Identification and Characterization of the Niddamycin Polyketide Synthase Genes from Streptomyces cælestis

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 cælestis 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 cælestis 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, mycaminose-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 ketoeductases [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 condensation 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 undefined 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 domains 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 K2SO4, 0.25 g; yeast extract, 4 g; RSM plates consist of (per liter) the following: sucrose, 103 g; K2SO4, 0.25 g; yeast extract, 4 g; Casamino Acids, 4 g; trypotine, 4 g; agar, 22 g; and H2O, 830 ml. After sterilization, 20 ml of a 2.5 M solution of MgCl2, 20 ml of a 50% glucose solution, 20 ml of a 2.5 M solution of CaCl2, 12.5 ml of a 2 M solution of Tris-HCl (pH 7.0), 0.2 ml of a 5,000× 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 KH2PO4, were added.

Strains, bacteriophage, and plasmids. Escherichia coli DH5α 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 S171 was obtained from Philippe Mazodier, Institut Pasteur, Paris, France. Bacteriophage T7 was obtained from Phil Youngman, University of Georgia, Athens, Ga. Plasmid pGEM-3Zf was purchased from Promega Corporation, Madison, Wis. Plasmids pUC19 (35), pKC1139 (6), and pNJ1 (31) have been described previously.

Strains, bacteriophage, and plasmids. Escherichia coli DH5α 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 S171 was obtained from Philippe Mazodier, Institut Pasteur, Paris, France. Bacteriophage T7 was obtained from Phil Youngman, University of Georgia, Athens, Ga. 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 Qiagen 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.
Construction of a Streptomyces caelestis genomic library. S. caelestis DNA was prepared by partially digesting 5 µg of DNA with SauIII and size selecting fragments of around 40 kb by electrophoresis through a 0.4% agarose gel. Cosmid pNmI arms were prepared by digestion of the vector with EcoRI, dephosphorylation with CIAP, and then digestion with BglII to generate one arm and digestion of the cosmid with HindIII, dephosphorylation with CIAP, and then digestion with BglII to generate the other. Ligation of 1 µg of cosmid arms to 1 µg of the size-selected S. caelestis DNA was in a 20-µl volume. Two microcentrifuers of this ligation mix was packaged by using GigapackII XL (Stratagen) 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'-CGGTSAAAGTCSAAACTCGG-3') and the AT-specific oligonucleotide (5'-GCRATCTCRRCTCGCARTG-3') were used in a PCR mixture containing Thermopol reaction buffer (New England Biolabs, Beverly, Mass.), 0.2 mM deoxynucleoside triphosphate mixture, 0.5 µg of genomic DNA, and 100 µn of the primers in a volume of 98 µl. After incubation at 96°C for 2 min, 2 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, min. The reaction generated a 900-bp DNA fragment which was gel purified

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).

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 deposed 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).

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 [γ-32P]ATP (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× SSC (1× 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× 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 SsrI and the resulting fragments were subcloned into vector pGEM-ZfI or pUC19. SsrI fragments that were larger than 8 kb were further digested with Smal and subcloned to generate plasmids that would be more suitable for sequencing. Two methods were used for DNA sequencing. One method used [α-35S]dCTP (DuPont, NEN Research Products) and the female DNA cycling sequence 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 Readiness kit (Perkin-Elmer Corporation), and the samples were run on 4.75% polyacrylamide–8.5 M urea gels on an Applied Biosystems 373 sequencer. All subclones were sequenced on both strands by using multilonging site primers to isolate 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. cattelis recipient cells were prepared by inoculating 30 ml of SGGP with 25 µl 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 pellets were washed once in SGGP, reconstituted, and resuspended in 1 ml of E. coli donor cells. The plates were inoculated at 30°C for 20 min and washed once in the same medium. After a final centrifugation, the pellets were resuspended in 500 µl of LB. Nucleopore membranes (25 mm, 0.2 µm; VWR) were placed onto an AS-1 agar plate and spotted with 20 µl 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 µl 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 6 µl of a 50-mg/ml solution of apramycin and 150 µl of T7 bacteriophage (1010 PFU) and returned to 30°C until transconjugants appeared (about 7 days).

Nucleotide sequence accession number. The GenBank accession number for the niddamycin PKS and portions of two flanking genes is AF016585.
analyzed by Southern hybridization. Banding patterns of SstI-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 did not cross-react (i.e., were false positives). Lanes containing cosmid DNA specifically referred to in the text are labeled.

The large SstI fragment was isolated from clones pCEL13f5 and pCEL18h5 and further digested with Smal 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. 2. Southern analysis of positively hybridizing cosmids. Nineteen cosmids that hybridized with the KS/AT probe were digested with SstI 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.](http://jb.asm.org/)

![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.](http://jb.asm.org/)
TABLE 1. Niddamycin PKS sequence coordinates

<table>
<thead>
<tr>
<th>ORF no.</th>
<th>Nucleotide coordinates</th>
<th>No. of amino acids</th>
<th>Homology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>627–13646</td>
<td>4340</td>
<td>PKS (loading domain)</td>
</tr>
<tr>
<td>2</td>
<td>13693–19209</td>
<td>1893</td>
<td>PKS (module 3)</td>
</tr>
<tr>
<td>3</td>
<td>19269–30230</td>
<td>3654</td>
<td>PKS (modules 4 and 5)</td>
</tr>
<tr>
<td>4</td>
<td>30294–35000</td>
<td>1569</td>
<td>PKS (module 6)</td>
</tr>
<tr>
<td>5</td>
<td>35000–40636</td>
<td>1879</td>
<td>PKS (module 7, TE)</td>
</tr>
<tr>
<td>6</td>
<td>1–396</td>
<td>132</td>
<td>P450 (N-terminal portion)</td>
</tr>
<tr>
<td>7</td>
<td>40626–41097</td>
<td>157</td>
<td>N-Methyltransferase (N-terminal portion)</td>
</tr>
</tbody>
</table>

have 68% identity to the product of ORF 3*, an N-methyltransferase of the tyml1-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, 406, and 410 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, GGAGG, and GGAG, respectively. nidA3 and nidA4 appear to begin with GTG codons, with the RBSs being GGAGA and GGAGG, respectively. The enzymatic motifs encoded by these genes (i.e., KS, AT, DH, ER, KR, and ACP) are 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 PKSs; 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 hydroxy groups at C-8, C-9, and C-10, 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 TVDTGSSSLV, which is highly conserved among KS domains (3, 10), is replaced by a glutamine residue. A TE domain with the conserved GxSxG and GdH motifs (10) was found at the carboxy-terminal end of NidG5 (module 7), identifying the last module of the PKS cluster.

All of the remaining seven KS domains of this cluster retain the conserved active-site cysteine residue. In addition, two highly conserved His303/347 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/DENDROGRAM 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...
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 macroclide which contains a methoxy group at C-4, was examined by feeding the producing organism, *Streptoverticillium kitasatoense*<sup>2</sup>-<sup>13</sup>C-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 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 acid sequence of ATs 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).
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 GxGxxGxxxxA (29). The corresponding region of the KR4 domain of the niddamycin cluster is DxTxxPxxExL (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 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 LxHx(α,β)GxGxG (2, 29, 32).

ACP domains were identified in 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 SsrI fragment containing AT- and DH-encoding sequences from the coding regions for module 3 (Fig. 2) was isolated from cosmid pCELI8He5, treated with Klenow fragment, and cloned into the EcoRV site of pKC1139 to generate pSK7595 (Fig. 7). This plasmid contains an apramycin resistance gene for selection, ortT 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 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 μg/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* nidi::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
niddamycin. In addition, no bioactivity was observed in the
supernatants were examined by TLC and TLC-bioautography
medium containing apramycin to generate single colonies.
Three of the isolates were examined for the production of
plasmid was lost.
lected only if a rare integration event took place before the
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 me-
dium 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 Rf
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. caele-
slis (pSK7595), nid std (niddamycin
standard), 10 μg of niddamycin. The Rf 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.

DISCUSSION

We have reported the isolation and characterization of the
PKS genes involved in the biosynthesis of the macrolide anti-
biotic 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
domain preceding the loading AT brings the number of poten-
tially 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 niddamy-
cin 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. There-
therefore, this PCR approach proved fruitful for generating a nidd-
amycin-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 eryth-
romycin PKS generated the same hybridization patterns, indi-
cating that these were the only type 1 PKS sequences present
in the strain. It has been observed that some strains of Strept-
tomyces, 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 Sst-digested S. caelestis DNA,
we identified seven cross-hybridizing bands which corre-
sponded to seven of the Sst fragments encoding the niddam-
cin PKS.

Sequence analysis of the isolated DNA fragments revealed a
PKS cluster comprising seven modules as expected for synthe-
sis 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 enzy-
matic 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 (KS5) 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 struc-
ture 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 KS5, 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
TLC-bioautography assay, indicating that the niddamycin PKS
cluster had been disrupted and that the cloned genes were
from the niddamycin pathway.
I. 2, 3, and 7 and the loading AT fall into the class of ATs which utilize malonate. The remaining ATs fall into the methy-lmalonate 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 methymalonyl group 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 properties have 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. Nannapeneni for synthesis of oligonucleotides, and P. Youngman for insightful suggestions.

REFERENCES


