Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway

Trygve Brautaset[1], Olga N Sekurova[1], Håvard Sletta[2], Trond E Ellingsen[2], Arne R Strøm[1], Svein Valla[1] and Sergey B Zotchev[1]

Background: The polyene macrolide antibiotic nystatin produced by *Streptomyces noursei* ATCC 11455 is an important antifungal agent. The nystatin molecule contains a polyketide moiety represented by a 38-membered macrolactone ring to which the deoxysugar mycosamine is attached. Molecular cloning and characterization of the genes governing the nystatin biosynthesis is of considerable interest because this information can be used for the generation of new antifungal antibiotics.

Results: A DNA region of 123,580 base pairs from the *S. noursei* ATCC 11455 genome was isolated, sequenced and shown by gene disruption to be involved in nystatin biosynthesis. Analysis of the DNA sequence resulted in identification of six genes encoding a modular polyketide synthase (PKS), genes for thioesterase, deoxysugar biosynthesis, modification, transport and regulatory proteins. One of the PKS-encoding genes, *nysC*, was found to encode the largest (11,096 amino acids long) modular PKS described to date. Analysis of the deduced gene products allowed us to propose a model for the nystatin biosynthetic pathway in *S. noursei*.

Conclusions: A complete set of genes responsible for the biosynthesis of the antifungal polyene antibiotic nystatin in *S. noursei* ATCC 11455 has been cloned and analyzed. This represents the first example of the complete DNA sequence analysis of a polyene antibiotic biosynthetic gene cluster. Manipulation of the genes identified within the cluster may potentially lead to the generation of novel polyketides and yield improvements in the production strains.

Addresses: [1]UNIGEN Center for Molecular Biology and Department of Biotechnology, Norwegian University of Science and Technology, N-7489 Trondheim, Norway. [2]SINTEF Applied Chemistry, SINTEF, N-7034 Trondheim, Norway.

Correspondence: Sergey B Zotchev
E-mail: sergey.zotchev@chembio.ntnu.no

Key words: gene cluster, polyene antibiotic, polyketide synthase

Received: **30 March 2000**
Revisions requested: **4 May 2000**
Revisions received: **5 May 2000**
Accepted: **10 May 2000**

Published: **23 May 2000**

Chemistry & Biology 2000, **7**:395–403

1074-5521/00/$ – see front matter

Introduction

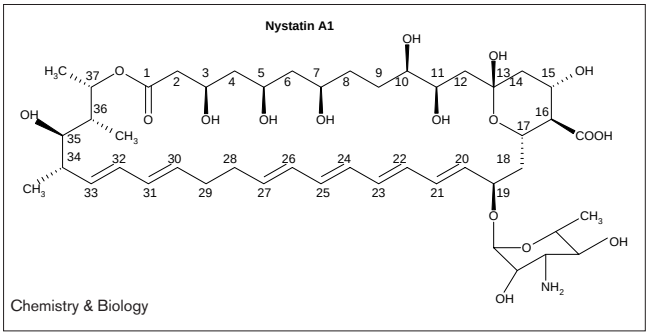
Polyketides are natural products, many of which have applied potential as pharmaceuticals. Examples of such polyketides include erythromycin (antibacterial), nystatin (antifungal), avermectin (antiparasitic), rapamycin (immunosuppressant) and daunorubicin (antitumor). The Gram-positive bacteria of the genus *Streptomyces* are the main producers of polyketides, and the genetics and biochemistry of polyketide biosynthesis in these organisms are relatively well characterized [1]. Macrolide polyketide compounds are formed via repeated condensations of simple carboxylic acids by modular (type I) polyketide synthases (PKSs) in a manner similar to fatty acid biosynthesis. The modular hypothesis proposed by Donadio *et al.* [2] suggested that type I PKSs are organized into repeated units (modules), each of which is responsible for one condensation cycle in the synthesis of a polyketide chain. This was proven to be correct by manipulations of type I PKS genes resulting in predictable changes in the chemical structures of macrolides [3]. Besides condensation of the next carboxylic acid onto the growing polyketide chain, ensured by the catalytic activity of a β-ketoacyl synthase (KS) domain, modules of PKSs type I may contain domains with β-ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) activities, which determine the reduced state of the incorporated extender unit. The acyltransferase (AT) and acyl carrier protein (ACP) domains present in each module are 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 probably also involved in cyclization of the final product. The type I PKSs therefore represent an assembly line for polyketide biosynthesis that can be manipulated by changing the number of modules, their specificities towards carboxylic acids, or by inactivating or inserting domains with reductive activities [3]. After the polyketide moiety is synthesized and cyclized to form a macrolactone ring, it is usually modified via hydroxylation,

glycosylation, methylation and/or acylation. These modifications are believed to be crucially important for the biological activities of macrolides.

The genes for macrolide antibiotics biosynthesis in *Streptomyces* are organized in clusters, making isolation of complete sets of such genes relatively straightforward. Indeed, exploitation of recombinant DNA technology makes it possible to isolate complete antibiotic biosynthetic gene clusters by screening gene libraries with DNA probes encoding PKS fragments [1]. The molecular cloning and complete DNA sequencing has been described for several macrolide antibiotics produced by streptomycetes, including those for avermectin, pikromycin and rapamycin [4–6]. The cloning and partial DNA sequencing of the gene cluster for the polyene macrolide antibiotic pimaricin has recently been reported [7].

Polyene macrolide antibiotics have 20–44-membered lactone rings containing three to eight conjugated double bonds. The latter structural feature most probably contributes to the mode of action of these compounds, determined to be an interaction with sterols present in the membranes of fungi [8]. Such an interaction leads to formation of polyene–sterol complexes capable of organizing themselves into transmembrane channels making the membrane permeable to water and ions, leading to cell death. Beside being antifungal agents, some of the polyene antibiotics have been shown to have antibacterial, antiviral and immunostimulating activities [9,10]. The antibiotic nystatin A1 (Figure 1) is produced by the Gram-positive bacterium *Streptomyces noursei* ATCC 11455, and is used as an antifungal agent. The chemical structure of nystatin suggests involvement of a type I PKS in the biosynthesis of its macrolactone ring. Furthermore, it can be predicted that the synthesis of the nystatin polyketide chain starts with acetyl-CoA, and proceeds through condensation of three methylmalonyl-CoA and 15 malonyl-CoA extender units. In this paper we describe the cloning and analysis of the complete nystatin biosynthetic gene cluster of *S. noursei*.

Figure 1



Structure of the nystatin A1 molecule according to Lancelin and Beau [36].

Results and discussion

Cloning of the nystatin biosynthetic gene cluster

In order to isolate the nystatin biosynthesis genes, we probed a *S. noursei* gene library with labelled DNA encoding a PKS. To obtain such a DNA probe, two degenerate oligonucleotide primers were designed, corresponding to conserved amino acid regions within KS and ACP domains of known modular PKSs (see the Materials and methods section). The polymerase chain reaction (PCR)-assisted amplification of a DNA fragment from the *S. noursei* genome with these primers was performed. Subsequent cloning and DNA sequence analysis of the resulting 0.7 kb PCR product (see the Materials and methods section) confirmed that it encodes part of a type I PKS. This DNA fragment was used for screening the *S. noursei* gene library constructed previously [11], and one recombinant phage, designated DASHII-N1, which hybridized to the probe, was isolated.

Preliminary DNA sequence analysis of the N1 DNA insert showed that it encodes a type I PKS. To confirm the involvement of the cloned DNA in nystatin biosynthesis, a 4.2 kb *Bam*HI DNA fragment from the N1 insert was used in a gene-disruption experiment as a part of the suicide conjugative plasmid pKO(4.2)–1 (Table 1). One of the resulting *S. noursei* disruption mutants, NCD3, carrying pKO(4.2)–1 integrated into its genome (data not shown) was unable to produce nystatin, thereby confirming the requirement of the identified PKS gene for nystatin biosynthesis. This fragment was later shown to be an internal part of the *nysC* gene (see below).

In order to clone a larger portion of the nystatin biosynthetic gene cluster, the DNA fragments derived from the ends of the DNA insert in phage N1, as well as those fragments from overlapping recombinant phages, were used as probes for screening the gene library. This 'chromosome walking' resulted in isolation of recombinant phages encompassing ~125 kb of the *S. noursei* genome (Figure 2).

DNA sequence and mutational analyses of the nystatin PKS genes

The complete DNA inserts from recombinant phages were subcloned in *Escherichia coli* and sequenced. Computer-assisted analysis of the DNA sequence (123,580 bp) of the cloned region led to identification of the genes shown in Figure 2 and listed in Table 2. Six genes, *nysA*, *nysB*, *nysC*, *nysI*, *nysJ* and *nysK*, encoding a type I PKS were identified. The amino acid (aa) sequences of the deduced products encoded by these genes were analysed by comparing them to the aa sequences of known type I PKSs. Because all six proteins were shown to share considerable homology with rifamycin and rapamycin PKSs [12,13],

Table 1

Bacterial strains and plasmids used in this study.

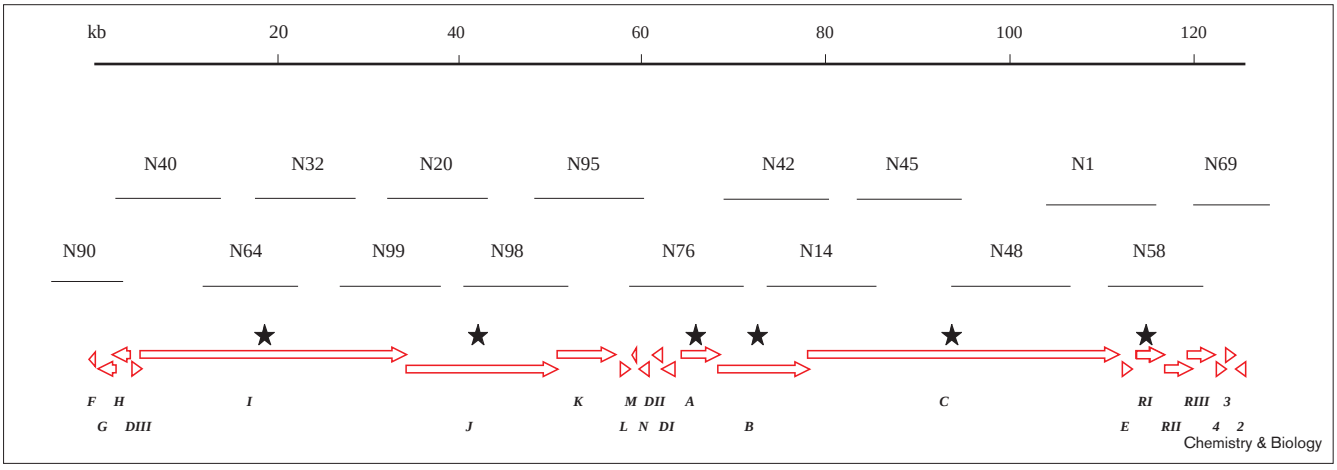
Strain or plasmid	Properties	Source or reference
E. coli XL-1 Blue MRA(P2)	Host for maintaining the gene library	Stratagene
E. coli DH5α	General cloning host	BRL
E. coli ET12567	Strain for intergeneric conjugation	[37]
S. noursei ATCC 11455	Wild type (WT), nystatin producer	ATCC
S. noursei NAD1	WT derivative with *nysA* disrupted by integration of pKNA1, nystatin nonproducer	This work
S. noursei NBD2	WT derivative with *nysB* disrupted by integration of pKNB1, nystatin nonproducer	This work
S. noursei NCD3	WT derivative with *nysC* disrupted by integration of pKO(4.2)–1, nystatin nonproducer	This work
S. noursei NID1	WT derivative with *nysI* disrupted by integration of pKNI1, nystatin nonproducer	This work
S. noursei NJD1	WT derivative with *nysJ* disrupted by integration of pKNJ1, nystatin nonproducer	This work
S. noursei NRD2	WT derivative with *nysRI* disrupted by integration of pNRD2, nystatin nonproducer	This work
pGEM3Zf(–)	ColEI replicon, Ap^R, 3.2 kb	Promega
pGEM11Zf(–)	ColEI replicon, Ap^R, 3.2 kb	Promega
pUZ8002	RK2 derivative, Km^R, Tc^R	D.H. Figurski*
pSOK201	pSG5 and ColEI replicons, *oriT* and Am^R, 7.1 kb	[11]
pPKS72	0.7 kb PCR product amplified from the *S. noursei* genome cloned in pUC18 (see the Materials and methods section)	This work
pL76S1.3	1.4 kb *Sal*I fragment of *nysA* cloned in pGEM3Zf(–)	This work
pKNA1	1.4 kb *Eco*RI–*Hin*dIII fragment from pL76S1.3 ligated with the 3.0 kb *Eco*RI–*Hin*dIII fragment from pSOK201	This work
pL76NB2.8	2.8 kb *Not*I–*Bam*HI fragment of *nysB* cloned in pGEM11Zf(–)	This work
pKNB1	2.8 kb *Bam*HI–*Hin*dIII fragment from pL76NB2.8 ligated with the 3.0 kb *Bam*HI–*Hin*dIII fragment from pSOK201	This work
pL64EB3.8	3.8 kb *Bam*HI–*Eco*RI fragment from N64 (internal for *nysI*) cloned in pGEM3Zf(–)	This work
pKNI1	3.8 kb *Eco*RI–*Hin*dIII fragment from pL64EB3.8 ligated with the 3.0 kb *Bam*HI–*Hin*dIII fragment from pSOK201	This work
pL20EB3.7	3.7 kb *Bam*HI–*Eco*RI fragment from N20 (internal for *nysJ*) cloned in pGEM3Zf(–)	This work
pKNJ1	3.7 kb *Eco*RI–*Hin*dIII fragment from pL20EB3.7 ligated with the 3.0 kb *Bam*HI–*Hin*dIII fragment from pSOK201	This work
pGEM4.2–1	4.2 kb *Bam*HI fragment of *nysC* cloned in pGEM11Zf(–)	This work
pKO(4.2)–1	4.2 kb *Eco*RI–*Hin*dIII fragment from pGEM4.2-1 ligated with the 3.0 kb *Eco*RI–*Hin*dIII fragment from pSOK201	This work
pNRD1	1.4 kb *Apa*I fragment of *nysRI* cloned in pGEM11Zf(–)	This work
pNRD2	1.4 kb *Eco*RI–*Hin*dIII fragment from pNRD1 ligated with the 3.0 kb *Eco*RI–*Hin*dIII fragment from pSOK201	This work

Am, apramycin; Ap, ampicillin; Km, kanamycin; Tc, tetracycline. *Personal communication.

presumptive functional analysis of the nystatin PKSs was based on the comparison to the former polypeptides. The predicted functional features of the NysA, NysB, NysC, NysI, NysJ and NysK proteins are shown in Figure 3.

The NysA protein most probably represents a loading module involved in the initiation of the nystatin aglycone biosynthesis. NysA contains a KS domain (KS^S) similar to that found in the presumed loading module of the pimaricin PKS PIMS0 [7]. Both the NysA and PIMS0 proteins are unusual in a sense that they represent the loading modules only, whereas in all other type I PKSs characterized so far, loading modules are fused to the first condensing module in multimodular polypeptides. The conserved active-site cysteine residue in both NysA and PIMS0 KS domains is replaced with a serine residue. In all the other known inactive KSs (KS^Q) in the loading modules of type I PKSs, the conserved cysteine is replaced by glutamine [14], except for the epothilone PKS loading module KS, which bears a Cys→Tyr replacement [15]. Because the aa sequence homology between the KS^S domains in NysA and PIMS0 (62.8%) is not higher than between the NysA KS^S and KS domains from other modular PKSs (ranging from 49.9% to 64.6%), it seems unlikely that these proteins have a recent common ancestor. It is tempting to speculate that the Cys→Ser replacement in the active sites of KS^S domains might have some significance for NysA and PIMS0 functioning as separate loading polypeptides. The latter might be related to the decarboxylase activity shown for the KS^Q domains to be largely

Figure 2



Gene organization within the *S. noursei* ATCC 11455 nystatin biosynthetic gene cluster. The inserts from the overlapping recombinant phages encompassing the cloned region are shown above the physical/genetic map. The *nys* genes are designated with capital letters in italics, other ORFs are numbered. Stars above certain ORFs indicate that these genes were successfully disrupted (see the text and Table 1 for details).

dependent on the presence of a glutamine residue in the active site [14]. It has been suggested that KS^Q domains in the loading modules of modular PKSs might provide the starter unit through the specific decarboxylation of enzyme-bound extenders [14]. Because the significance of the serine residue in the active sites of KS^S domains for their putative decarboxylase activity is not apparent, we are planning to address this question experimentally. Comparison of the NysA AT domain with its counterparts from different PKSs suggested that it is acetate specific, which correlates well with the structure of the nystatin molecule. Interestingly, NysA contains an apparently intact DH domain, which seems to serve no function, and might just have been retained in this protein in the process of evolution.

NysB apparently represents modules 1 and 2. The DH domains in both of these modules lack the conserved active-site motif $H(X_3)G(X_4)P$, due to large internal deletions, and therefore must be inactive. The AT domains identified within NysB display characteristic features of the propionate-specific AT domains (mAT) [13,16]. These features of NysB support the assumption about its involvement in first two elongation steps of the nystatin polyketide moiety biosynthesis.

The NysC protein, to our knowledge the largest bacterial polypeptide discovered to date, is composed of six modules apparently responsible for elongation steps 3–8 in the nystatin polyketide chain formation. All AT domains in NysC modules have features typical of the acetate-specific ATs (data not shown). This correlates well with the assumption that NysC incorporates six malonyl-CoA extenders (C-32 to C-21) into the nystatin aglycone. The only other protein with an organization similar to NysC, the RAPS2 PKS of *S. hygroscopicus*, comprising six modules and involved in rapamycin biosynthesis, is 873 aa shorter than NysC. This difference can be explained by the more compact structure of the RAPS2 domains and interdomain linkers, because the overall number of modules and domains, as well as their relevant positions in NysC and RAPS2 match perfectly. The end-to-end alignment of the NysC and RAPS2 aa sequences shows 49% identity — an unusually high number considering the lengths of these polypeptides. This, and the fact that rapamycin has an antifungal activity and initially was classified as an atypical polyene macrolide [17], might indicate that similarities between NysC and RAPS2 are due to a common ancestor from which the two proteins have evolved.

The NysI protein is probably responsible for elongation steps 9–14 of the nystatin polyketide backbone biosynthesis. The presence of an mAT domain in module 11 is consistent with incorporation of methylmalonyl-CoA extender at this elongation step. The DH domains in modules 10, 11, 12, 13 and 14 seem to be inactive due the large internal deletions encompassing the active-site motif (see above). The KR domain in module 13 of NysI lacks the conserved motif aSRrG, and thus appears to be inactive. The latter feature, together with inactive DH domain in module 11, most probably account for the presence of a six-membered ketalic ring (between C-13 and C-17) on the nystatin molecule (Figure 1). As judged from the organization of modules in NysJ, the latter is required for elongation steps 15–17 in nystatin macrolactone ring assembly. The DH domain in module 16 contains a large

Table 2

Putative genes identified in the nystatin biosynthetic gene cluster of *S. noursei*.

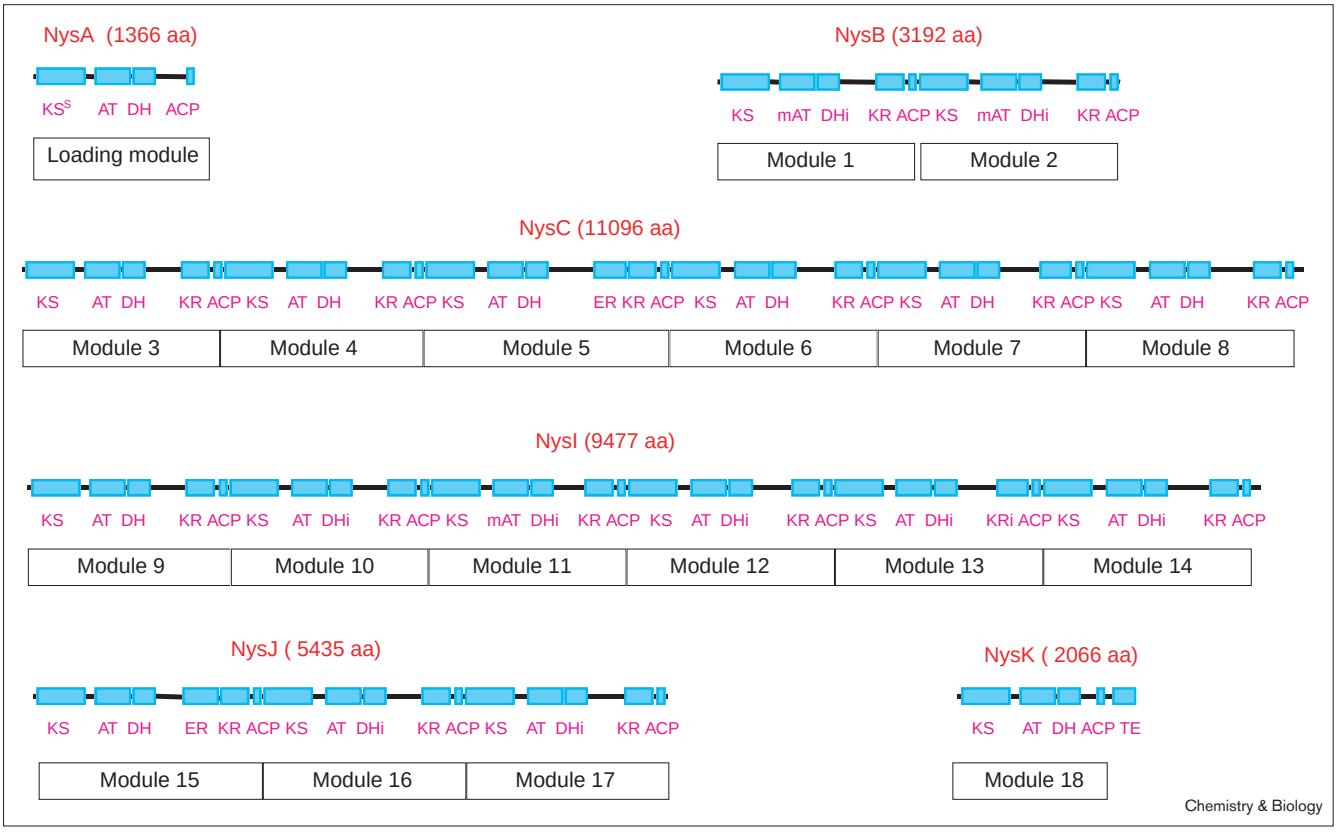
Designation	Product	Putative function
nysA	Type I PKS	Nystatin PKS (loading module)
nysB	Type I PKS	Nystatin PKS (modules 1 and 2)
nysC	Type I PKS	Nystatin PKS (modules 3–8)
nysDI	Glycosyltransferase	Attachment of mycosamine
nysDII	Aminotransferase	Mycosamine biosynthesis
nysDIII	GDP-mannose-4,6-dehydratase	Mycosamine biosynthesis
nysE	Thioesterase	Release of polyketide chain from PKS
nysF	4′-Phosphopantheteine transferase	Post-translational PKS modification
nysG	ABC transporter	Efflux of nystatin
nysH	ABC transporter	Efflux of nystatin
nysI	Type I PKS	Nystatin PKS (modules 9–14)
nysJ	Type I PKS	Nystatin PKS (modules 15–17)
nysK	Type I PKS	Nystatin PKS (module 18 + TE)
nysL	P450 monooxygenase	Hydroxylation at C-10
nysM	Ferredoxin	Electron transfer in P450 system
nysN	P450 monooxygenase	Oxidation of methyl group at C-16
nysRI	Transcriptional activator	Regulation of nystatin production
nysRII	Transcriptional activator	Regulation of nystatin production
nysRIII	Transcriptional activator	Regulation of nystatin production
ORF4	Transcriptional activator	Regulation
ORF3	Transcriptional repressor	Regulation
ORF2	Transcriptional activator	Regulation

internal deletion encompassing the active site, and the conserved histidine residue in the module 17 DH domain active site $H(X_3)G(X_4)P$ is replaced with tyrosine. The DH domains in modules 16 and 17 within NysJ therefore appear to be inactive. The ER domain localized in module 15 is probably responsible for the reduction of a double bond between C-8 and C-9. The final 18th module in the nystatin PKS system is represented by the NysK protein, which lacks a KR domain, and contains an apparently intact DH domain whose activity should not be required at the last condensation step. We note that a similar observation was made during analysis of the rifamycin PKS, in which intact DH domains in modules 6 and 7 seem to be nonfunctional in the producing organism [13]. A TE domain was identified at the carboxyl terminus of NysK, suggesting that in addition to the condensation of the last extender unit, this protein also participates in the release of the mature nystatin polyketide chain from the PKS complex.

To confirm the involvement of *nysA*, *nysB*, *nysI* and *nysJ* in nystatin biosynthesis, these genes were disrupted in *S. noursei* via homologous recombination using the conjugative suicide vectors pKNA1, pKNB1, pKNI1, and pKNJ1 (Table 1). All the disruption mutants obtained (Table 1) were unable to produce nystatin, suggesting that the genes mentioned above are essential for nystatin biosynthesis. Even though the polar effects of the disruptions mentioned above cannot be ruled out, such effects would imply that the genes affected are transcribed as part of a polycistronic mRNA together with the disrupted genes (see Figure 2). It seems highly unlikely, therefore, that the disrupted genes themselves have no role in nystatin biosynthesis.

Immediately downstream of the *nysC* gene, a coding sequence (*nysE*) for a TE was identified, providing yet another example of two TE activities being localized in the same antibiotic biosynthetic gene cluster. The 251 aa NysE protein shows a high degree of homology (45%) to the TEs encoded by the independent genes in the pikromycin and tylosin biosynthetic gene clusters of *Streptomyces venezuelae* and *Streptomyces fradiae* [6,17]. The NysE polypeptide, however, does not display substantial homology to the TE domains located at the carboxyl termini of several modular PKSs. The precise role for the putative TE encoded by *nysE* in nystatin biosynthesis remains uncertain. Both the pikromycin and tylosin biosynthetic gene clusters encode two TE activities, one of them embedded in the PKS [6,18]. Although the exact functions of these TEs are not clear, it was proposed that one of them might be a 'proof-reading' enzyme, clearing off certain substrates that would block further extension of the chain by PKS [19,20].

Figure 3



Functional organization of the nystatin PKS including the NysA, NysB, NysC, NysI, NysJ and NysK proteins. KS, ketosynthase; KS^S, ketosynthase with the Cys→Ser substitution in the active site; AT, acetate-specific acyltransferase; mAT, propionate-specific acetyltransferase; DH, dehydratase; DHi, inactive dehydratase; ER, enoyl reductase; KR, ketoreductase; KRi, inactive ketoreductase; ACP, acyl carrier protein.

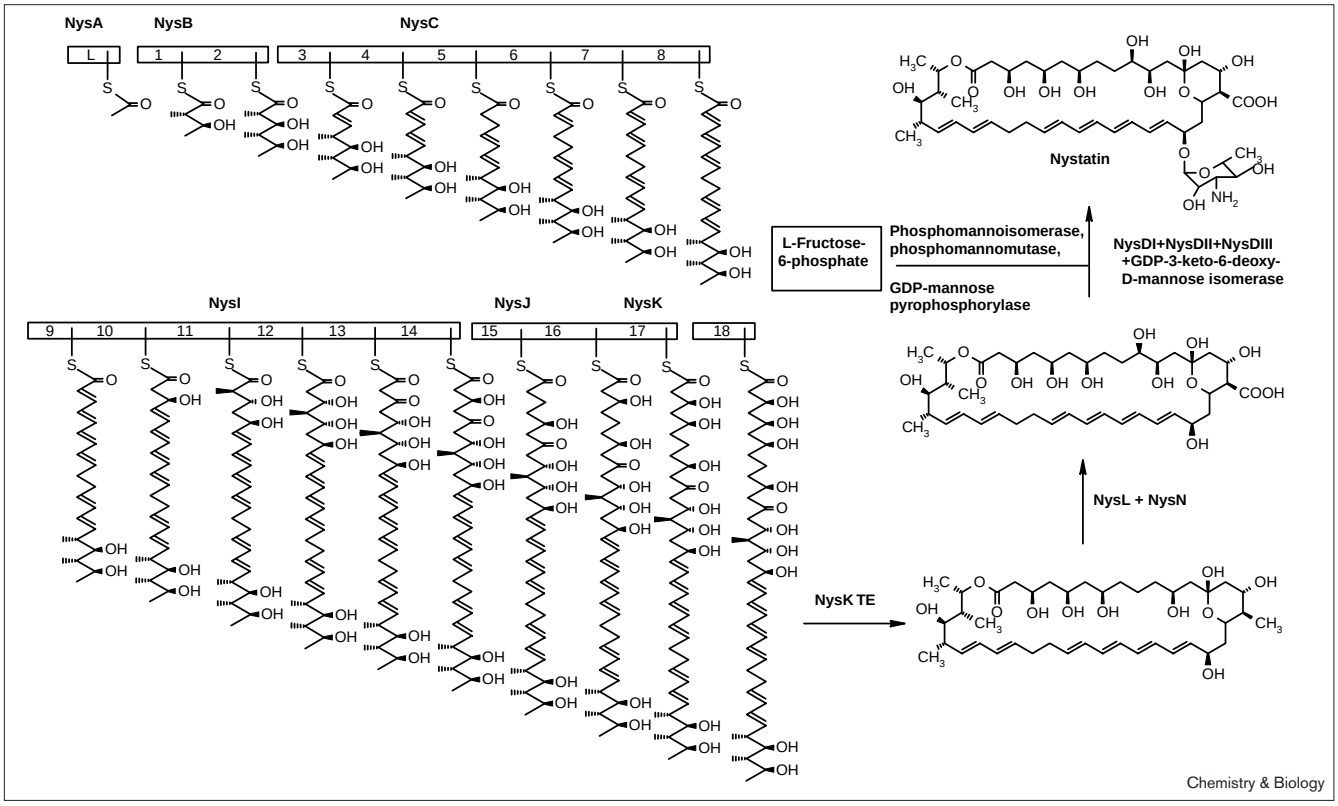
Genes in the regulatory region

The putative regulatory genes *nysRI*, *nysRII* and *nysRIII* encoding polypeptides of 966 aa, 953 aa and 927 aa, respectively, were identified downstream of *nysE*. The deduced products of these genes are homologous to the transcriptional regulators found in the vicinity of the cholesterol oxidase–cytochrome P450 operon of *Streptomyces* sp. SA-COO [21]. To confirm the involvement of *nysRI* in nystatin biosynthesis, this gene was disrupted in *S. noursei* with the suicide vector pNRD2 (Table 1). Analysis of the secondary metabolites produced by the corresponding mutant NRD2 revealed that it cannot synthesize nystatin. Because the *nysRII* and *nysRIII* genes are located just downstream of *nysRI*, and are transcribed in the same direction, it is plausible that the *nysRI* disruption has a polar effect on transcription of these genes.

Downstream of *nysRIII*, three ORFs (ORFs 2–4) were identified that might be involved in the regulation of nystatin biosynthesis. The deduced *ORF4* product of 210 aa is similar to the transcriptional activators of response regulator type, whereas a 253 aa polypeptide encoded by *ORF3* shows considerable homology to the transcriptional repressors of the DeoR family [22]. *ORF2*, located downstream of *ORF3*, and transcribed in the opposite direction, encodes a putative polypeptide of 354 aa that is similar to the transcriptional activators of the AsnC family [23]. Experiments aimed at mutational analysis of all the regulatory genes associated with the gene cluster are currently under way, and will provide the answer as to their involvement in the process of nystatin biosynthesis.

Sequencing of an additional 5 kb DNA from the phage N69 insert extending the cloned region to the right from *ORF2* (data not shown) identified several genes for peptide metabolism (T.B. and S.B.Z., unpublished observations) for which no role in nystatin biosynthesis could be assigned. This observation suggests that the right border of the nystatin biosynthetic gene cluster has been identified.

Figure 4



Proposed model for nystatin biosynthesis in *S. noursei*.

Putative mycosamine biosynthesis genes

Three genes presumably involved in biosynthesis and attachment of the mycosamine moiety were found in the cluster (Figure 2; Table 2). The 506 aa-long *nysDI* product shows considerable homology to the eukaryotic UDP-glucuronosyltransferases. The latter enzyme belongs to the UDP-glycosyltransferase family, and is involved in eliminating potentially toxic xenobiotics by the way of their glycosylation [24]. It seems likely that NysDI represents a glycosyltransferase responsible for the attachment of the deoxysugar moiety (mycosamine) to the nystatin aglycone at C-19. The deduced product of *nysDII* (352 aa) is highly similar to perosamine synthetases from different bacteria. The latter enzyme catalyzes conversion of GDP-4-keto-6-deoxy-D-mannose to 4-NH_2-4,6-dideoxy-D-mannose in the biosynthesis of perosamine, which constitutes the backbone structural unit of the lipopolysaccharide O-antigens in pathogenic bacteria [25]. Because the nystatin deoxysugar moiety mycosamine contains an amino group, it is plausible that NysDII represents an aminotransferase involved in mycosamine biosynthesis. The *nysDIII* gene encodes a 344 aa protein similar to GDP-mannose-4,6-dehydratases, and is therefore probably responsible for one of the initial steps in mycosamine biosynthesis. The fact that NysDIII more closely resembles the GDP-mannose-4,6-dehydratases than the TDP-glucose-4,6-dehydratases (Gdh) suggests that the mycosamine biosynthetic pathway in *S. noursei* differs from those of other antibiotics' deoxysugar moieties. This would also explain why our initial attempt to identify the nystatin gene cluster using a Gdh-specific DNA probe failed [11]. The presence of only two deoxysugar biosynthesis genes in the cluster might be explained by the key roles played by the corresponding enzymes in mycosamine biosynthesis. All other enzymes required for this deoxysugar formation can, in principle, be recruited from different (i.e. primary) metabolic pathways.

Modification and transport genes

Three genes encoding proteins presumably involved in modification of the nystatin molecule were identified between *nysK* and *nysDII* (Figure 2). Both the *nysL* and *nysN* genes encode P450 monooxygenases of 394 aa and 398 aa, respectively, that are probably responsible for hydroxylation of the nystatin polyketide moiety at C-10, and oxidation of the methyl group at C-16. Which protein is responsible for which reaction is not clear at the moment, and additional experiments are required for exact placement of NysL and NysN in the nystatin biosynthetic pathway. The *nysM* gene apparently encodes

a ferredoxin of 64 aa, which presumably constitutes a part of one or both P450 monooxygenase systems, and serves as an electron donor [26].

The *nysH* and *nysG* genes localized upstream of *nysDIII* encode 584 aa and 605 aa polypeptides, respectively. Both NysH and NysG display high degrees of similarity to transporters of the ABC family [27], and thus might be involved in ATP-dependent efflux of nystatin.

Downstream of *nysG*, a putative *nysF* gene was identified whose 245 aa product is homologous to the 4′-phosphopantheteine transferases. The latter enzyme carries out the post-translational modification of the ACP domains on the PKSs, which is required for their full functionality [28]. It seems likely, therefore, that the NysF protein functions in modification of the nystatin PKS and is important for the nystatin biosynthesis. DNA sequencing and analysis of the insert in phage N90, which extends the characterized region by ~10 kb to the left of *nysF*, identified genes for signal peptidase, and other proteins with no possible role in nystatin biosynthesis (T.B. and S.B.Z., unpublished observations). The putative *nysF* gene therefore marks the left border of the nystatin biosynthetic gene cluster in *S. noursei*.

Proposed model for the nystatin biosynthetic pathway of *S. noursei*

Based on the information derived from the analysis of the genes found in the nystatin biosynthetic gene cluster we propose a model for synthesis of this antibiotic in the producing organism *S. noursei* (Figure 4). The synthesis starts with loading of the acetyl-CoA onto the NysA protein, and proceeds through condensation of three methylmalonyl-CoA and 15 malonyl-CoA extender units by NysB, NysC, NysI, NysJ and NysK PKS. After cleavage of the mature polyketide chain from the PKS complex by the TE domain of NysK, the chain is cyclized to form the nystatin aglycone. The next two steps in the nystatin biosynthesis are probably accomplished by the NysL and NysN monooxygenases, which perform hydroxylation and oxidation of the macrolactone ring at C-10 and C-16, respectively.

Biosynthesis of mycosamine presumably starts with the L-fructose-6-phosphate, which is converted to GDP-D-mannose through the action of a phosphomannoisomerase, phosphomannomutase, and a GDP–mannose pyrophosphorylase. These enzymes are involved in primary metabolic pathways for cell-surface lipopolysaccharide biosynthesis in many bacteria [25], and the genes for the two former have been located on the chromosome of *S. coelicolor* A3(2) during the genome sequencing effort. It seems likely, therefore, that there is interplay between the primary and secondary metabolisms in *S. noursei* during mycosamine biosynthesis, although it is not apparent how the regulation of such genes is coordinated. The GDP-D-mannose probably serves as a substrate for the NysDIII protein, which converts it to the GDP-4-keto-6-deoxy-D-mannose. A GDP-3-keto-6-deoxy-D-mannose isomerase must perform the next step in mycosamine biosynthesis, probably followed by the NysDII-mediated amidation leading to formation of GDP-mycosamine. The NysDI protein, a putative glycosyltransferase, presumably completes the pathway by attaching the mycosamine moiety to the modified nystatin aglycone.

The model presented above is based on the assumption that both monooxygenases modify the nystatin macrolactone ring prior to its glycosylation. Whether this is true will be verified by the ongoing experiments on inactivation of the *nysL* and *nysN* genes, and structural analysis of the nystatin derivatives produced by the corresponding mutants.

Significance

Analysis of the *Streptomyces noursei* nystatin biosynthetic gene cluster carried out in this study provides useful information with regard to the genes governing synthesis of this important antifungal agent. Functional assignments for the proteins encoded within the cluster suggest a model for the nystatin biosynthetic pathway, and open possibilities for genetic manipulations with the aim of producing novel nystatin derivatives. Identification of several putative regulatory genes associated with the cluster, along with data on inactivation of one potential regulator, imply that they can be used for enhancing the yield of nystatin (and probably its derivatives) in fermentations.

Materials and methods

Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 1. *S. noursei* ATCC 11455 and its mutants were grown on solid ISP2 medium (Difco), and in liquid TSB medium (Oxoid). Intergeneric conjugation from *E. coli* ET12567 (pUZ8002) into *Streptomyces* strains was done as reported elsewhere [29], but with the 'heat shock' time reduced to 5 min. *E. coli* strains were grown and transformed as described by others [30], except for *E. coli* ET12567 (pUZ8002), which was maintained on media with 20 μg/ml chloramphenicol and 50 μg/ml kanamycin. Analysis of the nystatin production by *S. noursei* strains was performed in SAO-23 liquid medium [31].

DNA manipulations

Plasmid, phage and total DNA preparations, endonuclease digestions and ligations were performed as described previously [30,32]. DNA fragments were isolated from agarose gels using the QIAGEN Kit (QIAGEN GmbH, Germany), labelled with the use of the digoxygenin kit from Boehringer Mannheim, and used for Southern blot analysis according to the manufacturer's instructions. Genotypes of all disruption mutants obtained in the course of this study (Table 1) were verified by Southern blot analysis (data not shown). DNA sequencing was performed at QIAGEN GmbH, and the data were analyzed with the FramePlot 2.3 online program [33] and GCG software [34].

Amplification of a PKS-encoding DNA fragment used for screening of the S. noursei *gene library*

Degenerate primers used for amplification of a PKS-encoding DNA fragment corresponded to the conserved aa motifs in ACP and KS domains in known type I PKSs, and were designed according to the

codon usage table for *Streptomyces* [35]. The ACP oligonucleotide primer (sense) had the sequence 5′-GAG/C CTG/C GGC/G T/CTG/C GAC TCC/G CTG/C-3′, and the KS oligonucleotide primer (antisense) had the sequence 5′-G/CGA G/CGA G/ACA G/CGC C/GGT GTC G/CAC-3′. The 50 µl PCR mixture contained: 0.1 µg *S. noursei* ATCC 11455 genomic DNA, 25 pm each ACP and KS oligonucleotide primers, dNTPs (final concentration 350 µm), 1× PCR buffer from Expand High Fidelity PCR System (Boehringer Mannheim), and 1.5 U of the DNA polymerase mixture from the same system. The PCR was performed on the Perkin Elmer GeneAmp PCR System 2400 with the following program: 1 cycle of denaturation at 96°C (4 min), 35 cycles of denaturation/annealing/synthesis at 94°C (45 s) and 70°C (5 min), and 1 cycle of final annealing/extension at 72°C (7 min). The 0.7 kb DNA fragment obtained with this procedure was cloned in pUC18 with the use of SureClone Ligation Kit (Pharmacia). One of the resulting recombinant plasmids, pPKS72, was subjected to DNA sequence analysis, and later used for primary screening of the *S. noursei* gene library.

Accession numbers

The DNA sequence reported here was deposited in GenBank under the accession number AF263912.

Acknowledgements

We thank K. Haugan for constructing the *S. noursei* gene library, D.H. Figurski and D. MacNeil for the gift of plasmids and strains, R. Aune for running the fermentations, and E. Fjærvik for participation in discussions. C.R. Hutchinson's constant support and useful comments on the manuscript are greatly appreciated. This work was supported by the Norwegian Research Council, SINTEF, Norwegian University of Science and Technology, and Alpharma AS.

References

1. Hopwood, D.A. (1997). Genetic contributions to understanding polyketide synthases. *Chem. Rev.* **97**, 2465-2497.
2. Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J. & Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**, 675-679.
3. Katz, L. & McDaniel, R. (1999). Novel macrolides through genetic engineering. *Med. Res. Rev.* **19**, 543-558.
4. 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.
5. Schwecke, T., *et al*., & Leadlay, P.F. (1995). The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc. Natl Acad. Sci. USA* **92**, 7839-7843.
6. Xue, Y., Zhao, L., Liu, H.W. & Sherman, D.H. (1998). A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: architecture of metabolic diversity. *Proc. Natl Acad. Sci. USA* **95**, 12111-12116.
7. Aparicio, J.F., Colina, A.J., Ceballos, E. & Martín, J.F. (1999). The biosynthetic gene cluster for the 26-membered ring polyene macrolide pimaricin: a new polyketide synthase organization encoded by two subclusters separated by functionalization genes. *J. Biol. Chem.* **274**, 10133-10139.
8. Bolard, J. (1986). How do the polyene macrolide antibiotics affect the cellular membrane properties? *Biochim. Biophys. Acta* **864**, 257-304.
9. Ciftci, T., Borkman, T.A., McDaniel, L.E. & Schaffner, C.P. (1984). Comparative analysis of hexaene antibiotics. *J. Antibiot.* **37**, 876-884.
10. Abu-Salah, K.M. (1996). Amphotericin B: an update. *Br. J. Biomed. Sci.* **53**, 122-133.
11. Zotchev, S. B., Haugan, K., Sekurova, O., Sletta, H., Ellingsen, T.E. & Valla, S. (2000). Identification and partial characterization of the gene cluster governing biosynthesis of a novel antibacterial polyketide-derived antibiotic in *Streptomyces noursei* ATCC 11455. *Microbiology* **146**, 611-619.
12. Aparicio, J.F., *et al*., & Leadlay, P.F. (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.
13. Tang, L., Yoon, Y.J., Choi, C.Y. & Hutchinson, C.R. (1998). Characterization of the enzymatic domains in the modular polyketide synthase involved in rifamycin B biosynthesis by *Amycolatopsis mediterranei*. *Gene* **216**, 255-265.
14. Bisang, C., *et al*., & Leadlay, P.F. (1999). A chain initiation factor common to both modular and aromatic polyketide synthases. *Nature* **401**, 502-505.
15. Tang, L., Shah, S., Chung, L., Carney, J., Katz, L., Khosla, C. & Julien, B. (2000). Cloning and heterologous expression of the epothilone gene cluster. *Science* **287**, 640-642.
16. Haydock, S.F., *et al*., & Leadlay, P.F. (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. Omura, S. & Tanaka, H. (1984). Production, structure, and antifungal activity of polyene macrolides. In *Macrolide Antibiotics: Chemistry, Biology and Practice*. (Omura, S., ed.), pp.351-405, Academic Press, Inc., New York.
18. Merson-Davies, L.A. & Cundliffe, E. (1994). Analysis of five tylosin biosynthetic genes from the *tylIBA* region of the *Streptomyces fradiae* genome. *Mol. Microbiol.* **13**, 349-355.
19. Butler, A.R., Bate, N. & Cundliffe, E. (1999). Impact of thioesterase activity on tylosin biosynthesis in *Streptomyces fradiae*. *Chem. Biol.* **6**, 287-292.
20. Tang, L., Fu, H., Beltach, M.C. & McDaniel, R. (1999). Elucidating the mechanism of chain termination switching in the picromycin/methymycin polyketide synthase. *Chem. Biol.* **6**, 553-558.
21. Molnar, I. & Murooka, Y. (1993). Nucleotide sequence analysis of a region upstream of the cholesterol oxidase-cytochrome P450 operon of *Streptomyces* sp. SA-COO revealing repeating units coding for putative transmembrane and DNA-binding proteins. *J. Ferment. Bioeng.* **76**, 257-264.
22. Valentin-Hansen, P., Hoejrup, P. & Short, S. (1985). The primary structure of the DeoR repressor from *Escherichia coli* K-12. *Nucleic Acids Res.* **13**, 5927-5936.
23. Koelling, R. & Lother, H. (1985). AsnC: an autogenously regulated activator of asparagine synthetase A transcription in *Escherichia coli*. *J. Bacteriol.* **164**, 310-315.
24. Radominska-Pandya, A., Czernik, P.J., Little, J.M., Battaglia. E. & Mackenzie, P.I. (1999). Structural and functional studies of UDP-glucuronosyltransferases. *Drug Metab. Rev.* **31**, 817-899.
25. Trefzer, A., Salas, J.A. & Bechthold, A. (1999). Genes and enzymes involved in deoxysugar biosynthesis in bacteria. *Nat. Prod. Rep.* **16**, 283-289.
26. O'Keefe, D.P. & Harder, P.A. (1991). Occurrence and biological function of cytochrome P450 monooxygenases in the actinomycetes. *Mol. Microbiol.* **5**, 2099-2105.
27. Fath, M.J. & Kolter, R. (1993). ABC transporters: bacterial exporters. *Microbiol. Rev.* **57**, 995-1017.
28. Kealey, J.T., Liu, L., Santi, D.V., Betlach, M.C. & Barr, P.J. (1998). Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc. Natl Acad. Sci. USA* **95**, 505-509.
29. Flett, F., Mersinias, V. & Smith, C.P. (1997). High efficiency conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA restricting streptomycetes. *FEMS Microbiol. Lett.* **155**, 223-229.
30. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, (2nd edn). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
31. Sekurova O., Sletta H., Ellingsen T., Valla S. & Zotchev, S.B. (1999). Molecular cloning and analysis of a pleiotropic regulatory gene locus from the nystatin producer *Streptomyces noursei* ATCC 11455. *FEMS Microbiol. Lett.* **177**, 297-304.
32. Hopwood, D.A., *et al*., & Schrempf, H. (1985). *Genetic Manipulation of Streptomyces, A Laboratory Manual*. John Innes Foundation, Norwich.
33. Ishikawa, J. & Hotta, K. (1999). FramePlot: a new implementation of the Frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. *FEMS Microbiol. Lett.* **174**, 251-253.
34. Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 387-395.
35. Wright, F. & Bibb, M.J. (1992). Codon usage in the G+C-rich *Streptomyces* genome. *Gene* **113**, 55-65.
36. Lancelin, J.-M. & Beau, J.-M. (1989). Complete stereostructure of nystatin A1: a proton NMR study. *Tetrahedron Lett.* **30**, 4521-4524.
37. MacNeil, D.J., Gewain, K.M., Ruby, C.L., Dezeny, G., Gibbons, P.H. & MacNeil, T. (1992). Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**, 61-68.