<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Polyene macrolide biosynthesis in streptomycetes and related bacteria: recent advances from genome sequencing and experimental studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authors(s)</strong></td>
<td>Caffrey, Patrick; De Poire, Eimear; Sheehan, James; Sweeney, Paul</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2016-05</td>
</tr>
<tr>
<td><strong>Publication information</strong></td>
<td>Applied Microbiology and Biotechnology, 100 (9): 3893-3908</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Springer</td>
</tr>
<tr>
<td><strong>Item record/more information</strong></td>
<td><a href="http://hdl.handle.net/10197/7577">http://hdl.handle.net/10197/7577</a></td>
</tr>
<tr>
<td><strong>Publisher's statement</strong></td>
<td>The final publication is available at <a href="http://www.springerlink.com">www.springerlink.com</a></td>
</tr>
<tr>
<td><strong>Publisher's version (DOI)</strong></td>
<td>10.1007/s00253-016-7474-z</td>
</tr>
</tbody>
</table>
Accepted manuscript

Published as:
Polyene macrolide biosynthesis in streptomycetes and related bacteria: recent advances from genome sequencing and experimental studies

Patrick Caffrey†, Eimear De Poire, James Sheehan, Paul Sweeney

School of Biomolecular and Biomedical Science, University College Dublin, Belfield, Dublin 4, Ireland.

†Correspondence to: Patrick Caffrey
patrick.caffrey@ucd.ie
Tel: + 353 1 716 2618

Key words: Polyene macrolide biosynthesis, genome sequence
Abstract

The polyene macrolide group includes important antifungal drugs, to which resistance does not arise readily. Chemical and biological methods have been used in attempts to make polyene antibiotics with fewer toxic side effects. Genome sequencing of producer organisms is contributing to this endeavour, by providing access to new compounds and by enabling yield improvement for polyene analogues obtained by engineered biosynthesis. This recent work is also enhancing bioinformatic methods for deducing the structures of cryptic natural products from their biosynthetic enzymes. The stereostructure of candicidin D has recently been determined by NMR spectroscopy. Genes for the corresponding polyketide synthase have been uncovered in several different genomes. Analysis of this new information strengthens the view that protein sequence motifs can be used to predict double bond geometry in many polyketides.

Chemical studies have shown that improved polyanes can be obtained by modifying the mycosamine sugar that is common to most of these compounds. Glycoengineered analogues might be produced by biosynthetic methods, but polyene glycosyltransferases show little tolerance for donors other than GDP-α-D-mycosamine. Genome sequencing has revealed extending glycosyltransferases that add a second sugar to the mycosamine of some polyenes. NppY of Pseudonocardia autotrophica uses UDP-N-acetyl-α-D-glucosamine as donor whereas PegA from Actinoplanes caeruleus uses GDP-α-D-mannose. These two enzymes show 51% sequence identity and are also closely related to mycosaminyltransferases. These findings will assist attempts to construct glycosyltransferases that transfer alternative UDP- or (d)TDP-linked sugars to polyene macrolactones.

Introduction

Polyene macrolides are bioactive natural products that are mainly synthesised by streptomycetes and related actinobacteria. Some of these compounds are used as antifungal and antiparasitic drugs, with pimaricin, nystatin A1 and amphotericin B as prominent examples (Fig. 1). Polyene antibiotics have serious side effects such as nephrotoxicity, but are likely to remain important as medicines because resistance does not arise readily. Polyenes are also active against enveloped viruses and pathogenic prion proteins, and could be developed as treatments for diseases caused by these agents (Lemke et al., 2005).

Polyene macrolides have been a focus of scientific interest since their discovery in the 1950s. Their structure-activity relationships, modes of action and biosynthesis are still being actively investigated by chemists, biophysicists and biologists, with the need for non-toxic antifungals driving this work. Improved analogues have been made by chemical methods and by genetic manipulation of producer micro-organisms (Cereghetti and Carreira, 2006). There have been several comprehensive reviews on polyene biosynthesis over the past few years (Caffrey et al., 2008; Kong et al., 2013; Letek et al., 2014; Aparicio et al., 2016). The field advances rapidly and genome sequences of many polyene producers are now available. This minireview explores how this information can fill gaps in biosynthetic knowledge, enable overproduction of valuable analogues, and assist discovery of new related compounds. Progress in other areas of polyene research will also be summarised, to place the studies in bacterial genetics in context.
Recent insights into biological activities of polyene macrolides

Since efficient production of improved polyene analogs requires a deep understanding of their modes of action, some of the recent work on these aspects is briefly reviewed here. The antibiotic activity of polyene macrolides results from interactions with membrane sterols. Selective toxicity results from a high affinity for ergosterol, the predominant sterol in fungal and *Leishmania* cells, whereas weaker binding to cholesterol in mammalian membranes contributes to adverse side effects (Lemke et al., 2005). The mechanisms of antibiotic action are complex and still incompletely understood, even after five decades of research. However, recent studies with non-permeabilising amphotericin analogues suggest that transmembrane channels are not essential for antifungal activity, sequestering of ergosterol disrupts multiple cellular processes and may be sufficient (Szpielman et al., 2008; Gray et al., 2012). The new “sponge model” proposes that aggregates of amphotericin B align on the membrane surface and extract ergosterol from the lipid bilayer (Anderson et al., 2014). This activity is similar to that of methyl-β-cyclodextrins, which are also known to deplete membranes of cholesterol (Zidovetzki and Levitan, 2007). It is also known that amphotericin B somehow induces oxidative stress in sensitive cells, and formation of reactive oxygen species contributes to its killing effects (Mesa-Arango et al., 2012). These effects enable high concentrations of polyenes to inactivate fungal biofilms (Delattin et al., 2014). Fungal pathogens can become resistant by replacing ergosterol with alternative sterols that do not bind polyenes. However, this results in diminished virulence and has not posed a serious clinical problem (Vincent et al., 2013).

As well as antifungal and anti-parasitic activities, polyenes have other potentially interesting biological properties. Some of these result from sterol-binding, others from interactions with specific proteins. The anti-retroviral, anti-prion and immunomodulating activities have been discussed in previous reviews (Hartsel and Bolard, 1996; Abu-Salah et al., 1996; Mesa-Arango et al., 2012). Amphotericin B interacts with a structural motif in amyloid peptides that aggregate in brain tissue to cause Alzheimer’s disease (Hartsel and Weiland, 2003). This property is not thought to be of therapeutic value, but may be useful for detection of these destructive peptides (Smith et al., 2009).

Some recently discovered effects of polyenes result from interactions with intracellular membranes. In mammalian cells, amphotericin B counteracts the antiviral effects of the interferon-induced transmembrane (IFITM) proteins (Yu et al., 2015), making influenza A virus infections more severe (Roethl et al., 2011; Lin et al., 2013). High concentrations (125 μM) of allicin potentiate the effects of amphotericin B and nystatin on *Saccharomyces cerevisiae* by facilitating polyene penetration to internal vacuoles (Borjihan et al., 2009; Ogita et al., 2010). Combinations of allicin and low doses of amphotericin B were effective in treating experimental infections of hamsters with *Leishmania* (Corral et al., 2014a and 2014b). In contrast, manumycin antagonises the effects of amphotericin B on *Schizosaccharomyces pombe* by inhibiting exocytosis. This depletes the amount of ergosterol in the cytoplasmic membrane and reduces sensitivity to polyenes (Nishimura et al., 2014).

Nystatin and amphotericin B have been found to activate a quorum sensing mechanism that induces multicellularity and bacterial biofilm formation by *Bacillus subtilis* (Lopez et al., 2008). It has recently been suggested that sub-lethal concentrations of amphotericins could correct deficiencies caused by lack of ion channel proteins (Cioffi et al., 2015). This may lead to treatments for diseases like cystic fibrosis.
Genes for polyene macrolide biosynthesis

From about 2000, the biosynthetic gene clusters for pimaricin, nystatin, amphotericin B (Fig. 1) and candidin were cloned from *Streptomyces natalensis*, *Streptomyces noursei*, *Streptomyces nodosus* and *Streptomyces* FR008 (Caffrey et al., 2008). Sequencing revealed the primary structures of the modular polyketide synthases (PKSs) that assemble the macrolactone cores of these compounds. Most polyenes undergo two characteristic late modifications: a methyl branch is oxidised to a carboxyl group, and a D-mycosamine sugar is attached. In some tetraenes, the exocyclic carboxyl group can be converted to a carboxamide after mycosaminylation. The polyol chain may be modified further. Hydroxylation at C8 and C10 occurs in the biosyntheses of amphotericin B and nystatin A1, respectively. The 4,5 epoxide in pimaricin is formed by the action of a cytochrome P450 on an alkene group. In nystatin, the C35 hydroxyl group may be modified with L-digitoxose or L-mycarose sugars (Bruheim et al., 2004). For a few polyenes, the mycosaminyl residue is modified with a second sugar to form a disaccharide chain (De Poire et al., 2013). The p-aminobenzoyl moiety of some aromatic heptaenes is N-methylated (see below).

New and known polyenes have been isolated from previously unexplored sources, such as marine microorganisms and rare actinomycetes (Letek et al., 2014). Prolific producers have been found among symbionts of ants that culture and exploit fungal species (Haeder et al., 2009; Barke et al., 2010; Seipke et al., 2011a). Microbial genome sequencing is also contributing to discovery of new polyenes and to a more complete understanding of their biosynthesis. Candidin is frequently re-isolated and its biosynthetic gene cluster appears in several genomes (Jorgensen et al., 2009; Seipke et al., 2011b; Olano et al., 2014; Tang et al., 2015; Komaki et al., 2015; Liu et al., 2015a). Since the sequencing of the nystatin biosynthetic genes in *S. noursei* (Brautaset et al., 2000), at least three other closely related clusters have been found (Kim et al., 2009; Barke et al., 2012; Cui et al., 2015). Pimaricin is produced by *S. natalensis*, *Streptomyces gilvosporus*, *Streptomyces lydicus* and *Streptomyces chattanoogensis* (Aparicio et al., 2016). Genome sequences for *S. natalensis* (accession JRKI0000000.1) and *S. chattanoogensis* (Jiang et al., 2013) are now available. This information will give insights into tetraene biosynthesis. The genome of *S. nodosus* has been sequenced (Sweeney et al., 2016). Until recently, this organism was the only one known to synthesise amphotericin B. Another producer has now been identified as a *Penicillium* species (Svahn et al., 2015). It is not yet known how this fungal producer achieves self resistance.

Polyene biosynthetic gene clusters encode PKS, cytochrome P450, mycosaminylation, export and regulatory proteins (Caffrey et al., 2008). Genome sequences provide access to unlinked genes (not located within the main clusters) that contribute to polyene production. These include genes required for additional functions, such as phosphopantetheinylation of ACP domains, formation of polyketide precursors malonyl and methylmalonyl CoA, formation of GDP-α-D-mannose for mycosamine synthesis, carboxamide formation, and glycosylation with additional sugars (Stephens et al., 2013; Sweeney et al., 2015). This information should allow production to be optimised for valuable analogues, as high-yielding strains are essential for advancement of these compounds.

Studies on genomes of polyene producers may yield other potentially valuable natural products. As well as the amphotericin cluster, the *S. nodosus* genome contains 24 biosynthetic gene clusters, although 18 of these specify previously known compounds (Sweeney et al., 2016). It has been suggested that genomes of rare non-streptomycete actinomycetes may be a richer source of biosynthetic genes for new antibiotics and bioactive compounds (Jose and Jeukumar, 2013).
Polyene biosynthesis and polyketide stereochemistry

Knowledge of polyene biosynthesis has been valuable in investigations of how modular PKSs determine stereochemistry in their products. The availability of polyene PKS sequences assisted identification of the LDD motif characteristic of B-type ketoreductase domains, which form 3\text{\textalpha}-3-hydroxyacyl-ACP intermediates (Caffrey, 2003; Reid et al., 2003). This motif is absent from A-type ketoreductases that form 3\text{\textbeta}-3-hydroxyacyl-ACP intermediates. Ketoreductase domains also determine methyl stereochemistry after incorporation of branched extenders (Weissman et al., 1997; Valenzano et al., 2009; Zheng et al., 2013; Annaival et al., 2015). Additional amino acid residues have been identified that allow prediction of C-2 methyl as well as C-3 alcohol stereochemistry in an extended chain (Keatinge-Clay, 2007). Active KRs in modules that use methylmalonyl extenders have been further classified as A1, A2, B1 and B2 (Keatinge-Clay, 2007). These impose the same stereochemistry as that observed in (2\text{\textalpha}, 3\text{\textalpha}), (2\text{\textalpha}, 3\text{\textomega}), (2\text{\textomega}, 3\text{\textalpha}) and (2\text{\textomega}, 3\text{\omega}) isomers of 2-methyl-3-hydroxybutyrate. C1 and C2 KRs do not reduce the \beta-ketone but give (2\text{\textalpha})- and (2\text{\omega})-2-methyl-3-ketoacyl products, respectively. [The reasons for using both relative (l and d) and absolute (R and S) indictators when describing polyketide stereochemistry have been explained by Labonte and Townsend (2013).] Crystal structures have now been determined for seven PKS KR domains, including the A1 KR2 and A2 KR11 domains of the amphotericin PKS (Bonnett et al., 2013). These structural studies have given insights into how PKSs determine hydroxyl and methyl stereochemistry (Zheng et al., 2013; Keatinge-Clay 2016). Where a 2-methyl-3-ketoacyl intermediate is fully processed, the chirality at C-2 is determined by the enoylreductase domain. Kwan and Leadlay have identified a key amino acid residue that allows prediction of stereochemical outcome in this case (Kwan et al., 2008).

The stereospecificity motifs have been used to predict accurately the stereochemistry of many polyketides (Tietz and Mitchell, 2015; Essig et al., 2016). It is now known that the stereostructure of filipin agrees with that inferred from the sequence of the polyketide synthase (Gao et al., 2014; Ikeda et al., 2014). However, the stereostructure of candicidin D (Fig. 2) has recently been determined by NMR spectroscopy (Szwarc et al., 2015a) and this has revealed a striking failure of prediction methods. The hydroxyl-bearing C41 was the last chiral centre to have its stereochemistry assigned. The equivalent position in perimycin has yet to be determined. In the candidicidin PKS, module 2 is strongly predicted to form a (2\text{\textalpha}, 3\text{\textalpha}) 2-methyl-3-hydroxyacyl intermediate, yet the NMR structure indicates that the product is (2\text{\textomega}, 3\text{\omega}) (Fig. 3). To our knowledge this is the only clear example of a KR domain with an LDD motif that forms a 3\text{\textalpha}-3-hydroxyacyl chain. The structural basis for this anomaly may be of interest for future work. Protein engineering has shown that KR stereospecificity can be switched by changing amino acids other than the fingerprint residues (Siskos et al., 2005; Bailey et al., 2015).

In fatty acid and polyketide biosynthesis, dehydration of a 3\text{\textalpha}-3-hydroxyacyl chain gives a trans (E) double bond. The majority of PKS dehydratase domains form trans double bonds and are paired with B-type KRs. In polyketides, cis double bonds are less common. Some of these result from trans-cis isomerisation (Perlova et al., 2006; Vergnolle et al., 2011) or late modification (Palaniappan et al., 2008), others from dehydration of 3\text{\textalpha}-3-hydroxyacyl chains by the PKS. In these last cases, DH domains are paired with A-type KRs (Alhamadsheh et al., 2007; Labonte and Townsend, 2013).

There are two classes of aromatic heptaene, which differ slightly in double bond geometry. Candidicidin and perimycin are representative members of each class (Pawlak et al., 1995; Szwarc et al., 2015a) (Fig. 2). Levorin A1 is similar to candidicidin D whereas vacidin A and gedamycin are similar to perimycin (Sowinski et al., 1989a; Sowinski et al., 1989b; Sowinski et al., 1995; Szwarc et al., 2015b). Partricins A and B are identical to vacidin A
and gedamycin, respectively (Sowinski et al., 1995). Candicidin and levorin A1 are further distinguished by a methyl branch at C40. Some members of the other group are N-methylated on the p-aminobenzoyl moiety, namely perimycin, gedamycin and partricin B.

Candicidin D has two consecutive cis double bonds (C-26 to C-27 and C-28 to C-29) in the heptaene chain (Fig. 2). This correlates well with pairing of A-type KR and DH domains in modules 8 and 9 of the candicidin/FR008 PKS (Fig. 4). This is observed in all occurrences of the candicidin PKS. However, Actinokineospora enzanensis has an aromatic heptaene PKS (accession number = WP_018680914) in which A-type KRs appear in modules 7 and 8, indicating that cis double bonds in this as yet uncharacterized polyketide should be present in the same locations as those of perimycin (C-28 to C-29 and C-30 to C-31) (Fig. 2). This suggests that as microbial genome sequencing reveals increasing numbers of PKSs, it should be possible to infer initial double bond geometry in cryptic products, using existing prediction methods. However, cis alkenes formed as a result of post-PKS reactions remain unpredictable (Palaniappan et al., 2008; Vergnolle et al., 2011).

**Production of polyene analogues by chemical and biological methods**

Many polyene derivatives have been synthesised by chemical modification of the original natural products (Cereghetti and Carreira, 2006; Preobrazhenskaya et al., 2010). Some of these have improved antifungal selectivity and reduced toxicity. Engineering of PKS and late genes has delivered several analogues of pimaricin, amphotericin B, nystatin, candicidin and rimocidin. Much of this work has been reviewed previously (Caffrey et al., 2008; Kong et al., 2013; Letek et al., 2014; Aparicio, 2016). Of the biosynthetically engineered polyene analogues, the most promising are those in which the free carboxyl group is replaced by a methyl group. One example is 16-descarboxyl-16-methyl-amphotericin B, which is produced in high yields (100 mg/litre) by S. nodosus ðamphNM (Carmody et al., 2005; Murphy et al., 2010). The same compound has been synthesised by chemical modification of amphotericin B, but this approach is less efficient (Palacios et al., 2007). A nystatin heptaene counterpart of this analogue is biosynthesised at yields of 0.5 g/litre and is superior to amphotericin B in animal models of candidiasis (Brautaset et al., 2008).

Chemical methods can generate diverse array of polyene analogues. Engineered biosynthesis gives a smaller range, but has the potential to deliver a valuable compound in greater yield by fermentation. Burke and co-workers have developed a scalable chemical method for converting amphotericin B to valuable derivatives in which substituted ureas replace the free carboxyl group at C-16 (Davis et al., 2015). Chemical and biological methods can complement one another (Tevyashova et al., 2013). Murata and co-workers have made two new amphotericins by reducing the 7-keto group in biosynthetically engineered 7-oxoamphotericin B, and separating the two stereoisomers of 7-hydroxy-amphotericin B (Yamamoto et al., 2015).

While many potentially valuable polyene analogues have been developed, so far none has advanced into clinical medicine. Another problem with amphotericin B is its tendency to aggregate in aqueous solution, which further diminishes its selective toxicity. Monomers show antifungal specificity whereas aggregates are equally damaging towards fungal and mammalian cells (Szlinger-Richert et al., 2001). Liposomal amphotericins show reduced toxicity, possibly because they release monomers slowly (Lemke et al., 2005). These lipid formulations are expensive to manufacture. Non-aggregating derivatives have been synthesised by coupling two cholate residues to amphotericin B through a spermidine linker (Janout et al., 2015).
Engineering late modifications of polyene macrolactone rings

Genetic manipulation of late biosynthetic steps has delivered several new tetraenes. The first of these was 4, 5 de-epoxyproducerin (DEP), obtained by inactivating the PimD cytochrome P450 (Fig. 5A). Santos-Aberturas and co-workers (2015) have exploited further the late modifications of tetraenes in vivo and in vitro. In the pimD mutant host, the S. noursei NysL P450 (which hydroxylates C-10 in 10-deoxynystatin) formed a new analogue, 6-hydroxy-DEP, whereas AmphL (which hydroxylates C-8 in 8-deoxyscardicidin) did not act on DEP (Fig. 5A). The TetrK P450 hydroxylates C-4 of tetramycin A to form tetramycin B (Cao et al., 2012) (Fig. 5B). In a pimD mutant of S. natalensis, TetrK converted DEP to pimaricin, indicating that an alkane substrate is hydroxylated whereas an alkene is converted to an epoxide. The possibility of introducing epoxides into nystatins and amphotericins has been discussed (Santos-Aberturas et al., 2015). These efforts will be assisted by success with overproduction of PimD and NysL in E. coli in an active form. PimD has been crystallised and the three-dimensional structure has been determined (Volokhan et al., 2006; Kells et al., 2010).

Most polyene biosynthetic gene clusters contain a gene for a cytochrome P450 that catalyses conversion of a methyl branch to an exocyclic carboxyl group. This gene has been inactivated in the producers of amphotericin B, nystatin, candidicidin, and rimocidin (Caffrey et al., 2008; Brautaset et al., 2008; Chen et al., 2009). In all cases, replacing the exocyclic carboxyl group with a methyl group decreases haemolytic activity but not antifungal activity. Recently the gene for the ScnG P450 was inactivated in S. chattanoogensis, a pimaricin producer (Liu et al., 2015b; Qi et al., 2015). This gave 12-decarboxyl-12-methyl analogues of pimaricin and DEP, which were fungicidal and had low haemolytic activity. A third analogue was also obtained, 2-hydroxy-12-decarboxyl-12-methyl-DEP, presumably resulting from omission of the dehydration step in the last chain extension cycle. This analogue had neither antifungal nor haemolytic activity.

*Actinoplanes caeruleus* synthesizes the disaccharide-modified aromatic heptaene 67-121C (Wright et al., 1977). Analysis of late biosynthetic genes identified AceN as the counterpart of the AmphN P450 that forms the C-16 carboxyl group in amphotericin B (Stephens et al., 2013). The aceN gene was capable of weakly complementing *amphN* deletions in *S. nodosus ΔamphNM* strains, suggesting some substrate tolerance (Stephens et al., 2013). Apart from this, little is known about exocyclic carboxyl group formation. So far it has not been possible to overproduce any of the relevant cytochrome P450 enzymes in an active form, which would allow mechanistic studies in vitro. It remains unclear how two oxygen atoms are introduced into the methyl group. In biosynthesis of artemisinic acid, the CYP71AV1 P450 is capable of oxidising a methyl group to a carboxyl in three stages (methyl → hydroxymethyl → aldehyde → carboxylic acid) (Ro et al., 2006). However, it is now known that two more short-chain dehydrogenase-reductase (SDR) enzymes are required for efficient catalysis of the last two steps (Paddon and Keasling, 2014). The availability of increased numbers of genome sequences of polyene producers could reveal any conserved oxidoreductases that might be involved in forming free carboxyl groups in polyene macrolides.

Rimocidins and related tetraenes are active against trypanosomes and fungal pathogens (Rolon et al., 2006). The *Streptomyces diastaticus* var 108 polyene PKS can use a butyryl primer to form rimocidin or an acetyl primer to form CE108 (Fig. 5C). Escudero and co-workers (2015) knocked out the rimJ gene for a crotonyl CoA reductase-carboxylase that catalyses formation of both butyryl CoA and ethylmalonyl CoA. This led to formation of rimocidin analogues in which the starter was acetate and a methyl group replaced the ethyl branch introduced in the last cycle (Fig. 5C). However, loss of the ethyl chain reduced antifungal activity. A new rimocidin primed with a propionyl starter has been isolated from *Streptomyces mauvecolor* (Jeon et al., 2016).
In the course of genetic manipulation of *S. diastaticus* var 108, Malpartida and co-workers found that the exocyclic carboxyl groups of rimocidin and CE108 could be converted to carboxamides (Fig. 5D), giving analogues with improved antifungal activity (Seco et al., 2010). Two different enzymes were found to be capable of catalysing this new late modification, PsCA and PsCB, which are related to glutamine synthetases. PsCA converts rimocidin, CE108 and pimaricin to carboxamide derivatives whereas PsCB only catalyses this reaction with pimaricin. It has now been shown that PsCA converts DEP, 6-hydroxy-DEP, CE108D and lucensomycin to carboxamides (Santos-Aberturas et al., 2015; Escudero et al., 2015). Amphotericin B does not undergo this modification.

The Aparicio group have inactivated cytochrome P450 enzymes that function at a late stage in filipin biosynthesis, and generated analogues with increased antifungal activity (Payero et al., 2015).

**Glycosylation**

The mycosamine sugar of polyenes is synthesised from GDP-\(\alpha\)-\(\alpha\)-d-mannose (Nic Lochlainn and Caffrey, 2009; Nedal et al., 2007). GDP-\(\alpha\)-\(\alpha\)-d-mannose 4, 6 dehydratase catalyses formation of GDP-4-keto-6-deoxy-\(\alpha\)-\(\alpha\)-d-mannose, which is somehow isomerised to the 3-keto sugar (Fig. 6). The final step is a transamination catalysed by a mycosamine synthase. While the 3, 4 ketoisomerisation can occur spontaneously, it is not known whether the reaction is enzyme-catalyzed *in vivo* (Hutchinson et al., 2010). The Zotchev group found that NysDIII GDP-\(\alpha\)-\(\alpha\)-d-mannose 4, 6 dehydratase and NysDII GDP-mycosamine synthase were not sufficient to reconstitute GDP-mycosamine biosynthesis *in vitro* (Nedal et al., 2007). So far, bioinformatic analysis of genome sequences has not helped to resolve this issue. It has not yet been possible to identify a gene for a 3, 4 ketoisomerase enzyme within the *S. nodosus* genome. There are no obvious homologs of the 3, 4 isomerases that act on TDP-4-keto-6-deoxy-\(\alpha\)-\(\alpha\)-D-glucose in the biosynthesis of \(\alpha\)-mycaminose and other deoxysugars (Hutchinson et al., 2010; Sweeney et al., 2016).

In mycosamine synthase-deficient *amphDII* mutants of *S. nodosus*, the aglycone is modified with a neutral deoxyhexose. This is thought to be \(\alpha\)-d-rhamnose because some GDP-\(\alpha\)-\(\alpha\)-d-mannose 4, 6 dehydratases further convert their GDP-\(\alpha\)-\(\alpha\)-d-4-keto-6-dideoxymannose product to GDP-\(\alpha\)-\(\alpha\)-d-rhamnose *in vitro* (King et al., 2009). The AmphDI mycosaminyl transferase is capable of modifying the aglycone with this neutral sugar. The rhamnosyl analogue shows a ten-fold decrease in antifungal activity. In the aromatic heptaene perimycin (Fig. 2) and the pimaricin analogue JB1R-13 (Komaki et al., 2009), the amino sugar is \(\alpha\)-d-perosamine (4, 6-dideoxy-4-amino \(\alpha\)-d-mannose) rather than \(\alpha\)-d-mycosamine. The *perDII* perosamine synthase and *perDI* perosaminyltransferase genes have been cloned from the perimycin producer *Streptomyces cacaoi* (Hutchinson et al., 2010). Complementation of *amphDII* mutations with a *perDII* perosamine synthase gene led to formation of perosaminyl-amphoteronolides (Hutchinson et al., 2010; Stephens et al., 2012). Increased yields of the perosaminyl analogue were obtained by including an AmphDI-PerDI hybrid glycosyltransferase (GT) capable of recognising the amphotericin aglycone and GDP-\(\alpha\)-\(\alpha\)-d-perosamine (Hutchinson et al., 2010). From this it can be concluded that the perosamine and mycosamine synthases are specific for the 4-keto and 3-keto sugars respectively. Perosamine appears in many Gram negative lipopolysaccharides and several GDP-\(\alpha\)-\(\alpha\)-d-perosamine synthases have been sequenced, some of which have been biochemically characterised (Hutchinson et al., 2010). As genome sequencing uncovers increased numbers of polyene clusters, it may become possible to identify sequence motifs that distinguish perosamine and mycosamine synthases.
AmphDI, NysDI, FscMI, PimK and other mycosaminytransferases tolerate structural changes in their aglycone acceptors but show moderately strict specificity for NDP-sugar donors. A variety of amphoteronolide, nystatinolide and candicidinolide macrolactones have been generated by PKS engineering (Caffrey et al., 2008). Almost all were mycosamylated in vivo except for a pentaene related to amphotericin B (Murphy et al., 2010). Both NysDI and AmphDI mycosamylate candicidin and pimaricin aglycones in vivo, albeit less efficiently than the cognate GTs. PimK can act on candicidinolide, which is larger than its normal substrate (Lei et al., 2013; Liu et al., 2015b). An in vitro reverse glycosylation assay confirmed that AmphDI can recognise pimaricin and candicidin aglycones (Zhang et al., 2008). In contrast, two GTs that normally act on aromatic heptaenes do not recognise amphotericin aglycones in vivo. These are the PerDI perosaminyltransferase from S. cacaoi and the AceDI mycosaminyltransferase from A. caeruleus (Hutchinson et al., 2010; Stephens et al., 2013). These enzymes appear to have a lower tolerance for unnatural acceptors.

The sugar flexibility of AmphDI and NysDI was investigated in vitro using a panel of 21 synthetic and enzymatically generated NDP-sugars (Zhang et al., 2008). Of these, only D-mannose and L-gulose were loaded onto amphoteronolide aglycones. In vivo studies indicate that AmphDI can use GDP-α-D-rhamnose and GDP-α-D-perosamine inefficiently (see above). The FscMI mycosaminytransferase can use one of the GDP-ketosugar intermediates (Fig. 6) to modify its aglycone. Lei and co-workers have identified residues in the FscMI that are critical for activity (Lei et al., 2013).

Some nystatins and candidins have an additional L-digitoxose or L-mycarose sugar residue attached to C35 (Bruheim et al., 2004; Zielinski et al., 1979) (Fig. 7). These sugars occur in other natural products and the biosynthetic pathway has been elucidated, but the GT that modifies C35 in nystatin has not yet been identified. Sequencing of the S. noursei genome is in progress (NCBI BioProject Accession PRJNA272896) and should reveal the gene for this enzyme.

Chemical studies have given further insights into the importance of the sugar moiety of amphotericin B. Croatt and Carreira (2011) replaced mycosamine with 3, 6 dideoxy 3-amino-β-D-glucose to form 2′-epi-amphotericin B. This was as fungicidal as amphotericin B, but methylation of the epimeric 2′ hydroxyl group abolished activity. An analogue in which the mirror image of 3, 6 dideoxy 3-amino-β-D-glucose replaced mycosamine was also inactive (Croatt and Carreira, 2011). Burke and co-workers found that synthetic 2′-deoxy amphotericin B (lacking the 2′ hydroxyl on the mycosamine sugar) was active and had increased specificity for ergosterol-containing membranes (Wilcock et al., 2013). However, this analogue could not be chemically synthesised in large quantities (Davis et al., 2015). In principle, 2′-deoxy amphotericin B could be generated by biosynthetic engineering. An activated 2′-deoxymycosamine could be synthesised in vivo using genes for TDP-D-angolosamine formation (Schell et al., 2008). However, current knowledge of polyene GTs indicates that redesigning them to accept alternative dTDP-sugars could prove challenging.

Extending glycosyltransferases

A semi-synthetic analogue of amphotericin B, N-methyl-N-D-fructosyl-amphotericin B methyl ester (MFAME), showed increased water-solubility, altered aggregation properties, and reduced haemolytic activity compared to the parent compound (Cybulska et al., 2000; Szlindler-Richert et al., 2001). This led to efforts to engineer the biosynthesis of disaccharide-modified amphotericins. Before 2012, only one naturally occurring polyene was known to contain a disaccharide chain, 67-121C made by A. caeruleus (Fig. 8). A draft genome
sequence of this organism revealed the genes involved in late modification of the polyene macrolactone (Stephens et al., 2013). The extending GT gene pegA was not within the main cluster but was apparently located within a transposable element. Heterologous expression in other polyene producers revealed that the enzyme did not act on amphotericins but showed weak activity towards candicidins, which are more similar to the natural acceptor substrate (De Poire et al., 2013). PegA shows about 48% sequence identity with AmphDI, suggesting that extending GTs and mycosaminyltransferases have a close evolutionary relationship.

Kim and co-workers (2009) identified a disaccharide-modified nystatin NPP (nystatin-like Pseudonocardia polyene) in cultures of the rare actinomycete Pseudonocardia autotrophica. The yield of NPP was low, 2 mg per liter, but the group were able to purify enough material for structural analysis and characterisation of biological activities (Lee et al., 2012). The second sugar was an N-acetylglucosamine residue α1, 4 linked to the mycosamine sugar. NPP had 50% of the antifungal activity of nystatin A1, but was 10 times less haemolytic and 300 times more water-soluble. Inactivation of the ER domain of module 5 led to production of a heptaene analogue, which had a higher activity than NPP but was produced at a lower yield. Sequencing of the Ps. autotrophica genome identified the extending GT gene nppY located adjacent to the main cluster. This was 51% identical to PegA (Kim et al., 2015). Inactivation of nppY gave production of 10-deoxynystatin, suggesting that addition of the second sugar occurs before C-10 hydroxylation (Fig 8). This pathway was confirmed by extensive gene disruption and complementation experiments. Introduction of NppY into S. noursei did not result in production of NPP, presumably because in this host 10-deoxynystatin is rapidly hydroxylated and exported before the second sugar can be added. In S. noursei, C-10 hydroxylation of nystatin is the last modification in the pathway and does not proceed to completion when export proteins are inactivated (Sletta et al., 2005). Possibly intracellular accumulation of nystatin shifts the reaction equilibrium towards the 10-deoxy form.

Presumably the donor substrate for NppY is UDP-N-acetyl-α-D-glucosamine whereas that for PegA is GDP-α-D-mannose (Fig. 8). In 67-121C, the mannose residue is β 1, 4 linked to the mycosaminyl sugar, suggesting that PegA is an inverting GT. In NPP, the N-acetyl glucosamine is α 1, 4 linked, suggesting that NppY is a retaining GT. This is an exception to Klyne’s rule, which states that D sugars in natural products are β-linked whereas L sugars are α-linked, because the GTs involved are inverting (Rix et al., 2002). This rule applies when an NDP-sugar is the immediate donor. In the biosynthesis of some glycosylated non-ribosomal peptides like mannopeptimycins, α-linked D-mannosyl residues are transferred from a dolichol phosphate-β-D-mannose donor rather than directly from GDP-α-D-mannose (Magarvey et al., 2006). The GTs involved are inverting enzymes.

Another disaccharide-modified polyene has been obtained from a bacterial symbiont of the ant Acromyrex octospinosus (Seipke et al., 2011a). Ants cultivate fungal monocultures as a food source, or to build traps for insect prey. These ants also use streptomycete producers of filipins, nystatins, candidins and other polyenes to control fungal growth (Barke et al., 2012; Gao et al., 2014). One of these ant symbionts, Pseudonocardia sp. P1, produced a new analogue of nystatin that is modified with a second sugar. This was obtained in low amounts but LC MS analysis of the intact polyene and fragment ions revealed that the macrolactone contained a disaccharide rather than two monosaccharides. The identity of the second sugar is not known but it has a mass appropriate for glucose or mannose. The nypY gene for the extending GT has been identified by genome sequencing. NypY is 83% identical to NppY and 51% identical to PegA.

The nypY gene has also been expressed in heterologous hosts. NypY did not act on nystatins in S. noursei (Seipke et al., 2012) but in S. nodosus showed modest activity towards amphotericins A and B (De Poire et al.,
Slightly more efficient modification of 7-keto amphotericin B was observed in the \textit{S. nodosus} \textDelta KR16 mutant. Some NypY-catalysed hexosylation of candicidin was detected in \textit{Streptomyces griseus}. The highest yields of the putative disaccharide-modified polyenes amounted to only 5\% of the total polyene. It is still unclear whether the low conversion rates result from poor recognition of unnatural substrates, rapid export of monosaccharide-modified polyenes, or poor expression of the extending GT. It was not possible to detect disaccharide-modified derivatives of amphotericins lacking exocyclic carboxyl groups or mycosamine amino groups. It is conceivable that such analogues might be obtained in an \textit{in vitro} system in which enzyme and substrate concentrations might be optimised, and competing processes eliminated. However, so far it has not been possible to overproduce the NypY or PegA GTs in \textit{E. coli} in a soluble form (E. De Poire and P. Caffrey, unpublished data).

The fact that PegA normally modifies 67-121A with a mannose suggests that the enzyme catalyses addition of this same sugar to candicidin. NypY adds a second hexose to candicidin and amphotericin B. Further work will be required to determine whether NypY uses GDP-\alpha-D-mannose or UDP-\alpha-D-glucose as sugar donor, and whether it is a retaining or inverting GT. At the sequence level, NypY shows greater similarity to NppY, which is thought to use UDP-\alpha-D-N-acetylg glucosamine.

### Regulation of polyene biosynthesis and use of genome sequences to improve yields

Polyene biosynthetic gene clusters are directly controlled by regulatory proteins that respond to environmental signals and activate transcription. Regulation has been investigated for nystatin, tetramycin, and FR008 biosynthesis (Caffrey et al., 2008; Cui et al., 2015; Zhang et al., 2015), but the most detailed studies have been carried out on pimaricin (Aparicio et al., 2016).

In the pimaricin cluster, the PimR protein activates transcription of the gene for the pathway-specific activator PimM (Aparicio et al., 2016). This in turn regulates several promoters for PKS and late genes. Homologues of PimM are found in other polyene biosynthetic gene clusters. These include NysRIV, AmphRIV, and PteF, which can activate pimaricin production in \textit{pimM} mutants of \textit{S. natalensis}. Overproduction of PimM in the producers of amphotericin, nystatin and filipin boosted yields of all three polyenes. The DNA binding sites for PimM and PimR have been identified. In \textit{Streptomyces avermitilis}, which is capable of synthesising filipin, the binding site for PimM appears upstream from 97 operons. These are mostly involved in primary metabolism, but could indirectly contribute to polyene production. High-level polyene production may require global changes in gene expression. The availability of genome sequences of polyene producers will enable further investigation.

Polyene-producing streptomycetes contain genes for synthesis and detection of gamma-butyrolactones, quorum sensing compounds that activate morphological differentiation and secondary metabolite production. However, gamma butyrolactones have not been detected. A unique signalling molecule has been isolated from \textit{S. natalensis}. Structural characterisation revealed that pimaricin inducer (PI) is 2, 3-diamino-2, 3-bis (hydroxymethyl)-1, 4-butanediol (Recio et al., 2004). The genome sequence should assist investigation of how this molecule is biosynthesised. Several other diol compounds and glycerol boost production of many polyenes. This is a low cost approach towards increasing yields (Aparicio et al., 2016).

\textit{S. natalensis} synthesises an extracellular cholesterol oxidase that is essential for pimaricin production (Mendes et al., 2007). This enzyme is encoded within biosynthetic gene clusters for several tetraenes. Products of the sterol oxidation reaction may activate polyene production. Pimaricin production is repressed by phosphate, and also affected by pleiotropic regulators and oxidative stress.
The genome sequence of *S. nodosus* has been investigated with a view to increasing yields of the valuable amphotericin analogue 16-descarboxyl-16-methyl-amphotericin B. Overproduction of the pathway-specific activator AmphRIV increased yields on *Streptomyces* medium (Sweeney et al., 2016). No improvements were observed following overproduction of acetyl and propionyl CoA carboxylases, a methylmalonyl CoA epimerase or a phosphopantetheine transferase specific for ACP domains of modular synthases. These results were interpreted as meaning that the levels of the large amphotericin PKS proteins are limiting. Jiang and co-workers (2013) found that overproduction of phosphopantetheine transferase increased yields of pimaricin in *S. chattanoogensis*. The smaller pimaricin PKS are possibly more efficiently translated and abundant, so phosphopantetheinylation may be incomplete in wild type cells.

Genome sequencing revealed that *S. nodosus* contains no other modular PKS that might compete with amphotericin production (Sweeney et al., 2016). The tetramycin producer *Streptomyces ahygroscopicus* also contains biosynthetic genes for nystatin (Cui et al., 2015). Inactivation of the tetramycin activator TtmRIV resulted in increased yields of nystatin, possibly because of re-direction of precursor flux.

### New polyene producers from marine and terrestrial sources

A few non-glycosylated polyene-polyol macrolactones have been obtained from marine and terrestrial bacteria. *Streptomyces* CNQ343 produces bahamaolide hexaenes (Kim et al., 2012). The structurally related reedsmycins were identified after a draft genome sequence of *Streptomyces* sp. CHQ-64 revealed part of a polyene PKS (Che et al., 2015). *Marinosporus* strain CNQ-140 produces the pentaenes marinisporolides A and B (Kwon et al., 2009). Polyene-polylols undergo spontaneous *trans-cis* isomerisation after isolation. They show weak antifungal activity. In the case of bahamaolides, this results from inhibition of isocitrate lyase.

Polyene macrolactams have antibacterial, antifungal, antiparasitic and antitumour activities. Several producers of these compounds have been isolated and a few PKSs have been sequenced (Jorgensen et al., 2009b; Schulze et al., 2015). In some cases, *cis* double bonds are introduced by modules containing A-type KR and DH domains, while others result from isomerization that occurs at a later stage.

### Conclusions

The clinical need for non-toxic antifungal and antiparasitic drugs has not yet been met. However, there has been recent progress towards reaching this goal. Engineered polyene-producing micro-organisms may provide improved compounds directly, or starting materials for chemical modification or lipid formulation. Microbial genome sequencing is contributing to this area by uncovering biosynthetic genes for new and previously known polyene macrolides, and by enabling improvements in yields of potentially valuable polyene analogues. The regulation of polyene biosynthesis appears to be similar in different producer organisms. Genome sequencing is beginning to reveal the global changes in gene expression that are associated with high-level production.

Some of this work has given insights into the biosynthesis of other natural products. Bioinformatic methods allow prediction of polyketide stereostructures. Module 2 of the candididin PKS provides an intriguing exception to the otherwise reliable correlation between the LDD motif and B-type ketoreduction, which gives 3D-3-hydroxacyl-ACP intermediates. Despite this, the latest findings suggest that analysis of PKS sequences should allow prediction of double bond geometry in aromatic heptaenes.
Many of the steps in polyene macrolide biosynthesis have been studied experimentally but some of the late stages are still incompletely understood. These include the detailed mechanism by which the exocyclic carboxyl group is formed, and the 3, 4 isomerisation step in GDP-\(\alpha\)-D-mycosamine formation. At present these issues appear to be of academic interest but further investigation could reveal biotechnological relevance.

Addition of extra sugar residues to polyene macrolides is considered likely to improve water-solubility and other pharmacological properties. An interesting advance from genome sequencing has been the identification of genes for polyene extending GTs that catalyse attachment of a second sugar to the mycosamine residues of nystatins and 67-121 heptaenes. There is a high degree of sequence identity (51%) between NppY, which uses UDP-N-acetyl-\(\alpha\)-D-glucosamine and PegA, which uses GDP-\(\alpha\)-D-mannose. Both of these extending GTs are similar to enzymes that attach mycosamine directly to polyene aglycones. This information should facilitate redesign of mycosaminytransferases to use UDP- and dTDP-sugar donors. This could enable glycosylation engineering of improved polyenes such as 2′ deoxy amphotericin B, which have so far only been obtained in low yield by semi-synthesis.

Acknowledgements

Work in the authors’ laboratory has been supported by Science Foundation Ireland, grant number 09/RFP/GEN2132.

Compliance with ethical standards

The authors declare that they have no conflict of interest. This article does not contain any studies with human participants or animals performed by any of the authors.

References


17


Letek M, Mateos LM, Gil JA. 2014. Genetic analysis and manipulation as a way to produce more effective antifungal compounds. In Antimicrobial compounds TG Villa and P. Veiga-Crespo (eds.), Springer Verlag Berlin Heidelberg. doi: 10.1007/978-3-642-40444-3_7


Figure captions

Fig. 1 Structures of pimaricin, nystatin A1 and amphotericin B.

Fig. 2 Stereostructures of candicidin D and perimycin. The chirality at C-41 in perimycin has yet to be assigned.

Fig. 3 Comparison of the \( p \)-aminobenzoyl chain region of candicidin D with the predicted triketide product of module 2 of the candicidin PKS. Apart from C41, there is perfect agreement between the structure determined by NMR and the one predicted from PKS stereospecificity motifs.

Fig. 4 Formation of \( cis \) double bonds during biosynthesis of the candicidin polyketide macrolactone. The eighth and ninth PKS extension modules contain A-type KR\( s \) paired with DH domains. This is consistent with formation of \( cis \) (\( Z \)) double bonds as a result of dehydration of 3L-3-hydroxyacyl chains. Shaded domains in module 21 are inactive.

Fig. 5 Late modifications in tetraene biosynthesis. A. The PimD and TetrK cytochrome P450 enzymes convert DEP to pimaricin whereas NysL forms 6-hydroxy-DEP. B. TetrK normally functions to convert tetramycin A to tetramycin B. C. Use of butyryl and acetyl primers by the rimocidin PKS results in formation of rimocidin and CE108. Inactivation of RimJ crotonyl CoA reductase-carboxylase results in use of an acetyl starter and a methylmalonyl rather than an ethylmalonyl extender in the last PKS cycle. D. Pimaricin, DEP, 6-hydroxy-DEP, CE108D and lucensomycin have all been converted to carboxamides by PscA.

Fig. 6. Biosynthesis of mycosamine and perosamine.

Fig. 7. Glycosylation of the C35 hydroxyl group of nystatin. The equivalent position in the heptaene candidin is modified with \( L \)-digitoxose.

Fig. 8 Disaccharide formation in biosynthesis of 67-121C and NPP.
Fig. 1 Structures of pimaricin, nystatin A1 and amphotericin B.
Fig. 2 Stereostructures of candicidin D and perimycin. The chirality at C-41 in perimycin has yet to be assigned.
Stereochemistry predicted for B1 type KR in module 2

Structure determined by NMR

Fig. 3. Comparison of the $p$-aminobenzoyl chain region of candicidin D with the predicted triketide product of module 2 of the candicidin PKS. Apart from C41, there is perfect agreement between the structure determined by NMR and the one predicted from PKS stereospecificity motifs.
Fig. 4 Formation of cis double bonds during biosynthesis of the candicidin polyketide macrolactone. The eighth and ninth extension modules contain A-type KRαs paired with DH domains. This is consistent with formation of cis (Z) double bonds as a result of dehydration of 3L-3-hydroxyacyl chains. Shaded domains in module 21 are inactive.
A

4, 5-De-epoxy-pimaricin

PimD
or TetrK

NysL

4, 5-De-epoxy-
6-hydroxy-pimaricin

B

Tetramycin A

TetrK

C

Rimocidin

CE108

CE108D

D
Fig. 5 Late modifications in tetraene biosynthesis. A. The PimD and TetrK cytochrome P450 enzymes convert DEP to pimaricin whereas NysL forms 6-hydroxy-DEP. B. TetrK normally functions to convert tetramycin A to tetramycin B. C. Use of butyryl and acetyl primers by the rimocidin PKS results in formation of rimocidin and CE108. Inactivation of RimJ crotonyl CoA reductase results in use of an acetyl starter and a methylmalonyl rather than ethylmalonyl extender in the last PKS cycle. D. Pimaricin, DEP, 6-hydroxy-DEP, CE108D and lucensomycin have all been converted to carboxamides by PscA.
Fig. 6. Biosynthesis of mycosamine and perosamine
Fig. 7. Glycosylation of the C35 hydroxyl group of nystatin. The equivalent position in the heptaene candidin is modified with L-digitoxose.
Fig. 8 Disaccharide formation in biosynthesis of 67-121C and NPP