Microbial Ecology

New PCR Primers for the Screening of NRPS and PKS-I Systems in Actinomycetes: Detection and Distribution of These Biosynthetic Gene Sequences in Major Taxonomic Groups

A. Ayuso-Sacido and O. Genilloud

Centro de Investigatión Bńsica, Merck Research Laboratories, Merck Sharp and Dohme de España S.A., Josefa Valcńrcel 38, E-28027 Madrid, Spain

Received: 18 November 2003 / Accepted: 28 January 2004 / Online publication: 21 December 2004

Abstract

Nonribosomal peptide synthetases (NRPS) and type I polyketide synthases (PKS-I) are biosynthetic systems involved in the synthesis of a large number of important biologically active compounds produced by microorganisms, among others by actinomycetes. In order to assess the occurrence of these biosynthetic systems in this metabolically active bacterial group, we designed new PCR primers targeted to specifically amplify NRPS and PKS-I gene sequences from actinomycetes. The sequence analysis of amplified products cloned from two model systems and used to validate these molecular tools has shown the extreme richness of NRPS or PKS-I-like sequences in the actinomycete genome. When these PCR primers were tested on a large collection of 210 reference strains encompassing all major families and genera in actinomycetes, we observed that the wide distribution of these genes in the well-known productive *Streptomyces* species is also extended to other minor lineages where in some cases very few bioactive compounds have been identified to date.

Introduction

A broad range of biologically active polyketide and peptide compounds with applications in medicine, agriculture, and biochemical research are synthesized by type-I polyketide synthases (PKS-I) and nonribosomal peptide synthetases (NRPS). These structurally diverse metabolites include among others antibiotics (e.g., penicillins, vancomycin, and erythromycin), antifungals (e.g., nystatin), antitumor agents (e.g., ansamitocin, bleomycins), anthelmintics (e.g., avermectin) and immunosuppressive agents (e.g., rapamycin and FK506). PKS-I and NRPS biosynthetic systems have been extensively described not only in actinomycetes but also in myxobacteria [7] and cyanobacteria [12], among other bacterial taxa, and in filamentous fungi [8, 28, 37].

Structurally, both PKS-I and NRPS are multifunctional polypeptides encoded by a variable number of modules with multiple enzymatic activities. Each PKS-I module encodes at least three domains corresponding to a ketosynthase (KS), acyltransferase (AT), and acyl carrier protein (ACP) involved in the selection and condensation of the correct extender unit. They can also include additional enoylreductase, dehydratase, and ketoreductase activities involved in the reduction of the β-keto group formed in the condensation. All these domains are involved in a programmed synthesis of the new polyketide chain [1, 15]. The substrate specificity of each elongating carbon unit is determined by the AT domain where two main groups of acyl transferases, malonyl-CoA transferases and methyl-malonyl-CoA transferases, can be clearly distinguished at sequence level [16]. Similarly NRPS modules contain the activities corresponding to the condensation, adenylation, and thiolation steps involved in the recognition and condensation of the substrate. Additional domains (heterocyclase, N-methylase, epimerase, thioesterase, and reductase) are also present depending on the requirements for the substrate activation, elongation, and modification [25, 26, 44].

In the previous decades, natural products screening programs have concentrated an intense effort in the discovery of biologically active metabolites produced by actinomycetes. Whereas enormous progress has been made in the discovery and identification of the genetic organization and mechanism of biosynthesis of numerous compounds of commercial interest, very little is known about the distribution of these biosynthetic sys-

Correspondence to: O. Genilloud; E-mail: olga_genilloud@merck.com

 DOI: 10.1007/s00248-004-0249-6 • Volume 49, 10–24 (2005) •

tems in other actinomycetes, or even in other microbial taxa. Traditionally the selection of the microbial groups of interest has been performed on the basis of morphologic and taxonomic criteria. Nevertheless an early evaluation of these different taxonomic groups for the production of secondary metabolites would allow to focus the screening on the most metabolically talented groups. The detection of gene sequences involved in the synthesis of secondary metabolites to evaluate the biosynthetic potential has already been described [2, 12, 42, 43] in different taxonomic groups. This approach may represent an alternative way to focus the screening not only on microbial groups of taxonomic interest but also on wild-type isolates with the highest metabolic potential. In this work, we present the design and validation of degenerate PCR primers targeted to NRPS and PKS-I sequences specific for actinomycetes. These PCR primers were applied to evaluate the occurrence of NRPS and PKS-I sequences in a group of 210 reference actinomycetes representative of 32 genera. We discuss the usefulness of this approach for the rapid detection of PKS-I and NRPS genes in different actinomycete taxa as well as the frequency and the taxonomic distribution of these sequences that do not always correlate with the productivity of the groups in the study. To our knowledge this is the first time that such a systematic study has been performed to explore the biosynthetic potential across all major actinomycete lineages.

Methods

Bacterial Strains. The bacterial strains used in this study are listed in Table 3. All strains were grown at 28°C on YME agar medium (0.4% yeast extract, 1% malt extract, 0.4% glucose, and 0.2% Bacto-agar) and ATCC-2 liquid medium (0.5% yeast extract, 0.3% beef extract, 0.5% peptone, 0.1% dextrose, 0.2% starch from potato, 0.1% $CaCO_3$, and 0.5% NZamine E).

Design of Oligonucleotide Primers. Sequence comparison and analysis were carried out using programs from the University of Wisconsin GCG package (version 7.2). Alignments of ketosynthase, acyl transferase and adenylation sequences were performed using the multiple alignment program CLUSTAL W [46]. The specificity of degenerated oligonucleotide sequences was tested against all DNA sequences available in GenBank with the program FASTA. Melting temperatures (*T*m) were estimated using the formulas of Thomas and Dancis and the Lathe formulas [45]. Degenerate oligonucleotides A3 (5′GCSTACSYSATSTACACSTCSGG3′), A7R (5′SASGT CVCCSGTSCGGTAS3′), K1 (5′TSAAGTCSAACATCGG BCA3′), and M6R (5′CGCAGGTTSCSGTACCAGTA3′) were supplied by ECOGEN.

DNA Extraction. Total genomic DNA from the different microorganisms used in this study was recovered and purified as previously described [20].

PCR Amplification. DNA preparations were used as template DNA for *Taq* polymerase. Reactions were performed in a final volume of 50 μL containing 10% of extracted DNA, 0.4 μM of each primer, 0.2 mM of each of the four dNTPs (Roche), 5 μL extracted DNA, 1 U *Taq* polymerase (Appligene) with its recommended reaction buffer, and 10% DMSO. In order to find the correct annealing temperatures for each pair of primers, gradient PCRs were carried out in an Eppendorf Mastercycler Gradient. Subsequent amplifications were then performed in a Peltier Thermal Cycler PTC-200, according to the following profile: 5 min at 95°C and 35 cycles of 30 s at 95°C, 2 min at 55°C for K1F/M6R, 59°C for A3F/A7R or 58°C for K1F/M6R and A3F/A7R in multiplex PCR, and 4 min at 72°C, followed by 10 min at 72°C. Amplification products were analyzed by electrophoresis in 1% (w/v) agarose gels stained with ethidium bromide.

Cloning. PCR products were cloned using the TA Cloning kit (Invitrogen). The restriction analysis of the clones were performed using the *Eco*RI, *Pst*I, *Bgl*I, *Bss*HII, and *Mva*I restriction enzymes according to Sambrook and Russel [36] and were separated on 4–20% gradient polyacrylamide gels (Criterion Precast Gel 4–20% TBE, Bio-Rad).

DNA Sequencing. Cloned products were sequenced using universal primers M13R-28 and M13F-20 in an ABI PRISM DYE Terminator cycle sequencing KIT (PerkinElmer).

GenBank Accession Numbers. AY271628-AY271644: clones AAL1, AAL10, AAL14, AAL17, AAL18, AAL19, AAL20, AAL22, AAL24, AAL27, AAL29, AAL3, AAL30, AAL35, AAL7, AAL8, and AAL9; AY271645-AY271658: clones ASH17, ASH21, ASH26, ASH27, ASH31, ASH34, ASH38, ASH4, ASH41, ASH44, ASH53, ASH54, ASH66, and ASH8.

Data Analysis. Sequences were assembled using the GCG Fragment Assembly System. Alignment of the sequences obtained in this study were performed using the multiple alignment program CLUSTAL W [46] and tested against all DNA sequences available in GenBank with the program FASTA.

Results

Design of Degenerate PCR Primers for PKS-I Sequences. The design of degenerate PCR primers specific for PKS-I has been based on the alignment of

known DNA sequences of eleven modular PKS-I biosynthetic clusters characterized in actinomycetes and involved in the synthesis of the compounds rapamycin [3], rifamycin [39], avermectin [19], erythromycin [49], oleandomycin [41], niddamicin [21], pikromycin [51], FK506 [27], FK520 [50], nystatin [9], and pimaricin [4]. The position and extension of the KS and AT domains was deduced from the sequence of the rapamycin PKS-I cluster from *Streptomyces hygroscopicus* NRRL 5491.

From the analysis of the nucleotide sequence alignment of 106 KS domains, we observed a conserved region with enough sequence homology to design the KS-specific degenerate primer K1 (forward) (Fig. 1A). In order to ensure the specificity of the amplification, the design of the reverse primer was focused on the contiguous AT domains normally present in each module. We performed an amino acid sequence alignment of 100 AT domains, including the malonyl-CoA:ACP transacylase (MCAT) from *Escherichia coli* [40]. All the clusters in the analysis contain, at least, one or more methyl-malonyl-CoA transferase domains that can be distinguished at sequence level from malonyl-CoA transferase domains. In fact, around the active sites some conserved regions present divergent sequence motifs that distinguish malonyl-CoA transferase domains from methyl-malonyl-CoA transferase domains [16]. From the analysis of the nucleotide sequence alignment of the AT domains we have identified a conserved region, corresponding to the Gln-250 position, with enough sequence homology to design the degenerate PCR primer M6R specific for the widely distributed methyl-malonyl-CoA transferases (Fig. 1A).

Design of Degenerate PCR Primers Specific for NRPS. Several degenerate oligonucleotide PCR primers targeting conserved regions in the adenylation and thiolation domains have been already reported to detect the presence of NRPS sequences [22]. However, most of these oligonucleotide primers were designed based on known sequences from bacterial taxa such as *Bacillus*, *Lactobacillus*, and *Pseudomonas* with lower G+C content than actinomycetes [24, 33, 48]. In order to maximize the diversity of NRPS genes rescued by PCR amplification, we specifically designed PCR primers for actinomycete NRPS sequences. On the basis of known GenBank DNA sequences of six NRPS biosynthetic clusters characterized in actinomycetes and involved in the synthesis of cephamycin [13], vancomycin [34], balhimycin [30], actinomycin [38], pristinamycin [14], and chloroeremomycin [47], we designed a pair of degenerate PCR primers, A3 (forward) and A7R (reverse), derived from the conserved motifs A3 and A7 previously identified in NRPS adenylation domains [23] (Fig. 1B).

PCR Amplification of NRPS and PKS-I Sequences. Both pairs of primers A3F/A7R and K1F/M6R were tested using PCR annealing temperatures of 55°C and 59°C, respectively, defined as optimal for each case. Amplification products of the expected size ranges were obtained from the six reference actinomycete strains tested that were also known to produce different metabolites (Fig. 2A, B). When the primers were used in a multiplex PCR with these DNAs, the same amplification bands were obtained in all cases (Fig. 2C).

Sequence Analysis of A3F/A7R (NRPS) and K1F/M6R (PKS-I) PCR Products. Two model systems, the ACV gene cluster from *Amycolatopsis lactamdurans* NRRL 3802 [13] and the rapamycin gene cluster from *Streptomyces hygroscopicus* ATCC 29253 [3], were used to test the specificity of the pairs of primers to detect the presence of known NRPS and PKS-I sequences, respectively. The A3F/A7R amplification products derived from *Amycolatopsis lactamdurans* were cloned and 48 randomly chosen clones were analyzed by *Eco*RI–*Hinc*II, *Eco*RI–*Bgl*I, and *Eco*RI–*Pst*I restriction. Clones were grouped into 17 different restriction patterns and a representative clone from each group was sequenced. All the analyzed clones present sequence similarities with known adenylation domains that range between 61% and 99% (Table 1). Two clones (AAL20 and AAL18) contain sequences that correspond to the ACV1 and ACV3 modules of the cephamycin cluster. The remaining 15 clones contain new different adenylation sequences not previously described in this strain. After translation of the nucleotide sequences and alignment with the GrsA adenylation domain [10] we deduced the corresponding position of the amino-acid binding pocket. The amino-acid substrate of five of these 15 new sequences was predicted by comparison to the amino-acid binding pockets of previously reported adenylation domains [10] (Table 2).

Similarly, we cloned the two amplification bands obtained as predicted with the primers K1F/M6R from *Streptomyces hygroscopicus* NRRL 5491 total DNA. Eighty-three randomly chosen clones were analyzed by *Eco*RI–*Bss*HII and *Eco*RI–*Bgl*II restriction and grouped in 14 restriction patterns, and a clone from each group was sequenced. All the sequenced clones show a high similarity with known methyl-malonyl-CoA transferase domains that ranges between 61% and 99%. Four of these sequences (ASh26, ASh38, ASh53, and ASh54) correspond to four methyl-malonyl-CoA transferase domains of the rapamycin cluster (RAPM3, RAPM4, RAPM10, RAPM13). In contrast, the clone ASh41 corresponds to one of the four domains of the second PKS-I gene cluster previously reported in this strain [35]. The remaining nine clones contain new different methyl-malonyl-CoA transferase domains (Table 2).

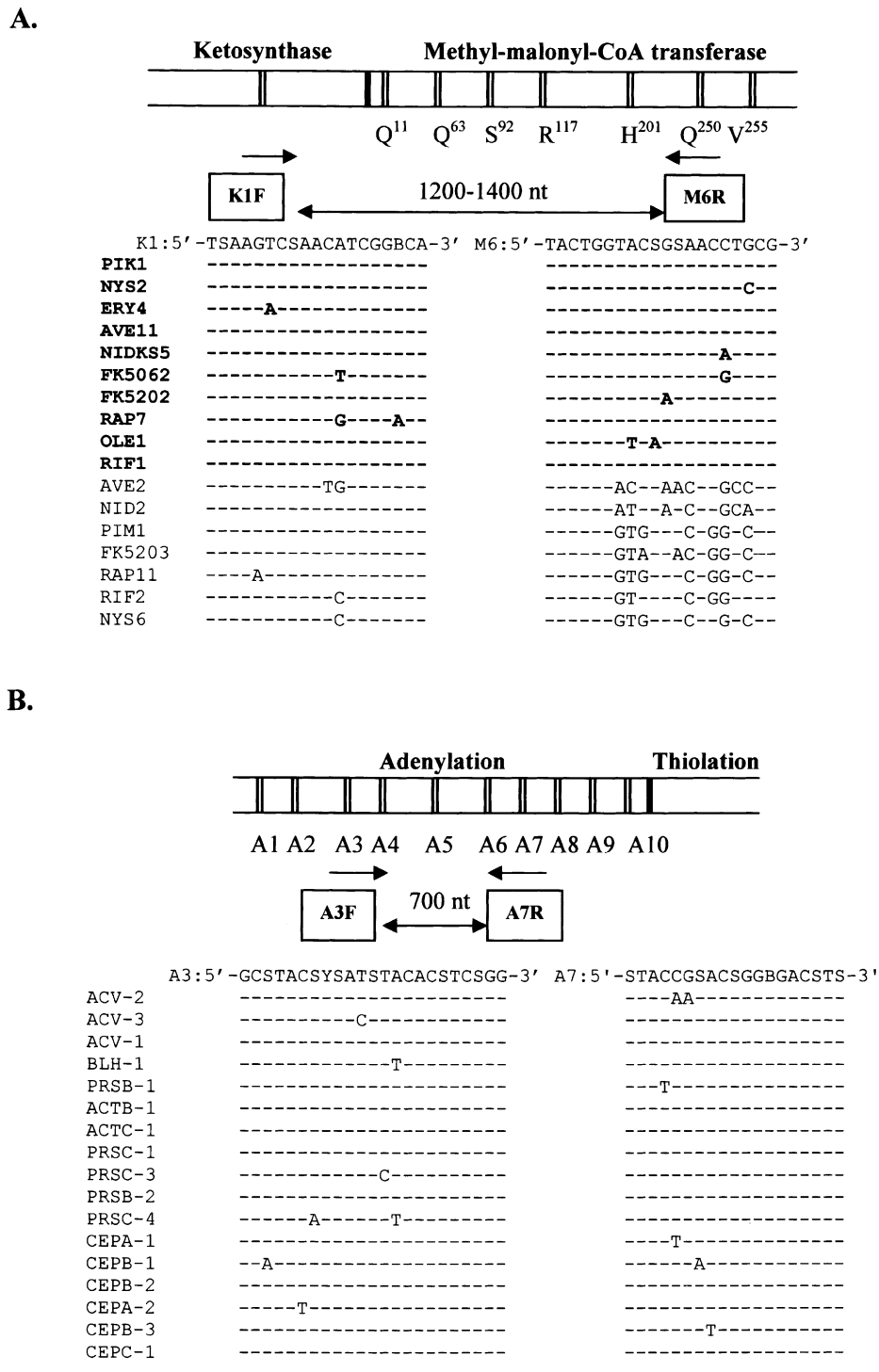

Figure 1. Degenerate primers targeted to NRPS and PKS-I sequences. Conserved motifs are indicated by vertical bars in the diagrams, and mismatches are shown in the aligned sequences. (**A**) PKS-I ketosynthase and methyl-malonyl-CoA transferase modules and annealing sites of the degenerated primers K1F and M6R. The size range of the amplified fragments (1200–1400 bp) depends on the variable interdomain region extension. Representative nucleotide sequence alignments of the methyl-malonyl-CoA transferase (in bold) and malonyl-CoA transferase domains are shown. Abbreviations, PIK: pikromycin; NYS: nystatin; ERY: erythromycin; RAP: rapamycin; OLE: oleandomycin; RIF: rifamycin; NID: niddamycin. (**B**) Conserved motifs in the NRPS adenylation domains and annealing sites of the degenerated primers A3F and A7R that amplify a 700-bp fragment. Representative alignment of the annealing sequences of known adenylation domains. Abbreviations, ACVN: alpha-aminoadipyl-L-cysteinyl-D-valine synthetase; BLH: balhimycin; ACT: actinomycin; PRS: pristinamycin; CEP: chloroeremomycin.

Detection and Distribution of NRPS and PKS-I Systems in Actinomycetes. In order to evaluate the biosynthetic potential of different actinomycete taxonomic groups, we applied our PCR primers to detect the presence of NRPS and PKS-I sequences in a collection of 210 reference strains representative of 33 different genera. NRPS sequences were found to be extensively distributed among these strains (79.5%), whereas PKS-I sequences were only detected in 56.7% of the analyzed strains (Tables 3 and 4).

NRPS and PKS-I sequences were detected in most of the strains of *Streptomyces* tested (97% and 79% respectively), but their occurrence varied very much in the remaining taxa. NRPS genes were present in most of the

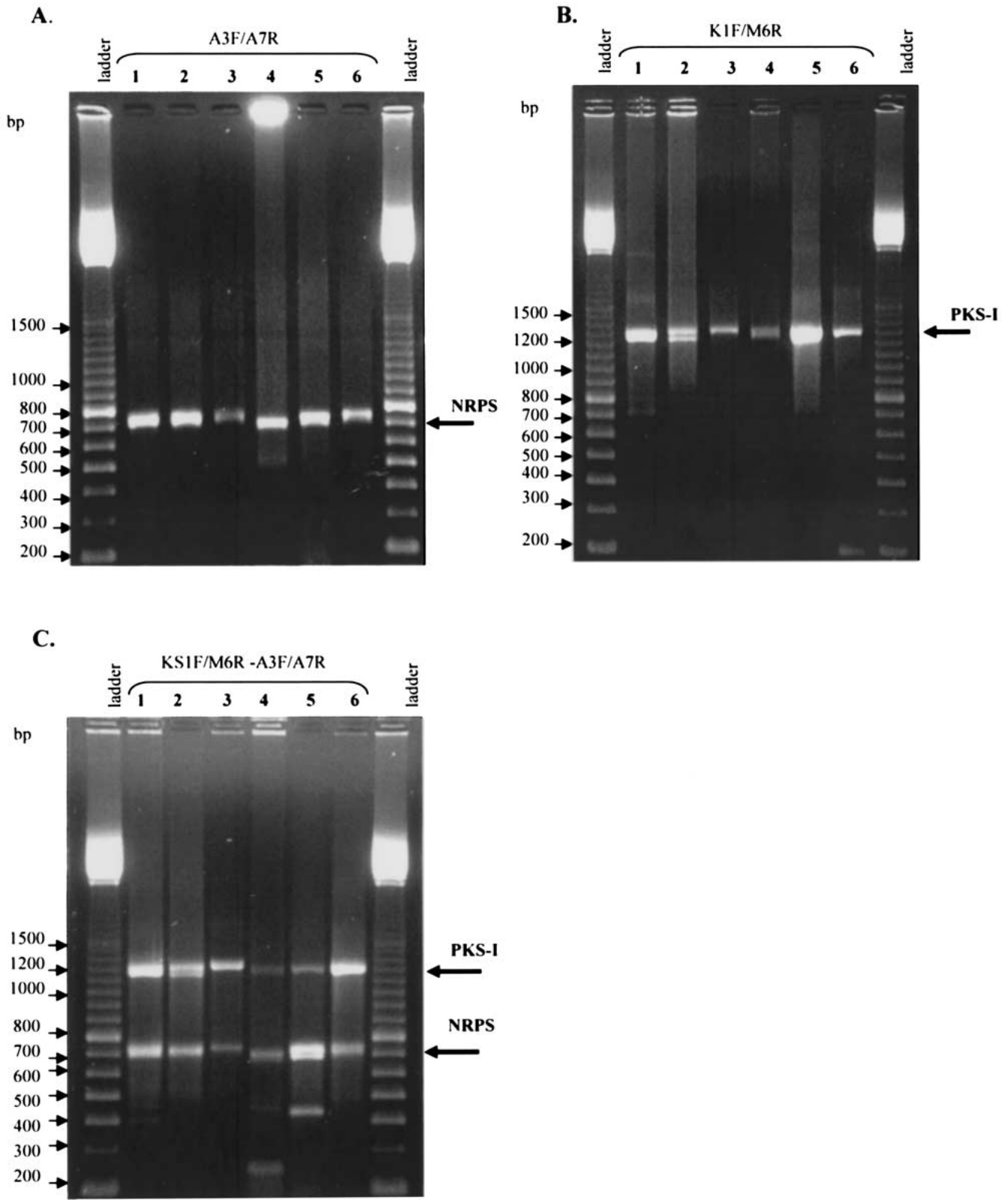

Figure 2. Agarose gel electrophoresis of PCR products from DNA isolated from representative actinomycete strains: (**A**) Selective amplification of 700–800 bp fragments using primers A3F/A7R specific for NRPS adenylation sequences. (**B**) Selective amplification of the 1200–1400 bp fragments using K1F/M6R specific primers for PKS-I ketosynthase and methyl-malonyl-CoA transferase sequences. (**C**) Multiplex PCR amplification with A3F/A7R and K1F/M6R pairs of primers. Lanes, 1: *Amycolatopsis lactamdurans* NRRL 3802; 2: *Streptomyces hygroscopicus* NRRL 5491; 3: *Micromonospora carbonacea* ATCC 27115; 4: *Kibdelosporangium aridum* ATCC 39323; 5: *Saccharopolyspora erythraea* NRRL 2338; 6: *Actinoplanes* sp. ATCC 33076.

Table 1. Characterization of adenylation sequences amplified from *Amycolatopsis lactamdurans* NRRL 3802 genomic DNA using primers A3F/A7R

	Position in amino-acid binding pocket[a]											
Clones	*235*	*236*	*239*	*278*	*299*	*301*	*322*	*330*	*Amino acid*	*Similarity/identity (%)*	*Description*	*Accession no.*
AAl 20	E	P	R	N	L	V	E	F	Aad	99/99	*A. lactamdurans* (ACV-1)	X57310
AAl 18	D	F	E	S	L	A	A	V	Val	99/99	*A. lactamdurans* (ACV-3)	X57310
AAl 29	D	F	W	N	V	G	M	V	Thr	68/57	*S. avermitilis* (Nrpsl-1)	AB070950
AAl 17	D	V	W	E	V	T	A	D	?	61/41	*S. avermitilis* (Nrpsl-3)	AB070950
AAl 8	D	A	W	Q	C	A	T	I	?	73/66	*S. avermitilis* (Nrps2-2)	AB070951
AAl 10	D	L	P	K	V	A	E	V	?	72/68	*S. avermitilis* (Nrps2-3)	AB070951
AAl 27	D	M	G	G	M	G	L	V	?	75/70	*S. avermitilis* (Nrps3-1)	AB070952
AAl 35	D	V	W	H	F	S	L	I	?	72/59	*S. avermitilis* (Nrps3-3)	AB070952
AAl 1	D	T	I	Q	L	G	V	I	Phe	63/44	*S. chrysomallus* (acmC-2)	AF204401
AAl 7	D	I	L	Q	T	G	L	I	Gly	68/55	*S. chrysomallus* (acmC-2)	AF204401
AAl 3	D	I	W	G	F	G	F	V	?	68/55	*S. coelicolor*(CDAPSI)	AL939115
AAl 9	D	L	Y	N	M	S	L	V	Cys	66/53	*S. verticillus* (blmIV)	AF210249
AAl 14	D	A	Y	F	W	G	V	T	Val	73/68	*S. pristinaespiralis* (snbC)	X98690
AAl 19	D	T	W	T	L	G	Y	V	?	62/48	*S. virginiae* (snbDE-2)	Y11547
AAl 22	D	L	Y	D	A	L	D	I	?	67/60	*S. cellulosum* (epoP)	AF210843
AAl 24	D	V	W	H	F	S	L	I	?	70/60	*S. viridochromogenes* (phsB)	Y17268
AAl 30	D	L	Y	D	L	S	G	I	?	66/55	*S. aurantiaca* (MtaC)	AF188287

[a]According to GrsA numbering.
?: Unknown substrate.
The eight-amino acids responsible for the substrate specificity of the adenylation domains were identified aligning the sequences with the PheA specific module of the Gramicidin S biosynthetic system.
Abbreviations, ACV: alpha-aminoadipyl-L-cysteinyl-D-valine synthetase; Nrps: nonribosomal peptide synthase; acm: actinomycin; CDA: calcium-dependent antibiotic; blm: bleomycin; epo: epothilone; phs: phosphinothricin tripeptide synthetase; Mta: myxothiazol.

members of the family *Micromonosporaceae* (50% to 100%), but the occurrence of PKS-I genes is genus dependent, PKS-I amplicons were obtained respectively in 68% and 60% of the *Micromonospora* and *Catenuloplanes* species, but only in 22% of the members of the genus *Actinoplanes*. Furthermore, PKS-I amplification products were not detected in the genus *Catellatospora*.

Among representatives of the *Nocardiaceae* family, NRPS and PKS-I sequences were detected respectively in 91% and 59% of the strains analyzed, in spite of the very low number of non–ribosomally synthesized peptide and polyketide compounds described for this taxon.

The NRPS and PKS-I amplicons were frequently found in the family *Pseudonocardiaceae* (76.3 and 52.6%,

Table 2. Methyl-malonyl-CoA transferase sequences amplified from *Streptomyces hygroscopicus* NRRL 5491 genomic DNA using the primers K1F/M6R

	Position in the methyl-malonyl-CoA binding pocket[a]			
Clones	Q^{63} ... S^{92}	*Similarity/identity (%)*	*Description*	*Accession no.*
ASh 38	VDVVQPASWAVMVSLAAVW...GHSQGEIAA	97/95	*S. hygroscopicus* (rapA-3)[|]	X86780
ASh 26	VDVVQPASWAVMVSLAAVW...GHSQGEIAA	99/98	*S. hygroscopicus* (rapA-4)[|]	X86780
ASh 54	VDVVQPASWAVMVSLAAVW...GHSQGEIAA	99/98	*S. hygroscopicus* (rapB-6)[†]	X86780
ASh 53	VDVVQPASWAMMVSLAAVW...GHSQGEIAA	99/99	*S. hygroscopicus* (rapC-3)[‡]	X86780
ASh 41	VDVVQPVLFTVMVALANVW...GHSQGEIAA	99/98	*S. hygroscopicus* (orf-4)	AF007101
ASh 31	VDVVQPVLFAVMVSLARVW...GHSQGEIAA	61/53	*S. hygroscopicus* (rapA-3)	X86780
ASh 34	VDVVQPALWAVMVSLAEVW...GHSQGEIAA	67/57	*S. avermitilis* (pks5-2)	AB070945
ASh 44	VDVAQPALFAVMVSLAELW...GHSQGEIAA	68/50	*S. avermitilis* (olmA-1)	AB070940
ASh 21	VDVVQPVLWAVMVSLAEVW...GHSQGEIAA	66/48	*A. mediterranei* (rifA-3)	AF040570
ASh 66	TEVLQPVLFAAMVSLAALW...GHSQGEMAA	68/56	*A. mediterranei* (rifA-1)	AF040570
ASh 4	VDVVQPVLFAVMVSLAKVW...GHSQGEIAA	67/54	*M. tuberculosis* (pks 7)	Z85982
ASh 8	VDVVQPVLFAVMVSLAKVW...GHSGEIAA	67/55	*M. megalomicea* (megAII-3)	AX112026
ASh 17	VDVVQPASWAVAVSLAGLW...GHSQGEIAA	65/57	*S. antibioticus* (oleAII-4)	AE007033
ASh 27	VDVVQPALWSVMVSLAALW...GHSQGEIAA	63/52	*A. verrucosospora* (pks-2)	AF411576

[a]According to MCAT numbering.
[|]RAPM3, [|]RAPM4, [†]RAPM10, and [‡]RAPM13.
The predicted amino acid sequences of the 14 clones (ASh) show the divergent motifs that correlate with the methyl-malonyl-CoA substrate specificity [16].
Abbreviations, rap: rapamycin; orf: open reading frame; pks: polyketide synthase; rif: rifamycin; olm: oleandomycin; meg: megalomycin.

Table 3. Distribution of NRPS and PKS-I in actinomycetes

Taxa	*Reported product*	*NRPS PCR amplification*	*PKS-I PCR amplification*
Actinomycetales			
Streptomycetaceae			
Streptomyces			
Streptomyces avermitilis NRRL 8165	Avermectin†	+	+
Streptomyces hygroscopicus NRRL 5491	Rapamycin†	+	+
Streptomyces griseus ATCC 6855		+	+
Streptomyces fradiae ATCC 10745		+	+
Streptomyces thermoviolaceus ATCC 1994		+	−
Streptomyces peucetius NRRL B3826		+	+
Streptomyces caelestis ATCC 15084		+	+
Streptomyces venezuelae ATCC 15439	Pikromycin† / Methymicin†	+	+
Streptomyces ambofaciens ATCC 23877	Spiramycin† / Netropsin	+	+
Streptomyces lividans ATCC 19844		+	+
Streptomyces diastaticus ATCC 3315		+	+
Streptomyces hygroscopicus ATCC 53110		+	+
Streptomyces thermotolerans ATCC 11416	Carbomycin†	+	+
Streptomyces platensis ATCC 13865		+	+
Streptomyces antibioticus ATCC 11891	Oleandomycin†	+	+
Streptomyces virginiae ATCC 13161	Streptogramin §	+	−
Streptomyces setonii ATCC 39116		+	−
Streptomyces lavendulae ATCC 14159		+	−
Streptomyces albidoflavus ATCC 25422		+	+
Streptomyces cinnamoneum ATCC 23897	Carbomycin† / Niddamycin†	+	+
Streptomyces cyaneus ATCC 14923		+	+
Streptomyces canescens ATCC 19736		+	+
Streptomyces coelicolor ATCC 23899		+	+
Streptomyces aeurofaciens MA 6538		+	+
Streptomyces violaceoruber ATCC 14980		+	+
Streptomyces fradiae DSM 41757	Tylosin†	+	+
Streptomyces pristinaespiralis DSM 40338	Pristinamycin §	+	+
Streptomyces natalensis DSM 40357	Pimaricin†	+	+
Streptomyces hygroscopicus var. *ascomyceticus* ATCC 14891	FK520† / Ascomycin†	−	+
Streptomyces sp. ATCC 53770	FK506†	+	+
Streptomyces antibioticus ATCC 8663	Dactinomycin / Actinomycin §	+	−
Streptomyces chattanoogensis ATCC 13358	Natamycin / Pimaricin†	+	+
Streptomyces chrysomallus ATCC 11523	Actinomycin C §	+	−
Micromonosporaceae			
Micromonospora			
Micromonospora carbonacea sub sp. *aurantiaca* ATCC 27115		+	+
Micromonospora sp. ATCC 10026		+	+
Micromonospora inyoensis NRRL 3292		+	+
Micromonospora megalomicea subsp. *niera* NRRL 3274	Megalomicin†	+	+
Micromonospora megalomicea subsp. *niera* NRRL 3275	Megalomicin†	+	+
Micromonospora rosaria ATCC 29337	Rosamicin†	+	+
Micromonospora sagamiensis subsp. *Nonreductans* ATCC 21803		+	+
Micromonospora sp. MA 7190		+	+
Micromonospora lacustris ATCC 21975	Rifamycins complex†	+	+
Micromonospora purpurea ATCC 15835		+	
Micromonospora chalcea ATCC 12452		+	+
Micromonospora sp. ATCC 10026		+	+
Micromonospora echinospora subsp. *ferruginea* ATCC 15836		+	+
Micromonospora echinospora subsp. *pallida* NRLL 2996		+	+
Micromonospora matsumotoensis DSM 44100		+	−
Micromonospora polytrota NRRL 12066		+	−
Micromonospora sp. DSM 43126		+	+
Micromonospora sp. DSM 43170		−	−

(continues)

Table 3. Continues

Taxa	*Reported product*	*NRPS PCR amplification*	*PKS-I PCR amplification*
Micromonospora inositola DSM 43819		+	−
Micromonospora purpureochromogenes DSM 43821		−	−
Micromonospora sp. DSM 43906		−	−
Micromonospora sp. DSM 43907	Actinomycins §	+	+
Micromonospora sp. DSM 43913		−	−
Micromonospora sp. DSM 43918		+	+
Micromonospora halophytica subsp. *niera* ATCC 33088		+	+
Micromonospora purpureochromogenes ATCC 27334		−	−
Micromonospora yulongenesis ATCC 43540		+	+
Micromonospora citrea ATCC 35571		−	+
Micromonospora aurantiaca ATCC 27029		−	+
Micromonospora sp. MA 7094	Rustmicin†	+	+
Micromonospora coerulea ATCC 27008		−	−
Actinoplanes			
Actinoplanes auranticolor ATCC 15330		+	−
Actinoplanes globisporus ATCC 23056		+	−
Actinoplanes sp. ATCC 33076	Ramoplanin §	+	+
Actinoplanes missouriensis ATCC 23342	Actaplanin §	+	−
Actinoplanes awdjinensis subsp. *mycoplanecinus* NRRL 11462		+	−
Actinoplanes caeruleus NRRL 5325		+	+
Actinoplanes deccanensis ATCC 21985		+	−
Actinoplanes italicus NRRL B-16722		+	−
Actinoplanes lobatus ATCC 15550		+	−
Actinoplanes missouriensis ATCC 14538		+	−
Actinoplanes philippiensis NRRL 2506		−	−
Actinoplanes rectiliniatus NRRL B-16090		−	−
Actinoplanes regularis ATCC 31517		+	−
Actinoplanes regularis ATCC 33986		−	−
Actinoplanes teichomyceticus ATCC 31121	Teicoplanins §	+	+
Actinoplanes utahensis NRRL 5614		+	+
Actinoplanes utahensis NRRL 12052		+	−
Actinoplanes yunnaensis NRRL B-2254		+	−
Dactylosporangium			
Dactylosporangium aurantiacum ATCC 23491		+	−
Dactylosporangium vescum ATCC 39499		+	+
Pilimelia			
Pilimelia anulata DSM 43039		−	−
Pilimelia columellifera subsp. *columellifera* DSM 43797		−	−
Pilimelia columellifera subsp. *pallida* DSM 43799		+	−
Pilimelia terevasa DSM 43040		+	+
Catellatospora			
Catellatospora tsunoense DSM 44100		+	−
Catellatospora citrea subsp. *citrea* DSM 44097		+	−
Catellatospora citrea subsp. *methionotrophica* DSM 44098		+	−
Catenuloplanes			
Catenuloplanes japonicus ATCC 31637		+	+
Catenuloplanes atrovinosus ATCC 700015		+	−
Catenuloplanes castaneus ATCC 700016		+	+
Catenuloplanes nepalensis ATCC 700017		+	+
Catenuloplanes crispus DSM 44128		−	−
Couchioplanes			
Couchioplanes caeruleus subsp. *azureus* ATCC 31157		+	−

(continues)

Table 3. Continues

Taxa	*Reported product*	*NRPS PCR amplification*	*PKS-I PCR amplification*
Couchioplanes caeruleus subsp. *caeruleus* ATCC 33937		+	+
Nocardiaceae			
Nocardia			
Nocardia rubra MA 1078		+	+
Nocardia rubra ATCC 13778		+	−
Nocardia asteroides ATCC 9969		+	+
Nocardia asteroides ATCC 9970		+	−
Nocardia farcinica ATCC 6846		+	+
Nocardia sylvodorifera ATCC 4919		+	+
Nocardia asteroides ATCC 7372		−	−
Nocardia asteroides ATCC 10904		+	+
Nocardia asteroides ATCC 8674		+	+
Nocardia uniformis subsp. *tsuyamanensis* ATCC 21806	Nocardicin A §	+	+
Nocardia sp. FERM P4447		+	+
Nocardia sp. ATCC 53695		+	+
Nocardia sp. ATCC 14559		+	+
Nocardia sp. NRRL 5646		+	+
Nocardia sp. NRRL 8050		+	−
Nocardia sp. NRRL 3385		+	−
Rhodococcus			
Rhodococcus rhodochrous ATCC 999		+	−
Rhodococcus rhodochrous ATCC 4273		+	−
Rhodococcus erythropolis DSM 43066		+	−
Rhodococcus sp. ATCC 13258		−	−
Rhodococcus sp. ATCC 13259		+	+
Rhodococcus equi IFO 3730		+	+
Pseudonocardiaceae			
Kutzneria			
Kutzneria albida ATCC 25243		+	+
Amycolatopsis			
Amycolatopsis alba DSM 44262		+	+
Amycolatopsis mediterranei ATCC 13685	Rifamycin†	+	+
Amycolatopsis lactamdurans ATCC 27382	Cephamycin C §	+	+
Amycolatopsis orientalis subsp. *orientalis* ATCC 19795	Vancomycin §	+	+
Amycolatopsis orientalis subsp. *orientalis* ATCC 21425		+	−
Amycolatopsis orientalis ATCC 43491	Vancomycin §	+	+
Amycolatopsis orientalis subsp. *orientalis* DSM 40046	Vancomycin §	+	+
Amycolatopsis azurea DSM 43854		+	−
Amycolatopsis methanolica DSM 44096		+	+
Amycolatopsis coloradensis DSM 44225	Avopartin / Teicoplanins §	+	−
Amycolatopsis japonica DSM 44213		+	+
Amycolatopsis thermoflava DSM 44574		+	+
Amycolatopsis sulphurea ATCC 27624		+	−
Amycolatopsis orientalis subsp. *orientalis* ATCC 21425		+	−
Amycolatopsis orientalis subsp. *orientalis* ATCC 35165		+	−
Amycolatopsis orientalis subsp. *lurida* ATCC 14930	Ristocetin §	+	−
Saccharomonospora			
Saccharomonospora viridis ATCC 15345		+	+
Saccharomonospora caesia ATCC 49227		+	+
Saccharomonospora internatus ATCC 33517		+	+
Saccharomonospora viridis ATCC 15386		+	−
Saccharomonospora glauca DSM 43769		−	−
Saccharopolyspora			
Saccharopolyspora erythraea NRRL 2338	Erythromycin†	+	+

(continues)

Table 3. Continues

Taxa	*Reported product*	*NRPS PCR amplification*	*PKS-I PCR amplification*
Saccharopolyspora spinosa NRRL 18537	Spinosin†	+	+
Saccharopolyspora spinosa NRRL 18395	Spinosin†	+	+
Pseudonocardia			
Pseudonocardia yunnanensis DSM 44253		+	+
Pseudonocardia saturnea NRRL B16172		−	−
Pseudonocardia autotrophica ATCC 35203		+	+
Pseudonocardia autotrophica DSM 43098		−	−
Pseudonocardia autotrophica DSM 43103		−	−
Pseudonocardia autotrophica ATCC 35204		−	−
Pseudonocardia nitrificans DSM 46012		−	+
Pseudonocardia petroleophila DSM 43098		−	−
Pseudonocardia halophobica DSM 43089		−	−
Pseudonocardia compacta ATCC 35407		−	−
Pseudonocardia thermophila ATCC 19285		−	−
Kibdelosporangium			
Kibdelosporangium aridum ATCC 39323		+	+
Nocardiopsaceae			
Nocardiopsis			
Nocardiopsis lucentensis ATCC 51300		−	−
Nocardiopsis antarticus ATCC 43517		+	+
Nocardiopsis dassonvillei subsp. *albirudida* DSM 40465		+	+
Nocardiopsis listeri ATCC 27442		−	+
Thermobifida			
Thermobifida fusta DSM 43792		+	−
Thermobifida alba DSM 43795		+	−
Actinosynnematacea			
Actinokineospora			
Actinokineospora riparia ATCC 49499		+	+
Lechevalieria			
Lechevalieria aerocolonigenes ATCC 23870		+	+
Lechevalieria flava DSM 43885		+	−
Saccharothrix			
Saccharothrix syringae DSM 43886		+	−
Saccharothrix coeruleofusta DSM 43679		+	+
Saccharothrix mutabilis subsp. *capreolus* DSM 40225	Capreomycin §	+	+
Saccharothrix mutabilis subsp. *mutabilis* DSM 43853		+	+
Saccharothrix mutabilis subsp. *mutabilis* ATCC 31520		+	+
Saccharothrix espanensis DSM 44229		+	−
Lentzea			
Lentzea waywayandensis DSM 44232		+	−
Lentzea albidocapillata ATCC 51859		+	−
Actinosynnema			
Actinosynnema pretiosum subsp. *pretiosum* ATCC 31281		+	+
Actinosynnema pretiosum subsp. *pretiosum* NRRL B16060		+	+
Thermomonosporaceae			
Actinomadura			
Actinomadura sp. MA 6867		−	−
Actinomadura hibisca MA 6994		−	−
Actinomadura cremea ATCC 33577		+	+
Actinomadura malachitica ATCC 27888		−	−
Actinomadura formosensis DSM 43997		+	−
Actinomadura echinospora ATCC 27300		+	−
Actinomadura livida ATCC 35578		+	+
Thermomonospora			
Thermomonospora chromogena DSM 43792		+	+
Streptosporangiaceae			
Nonomuraea			

(continues)

Table 3. Continued

Taxa	*Reported product*	*NRPS PCR amplification*	*PKS-I PCR amplification*
Nonomuraea africana DSM 43748		+	−
Nonomuraea pusilla MA 5414		−	−
Nonomuraea fastidiosa ATCC 33516		−	−
Nonomuraea ferruginea ATCC 35575		−	−
Nonomuraea pusilla ATCC 27296		+	+
Nonomuraea roseola ATCC 33579		+	−
Nonomuraea rubra ATCC 27031		+	+
Nonomuraea salmonea ATCC 33580		+	−
Microbispora			
Microbispora rosea subsp. *aerata* DSM 43176		+	−
Microbispora rosea subsp. *rosea* DSM 43164		+	−
Microbispora rosea subsp. *rosea* ATCC 33326		−	−
Microbispora rosea subsp. *rosea* ATCC 12950		−	−
Microtetraspora			
Microtetraspora fusca ATCC 23058		+	−
Microtetraspora glauca ATCC 23057		+	−
Microtetraspora niveoalba ATCC 27301		+	−
Planobispora			
Planobispora rosea ATCC 23866		+	+
Planobispora venezuelensis ATCC 23865		+	−
Streptosporangium			
Streptosporangium roseum ATCC 1 2428		+	+
Streptosporangium viridogriseum CECT 3305		−	−
Streptosporangium vulgare ATCC 33329		−	−
Glycomycetaceae			
Glycomyces			
Glycomyces rutgersensis DSM 43812		+	−
Glycomyces tenuis DSM 44171		−	−
Glycomyces harbinensis DSM 46494		+	−
Geodermatophilaceae			
Geodermatophilus			
Geodermatophilus obscurus DSM 43160		−	+
Geodermatophilus obscurus subsp. *dictyosporus* DSM 43161		−	−
Geodermatophilus obscurus subsp. *utahensis* DSM 43162		−	−

§: NRPS system, †: PKS-I system; only NRPS and PKS-I derived products are included (source: American Type Culture Collection strain database).

respectively) with a high incidence in the genera *Amycolatopsis*, *Saccharopolyspora*, and *Saccharomonospora*. In contrast, NRPS and PKS-I seem to be rare in the genus *Pseudonocardia*, despite the high number of strains analyzed.

Among members of the family *Actinosynnemataceae,* NRPS and PKS-I amplicons were detected respectively in 100% and 61.5% of the strains tested, although few representatives of the genera *Actinokineospora*, *Lechevalieria*, *Saccharothrix*, *Lentzea*, and *Actinosynnema* were evaluated.

The occurrence of NRPS and PKS-I amplicons is low in members of the family *Streptosporangiaceae* (65% and 20%, respectively), and they are rare or even absent among representatives of the families *Nocardiopsaceae, Thermomonosporaceae, Glycomycetaceae,* and *Geodermatophilaceae,* where few strains were examined.

Discussion

In this work we have presented the design and validation of new PCR primers specific for the detection of NRPS and PKS-I sequences in actinomycetes. Effectively, we have shown that all the PCR products obtained with the pair of NRPS primers A3/A7R correspond to adenylation modules, as confirmed from the sequence of clones containing *Amycolatopsis lactamdurans* amplification fragments. Furthermore, we have been able to detect and identify 17 different sequences from the 25 putative NRPS modules previously proposed in this strain from Southern hybridization results [43]. Two of these sequences correspond to the cephamycin ACV1 and ACV3 modules. The lack of amplification of the ACV2 module was probably due to the presence of two mismatches near the 3' end of the A7R primer sequence target (Fig. 1B) [18]. Similarly, all the sequenced clones containing *S. hygroscopicus* amplification fragments obtained with the pair of primers KS1/M6R were identified as ketosynthase/methyl-malonyl-CoA transferase modules. Only five of the 14 different sequences correspond to known PKS-I modules, the four modules RAPM3, RAPM4, RAPM10, and RAPM13 of the rapamycin biosynthetic cluster, and

Table 4. Frequency of positive A3F/A7R and K1F/M6R PCR amplifications in the different actinomycete taxonomic groups

Taxa	*Strains*	*A3F/A7R positive amplifications*	*%*	*K1F/M1R positive amplifications*	*%*
Actinomycetales	**210**	**167**	**79.5**	**119**	**56.7**
Streptomycetaceae	**33**	**32**	**97**	**26**	**78.8**
Streptomyces	33	32	97	26	78.8
Micromonosporaceae	**65**	**50**	**76.9**	**31**	**47.7**
Micromonospora	31	22	70.9	21	67.7
Actinoplanes	18	15	83.3	4	22.2
Dactylosporangium	2	2	100	1	50
Pilimelia	4	2	50	1	25
Catellatospora	3	3	100	0	0
Catenuloplanes	5	4	80	3	60
Couchioplanes	2	2	100	1	50
Nocardiaceae	**22**	**20**	**90.9**	**13**	**59**
Nocardia	16	15	93.8	11	68.8
Rhodococcus	6	5	83.3	2	33.3
Pseudonocardiaceae	**37**	**29**	**76.3**	**20**	**52.6**
Kutzneria	1	1	100	1	100
Amycolatopsis	16	16	100	10	62.5
Saccharomonospora	5	4	80	3	60
Saccharopolyspora	3	3	100	3	100
Pseudonocardia	11	3	27.3	3	27.3
Kibdelosporangium	1	1	100	1	100
Nocardiopsaceae	**6**	**4**	**66.7**	**3**	**50**
Nocardiopsis	4	2	50	3	75
Thermobifida	2	2	100	0	0
Actinosynnemataceae	**13**	**13**	**100**	**8**	**61.5**
Actinokineospora	1	1	100	1	100
Lechevalieria	2	2	100	1	50
Saccharothrix	6	6	100	4	66.7
Lentzea	2	2	100	0	0
Actinosynnema	2	2	100	2	100
Thermomonosporaceae	**8**	**5**	**62.5**	**3**	**37.5**
Actinomadura	7	4	57.1	2	28.6
Thermomonospora	1	1	100	1	100
Streptosporangiaceae	**20**	**13**	**65**	**4**	**20**
Nonomuraea	8	5	62.5	2	25
Microbispora	4	2	50	0	0
Microtetraspora	3	3	100	0	0
Planobispora	2	2	100	1	50
Streptosporangium	3	1	33.3	1	33.3
Glycomycetaceae	**3**	**2**	**66.7**	**0**	**0**
Glycomyces	3	2	66.7	0	0
Geothermatophilaceae	**3**	**0**	**0**	**1**	**33.3**
Geothermatophilus	3	0	0	1	33.3

the module PKSM4 from the second PKS-I system described in this strain [35]. However, the sequences corresponding to the rapamycin modules RAPM1, RAPM6, and RAPM7 or to the modules PKSM1 and PKSM3 from the second PKS-I cluster have not been recovered. With only one exception, all the sequenced clones corresponding to known modules contain perfect matches with the primers KSF/M6R. Interestingly, the rapamycin module RAPM3 presents one mismatch three nucleotides from the 5' end of the primer M6R sequence target. Although the same occurs in modules PKSM1 and PKSM3 from the second PKS-I cluster, none of these two modules were obtained by amplification. The remaining known modules that were not cloned contain at least one mismatch close to the 3' end of the KSF primer sequence target.

We have shown that both pairs of primers A3F/A7R (NRPS) and KSF/M6R (PKS-I) can amplify specifically and exclusively sequences from adenylation and methyl-malonyl-CoA transferase modules present in the actinomycete strains tested. Nevertheless, although the lack of amplification may be an indication of the absence of NRPS or PKS-I systems, we should also consider the occurrence of adenylation or methyl-malonyl-CoA transferase domains with less conserved sequences and therefore lower homology with the primers.

The high number of adenylation domains obtained from *Amycolatopsis lactamdurans* together with their high sequence diversity may suggest the presence of additional NRPS clusters, although no NRPS gene products have been reported in this strain besides those involved in the cephamycin biosynthesis. On the other hand, the sequence homology of the nine new PKS-I sequences cloned from *Streptomyces hygroscopicus* may suggest the occurrence of additional PKS-I clusters in this strain as well. Some of these new PKS-I modules could be related with the nigericin biosynthesis [35], although additional experiments would be necessary to determine if they correspond to related genes associated within a cluster. The occurrence of a high number of NRPS and PKS-I domains has also been found in the complete genome sequences of the actinomycetes *Streptomyces coelicolor* and *Streptomyces avermitilis*. From the eight PKS-I and eight NRPS clusters identified in the *Streptomyces avermitilis* genome, only three PKS-I clusters are known to be involved in the biosynthesis of avermectin, oligomycin, and a polyene macrolide biosynthesis, and no products were detected from the remaining PKS-I and NRPS clusters [29]. Similarly the whole genome of *Streptomyces coelicolor* contains five NRPS and two modular PKS-I clusters, of which only four NRPS clusters are known to be involved in the synthesis of known compounds [6]. The high occurrence of NRPS and PKS-I sequences with no identified secondary metabolite end products has also suggested their potential involvement in primary metabolism. This is the case of the adenylation domains involved in the incorporation of D-alanine into membrane-associated D-alanyl-lipoteichoic acid in *Lactobacillus casei* [17] and *Bacillus subtilis* [31], but not detected in actinomycetes. The involvement of several NRPS in the synthesis of siderophores could also be considered to lie on the borderline between primary and secondary metabolism [43]. In spite of this, none of the NRPS or PKS-I systems described in actinomycetes could suggest their role in primary metabolism. On the other hand, the high number of NRPS and PKS-I modules found in *Amycolatopsis lactamdurans* and *Streptomyces hygroscopicus* could be an evidence of the high potential of actinomycetes for producing a high number of secondary metabolites that may be expressed under different nutrient conditions.

When examining the large collection of reference actinomycetes, we have observed the extensive distribution of NRPS sequences among the actinomycete taxa tested, whereas PKS-I genes were concentrated in fewer genera. The latter results contrast those obtained when targeting only polyketide ketosynthase domains [1] (A. Anderson, personal communication; A. Ayuso-Sacido, unpublished results). In both studies KS genes were detected in 80% to 90% of the strains tested. This discrepancy may originate from the specificity of the primer for methyl-malonyl-CoA domains and from the sequences targeted that span two domains. The occurrence of the recently described AT-less type I PKS clusters [11, 32] could also account for these differences.

Both kinds of sequences occur more frequently in strains of the main producer groups such as the genus *Streptomyces*, or the families *Micromonosporaceae, Pseudonocardiaceae,* and *Actinosynnemataceae.* Within these taxa are included the producing strains of the NRPS compound vancomycin or the polyketide compounds erythromycin or spinosin, all members of the genera *Amycolatopsis* and *Saccharopolyspora*. Surprisingly, the same high incidence was observed in members of the genus *Saccharomonospora*, where NRPS and PKS-I products have not been described. These sequences were also detected in many other nonproducer strains as well as other strains belonging to less frequently isolated actinomycete genera. This is the case of the members of the families *Streptosporangiaceae, Thermomonosporaceae,* and *Nocardiaceae* where a high incidence of NRPS and PKS-I sequences has been detected despite the lack of NRPS and polyketide compounds described as produced by the analyzed strains. Further expression analysis of these biosynthetic systems could determine if these sequences are involved in the production of other metabolites not only in known actinomycete producers but also in other actinomycete strains where NRPS and PKS-I have been detected.

Similar extensive NRPS distributions were previously observed in cyanobacteria [12], and both NRPS and PKS-I systems have also been detected in phylogenetically distant taxa such as *Cryptosporidium parvum* [52] and *Myxococcus xanthus* [5]. There is no doubt that NRPS and PKS-I PCR screening in different bacterial taxa may help in the identification of unexplored groups with unknown genetic potential for the production of interesting metabolites.

Finally, the design of molecular tools for the rapid detection of other genes involved in the synthesis of secondary metabolites such as PKS-II, chalcone synthase (PKS-III), or isoprenoid synthetic genes would allow us to enlarge our previous knowledge of the genetic potential of actinomycetes for the synthesis of secondary metabolites.

Acknowledgments

A. Ayuso was recipient of a pre-doctoral fellowship from the MIT program, Ministerio de Ciencia y Tecnología, Spain. We are grateful to O. Salazar for his help on sequence data analysis and for providing many of the genome DNAs tested from the reference strain collection. We also thank A. Anderson and Z. An for sharing with us their results prior to publication, and especially A.

Anderson for her helpful comments and discussions during the preparation of the manuscript.

References

1. Anderson, AS, An, Z, Strohl, WR (2000) Polyketide antibiotics. In: Lederberg (Ed.) Encyclopedia for Microbiology, vol. 3, Academic Press, San Diego, pp 241–254
2. Anderson, AS, Clark, D, Gibbons, P, Sigmund, J (2002) The detection of diverse aminoglycoside phosphotransferases within natural populations of actinomycetes. J Ind Microbiol Biotechnol 29: 60–69
3. Aparicio, JF, Molnar, I, Schwecke, T, König, A, Haydock, SF, Khaw, LE, Staunton, J, Leadlay, PF (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
4. Aparicio, JF, Fouces, R, Mendes, VM, Oliver, N, Martín, JF (2000) A complex multienzyme system encoded by five polyketide synthase genes is involved in the biosynthesis of the 26-membered polyene macrolide pimaricin in *Streptomyces natalensis.* Chem Biol 7: 895–905
5. Arslanian, RL, Tang, L, Blough, S, Ma, W, Qiu, RG, Katz, L, Carney, JR (2002) A new cytotoxic epothilone from modified polyketide synthases heterologously expressed in *Myxococcus xanthus.* J Nat Prod 65: 1061–1064
6. Bentley, SD, Chater, KF, Cerdeño-Tńrraga, AM, Challis, GL, Thomson, NR, James, KD, Harris, DE, Quail, MA, Kieser, H, Harper, D, Bateman, A, Brown, S, Chandra, G, Chen, CW, Collins, M, Cronin, A, Fraser, A, Goble, A, Hidalgo, J, Hornsby, T, Howarth, S, Huang, CH, Kieser, T, Larke, L, Murphy, L, Oliver, K, O'Neil, S, Rabbinowitsch, E, Rajandream, MA, Rutherford, K, Rutter, S, Seeger, K, Saunders, D, Sharp, S, Squares, R, Squares, S, Taylor, K, Warren, T, Wietzorrek, A, Woodward, J, Barrell, BG, Parkhill, J, Hopwood, DA (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature 417: 141–147
7. Beyer, S, Kunze, B, Silakowski, B, Müller, R (1999) Metabolic diversity in myxobacteria: identification of the myxalamid and the stigmatellin biosynthetic gene cluster of *Stigmatella aurantiaca* Sg a15 and a combined polyketide–(poly)peptide gene cluster from the epothilone producing strain *Sorangium cellulosum* So ce90. Biochim Biophys Acta 1445: 185–195
8. Bingle, LE, Lazarus, CM (1999) Ketosynthase domain probes identify two subclasses of fungal polyketide synthase genes. Fungal Genet Biol 26: 209–223
9. Brautaset, T, Sekurova, ON, Sletta, H, Ellingsen, TE, Stronm, AR, Valla, S, Zotchev, S (2000) Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: analysis of the gene cluster and deduction of biosynthetic pathway. Chem Biol 7: 395–403
10. Challis, GL, Ravel, J, Townsend, CA (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem Biol 7: 211–224
11. Cheng, YQ, Tang, GL, Shen, B (2003) Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. Proc Natl Acad Sci USA 100: 3149–3154
12. Christiansen, G, Dittmann, E, Ordorika, LV, Rippka, R, Herdman, M, Börner, T (2001) Nonribosomal peptide synthase genes occur in most cyanobacterial genera as evidenced by their distribution in axenic strains of the PCC. Arch Microbiol 176: 452–458
13. Coque, JJ, Martin, JF, Calzada, JG, Liras, P (1991) The cephamycin biosynthetic genes pcbAB, encoding a large multidomain peptide synthetase, and pcbC of *Nocardia lactamdurans* are clustered together in an organization different from the same genes in *Acremonium chrysogenum* and *Penicillium chrysogenum.* Mol Microbiol 5: 1125–1133
14. de Crecy-Lagard, V, Saurin, W, Thibaut, D, Gil, P, Naudin, L, Crouzet, J, Blanc, V (1997) Streptogramin B biosynthesis in *Streptomyces pristinaespiralis* and *Streptomyces virginiae*: molecular characterization of the last structural peptide synthetase gene. Antimicrob Agents Chemother 41: 1904–1909
15. Donadio, S, Katz, L (1992) Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythraea.* Gene 111: 51–60
16. Haydock, SF, Aparicio, JF, Molnar, I, Schwecke, T, Khaw, LE, König, A, Marsden, AF, Galloway, IS, Staunton, J, Leadlay, PF (1995) Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases. FEBS Lett 374: 246–248
17. Heaton, MP, Neuhaus, FC (1992) Biosynthesis of D-alanyl-lipoteichoic acid: cloning, nucleotide sequence, and expression of the *Lactobacillus casei* gene for D-alanine-activating enzyme. J Bacteriol 174: 4707–4717
18. Igloi, GL (1998) Variability in the stability of DNA-peptide nucleic acid (PNA) single-base mismatched duplexes: real-time hybridization during affinity electrophoresis in PNA-containing gels. Proc Natl Acad Sci USA 95: 8562–8567
19. Ikeda, H, Nonomiya, T, Usami, M, Ohta, T, Omura, S (1999) Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis.* Proc Natl Acad Sci USA 96: 9509–9514
20. Innis, MA, Gelfand, DH, Sninsky, JJ, White, TJ (1990) PCR Protocols. A Guide to Methods and Applications. Academic Press San Diego
21. Kakavas, SJ, Katz, L, Stassi, D (1997) Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis.* J Bacteriol 179: 7515–7522
22. Khosla, C, Gokhale, RS, Jacobsen, JR, Cane, DE (1999) Tolerance and specificity of polyketide synthases. Annu Rev Biochem 68: 219–253
23. Konz, D, Marahiel, MA (1999) How do peptide synthases generate structural diversity? Chem Biol 6: 39–48
24. Lee, SY, Rhee, SK, Kim, CH, Suh, JW (1998) Rapid and efficient isolation of genes for biosynthesis of peptide antibiotics from gram-positive bacterial strain. J Microbiol Biotechnol 8: 310–317
25. Marahiel, MA (1997) Protein templates for the biosynthesis of antibiotics. Chem Biol 4: 561–567
26. Mootz, HD, Schwarzer, D, Marahiel, A (2002) Biosynthetic strategies of nonribosomal peptide synthetases. Chem Bio Chem 3: 490–504
27. Motamedi, H, Shafiee, A (1998) The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506. Eur J Biochem 256: 528–534
28. Nicholson, TP, Dawson, M, Lazarus, CM, Simpson, TJ, Cox, RJ (2001) Design and utility of oligonucleotide gene probes for fungal polyketide synthases. Chem Biol 8: 157–178
29. Omura, S, Ikeda, H, Ishikawa, J, Hanamato, A, Takahashi, C, Shinose, M, et al. (2001) Genome sequence of an industrial microorganism *Streptomyces avermitilis:* deducing the ability of producing secondary metabolites. Proc Natl Acad Sci USA 98: 12215–12220
30. Pelzer, S, Reichert, W, Huppert, M, Heckmann, D, Wohlleben, W (1997) Cloning and analysis of a peptide synthetase gene of the balhimycin producer *Amycolatopsis mediterranei* DSM 5908 and development of a gene disruption/replacement system. J Biotechnol 56: 115–128

31. Perego, M, Glaser, P, Minutello, A, Strauch, MA, Leopold, K, Fisher, W (1995) Incorporation of D-alanine into lipoteichoic acid and wall teichoic acid in *Bacillus subtilis*. J Biol Chem 26: 15598–15606
32. Piel, J (2002) A polyketide synthase–peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. Proc Natl Acad Sci USA 99: 14002–14007
33. Rajendran, N (1999) Identification and cloning of a gene locus encoding peptide synthase of *Pseudomonas fluorescens* by two sets of PCR primers. Z Naturforsch 54: 105–109
34. Recktenwald, J, Shawky, R, Puk, O, Pfennig, F, Keller, U, Wohlleben, W, Pelzer, S (2002) Nonribosomal biosynthesis of vancomycin-type antibiotics: a heptapeptide backbone and eight peptide synthetase modules. Microbiology 148: 1105–1118
35. Ruan, X, Stassi, D, Lax, SA, Katz, L (1997) A second type-I PKS gene cluster isolated from *Streptomyces hygroscopicus* ATCC 29253, a rapamycin-producing strain. Gene 203: 1–9
36. Sambrook, J, Russel, DW (2001) Molecular Cloning: A Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
37. Sauser, M, Lu, P, Sangari, R, Kennedy, S, Polishook, J, Bills, G, An, Z (2002) Estimating polyketide metabolic potential among nonsporulating fungal endophytes of *Vaccinium macrocarpon*. Mycol Res 106: 460–470
38. Schauwecker, F, Pfennig, F, Schroder, W, Keller, U (1998) Molecular cloning of the actinomycin synthetase gene cluster from *Streptomyces chrysomallus* and functional heterologous expression of the gene encoding actinomycin synthetase II. J Bacteriol 180: 2468–2474
39. Schupp, T, Toupet, C, Nathalie, E, Goff, S (1998) Cloning and sequence analysis of the putative rifamycin polyketide synthase gene cluster from *Amycolatopsis mediterranei*. FEMS Microbiol Lett 159: 201–207
40. Serre, L, Verbree, EC, Dauter, A, Stuitje, AR, Derewenda, ZS (1995) The *Escherichia coli* malonyl-CoA:acyl carrier protein transacylase at 1.5 Å resolution. J Biol Chem 270: 12961–12964
41. Shah, S, Xue, Q, Tang, L, Carney, JR, Betlach, M, McDaniel, R (2000) Cloning, characterization and heterologous expression of a polyketide synthase and P-450 oxidase involved in the biosynthesis of the antibiotic oleandomycin. J Antibiot 53: 502–508
42. Sigmund, JM, Clark, DJ, Rainey, FA, Anderson, AS (2003) Detection of eubacterial 3-hydroxy-3-methylglutaryl coenzyme A reductases from natural populations of actinomycetes. Microb Ecol 46: 106–112
43. Sosio, M, Bossi, M, Bianchi, A, Donadio, S (2000) Multiple peptide synthase gene clusters in actinomycetes. Mol Gen Genet 264: 213–221
44. Stachelhaus, T, Marahiel, MA (1995) Modular structure of genes encoding multifunctional peptide synthetases required for nonribosomal peptide synthesis. FEMS Microbiol Lett 125: 3–14
45. Stahl, DA, Amann, R (1991) Development and application of nucleic acid probes. In: Stackebrandt, E, Goodfellow, M (Eds.) Nucleic Acid Techniques in Bacterial Systematics, Wiley, Chichester, UK, pp 205–248
46. Thompson, JD, Higgins, DG, Gibson, TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680
47. Trauger, JW, Walsh, CT (2000) Heterologous expression in *Escherichia coli* of the first module of the nonribosomal peptide synthetase for chloroeremomycin, a vancomycin-type glycopeptide antibiotic. Prac Natl Acad Sci USA 97: 3112–3117
48. Turgay, K, Marahiel, MA (1994) A general approach for identifying and cloning peptide synthetase genes. Pept Res 7: 238–241
49. Weissman, KJ, Bycroft, M, Staunton, J, Leadlay, F (1998) Origin of starter units for erythromycin biosynthesis. Biochemistry 37: 11012–11017
50. Wu, K, Chung, L, Revill, WP, Katz, L, Reeves, CD (2000) The FK520 gene cluster of *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. Gene 251: 81–90
51. Xue, Y, Wilson, D, Sherman, DH (2000) Genetic architecture of the polyketide synthases for methylmycin and pikromycin series macrolides. Gene 245: 203–211
52. Zhu, G, LaGier, MJ, Stejskal, F, Millership, JJ, Cai, X, Keithly, JS (2002) *Cryptosporidium parvum*: the first protist known to encode a putative polyketide synthase. Gene 298: 79–89