamily protein
orf2	81	SAV_7559 (NP_828735), 81	92/98	hypothetical protein of unknown function
orf3	429	SGR_5621 (YP_001827133), 483	25/58	putative penicillin-binding protein
orf4	253	Francci3_2578 (YP_481669), 236	27/41	hypothetical protein of unknown function
idmA	252	OlmC (NP_824079), 252	64/78	type II thioesterase
idmB	529	SAV_2703 (NP_823879), 531	80/89	acyl-CoA carboxylase
idmC	516	MonT (AAO65793), 511	47/67	indanomycin efflux transporter
idmD	198	SACE_2167 (YP_001104398), 164	35/55	MarR regulatory protein
idmE	569	PlmT8 (AAAQ84148), 571	49/59	3-hydroxybutyryl-CoA dehydrogenase
idmF	458	CcrA2 (NP_823087), 445	88/95	crotonyl-CoA reductase
idmG	908	Tmn5 (BAE93719), 910	41/58	LuxR-family transcriptional regulator
idmH	145	Glr0227 (NP_923173), 180	35/50	cyclase
idmI	377	Dox7 (ABO15874), 379	59/72	prolyl-*S*-dehydrogenase
idmJ	496	CloN4 (AAN65233), 501	60/70	proline adenyltransferase
idmK	91	CouN5 (AAG29790), 89	42/65	proline carrier protein
idmL	4779	SAMR0465 (CAJ88175), 8154	52/64	PKS modules 1–3
idmM	3596	Orf16 (AAX98191), 7510	55/66	PKS modules 4–5
idmN	5214	HlsF (BAF02926), 5232	52/63	PKS modules 6–8
idmO	2175	Orf16 (AAX98191), 7510	55/68	PKS module 9
idmP	2547	SGR_6078 (YP_001827590), 2411	53/64	PKS modules 10–11
orf21	491	DR_1773 (NP_295496), 433	32/44	hypothetical protein of unknown function
orf22	146	SGR_424 (YP_001821936), 146	89/94	hypothetical protein of unknown function
orf23	133	SCP1.197 (NP_639793), 133	92/94	possible replication initiation protein

[a] The gene accession number is given in parentheses.

ments of the open reading frames are described in the following sections.

Disruption of the proline adenylyltransferase gene *idmJ*

To confirm that the identified gene cluster from *S. antibioticus* is involved in indanomycin biosynthesis, a deletion mutant of *idmJ* was generated using a PCR-targeted disruption method.[22] The deletion mutant, *S. antibioticus* Δ*idmJ*, was confirmed by PCR using primers flanking the targeted region (Figure S2) and by sequence analysis of the amplified product. The fermentation cultures of both mutant and wild-type *S. antibioticus* were evaluated for indanomycin production. Indanomycin was clearly present in the wild-type culture and was absent in that from the *idmJ* mutant. Production was restored by the *in trans* expression of *idmJ* in the deletion mutant (*S. antibioticus* CL1, Figure 3); this ruled out any detrimental polar effects from the mutation. The above results unambiguously establish that

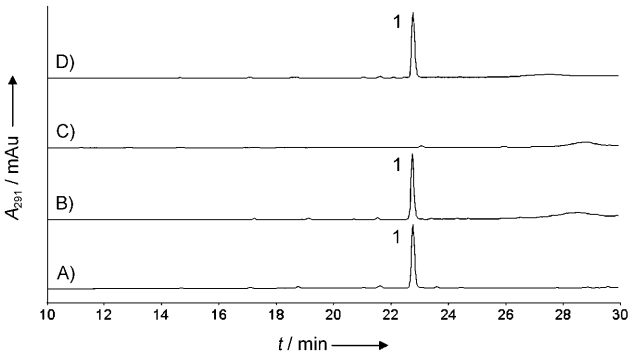


Figure 3. HPLC analysis of A) indanomycin standard and culture extracts from B) *S. antibioticus* NRRL 8167, C) *S. antibioticus* Δ*idmJ*, and D) *S.antibioticus* CL1.

idmJ is essential for indanomycin production and confirms that this locus was correctly assigned as participating in the biosynthesis of indanomycin.

Determination of *idm* gene cluster boundaries by gene inactivation

To establish the boundaries of the *idm* cluster, the regions corresponding to *orf3–orf4* and *orf21*, located upstream of *idmA* and downstream of *idmP*, respectively, were each replaced with an apramycin resistance cassette; this generated the mutants *S. antibioticus* Δ*orf3–4* and *S. antibioticus* Δ*orf21* (Supporting Information). At the upstream boundary of the cluster, *orf3* encodes a protein similar to penicillin-binding proteins, and *orf4* encodes a hypothetical protein with weak similarity to integral membrane proteins (Table 1). The other two genes at the upstream arm of the sequenced region encode a putative metallo-β-lactamase (*orf1*) and a hypothetical protein of unknown function (*orf2*). As for the open reading frames downstream of the proposed cluster, neither *orf21* nor *orf22* are similar to any protein of known function, whereas *orf23* is most similar to a putative replication initiation protein (Table 1). On the basis of sequence analysis alone, *orfs1–4* and *orfs21–23* do not appear to be involved in the production of indanomycin. Neither loss nor decreased production of indanomycin was observed with *S. antibioticus* Δ*orf3–4* or *S. antibioticus* Δ*orf21* (Supporting Information); this confirmed that the disrupted regions are not essential components of the *idm* gene cluster. It therefore appears that the sixteen genes encompassed between *idmA* and *idmP* encode all enzymes necessary for construction of indanomycin.

Pyrrole formation

As described above, three proteins are likely required for the formation of pyrrole-2-carboxylate from L-proline: a prolyl adenylyltransferase, a carrier protein, and a flavin-dependent L-prolyl-*S*-CP dehydrogenase (Figure 1 B). Adjacent on either side of *idmJ* are two genes encoding proteins likely needed for pyrrole generation. IdmI and IdmK both exhibit high identity and similarity to counterparts found for other pyrrole-containing metabolites (Table 1). IdmI, most similar to Dox7 from the pyrrolomycin biosynthetic cluster, is hypothesized to fulfill the role of the dehydrogenase.[23] The prolyl carrier protein encoded by *idmK* is most similar to CloN5 and CouN5, the carrier proteins involved in clorobiocin and coumermycin A1 biosynthesis, respectively.[14] The sequence similarity of IdmI, IdmJ, and IdmK to known enzymes mediating the formation of pyrrole-2-carboxylate strongly argues for their role in generation of indanomycin's pyrroloketo moiety. The pyrrolyl starter unit is likely handed off to the first module of the indanomycin PKS, initiating construction of the indanomycin backbone.

Organization of the indanomycin PKS and polyketide assembly

Five large open reading frames encoding subunits of a modular type I PKS are found immediately downstream of *idmIJK*, the genes ascribed to pyrrole biogenesis (Figures 2 and 4). The PKS subunits are designated as IdmL-P and contain eleven modules (Table 1 and Figure 4). The indanomycin PKS includes the predicted ten extension modules and harbors an unusual terminal module. The *idmL* gene encodes extension modules 1–3, *idmM* encodes the extension modules 4 and 5, *idmN* encodes extension modules 6–8, *idmO* encodes extension module 9, and *idmP* encodes extension module 10 and an incomplete terminal module (module 11; Figure 4).

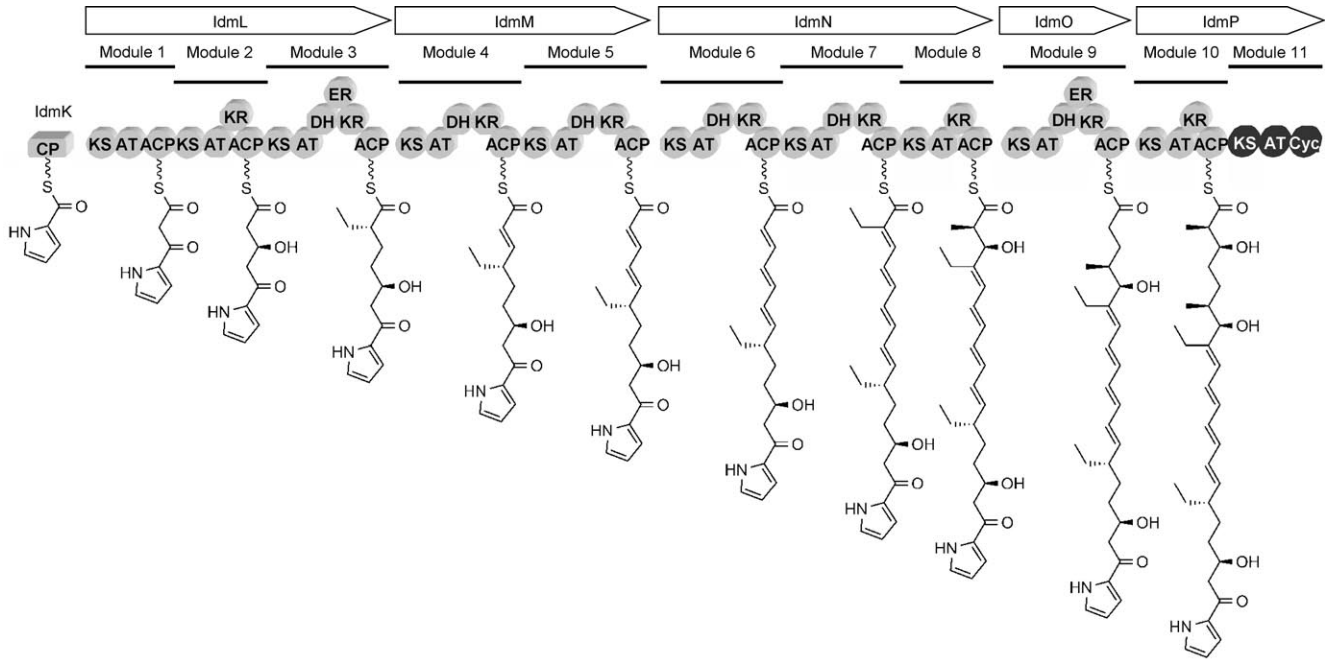


Figure 4. Deduced module and domain organization of the indanomycin PKS encoded by *idmL-idmP*, showing modules 1–11 and the predicted carrier protein-bound intermediates.

All eleven ketosynthase (KS) domains of the indanomycin PKS retain the highly conserved Cys-His-His catalytic triad required for activity.[24] Likewise, all ten identified acyl carrier proteins (ACPs) harbor the conserved serine needed for post-translational attachment of the 4′-phosphopantetheinyl cofactor.[24] The acyltransferase (AT) domain within each module determines the extension unit incorporated into the backbone of the growing polyketide. Substrate specificity of an AT domain can be predicted following examination of its amino acid sequence and by comparison to AT domains with known specificity.[25] The extension substrates predicted for the IdmL-P AT domains are consistent with the structure of the final metabolite (Figure 5). Sequence analysis revealed AT1, 2, 4, 5, 6 and 9

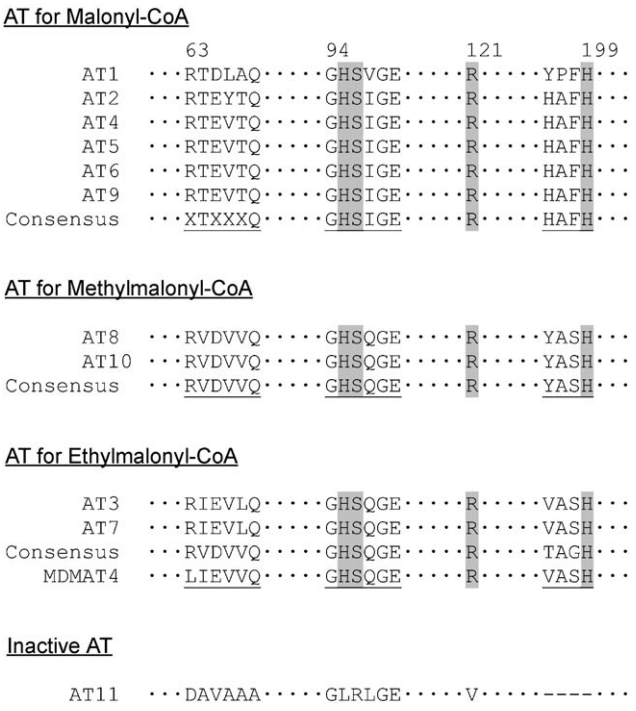


Figure 5. Alignment of the partial amino acid sequences of the AT domains from the indanomycin PKS (AT1-AT11) and the AT4 of meridamycin synthase MerB (MDMAT4). The underlined residues are the three motifs associated with substrate specificity. The four amino acids critical for catalytic activity are highlighted.

all contain the conserved motifs for the selection and loading of malonyl-CoA. The expected motifs for methylmalonyl-CoA specificity are present in AT8 and AT10 (Figure 5). These latter two domains are identical in both amino acid sequences and in the encoding nucleotide sequences, suggesting these domains arose by duplication during evolution of the indanomycin PKS. Although the motifs observed in AT3 and AT7 are not identical to the previously reported consensus motifs for loading ethylmalonyl-CoA,[25] they do agree with the ethylmalonyl-CoA-specific AT domain of the meridamycin PKS subunit, MerB (Figure 5);[26,27] this indicates that AT3 and AT7 also select ethylmalonyl-CoA. The critical active-site residues[24] of AT11 are absent and this domain is truncated. As such, AT11 is anticipated to be inactive as an acyltransferase.

The reductive loop of a PKS module consists of the ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) domains, and the presence of these domains is useful in predicting the structure created by a PKS. The KR domain determines the configuration of the resulting β-hydroxyl group, and ultimately that of the double bond following DH action.[28] Nine KR and six DH domains were found in the indanomycin PKS, and all contain the active site amino acid residues essential to catalytic function.[29] Further examination of the motifs dictating the stereochemical outcome of ketoreduction indicate KR3-7 and KR9 fit within the type B1 group to provide an alcohol intermediate of the *R* configuration and would lead to the formation of *trans*-double bonds.[28] KR2, KR8, and KR10 are predicted to result in the *S* configuration of the newly generated hydroxyl group and fall within the Type A1 family.[28] All the KR and DH products that remain in the final structure of indanomycin match the configurations predicted from this bioinformatic analysis. The two ER domains of the indanomycin PKS, ER3 and ER9, both contain the highly conserved NADP(H) binding motif GXGXXAXXXA, and thus appear to be active.[29]

The final module of IdmP is unusual (Figure 4). KS11 appears to be a catalytically competent domain, whereas AT11, as noted above, is likely not a functional acyltransferase. This module also lacks the obligatory ACP. Conspicuously absent from the indanomycin PKS is the thioesterase (TE) domain that is typically fused to the final module of the assembly line and is required for release of the full-length polyketide.[30] Instead of the typical thioesterase, a highly unusual domain is observed at the C-terminus of the final module in IdmP. This final domain, Cyc11, bears sequence similarity to members of the Δ^5-3-ketosteroid isomerase structural family, including marked similarity to the ring-opening epoxide cyclases implicated in the biosynthesis of the polyether ionophores nigericin and monensin: NigBI, NigBII, MonBI, and MonBII.[31–33] The epoxide-opening cyclases observed in the biosynthetic gene clusters of the polyether ionophores are thought to catalyze a cascade reaction coupling the opening of multiple epoxide rings to the generation of a polyether framework.[31] As discussed below, Cyc11 may be recruited to form the tetrahydropyran ring of indanomycin and participate in liberation of the PKS-bound product.

Neither the canonical integral thioesterase fused to the final PKS module nor any chain-releasing thioesterase candidates are observed within the *idm* cluster. IdmA, however, is closely related to a number of type II, or editing, thioesterases.[34] These editing thioesterases remove aberrant acyl intermediates during assembly of a metabolite's backbone by PKSs and NRPSs.[34–36] If IdmA is indeed utilized in the proofreading of IdmL-P as suggested by its sequence homology, another enzyme or domain dedicated to liberation of the full-length polyketide from IdmP is required, but the identity of this domain is not readily apparent. Therefore, the precise role of IdmA in the context of indanomycin biosynthesis remains unclear until further analysis establishes its biochemical function.

Biosynthesis of the extender units for the PKS

Three genes present in the *idm* cluster presumably assist in supplying the available pool of ethylmalonyl-CoA and methylmalonyl-CoA for efficient production of indanomycin. One of these, *idmB,* encodes the β-subunit for an acyl-CoA carboxylase, and has the greatest sequence similarity to methylmalonyl-CoA decarboxylase from *Streptomyces avermitilis* (Table 1; 80% identity, 89% similarity). Significant similarity is also demonstrated to other acyl-CoA carboxylases, including an acetyl-CoA carboxylase from *Streptomyces antibioticus* ATCC 11891 (CAG14973.1; 86% similarity and 78% identity) and an acetyl-CoA carboxylase from *Streptomyces aureofaciens* (CAG14971.3; 85% similarity and 77% identity). Only the β-subunit of the acyl-CoA complex is present within the indanomycin cluster, and it is possible it forms an active complex with the partnering α and ε subunits encoded elsewhere within *S. antibioticus*.[37,38] The amino acid side chain at position 422 of the *Streptomyces coelicolor* propionyl-CoA carboxylase is largely responsible for selection of the acyl-CoA substrate.[39] Acyl-CoA carboxylases possessing a Cys, Asp, or Ala at position 422 have a greater specificity for propionyl-CoA. The acyl-CoA carboxylases with an Ile or Leu at the position corresponding to 422 can accept, with equal affinity, acetyl-, propionyl-, and butyryl-CoA.[37,39] IdmB, which possesses an Ile at the specificity-determining position, is most likely a more permissive acyl-CoA carboxylase, and may supply both methylmalonyl-CoA and ethylmalonyl-CoA to the indanomycin PKS.

Two genes present in the indanomycin cluster appear to be dedicated to the production of ethylmalonyl-CoA: *idmE* and *idmF*. A crotonyl-CoA reductase/carboxylase, encoded by *idmF* (Table 1), may have the critical role of supplying ethylmalonyl-CoA by the reductive carboxylation of crotonyl-CoA.[40] The gene *idmE* encodes a protein with similarity to numerous proteins annotated as 3-hydroxyacyl-CoA or 3-hydroxybutyryl-CoA dehydrogenases (Table 1). IdmE may recruit acetoacetyl-CoA toward the generation of the crotonyl-CoA substrate for IdmF, however, this would require coordination with a dehydratase encoded elsewhere in the *S. antibioticus* genome. Both a crotonyl-CoA reductase/carboxylase and a 3-hydroxybutyryl-CoA dehydrogenase are observed in the concanamycin and lasalocid biosynthetic gene clusters; this suggests a more widespread strategy to generate a sufficient pool of ethylmalonyl-CoA for polyketide production.[41–43] The indanomycin biosynthetic machinery may, in fact, have the flexibility to adjust for the source of ethylmalonyl-CoA according to the available nutrients. The role of IdmB in permitting the direct utilization of butyryl-CoA is supported by the efficient incorporation of butyrate into indanomycin (~22%).[13] Alternatively, IdmE and IdmF permit siphoning of the cell's crotonyl-CoA pool toward ethylmalonyl-CoA production as needed.

Self resistance, regulation, and polyketide tailoring

Two regulatory genes are found in the *idm* cluster (Table 1). The first of these, *idmD,* encodes a protein that exhibits significant homology to MarR transcriptional regulators. The MarR-like proteins, widely distributed among prokaryotes, usually serve as transcriptional repressors involved in the regulation of genes related to degradation and detoxification of toxic compounds and genes involved in antibiotic resistance.[44,45] Immediately downstream of *idmD* resides *idmC,* encoding a protein similar to the efflux pump MonT (Table 1) that mediates resistance to monensin, another ionophore antibiotic.[46] Therefore, IdmC likely has a comparable role in the resistance or export of indanomycin. Overlap of *idmD* with *idmC,* suggests a bicistronic transcript with *idmD* controlling the expression of *idmC*.

The second regulatory protein, IdmG, appears to be a LuxR-family transcriptional regulator. The LuxR family of regulators are activated upon binding an *N*-acyl homoserine lactone modulator and often regulate aspects of secondary metabolism.[47,48] LuxR regulators are present in multiple antibiotic gene clusters and often appear as positive regulators of antibiotic biosynthesis, yet there are reports of negative regulation.[47–49] A rare TTA codon encodes Leu76 of IdmG, suggesting its expression is controlled by the *bldA* gene product, a $tRNA^{Leu}$ specific for the TTA codon in *Streptomyces*.[50] Further analysis is required to determine the precise roles of IdmG and *bldA* in indanomycin biosynthesis.

The protein encoded by *idmH* is similar to a number of putative proteins (Table 1). A detailed sequence analysis of IdmH did reveal similarity to several proteins annotated as SnoaL-like polyketide cyclases. The SnoaL-like cyclases catalyze an intramolecular aldol reaction to introduce the fourth ring into the anthracycline antitumor antibiotics.[51,52] SnoaL is a member of a structural family with an α+β fold that includes enzymes of diverse catalytic activities, including Δ^5-3-ketosteroid isomerase, scytalone dehydratase, and limonene-1,2-epoxide hydrolases.[53–55] The individual enzymes of this structural family often lack significant sequence similarity to each other.[51] Given these considerations, IdmH may be assigned to the cyclization of either of indanomycin's defining rings, as discussed below.

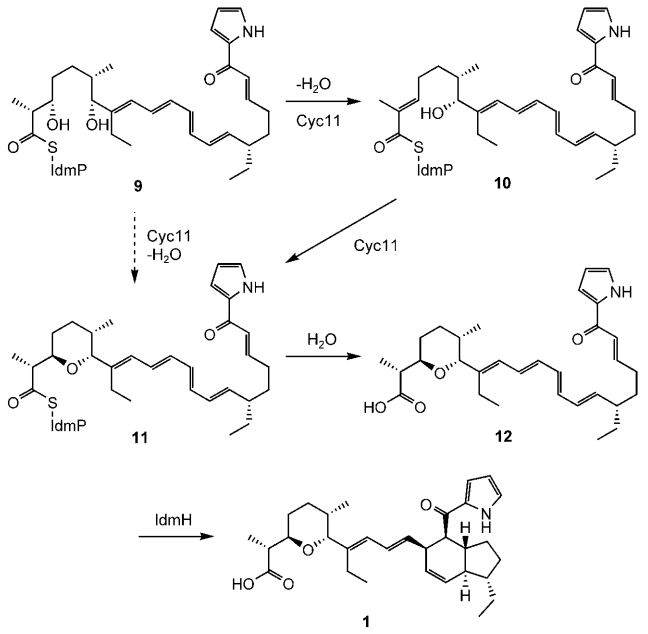
Maturation of the indanomycin PKS product

A small number of post-assembly line modifications are needed to produce the final structure of indanomycin. Inspection of the molecule reveals a lack of post-assembly line oxidations and there is no further decoration of the core scaffold by glycosylation, methylation, or other functional group transfer. In accordance with this observation, the *idm* cluster is relatively minimal in its composition. Maturation of the IdmL-P product does incorporate the construction of two separate structural features: a tetrahydropyran and an indane ring system. Aside from the genes assigned to the megasynthetase assembly line, precursor supply, or regulation, there are few candidates to be considered for co- or post-assembly line processing.

The tetrahydropyran ring could logically arise following the Michael addition of an oxygen nucleophile upon an α,β-unsaturated carbonyl system in a manner similar to that proposed for tetrahydropyran rings found in nigericin and tetronomycin.[33,56] This proposed 6-*exo*-trig cyclization yielding the indanomycin tetrahydropyran is mechanistically similar to the 5-

exo-trig cyclization catalyzed by NonS to provide the nonactin's tetrahydrofuran ring.[57] NonS, however, is similar to enoyl-CoA hydratases, and no such homologue resides in the *idm* gene cluster.[57] The lack of a DH domain in module 10 is inconsistent with a Michael addition to produce the tetrahydropyran, and it is expected to leave an alcohol β to the thioester (at C-3) in the predicted *S* configuration. The resulting cyclic ether is of the *R* configuration at C-3, opening the possibility that a direct nucleophilic substitution leads to the cyclic ether.

The role of the incomplete eleventh module in IdmP, including the KS and the truncated AT, is enigmatic. The KS domains required for the indanomycin backbone are accounted for within modules 1–10, and there is no clear need for an eleventh. The terminal cyclase domain within module 11, as noted earlier, is similar to ring-opening epoxide cyclases of polyether ionophore biosynthesis. This similarity suggests module 11 installs the tetrahydropyran moiety while the full-length polyketide is still tethered to the PKS (Scheme 2). This first requires



Scheme 2. Proposed biosynthetic pathway leading to formation of the tetrahydropyran and indane rings of indanomycin.

dehydration of the C-3 alcohol in **9** to provide an α,β-unsaturated intermediate **10**, presumably by Cyc11 or by the noncolinear action of the upstream module's DH9 domain. Next, Cyc11 could facilitate the nucleophilic attack upon C-3 by the C-7 alcohol to provide the cyclic ether **11** (Scheme 2). Alternatively, Cyc11 could mediate a direct nucleophilic displacement at C-3 in **9** to eliminate water and generate **11**. The lack of a thioesterase integrated adjacent to the final module implies an unusual mechanism for release of the polyketide from the PKS. It is possible that generation of the tetrahydropyran **11** then permits entry of water into the Cyc11 active site followed by hydrolysis. The role of this final module, including the biochemical activities of KS11 and Cyc11 is the subject of an ongoing investigation.

The reasonable mechanisms available to account for introduction of the indane ring raise an intriguing question: could indane ring formation proceed by a Diels–Alder cycloaddition? "Diels–Alderases" have been proposed in the biosynthesis of many metabolites, including indanomycin and related pyrroloketoindanes.[58,59] To date, however, only one enzyme has been isolated for which the activity reported is consistent with a Diels–Alder reaction: lovastatin nonaketide synthase (LNKS or LovB).[58,60] It still remains to be determined how LovB and other putative "Diels–Alderases" catalyze their respective reactions. One possibility is that they simply sequester the substrate and bind it in a conformation enabling a facile [4+2] cycloaddition without direct participation by the enzyme. An alternative strategy, and one observed for the Diels–Alder catalytic antibodies, achieves catalysis both by proximity and activation of the dienophile through hydrogen bonding.[61,62]

The indanomycin PKS lacks a DH domain in the second module, which would ultimately correspond to the double bond between C-19 and C-20. An alkene at this position is necessary to present the dienophile **12** for a proposed [4+2] cycloaddition to produce the indane ring system (Scheme 2). If the alcohol at the position corresponding to C-19 is retained, dehydration is needed either during extension by the PKS or following release from the megasynthetase. One scenario is that the enzymatic activity could result from a neighboring module's DH domain, for example, DH3, to introduce the alkene, by using a tactic similar to that observed during epothilone biosynthesis.[63]

A candidate for the indane cyclase is IdmH (Scheme 2). The sequence similarity of IdmH to SnoaL and Δ^5-3-ketosteroid isomerase indicates that it falls in the same structural family as the Cyc11 domain of IdmP.[53–55] Although IdmH and IdmP-Cyc11 may very well share a common fold, they do not demonstrate any significant amino acid sequence similarity to each other; this suggests distinct evolutionary origins for the two proteins. Additional genetic and biochemical characterization of IdmH and IdmP-Cyc11 is required to definitively assign the functions of the indane cyclase and the tetrahydropyran synthase and to ascertain the timing of these two modifications relative to each other.

Conclusions

At first glance, indanomycin and the related pyrroloketoindanes are a family of metabolites harboring a relatively simple structure, and one might initially expect a relatively straightforward biosynthetic pathway. The identification of the indanomycin biosynthetic gene cluster from *S. antibioticus* described in this work, however, provides an NRPS-PKS assembly line with an unusual terminal module and raises several questions pertaining to release of the PKS-bound product and backbone modifications. The indanomycin biosynthetic system identified here thus presents an ideal platform to examine the mechanisms involved in generation of a tetrahydropyran, an indane

ring system, and a novel mechanism to liberate a metabolite from a PKS.

Experimental Section

Bacterial strains, plasmids and growth medium: Unless specified, common chemicals, restriction enzymes, DNA ligase and other materials were purchased from standard commercial sources and used as provided. *Streptomyces antibioticus* NRRL 8167 (*S. antibioticus*) was obtained from the National Center for Agricultural Utilization Research. Strains and plasmids used in this study are listed in Table S1 in the Supporting Information, and primers used in this study are listed in Table S2. The plasmid pSET1520 was a generous gift of Prof. Kevin A. Reynolds (Portland State University, USA). The plasmid pGM160 was kindly provided by Prof. Günther Muth (University of Tübingen, Germany), and authentic indanomycin was provided by Roche. All *Escherichia coli* (*E. coli*) strains were grown using Luria-Bertani liquid or solid medium with the appropriate antibiotic(s). Kanamycin (50 $\mu g\,mL^{-1}$), apramycin (50 $\mu g\,mL^{-1}$), ampicillin (100 $\mu g\,mL^{-1}$), thiostrepton (20 $\mu g\,mL^{-1}$), nalidixic acid (25 $\mu g\,mL^{-1}$) and chloramphenicol (30 $\mu g\,mL^{-1}$) were used in this research for the selective growth of either *E. coli* or *Streptomyces* strains as appropriate. ISP-3 agar (oatmeal agar) was used for the growth and sporulation of *S. antibioticus*. Media used for genetic manipulation of *S. antibioticus* or production of indanomycin were prepared as described previously.[13,64]

Evaluation of indanomycin production by *S. antibioticus*: Spores from an *S. antibioticus* colony grown on ISP-3 medium were used to inoculate seed medium (20 mL) in an Erlenmeyer flask (250 mL) and the resulting culture was grown at 28 °C and 200 rpm for 48 h. After this time, seed culture (0.5 mL) was used to inoculate an Erlenmeyer flask (250 mL) containing fermentation medium (25 mL). The fermentation culture was grown at 28 °C and 200 rpm for 120 h. The *S. antibioticus* fermentation broth was centrifuged at 23 °C and 3500 x *g* for 10 min. The mycelium was then extracted with methanol (0.5× volume). High performance liquid chromatography (HPLC) analysis was performed using a Beckman–Coulter System Gold® instrument and a Phenomenex Synergi 4 μm MAX-RP column (250×4.6 mm). A gradient was employed from 20% solvent B (acetonitrile) in solvent A (50 mM $NaHCO_3$) to 80% solvent B in solvent A over 30 min at 1 $mL\,min^{-1}$. Absorbance was monitored at 291 nm. Indanomycin eluted with a retention time of ~22.5 min. LC-MS analysis was performed at the Georgia Institute of Technology Bioanalytical Mass Spectrometry facility with a Micromass Quattro LC Mass Spectrometer.

PCR amplification of a discrete proline adenylyltransferase gene: Genomic DNA of *S. antibioticus* was used as the template for the amplification of a discrete proline adenylyltransferase by polymerase chain reaction using a pair of degenerate primers, CL1-F and CL1-R. The PCR reactions (50 μL) contained *S. antibioticus* genomic DNA (100 ng), Pfu (1.25 U), dNTP's (0.8 mM), and the primers CL1-F and CL1-R (each at 0.5 μM). The reactions were performed with the following reaction sequence: an initial denaturation at 96 °C for 4 min, followed by 30 extension cycles (96 °C for 45 s, then 50 °C for 45 s, then 72 °C for 30 s), and a final segment at 72 °C for 10 min. The resulting PCR product was cloned using the Zero Blunt TOPO Cloning Kit (Invitrogen) and analyzed by DNA sequencing.

Cloning and sequencing of the *idm* biosynthetic gene cluster: A genomic fosmid library was constructed using pCC1FOS (Epicenter, Madison, WI, USA) according to protocol. 2049 clones were probed using a digoxigenin-labeled *idmJ* fragment and detected with the DIG Luminescent Detection Kit (Roche Applied Science). The identified fosmid and overlapping fosmids were subcloned using the TOPO Shotgun Subcloning Kit (Invitrogen). Sequencing of the shotgun library was performed by Functional Biosciences, Inc. (Wisconsin, USA). DNA sequences were assembled using ContigExpress of Vector NTI (Invitrogen), and analyzed using Frameplot 4.0beta (http://nocardia.nih.go.jp/fp4/), the BLAST algorithm and NRPS-PKS (http://www.nii.res.in/nrps-pks.html).[65–67]

Disruption of *idmJ*: Disruption of *idmJ* followed the method for PCR-targeted gene replacement (see the Supporting Information).[22] A 1.4 kbp fragment containing the apramycin resistance cassette (*aac*(3)*IV*) and *oriT* was amplified from pIJ773 by PCR using the primers CL2-F and CL2-R. The resulting PCR product was used for the in-frame replacement of the *idmJ* fragment in KR1H6 according to protocol to provide KR1H6d. A 4.2 kbp fragment containing disrupted *idmJ* with flanking 1.4 kbp sequences was amplified by PCR from KR1H6d using the primers CL3-F and CL3-R. The resulting PCR product was cloned using Zero Blunt TOPO Cloning Kit, yielding pCL162 and was confirmed by DNA sequencing. Digestion of pCL162 with HindIII yielded a 4.2 kbp fragment that was ligated into pGM160 to provide pCL161. *S. antibioticus* was grown in 148G medium containing 0.5% glycine at 30 °C, 220 rpm for 20–30 h. *S. antibioticus* protoplasts were prepared and transformed with pCL161 isolated from *E. coli* ET12567 according to protocol.[64] A double crossover colony of the phenotype $apra^R tsr^S$ was selected after forced homologous recombination at 37 °C to provide *S. antibioticus* T1d1. The allelic replacement of *idmJ* by the apramycin resistance cassette and *oriT* in *S. antibioticus* T1d1 was confirmed by PCR amplification and sequencing of the resulting product. A nonpolar, unmarked deletion mutant of *idmJ* was also constructed. FLP-recombinase mediated excision of pCL161 was utilized to remove the disruption cassette,[22] leaving a 81 bp scar sequence and generating pCL160. Protoplasts of *S. antibioticus* T1d1 were prepared and transformed with pCL160 and a double crossover colony of the phenotype $apra^S tsr^S$ was selected as described above to provide *S. antibioticus* Δ*idmJ*. The deletion mutant was confirmed by PCR amplification and sequence analysis of the isolated product.

Genetic complementation of *S. antibioticus* Δ*idmJ*: The *idmJ* gene was amplified using the primers CL5-F and CL5-R. The resulting PCR product was cloned into pCR4Blunt-TOPO, generating pA1, prior to excision with NdeI and EcoRI and subsequent ligation into pSET1520.[68] The resulting pSET1520A was introduced into *S. antibioticus* Δ*idmJ* by protoplast transformation to provide *S. antibioticus* CL1.

Nucleotide sequence accession numbers: The sequences reported here have been deposited into the GenBank database under the accession number FJ545274.

Acknowledgements

This work was supported by the Camille and Henry Dreyfus Foundation and the Georgia Institute of Technology. A GAANN predoctoral fellowship was provided to K.R..

Keywords: biosynthesis · indanomycin · ionophores · polyketides · pyrroloketoindanes

[1] M. Izumikawa, H. Komaki, J. Hashimoto, M. Takagi, K. Shin-ya, *J. Antibiot.* **2008**, *61*, 326–329.
[2] S. H. Larsen, L. D. Boeck, F. P. Mertz, J. W. Paschal, J. L. Occolowitz, *J. Antibiot.* **1988**, *41*, 1170–1177.
[3] C. M. Liu, T. E. Hermann, M. Liu, D. N. Bull, N. J. Palleroni, B. L. Prosser, J. W. Westley, P. A. Miller, *J. Antibiot.* **1979**, *32*, 95–99.
[4] S. Miao, M. R. Anstee, V. Bauchwal, A. Park, *Tetrahedron Lett.* **1995**, *36*, 5699–5702.
[5] N. V. Murenets, M. R. P. Kudinova, T. N. Korobkova, T. N. Brobysheva, N. A. Klyuev, I. V. Yartseva, T. A. Ivanova, L. P. Ivanitskaya, *Antibiot. Med. Biotekhnol.* **1987**, *32*, 811–814.
[6] P. Toth, V. Széll, I. Koczka, G. Horváth, I. M. Szabó, M. Szayly, G. Ambrus, J. Bedo, J. Berdy, A. Jekkel, N. V. Makk, *Hung. Teljes*, HU 49 909 (A2) **1989** [Chem. Abstr. **1990**, 113, 130 720].
[7] J. W. Westley, R. H. Evans, Jr., L. H. Sello, N. Troupe, C. M. Liu, J. F. Blount, *J. Antibiot.* **1979**, *32*, 100–107.
[8] C. Gumila, M. L. Ancelin, G. Jeminet, A. M. Delort, G. Miquel, H. J. Vial, *Antimicrob. Agents Chemother.* **1996**, *40*, 602–608.
[9] D. Zhang, M. G. Nair, M. Murry, Z. Zhang, *J. Antibiot.* **1997**, *50*, 617–620.
[10] J. W. Westley, *Adv. Appl. Microbiol.* **1977**, *22*, 177–223.
[11] A. Qureshi, C. S. Stevenson, C. L. Albert, R. S. Jacobs, D. J. Faulkner, *J. Nat. Prod.* **1999**, *62*, 1205–1207.
[12] N. U. Sata, N. Fusetani, *Tetrahedron Lett.* **2000**, *41*, 489–492.
[13] K. E. Roege, W. L. Kelly, *Org. Lett.* **2009**, *11*, 297–300.
[14] F. Pojer, S.-M. Li, L. Heide, *Microbiology* **2002**, *148*, 3901–3911.
[15] M. G. Thomas, M. D. Burkart, C. T. Walsh, *Chem. Biol.* **2002**, *9*, 171–184.
[16] H. Xu, Z. X. Wang, J. Schmidt, L. Heide, S. M. Li, *Mol. Genet. Genomics* **2002**, *268*, 387–396.
[17] M. A. Marahiel, T. Stachelhaus, H. D. Mootz, *Chem. Rev.* **1997**, *97*, 2651–2673.
[18] Y. Q. Cheng, G. L. Tang, B. Shen, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3149–3154.
[19] H. D. Mootz, M. A. Marahiel, *J. Bacteriol.* **1997**, *179*, 6843–6850.
[20] G. L. Challis, J. Ravel, C. A. Townsend, *Chem. Biol.* **2000**, *7*, 211–224.
[21] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, 493–505.
[22] B. Gust, G. L. Challis, K. Fowler, T. Kieser, K. F. Chater, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1541–1546.
[23] X. Zhang, R. J. Parry, *Antimicrob. Agents Chemother.* **2007**, *51*, 946–957.
[24] S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson, L. Katz, *Science* **1991**, *252*, 675–679.
[25] C. D. Reeves, S. Murli, G. W. Ashley, M. Piagentini, C. R. Hutchinson, R. McDaniel, *Biochemistry* **2001**, *40*, 15464–15470.
[26] W. He, J. Lei, Y. Liu, Y. Wang, *Arch. Microbiol.* **2008**, *189*, 501–510.
[27] Y. Sun, H. Hong, M. Samboskyy, T. Mironenko, P. F. Leadlay, S. F. Haydock, *Microbiology* **2006**, *152*, 3507–3515.
[28] A. T. Keatinge-Clay, *Chem. Biol.* **2007**, *14*, 898–908.
[29] S. Donadio, L. Katz, *Gene* **1992**, *111*, 51–60.
[30] J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **2001**, *18*, 380–416.
[31] A. R. Gallimore, C. B. Stark, A. Bhatt, B. M. Harvey, Y. Demydchuk, V. Bolanos-Garcia, D. J. Fowler, J. Staunton, P. F. Leadlay, J. B. Spencer, *Chem. Biol.* **2006**, *13*, 453–460.
[32] N. C. Ha, M. S. Kim, W. Lee, K. Y. Choi, B. H. Oh, *J. Biol. Chem.* **2000**, *275*, 41100–41106.
[33] B. M. Harvey, T. Mironenko, Y. Sun, H. Hong, Z. Deng, P. F. Leadlay, K. J. Weissman, S. F. Haydock, *Chem. Biol.* **2007**, *14*, 703–714.
[34] M. Kotowska, K. Pawlik, A. R. Butler, E. Cundliffe, E. Takano, K. Kuczek, *Microbiology* **2002**, *148*, 1777–1783.
[35] B. S. Kim, T. A. Cropp, B. J. Beck, D. H. Sherman, K. A. Reynolds, *J. Biol. Chem.* **2002**, *277*, 48028–48034.
[36] D. Schwarzer, H. D. Mootz, U. Linne, M. A. Marahiel, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14083–14088.
[37] L. Diacovich, S. Peiru, D. Kurth, E. Rodriguez, F. Podesta, C. Khosla, H. Gramajo, *J. Biol. Chem.* **2002**, *277*, 31228–31236.
[38] E. Rodriguez, C. Banchio, L. Diacovich, M. J. Bibb, H. Gramajo, *Appl. Environ. Microbiol.* **2001**, *67*, 4166–4176.
[39] L. Diacovich, D. L. Mitchell, H. Pham, G. Gago, M. M. Melgar, C. Khosla, H. Gramajo, S. C. Tsai, *Biochemistry* **2004**, *43*, 14027–14036.
[40] T. J. Erb, I. A. Berg, V. Brecht, M. Muller, G. Fuchs, B. E. Alber, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10631–10636.
[41] Y. A. Chan, A. M. Podevels, B. M. Kevany, M. G. Thomas, *Nat. Prod. Rep.* **2009**, *26*, 90–114.
[42] S. F. Haydock, A. N. Appleyard, T. Mironenko, J. Lester, N. Scott, P. F. Leadlay, *Microbiology* **2005**, *151*, 3161–3169.
[43] L. Smith, H. Hong, J. B. Spencer, P. F. Leadlay, *ChemBioChem* **2008**, *9*, 2967–2975.
[44] G. Fiorentino, R. Ronca, R. Cannio, M. Rossi, S. Bartolucci, *J. Bacteriol.* **2007**, *189*, 7351–7360.
[45] A. S. Seoane, S. B. Levy, *J. Bacteriol.* **1995**, *177*, 3414–3419.
[46] M. Oliynyk, C. B. W. Stark, A. Bhatt, M. A. Jones, Z. A. Hughes-Thomas, C. Wilkinson, Z. Oliynyk, Y. Demydchuk, J. Staunton, P. F. Leadlay, *Mol. Microbiol.* **2003**, *49*, 1179–1190.
[47] Y.-A. Hur, S.-S. Choi, D. H. Sherman, E.-S. Kim, *Microbiology* **2008**, *154*, 2912–2919.
[48] S. McGowan, M. Sebaihia, S. Jones, B. Yu, N. Bainton, P. F. Chan, B. Bycroft, G. S. Stewart, P. Williams, G. P. Salmond, *Microbiology* **1995**, *141*, 541–550.
[49] Y. T. Horng, S. C. Deng, M. Daykin, P. C. Soo, J. R. Wei, K. T. Luh, S. W. Ho, S. Swift, H. C. Lai, P. Williams, *Mol. Microbiol.* **2002**, *45*, 1655–1671.
[50] B. K. Leskiw, E. J. Lawlor, J. M. Fernandez-Abalos, K. F. Chater, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2461–2465.
[51] P. Kallio, A. Sultana, J. Niemi, P. Mäntsälä, G. Schneider, *J. Mol. Biol.* **2006**, *357*, 210–220.
[52] A. Sultana, P. Kallio, A. Jansson, J. S. Wang, J. Niemi, P. Mantsala, G. Schneider, *EMBO J.* **2004**, *23*, 1911–1921.
[53] M. Arand, B. M. Hallberg, J. Zou, T. Bergfors, F. Oesch, M. J. van der Werf, J. A. de Bont, T. A. Jones, S. L. Mowbray, *EMBO J.* **2003**, *22*, 2583–2592.
[54] S. W. Kim, S. S. Cha, H. S. Cho, J. S. Kim, N. C. Ha, M. J. Cho, S. Joo, K. K. Kim, K. Y. Choi, B. H. Oh, *Biochemistry* **1997**, *36*, 14030–14036.
[55] T. Lundqvist, J. Rice, C. N. Hodge, G. S. Basarab, J. Pierce, Y. Lindqvist, *Structure* **1994**, *2*, 937–944.
[56] Y. Demydchuk, Y. Sun, H. Hong, J. Staunton, J. B. Spencer, P. F. Leadlay, *ChemBioChem* **2008**, *9*, 1136–1145.
[57] A. J. Woo, W. R. Strohl, N. D. Priestley, *Antimicrob. Agents Chemother.* **1999**, *43*, 1662–1668.
[58] W. L. Kelly, *Org. Biomol. Chem.* **2008**, *6*, 4483–4493.
[59] E. M. Stocking, R. M. Williams, *Angew. Chem.* **2003**, *115*, 3186–3223; *Angew. Chem. Int. Ed.* **2003**, *42*, 3078–3115.
[60] K. Auclair, A. Sutherland, J. Kennedy, D. J. Witter, J. P. Van den Heever, C. R. Hutchinson, J. C. Vederas, *J. Am. Chem. Soc.* **2000**, *122*, 11519–11520.
[61] F. E. Romesberg, B. Spiller, P. G. Schultz, R. C. Stevens, *Science* **1998**, *279*, 1929–1933.
[62] J. Xu, Q. Deng, J. Chen, K. N. Houk, J. Bartek, D. Hilvert, I. A. Wilson, *Science* **1999**, *286*, 2345–2348.
[63] L. Tang, S. Ward, L. Chung, J. R. Carney, Y. Li, R. Reid, L. Katz, *J. Am. Chem. Soc.* **2004**, *126*, 46–47.
[64] T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood, *Practical Streptomyces Genetics*, John Innes Foundation, Norwich, **2000**.
[65] S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, *Nucleic Acids Res.* **1997**, *25*, 3389–3402.
[66] M. Z. Ansari, G. Yadav, R. S. Gokhale, D. Mohanty, *Nucleic Acids Res.* **2004**, *32*, W405–413.
[67] J. Ishikawa, K. Hotta, *FEMS Microbiol. Lett.* **1999**, *174*, 251–253.
[68] N. Palaniappan, M. M. Alhamadsheh, K. A. Reynolds, *J. Am. Chem. Soc.* **2008**, *130*, 12236–12237.

Received: December 15, 2008
Published online on March 19, 2009

