J Ind Microbiol Biotechnol (2009) 36:1425–1434
DOI 10.1007/s10295-009-0629-5

ORIGINAL PAPER

Identification of functionally clustered nystatin-like biosynthetic genes in a rare actinomycetes, *Pseudonocardia autotrophica*

Byung-Gyun Kim · Mi-Jin Lee · Jiyoon Seo · Young-Bin Hwang · Mi-Yeon Lee · Kyuboen Han · David H. Sherman · Eung-Soo Kim

Received: 23 June 2009 / Accepted: 27 July 2009 / Published online: 21 August 2009

Abstract The polyene antibiotics, including nystatin, pimaricin, amphotericin, and candicidin, comprise a family of very valuable antifungal polyketide compounds, and they are typically produced by soil actinomycetes. Previously, using a polyene cytochrome P450 hydroxylase-specific genome screening strategy, *Pseudonocardia autotrophica* KCTC9441 was determined to contain genes potentially encoding polyene biosynthesis. Here, sequence information of an approximately 125.7-kb contiguous DNA region in five overlapping cosmids isolated from the *P. autotrophica* KCTC9441 genomic library revealed a total of 23 open reading frames, which are presumably involved in the biosynthesis of a nystatin-like compound tentatively named NPP. The deduced roles for six multimodular polyketide synthase (PKS) catalytic domains were found to be highly homologous to those of previously identified nystatin biosynthetic genes. Low NPP productivity suggests that the functionally clustered NPP biosynthetic pathway genes are tightly regulated in *P. autotrophica*. Disruption of a NPP PKS gene completely abolished both NPP biosynthesis and antifungal activity against *Candida albicans*, suggesting that polyene-specific genome screening may constitute an efficient method for isolation of potentially valuable previously identified polyene genes and compounds from various rare actinomycetes widespread in nature.

Keywords Polyene · Polyketide · *Pseudonocardia* · Nystatin · Antifungal

Introduction

The polyene antifungal antibiotics, produced principally by gram-positive soil actinomycetes, comprise a family of type I polyketide macrolide ring compounds with 20- to 40-carbon backbones, containing three to eight conjugated double bonds [15, 40]. The primary antifungal mechanism by which these polyene antibiotics exert their effects is believed to involve specific binding to the ergosterol present in the fungal membrane, and the formation of channels that allow for the leakage of cellular K^+ and Mg^{2+}, eventually culminating in the death of the fungal cell [4, 35]. Although polyene compounds are limited with regard to their clinical use due largely to their high toxicity and side effects, owing to their superior antifungal activities, these compounds are still being considered in the further development of improved antifungal drugs [6, 8, 11, 15, 30]. Recently, polyene biosynthetic gene clusters from nystatin, amphotericin, pimaricin, and

B. G. Kim and M. J. Lee contributed equally.

Electronic supplementary material The online version of this article (doi:10.1007/s10295-009-0629-5) contains supplementary material, which is available to authorized users.

B.-G. Kim · M.-J. Lee · J. Seo · Y.-B. Hwang · M.-Y. Lee · E.-S. Kim (✉)
Department of Biological Engineering,
Inha University, Incheon 402-751, Korea
e-mail: eungsoo@inha.ac.kr

K. Han
Hanson Biotech Co., Ltd. 201 IACRI,
Han Nam University, Daejeon 306-791, Korea

D. H. Sherman
Life Sciences Institute and Department of Medicinal Chemistry,
University of Michigan, Ann Arbor, MI 48109-2216, USA

candicidin have been cloned and characterized [1–3, 5, 9, 10, 39]. Based on the complete sequences of polyene biosynthetic genes, highly similar multi-modular (or type I) polyketide synthase (PKS) genes and post-PKS modification genes have been identified in the clusters [31]. The polyene type I PKSs are organized into repeated units called modules, each of which is responsible for one condensation cycle in the synthesis of a polyketide chain. Each module contains various activity domains, such as a β-ketoacyl synthase (KS) domain responsible for carbon condensation, as well as several domains determining the reduced state of the incorporated extender unit, including β-ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) domains [14, 19, 25]. The acyltransferase (AT) and acyl carrier protein (ACP) domains present in each module are believed to be responsible for the choice of extender unit and retention of the growing polyketide chain on the PKS, respectively. Upon completion of synthesis, the polyketide chain is released from the PKS via the action of a thioesterase (TE), which is likely also involved in cyclization of the final product. In order to completely biosynthesize polyene compounds, formation of the polyketide backbone should be followed by post-PKS modifications, including glycosylation and regiospecific hydroxylation with cytochrome P450 hydroxlases (CYPs) [24, 26, 28, 29]. Genes encoding polyene-specific CYPs have been located in all of the previously characterized polyene gene clusters, including those for nystatin, amphotericin, pimaricin, and candicidin [1, 5, 9, 10, 24, 26, 29, 39]. Previously, we found highly conserved regions specific to only polyene CYP genes and reported a novel genomics-guided approach for the isolation of a previously unknown polyene cluster using polyene CYP-specific PCR screening [17, 23, 37]. This genomics-guided polyene screening method successfully led to the isolation of a rare actinomycetes strain called *Pseudonocardia autotrophica* KCTC9441 containing some previously unknown polyene genes in the chromosome [23]. Here we report the isolation, characterization, and functional expression of the entire biosynthetic genes of an approximately 125.7-kb contiguous DNA region in five overlapping cosmids isolated from the *P. autotrophica* KCTC9441 genomic library. The deduced roles for six multi-modular type I PKS catalytic domains were found to be highly similar compared to previously identified nystatin biosynthetic genes. The production of Nystatin-like *Pseudonocardia* Polyene (tentatively named NPP) by *P. autotrophica* KCTC9441 was confirmed by a HPLC assay, antifungal bioassay, and gene disruption, suggesting that the polyene-specific genome screening approach may constitute an efficient method for isolation of potentially valuable polyene biosynthetic gene clusters from various rare actinomycetes in nature.

Methods

Bacterial strains, plasmids, and cultivation conditions

P. autotrophica KCTC9441, purchased from the Korean Collection for Type Cultures (KCTC), was grown routinely in GSMY medium (glucose 0.7%, yeast extract 0.45%, malt extract 0.5%, soluble starch 1.0%, and calcium carbonate 0.005%) at 30°C for sporulation [20]. *P. autotrophica* KCTC9441 spores were resuspended and stored in a sterile 20% glycerol solution at −20°C. For total DNA isolation, spore suspensions were inoculated into 25 ml of YEME liquid media and cultured for 2 days at 30°C. The total DNA isolation method was previously described elsewhere [20]. *Escherichia coli* DH5α strain was primarily used for DNA cloning and plasmid propagation. *E. coli* XL-1 Blue MR strain was used for cosmid library construction. *E. coli* ET12567/pUZ8002 (*dam*$^-$, *dcm*$^-$, *hrdM*) was used as the transient host for *E. coli–Streptomyces* conjugation. All *E. coli* strains were cultured at 37°C in Luria broth or on Luria agar supplemented with the appropriate antibiotics when needed [20, 32].

Cloning and sequence analysis of the NPP gene cluster

Total genomic DNA of *P. autotrophica* KCTC9441 grown on GSMY medium was prepared as described [20]. A cosmid library was prepared using *P. autotrophica* KCTC9441 genomic DNA partially digested with *Sau*3AI and a commercially available Supercos-1 cosmid system (Stratagene, USA) according to the manufacturer's protocol. The cosmid library was initially screened by polymerase chain reaction (PCR) using polyene CYP-specific primer pairs, and also the generated positive cosmid clone (named pESK601) containing 34.5 kb of the *npp* central region in *P. autotrophica* KCTC9441 [23]. Using each end fragment of pESK601 as a probe, four additional overlapping cosmid clones (named pESK611, 612, 621, and 622) were further isolated and completely sequenced at Genotech (Daejeon, Korea). DNA sequences were assembled using BLAST searches on the National Center for Biotechnology Information (NCBI) server, and also analyzed using the web-based program Frameplot 2.3.2 (http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl). In addition, PKS amino acid sequence domain data were also analyzed using the Modular Polyketide Synthase Database (http://linux1.nii.res.in/~pksdb/DBASE/page.html). DNA sequence data reported here was deposited in GenBank under the accession number EU108007.

HPLC analysis for NPP production and antifungal bioassay against *Candia albicans*

Extracts from *P. autotrophica* KCTC9441 cultured at 28°C in triple-baffled flasks containing 250 ml YEME for 5 days

were prepared by extraction with an equal volume of buthanol, followed by concentration and methanol resuspension. A Shimadzu SPD M10A (Shimadzu, Japan) with a reverse-phase C-18 column (5 µm particles, 4.6 × 150 mm, Phenomenex, USA) was used for the assay. The sample injection volume was 20 µl, and the run time was fixed at 20 min [18]. The column was kept at 25°C, and an isocratic elution system was maintained with 0.05 M ammonium acetate (pH 3.8) mixed with acetonitrile at a ratio of 65:35. The flow rate was kept constant at 1.0 ml/min, and the polyene compounds were detected using a photo diode array (PDA) detector [18]. The major NPP peak sample for MS analysis was prepared by dissolving purified compounds in MeOH to a final concentration of approximately 1 mg/µl. The MS analysis was performed using the Varian 1,200 L Quadruple LC/MS system. The Q mass spectrometer was operated with the electrospray ionization source in positive ionization mode. Drying gas flow was 0.4 ml/min, drying gas pressure 20 psi, and nebulizer pressure was 51 psi. Drying gas temperature was 300°C, and the fragmentor (capillary) voltage was 52 V. NPP production was assayed using YM medium (yeast extract 0.3%, malt extract 0.3%, peptone 0.5%, glucose 1.0%) and a paper disc-agar diffusion method using 12 *Candida albicans* as the indicator [21].

Inactivation of a putative NPP biosynthetic gene

The *nppA* encoding NPP PKS gene was inactivated using a temperature-sensitive suicide vector, pKC1139 [20]. A 6,548-bp *Sph*I–*Eco*RI fragment containing *nppN, nppDII, nppDI*, and *nppA* genes was isolated from the pESK601 cosmid, and a thiostrepton (*tsr*) resistance gene was subsequently inserted into the *nppA*, followed by ligation into pKC1139, resulting in pConstV (Fig. 4). The pConstV was first introduced into *E. coli* ET12567/pUZ8002 via electroporation, followed by introduction into *P. autotrophica* KCTC9441 by conjugation [20]. After incubation at 28°C for 16 h, each plate was overlaid with 1 ml of sterile water containing both thiostrepton and apramycin at a final concentration of 50 µg/ml each. Positive conjugants were then further incubated at 37°C to induce homologous recombination. One out of 48 exconjugants growing on only GSMY/*tsr* (not on the GSMY/*tsr* + *apr* plate), which presumably went through homologous double crossover recombination, was selected and confirmed via Southern blot hybridization. Total DNAs from *P. autotrophica* KCTC9441 and *P. autotrophica* KCTC9441-001 were completely digested with *Nar*I, and then hybridized using two PCR-amplified probes: 1,073 bp *tsr*-specific (forward primer: ggatcctgatcaaggcgaatac, reverse primer: atgcattgatcatcactgacg) and 550 bp *nppDII*-specific (forward primer: gcagtgatgatcttgttggc, reverse primer: agggagtcacgatgtccttc) probes.

Results and discussion

Cloning and sequencing of the NPP biosynthetic gene cluster in *P. autotrophica* KCTC9441

Previously, a *P. autotrophica* KCTC9441 cosmid clone named pESK601 containing an approximately 35-kb insert DNA was isolated and completely sequenced, revealing a total of seven complete and two partial open reading frames (ORFs) [23]. A DNA database search revealed that they were a part of a unique polyene biosynthetic gene cluster, but highly homologous to a previously characterized nystatin cluster in *S. noursei* [5]. These ORFs were assigned as *npp* (Nystatin-like *Pseudonocardia* Polyene) genes with a nystatin-like nomenclature, thus giving *nppI*, *nppJ*, *nppK*, *nppL*, *nppN*, *nppDII*, *nppDI*, *nppA*, and *nppB*. To isolate the entire NPP biosynthetic gene cluster, more than 1,400 cosmids from the *P. autotrophica* KCTC9441 genomic DNA library were further screened using pESK601 as a probe, providing four additional overlapping cosmid clones (named pESK611, pESK612, pESK621, and pESK622). Complete sequencing of these positive cosmids (a total of 125.7-kb contiguous DNA) revealed six open reading frames (ORFs) encoding typical type I modular PKS genes as well as 17 ORFs for post-PKS modification, mycosamine biosynthesis, export, and regulation (Table 1; Fig. 1).

Organization of the overall NPP-encoded PKS genes and catalytic domains

Biosynthesis of the NPP polyketide backbone structure is believed to be catalyzed by six ORFs (*nppA, nppB, nppC, nppI, nppJ,* and *nppK*) that encode multi-modular PKSs. Because all six deduced products were shown to share considerable identity with nystatin PKSs, a presumptive functional analysis of the NPP PKSs was based on a comparison to the nystatin polypeptides [39]. Two three-PKS-gene units, separated by a 5,219 bp DNA fragment including *nppL, nppN, nppDI,* and *nppDII*, have a total of 19 PKS modules: a single loading module in *nppA*, 2 modules in *nppB*, 6 modules in *nppC*, 6 modules in *nppI*, 3 modules in *nppJ*, and a single termination module in *nppK*. Within the 19 PKS modules are various functional domains including 19 KS/AT/ACP domains, 18 KR domains, 10 DH domains, 2 ER domains, and 1 TE domain (Table 2; Fig. 2). All active site motifs within individual catalytic domains were highly conserved in each of the 19 NPP PKS modules [38].

Organization of KS/AT/ACP domains in NPP PKS genes

All 19 KS/AT/ACP domains in NPP PKS modules possess the expected set of conserved consensus sequences. It has

Table 1 Sizes, probable functions, and similarities of 23 complete ORFs present in a polyene gene cluster in *P. autotrophica* KCTC9441

Protein	aa	Proposed function	Homologs, origins	Identities, positives (%)
NppF	224	Phosphopantetheinyl transferase	NysF, *Streptomyces noursei*	63.73
NppG	593	ABC transporter	NysG, *Streptomyces noursei*	75.83
NppH	585	ABC transporter	NysH, *Streptomyces noursei*	76.84
NppDIII	345	GDP-mannose-4,6-dehydratase	NysDIII, *Streptomyces noursei*	89.93
NppI	9500	Type I PKS	NysI, *Streptomyces noursei*	73.80
NppJ	5426	Type I PKS	AmphJ, *Streptomyces noursei*	61.67
NppK	2032	Type I PKS	NysK, *Streptomyces noursei*	64.69
NppL	398	P450 mono oxygenase	NysL, *Streptomyces noursei*	68.77
NppN	414	P450 mono oxygenase	NysN, *Streptomyces noursei*	74.83
NppDII	353	Aminotransferase	NysDII, *Streptomyces noursei*	82.90
NppDI	490	Glycosyltransferase	NysDI, *Streptomyces noursei*	79.85
NppA	1098	Type I PKS	NysA, *Streptomyces noursei*	62.69
NppB	3176	Type I PKS	NysB, *Streptomyces noursei*	71.78
NppC	10877	Type I PKS	NysC, *Streptomyces noursei*	71.77
NppE	275	Thioesterase	Discrete thioesterase, *Streptomyces nodosus*	77.83
NppRI	947	Regulation	NysRI, *Streptomyces noursei*	45.58
NppRII	766	Regulation	FscRIV, *Streptomyces sp. FR-008*	35.44
NppRIII	926	Regulation	NysRIII, *Streptomyces noursei*	38.48
NppO	527	Decarboxylase	Putative acyl CoA decarboxylase, *Streptomyces griseus*	77.85
NppRIV	214	Regulation	ORF4, *Streptomyces noursei*	45.66
NppM	139	Ferredoxin	4Fe-4S ferredoxin, *Kineococus radiotolerans SRS30216*	35.45
NppRV	964	Regulation	FsCRII, *Streptomyces sp. FR-008*	35.46
NppRVI	685	Regulation	FsCRII, *Streptomyces sp. FR-008*	32.42

[a] Amino acid identity in percent

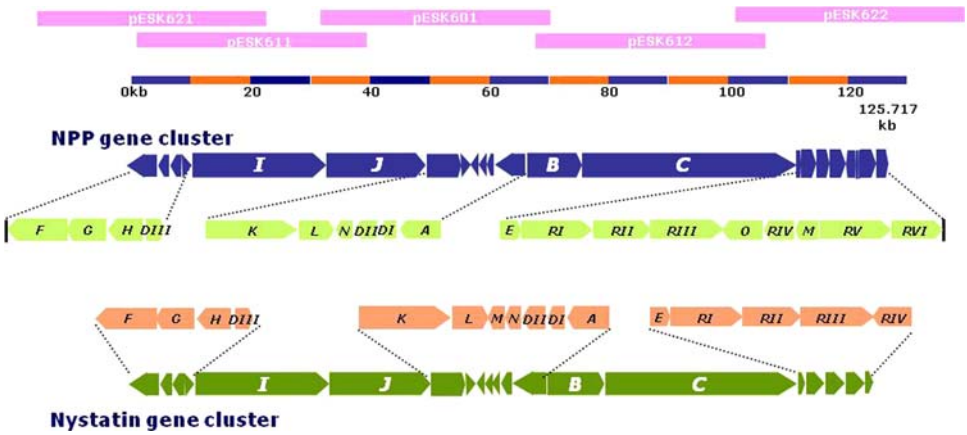

Fig. 1 Gene organization and comparison of the *P. autotrophica* KCTC9441 NPP biosynthetic gene cluster with gene organization in the *S. noursei* nystatin biosynthetic gene cluster. The inserts from the overlapping cosmids encompassing the cloned region are shown above the physical/genetic map. *npp* and *nys* genes are designated with *capital letters* in *italics*

been reported that the conserved amino acid sequence pattern in AT domains determines the extender unit specificity, such as malonyl-specific AT (AT) and methylmalonyl-specific AT (mAT) [13, 16, 33]. Based on the conserved sequence pattern analysis, all 19 AT domains in NPP PKS genes were classified; AT domains in modules 1, 2, and 11 for methylmalonyl extenders and AT domains in the remaining modules for malonyl extenders (Supplementary Fig. 1). The order of AT domains in NPP PKS was identical to the selection sequence of the expected extender unit for the nystatin polyene structure (Supplementary Fig. 1).

The NppA encoding a loading module contains a KS domain (KS^S) similar to that found in the presumed loading module of the nystatin PKS NysA [5]. Like other polyene loading modules such as NysA, AmphA, and PimS0, NppA is also unusual in the sense that it represents the loading module only, whereas most type I PKS loading modules are fused to the first condensing module in multimodular

Table 2 Sequence coordinates of the NPP PKS gene cluster

ORF	Proposed function of sequence similarities detected						
nppA							
Loading module	KS	AT				ACP	
nppB							
Module 1	KS	mAT			KR	ACP	
Module 2	KS	mAT			KR	ACP	
nppC							
Module 3	KS	AT	DH		KR	ACP	
Module 4	KS	AT	DH		KR	ACP	
Module 5	KS	AT	DH	ER	KR	ACP	
Module 6	KS	AT	DH		KR	ACP	
Module 7	KS	AT	DH		KR	ACP	
Module 8	KS	AT	DH		KR	ACP	
nppI							
Module 9	KS	AT	DH		KR	ACP	
Module 10	KS	AT			KR	ACP	
Module 11	KS	mAT			KR	ACP	
Module 12	KS	AT			KR	ACP	
Module 13	KS	AT			KR[a]	ACP	
Module 14	KS	AT			KR	ACP	
nppJ							
Module 15	KS	AT	DH	ER	KR	ACP	
Module 16	KS	AT			KR	ACP	
Module 17	KS	AT	DH[a]		KR	ACP	
nppK							
Module 18	KS	AT	DH[a]		KR	ACP	TE

[a] Predicted to be inactive

ACP acyl-carrier protein, *AT* acyltransferase incorporating malonate, *DH* β-hydroxy-acyl-ACP dehydratase, *ER* enoylreductase, *KR* β-ketoacyl-ACP reductase, *KS* β-ketoacyl ACP synthase, *mAT* acyltransferase incorporating methylmalona

polypeptides [5, 9]. The conserved active-site cysteine residues in NppA, NysA, AmphA, and PimS0 KS domains are replaced with serine residues. In all other known inactive KS^S (KS^Q) in the loading modules of type I PKSs, the conserved cysteine is replaced by glutamine, except for the epothilone PKS loading module KS, which bears a Cys → Tyr replacement [5]. Although it was previously speculated that the Cys → Ser replacement in the active sites of KS^S domains might have some significance for NysA functioning as separate loading polypeptides [39], this aspect needs to be further characterized.

Organization of KR/DH/ER domains in NPP PKS genes

Within 19 NPP PKS modules, 18 KR domains, 10 DH domains, and 2 ER domains were identified. The NPP KR domains exhibited 30–84% amino acid identity to each other. All 16 KR domains in NPP PKS modules showed the two conserved active site motifs GXGXXG(A)XXXA and LXS(G)RXG(T,A). Like the KR13 in nystatin PKS module, the NPP KR13 is believe to be nonfunctional due to its aberrant second active site of LXAPSA ([5], Supplementary Fig. 1). Although the NPP KR12 has Ala instead of Ser or Gly at the third position in the second conserved active site, it is believed to be still functional due to the structural relatedness of these three amino acids.

The NPP DH domains exhibited 38–75% amino acid identity to each other. All DH domains bear the typical DH active site motif (LXXHXXXGXXXXP) [38], but two that were predicted to be inactive due to aberrant active site sequences, including LXXPXXDXXXXP in DH17 in module 17 and LXXEXXXGXXXXP in DH18 in module 18 (Supplementary Fig. 1). Although both the NPP and nystatin PKS modules contain the same number of presumably active DH domains, interestingly, there are eight inactive DH domains present in the nystatin PKS genes, in contrast with only two inactive DH domains in NPP PKS.

The two ER domains in NPP PKS modules, ER5 of NppC and ER15 of NppJ, showed 71% amino acid identity to each other, and contained the conserved active motif of GGVGXAAXQXA (Supplementary Fig. 1). The two ER domains localized in modules 5 and 15 are likely responsible for the reduction of a double bond between C-28/C-29 and C-8/C-9 of NPP, respectively. A TE domain was identified at the carboxyl terminus of NppK, suggesting that in addition to the condensation of the last extender unit, this protein also participates in the release of the mature NPP polyketide chain from the PKS complex [5].

NPP genes involved in post-PKS tailoring functions

Three genes (*nppDI, nppDII,* and *nppDIII*) are presumably involved in biosynthesis and attachment of the mycosamine moiety; they are highly homologous to genes previously characterized in the nystatin gene cluster [5]. The 490 aa-encoding *nppDI* is believed to encode glycosyltransferase, which is responsible for the attachment of the deoxysugar moiety (mycosamine) to the NPP aglycone at C-19. The 353 aa-encoding *nppDII* is believed to encode an aminotransferase, presumably involved in attachment of an amino group to the deoxysugar moiety for mycosamine biosynthesis. The 345 aa-encoding *nppDIII,* similar to the GDP-mannose-4,6-dehydratases gene, is believed to be responsible for one of the initial steps in mycosamine biosynthesis. Since these same three genes involved in mycosamine biosynthesis for both nystatin and NPP are located within the clusters, the other genes necessary for completion of mycosamine biosynthesis should be provided from outside of the cluster [3, 27].

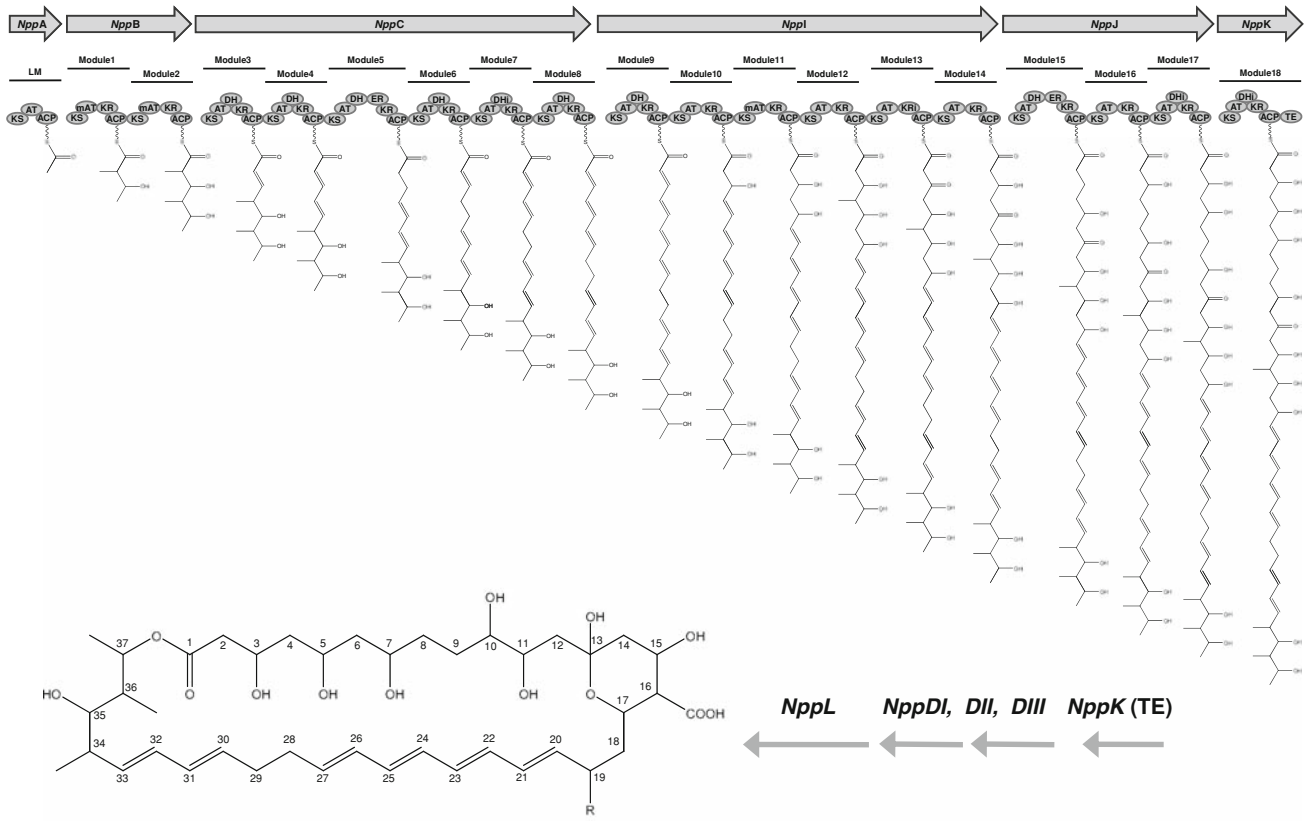

Fig. 2 A schematic representation of the modular organization of NPP PKS with catalytic domains and proposed NPP biosynthetic pathway in *P. autotrophica* KCTC9441. The sequence-based putative NPP structure is suggested to be identical to the Nystatin A1 aglycon with yet uncharacterized sugar(s) designated as *R*

There are two genes (404 aa-encoding *nppL* and 413 aa-encoding *nppN*) encoding cytochrome P450 hydroxylase (CYP), whose counterparts in the nystatin cluster are responsible for hydroxylation of the NPP polyketide moiety at C-10 and oxidation of the methyl group at C-16, respectively. In the nystatin CYP gene cluster, there is a ferredoxin-encoding *nysM* located just downstream of *nysN*, implying that NysM is required for oxidation of the methyl group at C-16 of nystatin. Interestingly, however, *nppM*, a putative counterpart of *nysM*, is absent downstream of *nppN,* yet located within the regulatory genes localized at the right fringe of the *npp* cluster (Fig. 1), whose biological significance needs to be further characterized.

The 585 aa-encoding *nppH* and 593 aa-encoding *nppG* genes located just upstream of *nppDIII* exhibit significant similarity to the ABC transporter family, and are presumably involved in ATP-dependent efflux of NPP [5]. Downstream of *nppG*, a putative *nppF* gene whose 225 aa products are homologous to the 4′-phosphopantheteine transferases was identified. NppF is believed to carry out post-translational modification of the ACP domains on the PKSs, which is required for their full functionality [5, 36]. It is likely, therefore, that the NppF protein functions in modification of the NPP PKS and is important for NPP biosynthesis.

A 275 aa-encoding *nppE* showed more than 77% aa identity to its counterparts of *nysE,* suggesting it may be involved in the expected functions of proof-reading type II thioesterase [5, 22]. Interestingly, however, the counterpart of a 527 aa-encoding *nppO,* which presumably encodes an acyl-CoA decarboxylase, was not found in the nystatin cluster (Fig. 1). Finally, a total of six putative regulatory genes (*nppRI, nppRII, nppRIII, nppRIV, nppV, and nppVI*), located at the right fringe of the cluster (Fig. 1) were suggested to be involved in NPP biosynthesis, on the bases of high aa similarities and previously characterized regulatory functions in both nystatin and FR008 biosynthesis [12, 34].

NPP production and antifungal bioassay against *Candia albicans*

Although a complete set of previously unknown polyene biosynthetic gene clusters was identified in the *P. atutotrophica* chromosome, it was not certain whether these

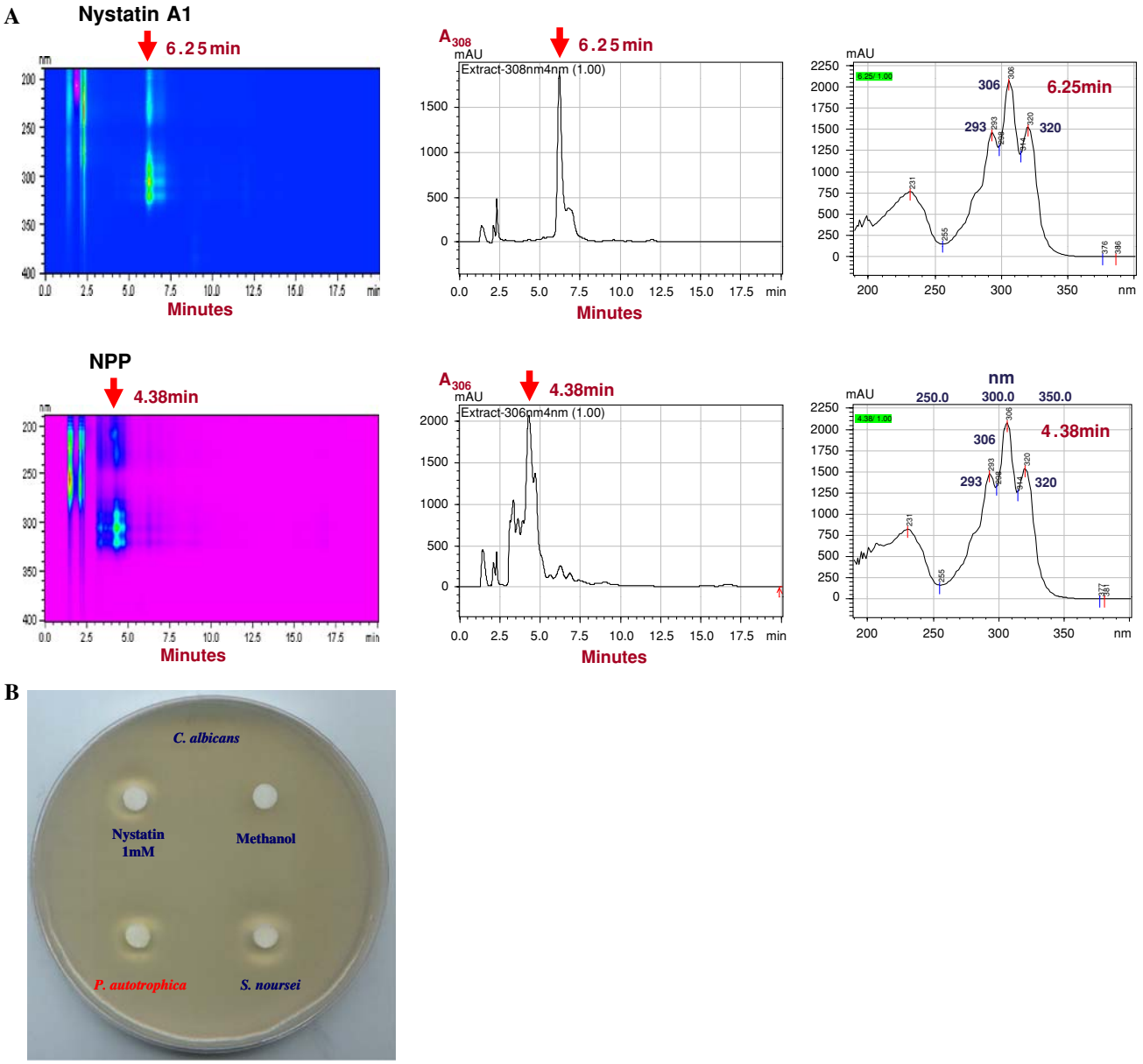

Fig. 3 **a** Comparison of PDA-HPLC analyses and UV spectrums of Nystatin A1 and butanol-extracted broth from *P. autotrophica* KCTC9441. **b** Antifungal bioassay for NPP production

genes could be functionally expressed in a culture for production of the actual polyene compound. In order to verify NPP production, *P. atutotrophica* was inoculated into various polyene production media including SAO-23 (nystatin medium), FD (amphotericin medium), and SPG (pimaricin production medium for *S. natalensis*), as well as several typical *Streptomyces* culture media including R2YE, YEME, and NDYE [5]. Among all the tested culture media, a very low (approximately less than 1 mg/l) NPP production was only detected from the *P. autotrophica* extracts cultured in YEME medium using a PDA-HPLC assay and a paper disc-agar diffusion antifungal bioassay using *C. albicans* (Fig. 3). It is thus implied that the *npp* biosynthetic pathway gene expression is tightly regulated and highly dependent on the ingredients of the medium. Since the PDA-HPLC chromatogram of NPP showed earlier elution time than that of nystatin A1 in a reverse phase HPLC column assay (Fig. 3), NPP might contain the same nystatin-like tetraene aglycon structure with different post-PKS modifications, such as different glycosylation patterns. Unlike the nystatin A1 molecular weight of 925.5035, the ESI-MS analysis of the major NPP compound resulted in the [Mw + H] value of 1,129.57, implying that NPP is believed to be structurally different from nystatin A1 (supplementary Fig. 2). The complete structural elucidation of NPP, which is currently hampered by its extremely low

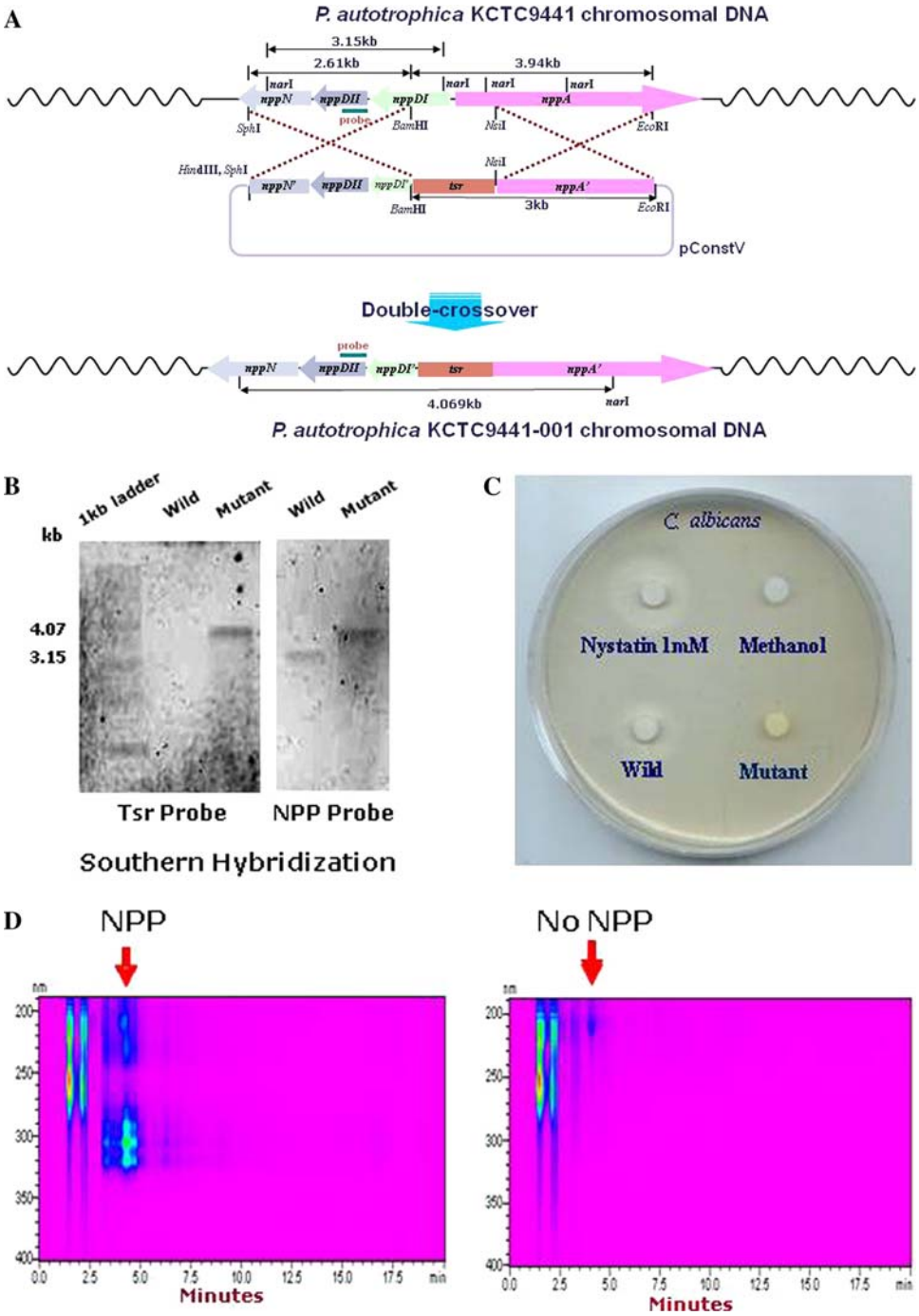

Fig. 4 **a** Model explaining the generation of the *P. autotrophica* KCTC9441-001 mutant. The *Nar*I restriction pattern before and after disruption is shown. The probe is indicated by *thick lines*. **b** Southern hybridization of the *Nar*I digested chromosomal DNA of wild-type *P. autotrophica* KCTC9441 and mutant *P. autotrophica* KCTC9441-001. **c** Comparison of antifungal activity against *C. albicans* of polyene between wild-type and mutant. **d** Comparison of PDA-HPLC analysis of butanol-extracts from wild-type and mutant

productivity as well as mixture production, is currently under investigation via *P. autotrophica* culture condition optimization as well as NPP overproduction strain improvement.

Confirmation of the identity of the NPP biosynthetic gene cluster by gene disruption

The organization of the *npp* gene cluster, particularly the PKS domain motif arrangement that was isolated and sequenced from *P. atutotrophica*, is consistent with its role in NPP biosynthesis. To verify its identity experimentally, we sought to confirm the function of this locus using a gene disruption approach. Inactivation of *nppA*, which is presumably involved in the biosynthesis of the polyketide backbone of NPP, was performed using the temperature-sensitive suicide vector pKC1139 (Fig. 4a). Construction of the *nppA* mutant (*P. autotrophica* KCTC9441-001) was confirmed by Southern hybridization. The expected size of the 3.15-kb band was observed in Southern hybridization

using a *nppA* gene probe and genomic DNA samples isolated from the wild-type *P. autotrophica* KCTC9441, while the expected size (4.07 kb) PCR-amplified bands were observed in genomic DNA samples isolated from *P. autotrophica* KCTC9441-001 (Fig. 4b). In addition, the expected size (4.07 kb) band observed in Southern hybridization using a *tsr* gene probe was observed only in *P. autotrophica* KCTC9441-001 (Fig. 4b), implying that *nppA* was specifically disrupted as expected. Both *P. autotrophica* KCTC9441 wild-type and *P. autotrophica* KCTC9441-001 strains were grown in YEME culture conditions for NPP production, followed by an antifungal bioassay and HPLC quantification. The absence of antifungal activity against *C. albicans* (Fig. 4c) as well as the absence of NPP in the PAD-HPLC assay (Fig. 4d) using extracts of the *P. autotrophica* KCTC9441 strain under the same culture conditions provide strong evidence that the *NPP* cluster isolated here does indeed encode enzymes responsible for NPP biosynthesis in *P. autotrophica* KCTC9441.

Concluding remarks

The polyene antibiotics comprise a family of valuable antifungal polyketide compounds, given their superior antifungal activities. These compounds are considered a promising structural family for novel and improved pharmacokinetic activities via biosynthetic pathway engineering as well as combinatorial biosynthesis approaches. Through genomic DNA library screening using a polyene-specific CYP gene probe, an approximately 125.7-kb contiguous DNA region in five overlapping cosmids isolated from the *P. autotrophica* KCTC9441 genomic library was identified, revealing a total of 23 ORFs that are presumably involved in Nystatin-like *Pseudonocardia* Polyene (NPP) biosynthesis. The isolation and analysis of the NPP biosynthetic gene cluster carried out in this study provide useful information with regard to the previously unknown genes governing polyene synthesis in a rare actinomycetes species such as *P. autotrophica*. Functional assignments for the proteins encoded within the cluster also open possibilities for genetic manipulations with the aim of polyene combinatorial biosynthesis in a genetically less-characterized rare actinomycetes species [7]. These results suggest that the polyene CYP-specific PCR screening approach may well constitute an efficient method for isolating potentially valuable polyene-producing actinomycetes for so-called "genomics-based in silico drug development."

Acknowledgments The authors appreciate Professor Sergey B. Zochev at Norwegian University of Science and Technology who kindly provided the authentic nystatin sample as well as some critical comments. This work was financially supported by a 21C Frontier R&D program grant from the Korean Ministry of Education, Science and Technology, and also NIH grant GM076477 to DHS.

References

1. Aparicio JF, Colina AJ, Ceballos E, Martín JF (1999) The biosynthetic gene cluster for the 26-membered ring polyene macrolide pimaricin. J Biol Chem 274:10133–10139
2. Aparicio JF, Fouces R, Mendes MV, Olivera 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
3. Aparicio JF, Caffrey P, Gil JA, Zotchev SB (2002) Polyene antibiotic biosynthesis gene clusters. Appl Microbiol Biotechnol 61:179–188
4. Bolard J (1986) How do the polyene macrolide antibiotics affect the cellular membrane properties? Biochim Biophys Acta 864:257–304
5. Brautaset T, Sekurova ON, Sletta H, Ellingsen TE, Strom AR, Valla S, Zotchev SB (2000) Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway. Chem Biol 7:395–403
6. Brautaset T, Bruheim P, Sletta H, Hagen L, Ellingsen TE, Strom AR, Valla S, Zotchev SB (2002) Hexaene derivatives of nystatin produced as a result of an induced rearrangement within the *nysC* polyketide synthase gene in *Streptomyces noursei* ATCC 11455. Chem Biol 9:367–373
7. Braustaset T, Sletta H, Nedal A, Borgos SEF, Degnes KF, Bakke I, Volokhan O, Sekurova ON, Treshalin ID, Mirchink EP, Dikiy A, Ellingsen TE, Zotchev SB (2008) Improved antifungal polyene macrolides via engineering of the nystatin biosynthetic genes in *Streptomyces noursei*. Chem Biol 15:1198–1206
8. Bruheim P, Borgos SEF, Tsan P, Sletta H, Ellingsen TE, Lancelin JM, Zotchev SB (2004) Chemical diversity of polyene macrolides produced by *Streptomyces noursei* ATCC 11455 and recombinant strain ERD44 with genetically altered polyketide synthased *NysC*. Antimicrol Agents Ch 48:4120–4129
9. Caffrey P, Lynch S, Flood E, Finnan S, Oliynyk M (2001) Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes. Chem Biol 8:713–723
10. Campelo AB, Gil JA (2002) The candicidin gene cluster from *Streptomyces griseus* IMRU 3570. Microbiology 148:51–59
11. Carmody M, Byrne B, Murphy B, Breen C, Lynch S, Flood E, Finnan S, Caffrey P (2004) Analysis and manipulation of amphotericin biosynthetic genes by means of modified phage KC515 transduction techniques. Gene 343:107–115
12. Chen S, Huang X, Zhou X, Bai L, He J, Jeong KJ, Lee SY, Deng Z (2003) Organization and mutation analysis of a complete FR-008/candicidin gene cluster encoding a structurally related polyene complex. Chem Biol 10:1065–1076
13. Del Vecchio F et al (2003) Active-site residue, domain and module swaps in modular polyketide synthases. J Ind Microbiol Biotechnol 30:489–494
14. Floss HG (2006) Combinatorial biosynthesis–potential and problems. J Biotechnol 124:242–257
15. Gupte M, Kulkarni P, Ganguli BN (2002) Antifungal antibiotics. Appl Microbiol Biotechnol 58:46–57
16. Haydock SF, Appleyard AN, Mironenko T, Lester J, Scott N, Leadlay PF (2005) Organization of the biosynthetic gene cluster for the macrolide concanamycin A in *Streptomyces neyagawaensis* ATCC. Microbiology 151:3161–3169
17. Hwang YB, Lee MY, Park HJ, Han K, Kim ES (2007) Isolation of putative polyene-producing actinomycetes strains via PCR-based genome screening for polyene-specific hydroxylase genes. Proc Biochem 42:102–107

18. Jonsbu E, Mcintyre M, Nielsen J (2002) The influence of carbon sources and morphology on nystatin production by *Streptomyces noursei*. J Biotechnol 95:133–144
19. Katz L, Donadio S (1993) Polyketide synthesis: prospects for hybrid antibiotics. Annu Rev Microbiol 47:875–912
20. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical *Streptomyces* Genetics. A laboratory mannual. John Innes Foundation, Norwich
21. Kojima I, Cheng YR, Mohan V, Demain AL (1995) Carbon source nutrition of rapamycin biosynthesis in *Streptomyces hygroscopicus*. J Ind Microbiol 14:436–439
22. Kotowska M, Pawlik K, Butler AR, Cundli VE, Takano E, Kuczek K (2002) Type II thioesterase from *Streptomyces coelicolor* A3(2). Microbiology 148:1777–1783
23. Lee MY, Myong JS, Park HJ, Han K, Kim ES (2006) Isolation and partial characterization of a cryptic polyene gene cluster in *Pseudonocardia autotrophica*. J Ind Microbiol Biotechnol 33:84–87
24. Mendes MV, Recio E, Fouces R, Luiten R, Martin JF, Aparicio JF (2001) Engineered biosynthesis of novel polyenes: a pimaricin derivative produced by targeted gene disruption in *Streptomyces natalensis*. Chem Biol 8:635–644
25. Menzella HG, Reid R, Carney JR, Chaandran SS, Reisinger SJ, Patel KG, Hopwood DA, Santi DV (2005) Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes. Nat Biotechnol 23:1171–1176
26. Munro AW, Lindsay JG (1996) Bacterial cytochromes P-450. Mol Microbiol 20:1115–1125
27. Nedal A, Zotchev SB (2004) Biosynthesis of deoxyaminosugars in antibiotic-producing bacteria. Appl Microbiol Biotechnol 64:7–15
28. Nelson DR (1999) Cytochrome P450 and the individuality of species. Arch Biochem Biophys 369:1–10
29. O'Keefe DP, Harder PA (1991) Occurrence and biological function of cytochrome P450 monooxygenases in actinomycetes. Mol Microbiol 5:2099–2105
30. Resat H, Sungur FA, Baginski M, Borowski E, Aviyente V (2000) Conformational properties of amphotericin B amide derivatives—impact on selective toxicity. J Comput Aided Mol Des 14:689–703
31. Rodriguez E, McDaniel R (2001) Combinatorial biosynthesis of antimicrobials and other natural products. Curr Opin Microbiol 4:526–534
32. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory mannual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
33. Schwecke T, Aparicio JF, Molnar I, Koing A, Khaw LE, Haydock SF, Oliynyk M, Caffrey P, Cortes J, Lester JB, Bohm GA, Staunton J, Leadlay PF (1995) The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. Proc Natl Acad Sci USA 92:7839–7843
34. Sekurova ON, Brautaset T, Sletta H, Borgos SE, Jakobsen ØM, Elingsen TE, Strøm AR, Valla S, Zotchev SB (2004) In vivo analysis of the regulatory genes in the nystatin biosynthetic gene cluster of *Streptomyces noursei* ATCC 11455 reveals their differential control over antibiotic biosynthesis. J Bacteriol 186:1345–1354
35. Seo YW, Cho KW, Lee HS, Yoon TM, Shin JH (2000) New polyene macrolide antibiotics from *Streptomyces* sp. M90025. J Microbiol Biotechnol 10:176–180
36. Volokhan O, Sletta H, Sekurova ON, Ellingsen TE, Zotchev SB (2005) An unexpected role for the putative 4′-phosphopantetheinyl transferase-encoding gene *nysF* in the regulation of nystatin biosynthesis in *Streptomyces noursei* ATCC 11455. FEMS Microbiol lett 249:57–64
37. Zazopoulos E, Huang K, Staffa A, Liu W, Bachmann BO, Nonaka K, Ahlert J, Thorson JS, Shen B, Farnet CM (2003) A genomics-guided approach for discovering and expressing cryptic metabolic pathways. Nat Biotechnol 21:187–190
38. Zirkle R, Black TA, Gorlach J, Ligon JM, Molnar I (2004) Analysis of a 108-kb region of the *Saccharopolyspora spinosa* genome covering the obscurin polyketide synthase locus. DNA Sequence 15:123–134
39. Zotchev SB, Haugan K, Sekurova O, Sletta H, Ellingsen TE, Valla S (2000) Identification of a gene cluster for antibacterial polyketide-derived antibiotic biosynthesis in the nystatin producer *Streptomyces noursei* ATCC 11455. Microbiology 146:611–619
40. Zotchev SB (2003) Polyene macrolide antibiotics and their applications in human therapy. Curr Med Chem 10:211–224