ELSEVIER

Chemistry & Biology 8 (2001) 713–723

Chemistry & Biology

www.elsevier.com/locate/chembiol

Research Paper

Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes

Patrick Caffrey [a,b,*], Susan Lynch [a], Elizabeth Flood [a], Shirley Finnan [a], Markiyan Oliynyk [c]

[a] *Department of Industrial Microbiology, University College Dublin, Dublin 4, Ireland*
[b] *Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Dublin 4, Ireland*
[c] *Department of Biochemistry, University of Cambridge, Cambridge, UK*

Received 8 March 2001; revisions requested 27 April 2001; revisions received 8 May 2001; accepted 10 May 2001
First published online 8 June 2001

Abstract

Background: The polyene macrolide amphotericin B is produced by *Streptomyces nodosus* ATCC14899. Amphotericin B is a potent antifungal antibiotic and has activity against some viruses, protozoans and prions. Treatment of systemic fungal infections with amphotericin B is complicated by its low water-solubility and side effects which include severe nephrotoxicity. Analogues with improved properties could be generated by manipulating amphotericin biosynthetic genes in *S. nodosus*.

Results: A large polyketide synthase gene cluster was cloned from total cellular DNA of *S. nodosus*. Nucleotide sequence analysis of 113 193 bp of this region revealed six large polyketide synthase genes as well as genes for two cytochrome P450 enzymes, two ABC transporter proteins, and genes involved in biosynthesis and attachment of mycosamine. Phage KC515-mediated gene disruption was used to show that this region is involved in amphotericin production.

Conclusions: The availability of these genes and the development of a method for gene disruption and replacement in *S. nodosus* should allow production of novel amphotericins. A panel of analogues could lead to identification of derivatives with increased solubility, improved biological activity and reduced toxicity.

Keywords: Amphotericin B; Polyketide synthase; Gene disruption; *Streptomyces nodosus*

1. Introduction

The polyene macrolide amphotericin B is a medically important antifungal antibiotic that is produced by *Streptomyces nodosus* [1]. Polyenes disrupt eukaryotic cell membranes after complexing with sterols to form channels that allow loss of small molecules and ions and ultimately cause cell death [2]. Amphotericin B is selectively toxic towards fungal cells because it shows a high affinity for ergosterol, the predominant sterol in fungal cell membranes. However, the drug also interacts to a lesser extent with cholesterol in mammalian cell membranes and has many severe side effects, especially nephrotoxicity. In spite of these adverse effects, amphotericin B is still the most important antibiotic for the treatment of serious systemic fungal infections in humans.

In addition to its antifungal properties, amphotericin B displays a number of other biological activities that may result from its interaction with sterols [3]. The drug inhibits infection of cultured cells by human immunodeficiency virus (HIV) [4]. The envelopes of these virus particles have a higher cholesterol:phospholipid ratio than host cell membranes [5]. Amphotericin B also delays the onset of prion disease symptoms in animal models [6]. Polyenes may interfere with formation of abnormal isoforms of prion proteins during trafficking of sterol-rich membrane microdomains that contain these glycosylphosphatidylinositol-anchored proteins [7]. Amphotericin B is also active against *Leishmania*, a protozoal parasite that contains ergosterol precursors in its membranes [3].

The problems associated with amphotericin B are re-

* Correspondence: Patrick Caffrey; E-mail: patrick.caffrey@ucd.ie

1074-5521/01/$ – see front matter
PII: S1074-5521(01)00046-1

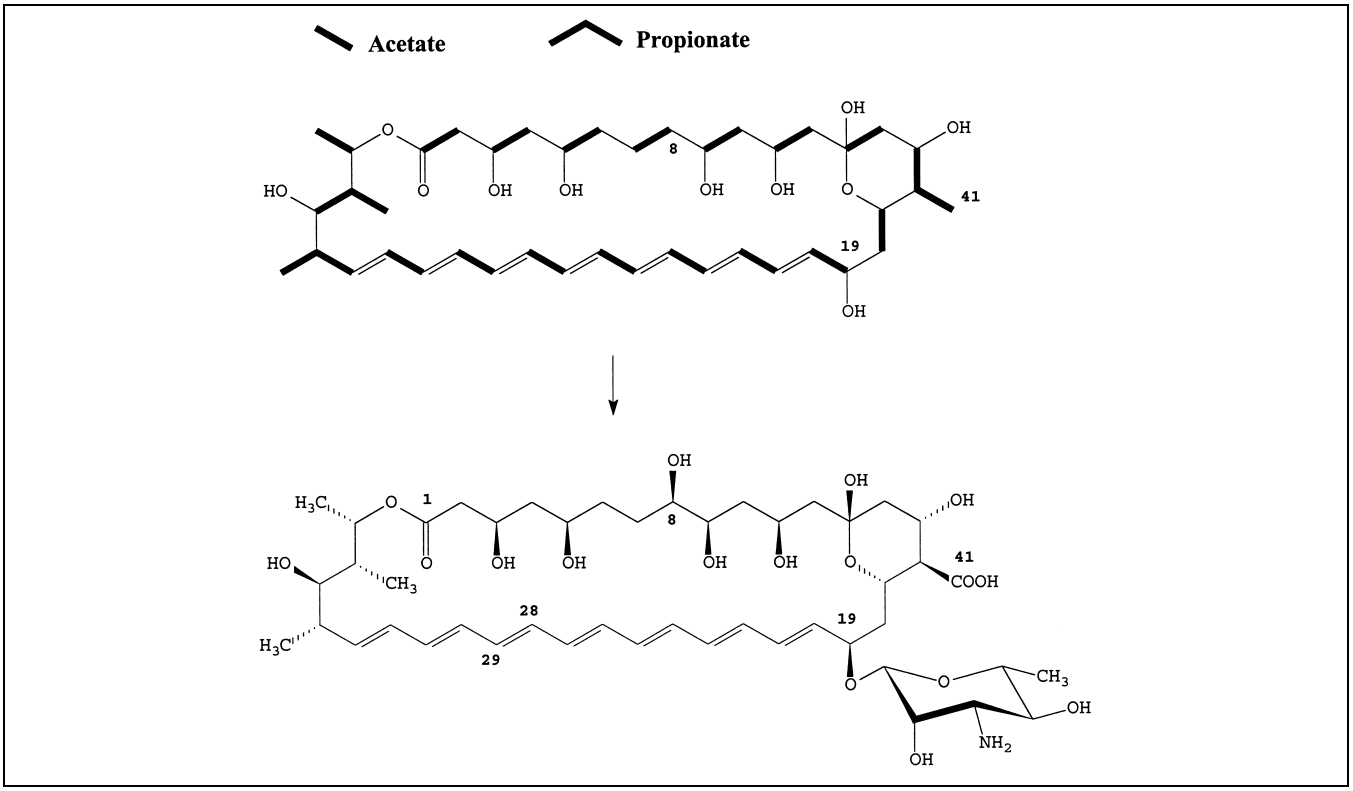

Fig. 1. Biosynthesis of amphotericin B. The polyketide precursor is assembled from acetate and propionate units [23]. After cyclisation, the macrolactone ring is hydroxylated at C8, the hydroxyl group at C19 is mycosaminylated and the propionate-derived C41methyl group is oxidised to form the exocyclic carboxyl group. The order in which these modifications occur is unknown. In the co-metabolite amphotericin A, the double bond between C28 and C29 is reduced.

duced when liposomal formulations of the drug are administered [1]. Chemical modification has also yielded bioactive derivatives with increased water-solubility and reduced toxicity [3]. Further improvements could be made by investigating a wider range of analogues. These could be generated by manipulating amphotericin biosynthesis in *S. nodosus*.

The carbon chains of polyenes are assembled from acetate and propionate units by polyketide synthases (PKSs) [8]. Polyketide biosynthesis begins with the decarboxylative condensation of a dicarboxylic acid extender unit with a starter acyl unit to form a β-ketoacyl intermediate. With complex macrolide polyketides, the β-ketone group may be converted to an alcohol, enoyl or methylene group before another extender unit is added in the next cycle of chain extension. The final product is a unique functionalised acyl chain that may cyclise and undergo further modification by hydroxylation, methylation and glycosylation.

Studies on erythromycin biosynthetic genes first revealed that complex PKSs contain a synthase unit or module for every cycle of chain extension [9,10]. Within each module, an acyl transferase (AT) selects the extender unit and a ketosynthase (KS) and an acyl carrier protein (ACP) carry out the condensation. Ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) enzymes may also be present that determine the extent of β-carbonyl group processing. These enzymes exist as domains within giant multienzyme polypeptides that may contain multiple modules [11].

Biosynthetic gene clusters have now been investigated for several macrolide polyketides. These include avermectin [12], oleandomycin [13], rapamycin [14], FK506 [15] and rifamycin [16]. Manipulation of these genes is a promising approach towards production of novel compounds with improved properties. These studies have now been extended to polyene PKSs. Criado et al. sequenced a *Streptomyces griseus* gene for a *p*-aminobenzoic acid synthase that functions in synthesis of the starter unit for the aromatic heptaene candicidin [17]. Hu et al. cloned the genes responsible for synthesis of FR-008, a polyene that is identical in structure to candicidin except for the presence of an unusual sugar in place of mycosamine [18]. Aparicio et al. cloned and sequenced the genes responsible for the 26-membered tetraene pimaricin [19,20]. The nystatin biosynthetic gene cluster has been completely sequenced by Brautaset et al. [21].

The structure of amphotericin B [22] suggests that the polyketide precursor is synthesised from the following sequence of starter and extender units: $AP_2A_8PA_7$, where A is acetate and P is propionate (Fig. 1). This has been confirmed by chemical labelling experiments carried out

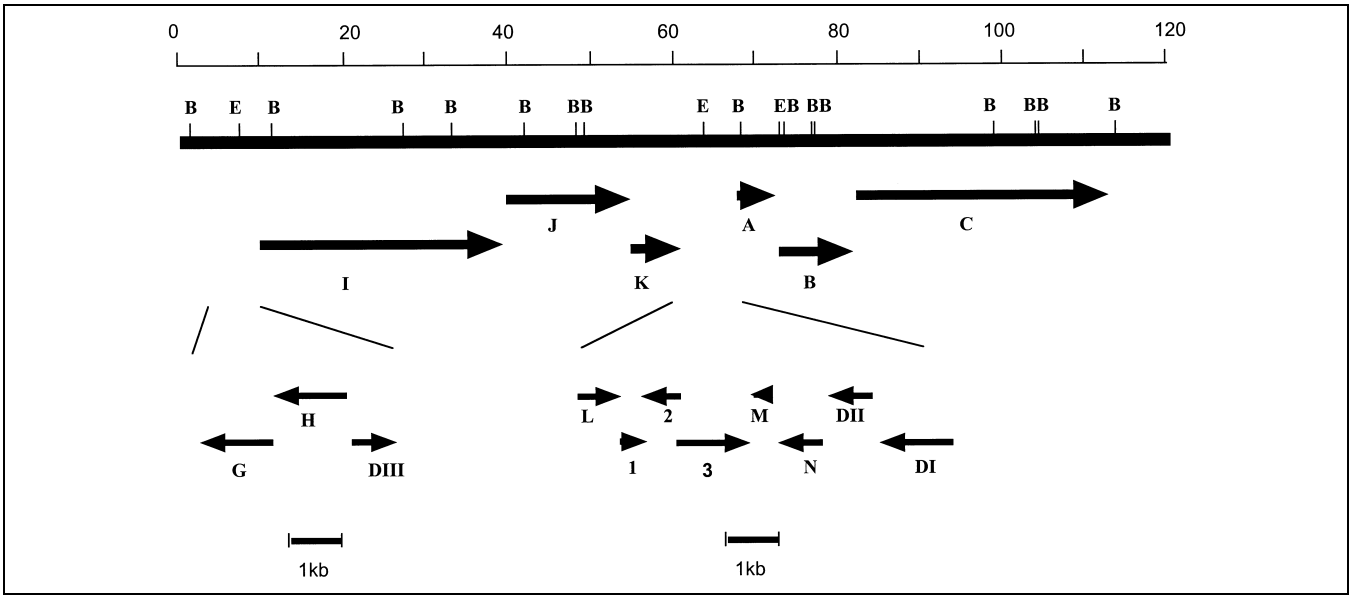

Fig. 2. Organisation of the amphotericin biosynthetic gene cluster. *Eco*RI and *Bam*HI restriction sites are denoted by E and B. Gene designations and proposed functions are listed in Table 1.

by Rawlings and coworkers [23]. After cyclisation, the macrolactone core undergoes three post-polyketide modifications. These are addition of a mycosamine sugar, hydroxylation at C8 and oxidation of a methyl branch (C41) to a carboxyl group. It is not known whether these late steps occur in a preferred order. *S. nodosus* also produces a co-metabolite amphotericin A. This is identical to amphotericin B except that the C28–C29 double bond is reduced.

In this paper we report the cloning and sequencing of amphotericin biosynthetic genes from *S. nodosus*. In addition we show that phage KC515 transduction can be used to introduce DNA into this organism for gene disruption and replacement. This provides the prerequisites for engineered biosynthesis of amphotericin analogues.

2. Results and discussion

2.1. Cloning and sequencing of amphotericin biosynthetic genes

The strategy for cloning the amphotericin biosynthetic genes involved the use of a heterologous probe to isolate PKS genes from *S. nodosus*. A cosmid library was constructed from total cellular DNA. Positive clones were identified using a probe derived from the *eryAII* gene [9,10]. Cosmid DNA was purified from these hybridising clones and characterised by restriction analysis with *Eco*RI, *Kpn*I and *Bam*HI. Overlapping clones were identified and initial contigs were extended by chromosome walking experiments.

Table 1
Amphotericin biosynthetic genes in *S. nodosus*

Gene	Protein	Size, amino acids
amphG	ABC transporter	606
amphH	ABC transporter	607
amphDIII	GDP-mannose-4,6-DH	344
amphI	PKS protein, modules 9–14	9510
amphJ	PKS protein, modules 15–17	5643
amphK	PKS protein, module 18-TE	2034
amphL	Cytochrome P450 monooxygenase	396
ORF1	Hypothetical protein	169
ORF2	Hypothetical protein	284
ORF3	Hypothetical protein	533
amphM	Ferredoxin	65
amphN	Cytochrome P450 monooxygenase	399
amphDII	Sugar aminotransferase	352
amphDI	Glycosyl transferase	483
amphA	PKS protein, loading module	1412
amphB	PKS protein, modules 1 and 2	3190
amphC	PKS protein, modules 3–8	10910

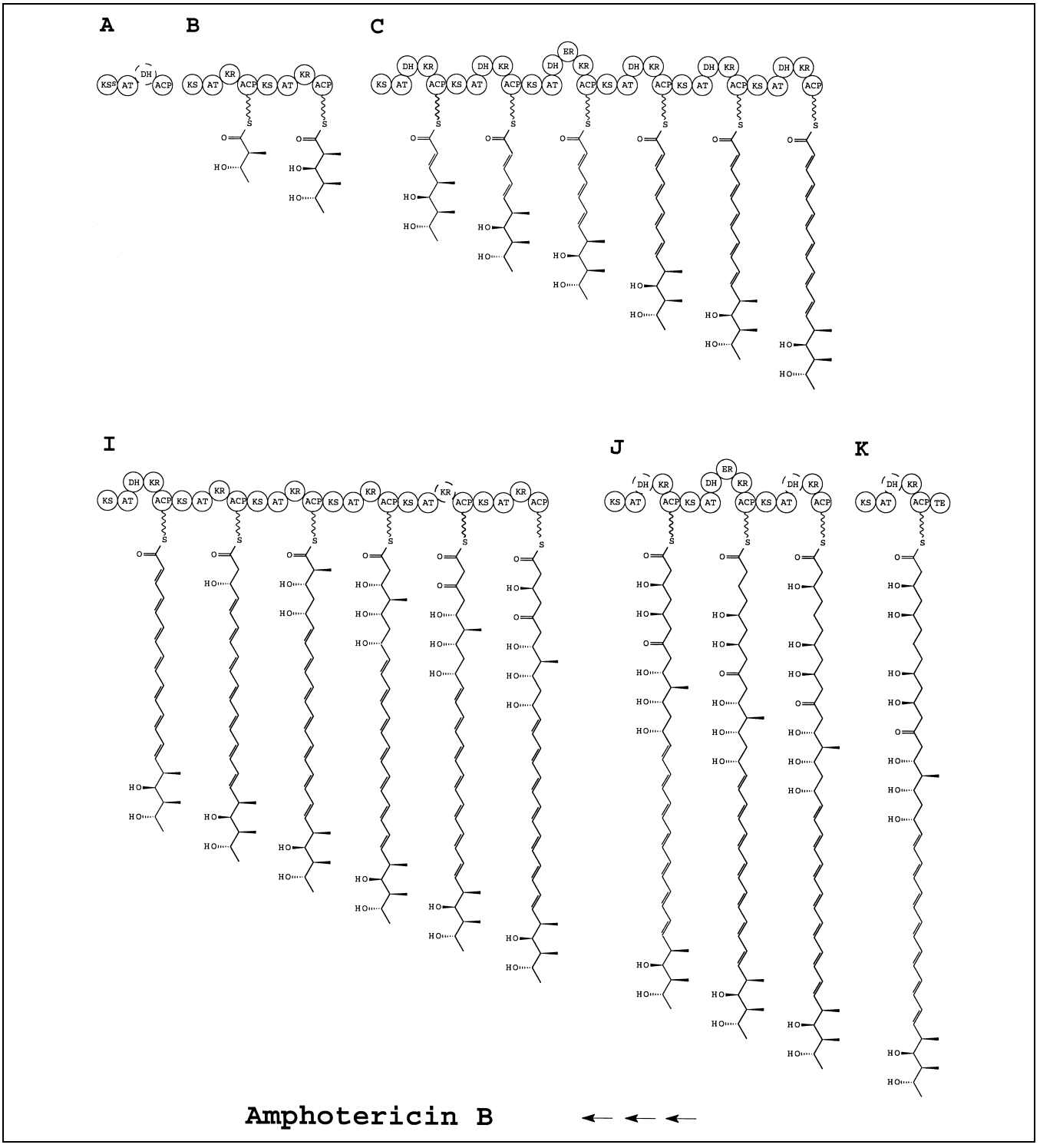

Fig. 3. Modular structure of the amphotericin PKS and intermediates in polyene biosynthesis. Non-functional DH and KR domains are shown as dashed circles. The ER domain in module 5 is thought to be partially active. Reduction of the enoyl group in this cycle would lead to production of amphotericin A.

A series of cosmids was assembled that represented a stretch of DNA containing approximately 100 kb of PKS genes plus at least 15 kb of flanking sequence at either end. This was large enough to accommodate the 18 extension modules anticipated for the amphotericin PKS. Each region of this DNA was represented by at least two independent cosmid clones. This indicated that the cloned DNA was not rearranged.

Preliminary sequencing indicated that the cloned DNA encoded a PKS that was capable of assembling the poly-

ketide precursor of amphotericin B. Complete sequencing of the cluster was then undertaken. Three cosmids representing most of the cluster were sequenced completely. Fragments from cosmids containing both flanking regions were also sequenced.

The amphotericin cluster (Fig. 2, Table 1) is very similar to the biosynthetic gene cluster for the structurally related polyene nystatin [21]. Analogous genes appear in the same order in both clusters. The corresponding genes show a high degree of sequence identity. This suggests a relatively recent common ancestry.

2.2. PKS genes

The amphotericin PKS contains bimodular, hexamodular, trimodular and single module proteins as shown in Fig. 3. It is likely that nascent polyketide chains are more efficiently transferred between successive modules that are covalently joined. The largest PKS proteins sequenced to date contain six modules, which may be the upper limit. No heptamodular PKS proteins have yet been identified. The fidelity and processivity of bacterial protein synthesis [24] may not be high enough to allow sufficiently accurate translation of PKS genes of this size.

The *amphA* gene encodes a loading module with the domain structure KS^S–AT–DH–ACP. Direct linkage of a DH to an ACP domain is unusual. The AT domain has the signature sequence characteristic of a malonyl transferase [25] and probably loads malonyl groups onto the ACP domain. The KS^S domain has a serine residue in place of the conserved active site cysteine. This domain may act as a decarboxylase that acts on malonyl-ACP to generate acetyl starter units.

KS domains are converted to potent decarboxylases when the active site cysteine is replaced with glutamine [26,27]. KS^Q enzymes appear in loading modules for other macrolide PKSs and also in aromatic PKSs. KS^Q domains convert malonyl or methylmalonyl groups to acetyl or propionyl starter units. This may represent an efficient means of delivering primers to KS1 (or KSα) and may also allow for stricter control of starter unit selection [26]. It is uncertain whether KS^S domains are equally efficient at generating primers. The active site cysteine-161 of the KS domain of rat fatty acid synthase (FAS) has been replaced with various amino acids [27]. The cysteine → serine change gave a mutant enzyme that retained a low residual condensation activity and had only a weak decarboxylase activity. In polyene PKSs the activity of the decarboxylase may have been attenuated so the rate of primer delivery matches the overall rate of polyene synthesis by the extension modules. The amphotericin KS^S domain retains the residues corresponding to histidine-303, aspartate-311, glutamate-314, lysine-335 and histidine-340 in the active site of *Escherichia coli* KSII [28]. In the evolution of the loading module the cysteine → serine change requires only a single point mutation and may be sufficient to give an adequate decarboxylase. However, this change does not completely uncouple decarboxylation from C–C bond formation in rat FAS. The rat mutant KS^S generates β-ketobutyryl groups that are off-loaded onto CoA [27]. This would represent a wasteful side reaction of AmphA since these starter units do not appear in the major polyene products made by *S. nodosus*. These considerations indicate that conversion of the KS^S domain to a KS^Q domain could improve the efficiency of the AmphA loading module and increase the yield of polyenes produced.

The DH domain in AmphA is presumably redundant since it would not normally encounter a β-hydroxyacyl-ACP substrate.

The AmphB protein contains the first two extension modules. As anticipated, these modules both contain methylmalonate-specific AT domains and reductive loops containing only KR domains. These modules are similar to modules 1 and 2 of 6-deoxyerythronolide B synthase (DEBS) [9,10] except that the amphotericin module 1 generates a 2*S*,3*S*-2-methyl-3-hydroxy butanoic acyl thioester whereas DEBS module 1 generates a 2*S*,3*R*-2-methyl-3-hydroxy pentanoic acyl chain. Amphotericin module 11 also generates a 2*S*,3*S*-2-methyl-3-hydroxyacyl chain. However, sequence alignments provide no obvious clues as to how KRs select a particular stereochemistry at C2 in the extended β-ketoacyl chain or how the stereochemical outcome of the ketoreduction is determined.

AmphC is a hexamodular protein that assembles most of the polyene unit of amphotericin. Modules 3–8 all contain a DH–KR reduction loop except module 5 which has DH–ER–KR. The interdomain region preceding ER5 is 45 amino acids shorter than its counterpart in the NysC protein and similar complete reduction loops in other PKS proteins. The shorter interdomain may restrict movement of the amphotericin ER5 domain so that some nascent chains are transferred from ACP5 to KS6 before enoyl reduction takes place. This kinetic competition might explain why *S. nodosus* produces amphotericin B as well as amphotericin A, in which the C28–C29 double bond is reduced. We cannot yet exclude the possibility that the PKS produces exclusively the polyketide precursor of amphotericin A, with amphotericin B being produced as a result of reintroduction of the C28–C29 double bond by a specific desaturase. However, the shortened linker domain favours the hypothesis that two polyketides are made as a result of a lapse in programming fidelity that results from a sluggish ER domain. A number of studies have found that reduction domains can fail to act when inserted into recombinant PKS modules. The ER domain from DEBS did not operate when the entire reductive loop from module 4 was relocated into a recombinant trimodular tetraketide synthase [29]. An equivalent reductive loop from the rapamycin PKS module 1 was fully functional in the same context [30]. DEBS KR6 was also nonfunctional when placed in module 2 of the DEBS1-TE triketide lactone synthase [31]. These studies suggest that

sequences flanking reduction domains can affect their action during cycles of chain extension.

The relative levels of amphotericins A and B could be influenced by the overall rate of flux through the PKS assembly line. Complete inactivation of ER5 by gene replacement should give an *S. nodosus* strain that exclusively synthesises amphotericin B, the more active antifungal agent.

It is remarkable that the *amphC* gene is stable in spite of its repetitive nature. The *amphC* coding sequences for the various KS domains are very similar and contain stretches of identical sequence of up to 250 bp. These regions appear to be limited enough to prevent instability resulting from excision of DNA after homologous recombination between direct repeats.

The *amphI* gene encodes another hexamodular protein that incorporates modules 9 to 14. The protein is similar to the NysI protein. Module 11 has a methylmalonate-specific AT domain and incorporates the methyl branch that is oxidised to form the exocyclic carboxyl group. Only a single non-functional reduction domain is present in the entire protein. Module 13 has an inactive KR domain in which three amino acids have been deleted from the NADPH-binding motif.

The *amphJ* gene encodes modules 15 to 17. Modules 15 and 17 both contain DH–KR reductive loops. The structure of the polyketide suggests that neither of these DH domains is functional. However, DH15 contains the conserved HXXXGXXXXP motif found in active DH domains [32]. It is possible that the interdomain linker regions constrain domain movements so that access to the substrate is denied. DH17 is clearly inactive since the active site histidine is replaced by an arginine residue. Module 16 contains a complete DH–ER–KR reductive loop. The main difference between the amphotericin and nystatin PKSs is that the reductive loops are different in modules 15 and 16. In the Nys J protein module 15 contains a complete reductive loop whereas modules 16 and 17 contain DH–KR reductive loops in which the DH domains must be non-functional although they contain conserved active site motifs.

The presence of these non-functional DH domains suggests that AmphJ could have arisen from the DNA coding for modules 4–6 (Fig. 3) whereas NysJ could have arisen from modules 5–7. A duplicated *amphC* or *nysC* gene may have undergone deletions to remove modules on either side of module 5. The final structures of the *amphJ* and *nysJ* genes may not have stabilised until after the divergence of the amphotericin and nystatin producers.

The *amphK* gene encodes module 18 and the chain-terminating thioesterase domain. Module 18 also contains a DH domain that is apparently non-functional.

2.3. Cytochrome P450 enzymes

The products of the *amphL* and *amphN* genes resemble cytochrome P450 enzymes. Both proteins contain the two sequences that are conserved in other P450s. These are the O_2-binding motif and the region containing a cysteine residue that coordinates the haem iron [33]. The *amphN* gene is adjacent to *amphM* which encodes a ferredoxin. The AmphL and AmphN proteins probably carry out post-polyketide modifications during biosynthesis of amphotericin. It is likely that one catalyses the hydroxylation at C8 and the other functions in oxidation of the C41 methyl branch to a carboxyl group. During bile acid biosynthesis, sterol 27-hydroxylase converts the C27 methyl group of cholesterol first to a CH_2OH group and then to a carboxyl group [34]. There is therefore a precedent for a single P450 oxidising a methyl group all the way to a carboxyl group.

The AmphN and NysN proteins show 84% sequence identity whereas the the AmphL and NysL proteins show 71% sequence identity. It is likely that the more homologous enzymes (AmphN and NysN) are involved in formation of the carboxyl groups. AmphL and NysL probably hydroxylate their macrolactone rings at C8 and C10 respectively.

The pimaricin cluster also contains two genes for cytochrome P450s, *pimD* and *pimG* [20]. PimG is similar to AmphN and NysN whereas PimD is similar to AmphL and NysL. Disruption of PimD led to synthesis of 4,5 deepoxypimaricin [35]. This supports the view that AmphL, NysL and PimD oxidise carbon atoms within the polyol regions whereas AmphN, NysN and PimG probably form carboxyl groups.

Attempts were made to disrupt the *amphN* gene in an effort to test this hypothesis. However, two different recombinant KC515 phage containing internal fragments of *amphN* failed to disrupt the chromosomal copy of the gene. A single thiostrepton-resistant lysogen was obtained that contained a large chromosomal deletion rather than a clean disruption. This mutant did not produce a polyene (E. Flood and P. Caffrey, unpublished results). It is possible that disruption of some late genes will lead to synthesis of toxic or insoluble products that can not be exported from the cell. Further attempts to disrupt *amphN* are in progress. Successful disruption of this gene could result in biosynthesis of a valuable amphotericin analogue with a methyl group in place of the exocyclic carboxyl group. Chemical modification of amphotericin B has shown that suppression of charge on the free carboxyl group improves specificity towards ergosterol-containing fungal membranes, although ester derivatives were neurotoxic [36,37].

Another interesting experiment would be to determine whether AmphL can hydroxylate nystatin at C8 and whether NysL can hydroxylate amphotericin B at C10. Introduction of an extra hydroxyl group could improve solubility of either polyene. However it may be that both P450s recognise a region of the macrolactone ring consisting of three methylene groups between two carbon

atoms bearing hydroxyl groups, in which case only a single hydroxylation would be possible.

2.4. Glycosylation

The AmphDI protein shows weak homology to UDP-glucuronate transferases and probably functions in addition of mycosamine to the aglycone core of amphotericin.

Like the nystatin and pimaricin clusters, the amphotericin cluster contain a gene (*amphDIII*) for a GDP-mannose-4,6-DH. This suggests that the biosynthetic pathway to mycosamine (3,6-dideoxy-3-amino-D-mannose) involves isomerisation of GDP-6-deoxy-4-ketomannose to GDP-6-deoxy-3-ketomannose, followed by an amino transfer to give GDP-mycosamine. This final step is probably catalysed by AmphDII protein which is similar to aminotransferases involved in biosynthesis of perosamine (4,6-dideoxy-4-aminomannose).

There is no gene common to the amphotericin, nystatin and pimaricin clusters that is likely to encode a GDP-6-deoxy-4-ketomannose-3,4-isomerase. This is remarkable because all of the other biosynthetic genes appear to be located close to the PKS genes. It has been suggested that the *eryCII* gene encodes a dTDP-6-deoxy-4-ketoglucose-3,4-isomerase that functions in the biosynthesis of desosamine [38]. Homologous genes have been found in clusters for other macrolides that are glycosylated with desosamine, mycaminose or daunosamine [38,39]. Biosynthesis of all of these amino sugars involves 3,4-isomerisation of dTDP-6-deoxy-4-ketoglucose [38]. These putative isomerases are similar to cytochrome P450 monooxygenases but lack the conserved cysteine residue that coordinates the haem iron.

The polyene biosynthetic genes may provide a clue why proteins resembling cytochrome P450s should act as dTDP-ketosugar isomerases. The hemiketalic portion of polyenes shows some similarity in overall shape to a deoxy ketosugar hemiacetal. Hemiketals also form in other macrolides. It is therefore possible that the EryCII protein and its homologues arose from cytochrome P450s that act on macrolides, retaining a binding site for a hemiacetal but not the iron haem group. The prosthetic group would not be required for the more facile isomerisation reaction. The putative dTDP-ketosugar isomerases show a weak similarity to AmphN which presumably binds to the hemiketal part of the amphotericin precursor to oxidise the C41 methyl group.

The structure of perimycin A also suggests a link between P450s and isomerases. This aromatic heptaene is not modified by P450s and has a methyl branch in place of the exocyclic carboxyl group found in most other glycosylated polyenes [40]. Perimycin A is also unusual among polyenes in that it is glycosylated with perosamine rather than mycosamine. Perosamine is synthesised from GDP-6-deoxy-4-ketomannose without isomerisation. The perimycin producer therefore lacks both the P450 and the ketosugar isomerase. This striking coincidence is consistent with the notion that a GDP-mannose-6-deoxy-4-ketomannose-3,4-isomerase may also have evolved from a P450 which is absent from the perimycin producer. We are investigating perimycin biosynthesis to determine whether there is a significant association between glycosylation with perosamine and absence of a free carboxyl group.

The isomerisation step in mycosamine biosynthesis is one of the last remaining questions in polyene biosynthesis. It is likely that answers will come from further sequencing and gene disruptions together with rigorous chemical analysis of polyene analogues.

2.5. Export

The AmphG and AmphH proteins show strong homology to ATP-binding cassette (ABC) transporters. There is only 25% sequence identity between AmphG and AmphH. It is unclear why two different transporter proteins are present. The nystatin and pimaricin transporters gene clusters also contain two transporters each, NysG and NysH [20], and PimA and PimB [21]. The most plausible start codon for the *amphG* reading frame occurs 17 bp before the stop codon for *amphH*. The 3′ end of the *nysH* gene also overlaps the apparent 5′ end of *nysG* by 20 bp and *pimA* similarly overlaps *pimB* by 23 bp.

Many bacterial ABC transporters are thought to function as homodimers [41]. It is possible that AmphG and AmphH associate to form a heterodimer that exports amphotericins from the producing cell. The polyene exporter would then resemble the 170-kDa P-glycoprotein that confers multidrug-resistance on eukaryotic tumour cells. The P-glycoprotein has two transmembrane domains (each composed of six α-helices) and two nucleotide-binding domains [41]. This protein is capable of exporting a wide range of structurally and functionally unrelated compounds. It remains to be seen whether the amphotericin transporter will export all of the analogues that could be generated by engineered biosynthesis.

An ABC transporter confers self-resistance to the ionophore tetronasin on the producing organism *Streptomyces longisporoflavus* [42]. This transporter is thought to actively pump ions back into cells that have been permeabilised by tetronasin. Polyene-producing streptomycetes could require a similar self protection system if even trace levels of exogenous sterols were to enter their membranes. An alternative role for one of the ABC transporters could be to afford this protection.

2.6. Additional genes

Open reading frames (ORFs) 1–3 have no homologues in the nystatin or pimaricin clusters. It is therefore unlikely that these ORFs have an important role in polyene biosynthesis. ORF1 encodes a 170-amino acid protein with 39% sequence identity with a 13-kDa protein found in the

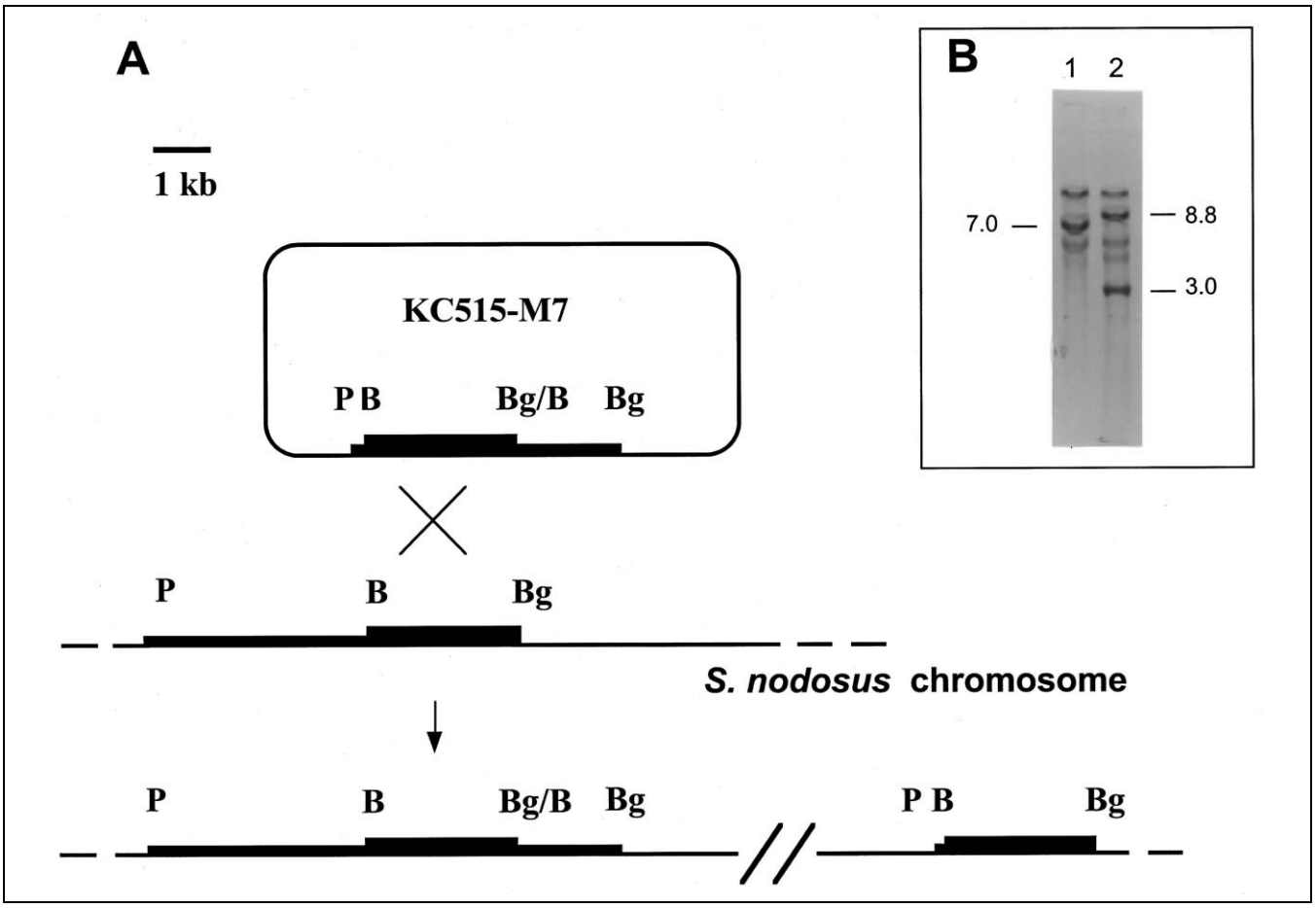

Fig. 4. Targeted disruption of the *amphC* gene. A: Schematic diagram showing integration of recombinant phage KC515-M7 DNA into the *S. nodosus* chromosome by homologous recombination. *Pst*I, *Bam*HI and *Bgl*II restriction sites are denoted by P, B and Bg. Bg/B is a hybrid *Bgl*II–*Bam*HI site. The phage carries a 3.0-kb *Bam*HI–*Bgl*II fragment from within the *amphC* gene (black rectangle). Integration results in loss of a 7-kb *Bgl*II–*Pst*I chromosomal fragment and formation of 8.8-kb and 3-kb *Bgl*II–*Pst*I fragments. These fragments are drawn as heavy lines. B: Analysis of chromosomal DNA by Southern hybridisation with 3-kb *Bgl*II–*Bam*HI fragment probe. Lanes 1 and 2 contain *Bgl*II–*Pst*I digests of chromosomal DNA from wild-type *S. nodosus* and the disruption mutant, respectively. The positions of the 7.0-, 8.8- and 3.0-kb hybridising bands are indicated. The probe fragment includes part of the coding sequence for a highly conserved KS domain and hybridises less strongly to other fragments even at high stringency. A faint 4.8-kb band in lane 2 may be derived from free phage DNA that loops out and replicates in a small proportion of the population.

mini-circle, a 2527-bp transposable element from *Streptomyces coelicolor* [43]. ORF1 may represent the remnants of an integrated mini-circle because neither of the other two mini-circle protein genes are present. ORF2 shows 38% sequence identity and 55% similarity to a hypothetical protein from the plant pathogenic bacterium *Xylella fastidiosa* (accession number AAF82888). ORF3 shows 55% identity with a hypothetical membrane protein from *S. coelicolor* (accession number CAB44413.1).

Sequence analysis of the regions flanking *amphH* and *amphC* has been carried out. No genes with an obvious role in amphotericin biosynthesis were identified downstream from *amphG*. The nystatin cluster has a candidate gene for a phosphopantetheinyl transferase in this region but no homologous gene was found within 5 kb of *amphG*. Sequence analysis of the region beyond *amphC* is still incomplete although genes for phosphomannomutase, phosphomannoisomerase and GDP-mannose synthetase have been found in this region (M. Oliynyk and P. Caffrey, unpublished results). These function in synthesis of GDP-mannose from fructose 6-phosphate, an intermediate in glycolysis. This region also contains a number of ORFs whose translation products resemble hypothetical proteins of unknown function. Gene disruption experiments are in progress to ascertain whether these have an essential role in amphotericin biosynthesis.

2.7. Disruption of PKS genes

The ϕC31-derived actinophage KC515 was found to form plaques on *S. nodosus*. This suggested that KC515-mediated transduction could be used for introduction of DNA for gene disruption and replacement. A 5182-bp *Bam*HI fragment from within the *amphC* gene (nucleotides

100297–105480 in the amphotericin cluster sequence) was subcloned into pUC118. A 3012-bp *Bgl*II–*Bam*HI subfragment (nucleotides 102468–105480) was excised from pUC118 using *Bgl*II and *Pst*I and subcloned between the *Bam*HI and *Pst*I sites of KC515. This 3012-bp region contains the coding sequence for part of module 7 (half of KS7 plus AT7-DH7). The recombinant phage, KC515-M7, was used to infect *S. nodosus* and lysogens were obtained by selecting for the thiostrepton resistance gene within the prophage DNA. Genomic DNA from a typical lysogen was digested with several restriction enzymes and analysed by Southern hybridisation (e.g. Fig. 4). This revealed that the phage had integrated into the PKS gene. The disruption mutant, *S. nodosus* DM7, did not produce amphotericin B, confirming that the biosynthetic gene cluster for this polyene had been cloned. Further evidence supporting this conclusion was obtained from a second gene disruption experiment.

A 2013-bp *Kpn*I fragment from within the *amphI* gene (nucleotides 9265 to 11278) was subcloned into pUC118. A 1667-bp *Kpn*I–*Bgl*II fragment of this plasmid contained the coding sequence (nucleotides 9265 to 10918) for the KR and ACP of module 9 and approximately half of the KS of module 10. This DNA was excised as a *Bgl*II–*Pst*I fragment and subcloned between the *Bam*HI and *Pst*I sites of KC515. The recombinant phage, KC515-M9, was used to infect *S. nodosus* and thiostrepton-resistant lysogens were selected. Analysis of genomic DNA from a typical lysogen indicated that integration of a phage had disrupted the PKS gene (data not shown). This mutant, *S. nodosus* DM9, also failed to produce amphotericin B.

These results show that the cloned genes function in amphotericin biosynthesis. The availability of at least one method for manipulating these *S. nodosus* genes should allow rational design of novel polyenes.

3. Significance

Amphotericin B is the most important drug for the treatment of life-threatening fungal infections in humans. The incidence of these infections has increased over the past 20 years [44]. There is an urgent need to develop amphotericin derivatives with reduced toxicity.

We have cloned the amphotericin biosynthetic genes from *S. nodosus* and shown that phage KC515 can be used to introduce DNA into this organism for gene disruption and replacement. This should allow engineered biosynthesis of improved antifungal drugs. These novel compounds could also be investigated as potential antiviral, antiprion or antiparasitic agents.

The amphotericin PKS consists of 18 extension modules and is one of the largest for which a sequence is available. The amphotericin system can now be used to engineer the biosynthesis of libraries of larger macrolide structures for drug discovery.

4. Materials and methods

4.1. Bacterial strains

S. nodosus ATCC 14899 was used as the source of DNA in the construction of a genomic library. *E. coli* strain XL1-Blue MR was used as a host for propagation of cosmids and plasmids. *Streptomyces lividans* 66 (John Innes strain 1326) was used for propagation of phage KC515 [45].

4.2. Assays for amphotericin B production

For amphotericin production, *S. nodosus* was grown on fructose–dextrin medium [23] (fructose 20 g/l, dextrin 60 g/l, soya flour 30 g/l $CaCO_3$ 10 g/l (pH 7.0)). Samples of the culture were withdrawn at 12-h intervals and supernatants were assayed for amphotericin by bioassay using *Saccharomyces cerevisiae* NCYC 1006 as an indicator organism. Alternatively, 100-µl samples of culture were mixed with 900-µl volumes of dimethylsulphoxide and sonicated for 20 min. The extract was centrifuged and the supernatant was diluted with methanol. The absorption spectrum was measured in the wavelength range 190–450 nm. Amphotericin B gives four specific absorption peaks at 346, 364, 382 and 405 nm [23]. Amphotericin A absorbs at 280, 292, 305 and 320 nm. Amphotericin B standard was obtained from Sigma.

4.3. Genetic procedures

Total cellular DNA was isolated from *S. nodosus* as described [46]. The DNA was partially digested with *Sau*3A and fragments in the 35–45-kb size range were isolated by sucrose density gradient centrifugation [47]. These fragments were ligated to *Bam*HI-cut dephosphorylated pWE15 cosmid vector and packaged into lambda phage particles using the Stratagene Gigapack II system. The library was plated on *E. coli* XL1-Blue MR and clones were screened by colony hydridisation using a probe derived from a DNA fragment containing the entire *eryAII* gene sequence [32]. Labelling of probes with digoxigenin-dUTP and detection of hybridising DNA were carried out using a Boehringer Mannheim DIG DNA labelling and detection kit. For Southern hybridisation, DNA fragments were transferred to nylon membranes by capillary transfer [47].

Cosmid and plasmid DNA was isolated using the alkaline-SDS procedure [47]. Restriction enzyme digestions and ligations were carried out using standard procedures. Competent *E. coli* cells for transformation were prepared by the calcium chloride method [47].

Methods for preparation of streptomycete spores and protoplasts were those described by Hopwood et al. [46]. Phage KC515 DNA was isolated using the small-scale method [46]. DNA fragments ligated to KC515 DNA were introduced into *S. lividans* 66 protoplasts by transfection [46]. Recombinant phage were identified by PCR using primers specific for the cloned insert.

PCRs were carried out with *Taq* polymerase as described [47] using a Perkin Elmer GeneAmp 2400 thermocycler. The oligonucleotide primers were synthesised by GenoSys, Cambridge, UK.

To obtain disruption mutants, recombinant phage were plated on *S. nodosus* spores to give near confluent lysis. After overnight incubation, plates were overlaid with soft nutrient agar containing 50 μg thiostrepton per ml. Thiostrepton-resistant lysogens resulting from integration of the phage were streaked on tryptone soya (TS) agar containing thiostrepton.

4.4. DNA sequencing

DNA sequencing was carried out at Biotica Technology Ltd., Cambridge, UK. Intact cosmids were partially digested with *Sau*3A1 and fragments in the size range 1 to 2 kb were purified by preparative agarose gel electrophoresis and subcloned into *Bam*HI-digested plasmid pBCSK+ (Stratagene). Plasmid subclones were isolated using Qiagen columns. DNA concentrations were measured using a Pharmacia Gene Quant spectrophotometer. Automated sequencing was done on double-stranded DNA templates by the dideoxynucleotide chain termination method with an Applied Biosystems model 373A sequencer. Sequence data obtained from single random subclones was assembled into a single continuous sequence and edited using the GAP4.1 program of the STADEN gene analysis package [48]. DNA and protein sequence homology searches of data bases were done using the BLAST programme [49].

The DNA sequence has been deposited in the GenBank database under the accession number AF357202.

Acknowledgements

Work in the authors' laboratory was supported by grants from the Wellcome Trust and Enterprise Ireland. We thank Professor P.F. Leadlay for helpful discussions, Dr. Jesus Cortes for advice on library construction and Dr. Lake Ee Khaw for advice on the use of the phage KC515 system.

References

[1] K.M. Abu-Salah, Amphotericin B: an update, Br. J. Biomed. Sci. 53 (1996) 122–133.
[2] J. Bolard, How do polyene macrolide antibiotics affect cellular membrane properties?, Biochim. Biophys. Acta 864 (1986) 257–304.
[3] S. Hartsel, J. Bolard, Amphotericin B: new life for an old drug, Trends Pharmacol. Sci. 17 (1996) 445–449.
[4] C.P. Schaffner, O.J. Plescia, D. Pontani, D. Sun, A. Thornton, R.C. Pandey, P.S. Sarin, Anti-viral activity of amphotericin B methyl ester: inhibition of HTLV-III replication in cell culture, Biochem. Pharmacol. 35 (1986) 4110–4113.
[5] R.C. Aloia, H. Tian, F.C. Jensen, Lipid composition and fluidity of the human immunodeficiency virus envelope and host cell plasma membranes, Proc. Natl. Acad. Sci. USA 90 (1993) 5181–5185.
[6] M. Pocchiari, S. Schmittinger, C. Masullo, Amphotericin B delays the incubation period of scrapie in intracerebrally inoculated hamsters, J. Gen. Virol. 68 (1987) 219–223.
[7] A. Mange, O. Milhavet, H.E.M. McMahon, D. Casanova, S. Lehman, Effect of amphotericin B on wild-type and mutated prion proteins in cultured cells: putative mechanism of action in transmissible spongiform encephalopathies, J. Neurochem. 74 (2000) 754–762.
[8] J.F. Martin, Biosynthesis, regulation and genetics of polyene macrolide antibiotics, in: S. Omura (Ed.), Macrolide Antibiotics: Chemistry, Biology and Practice, Academic Press, New York, 1986, pp. 405–424.
[9] J. Cortes, S.F. Haydock, G.A. Roberts, D.J. Bevitt, P.F. Leadlay, An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*, Nature 348 (1990) 176–178.
[10] S. Donadio, M.J. Staver, J.B. McAlpine, S.J. Swanson, L. Katz, Modular organisation of genes required for complex polyketide biosynthesis, Science 252 (1991) 675–679.
[11] P. Caffrey, D.J. Bevitt, J. Staunton, P.F. Leadlay, Identification of DEBS1, DEBS2 and DEBS3, the multienzyme polypeptides of the erythromycin-producing polyketide synthase from *Saccharopolyspora erythraea*, FEBS Lett. 304 (1992) 225–228.
[12] D.J. MacNeil, J.L. Occi, K.M. Gewain, T. MacNeil, P.H. Gibbons, C.L. Ruby, S.J. Danis, Complex organisation of the *Streptomyces avermitilis* genes encoding the avermectin polyketide synthase, Gene 115 (1992) 119–125.
[13] D.G. Swan, A.M. Rodriguez, C. Vilches, C. Mendez, J.A. Salas, Characterisation of a *Streptomyces antibioticus* gene encoding a type I polyketide synthase gene which has an unusual coding sequence, Mol. Gen. Genet. 242 (1994) 358–362.
[14] T. Schwecke, J.F. Aparicio, I. Molnar, A. Konig, L.-E. Khaw, S.F. Haydock, M. Oliynyk, P. Caffrey, J. Cortes, J.B. Lester, G.A. Bohm, J. Staunton, P.F. Leadlay, The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin, Proc. Natl. Acad. Sci. USA 92 (1995) 7839–7843.
[15] H. Motamedi, A. Shafiee, The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506, Eur. J. Biochem. 256 (1998) 528–534.
[16] P.R. August, L. Tang, Y.J. Yoon, S. Ning, R. Muller, T.W. Yu, M. Taylor, D. Hoffmann, C.G. Kim, X.H. Zhang, C.R. Hutchinson, H.G. Floss, Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S699, Chem. Biol. 5 (1998) 69–79.
[17] L.M. Criado, J.F. Martin, J.A. Gil, The *pab* gene of *Streptomyces griseus*, encoding *p*-aminobenzoic acid synthase, is located between genes possibly involved in candicidin biosynthesis, Gene 126 (1993) 135–139.
[18] Z. Hu, K. Bao, X. Zhou, Q. Zhou, D.A. Hopwood, T. Kieser, Z. Deng, Repeated polyketide synthase modules involved in biosynthesis of a heptaene macrolide by *Streptomyces* sp. FR-008, Mol. Microbiol. 14 (1994) 163–172.
[19] J.F. Aparicio, A.J. Colina, E. Ceballos, J. Martin, The biosynthetic gene cluster for the 26-membered ring polyene macrolide pimaricin: a new polyketide synthase organisation encoded by two subclusters separated by functionalisation genes, J. Biol. Chem. 274 (1999) 10133–10139.
[20] J.F. Aparicio, R. Fouces, M.V. Mendes, N. Olivera, J.F. Martin, 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 (2000) 895–905.
[21] T. Brautaset, O.N. Sekurova, H. Sletta, T.E. Ellingsen, A.R. Strom, S. Valla, S.B. Zotchev, Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 1145: analysis of the gene cluster and deduction of the biosynthetic pathway, Chem. Biol. 7 (2000) 395–403.
[22] W. Mechlinski, C.P. Schaffner, P. Ganis, G. Avitabile, Structure and absolute configuration of the polyene macrolide antibiotic amphotericin B, Tetrahedron Lett. 44 (1970) 3873–3876.
[23] C.M. McNamara, S. Box, J.M. Crawforth, B.S. Hickman, T.J. Nor-

wood, B.J. Rawlings, Biosynthesis of Amphotericin B, J. Chem. Soc. Perkin Trans. 1 (1998) 83–87.
[24] C.G. Kurland, Translational accuracy and the fitness of bacteria, Annu. Rev. Genet. 26 (1992) 29–50.
[25] S.F. Haydock, J.F. Aparicio, I. Molnar, T. Schwecke, A. Konig, A.F.A. Marsden, I.S. Galloway, J. Staunton, P.F. Leadlay, Divergent sequence motifs correlated with the substrate specificity of (methyl)-malonyl-CoA: acyl carrier protein transacylase domains in modular polyketide synthases, FEBS Lett. 374 (1995) 246–248.
[26] C. Bisang, P.F. Long, J. Cortes, J. Westcott, J. Crosby, A.L. Matharu, R.J. Cox, T.J. Simpson, J. Staunton, P.F. Leadlay, A chain initiation factor common to both aromatic and modular polyketide synthases, Nature 401 (1999) 502–505.
[27] A. Witkowska, A.K. Joshi, Y. Lindqvist, S. Smith, Conversion of a β-ketoacyl synthase to a malonyl decarboxylase by replacement of the active site cysteine with glutamine, Biochemistry 38 (1999) 11643–11650.
[28] W. Huang, J. Jia, P. Edwards, K. Dehesh, G. Schneider, Y. Lindqvist, Crystal structure of β-ketoacyl-acyl carrier protein synthase II from *E. coli* reveals the molecular architecture of condensing enzymes, EMBO J. 17 (1998) 1183–1191.
[29] R. McDaniel, C.M. Kao, H. Fu, P. Hevezi, C. Gustafsson, M. Betlach, G. Ashley, D.E. Cane, C. Khosla, Gain of function mutagenesis of a modular polyketide synthase, J. Am. Chem. Soc. 119 (1997) 4309–4310.
[30] C.M. Kao, M. McPherson, R. McDaniel, H. Fu, D.E. Cane, C. Khosla, Gain of function mutagenesis of a modular polyketide synthase. 2. Engineered biosynthesis of eight-membered ring teraketide lactone, J. Am. Chem. Soc. 119 (1997) 11339–11340.
[31] R. McDaniel, C.M. Kao, S.J. Hwang, C. Khosla, Engineered intermodular and intramodular polyketide synthase fusions, Chem. Biol. 4 (1997) 667–674.
[32] D.J. Bevitt, J. Cortes, S.F. Haydock, P.F. Leadlay, 6-Deoxyerythronolide B synthase 2 from *Saccharopolyspora erythraea*. Cloning of the structural gene, sequence analysis and inferred domain structure of a multifunctional enzyme, Eur. J. Biochem. 204 (1992) 39–49.
[33] T.L. Poulos, B.C. Finzel, A. Howard, High-resolution crystal structure of cytochrome $P450_{CAM}$, J. Mol. Biol. 195 (1987) 687–700.
[34] D.W. Russell, K.D.R. Setchell, Bile acid biosynthesis, Biochemistry 31 (1992) 4737–4749.
[35] M.V. Mendes, E. Recio, R. Fouces, R. Luiten, J.F. Martin, J.F. Aparicio, Engineered biosynthesis of novel polyenes: a pimaricin derivative produced by targeted gene disruption in *Streptomyces natalensis*, Chem. Biol. 8 (2001) 637–646.
[36] M. Cheron, B. Cybulska, J. Mazerski, J. Gryzbowska, A. Czerwinski, E. Borowski, Quantitative structure–activity relationships in amphotericin B derivatives, Biochem. Pharmacol. 37 (1988) 827–836.
[37] R.M. Parmegiani, B. Loebenberg, B. Antonacci, T. Yarosh-Tomaine, R. Scupp, J.J. Wright, P.J.S. Chiu, G.H. Miller, Comparative in vitro and in vivo evaluation of *N*-D-ornithyl amphotericin B methyl ester, amphotericin B methyl ester, and amphotericin B, Antimicrob. Agents Chemother. 31 (1987) 1756–1760.
[38] K. Salah-Bey, M. Doumith, J.-M. Michel, S. Haydock, J. Cortes, P.F. Leadlay, M.C. Raynal, Targeted gene inactivation for the elucidation of deoxysugar biosynthesis in the erythromcyin producer *Saccharopolyspora erythraea*, Mol. Gen. Genet. 257 (1998) 542–553.
[39] T.M. Hallis, H.-W. Liu, Learning nature's strategies for making deoxy sugars: pathways, mechanisms, and combinatorial applications, Acc. Chem. Res. 32 (1999) 579–588.
[40] S. Omura, H. Tanaka, Production, Structure and antifungal activity of polyene macrolides, in: S. Omura (Ed.), Macrolide Antibiotics, Chemistry, Biology and Practice, Academic Press, New York, 1986, pp. 351–404.
[41] H. Bolhuis, H.W. Van Veen, B. Poolman, A.J.M. Driessen, W.N. Konings, Mechanisms of multidrug transporters, FEMS Microbiol. Rev. 21 (1997) 55–84.
[42] K.J. Linton, H.N. Cooper, I.S. Hunter, P.F. Leadlay, An ABC transporter from *Streptomyces longisporoflavus* confers resistance to the polyether ionophore tetronasin, Mol. Microbiol. 11 (1994) 777–785.
[43] D.J. Henderson, D.J. Lydiate, D.A. Hopwood, Structural and functional analysis of the mini-circle, a transposable element of *Streptomyces coelicolor* A3(2), Mol. Microbiol. 3 (1989) 1307–1318.
[44] N.H. Georgopapakadou, T.J. Walsh, Antifungal agents: chemotherapeutic targets and immunologic strategies, Antimicrob. Agents Chemother. 40 (1996) 279–291.
[45] M.R. Rodicio, C.J. Bruton, K.F. Chater, New derivatives of the *Streptomyces* temperate phage ϕC31 useful for the cloning and functional analysis of *Streptomyces* DNA, Gene 34 (1985) 283–292.
[46] D.A. Hopwood, M.J. Bibb, K.F. Chater, T. Kieser, C.J. Bruton, H.M. Kieser, D.J. Lydiate, C.P. Smith, J.M. Ward, H. Schrempf, Genetic manipulation of *Streptomyces*: a Laboratory Manual, John Innes Foundation, Norwich, 1985.
[47] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
[48] R. Staden, The Staden sequence analysis package, Mol. Biotechnol. 5 (1996) 233–241.
[49] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipmann, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.