Understanding and manipulating antibiotic production in actinomycetes

Mervyn J. Bibb*1
*Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, U.K.

Abstract
Actinomycetes are prolific producers of natural products with a wide range of biological activities. Many of the compounds that they make (and derivatives thereof) are used extensively in medicine, most notably as clinically important antibiotics, and in agriculture. Moreover, these organisms remain a source of novel and potentially useful molecules, but maximizing their biosynthetic potential requires a better understanding of natural product biosynthesis. Recent developments in genome sequencing have greatly facilitated the identification of natural product biosynthetic gene clusters. In the present article, I summarize the recent contributions of our laboratory in applying genomic technologies to better understand and manipulate natural product biosynthesis in a range of different actinomycetes.

Introduction
Actinomycetes are Gram-positive mycelial bacteria of terrestrial and marine origin. For bacteria, they possess relatively large genomes (up to ~12 Mb in size) consisting of DNA with a relatively high GC content (up to 74 mol% GC). They are also characterized by a complex process of morphological differentiation that results in sporulation, reflecting their major route of dispersal in the environment.

Actinomycetes are well known as prolific producers of a wide variety of natural products (also referred to in the present article as specialized metabolites) with an equally wide range of biological activities. These include antibacterial, antifungal, anti-cancer, anti-helminthic, herbicidal and immunosuppressive activities, and many of these compounds are used in medicine and in agriculture [1]. Strikingly, these organisms produce approximately two-thirds of all known antibiotics of microbial origin [2], and currently approximately half of all clinically used classes of antibiotics (encompassing approximately 50 different compounds) are derived from actinomycetes.

The production of these generally structurally complex molecules occurs via biosynthetic pathways encoded by their corresponding gene clusters, which can vary in size from less than 10 kb to over 120 kb. Advances in the molecular genetics of actinomycetes has greatly facilitated the cloning, sequencing and characterization of natural product gene clusters, providing new insights into the enzymology involved in the synthesis of these specialized metabolites, and into the underlying mechanisms that regulate their synthesis. The production of these molecules is generally not a constitutive process, but occurs in a growth-phase-dependent manner and is frequently co-ordinately regulated with the onset of morphological differentiation, and indeed mutations exist that abolish both processes [3–5].

Over the last 30 or so years, a variety of approaches have been devised to isolate entire natural product biosynthetic gene clusters, but they are generally both labour-intensive and lengthy. Recently, we have taken advantage of developments in genome sequencing to markedly facilitate and expedite this process. The procedure that we have developed, termed genome scanning, is outlined in Figure 1 and commences with the isolation of high-molecular-mass genomic DNA from the producing strain. Some of this DNA is used to generate large insert cosmid or PAC (P1-derived artificial chromosome) libraries in *Escherichia coli*, and some is used for high-throughput sequencing. The resulting nucleotide contigs (often approximately 500 per genome with current
Illumina and Roche technologies) are then interrogated using tBLASTn [6] (http://blast.ncbi.nlm.nih.gov/Blast.cgi) for those encoding enzymes potentially involved in the biosynthesis of the molecule of interest. With the identified nucleotide sequence in hand, high-fidelity hybridization probes are derived by use of PCR and used to screen the large insert libraries. Once candidate clones are obtained, they are generally sequenced, and those likely to contain the entire biosynthetic gene cluster of interest transferred to an appropriate and genetically tractable host for heterologous expression. The gene cluster can then be analysed functionally and/or genetically tractable host for expression and further analysis and manipulation.

**Figure 1** | Procedure used for genome scanning
1. Genomic DNA is isolated from the strain of interest. 2. Some of that DNA is used to construct large-insert cosmid, BAC or PAC libraries, whereas some is subjected to high-throughput sequencing. 3. The resulting contigs are then interrogated for sequences likely to be involved in production of the compound of interest using tBLASTn. 4. This sequence information is used to design PCR primers for amplification of high-fidelity probes from the genomic DNA that are then hybridized against the large insert libraries. 5. Clones containing the gene cluster of interest are transferred to a genetically tractable host for expression and further analysis and manipulation.

**Lanthipeptides: feedforward mechanisms that regulate antibiotic production**

One well-studied class of RiPPs are the lanthipeptides, which are characterized by the possession of Lan (lanthionine) and/or MeLan (3-methyl-lanthionine) bridges [8]. They are encoded by a structural gene (generically named *lanA*) that encodes a precursor peptide with an N-terminal leader sequence followed by the core peptide that is retained in the mature compound. The precursor peptide undergoes modification and processing before the mature product is exported from the cell, coincident with, or followed by, leader peptide removal. The N-terminal leader sequence may have roles in directing export, retaining the precursor peptide in an inactive state until export and recruiting modifying enzymes. Bridge formation occurs via the dehydration of serine and threonine residues, followed by cyclization with cysteine residues, resulting in Lan and MeLan bridges respectively (Figure 2). These non-proteogenic amino acids constrain the molecule into a defined structural conformation that presumably promotes target specificity while resisting proteolytic degradation. Lanthipeptides are subdivided into four different classes (I–IV [8]) depending on the biosynthetic enzymes that install the Lan and MeLan motifs, and lanthipeptides that exhibit antimicrobial activity are referred to as lantibiotics. We have used genome scanning to analyse, in detail, the biosynthesis of two lantibiotics: microbisporicin and planosporicin.

**Microbisporicin**

Microbisporicin (Figure 2) is a potent 23-amino-acid lantibiotic made by *Microbispora corallina* NRRL 30420 [9]. It inhibits cell wall biosynthesis in a wide range of Gram-positive bacteria, selectively inhibiting peptidoglycan synthesis in *Staphylococcus aureus* and *Bacillus megaterium* probably by binding to the immediate precursor for cell wall biosynthesis, lipid II. It is active against methicillin-resistant and vancomycin-intermediate-resistant strains of *S. aureus* and, unusually for a lantibiotic, also against some Gram-negative species. Microbisporicin, under the commercial name NA1-107, is currently in late pre-clinical-phase trials and has demonstrated superior efficacy in animal models of multidrug-resistant infections compared with the drugs of last resort, i.e. linezolid and vancomycin. Interestingly, no microbisporicin-resistant mutants were observed during these studies [10]. The molecule has one MeLan and three Lan bridges and an S-[(Z)-2-aminovinyl]-D-cysteine residue at its C-terminus. It also contains two unusual modifications: 5-chlorotryptophan and 3,4-dihydroxyproline. An immature form of the lantibiotic lacking these two latter modifications shows markedly reduced levels of antibiotic activity.

To obtain a better understanding of microbisporicin biosynthesis, which could potentially enable the development of variants with improved clinical characteristics, we carried out genome scanning and identified a cosmid clone containing a cluster of 20 genes possibly involved in biosynthesis [11] (Figure 3). Although attempts to express the gene cluster in
several *Streptomyces* species failed, heterologous expression in the more closely related strain *Nonomuraea* ATCC 39727 was achieved, and, together with the construction of a *mibA* deletion mutant in *M. corallina*, served to confirm the identity and extent of the microbisporicin biosynthetic gene cluster. Bioinformatic analysis revealed genes likely to be involved in the introduction of the Lan and MeLan bridges (*mibB* and *mibC*, encoding a dehydratase and cyclase respectively), and the aminovinyl-cysteine moiety (*mibD*). In addition, it revealed genes apparently encoding a flavin-dependent tryptophan halogenase (*mibH*) and flavin reductase (*mibS*), and their role in chlorinating the tryptophan in microbisporicin was confirmed by mutational analysis. It also identified a gene encoding a cytochrome P450 presumably required for proline hydroxylation. Intriguingly, the analysis also revealed three putative regulatory genes: *mibR* encodes a protein with a C-terminal helix-turn-helix domain found in the LuxR family of transcriptional activators; *mibX* encodes a member of the ECF (extracytoplasmic function) class of σ factors (ECF σ factors generally respond to cell envelope stress by directing the corresponding RNA polymerase holoenzyme to particular promoter sequences); and *mibW* is predicted to encode an anti-σ factor with six transmembrane helices and a 73-amino-acid cytoplasmic N-terminal tail that probably sequesters MibX at the membrane until its release is prompted by an appropriate signal. Indeed, MibX and MibW were shown to interact *in vivo* using an *E. coli* two-hybrid assay. This was the first time that genes encoding an ECF σ factor–anti-σ factor pair had been found in an antibiotic biosynthetic gene cluster, and the individual deletion of *mibR* and *mibX* showed that both genes were required for transcription of the biosynthetic gene cluster and for microbisporicin production. High-resolution S1 nuclease protection experiments were used to identify six transcriptional start sites within the *mib* gene cluster [12]. Strikingly, five of these promoter elements possess –10 and –35 sequences that resemble closely the consensus sequence for the ECF family of σ factors. In contrast, the promoter sequence responsible for transcription of the operon required for the production of the immature and less active form of the lantibiotic lacking the chlorinated tryptophan and hydroxylated proline residues possessed sequences likely

---

**Figure 2** | Structures of compounds and structural motifs

For the Lan and MeLan bridges, filled orange circles represent the peptide backbone. In the lantibiotics and linaridins, residues are unmodified (white), dehydrated (purple), dehydrated and cyclized residues (blue) or possess other modifications (green). Abbreviations: Abu, aminobutyric acid; Abu-S-Ala, methyl-lanthionine; Aile, allo-isoleucine; Ala-S-Ala, lanthionine; Dha, dehydroalanine; Dhb, dehydrobutyryl; Me, methyl.
Figure 3 | Gene clusters cloned and analysed by genome scanning

Gene clusters required for (from top to bottom) microbisporicin, planosporicin, cypemycin, grisemycin, bottromycin and tunicamycin biosynthesis. For the RIPPs, genes encoding precursor peptides are in green, and genes potentially conferring immunity are in orange. For all clusters, genes encoding proteins involved in biosynthesis are shown in dark blue, and those likely to be involved in transport are shown in light blue. Confirmed and possible regulatory genes are shown in red. Sizes are indicated in kb.

---

to be recognized by the holoenzyme containing the major and essential σ factor HrdB. Mutational and qRT-PCR (quantitative reverse transcription–PCR) analyses combined with MALDI–TOF-MS led to a model for the regulation of microbisporicin biosynthesis in which a low level of transcription of \( mibR \) from an additionally identified, but uncharacterized, upstream promoter is initiated upon nutrient limitation. This results in the production of a low level of the immature and much less active form of the antibiotic. The exported version of the immature compound then functions in a feedforward mechanism to generate high levels of production of the lantibiotic by eliciting the release of MibX from MibW, either by interacting directly with the anti-σ factor or by interacting with lipid II thus evoking cell envelope stress that is sensed by MibW. The MibX released then activates transcription at its five cognate promoters, which include those for \( mibR \) and \( mibX \) itself, thus eliciting a positive-feedback loop.

Sequence analysis also revealed the presence of three pairs of genes (\( mibEF \), \( mibTU \) and \( mibYZ \)) encoding ABC (ATP-binding cassette) transporters. Although deletion of \( mibTU \) had no apparent effect on microbisporicin production, deletion of \( mibEF \) essentially abolished it [11]. Interestingly, BLASTp [6] (http://blast.ncbi.nlm.nih.gov/Blast.cgi) analyses revealed that, at that time, the homologues most similar to MibEF were components of ABC transporters from low-GC Gram-positive bacteria believed to play a role in conferring producer immunity to lantibiotics by actively transporting the modified peptide away from its lipid II target in the membrane. It is tempting to conclude that \( mibEF \) encode an immunity mechanism towards microbisporicin, and that, in its absence, a failsafe mechanism operates to prevent potentially lethal microbisporicin production.

Planosporicin

Planosporicin (Figure 2) is a 24-amino-acid lantibiotic made by \( \text{Planomonospora alba} \) NRRL 18924 that also inhibits cell wall biosynthesis in Gram-positive bacteria, probably by binding to lipid II. Its overall structure is similar to that of microbisporicin, with one MeLan and four Lan bridges, but instead of a C-terminal aminovinyl-cysteine residue, it terminates with a Lan [13]. It also lacks the chlorinated tryptophan and hydroxylated proline residues found in microbisporicin, and is not as potent as an antibiotic.

Genome scanning was used initially to identify the likely planosporicin biosynthetic gene cluster in a cosmid library of \( \text{P. alba} \) genomic DNA [14]. Heterologous expression in \( \text{Nonomuraea ATCC 39727} \) and deletion of \( pspA \) encoding the precursor peptide in the natural producer, combined with MALDI–TOF-MS, were employed to confirm its identity, and to define a minimal set of 15 genes required for planosporicin biosynthesis. In addition to the expected biosynthetic genes (\( pspB \) and \( pspC \), homologues of \( mibB \) and \( mibC \) respectively), homologues of \( mibR \), \( mibX \) and \( mibY \) (\( pspR \), \( pspX \) and \( pspY \) respectively) were also identified (Figure 3). High-resolution S1 nuclease protection analysis was used to locate four transcriptional start sites within the \( psp \) gene cluster, three of which were preceded by ECF promoter consensus sequences [15]. Reminiscent of the situation found with microbisporicin, the remaining promoter element, which is responsible for transcribing the \( pspABCTUV \) operon encoding all of the apparently
required biosynthetic enzymes and an ABC transporter, possessed instead sequences likely to be transcribed by the RNA polymerase holoenzyme containing HrdB. To establish transcriptional dependencies within the psp gene cluster, a series of transcriptional fusions were made in which the four psp promoters identified were cloned upstream of the β-glucuronidase reporter gene (gusA). These constructs were then introduced into Streptomyces coelicolor derivatives constitutively expressing either pspR or pspX. The resulting β-glucuronidase assays confirmed that PspX did indeed activate directly the expression of the promoter elements with the ECF consensus motifs, whereas PspR activated transcription from the promoter of the pspABC'TUV operon [15]. This again led us to propose a model of feedforward regulation. We suggest that, before the onset of planosporicin production, a basal level of expression of pspXW leaves the system poised for activation, with the PspX σ factor sequestered at the membrane by PspW. At the onset of nutrient limitation, a low level of transcriptional activation of the pspABC'TUV operon occurs, possibly mediated by the intracellular signalling molecule ppGpp [4], resulting in the production of a small amount of the lantibiotic. The resulting extracellular planosporicin then interacts either directly with PspW or with lipid II causing a low level of inhibition of peptidoglycan biosynthesis that is perceived by the anti-σ factor; in either case, PspX is released from PspW, resulting in high-level expression of the remaining three psp operons. This, in turn, results in high level expression of pspR and further activation of transcription of the pspABC'TUV operon and, consequently, high levels of planosporicin production.

The mechanism of autoinduction proposed for planosporicin, and indeed for microbisporicin, led us to wonder whether these molecules, in addition to possessing antibiotic activity, could function as signalling molecules to co-ordinate their own production in the wider population. To assess this possibility, we deleted pspW, resulting in markedly higher and apparently constitutive planosporicin production (Figure 4A). Interestingly, when the pspW mutant was grown adjacent to its parent, the constitutively produced lantibiotic was able to elicit precocious production in the wild-type strain [15] (Figure 4B). Similar observations were made using
culture supernatants from the pspW mutant (with a pspA deletion mutant serving as a negative control) (Figure 4C). The similar lantibiotic microbisporicin, and a range of other antibiotics with different mechanisms of action, failed to induce planosporicin biosynthesis. Thus planosporicin, and we suspect other actinomycete lantibiotics, and possibly other classes of actinomycete antibiotics, can trigger their co-ordinated production throughout the colony and nearby population, presumably in an attempt to achieve ecologically effective concentrations.

Initial attempts to express the planosporicin biosynthetic gene cluster in the genetically tractable S. coelicolor failed, as it had previously for microbisporicin. Given the marked increase in production elicited by deletion of pspW in P. alba, the mutant gene cluster was introduced into the S. coelicolor derivative M1152 that had been specifically engineered for improved heterologous expression [16]. Planosporicin production could now indeed be detected in the heterologous host, and the level of production was increased further by simultaneously expressing pspR from the constitutive ermE* promoter [15]. A similar stimulatory effect was observed in P. alba upon constitutive expression of pspR or pspX. A frequent stumbling block in the development of a natural product as it progresses toward commercialization is the provision of sufficient material for preclinical and clinical trials. In addition to providing fundamental insights into the regulation of antibiotic production in actinomycetes, these results illustrate how knowledge-based approaches can be used to effect marked increases in productivity, either by deleting negatively acting regulating genes or by overexpressing those encoding transcriptional activators, either in the natural producer or in a heterologous host.

Linaridins: a novel family of RiPPs derived by novel enzymology

Cypemycin

Cypemycin (Figure 2) is an extensively modified linear 22-amino-acid peptide produced by Streptomyces sp. OH-4156 with potent in vitro activity against mouse leukaemia cells and Micrococcus luteus [17]. Cypemycin does not contain Lan or Melan bridges, but exhibits some of the structural features of lantibiotics, notably four dehydrated threonine residues [Dhb (dehydrobutyrine) residues] and a C-terminal S-[(Z)-2-aminovinyl]-d-cysteine residue. Consequently, it was initially classified as a member of the lantibiotic family of RiPPs. Cypemycin also possesses two allo-isoleucine residues and an N-terminal N,N-dimethylalanine residue, both unique amino acid modifications. To gain a better understanding of cypemycin biosynthesis, we carried out genome scanning to identify the gene (cypA) encoding the precursor peptide using the predicted amino acid sequence of the unmodified core peptide as an in silico probe [18]. The resulting cypA and flanking nucleotide sequences were then used as a hybridization probe to screen an E. coli cosmid library of Streptomyces sp. OH-4156 genomic DNA, and hybridizing clones transferred by conjugation to Streptomyces venezuelae for initial activity screening using M. luteus as the assay organism (comparison of the two Streptomyces genome sequences suggests that the two strains are closely related). Production of cypemycin by some of the exconjugants was confirmed by MALDI–TOF-MS. Subsequent sequence analysis of one of the producing cosmid clones revealed, remarkably, that none of the enzymes apparently required for cypemycin production were orthologues of those involved in lantibiotic biosynthesis, despite the presence of the four Dhb residues and the C-terminal aminovinyl-cysteine moiety. A putative minimal gene set for cypemycin biosynthesis was identified (Figure 3), subcloned after introducing convenient flanking restriction sites by PCR-targeting, and transferred to S. coelicolor M1146 (another derivative of S. coelicolor engineered for heterologous expression of natural product gene clusters [16]) for further analysis. In-frame deletion mutations were then made by PCR-targeting in all of the genes contained within the minimal gene set, and their phenotypes analysed by bioassays against M. luteus and by MALDI–TOF- and ESI-MS, resulting in a putative novel biosynthetic pathway (Figure 5). Although sequence comparisons failed to reveal the dehydratase required for the formation of the Dhb residues, mutational analyses revealed that this activity must reside with CypH and/or CypL, both of...
which are essential for cypemycin biosynthesis. Since neither possess homologues of known function, this dehydration event may well occur via a novel enzymatic mechanism. The occurrence of a cysteine residue at position 19 in the core peptide region, rather than the expected serine residue usually required for the formation of the aminovinyl-cysteine, initially led us to think that formation of the aminovinyl-cysteine residue in cypemycin occurred via an unprecedented mechanism. Analysis of the ΔcypD mutant and an in vitro enzyme assay revealed that CypD is solely involved in decarboxylation of the C-terminal cysteine residue (Cys\(^{19}\)). Alkylation of the peptide produced by the cypD mutant with iodoacetamide and subsequent ESI analysis confirmed that Cys\(^{19}\) had been converted into Dha (dehydroalanine). Because the chemistry behind this reaction is similar to the dehydration of threonine to Dhb, we proposed that cysteine dehilation is also catalysed by CypH and/or CypL (Figure 5). Although they share little sequence similarity, CypD, which belongs to the distinct HFCD (homo-oligomeric flavin-containing cysteine decarboxylase) family of decarboxylases, presumably decarboxylates the C-terminal cysteine residue in a manner similar to LanD enzymes such as MibD. The biosynthesis of cypemycin thus represents an intriguing example of convergent evolution, where enzymes apparently evolutionarily unrelated to those involved in lanthipeptide biosynthesis have come together to create the same structural motifs.

**Grisemycin and other linaridins**

Bioinformatic analysis identified ten additional gene clusters that contained a cypL homologue together with a gene with full or partial homology with cypH and, with one exception, a gene encoding a putative precursor peptide [18]. These clusters are thus expected to be capable of producing linear (non-cyclized) dehydrated peptides. We thus proposed that cypemycin is the founding member of a new class of RiPPs: the linaridins. The gene clusters are phylogenetically widespread, occurring in different phyla of bacteria and archaea, suggesting that they play an important adaptive role in microbial physiology. Although cypemycin has antibiotic activity against *M. luteus*, it is inactive against many other bacteria and fungi, and the primary role of cypemycin, potentially the linaridin family of peptides, may be to function as extracellular signalling molecules rather than as antimicrobial agents.

One of the linaridin gene clusters identified was found in the genome sequence of *Streptomyces griseus*. Heterologous expression in *S. coelicolor* M1146 combined with MALDI-TOF- and ESI-MS resulted in the isolation of the 19-amino-acid linaridin grisemycin [19] (Figure 2). Grisemycin possesses at least three of the post-translationally modified residues found in cypemycin, but lacks a homologue of CypL, the gene putatively involved in the formation of the allo-isoleucine residues of cypemycin. Although neither the grisemycin (Figure 3) nor the cypemycin gene clusters contain putative regulatory genes, further studies in *S. coelicolor* revealed that the production of both linaridins is regulated by the transcriptional activator AdpA [19], a master regulatory protein that links \(\gamma\)-butyrolactone signalling to the expression of natural product gene clusters, and morphological differentiation, in many streptomycetes [4,5].

**Bottromycins: a precursor peptide with a follower, but no leader sequence**

The bottromycins [20–22] are a family of unique peptides produced by *Streptomyces bottropensis* DSM 402626-8 (Figure 2). They inhibit MRSA (methicillin-resistant *Staphylococcus aureus*) and VRE (vancomycin-resistant enterococci) by selectively blocking the binding of aminocacyl-tRNA to the A site of bacterial ribosomes. Bottromycins are structurally unusual, containing a macro lactamidine, several non-proteinogenic amino acids and a thiazole, and had been assumed to be produced by a non-ribosomal peptide synthetase. On the basis of earlier feeding experiments with labelled precursors, we wondered whether the bottromycins might instead originate from a ribosomally synthesized precursor that contained the core peptide sequence GPVVVFDC. The corresponding precursor peptide would then undergo a series of unusual post-translational modifications, including \(\beta\)-methylolation of phenylalanine, valine and proline residues, and proteolytic macro lactamidine formation. To assess this possibility, we carried out genome scanning of *S. bottropensis* [23]. tBLASTn searches of the resulting contig sequences using the putative bottromycin core peptide sequence identified a 10.2 kb DNA fragment that contained the core sequence embedded in a putative precursor peptide of 44 amino acids. Preliminary sequence analysis indicated that the contig did not contain the entire bottromycin biosynthetic gene cluster. To identify flanking contigs, amino acid sequences derived from partial PCSs (protein-coding sequences) present at each end of the contig were used to search the NCBI non-redundant protein sequence database using BLASTp. The complete amino acid sequences of the two proteins in the NCBI database with highest similarity to the sequences of the two protein fragments from the initially identified contig were then used to search the *S. bottropensis* contig database using tBLASTn, potentially identifying the flanking contigs. The contiguous nature of the three DNA fragments was confirmed by PCR using primers designed to anneal close to the ends of adjacent fragments and *S. bottropensis* genomic DNA as a template. This procedure, which we termed ‘bridge by blast’, was repeated until the entire sequence of the putative bottromycin biosynthetic gene cluster had been obtained, contained within a 44.5 kb contiguous region of the *S. bottropensis* genome (Figure 3). Deletion of the gene encoding the putative precursor peptide abolished the production of bottromycins in *S. bottropensis*, confirming the identity of the gene cluster. In contrast with other post-translationally modified ribosomal peptide antibiotics, the bottromycin precursor peptide lacks an N-terminal ‘leader’ sequence. Instead, it contains a C-terminal ‘follower’ sequence that is removed during post-translational processing and that is presumably required for correct maturation of the antibiotic.
Comparative sequence analyses of the proteins encoded by the genes flanking the precursor peptide gene identified ten enzymes that are likely to catalyse the post-translational modifications required for bottromycin assembly [23]. Gene clusters extremely similar to the S. bottropensis bottromycin biosynthetic gene cluster were also identified in the plant pathogen Streptomyces scabies 87.22 [24], the marine ascidian-derived Streptomyces spp. WM1B272 [25] and Streptomyces spp. BC16019 [26], and the production of bottromycins by each of these species confirmed. These findings led to a plausible pathway for bottromycin biosynthesis and set the stage for engineered biosynthesis of novel analogues with potentially improved clinical properties.

Tunicamycins: genome scanning reveals an unusual biosynthetic pathway

The tunicamycins are fatty acyl nucleoside antibiotics first isolated from Streptomyces lysosuperficus in 1971 [27,28] and later from Streptomyces chartreusis NRL3882. Their structures consist of an unusual 11-carbon aminodialdose core (tunicamine) to which uracil and GlcNAc (N-acetylglucosamine) are attached, alongside a range of amide-linked unsaturated fatty acids (Figure 2). The tunicamycins are potent inhibitors of bacterial cell wall biosynthesis that target MraY, which catalyses the formation of the peptidoglycan precursor undecaprenyl-pyrophosphoryl-N-acetylmuramoyl pentapeptide (lipid I). Unfortunately, they also inhibit eukaryotic protein N-glycosylation by binding to the active site of GPT (UDP-GlcNAc:dolichyl phosphate GlcNAc-1-phosphate transferase) and blocking production of the lipid-linked precursor Dol-PP-GlcNAc (dolichyl-GlcNAc-1-phosphate transferase) and blocking production of the lipid-linked precursor Dol-PP-GlcNAc (dolichylpyrophosphoryl-N-acetylglucosamine) required for glycosylation. The resulting toxicity has prevented the development of the tunicamycins as potentially clinically useful antibiotics. To gain an understanding of their biosynthesis, we carried out genome scanning to identify the tunicamycin biosynthetic gene cluster with the ultimate aim of engineering the biosynthetic pathway to produce derivatives that retain the ability to inhibit MraY, but that no longer prevent protein N-glycosylation [29].

Inspection of the structures of the tunicamycins suggested a role for a glycosyltransferase, a deacetylase and an acyl carrier protein in their biosynthesis. tBLASTn searches of the contig database of S. chartreusis with representatives of each of these classes of proteins identified a single contig that contained genes encoding the C-terminal portion of a glycosyltransferases and complete sequences for each of the other two types of protein. The ‘bridge-by-blast’ procedure outlined above for bottromycin was used to identify the adjacent contig containing the N-terminal coding sequence of the glycosyltransferase (confirmed by PCR analysis of genomic DNA). This led to the identification of 14 genes contained within a 12 kb region of the S. chartreusis genome that appeared to encode the entire tunicamycin biosynthetic pathway (Figure 3). This was subsequently confirmed by heterologous expression of this putative minimal gene set in S. coelicolor M1146 after identifying a cosmid clone containing the gene cluster. This led to the first informed proposal for a biosynthetic pathway for the tunicamycins, and subsequently revealed that tunicamycin biosynthesis occurs via unusual exo-glycal intermediates not encountered previously in the production of a specialized metabolite [30].

FK506 (tacrolimus): the use of PAC-based vectors to clone large biosynthetic gene clusters

The gene clusters described above all fall well within the size constraints (approximately 42 kb) of the in vitro packaging procedure used to make cosmid libraries. However, many natural product gene clusters, especially those encoding Type I polyketide synthases and/or non-ribosomal peptide synthetases, exceed this limit. One such gene cluster is that for the Type I polyketide FK506 (also known as tacrolimus) produced by Streptomyces tsukubaensis NRRL 18488 [31]. FK506 (Figure 2) is a high value calcineurin-inhibiting immunosuppressant [32,33] with annual sales of approximately US$2 billion. To clone and analyse such gene clusters requires the use of vectors, such as BAC (bacterial artificial chromosome) or PAC vectors, that can accommodate larger segments of genomic DNA. Using a conjugative and integrative PAC vector, pESAC13, developed by Margherita Sosio and Stefano Donadio (http://www.biost.com/page/PAC.aspx) and a tri-parental conjugation procedure based on one first developed at Diversa Corporation (M.J. Bibb, unpublished work), we were able to clone and heterologously express the 83.5 kb FK506 biosynthetic gene cluster contained within a genomic fragment of S. tsukubaensis DNA of approximately 110 kb in several S. coelicolor strains [34]. FK506 yields were increased approximately 5-fold (from 1.2 mg·l−1 to 5.5 mg·l−1) in S. coelicolor M1146 containing the FK506 PAC after overexpressing the pathway-specific regulatory gene fkbN.

The use of such conjugative PAC vectors is applicable to other large natural product gene clusters of biomedical and/or biotechnological interest, and, in addition to providing a relatively simple means to their isolation, it should also markedly assist in their further analysis and manipulation.

Engineered strains for heterologous expression of biosynthetic gene clusters

Much of the work described above was greatly facilitated by the development of a series of S. coelicolor strains (including M1146 and M1152 mentioned previously) specifically engineered for improved expression of heterologous gene clusters [16]. To remove potentially competitive sinks of carbon and nitrogen, and to provide a host devoid of antibiotic activity, we deleted four endogenous gene clusters.
from *S. coelicolor* M145 (a plasmid-free derivative of the wild-type strain); those encoding the biosynthesis of actinorhodin, the prodiginines, coelstycin P and the calcium-dependent antibiotic CDA. We then introduced point mutations into the *rpoB* and *rpsL* genes of the quadruple deletion mutant that had been shown previously to pleiotropically increase the level of specialized metabolite production [35–37]. Introduction of the native actinorhodin gene cluster and of gene clusters for the heterologous production of chloramphenicol and congoxamycin into these engineered derivatives revealed dramatic increases in antibiotic production compared with the parental strain. In addition to lacking antibacterial activity, the engineered strains possess relatively simple extracellular metabolite profiles. When combined with LC and MS, we believe that these genetically engineered strains will markedly facilitate the discovery of new compounds by heterologous expression of cloned gene clusters, particularly the numerous cryptic specialized metabolic gene clusters that are prevalent within actinomycete genome sequences. Recent genome sequencing has revealed that actinomycetes (and other groups of differentiating micro-organisms) contain a large number of specialized metabolite gene clusters with no known products (termed cryptic gene clusters), and thus these organisms have the potential to produce many more natural products than previously thought [38]. Many of these cryptic gene clusters are not expressed at significant levels in their natural hosts under typical laboratory screening conditions, but their expression can be activated by genetic manipulation of pathway-specific regulatory genes (see, e.g., [39,40]). However, when a strain is difficult to culture or genetically intractable, then heterologous expression of cloned gene clusters, particularly marked increases in antibiotic production compared with the engineered strains possess relatively simple extracellular metabolite profiles. When combined with LC and MS, we believe that these genetically engineered strains will markedly facilitate the discovery of new compounds by heterologous expression of cloned gene clusters, particularly the numerous cryptic specialized metabolic gene clusters that are prevalent within actinomycete genome sequences. Recent genome sequencing has revealed that actinomycetes (and other groups of differentiating micro-organisms) contain a large number of specialized metabolite gene clusters with no known products (termed cryptic gene clusters), and thus these organisms have the potential to produce many more natural products than previously thought [38]. Many of these cryptic gene clusters are not expressed at significant levels in their natural hosts under typical laboratory screening conditions, but their expression can be activated by genetic manipulation of pathway-specific regulatory genes (see, e.g., [39,40]). However, when a strain is difficult to culture or genetically intractable, then heterologous expression may be the only efficient approach for the identification of the products of these newly discovered gene clusters.

**Future perspectives**

Recent advances in genome sequencing and our improving ability to genetically manipulate streptomycetes and related actinomycetes provides new opportunities for the analysis and manipulation of natural product biosynthesis in these medically and agriculturally important micro-organisms. Moreover, the realization that these organisms contain numerous cryptic gene clusters and that they have the potential to produce many more specialized metabolites than previously thought provides new opportunities for the discovery of novel natural products with potential clinical, biotechnological and agricultural applications. Arguably, the application of genome scanning and genome mining not only to actinomycetes, but also to a wide range of micro-organisms and plants, may lead to a new era of natural product discovery. The prospects are indeed exciting.

**Acknowledgements**

I thank the many very talented students, postdoctoral fellows and other researchers who have, over the years, contributed to the research in my laboratory that led to this award. For the work described in this article, I thank particularly Govind Chandra, Jan Claesen, Lucy Foulston, Juan Pablo Gomez-Escribano, Andy Hesketh, Lorena Martinez-Fernandez, Mike Naldrett, Gerhard Saalbach and Emma Sherwood, and our collaborators Greg Challis, Ben Davis, Lijiang Song, Hua Wang and Filip Wyszynski.

**Funding**

Work in my laboratory was funded by the Biotechnology and Biological Sciences Research Council and the European Commission.

**References**

2. Bérdy, J. (2012) Thoughts and facts about antibiotics: where we are now and where we are heading. J. Antibiot. 65, 385–395

©The Authors Journal compilation ©2013 Biochemical Society


Received 4 September 2013

doi:10.1042/BST20130214