

Identification and Characterization of the Niddamycin Polyketide Synthase Genes from *Streptomyces caelestis*

STEPHAN J. KAKAVAS, LEONARD KATZ, AND DIANE STASSI*

Antibacterial Discovery Division, Abbott Laboratories, Abbott Park, Illinois 60064

Received 29 August 1997/Accepted 26 September 1997

The genes encoding the polyketide synthase (PKS) portion of the niddamycin biosynthetic pathway were isolated from a library of *Streptomyces caelestis* NRRL-2821 chromosomal DNA. Analysis of 40 kb of DNA revealed the presence of five large open reading frames (ORFs) encoding the seven modular sets of enzymatic activities required for the synthesis of a 16-membered lactone ring. The enzymatic motifs identified within each module were consistent with those predicted from the structure of niddamycin. Disruption of the second ORF of the PKS coding region eliminated niddamycin production, demonstrating that the cloned genes are involved in the biosynthesis of this compound.

Niddamycin is a macrolide antibiotic which is able to bind 50S ribosomal subunits to inhibit protein synthesis. The compound was first discovered as a secondary metabolite of *Streptomyces djakartensis* (16) and was later found to be produced by *Streptomyces caelestis* NRRL-2821 (11a). The structure of niddamycin (Fig. 1) suggests that the polyketide backbone of the macrolide ring is formed through the ordered condensation of carboxylic acid residues derived from acetate, propionate, butyrate, and perhaps glycolate (24). The disaccharide, mycaminoise-isobutyrylmycarose, is attached to the macrolide ring at C-5.

Macrolides belong to a class of molecules referred to as complex polyketides, which are synthesized on large, multifunctional enzymes called polyketide synthases (PKSs). The synthesis of polyketides is mechanistically similar to that of fatty acids; however, a greater variety of starter and extender carboxylic acid residues are incorporated into the growing polyketide chain, and the β -keto groups formed after each condensation step undergo various degrees of reduction (15, 20).

PKSs, in general, contain all of the enzymatic activities necessary for the sequential condensation of acyl thioesters (β -ketoacyl acyl carrier protein synthases [KS]), acyltransferases [AT], and acyl carrier proteins [ACP], the subsequent reduction of the β -keto groups (dehydratases [DH]), enoylreductases [ER], and ketoreductases [KR]), and the release of the completed chain from the PKS (thioesterases [TE]). Analysis of the erythromycin PKS genes revealed that these enzymatic domains are organized into modules, each of which is responsible for one round of condensation and reduction (5, 7, 9, 10). As a result, there is a direct correlation between the number of modules contained within the erythromycin PKS and the length of the polyketide chain. In addition, the genetic order of the erythromycin PKS modules was found to be colinear with the order of biochemical reactions, allowing directed genetic alterations which produce predicted novel erythromycin derivatives (9, 11).

The polyketide portion of the 16-membered macrolide niddamycin is predicted to be synthesized by a complex (type 1) PKS (15) comprising seven modules, each catalyzing one con-

densation reaction. It had previously been suggested that the choice of the extender coenzyme A (CoA)-thioester is determined by the AT domain contained in each module (9). Sequence comparisons of AT domains for fatty acid and polyketide synthases have revealed specific sequence motifs for malonyl- and methylmalonyl-ATs (m- and mmATs) (13). The chemical structure of niddamycin (Fig. 1) suggests that the PKS should contain ATs specific not only for malonyl-CoA (modules 1, 2, 3, and 7) and methylmalonyl-CoA (module 4) but also for the more rarely encountered ethylmalonyl-CoA (module 5) and for an as yet undetermined CoA derivative which results in the insertion of hydroxymalonate into the growing chain (module 6). In addition, the structure predicts the enzymatic motifs which should be present in each module to give the corresponding reduction state of the β -carbonyl group formed after each condensation.

This study describes the isolation and characterization of the PKS genes responsible for niddamycin biosynthesis, including those AT sequences which may result in the introduction of unusual side chains to the macrolide ring.

MATERIALS AND METHODS

Media and reagents. SeaKem-agarose was obtained from FMC BioProducts, Rockland, Maine. Bacto Soytone, soluble starch, yeast extract, Bacto Agar, and Antibiotic Medium 11 (AM11) were obtained from Difco Laboratories, Detroit, Mich. SGGP medium (34) and AS-1 medium (4) have been described previously. SCM medium contains (per liter) the following: Bacto Soytone, 20 g; soluble starch, 15 g; morpholinepropanesulfonic acid 10.5 g; yeast extract, 1.5 g; and CaCl_2 , 0.1 g. R3M plates consist of (per liter) the following: sucrose, 103 g; K_2SO_4 , 0.25 g; yeast extract, 4 g; Casamino Acids, 4 g; tryptone, 4 g; agar, 22 g; and H_2O , 830 ml. After sterilization, 20 ml of a 2.5 M solution of MgCl_2 , 20 ml of a 50% glucose solution, 20 ml of a 2.5 M solution of CaCl_2 , 12.5 ml of a 2 M solution of Tris-HCl (pH 7.0), 0.2 ml of a 5,000 \times concentrated trace element solution (14), 2.5 ml of a 1 M solution of NaOH, and 0.37 ml of a 0.5 M solution of KH_2PO_4 were added.

Strains, bacteriophage, and plasmids. *Escherichia coli* DH5 α cells and *E. coli* XL1-Blue MR cells were purchased from BRL/Life Technologies, Gaithersburg, Md., and Stratagene Cloning Systems, La Jolla, Calif., respectively. *E. coli* S17-1 was obtained from Philippe Mazodier, Institut Pasteur, Paris, France. Bacteriophage T7 was obtained from Phil Youngman, University of Georgia, Athens. Plasmid pGEM-3Zf was purchased from Promega Corporation, Madison, Wis. Plasmids pUC19 (35), pKC1139 (6), and pNJ1 (31) have been described previously.

Nucleic acid isolation and manipulation. Plasmid isolation from *E. coli* was performed by using a Qiagen Midi kit and Qiaprep Spin Plasmid kit (Qiagen GmbH, Hilden, Germany). DNA cloning and manipulations were performed by using standard procedures (27). Restriction enzymes, T4 DNA ligase, calf intestinal alkaline phosphatase (CIAP), and the Klenow fragment of DNA polymerase I were purchased from BRL/Life Technologies. Hybond-N membranes were obtained from Amersham Corporation, Arlington Heights, Ill.

* Corresponding author. Mailing address: Abbott Laboratories, D-47P AP-9A, 100 Abbott Park Road, Abbott Park, IL 60064. Phone: (847) 938-7389. Fax: (847) 938-3403. E-mail: stassid@ugene1.pprd.abbott.com.

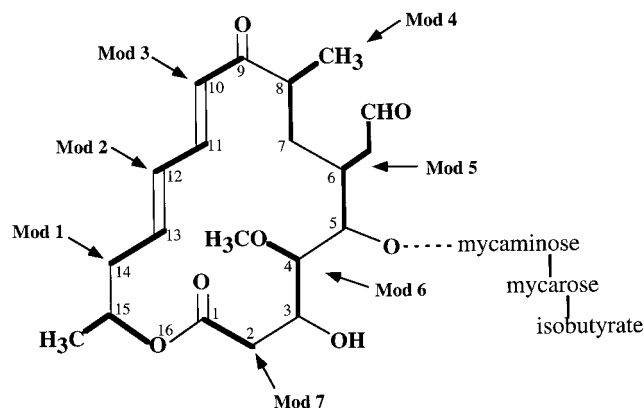


FIG. 1. Structure of niddamycin. The atoms of the macrolide ring are numbered. The thick lines indicate carbon atoms contributed by the indicated modules (Mod).

Construction of a *Streptomyces caelestis* genomic library. *S. caelestis* DNA was prepared by partially digesting 5 μ g of DNA with *Sau*III and size selecting fragments of around 40 kb by electrophoresis through a 0.4% agarose gel. Cosmid pNJ1 arms were prepared by digestion of the vector with *Eco*RI, dephosphorylation with CIAP, and then digestion with *Bgl*II to generate one arm and digestion of the cosmid with *Hind*III, dephosphorylation with CIAP, and then digestion with *Bgl*II to generate the other. Ligation of 1 μ g of cosmid arms to 1 μ g of the size-selected *S. caelestis* DNA was in a 20- μ l volume. Two microliters of this ligation mix was packaged by using GigapackII XL (Stratagene) as instructed by the manufacturer, and *E. coli* XL1-Blue MR cells were hosts for infection. Individual colonies were transferred to 30 96-well plates (Costar, Cambridge, Mass.).

***S. caelestis* PKS probe.** A PKS-specific probe was generated by PCR amplification of *S. caelestis* genomic DNA, with degenerate primers designed from KS and AT sequences in the GenBank/EMBL database. The KS-specific oligonucleotide (5'-CGGTSAAGTCSAACATCGG-3') and the AT-specific oligonucleotide (5'-GCRATCTCRCCCTGCGARTG-3') were used in a PCR mixture containing ThermoPol reaction buffer (New England Biolabs, Beverly, Mass.), 0.2 mM deoxynucleoside triphosphate mixture, 0.5 μ g of genomic DNA, and 100 pmol of the primers in a volume of 98 μ l. After incubation at 96°C for 2 min, 2 U of Vent DNA polymerase (New England Biolabs) was added, and the mixture was cycled 25 times in a Perkin-Elmer Cetus 9600 thermocycler at 96°C for 30 s, 50°C for 2 min, and 72°C for 4 min, followed by a final incubation at 72°C for 15 min. The reaction generated a 900-bp DNA fragment which was gel purified from a 1% agarose gel by using a Prep-A-Gene kit (Bio-Rad Laboratories, Hercules, Calif.).

Hybridization of the PKS probe. The *S. caelestis* library was transferred from 96-well plates to Hybond-N membranes for hybridization. The membranes were soaked in prehybridization solution (27) for at least 2 h at 65°C. About 50 ng of the KS/AT probe was labeled with [α -³²P]dCTP (DuPont, NEN Research Products), using the Megaprime DNA labeling system (Amersham) as instructed by the manufacturer. The radiolabeled probe was added to the prehybridization solution and incubated at 65°C for 16 to 20 h. The membranes were then washed twice in 1 \times SSC (1 \times SSC is 0.15 M NaCl plus 0.015 M sodium citrate)-0.1% sodium dodecyl sulfate at room temperature and once in 0.1 \times SSC-0.1% sodium dodecyl sulfate at 65°C for 20 min. The membranes were dried and placed on 35-by 43-cm phosphor screens (Molecular Dynamics, Sunnyvale, Calif.) for 2 to 4 h. Imaging of the screens was on a Molecular Dynamics PhosphorImager 425.

DNA sequencing. Cosmids that hybridized to the KS/AT probe were digested with *Sst*I, and the resulting fragments were subcloned into vector pGEM-3Zf or pUC19. *Sst*I fragments that were larger than 8 kb were further digested with *Sma*I and subcloned to generate plasmids that would be more suitable for sequencing. Two methods were used for DNA sequencing. One method used [α -³²P]dCTP (DuPont, NEN Research Products) and the fmole DNA cycle sequencing system (Promega). Samples were run on 6% polyacrylamide-8 M urea gels, using a Sequi-Gen II sequencing apparatus (Bio-Rad). The second method used an ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction kit (Perkin-Elmer Corporation), and the samples were run on 4.75% polyacrylamide-8.3 M urea gels on an Applied Biosystems 373 sequencer. All subclones were sequenced on both strands by using multicloning site primers to initiate the sequences and internal primers to extend the sequences.

DNA sequence data was analyzed by using the Wisconsin sequence analysis package programs (Genetics Computer Group, Madison, Wis.) (8). Database searches to identify homologs to the deduced amino acid sequences were performed with BLAST (1).

Conjugation. *S. caelestis* recipient cells were prepared by inoculating 30 ml of SGGP with 25 μ l of spores and incubating the culture overnight at 30°C on a rotary shaker. The culture was then centrifuged at a relative centrifugal force of 2,190 for 15 min. The pelleted cells were washed once in SGGP, recentrifuged, and resuspended in 2 ml of SGGP. *E. coli* donor cells, S17-1 and S17-1(pSK7595), were prepared by growing overnight cultures in 2 ml of LB (with apramycin [50 μ g/ml] where appropriate) at 37°C on a rotary shaker. Two hundred microliters of the *E. coli* cultures was inoculated into 30 ml of fresh medium and grown for about 3 h (to mid-log phase) at 37°C. The cells were pelleted by centrifugation for 20 min and washed once in the same medium. After a final centrifugation, the pellets were resuspended in 300 μ l of LB. Nucleopore membranes (25 mm, 0.2 μ m; VWR) were placed onto an AS-1 agar plate and spotted with 50 μ l each of donor and recipient cells. The plates were incubated at 30°C for 3 h, and then each membrane was placed into a 50-ml conical tube containing 5 ml of SGGP and vortexed for 30 s to dislodge the cells. The cells were centrifuged for 20 min, and the pellets were resuspended in 100 μ l of SGGP. Cells were then plated onto AS-1 plates and incubated overnight at 30°C. The next day, the plates were overlaid with 2 ml of distilled H₂O containing 30 μ l of a 50-mg/ml solution of apramycin and 150 μ l of T7 bacteriophage ($\approx 10^{10}$ PFU) and returned to 30°C until transconjugants appeared (about 7 days).

Metabolite analysis. Culture supernatants of cells grown for 3 days at 30°C in SCM medium were adjusted to pH 9.0 with NH₄OH and extracted two times with equal volumes of ethyl acetate. The organic phases were pooled and concentrated. Extracts were evaluated on silica gel thin-layer chromatography (TLC) plates (60F-254; Merck), which were developed with isopropyl ether-methanol-NH₄OH (75:35:2). Compounds were visualized by spraying the plates with anisaldehyde-sulfuric acid-ethanol (1:1:9) and heating the plates with a hot air gun until color developed. Biological activity was assessed by placing unstained TLC plates onto AM11 agar plates seeded with *Staphylococcus aureus* as the indicator strain and incubating the plates overnight at 37°C to develop zones of inhibition.

Nucleotide sequence accession number. The GenBank accession number for the niddamycin PKS and portions of two flanking genes is AF016585.

RESULTS

Cloning and sequencing of the niddamycin PKS cluster in *S. caelestis*. A strategy was devised to isolate the PKS genes of the niddamycin biosynthetic pathway by using sequence conservation in the KS and AT regions encoded by previously sequenced PKS genes (5, 10). A pair of degenerate primers spanning conserved regions was designed from sequences deposited in GenBank (see Materials and Methods) with the expectation of amplifying, from *S. caelestis* chromosomal DNA, most or all of the genes for the seven KS/AT regions predicted to be found in the niddamycin PKS. The product of the PCR ran as a 900-bp fragment on an agarose gel (data not shown) and was subsequently cloned into M13 for sequencing. Seven clones were analyzed, and all contained the same sequence, which was later found to correspond to the KS/AT region in module 5 of the niddamycin cluster (see below). This PCR product was used to probe chromosomal DNA to determine whether it had sufficient homology to hybridize to the remaining KS/AT domains of the niddamycin biosynthetic cluster and therefore could be used to identify cosmids in a library of *S. caelestis* DNA which encoded the remainder of the pathway. When DNA from *S. caelestis* was digested with *Sst*I and probed with the PCR fragment, at least seven hybridizing bands were visible (data not shown). A second Southern hybridization, but with a 1.45-kb *Sst*I-*Msc*I fragment (coordinates 10793 to 12244; GenBank accession number M63677) encoding part of the KS domain from module 5 of the erythromycin PKS, gave the same banding pattern, suggesting that these bands may represent all of the type 1 PKS genes present in *S. caelestis*.

To construct the *S. caelestis* library, a partial *Sau*III digest of genomic DNA which had been size selected for 30-kb fragments was ligated to cosmid pNJ1 arms and propagated in *E. coli* XL-1 Blue MR cells. The library of 2,880 clones was screened by colony blotting with the *S. caelestis* KS/AT probe described above. Nineteen clones gave a strong hybridization signal, and restriction digests of the positive clones were ana-

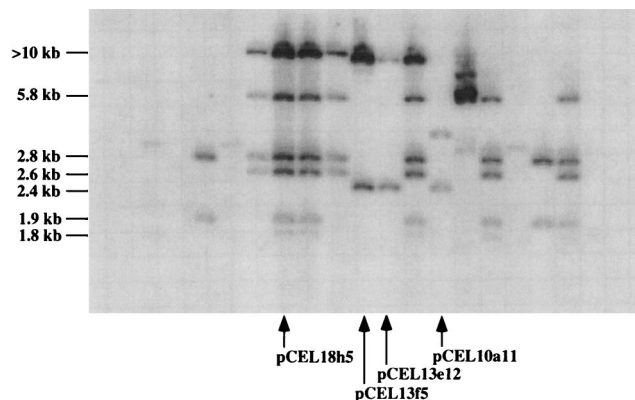


FIG. 2. Southern analysis of positively hybridizing cosmids. Nineteen cosmids that hybridized with the KS/AT probe were digested with *Sst*I and analyzed by Southern hybridization using a KS/AT probe generated from *S. caelestis* DNA. Three of the cosmids did not cross-react (i.e., were false positives). Lanes containing cosmid DNA specifically referred to in the text are labeled.

lyzed by Southern hybridization. Banding patterns of *Sst*I-digested cosmid DNAs probed with the same DNA indicated that 16 of the 19 clones which were positive in colony blots were also positive by Southern analysis and that 15 of these clones contained overlapping inserts (Fig. 2). Three of the clones contained six hybridizing bands, while the other 12 overlapping clones contained from two to five of the same hybridizing bands. Clone pCEL18h5, with six hybridizing bands, was chosen for sequence analysis. Two other clones, pCEL13f5 and pCEL13e12, not only shared a large hybridizing fragment (>10 kb) with pCEL18h5 but also contained a unique 2.4-kb fragment, indicating that these cosmids could contain PKS sequences not present in pCEL18h5. Another related clone, pCEL10a11, contained the 2.4-kb fragment but not the larger fragment, and so it was not analyzed further. The

large *Sst*I fragment was isolated from clones pCEL13f5 and pCEL18h5 and further digested with *Sma*I for Southern analysis. Probing indicated that both clones contained hybridizing bands of 2.1, 2.2, and 3.2 kb, but clone pCEL13f5 also contained a 1.9-kb hybridizing band not found in clone pCEL18h5. This was further evidence that clone pCEL13f5 shared DNA regions with clone pCEL18h5 but also contained neighboring PKS sequences. For this reason, clone pCEL13f5 was also analyzed along with clone pCEL18h5.

Restriction fragments from cosmids pCEL18h5 and pCEL13f5 (Fig. 3) were subcloned into pUC19 or pGEM-3Zf and sequenced as described in Materials and Methods. Orientation of the subcloned fragments was determined by using the parent cosmid as the template for primers that annealed to the 5' and 3' ends of each insert, generating upstream and downstream sequences. These could then be aligned with sequences of the individual subclones to order them within the cosmid. Each cosmid was found to contain approximately 30 kb of insert DNA, with around 10 kb of overlapping sequence. In all, 41.1 kb of DNA was sequenced.

Organization of the niddamycin gene cluster. DNA sequence data obtained from cosmids pCEL18h5 and pCEL13f5 were analyzed for open reading frames (ORFs) by using a *Streptomyces* codon usage table (33). Five large ORFs (ORFs 1 to 5) spanning 40,012 bp of DNA were revealed (see below). These large ORFs were bounded by smaller, partially sequenced ORFs (ORFs 6 and 7) located at the 5' and 3' ends of ORFs 1 to 5, respectively (Fig. 3). The deduced N-terminal 132-amino-acid sequence encoded by ORF 6 was found, by BLAST search, to have 61% identity to the N-terminal region of the product of the *tylI* gene, which encodes a cytochrome P450 hydroxylase in the biosynthetic pathway of the 16-membered macrolide tylosin (19). Like the gene containing ORF 6, *tylI* in *Streptomyces fradiae* is also located immediately 5' to, and reads in the opposite direction from, the genes encoding the PKS for tylosin production. The deduced N-terminal sequence encoded by ORF 7 (157 amino acids) was found to

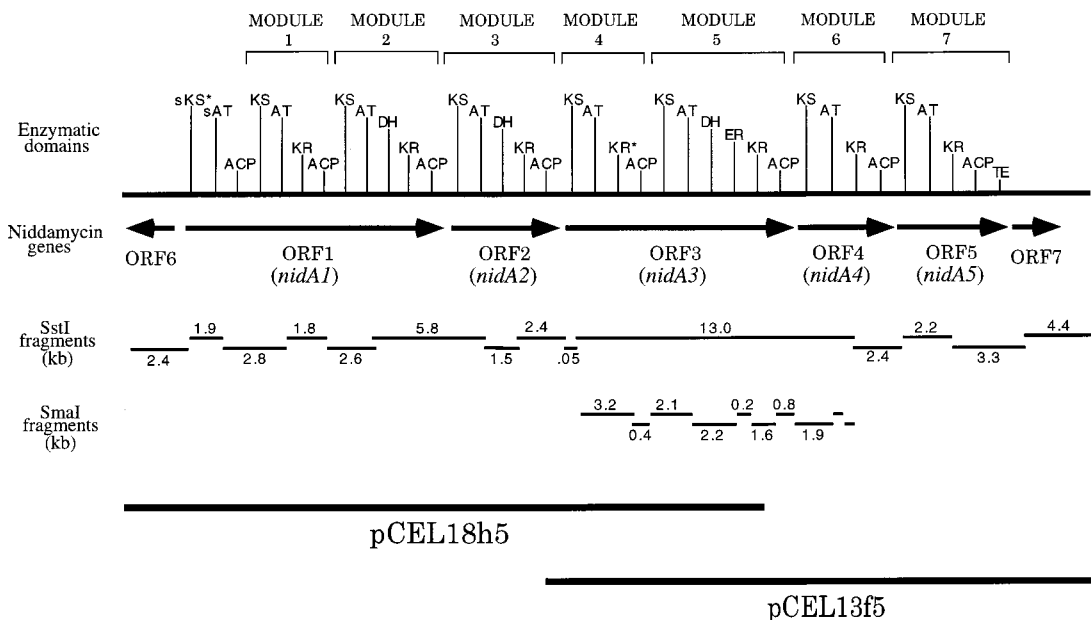


FIG. 3. Organization of the niddamycin PKS genes. Cosmids pCEL18h5 and pCEL13f5 were isolated from a genomic library of *S. caelestis* DNA. DNA sequence was obtained from subcloned restriction fragments indicated by narrow lines and from cosmid DNA to link the restriction fragments. ORFs, enzymatic domains, and modules were identified from DNA sequence analysis. Domains labeled with asterisks are probably inactive.

TABLE 1. Niddamycin PKS sequence coordinates

ORF no.	Nucleotide coordinates	No. of amino acids	Homology
6	1–396	132	P450 (N-terminal portion)
1	627–13646	4340	PKS (loading domain, modules 1 and 2)
	690–1904	405	KS* (inactive)
	2235–3242	336	ATs
	3312–3572	87	ACPs
	3645–4916	424	KS
	5214–6215	334	AT
	6900–7442	181	KR
	7695–7955	87	ACP
	8025–9296	424	KS
	9594–10595	334	AT
	10650–11225	192	DH
	12219–12770	184	KR
	13068–13325	86	ACP
2	13693–19209	1839	PKS (module 3)
	13795–15066	424	KS
	15379–16380	334	AT
	16423–16992	190	DH
	17839–18393	185	KR
	18676–18933	86	ACP
3	19269–30230	3654	PKS (modules 4 and 5)
	19374–20645	424	KS
	20973–22007	345	AT
	22530–23057	176	KR* (inactive)
	23322–23579	86	ACP
	23646–24926	427	KS
	25293–26315	341	AT
	26355–26909	185	DH
	27789–28835	349	ER
	28842–29393	184	KR
	29679–29936	86	ACP
4	30294–35000	1569	PKS (module 6)
	30423–31691	423	KS
	32013–33038	342	AT
	33786–34337	184	KR
	34593–34853	87	ACP
5	35000–40636	1879	PKS (module 7, TE)
	35120–36388	423	KS
	36692–37723	344	AT
	38633–39184	184	KR
	39455–39712	86	ACP
	39824–40636	271	TE
7	40626–41097	157	N-Methyltransferase (N-terminal portion)

have 68% identity to the product of ORF 3*, an N-methyltransferase of the *tylMI*-encoded region of the tylosin biosynthetic cluster (12). Like ORF 7, ORF 3* is downstream of and reads in the same direction as the genes encoding PKS; however, it is not immediately adjacent to the genes encoding PKS. Based on sequence alignments and distance from potential ribosome binding sites (RBSs) (30), ORFs 6 and 7 appear to use ATG as the start codon (Table 1), with the putative RBSs being GGAGA and GGAGG, respectively. In addition, ORF 7 may be translationally coupled to ORF 5, since the putative ATG start codon of ORF 7 begins 11 bases upstream of the TGA stop codon of ORF 5.

The larger ORFs (ORFs 1 to 5) were found to encode a PKS comprising proteins with molecular masses of 458, 194, 379,

167, and 198 kDa, respectively. The genes corresponding to these ORFs were designated *nidA1* through *nidA5*, respectively. *nidA1*, *nidA2*, and *nidA5* appear to begin with ATG codons (Table 1) and have the putative RBSs GGAGG, GGGGG, and GGAG, respectively. *nidA3* and *nidA4* appear to begin with GTG codons, with the RBSs being GGAGA and GGAAG, respectively. The enzymatic motifs encoded by these genes (i.e., KS, AT, DH, ER, KR, and ACP) were delineated by BLAST searches which identified homology to other PKS genes deposited in the EMBL, GenBank, and Swissprot databases. As is characteristic of type I PKSs, the motifs within each ORF product are organized into modules. Seven modules, corresponding to the seven condensation reactions predicted to be required for niddamycin biosynthesis, were found in the PKS; NidA1 and NidA3 contain two modules each, whereas NidA2, NidA4, and NidA5 each contain one module (Fig. 3). In addition, the sets of motifs contained within each module were also consistent with the extent of β -carbon processing at the corresponding positions of the niddamycin ring (Fig. 1 and 3). Besides the KS, AT, and ACP domains found in all modules, modules 2 and 3 have the KR and DH domains required for the formation of the double bonds found between carbons 12 and 13 and carbons 10 and 11, respectively. Module 5 has a full complement of domains (KS, AT, DH, ER, KR, and ACP) to produce the methylene group at C-7. Modules 4, 6, and 7 each have KR domains which would result in hydroxyl groups at C-9, C-5, and C-3, respectively. However, sequence analysis of the KR in module 4 indicates that it is inactive (see below), resulting in the keto group at C-9. A TE domain with the conserved GxSxG and GdH motifs (10) was found at the carboxy-terminal end of NidG5 (module 7), identifying the end of the PKS cluster.

In contrast to the erythromycin PKS, where the N-terminal enzymatic motifs are a loading AT and ACP, the loading AT and ACP of the niddamycin PKS follow a KS domain which is not predicted to be necessary for the initiation of polyketide synthesis. Homology of this motif with the other KS domains in the cluster ranges from only 45 to 47% identity, whereas homologies among the KS domains of modules 1 to 7 range from 70 to 92% identity. More importantly, sequence analysis indicates that this KS is enzymatically nonfunctional, as a critical cysteine residue in the motif TVDTGCSSSLV, which is highly conserved among KS domains (3, 10), is replaced by a glutamine residue at position 173, forming a KS^Q domain (Fig. 4). Interestingly, it has recently been reported that KS^Q domains are also found at the N termini of the PKSs which synthesize the other 16-membered macrolides, carbomycin, spiramycin, and tylosin (17). The function of this motif, if any, remains to be determined.

All of the remaining seven KS domains of this cluster retain the conserved active-site cysteine residue. In addition, two highly conserved His^{308/347} residues (3, 11) are present in all of the niddamycin KS domains, including the unusual starter KS (Fig. 4).

The niddamycin AT domains show more sequence variability than the KS domains. For example, AT2 and AT6 only have 30% amino acid identity, whereas AT2 and AT3 have 95% amino acid identity. It has been demonstrated that mATs and mmATs fall into distinct families based on amino acid sequence and can be distinguished by conserved motifs (28). As a result, it may be possible to predict AT substrate specificity on the basis of primary sequence. A PILEUP/DENDRO GRAM analysis comparing the niddamycin ATs with those from the rapamycin and erythromycin PKS clusters is shown in Fig. 5. This analysis demonstrates that the loading AT, as well as those of modules 1, 2, 3, and 7, fall into the group of ATs

KS	171	305	344	
NidKS1	TACSSSLVAL	-----VEAHGTGTAL	-----NIGHTQAAAG	
NidKS2	TACSSSLVAL	-----VEAHGTGTPL	-----NIGHTSQAAG	
NidKS3	TACSSSLVAL	-----VEAHGTGTPL	-----NIGHTQAAAG	
NidKS4	TACSSSLVAL	-----VEAHGTGTPL	-----NIGHTQAAAG	
NidKS7	TACSSSLVAL	-----VEAHGTGTPL	-----NIGHTQAAAG	
NidKS6	TACSSSLVAL	-----VEAHGTGTPL	-----NIGHTQAAAG	
NidKS5	TACSASLTAL	-----VEAHGTGTHL	-----NIGHTQAAAG	
NidKSs*	TGQSSSLVAV	-----VELHGTGTPA	-----NIGHTLEGAAG	
	*			
DH	34			
NidDH2	VSLAAHPWLA	DHTVAGAVLF	PGTAFADLLL	
NidDH5	IGVDEFPWMA	DHTLLGAVLL	PGAAFADLAL	
NidDH3	LSARALPWLA	DHLVWDRGVA	PGTAVLEIVL	
		*		
KR	1			
NidKR4*	GTVLVVGDS	TVPARLLRSL	LD.DGAERVV	
NidKR6	GTILITGGT	ALGTHIATWL	AHK.GAKHLI	
NidKR7	GTILITGGT	ALGTHIATWL	AHK.GAKHLI	
NidKR3	GTVLITGGT	ALAAVTARHL	VARHGARHLL	
NidKR5	GTVLITGGT	TIGSRITARHL	VTRHGVRRLL	
NidKR2	GTVLLTGGT	ALGRRVAVHL	ARRHGVRRLM	
NidKR1	GTVLITGGT	ALGSQVARRL	AL.AGAPHLL	
ER	154			
NidER5	GLVALAGLRR	GETVLVHAAA	GGVMAAVQI	ARHLGAQVLA
EryER4	ALHDLAGLRA	QQSVLIHAAA	GGVMAAVAL	ARRAGAEVLA
ACP	35			
NidACP4	STRAFRELGF	ESLTAVELRG	RLAEATGLTL	
NidACP7	ARASFRLDG	DSLAAVSLRD	GLAEATGLEL	
NidACP6	HDVAFKDLGF	DSLAAVKMRT	RLRETTGLDL	
NidACP3	AERSPKDAGF	DSLTAVDLRN	RLNARTGLRL	
NidACP5	SDRAFREAGF	DSLTAVELRN	RLAAATGLRL	
NidACP2	PERPFHEIGF	DSLATLELRN	RLGRLVGLRL	
NidACP1	AETAFFRAGGF	DSLTLVELRN	RLASATGLRL	
NidACPs	LDRPFTSQGL	DSMTAVELAG	LLGTAAGVAL	
		*		
TE	101			
NidTE	PFALAGHSAG	ANVAYALAAY	-----DVPGNHFTML	
EryTE	PFVVAGHSAG	ALMAYALATE	-----AVPGDHFTMV	
		*		

FIG. 4. Conserved active sites of the niddamycin PKS domains. Deduced protein sequences were aligned by using the PILEUP/PRETTY program. Only the regions containing the proposed active sites are shown. The numbers correspond to the amino acid position from the first residue in the first motif listed. The boldfaced amino acids are conserved residues or motifs important for the function of each domain. Active-site residues are marked with asterisks. Motifs important for functioning of the KR and ER domains are underlined. For comparative purposes, the erythromycin ER from module 4 and the erythromycin TE are included.

that use malonyl-CoA as the substrate. Based on the structure of niddamycin (Fig. 1), the acetate-derived methyl group at C-15 and the lack of methyl groups at C-14, C-12, C-10, and C-2 are consistent with those ATs being malonyl-CoA specific. On the other hand, the ATs of modules 4, 5, and 6 fall into the methylmalonyl-specific group. The methyl group at C-8 is consistent with AT4 being an mmAT. AT5, however, presumably uses ethylmalonyl-CoA as a substrate, resulting in the ethyl side chain at C-6. The fact that it does not segregate into a unique group in the dendrogram indicates that methylmalonyl and ethylmalonyl ATs may be relatively similar. The AT domain in module 6 also falls into the methylmalonyl family, but its activity results in a hydroxyl or perhaps a methoxy group at C-4. The biosynthesis of leucomycin, a 16-membered macroide which contains a methoxy group at C-4, was examined by feeding the producing organism, *Streptovericillium kitasatoensis*, 2-¹³C-labeled precursors (23, 24). Labeled malonate was not incorporated at carbons 3 and 4, indicating that malonyl-CoA was not a substrate for the corresponding AT. Labeled glycerol, however, was incorporated at C-4, suggesting that glycolate may be the substrate for the AT of the module corresponding to the C-4 position of leucomycin. It is conceivable

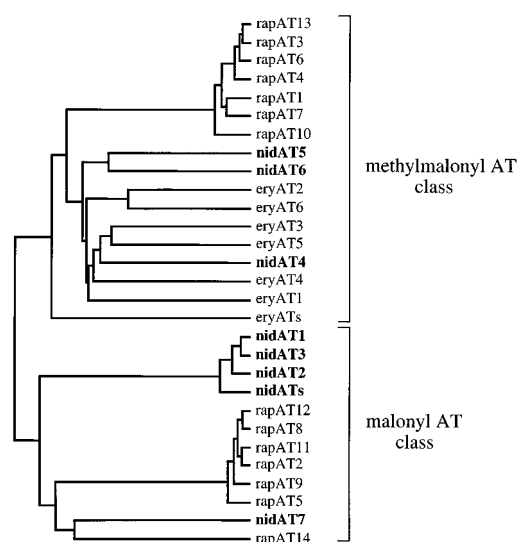


FIG. 5. Dendrogram analysis of AT domains. The dendrogram shows the sequence similarity of AT domains from the niddamycin, erythromycin, and rapamycin PKS clusters. The niddamycin AT domains are boldfaced. The ATs are grouped according to the substrate specificities of the rapamycin and erythromycin ATs (malonyl-CoA or methylmalonyl-CoA).

that the AT in module 6 of the niddamycin pathway uses the same or a similar substrate.

A closer examination of the amino acid sequence (Fig. 6) of the niddamycin ATs shows that the active-site residues Gln⁶, Ser⁹¹ (of the GHSXG motif where Ser is involved in the formation of the acyl-enzyme intermediate), Arg¹¹⁶, and His¹⁹²

MALONYL AT CLASS

NidAT1	1	58	89
NidAT3	LLFTGQGAQH	--RTEYTQTALYRTEL
NidAT2	LLFTGQGAQH	--RTEYTQTALYRTEL
NidATs	LLFTGQGAQH	--RTEYTQTALYRTEL
NidAT7	PLFSGQGAQR	--RTEYTQTALYRTEL
	*		*
con. acet.		ETGYAQ.A.FGLL
		GHSXG

NidAT1	113	187
NidAT3	ITARAHLMGQ	--VSHAFHSALM
NidAT2	ITARAHLMGQ	--VSHAFHSALM
NidATs	ITARAHLMGQ	--VSHAFHSALM
NidAT7	VAARGRLMQR	--VSHAFHSALM
	*	*

METHYLMALONYL AT CLASS

NidAT5	1	58	89
NidAT6	FVFPQGGSQW	--RVDVVMVSLAAVW
NidAT4	FVFPQGGSQW	--RVDVVMVSLAAVW
	*		*
con. prop.		RVDVVM.S.AAHW
		GHSQG

NidAT5	113	190
NidAT6	VALRSRAWLG	--VDfAGHSPOV
NidAT4	IAGRSRLWGR	--VDfAGHSPOV
	*	*

FIG. 6. Alignment of the sequences of conserved motifs found in AT domains. Sequences were aligned by using the PILEUP/PRETTY program. Boldfaced amino acids are conserved residues or motifs important for the function of AT domains. Active-site residues are marked with asterisks. Proposed consensus sequences for motifs of ATs specific for acetate and propionate (13) are indicated.

are present in all mATs and mmATs (3, 10). In addition, motifs which have been proposed to distinguish mATs from mmATs (13) seem to be more divergent in the niddamycin cluster (Fig. 6, residues 58 to 77); however, these same stretches of amino acids, though they have different sequences, still seem to be specific for mATs and mmATs within the niddamycin sequence.

KR domains were found in all seven modules. Homologies between the KR domains ranged from 31% amino acid identity between KR4 and KR6 to 99% amino acid identity between KR6 and KR7. It has been observed that modules corresponding to unreduced β -carbonyls in the polyketide chains of erythromycin (9), rapamycin (3), and spiramycin (17) still contain KR motifs with, in most cases, obvious inactivating mutations. An inactive KR in module 4 would be consistent with the presence of the keto group at C-9 of niddamycin and would be another example of this type of organization. The consensus NADP(H) binding site of a KR is GxGxxGxxxA (29). The corresponding region of the KR4 domain of the niddamycin cluster is DxTxxPxxxL (Fig. 4), indicating that the KR is non-functional. All of the other KR domains in the cluster contain the expected NADP(H) binding motif with the exception of KR2. This KR is predicted to be active; however, a Cys residue is present at the first position of the NADP(H) binding site rather than a Gly. How this amino acid change would affect KR activity is uncertain.

DH domains were identified in modules 2, 3, and 5. Hydroxyl groups formed by KRs at C-13 and C-11 are reduced by DHs to form the double bonds between carbons 12 and 13 and carbons 10 and 11, corresponding to modules 2 and 3, respectively. Formation of the methylene group at C-7 (corresponding to module 5) also requires DH activity. DH2 and DH3 have 39% amino acid identity, DH2 and DH5 have 45% identity, and DH3 and DH5 have 38% identity. The highly conserved residues His⁴⁵, Gly⁴⁹, and Pro⁵⁴ (5, 10) were present in DH2 and DH5, but, surprisingly, in DH3 there is an Asp in place of the Gly residue (Fig. 4). Based on the niddamycin structure, this DH domain should be active, and so this divergence from consensus may not adversely affect the activity of the enzyme.

The only ER domain present in the PKS cluster is in module 5. An amino acid sequence comparison of ER5 with ER domains from the products of the erythromycin and rapamycin PKS genes (Fig. 4) reveals the presence of the putative NADP(H) binding motif LxHxg(a)xGGVG (2, 29, 32).

ACP domains were identified in each of the modules and after the loading AT, as expected. Homologies between the ACP domains range from 32 to 67% amino acid identity. The pantotheine-binding Ser⁴⁶ residue (10) in the GFDSL motif was present in all of the ACP domains (Fig. 4).

Gene disruption of the niddamycin PKS cluster. A gene disruption experiment was performed with *S. caelestis* to demonstrate that the PKS genes isolated from the genomic library were involved in niddamycin biosynthesis. A 1.5-kb *Sst*I fragment containing AT- and DH-encoding sequences from the coding regions for module 3 (Fig. 2) was isolated from cosmid pCEL18h5, treated with Klenow fragment, and cloned into the *Eco*RV site of pKC1139 to generate pSK7595 (Fig. 7). This plasmid contains an apramycin resistance gene for selection, *ori*T for conjugal transfer, and both *E. coli* and *Streptomyces* origins of replication. The *Streptomyces* origin of replication from *Streptomyces ghanaensis* (21) functions only at temperatures below 34°C, and so under selective pressure at a nonpermissive temperature, the plasmid must integrate into the chromosome through homologous recombination to be maintained. Plasmid pSK7595 was first transformed into *E. coli* S17-1, the donor strain carrying the conjugation locus RP4. To transfer the

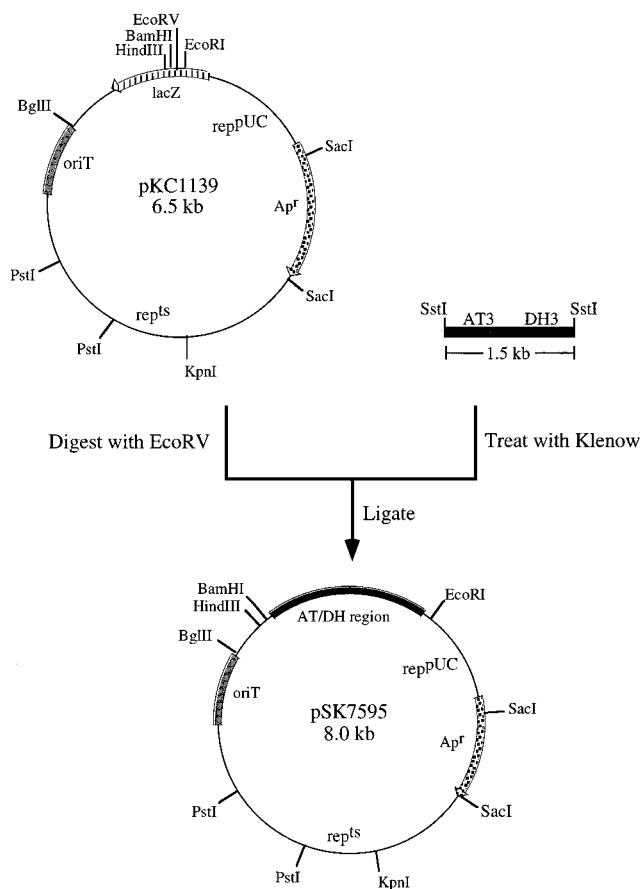


FIG. 7. Construction of pSK7595. The vector pKC1139 was used to construct the conjugation vector pSK7595. The *ori*T gene allows for conjugal transfer from *E. coli* strains containing the RP4 locus. The apramycin resistance gene (*Ap*^r) allows for selection of the plasmid. The temperature-sensitive *Streptomyces* origin of replication (*rep*^{ts}) and the *E. coli* origin of replication (*rep*^{pUC}) are shown. The 1.5-kb *Sst*I fragment containing AT- and DH-encoding sequences was cloned into the *Eco*RV site of pKC1139 (black bar).

plasmid to the niddamycin-producing strain, S17-1 transformants were incubated with *S. caelestis* cells as described in Materials and Methods. As a negative control, S17-1 cells without plasmid were also incubated with *S. caelestis* cells. The cultures were challenged with apramycin to eliminate non-plasmid-containing *S. caelestis* cells. Nalidixic acid (50 μ g/ml) is routinely used in *Streptomyces-E. coli* conjugation experiments to eliminate the *E. coli* donor cells; however, *S. caelestis* was found to be sensitive to this antibiotic. Therefore, the conjugation cultures were challenged with bacteriophage T7 in order to remove the *E. coli* cells. Cultures were incubated for 7 days at 30°C, and in each of two separate experiments, only one *S. caelestis* colony was recovered.

One putative transconjugant, *S. caelestis*(pSK7595), was grown at 37°C in apramycin-containing medium to force integration of the temperature-sensitive plasmid. Southern analysis of genomic DNA indicated that the plasmid was present and had integrated into the appropriate location in the chromosome (Fig. 8). This strain was designated *S. caelestis* nid::SK7595. Surprisingly, the original transconjugant isolated at 30°C and never grown at a nonpermissive temperature also showed the pattern expected for integrated plasmid, indicating that this plasmid was not replicating at the permissive temperature in *S. caelestis*. This may also explain the extremely low

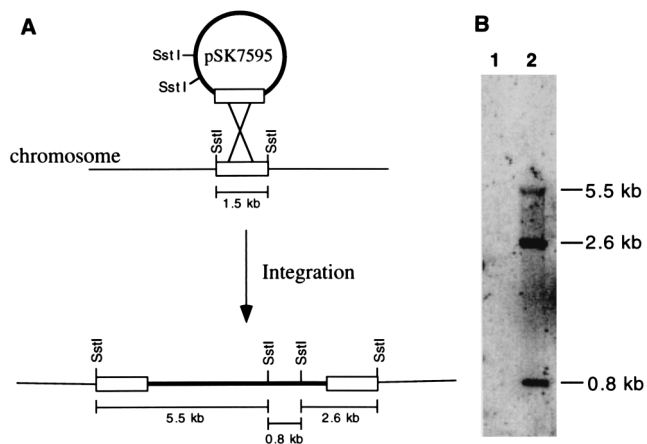


FIG. 8. Disruption of the niddamycin PKS locus. (A) Diagrammatic representation of the integration of pSK7595 into the *S. caelestis* chromosome. The region of homology between the plasmid and chromosome is represented by an open box. pKC1139 DNA is represented by a heavy line. *SstI* sites are indicated. (B) Southern analysis of *SstI*-digested genomic DNA probed with vector pKC1139. Lane 1, *S. caelestis* NRRL-2821 (wild type); lane 2, *S. caelestis*(pSK7595). DNA fragment sizes based on migration of a 1-kb ladder (BRL) on the same gel are indicated at the right.

frequency of conjugation since transconjugants could be selected only if a rare integration event took place before the plasmid was lost.

Spores of *S. caelestis* *nid::SK7595* were plated onto R3M medium containing apramycin to generate single colonies. Three of the isolates were examined for the production of niddamycin. Cells were grown in apramycin-containing medium at 30°C (data not shown) and 37°C for 4 days, and supernatants were examined by TLC and TLC-bioautography (Fig. 9). None of the *S. caelestis* *nid::SK7595* isolates, grown at either temperature, produced a spot on the TLC plate at the R_f of niddamycin. In addition, no bioactivity was observed in the

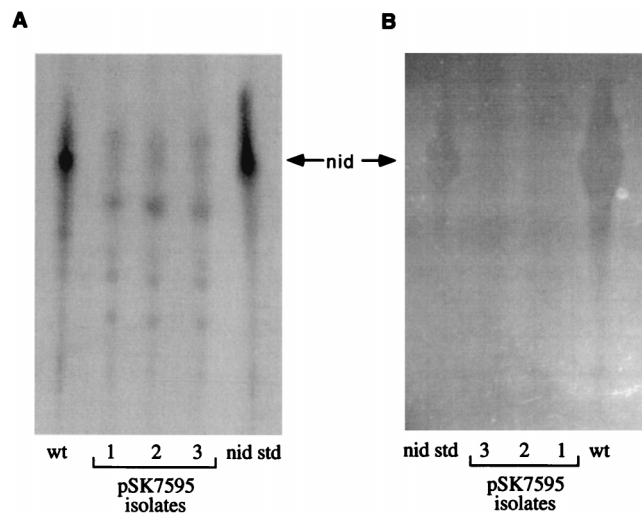


FIG. 9. Niddamycin production in *S. caelestis*(pSK7595). (A) TLC analysis of ethyl acetate extractions of culture supernatants. Lanes: wt (wild type), *S. caelestis* NRRL-2821; 1 to 3, isolates of *S. caelestis*(pSK7595). nid std (niddamycin standard), 10 µg of niddamycin. The R_f of niddamycin (nid) is indicated by the arrow. (B) TLC-bioautography analysis of ethyl acetate extractions of culture supernatants. The lane assignments and niddamycin location are identified as in panel A. Dark spots are zones of inhibition of *S. aureus* due to the presence of niddamycin.

TLC-bioautography assay, indicating that the niddamycin PKS cluster had been disrupted and that the cloned genes were from the niddamycin pathway.

DISCUSSION

We have reported the isolation and characterization of the PKS genes involved in the biosynthesis of the macrolide antibiotic niddamycin. The initial approach to the isolation of this cluster was to PCR amplify AT regions based on amino acid sequences which are found to be highly conserved among PKSs. We generated primers corresponding to regions in the KS and AT domains with the expectation of amplifying DNA corresponding to portions of all seven domains hypothesized to be present in the cluster. (The subsequent discovery that a KS^O domain preceding the loading AT brings the number of potentially amplifiable regions to 8.) The PCR resulted in the coding region for only one AT, that from module 5, being amplified. On reexamination of the homology of the niddamycin nucleotide sequences to the primers, it was found that there was only one mismatch near the 5' end of the KS primer among the eight niddamycin KS nucleotide sequences but a much higher frequency of mismatches at the 3' end of the AT primer. For example, there were six mismatches each for five of the niddamycin AT nucleotide sequences, three mismatches for one and one for another. The coding region for module 5, which was successfully amplified, had no mismatches. Therefore, this PCR approach proved fruitful for generating a niddamycin-specific probe to isolate the remaining coding regions of the cluster but not for simultaneously amplifying the coding regions for multiple domains.

Probing the *S. caelestis* genome with either the PCR-generated probe or a probe from a conserved region of the erythromycin PKS generated the same hybridization patterns, indicating that these were the only type 1 PKS sequences present in the strain. It has been observed that some strains of *Streptomyces*, such as *Streptomyces hygroscopicus* 29253, contain multiple PKS clusters (18, 26), which can complicate isolation of a particular cluster by hybridization. This does not appear to be the case in *S. caelestis*. When the PCR-generated probe was used for Southern analysis of *SstI*-digested *S. caelestis* DNA, we identified seven cross-hybridizing bands which corresponded to seven of the *SstI* fragments encoding the niddamycin PKS.

Sequence analysis of the isolated DNA fragments revealed a PKS cluster comprising seven modules as expected for synthesis of niddamycin. The modules are encoded by five ORFs, with two modules each encoded by ORFs 1 and 3 and one each encoded by ORFs 2, 4, and 5. The organization of the enzymatic domains contained within each module is also consistent with what would be expected for synthesis of the macrolide ring of niddamycin. An unexpected KS domain (KS^O) with a glutamine replacing a conserved cysteine residue was found at the amino terminus of the PKS. The function of this domain, if any, remains unclear. Platenolide A is the polyketide backbone of the macrolide antibiotic spiramycin and is identical in structure to the polyketide backbone of niddamycin. A genetic map of the spiramycin PKS was recently published (17), and the organization and domain content of the modules, including the unusual KS^O , are the same as those found in this study for the niddamycin cluster. The degree of genetic relatedness of these clusters awaits the release of the spiramycin PKS nucleotide sequence.

Previous work has suggested that AT sequences cluster into families based on substrate specificity (13, 28). As predicted from the structure of niddamycin, the AT domains of modules

1, 2, 3, and 7 and the loading AT fall into the class of ATs which utilize malonate. The remaining ATs fall into the methylmalonate class. The methyl group at C-8 is consistent with AT4 being an mmAT. AT5 and AT6, however, are predicted to utilize, respectively, ethylmalonyl-CoA and a CoA derivative of unknown structure which would result, perhaps by further modification, in the methoxy group at C-4. It should also be noted that the loading AT for the erythromycin PKS also falls within the methylmalonate class even though the substrate for this AT is propionyl-CoA. Thus, it seems that the methylmalonyl grouping may encompass ATs which recognize a broader range of substrates than previously predicted. As a result, it may be difficult to assign substrate specificity to ATs not assigned to known PKS clusters or to those whose modular positions have not been determined within a known cluster solely on the basis of its homology with other members of the methylmalonyl class.

Domain replacements within the erythromycin PKS have been shown to result in the production of novel bioactive compounds (11, 22, 25). The ethylmalonyl AT domain (AT5) and the methoxy AT domain (AT6) of the niddamycin cluster could conceivably be used to replace the mmAT domains in the erythromycin PKS to generate erythromycin derivatives with novel polyketide backbone structures that would be difficult to produce by chemical methods.

ACKNOWLEDGMENTS

We thank R. Summers for KS/AT primers, M. Staver for assisting in genomic library construction, T. Kavanaugh and S. Nannapeni for synthesis of oligonucleotides, and P. Youngman for insightful suggestions.

REFERENCES

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Amy, C. M., A. Witkowski, J. Naggert, B. Williams, Z. Randhawa, and S. Smith. 1989. Molecular cloning and sequencing of cDNAs encoding the entire rat fatty acid synthase. *Proc. Natl. Acad. Sci. USA* **86**:3114–3118.
- Aparicio, J. F., I. Molnar, T. Schwecke, A. Konig, S. F. Haydock, L. E. Khaw, J. Staunton, and P. Leadlay. 1996. Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase. *Gene* **169**:9–16.
- Baltz, R. H. 1980. Genetic recombination in *Streptomyces fradiae* by protoplast fusion and cell regeneration. *Dev. Ind. Microbiol.* **21**:43–54.
- Bevitt, D. J., J. Cortes, S. F. Haydock, and P. F. Leadlay. 1992. 6-Deoxyerythronolide B synthase 2 from *Saccharopolyspora erythraea*. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme. *Eur. J. Biochem.* **204**:38–49.
- Bierman, M., R. Logan, K. O'Brien, E. T. Seno, R. N. Rao, and B. E. Schoner. 1992. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* **116**:43–49.
- Cortes, J., S. F. Haydock, G. A. Roberts, D. J. Bevitt, and P. F. Leadlay. 1990. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature* **348**:176–178.
- Devereux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387–395.
- Donadio, S., M. J. Staver, J. B. McAlpine, S. J. Swanson, and L. Katz. 1991. Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**:675–679.
- Donadio, S., and L. Katz. 1992. Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythraea*. *Gene* **111**:51–60.
- Donadio, S., J. B. McAlpine, P. J. Sheldon, M. Jackson, and L. Katz. 1993. An erythromycin analog produced by reprogramming polyketide synthesis. *Proc. Natl. Acad. Sci. USA* **90**:7119–7123.
- Fager, E. E., T. J. Oliver, J. F. Prokop, and A. C. Sinclair. November 1965. U.S. patent 3218239.
- Gandecha, A. R., S. L. Large, and E. Cundliffe. 1997. Analysis of four tylosin biosynthetic genes from the *tylLM* region of the *Streptomyces fradiae* genome. *Gene* **184**:197–203.
- Haydock, S., J. F. Aparicio, I. Molnar, T. Schwecke, A. Konig, A. F. A. Marsden, I. S. Galloway, J. Staunton, and P. F. Leadlay. 1995. Divergent structural motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acylcarrier protein transacylase domains in the modular polyketide synthases. *FEMS Lett.* **374**:246–248.
- Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schempf. 1985. Genetic manipulation of *Streptomyces*: a laboratory manual. The John Innes Foundation, Norwich, England.
- Hopwood, D. A., and D. H. Sherman. 1990. Molecular genetics of polyketide and its comparison to fatty acid biosynthesis. *Annu. Rev. Genet.* **24**:37–66.
- Huber, V. G., K. H. Wallhauser, L. Fries, A. Steigler, and H.-L. Weidenmuller. 1962. Niddamycin, ein neues makrolid-antibiotikum. *Arzneim. Forsch.* **12**:1191–1195.
- Kuhstoss, S., M. Huber, J. R. Turner, J. W. Paschal, and R. N. Rao. 1996. Production of a novel polyketide through the construction of a hybrid polyketide synthase. *Gene* **183**:231–236.
- Lomovskaya, N., L. Fonstein, X. Ruan, D. Stassi, L. Katz, and C. R. Hutchinson. 1997. Gene disruption and replacement in the rapamycin-producing *Streptomyces hygroscopicus* strain ATCC 29253. *Microbiology* **143**:875–883.
- Merson-Davies, L. A., and E. Cundliffe. 1994. Analysis of five tylosin biosynthetic genes from the *tylIBA* region of the *Streptomyces fradiae* genome. *Mol. Microbiol.* **13**:349–355.
- Molnar, I., J. F. Aparicio, S. F. Haydock, L. E. Khaw, T. Schwecke, and D. O'Hagan. 1993. Biosynthesis of fatty acid and polyketide metabolites. *Nat. Prod. Rep.* **10**:593–624.
- Muth, G., B. Nussbaumer, W. Wohlleben, and A. Puhler. 1989. A vector system with temperature-sensitive replication for gene disruption and mutational cloning in streptomyces. *Mol. Gen. Genet.* **219**:341–348.
- Oliynyk, M., M. J. B. Brown, J. Cortes, J. Staunton, and P. F. Leadlay. 1996. A hybrid modular polyketide synthase obtained by domain swapping. *Chem. Biol.* **3**:833–839.
- Omura, S., A. Nakagawa, H. Takeshima, K. Atsumi, J. Miyazawa, F. Piriou, and G. Lukacs. 1975. Biosynthetic studies using ¹³C enriched precursors on the 16-membered macrolide antibiotic leucomycin A₃. *J. Am. Chem. Soc.* **97**:6600–6602.
- Omura, S., K. Tsuzuki, A. Nakagawa, and G. Lukacs. 1983. Biosynthetic origin of carbons 3 and 4 of leucomycin aglycone. *J. Antibiot.* **36**:611–613.
- Ruan, X., A. Pereda, D. L. Stassi, D. Zeidner, R. G. Summers, M. Jackson, A. Shivakumar, S. Kakavas, M. J. Staver, S. Donadio, and L. Katz. 1997. Acyltransferase domain substitutions in the erythromycin polyketide synthase yield novel erythromycin derivatives. *J. Bacteriol.* **179**:6416–6425.
- Ruan, X., D. Stassi, S. Lax, and L. Katz. A second type-I PKS gene cluster isolated from *Streptomyces hygroscopicus* ATCC 29253, a rapamycin-producing strain. *Gene*, in press.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Schwecke, T., J. F. Aparicio, I. Molnar, A. Konig, L. E. Khaw, S. F. Haydock, M. Oliynyk, P. Caffrey, J. Cortes, J. B. Lester, G. A. Bohm, J. Staunton, and P. F. Leadlay. 1995. The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc. Natl. Acad. Sci. USA* **92**:7839–7843.
- Scrutton, N. S., A. Berry, and R. N. Perham. 1990. Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature* **343**:38–42.
- Strohl, W. R. 1992. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res.* **20**:961–974.
- Tuan, J. S., J. M. Weber, M. J. Staver, J. O. Leung, S. Donadio, and L. Katz. 1990. Cloning of genes involved in erythromycin biosynthesis from *Saccharopolyspora erythraea* using a novel actinomycete-*Escherichia coli* cosmid. *Gene* **90**:21–29.
- Witkowski, A., V. S. Rangan, Z. I. Randhawa, C. M. Amy, and S. Smith. 1991. Structural organization of the multifunctional animal fatty-acid synthase. *Eur. J. Biochem.* **198**:571–579.
- Wright, F., and M. J. Bibb. 1992. Codon usage in the G+C-rich *Streptomyces* genome. *Gene* **113**:55–65.
- Yamamoto, H., K. H. Maurer, and C. R. Hutchinson. 1986. Transformation of *Streptomyces erythraeus*. *J. Antibiot.* **39**:1304–1313.
- Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**:103–119.