Analysis of the angucycline biosynthetic gene cluster in Streptomyces sp. QL37 and implications for lugdunomycin production

Heul, H.U. van der

Analysis of the angucycline biosynthetic gene cluster in *Streptomyces* sp. QL37 and implications for lugdunomycin production

Helga U. van der Heul

Provided by thesis specialist Ridderprint, ridderprint.nl
Printing: Ridderprint
Cover illustration: Helga U. van der Heul
Layout and design: Camiel Lemmens, persoonlijkproefschrift.nl

Analysis of the angucycline biosynthetic gene cluster in *Streptomyces* sp. QL37 and implications for lugdunomycin production

Proefschrift

ter verkrijging van

de graad van doctor aan de Universiteit Leiden,

op gezag van rector magnificus prof.dr.ir. H. Bijl,

volgens besluit van het college voor promoties

te verdedigen op woensdag 21 december 2022

klokke 15:00 uur

door

Hendrikje Ursula van der Heul

Geboren te Schiedam, Nederland

in 1990

Promotor:	Prof. dr. G.P. van Wezel
Co-promotor:	Dr. S.S. Elsayed
Promotiecommissie:	Prof. dr. A.H. Meijer Prof. dr. V. van Noort Prof. dr. D. Claessen Prof. dr. ir. A. J. Minnaard (Rijksuniversiteit Groningen) Prof. dr. J. Masschelein (Katholieke universiteit Leuven, België)

CONTENTS

General Introduction

The biology of *Streptomyces*

Most of the antibiotics used in the clinic are based on microbial natural products (Demain, 2014, Hutchings *et al.*, 2019, Wright, 2014). Especially *Streptomyces* and other genera of the Actinobacteria produce a plethora of bioactive natural products that are used in the pharmaceutical industry and agricultural sector; not only as antibiotics, but also as anti-cancer, anti-fungal, anti-viral agents or immunosuppressants (Barka *et al.*, 2016, Berdy, 2005). Streptomycetes are Gram-positive bacteria with a G+C rich genome that undergo a complex morphological development. Antibiotic production by streptomycetes is generally linked to their life cycle; this starts with the germination of a spore, initiating the growth of vegetative hyphae (Flärdh & Buttner, 2009). Via tip extension and branching, these grow out to form an intricate vegetative mycelium. Under stress conditions, such as nutrient depletion, new hyphae are formed on top of the vegetative mycelium that erect into the air, and are therefore called aerial hyphae. During this transition, a part of the vegetative mycelium undergoes programmed cell death, whereby nutrients are released from the lysed cell and used for the synthesis of the aerial hyphae (Chater & Chandra, 2006, Barka *et al.*, 2016, Flärdh & Buttner, 2009). Antibiotics are mostly produced during this stage of development, probably to protect the host against other microbes that seek to scavenge the released nutrients (Bibb, 2005, Tenconi *et al.*, 2020). The aerial hyphae further develop into spores that can then be spread in the environment and start a new lifecycle (Flärdh & Buttner, 2009).

Antibiotic discovery

In the Golden Age of antibiotic discovery (1940s-1960s), many clinically used antibiotics have been discovered from *Streptomyces* (Hutchings *et al.*, 2019). However, by the 1970s many strains and compounds were readily found again, a process called replication. Eventually, the return of investment of high-throughput screening in combination with targeted based approaches became too low (Kolter & van Wezel, 2016, Silver, 2011). Therefore, large industrial companies stopped investing in the discovery of novel antibiotics (Payne *et al.*, 2007, Brown & Wright, 2016). While drug discovery is failing, antimicrobial resistance (AMR) is on the rise; thus, pathogenic bacteria that once seemed like a thing from the past, are once more posing a huge threat to human health. These multi-drug resistant (MDR) pathogens include the so-called ESKAPE pathogens (*Enteroccocis faecium*, *Staphyloccocus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumanni*,

Pseudomonas aeruginosa, and *Enterobacter* species) and *Mycobacterium tuberculosis* (MDR-TB) (Rice, 2008, Ventola, 2015).

A possible solution came via next-generation sequencing, which revived the interest in antibiotic discovery in the early 21st century (Katz & Baltz, 2016, Bentley *et al.*, 2002). The production of antibiotics and other natural products is mediated by a cluster of genes that code for enzymes required for the biosynthesis of the compound(s). Such a cluster is called a biosynthetic gene cluster (BGC) (Medema *et al.*, 2015). Sequencing of *Streptomyces* genomes revealed that scientists had perhaps only scratched the surface in terms of the number of BGCs, whereby each *Streptomyces* species contains many more BGCs than initially thought. Even the model organism *Streptomyces coelicolor*, which had been worked on by thousands of scientists worldwide, had many BGCs that had never been seen (Bentley *et al.*, 2002). Until the genome sequence was published, no fewer than four antibiotics had been identified, namely undecylprodigiosin (Red), actinorhodin (Act), calcium dependent antibiotic (Cda) and the plasmid-encoded methylenomycin (Mmy). However, the genome sequence revealed in total 22 BGCs (Bentley *et al.*, 2002, Hoskisson & Seipke, 2020). One of these, the cryptic *cpk* cluster, is repressed by the global regulator DasR (Rigali *et al.*, 2008). It was later shown that this BGC specifies a novel antibiotic, known as coelimycin (Gomez-Escribano *et al.*, 2012). This is a well-know example of the principle of cryptic or silent BGCs, which started the postgenomic era of antibiotic discovery. Still, the major question we need to address is, are cryptic BGCs the answer to the AMR crisis? Can we really expect new chemistry, and in the end new drugs from them?

Activation of silent biosynthetic gene clusters

The reason why many BGCs are silent in the laboratory is most likely that the signals needed to activate them are not yet understood, and are to be found in the natural habitat, often the soil or a eukaryotic host. Indeed, streptomycetes are found everywhere, in the soil, sea, plants, animals and humans, and are part of a community with their own signals that can trigger antibiotic production in the bacterium (van der Meij *et al.*, 2017, van Bergeijk *et al.*, 2020). Regulatory networks play a key role in this phenomenon as it links environmental signals, such as the availability of sugars, nitrogen and phosphate supply to the activation of a BGC (van der Heul *et al.*, 2018, Liu *et al.*, 2013, van Bergeijk *et al.*, 2020).

All BGCs are regulated by one or more pleiotropic and cluster-situated regulators (Bibb, 2005, van der Heul *et al.*, 2018). These may either activate the transcription of the BGCs, or conversely, block the transcription of the BGC and thus act as repressors. External molecules can bind to these regulators and function as allosteric effectors, and thereby inhibit or activate the DNA binding of the activators or repressors (Liu *et al.*, 2013). One well-known example is the binding of glucosamine-6-phosphate to the pleiotropic regulator DasR, which releases DasR from the DNA, upon which the expression of the actinorhodin (Act), undecylprodigiosin (Red), coelimycin and calcium-dependent antibiotic (Cda) BGC is activated in *S. coelicolor* (Fillenberg *et al.*, 2016, Swiatek *et al.*, 2012, Rigali *et al.*, 2008).

One of the methods to activate antibiotic production is to find the specific elicitor (Zhu *et al.*, 2014b, van Bergeijk *et al.*, 2020, van Bergeijk *et al.*, 2022). A classical method is OSMAC (One Strain Many Compounds), whereby streptomycetes are grown at various conditions to activate multiple BGCs (Machushynets *et al.*, 2019). A recently developed method to activate silent BGCs in a more targeted manner is the high throughput elicitor screening method (HiTES). In this method a reporter gene is inserted in a BGC of interest and screened against a library of elicitors (Seyedsayamdost, 2014).

Genome mining

Genome mining using specific algorithms to predict BGCs has become a routine method (Medema *et al.*, 2011). These include antiSMASH, BAGEL, ClustScan, CLUSEAN, NP searcher, and PRISM (Medema *et al.*, 2011, de Jong *et al.*, 2006, Starcevic *et al.*, 2008, Weber *et al.*, 2009, Medema *et al.*, 2014, Li *et al.*, 2009, Skinnider *et al.*, 2015, Lee *et al.*, 2020). These detection methods are based on finding core genes of the known BGC classes, such as the non-ribosomal peptides (NRPS), Polyketide synthases type I, -II, -III (PKS I, PKS II, PKS IIII) and ribosomally synthesised and post-translationally modified peptides (RiPPs) (Medema *et al.*, 2021, Skinnider *et al.*, 2017, Helfrich *et al.*, 2019, Chevrette *et al.*, 2017). Based on the presence of accessory genes, such as transporters and modification genes, a border is predicted (Medema *et al.*, 2011). Prediction of BGCs is also based on comparison with known orthologous BGCs. However, these tools overlook BGCs of unknown compound classes. Therefore, it is important to develop new genomic mining algorithms. An example is the improvement on the identification

of novel subclasses of RiPPs using the machine-learning algorithm decRiPPter (Data-driven Exploratory Class-independent RiPP TrackER), which was developed in our laboratory and implemented in the novel version of antiSMASH, version 6.0 (Blin *et al.*, 2021, Kloosterman *et al.*, 2020a, Kloosterman *et al.*, 2020b).

BGCs with high similarity to known gene clusters are often discarded as they are likely to lead to the production of known molecules (Martinet *et al.*, 2020). However, while such BGCs may encode similar enzymes that catalyse known enzymatic reactions, they may still use different substrates and thus lead to new chemical diversity (Martinet *et al.*, 2020, Yan *et al.*, 2016). In addition, similar BGCs may be differentially expressed amongst different strains, due to the absence or presence of certain regulatory elements (Amos *et al.*, 2017). Thus, when an interesting, but silent BGC is found in one strain, it is possible to look for a similar BGC with different regulatory elements in another strain, where the cluster may be active. This offers more possibilities to identify the natural product that is related to the BGC.

Expanding the chemical space

A means to find novel antibiotics or other bioactive compounds is the screening of thousands of extracts derived from natural sources and study the present chemistry, which is the field of metabolomics (Zdouc *et al.*, 2021, Wu *et al.*, 2015b). However, the analysis of such extracts is challenging due to their complexity; they contain primary and secondary metabolites and residues from the growth medium (Wolfender *et al.*, 2019). A valuable development in the detection of natural products is high resolution mass spectrometry (MS). This method allows the efficient detection of molecules that would have been missed using the classical methods, such as thin layer chromatography (TLC) and liquid chromatography (LC) combined with UV detection, thus broadening the chemical space. The LC-MS data of extracts derived from different growth conditions can be used to prioritise metabolites for further study by applying uni-and multivariate statistical analyses (Beniddir *et al.*, 2021, Machushynets *et al.*, 2019, Wu *et al.*, 2016a). However, LC-MS data only comprise a list of masses and obtaining more information on structural features of the molecule of interest still requires isolation of the molecule combined with structural elucidation using nuclear magnetic resonance spectroscopy (NMR)(Rogers *et al.*, 2019). In addition, it remains difficult to dereplicate known compounds and identify yet unknown ones. For this reason,

the Global Natural Products Social Molecular Networking (GNPS) platform was developed, which is based on the data obtained from tandem mass spectrometry (MS/MS) fragmentation spectra. This tool made a huge contribution to the field of metabolomics, clustering molecules with similar MS/MS fragmentation patterns and therefore molecules with similar structures (Nothias *et al.*, 2020, Nguyen *et al.*, 2013, Schmid, 2020). An important implementation is the comparison of the MS/MS data with spectral libraries (Wang *et al.*, 2016). This allows rapid dereplication of known natural products and - more importantly - identification of yet unknown analogues of already known compounds [30, 50].

Linking Natural products with their cognate BGCs

Fermentation under different growth conditions in combination with GNPS networking and transcriptomics or proteomics can readily facilitate the linkage of natural products with their cognate BGCs (Du & van Wezel, 2018, Gubbens *et al.*, 2014, Wu *et al.*, 2015a, Amos *et al.*, 2017). The identified BGC may in turn be genetically modified and/ or expressed heterologous in an optimised *Streptomyces* host, so as to increase the production of the natural product of interest for upscaling, structural elucidation and bioactivity assays (Li *et al.*, 2019, Jones *et al.*, 2013, Gomez-Escribano & Bibb, 2011, Zheng *et al.*, 2007, Hutchinson & Colombo, 1999). The characterisation of a BGC may help in the discovery of the biosynthetic pathway leading to the synthesis of natural product and thereby the discovery of novel modification enzymes that may be used to synthesise new natural product analogues which are more bioactive or can overcome antibiotic resistance mechanisms (Xiao *et al.*, 2020, Kim *et al.*, 2015a, Baltz, 2018, Hulst *et al.*, 2021).

Scope of this thesis

In the attempt to find novel antibiotics, a strain isolated from the Qinling mountains in China was screened using the OSMAC method. This led to the discovery of lugdunomycin (Wu *et al.*, 2019), which is produced by *Streptomyces* sp. QL37 (Zhu *et al.*, 2014b) . The molecule is characterised by its heptacyclic ring, a spiroatom, benzaza[4,3,3]propellant moiety, and two all carbon stereocenters (Wu *et al.*, 2019). Lugdunomycin is derived from the angucyclines, molecules with a benz[a] anthracene backbone that form one of the largest classes of natural products and are produced by polyketide type II synthases (Kharel *et al.*, 2012). Angucyclines are characterised by their benz[a]anthracene core with an additional carbohydrate

residue. Lugdunomycin is derived from the angucyclinones; angucyclines without the sugar residue (Mikhaylov *et al.*, 2021, Wu *et al.*, 2019). In this thesis we further refer to these natural products as angucyclines, which are well-known for their antibacterial and anti-cancer activities. However, their toxicity makes them less attractive for clinical application, which may be solved by creating chemical analogues (Mikhaylov *et al.*, 2021). A biosynthetic pathway was proposed for the production of lugdunomycin (Figure 1). This starts with one molecule of acetyl-CoA and nine molecules of malonyl-CoA. The minimal polyketide machinery, consisting of the heterodimer ketosynthase alpha (KSα), ketosynthase beta (KSβ) and acyl carrier protein (ACP) use these primary metabolites for the synthesis of a decaketide. Ketoreductases and cyclases then catalyse this carbon chain to generate the basic angucycline framework. The first stable intermediate is the angucycline UWM6 or its dehydrated form prejadomycin. Post-PKS tailoring enzymes, including a methyltransferase, reductases and oxygenases subsequently modify the angucyclines to a final product (Kharel *et al.*, 2012). The hypothesised biosynthesis pathway of lugdunomycin itself requires two key steps, namely (I) the rearrangement of an angucycline by opening of the C-ring of an angucycline backbone through a Baeyer-Villiger reaction; and (II) a final reaction wherein the rearranged angucycline reacts with a dienophile in an intermolecular Diels-Alder reaction (Wu *et al.*, 2019). Though angucyclines are produced in high quantities under most conditions, lugdunomycin itself is produced in very small amounts. More insights into the mode of action and biosynthesis of lugdunomycin is important to further augment our understanding of angucyclines in general, and of lugdunomycin in particular.

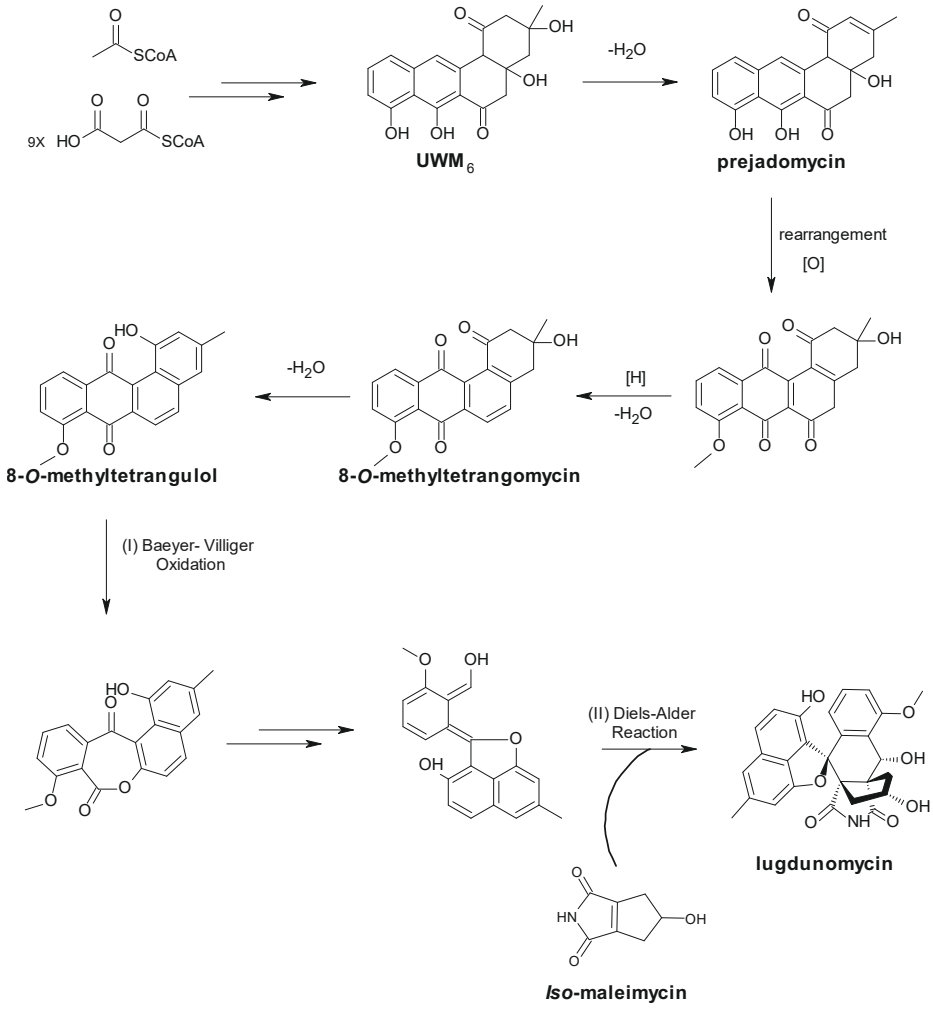


Figure 1 Postulated simplified lugdunomycin biosynthesis pathway
The exceptional transformation involves a C-ring cleavage between bond C6a–C7 in the angucycline framework and a Diels-Alder reaction between an angucycline derived diene and the dienophile *iso*-maleimycin (Wu *et al.*, 2019).

The aim of this thesis is to obtain novel insights into the transcriptional control of the lugdunomycin BGC and into its biosynthesis pathway, and to apply this knowledge to increase the production of lugdunomycin.

Chapter 2 reviews the regulatory networks that control antibiotic production in Actinobacteria. The includes specific and global control of BGCs, the impact of

primary metabolism (such as the availability of nitrogen, phosphate and carbon sources) on antibiotic production, the link between the complex life cycle and antibiotic production and the involvement of specialised metabolites in the regulation of BGCs (van der Heul *et al.*, 2018).

In **Chapter 3** the BGC governing lugdunomycin biosynthesis is characterised. The *lug* gene cluster is a polyketide type II BGC comprising 28 genes. The importance of this gene cluster in lugdunomycin biosynthesis was verified using gene deletion experiments and the extent of the cluster was assessed using temporal RNA-seq. Furthermore, bioinformatic comparisons were used to find other *lug*-type gene clusters in other *Streptomyces* and *Kitasatospora* strains and we evaluated whether the presence of a *lug*-type gene cluster in a strain is correlated to phylogeny.

The functional role of the *lug* gene cluster in lugdunomycin biosynthesis was further examined by heterologous expression in **Chapter 4.** A paired-omics approach, including genomics, metabolomics and transcriptomics, was applied to find a missing link in lugdunomycin biosynthesis.

The opening of the C-ring of the angucyclines is an important step in lugdunomycin biosynthesis. It was hypothesised that a C-ring rearranged angucycline is used as a substrate in the final reactions of lugdunomycin biosynthesis, presumably via a Baeyer-Villiger oxidation. In **Chapter 5** the five oxygenases encoded by the *lug* gene cluster have been studied in detail, using gene deletion strategies combined with metabolomics and GNPS networking tools. A primary focus was to discover the oxygenases required for cleavage of the angucycline C-ring.

The *lug* gene cluster contains five regulatory genes, which are studied in **Chapter 6**. Gene deletion experiments in combination with metabolomics and proteomics were applied to identify the role of the five *lugR* genes in the transcriptional control of the *lug* cluster, and identify the key transcriptional activators. Over-expression of the regulators was used as an approach to increase the production of angucyclines and in particular of lugdunomycin.

Chapter 7 presents a summarising discussion of the main findings of this thesis, including conceptual ideas on how to proceed with studies of the biosynthesis of lugdunomycin, aimed at obtaining a full picture of the mode of action of this exciting molecule.

Regulation of antibiotic production in Actinobacteria: new perspectives from the post-genomic era

Helga van der Heul[1]
Bohdan Bilyk[2]
Kenneth J. McDowall[2]
Ryan F. Seipke[2]
Gilles P. van Wezel[1]

This chapter is published as:
van der Heul, H.U., Bilyk, B.L., McDowall, K.J., Seipke, R.F., and van Wezel, G.P. (2018) Regulation of antibiotic production in Actinobacteria: New perspectives from the post-genomic era. *Nat Prod Rep* **35**: 575–604.

ABSTRACT

The antimicrobial activity of many of their natural products has brought prominence to the Streptomycetaceae, a family of Gram-positive bacteria that inhabit both soil and aquatic sediments. In the natural environment, antimicrobial compounds are likely to limit the growth of competitors, thereby offering a selective advantage to the producer, in particular when nutrients become limited and the developmental programme leading to spores commences. The study of the control of this secondary metabolism continues to offer insights into its integration with a complex lifecycle that takes multiple cues from the environment and primary metabolism. Such information can then be harnessed to devise laboratory screening conditions to discover compounds with new or improved clinical value. Here we provide an update of the review we published in NPR in 2011. Besides providing the essential background, we focus on recent developments in our understanding of the underlying regulatory networks, ecological triggers of natural product biosynthesis, contributions from comparative genomics and approaches to awaken the biosynthesis of otherwise silent or cryptic natural products. In addition, we highlight recent discoveries on the control of antibiotic production in other Actinobacteria, which have gained considerable attention since the start of the genomics revolution. New technologies that have the potential to produce a step change in our understanding of the regulation of secondary metabolism are also described.

Covering: 2000-2018

INTRODUCTION

Streptomyces species are renowned for their ability to produce a multitude of bioactive secondary metabolites, some of which have been co-opted clinically as a source of antibacterial, anticancer, antifungal, antiparasitic and immunosuppressive agents (Baltz, 2008, Barka *et al.*, 2016, Caffrey *et al.*, 2008, Hopwood, 2007, Olano *et al.*, 2009). The secondary metabolites produced by this taxon offer a chemical diversity that greatly exceeds that of libraries of compounds synthesized chemically and have been pre-selected through millions of years of evolution to interact effectively with biological targets. With the development of numerous approaches for counter selecting compounds with activities that have been previously characterised and in the case of antibiotics might have been rendered ineffective by the emergence of resistance, natural products are being revisited as a potential source of new pharmaceuticals (Harvey *et al.*, 2015, Lewis, 2013). The biological role of antibiotics has been a topic of some debate. Whilst antibiotics in the natural habitat are typically regarded as weapons, in the same way as they are used in the clinic (Abrudan *et al.*, 2015, Raaijmakers & Mazzola, 2012, Ratcliff & Denison, 2011), it has been argued that at least some could function primarily in cell communication and signalling (Davies *et al.*, 2006, Linares *et al.*, 2006, Romero *et al.*, 2011). The latter view was based largely on the believe that compounds with antibiotic activity are unlikely to reach concentrations in the soil that block growth, as defined by the minimal inhibitory concentration (MIC). However, selection for resistance occurs even at concentrations far below the MIC and antibiotic sensitive strains are demonstrably disadvantaged in competing for growth (Andersson & Hughes, 2014, Gullberg *et al.*, 2014, Westhoff *et al.*, 2017). The majority of the antibiotics that are used in the clinic are produced by Actinobacteria, which are high G + C, Gram-positive bacteria. Of the Actinobacteria, perhaps the most prolific antibiotic producers are members of the genus *Streptomyces*, which belong to the family Streptomycetaceae (Barka *et al.*, 2016, Labeda *et al.*, 2012, Ludwig *et al.*, 2012).

Streptomycetes are found in environments with varying nutrient supply, and metabolise a variety of carbon, nitrogen and phosphate sources. To respond appropriately to the challenges imposed by the environment, the genome of the model streptomycete *S. coelicolor* harbours a staggering 700 regulatory genes (Bentley *et al.*, 2002). Streptomycetes have a multicellular life cycle,

which culminates in sporulation. The reader is referred elsewhere for details of this process (Chater, 2011, Chater & Losick, 1997, Claessen *et al.*, 2014, Flärdh & Buttner, 2009, Hopwood, 1999). In brief, streptomycetes grow as non-motile, vegetative hyphae to produce a network of interwoven filaments called vegetative mycelium. When reproduction is required, for example at the time when nutrients run out, the vegetative mycelium acts as a substrate for newly formed aerial hyphae that eventually differentiate into chains of unigenomic exospores. Genes required for the transition from vegetative to aerial growth are typically referred to as *bld* genes, referring to their bald phenotype, due to their failure to produce the fluffy white aerial hyphae (Merrick, 1976). Mutants that produce aerial hyphae but no spores are referred to as *whi* mutants, for their white phenotype caused by the lack of the grey spore pigment (Chater, 1972). Many of the *bld* and *whi* mutants that had been isolated in the 1970s by phenotypic screening have later been identified by genetic complementation experiments, and they have been instrumental in providing better insights into the regulatory cascades that control morphological differentiation. For details we refer the reader to excellent reviews elsewhere (Barka *et al.*, 2016),(Flärdh & Buttner, 2009),(Chater *et al.*, 2010, Chater & Chandra, 2006, Hopwood, 2006, Kelemen & Buttner, 1998). Production of bioactive compounds is typically linked to the developmental lifecycle, and antibiotics are presumably produced to safeguard the nutrient supply during developmental growth (Bibb, 2005, Liu *et al.*, 2013, van Wezel & McDowall, 2011). Streptomycetes produce an arsenal of degradative enzymes (e.g. glycosyl hydrolases, lipases and proteases), which combined with the production of antibiotics and the ability to form desiccation-resistant exospores has facilitated their success in a multitude of soil environments and sediments including those of marine and freshwater ecosystems. The competitive attributes possessed by streptomycetes have not gone unutilised by higher organisms. For instance, it has become clear that many insects, animals and plants engage in protective symbioses with antibiotic-producing *Streptomyces* species (reviewed in ref. (Seipke *et al.*, 2012) and (van der Meij *et al.*, 2017)). However, not all interactions between streptomycetes and higher organisms are beneficial – a minority of species produce a cellulose synthase inhibitor called thaxtomin and a coronafacic acid-like phytotoxin, which lead to the development of scab diseases on potato and other tap-root crops (Bignell *et al.*, 2010, Loria *et al.*, 2008). Over the past 50 years, *S. coelicolor* has been the major model for the study of antibiotic production and its control. Early on it was apparent that this strain produced numerous natural products, including

actinorhodin (Act (Rudd & Hopwood, 1979)), undecylprodigiosin (Red (Feitelson *et al.*, 1985)), the calcium-dependent antibiotic (Cda (Hopwood & Wright, 1983)) and plasmid-encoded methylenomycin (Mmy (Wright & Hopwood, 1976)). The genes that encode the machinery for the production of these respective antibiotics are clustered together in 'biosynthetic gene clusters' (BGCs), which typically also harbour resistance gene(s) and one or more transcriptional regulators that control biosynthesis. Sequencing of the *S. coelicolor* genome was a landmark event that revealed an unexpected potential for the production of hitherto unidentified or cryptic natural products (Bentley *et al.*, 2002),with more than 20 BGCs specifying a diverse range of secondary metabolites (Challis & Hopwood, 2003, van Keulen & Dyson, 2014). One of these is a so-called cryptic polyketide antibiotic (later named coelimycin), which is only produced under specific growth conditions (Gomez-Escribano *et al.*, 2012, Pawlik *et al.*, 2007). Sequencing of other model Actinobacteria revealed a similar picture, with some species harbouring more than 50 different BGCs (Cao *et al.*, 2016, Cruz-Morales *et al.*, 2013, Ikeda *et al.*, 2003, Ohnishi *et al.*, 2008, Oliynyk *et al.*, 2007, Udwary *et al.*, 2007). Thus, the potential of Actinobacteria as producers of bioactive molecules was found to be much greater than was initially thought. This prompted the sequencing and analysis of the genomes of a large array of species to identify novel BGCs (reviewed in ref. (Gross, 2009, Medema *et al.*, 2011, Nett *et al.*, 2009, Zerikly & Challis, 2009)). plus the development of approaches to induce the production of natural products under laboratory conditions (Baltz, 2008), (Ochi *et al.*, 2014, Rutledge & Challis, 2015, Yoon & Nodwell, 2014, Zhu *et al.*, 2014a). The identification of BGCs is now relatively routine using bioinformatics tools, such as antiSMASH (Medema *et al.*, 2011), CLUSEAN (Weber *et al.*, 2009) and PRISM (Skinnider *et al.*, 2017). Available also are tools for the identification of BGCs corresponding to specific classes of natural product, e.g. NRPSPredictor for nonribosomal peptides (Rottig *et al.*, 2011), BAGEL for bacteriocins and lantibiotics (de Jong *et al.*, 2010) and SEARCHPKS for polyketides (Yadav *et al.*, 2003). For a comprehensive overview of the available bioinformatic tools for genome mining we refer the reader to excellent reviews elsewhere (Chavali & Rhee, 2017, Ziemert *et al.*, 2016). This review is intended to be an update to our comprehensive review on the same subject published in this journal in 2011 (van Wezel & McDowall, 2011).The broad subject is covered, but in the interest of limiting duplicated content, the reader is often referred to our previous review. Here, the focus lies on recent insights into the regulation of natural product biosynthesis in streptomycetes, based on the literature from the

period of 2011–2017. The article focuses on both pleiotropic and cluster-situated regulators, highlighting recent discoveries. We thereby give specific attention to the control of antibiotic production in other Actinobacteria. We also provide an update on our understanding of the links between primary and secondary metabolism and ecological triggers that stimulate natural product biosynthesis, and outline methodology that could be used to activate silent or cryptic natural product biosynthetic pathways.

TRANSCRIPTIONAL REGULATION BY CLUSTER-SITUATED REGULATORS

Over the last several decades, investigations into the regulation of the antibiotics produced by *S. coelicolor* (Act, Red, Cda, Mmy and coelimycin) and that of streptomycin biosynthesis by *S. griseus* have established key aspects of the regulation of secondary metabolism in *Streptomyces*. For details we refer to reviews elsewhere (Bibb, 2005, Liu *et al.*, 2013, van Wezel & McDowall, 2011). The regulation of secondary metabolism is complex and frequently involves pleotropic global regulators that either directly activate or repress biosynthetic genes or do so via cluster-situated repressors or activators. A plethora of regulatory proteins is involved in the control of antibiotic production, across a broad range of regulator families. and cross-regulation results in a highly complex regulatory network. This is necessary to correctly interpret the environmental signals and translate them into appropriate transcriptional responses, so as to time the production of natural products, often closely connect to development. The different families of transcriptional regulators known to be involved in the control of antibiotic production, and some well-studied examples, are provided in Table 1.

The regulation of the BGCs for actinorhodin (Act; controlled by ActII-ORF4), undecylprodigiosin (Red, controlled by RedD) and calcium-dependent antibiotic (Cda, controlled by CdaR) of *S. coelicolor* and for streptomycin (Str, controlled by StrR) are the most well-studied examples of cluster-situated regulators (CSRs). ActII-ORF4, CdaR and RedD belong to the SARP family of *Streptomyces* antibiotic regulatory proteins (Wietzorrek & Bibb, 1997), while StrR unusually belongs to the family of ParB-Spo0J proteins, most of which are involved in DNA segregation and sporulation (Autret *et al.*, 2001). All available evidence supports the conclusion that the cellular level of a cluster-situated regulator dictates the level of transcription

of its cognate BGC, which correlates closely with the level of production of the corresponding natural product (Gramajo *et al.*, 1993, Tomono *et al.*, 2005). Indeed, the timing of Red production fully depends on the promoter that drives the transcription of *redD*, allowing its use as a transcriptional reporter system (van Wezel *et al.*, 2000b). Thus, the ultimate factor deciding whether or not a BGC is expressed is its CSR(s). While ActII-ORF4 and StrR act as single CSRs within their respective BCGs, production of RedD is in turn controlled by RedZ (Guthrie *et al.*, 1998, White & Bibb, 1997), which is related to the response regulators (RR) of prokaryotic two-component systems (TCS) but 'orphaned', i.e. not genetically linked to a histidine kinase (Hutchings, 2007). It is becoming increasingly clear that the presence of multiple CSRs is more often the rule than the exception with each regulator effecting control of a subsets of genes or contributing to a hierarchical cascade. The latter is exemplified by the BGCs specifying polyene antifungal compounds such as amphotericin, nystatin, natamycin (pimaricin) and candicidin (Carmody *et al.*, 2004, Sekurova *et al.*, 2004, Vicente *et al.*, 2014, Zhang *et al.*, 2015b). It has been assumed and, in some cases, shown that many regulators are responsive to small molecule signals. It has been assumed and in some cases shown that many regulators are responsive to small molecule signals. Regulators responsive to autoregulatory molecules such as γ-butyrolactones are well known (Takano, 2006, Willey & Gaskell, 2011), and feedback control by biosynthetic intermediates over production or export has been demonstrated for jadomycin, Act and simocyclinone biosynthesis (Tahlan *et al.*, 2007, Wang *et al.*, 2009, Willems *et al.*, 2008). However, the identity of the ligands/signals perceived by both pleiotropic and CSRs is a major question within the field, and if answered could lead to a revolution in chemical genetic tools for the stimulation of natural product biosynthesis, and thus drug discovery.

Table 1a Major families of regulators involved in the control of antibiotic production. Representative examples and their host and target are indicated

Family[d]	Example	Host[a]	Control[b]	Target BGC[c], comment	Reference
SARP	ActII-ORF4, RedD, CdaR	*S. coelicolor*	(+)	Act, Red, Cda, respectively	(Wietzorrek & Bibb, 1997)
	AfsR	*S. coelicolor*	(+)	Activates transcription of AfsS	(Tanaka *et al.*, 2007)
	FarR3/ Far4	*S. lavendulae*	(+,-)	Indigoidine, nucleoside and D-cycloserine	(Kurniawan *et al.*, 2014)
StrR (ParB-Spo0J)	StrR	*S. griseus*	(+)	Streptomycin	(Vujaklija *et al.*, 1993)
	Tei15*	*Actinoplanes teichomyceticus*	(+)	Teicoplanin	(Horbal *et al.*, 2014, Horbal *et al.*, 2012)
	Dbv4	*Nonomureae* sp. ATCC39727	(+)	A40926	(Alduina *et al.*, 2007, Lo Grasso *et al.*, 2015)
LAL	FscRI	*S. albus*	(+)	Candicidin and antimycin	(McLean *et al.*, 2016)
	AveR	*S. avermitilis*	(+,-)	Avermectin and oligomycin	(Guo *et al.*, 2010)
	Dbv3	*Nonomureae* sp. ATCC39727	(+)	A40926	(Lo Grasso *et al.*, 2015)
TetR	AtrA	*S. griseus*	(+)	Global regulator	(Hirano *et al.*, 2008),
	ArpA	*S. griseus*	(-)	GBL-receptor, repressor of *adpA*	(Ohnishi *et al.*, 1999)
	ScbR	*S. coelicolor*	(+,-)	GBL-receptor	(Yang *et al.*, 2015)
AraC/XylS	AdpA	*S. griseus*	(+)	Activates StrR expression	(Akanuma *et al.*, 2009)

Table 1a Major families of regulators involved in the control of antibiotic production. Representative examples and their host and target are indicated *(continued)*

Family[d]	Example	Host[a]	Control[b]	Target BGC[c], comment	Reference
GntR	DasR	*S. coelicolor*	(+,-)	Global regulator of antibiotic production; effector molecule is N-acetylglucosamine	(Rigali *et al.*, 2008, Swiatek-Polatynska *et al.*, 2015)
c-AMP receptor protein	Crp	*S. coelicolor*	(+)	Regulator coordinating development, primary and secondary metabolism	(Gao *et al.*, 2012)
Orphan RR	RedZ	*S. coelicolor*	(+)	Red	(Guthrie *et al.*, 1998)
	GlnR	*S. coelicolor*	(+)	Act and Red	(Tiffert *et al.*, 2011, Tiffert *et al.*, 2008)
TCS	AbsA1/AbsA2	*S. coelicolor*	(-)	Act, Red, Cda	(McKenzie & Nodwell, 2007)
	AfsQ1/2	*S. coelicolor*	(+)	Act, Red, Cda; responds to nitrogen	(Shu *et al.*, 2009)
	PhoRP	*S. coelicolor*	(+,-)	Act; global regulator	(Santos-Beneit, 2015, Seipke, 2015)
	DraR/K	*S. coelicolor*	(+,-)	Act, Red, coelimycin, responds to high concentrations of nitrogen.	(Yu *et al.*, 2012)
	OsdR/K	*S. coelicolor*	(+)	Act, responds to oxygen level	(Urem *et al.*, 2016b)

Table 1b Major families of regulators involved in the control of antibiotic production. Representative examples and their host and target are indicated

Family[d]	Example	Host[a]	Control[b]	Target BGC[c], comment	Reference
ROK	Rok7B7	*S. coelicolor*	(+,-)	Act, Red, Cda; CCR	(Swiatek *et al.*, 2013, Park *et al.*, 2009)
σ Factor	MibX/MibW	*Microbispora corallina*	(+)	Microbisporicin	(Fernandez-Martinez *et al.*, 2015, Foulston & Bibb, 2011)
	Sigma(AntA)	*S. albus*	(+)	Antimycin.	(Seipke *et al.*, 2014)
BldB	BldB	*S. coelicolor*	(+)	Antibiotic production, development and CCR.	(Pope *et al.*, 1998, Pope *et al.*, 1996, Eccleston *et al.*, 2002)
tRNA	BldA	*Streptomyces species*		leucine-tRNA for UAA codon. Translational control of antibiotic production and morphogenesis	(Hackl & Bechthold, 2015)
XRE	MmyB	*S. coelicolor*	(+)	Methylenomycin B; controlled by furans	(O'Rourke *et al.*, 2009, Xu *et al.*, 2012)
Wbl (WhiB- like protein)	WblA	*S. coelicolor*	(-)	Pleiotropic regulator of antibiotic production and development	(Kang *et al.*, 2007)
LacI	AcrC	*Actinoplanes* s p. SE50/110	(-)	Acarbose	(Wolf *et al.*, 2017b)
LmbU		*S. lincolnensis*	(+,-)	Lincomycin	(Hou *et al.*, 2018, Ju *et al.*, 2018)

Table 1b Major families of regulators involved in the control of antibiotic production. Representative examples and their host and target are indicated *(continued)*

Family[d]	Example	Host[a]	Control[b]	Target BGC[c], comment	Reference
Lrp/AsnC	SCO3361	*S. coelicolor*	(+)	Act; control by amino acids	(Liu *et al.*, 2017)
NsdA	*NsdA*	*S. coelicolor*	(-)	Act, Cda, Mmy	(Li *et al.*, 2006)
IclR	NdgR	*S. coelicolor*	-	Act; dependent on amino acids.	(Yang *et al.*, 2009)
MarR	DptR3	*S. roseosporus*	+	Daptomycin	(Zhang *et al.*, 2015c)

[a] *Streptomyces* abbreviated with 'S.' [b] Activation indicated by +, repression by -
[c]Act, a ctinorhodin; Cda, calcium-dependent antibiotic; Red, prodiginines; Mmy, methylenomycin.[d] LAL, large ATP-binding regulators of the LuxR family (In the text mentioned as LuxR); XRE, xenobiotic response element.

Pathway-specific regulation: streptomycin and actinorhodin as paradigms

The first complete regulatory pathway leading to activation of a BCG was described for Str in *S. griseus* (Horinouchi, 2007). Transcription of StrR, which as mentioned above is the corresponding CSR, is activated by the pleiotropic regulator AdpA (A-factor-dependent protein; (Ohnishi *et al.*, 1999), whose transcription depends on the accumulation of the γ-butyrolactone 2-isocapryloyl-3*R*-hydroxymethyl-γ-butyrolactone, better known as A-factor. The hormone-like compound binds to ArpA (Onaka *et al.*, 1995), which acts as a repressor of *adpA* transcription (Onaka & Horinouchi, 1997). AdpA also activates morphological differentiation, and thus plays a key role in the coordination of chemical and morphological differentiation (Ohnishi *et al.*, 2005, Akanuma *et al.*, 2009). A-factor is synthesized by the enzyme AfsA (Kato *et al.*, 2007). The role of A-factor in the control of antibiotic biosynthesis is further discussed in Section 9.

The transcription of *strR* is subject to multi-level control, and in particular by the pleiotropic regulator AtrA (Hirano *et al.*, 2008, Hong *et al.*, 2007), which has an orthologue in *S. coelicolor* that activates transcription of *act*II-ORF4, the CSR within the *act* cluster (Uguru *et al.*, 2005). Binding of AtrA *in vivo* within the vicinity of the *act*II-ORF4 promoter has recently been confirmed by chromatin immunoprecipitation in combination with DNA sequencing (ChIP-seq) (McDowall *et al*, unpubl. data). Compared to what is known about *strR*, the control of *act*II-ORF4 is complex with many transcription factors reported to control its expression directly. Numerous direct and indirect regulators have been identified (Liu *et al.*, 2013),(van Wezel & McDowall, 2011). Some of the most recent examples are summarized in Table 2. For some of these transcription factors, binding has been demonstrated *in vivo* by ChIP-based approaches. In addition to AtrA, these include DasR (Swiatek-Polatynska *et al.*, 2015), a member of the GntR family that controls the uptake and metabolism of N-acetylglucosamine (GlcNAc) and the degradation of chitin to GlcNAc (Colson *et al.*, 2007, Rigali *et al.*, 2006), AbsA2 (McKenzie & Nodwell, 2007), the response regulator of the AbsA TCS, which negatively controls antibiotic production in *S. coelicolor* (Brian *et al.*, 1996, Champness *et al.*, 1992), AbrC3 (Rico *et al.*, 2014a), a response regulator of a TCS that is atypical in having two histidine kinases (Yepes *et al.*, 2011), and Crp (Gao *et al.*, 2012), the cyclic AMP receptor protein, which is perhaps best known for mediating carbon catabolite repression of the *lac* operon in *E. coli* (Gorke & Stülke, 2008), controls diverse cellular processes in many bacteria (Korner *et al.*, 2003), and is a key regulator of

secondary metabolism as well as spore germination and colony development in *S. coelicolor* (Piette *et al.*, 2005). In addition to direct regulation, the expression of *act*II-ORF4 is dependent on *relA* (Chakraburtty & Bibb, 1997), which is required for induction of the stringent response. The stringent response enables bacteria to survive sustained periods of nutrient deprivation by enhancing the transcription of numerous genes required to survive stress, while lessening transcription of genes, such as those specifying stable RNAs, whose products are required in significantly reduced amounts during periods of slowed growth (Kang *et al.*, 1998, Sun *et al.*, 2001). Whilst the signals transduced by Crp and the stringent response are well described, the signals sensed or transduced by most of the transcription factors that bind the *act*II-ORF4 promoter remain to be elucidated. An exception is DasR, which is a receptor for glucosamine-6-phosphate (GlcN-6P), an intermediate in GlcNAc metabolism, and derivatives (Rigali *et al.*, 2006). The binding of GlcN-6P by DasR reduces its affinity for DNA, which de-represses the expression of genes that facilitate the degradation of chitin to GlcNAc and its uptake and metabolism (Colson *et al.*, 2007, Rigali *et al.*, 2006). Links between DasR and AtrA are described later in this review (Section 5.3).

Table 2a Recently discovered transcriptional regulators that control antibiotic production in *S. coelicolor.* Orthologues also studied *in S. avermitilis* or *S. venezuelae* are indicated

Gene	ID[a]	Functions(s) of the regulators(s)[b]	Ref.
Regulators known to directly control antibiotic BGCs			
mtrAB	SCO3013/2 SVEN2756/5	TCS; MtrA activates *act*II-ORF4 and *redZ* and links proction to development.	(Som *et al.*, 2017)
draRK	SCO3063/2; SAV3481/0	TCS; regulator of *actII-ORF4* and *kasO* in *S. coelicolor* and of *olmRI* in *S. avermitilis*. Impacts Red and Ave production in *S. coelicolor* and *S. avermitilis*, resp.	(Yu *et al.*, 2012)
	SCO3361	Lrp/AsnC family positive regulator for Act production. Binds to *act*II-ORF4 (EMSA)	(Liu *et al.*, 2017)
Crp	SCO3571	Regulator of primary and secondary metabolism; activates *act*II-ORF4, *cdaR* and *cpkA* (Chip-seq).	(Gao *et al.*, 2012)
glnR	SCO4159; SAV4042	Activator of *act*II-ORF4 and repressor of *redZ* in *S. coelicolor* (EMSA). Activator of *aveR* (avermectin) and repressor of olmRI/olmRII (oligomycin) in S. avermitilis (EMSA).	(He *et al.*, 2016)
abrC1C2C3	SCO4596	Atypical TCS with two kinase (C1 and C2); response regulator AbrC3 is a transcriptional activator of *act*II-ORF4 (ChIP-chip); impacts Red production.	(Rico *et al.*, 2014a)
lexA	SCO5803	Global regulator of the DNA damage response; Repressor of *act*II-ORF4 (EMSA).	(Iqbal *et al.*, 2012)
	SCO6256	GntR family regulator of antibiotic production. Direct activator of *cdaR* and indirect repressor of Act production (EMSA).	(Yu *et al.*, 2016a)
scbR2	SCO6286	Activator of actII-ORF4, *redD*, *redZ* and *cdaR*, repressor of *cpkO* and SCO6268 (*cpk* cluster) (Chip-seq, EMSA).	(Li *et al.*, 2017b)

Table 2b Recently discovered transcriptional regulators that control antibiotic production in *S. coelicolor.* Orthologues also studied in *S. avermitilis* or *S. venezuelae* are indicated

Gene	ID	Functions(s) of the regulators(s)	Ref.
Regulators in pathway with missing link to antibiotic gene clusters			
ohkA	SCO1596 SAV6741	Orphan HK; plays global role in antibiotic biosynthesis, by influencing precursor supply, pleiotropic and pathway-specific antibiotic regulators.	(Lu *et al.*, 2011)
abrA1A2	SCO1744/5	TCS; represses Act, Red and Cda production and morphological differentiation.	(Rico *et al.*, 2014b)
	SCO2140	Lrp/AsnC family protein. Indirectly regulates ACT and CDA production or cooperate with other transcriptional regulators involved in production of these antibiotics (EMSA).	(Yu *et al.*, 2016b)
aor1	SCO2281	Orphan response regulator; upregulates Act, Red and Cda production and downregulates sigB, thus linking antibiotic production to osmotic stress response.	(Antoraz *et al.*, 2017)
stgR	SCO2964	LTTR; Negative regulator for Act and Red production trough upregulation of *act*II-*ORF4* and *redZ*, respectively. Exact regulatory cascade remains unknown.	(Mao *et al.*, 2013)
sigT	SCO3892	ECF sigma factor; required for normal Act production under nitrogen limitation.	(Feng *et al.*, 2011)
cmdABCDEF	SCO4126–SCO4131	Operon for membrane proteins; affects differentiation and causes increased production of Act.	(Xie *et al.*, 2009)
phoU	SCO4228	Activates Act and Red production. Exact regulatory cascade unknown	(Martin-Martin *et al.*, 2017)

[a]SCO, *S. coelicolor*; SAV, *S. avermitilis*; SVEN, *S. venezuelae*; see StrepDB for the full annotation (http://strebdb.Streptomyces.org.uk). [b]Experimental evidence presented between brackets (EMSA, Electrophoretic shift assay; ChIP-Seq, chromosome immunoprecipitation combined with next generation sequencing).

In addition to AraC and AbsA, several other TCSs regulate secondary metabolism in *S. coelicolor* and other actinobacteria (Rodriguez *et al.*, 2013, Shu *et al.*, 2009, Urem *et al.*, 2016b, Yu *et al.*, 2012). TCSs are the major signal-transduction systems of bacteria and enable them to monitor and adapt to environmental changes (Stock *et al.*, 2000, Whitworth, 2012). Streptomycetes harbour a large number of TCSs, which likely reflects the changing and variable nature of their natural habitats (Bentley *et al.*, 2002, Hutchings *et al.*, 2004, Rodriguez *et al.*, 2013). The PhoRP TCS system is ubiquitous in bacteria and senses phosphate and regulates its assimilation. PhoRP plays a major role in the control of antibiotic production in streptomycetes (Martin & Liras, 2012, Sola-Landa *et al.*, 2005, Sola-Landa *et al.*, 2003). Similar has been found for the AfsQ1/2 TCS, which controls the biosynthesis of Act, Red and Cda in response to nitrogen limitation (Shu *et al.*, 2009) via what appears to be direct interaction with the promoter regions of *act*II-ORF4*, redZ* (which activates *redD*) and *cdaR*, respectively. The AfsQ1/2 TCS is closely related to CseBC, which responds to cell-envelope stress (Hutchings, 2007). Recently, it was shown that the DraRK TCS, which responds to high concentrations of nitrogen (Yu *et al.*, 2012), and the OsdRK TCS, which is oxygen-responsive, are similar in function to the system controlling dormancy in mycobacteria (Daigle *et al.*, 2015, Urem *et al.*, 2016b), and are both required for Act production. Interestingly, in the absence of a functional DraRK system the production of Cpk and Red increases (Yu *et al.*, 2012). The AbsA system has been exploited to improve the chance of success during screening of streptomycetes for new antibiotics by overexpression of the *S. coelicolor* homologue in other streptomycetes; this led among others to the induction of pulvomycin production in *S. flavopersicus* . Cross-talk between the different regulatory networks is discussed in Sections 5 and 6.

Cross-regulation of disparate BGCs by cluster-situated regulators

It is well established that a CSR usually binds to promoter sequence(s) and either activates or represses genes only within its cognate BGC. For examples see Tables 1 and 2. However, this is not strictly true for all CSRs. Recently, the PAS-LuxR family cluster-situated regulator within the candicidin BGC was shown to not only activate 16 out of the 21 genes in the gene cluster, but also to be required for expression of the antimycin BGC (McLean *et al.*, 2016, Zhang *et al.*, 2015b). Thus, antimycin and candicidin biosynthesis are co-ordinately controlled by FscRI in *S. albus* (McLean *et al.*, 2016). A similar observation was made in *S. avermitilis,* where PteF, a member of PAS-LuxR family and cluster-situated activator of the filipin BGC,

was proposed to cross-regulate the production of oligomycin (Vicente *et al.*, 2015). Thus, evidence is accumulating, at least for PAS-LuxR family regulators, that they may not in fact simply be CSRs but act more broadly to co-ordinately control the biosynthesis of multiple compounds. This is likely rooted in the flexible inverted repeat the family of regulators appears to bind to both *in vitro* and *in vivo* (McLean *et al.*, 2016, Santos-Aberturas *et al.*, 2011). It is an obvious and attractive hypothesis that production of secondary metabolites with antimicrobial properties or subsets thereof should be coordinated, so as to maximise any synergistic activity and minimise the development of resistance to the agents produced.

THE IMPACT OF PHOSPHATE AVAILABILITY ON SECONDARY METABOLISM

The impact of phosphate availability on bacterial physiology and gene expression in particular has been intensely studied in *Streptomyces* species and other bacteria (McDowall *et al.*, 1999, Martinez-Castro *et al.*, 2013, Doull & Vining, 1990, Chouayekh & Virolle, 2002). Expression of a suite of genes involved in phosphate management termed the *pho* regulon is controlled by the PhoRP TCS (Fabret *et al.*, 1999, Hutchings *et al.*, 2004, Santos-Beneit, 2015). During phosphate starvation, the membrane-bound sensor kinase, PhoR, undergoes autophosphorylation and transfers its phosphate group to the response regulator, PhoP (Sola-Landa *et al.*, 2003, Fernandez-Martinez *et al.*, 2012) (Figure 1). The phosphorylated form of PhoP (PhoP-P) binds to a well conserved DNA motif called a PHO box and can either activate or repress expression of genes within the *pho* regulon (Sola-Landa *et al.*, 2005). During growth in phosphate replete conditions, PhoR is prevented from phosphorylating PhoP via physical interaction with the phosphate-specific transport (Pst) system, a high-affinity phosphate transport system whose production is activated by PhoR (Diaz *et al.*, 2005, Santos-Beneit *et al.*, 2008, Sola-Landa *et al.*, 2005). This interaction creates a regulatory loop in which the Pst system is produced at a low level during conditions of phosphate sufficiency. When phosphate levels drop, PhoR is released and phosphorylates PhoP, which then activates transcription of genes within the Pst system and the other genes within the pho regulon (Sola-Landa *et al.*, 2005). The precise signal that frees PhoR to phosphorylate PhoP is unknown, but it is known that the switch is reversible.

It has been known for some 15 years that deletion of *phoP* can lead to earlier and increased production of antibiotics (Sola-Landa *et al.*, 2003). This phenomenon was covered in our previous review (van Wezel & McDowall, 2011) and for *S. coelicolor* was rooted in destabilization of a negative regulatory loop involving the AfsKRS system (Horinouchi, 2003, Umeyama *et al.*, 2002). AfsR is a transcription factor related to SARPs that when phosphorylated by AfsK activates transcription of the gene encoding AfsS, a small sigma factor-like protein required for antibiotic biosynthesis in *S. coelicolor* (Santos-Beneit *et al.*, 2009, Hong *et al.*, 1991, Matsumoto *et al.*, 1994, Tanaka *et al.*, 2007). In the proposed regulatory loop, PhoP represses the production of AfsS and AfsR represses the production of PhoRP and the Pts system (Santos-Beneit *et al.*, 2009). However, recently PhoP was shown to in fact be an activator of *afsS* transcription in experiments using a full panel of *phoP, afsR* and *afsR/phoP* mutants and a suite of synthetic promoters engineered to prevent AfsR binding but not PhoP binding (Santos-Beneit *et al.*, 2011). In a revised model, PhoP hinders higher activation of *afsS* transcription by AfsR by outcompeting AfsR for binding to the *afsS* promoter (Figure 1) (Santos-Beneit *et al.*, 2011, Santos-Beneit *et al.*, 2009).

A series of ChIP-Chip experiments were conducted with *S. coelicolor*, which provided genome-wide insight into the role of PhoPR in controlling secondary metabolism (Allenby *et al.*, 2012). These revealed that PhoP serves as a master regulator of secondary metabolism during phosphate starvation, whereby it transiently represses pleotropic activators of antibiotic production and regulators of morphological development, namely *bldA,* which specifies the leucine tRNA corresponding to the rare UUA codon, and *scbAR*, which encodes the γ-butyrolactone regulatory system of *S. coelicolor* that positively influence morphological development, and Act and Red biosynthesis (Takano *et al.*, 2005, Takano *et al.*, 2001). Interestingly, the ScbAR system also indirectly controls the gene expression of *scbR2* whose gene product activates *afsK* expression (Yang *et al.*, 2015), which is the cognate sensor kinase responsible for activating the global regulator of secondary metabolism, AfsR (mentioned above). Thus, although PhoP activates expression of *afsS,* it also indirectly represses transcription of *afsK*, which means AfsR remains unphosphorylated and inactive (Figure 1).

Although there are only a handful of example thus far, it is clear that in addition to controlling pleotropic regulators, PhoP can also act directly upon BGCs. For

example, in *S. coelicolor,* PhoP negatively regulates the biosynthesis of Cda by repressing the *cdaR* gene (Allenby *et al.*, 2012). Interestingly, the inverse seems to be the case for the BGC specifying coelimycin where there are three PHO boxes within the DNA sequence of two structural genes and expression of the gene cluster appears to be PhoP-dependent (Allenby *et al.*, 2012). Direct regulation of biosynthetic pathways by PhoP is not a peculiarity of *S. coelicolor,* as PhoP was recently shown to negatively regulate avermectin biosynthesis by repressing the expression of *aveR,* which encodes a cluster-situated activator (Yang *et al.*, 2015).

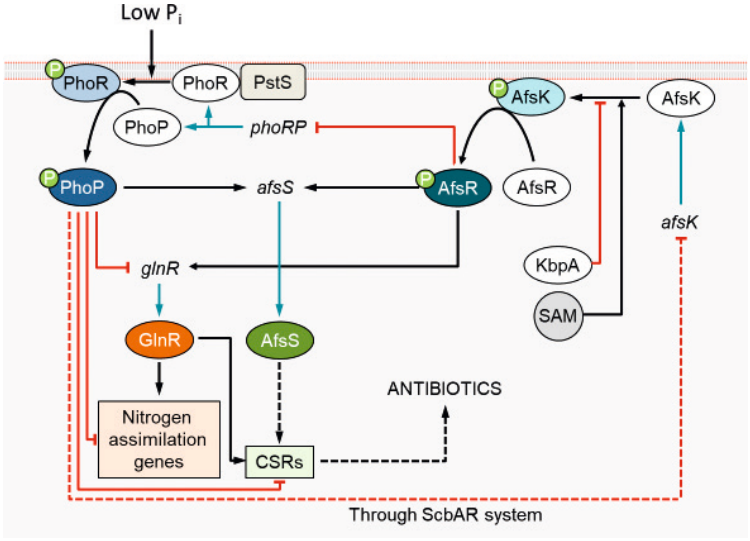


Figure 1 The PhoRP and AfsKRS systems and their interplay in regulation of nitrogen metabolism and antibiotic production.
Black arrows indicate activation and red bars indicate repression, cyan arrows indicate expression of genes. During growth under phosphate deplete conditions, the global regulator PhoP is activated by the membrane-bound sensor kinase, PhoR. Activated PhoP acts directly upon BGCs by modulating expression of CSRs or other transcription factors, such as *glnR*, which controls expression of nitrogen metabolism genes and *afsS*, part of AfsKRS regulatory system. PhoP may directly inhibit expression of nitrogen assimilation genes and has an indirect negative impact (through ScbAR system) on expression of afsK. KbpA and S-adenosyl-L-methionine (SAM) can also modulate the activity of AfsK. The membrane associated kinase, AfsK, in turn, activates AfsR. AfsR interacts with the PhoP in several ways: it can directly repress expression of the *phoRP* regulon, compete for activation of *afsS* or as activator of *glnR* expression can upregulate expression of the genes responsible for nitrogen assimilation.

REGULATIN OF SECONDARY METABOLISM BY NITROGEN

The uptake and incorporation of nitrogen is essential for anabolism of amino acids, nucleic acids and peptidoglycan, among other important macromolecules. *S. coelicolor* can utilise diverse nitrogen sources including ammonia, nitrate, nitrite, urea, amino sugars and amino acids (Reuther & Wohlleben, 2007, Tiffert *et al.*, 2008, Wang & Zhao, 2009). Assimilation of nitrogen results in the production of glutamate and glutamine, which act as the primary nitrogen donors within the cell (Reitzer & Schneider, 2001). Like other bacteria, *Streptomyces* species possess a sophisticated regulatory system that enables adaptation to nitrogen availability. Many studies have indicated that the source of nitrogen can influence the production of secondary metabolites. The production of most of the secondary metabolites is reduced by nitrogen sources that are favourable for growth (Merrick & Edwards, 1995, Sanchez & Demain, 2002). This is presumably because utilization of a high-quality nitrogen source (e.g. ammonium) causes more of the available carbon to be consumed for growth and generation of biomass and thus ultimately less carbon is available for secondary metabolism when starvation occurs. Although the above has been known for a long time, the underpinning molecular detail has taken longer to elucidate. The global regulator controlling nitrogen metabolism is GlnR, which is an orphan response regulator without a cognate sensor kinase (Figure 1) (Tiffert *et al.*, 2008, Wray *et al.*, 1991). Deletion of *glnR* in *S. coelicolor* blocks production of Act and Red (Tiffert *et al.*, 2011). GlnR-mediated regulation of Act and Red production was assumed to be indirect until a recent study demonstrated otherwise. *In vitro* DNA binding and DNaseI footprinting studies showed that GlnR binds the promoter sequence of CSRs within these BGCs (*act*II-*ORF4* and *redZ*, respectively), implying that GlnR regulation is direct (He *et al.*, 2016). In the same study, direct regulation of CSRs of avermectin and oligomycin biosynthesis (*aveR* and *olmRI/RII,* respectively) by GlnR in *S. avermitilis* was also demonstrated; thus, direct regulation of a subset of natural product BGCs by GlnR is likely to be universal (He *et al.*, 2016). Several studies have recently been conducted that have enhanced the understanding of nitrogen metabolism and its interconnectedness with phosphate and carbon utilization. These connections and their implications for secondary metabolism are further discussed in Section 6.

CONTROL OF ANTIBIOTIC PRODUCTION BY THE CARBON SOURCE

Carbon catabolite repression and the control of antibiotic production

In the natural environment, the availability of high-energy carbon sources, for instance, glucose, promotes vegetative growth and suppresses morphological and chemical differentiation (Hostalek, 1980, Sanchez *et al.*, 2010). Examples of antibiotics whose production is repressed by glucose include Act in *S. coelicolor* (Kim *et al.*, 2001, Lee *et al.*, 2009), chloramphenicol in *S. venezuelae* (Bhatnagar *et al.*, 1988), Str in *S. griseus* (Demain & Inamine, 1970), and erythromycin in *Saccharopolyspora erythraea* (Escalante *et al.*, 1982, Bermudez *et al.*, 1998). Like in most bacteria, carbon utilization by streptomycetes is controlled by carbon catabolite repression (CCR), which ensures that high-energy carbon sources such as glucose, fructose or TCA cycle intermediates are utilized preferentially over energetically less favourable ones, such as lactose, glycerol or mannitol. The best studied system is CCR by glucose, which is often referred to as glucose repression (Titgemeyer & Hillen, 2002, Deutscher *et al.*, 2006, Warner & Lolkema, 2003, Goerke & Stulke, 2008).

In most bacteria, glucose is transported through the phosphoenolpyruvate-dependent phosphotransferase system or PTS. The PTS encompasses Enzyme I (EI) and phosphocarrier protein HPr in combination with carbohydrate-specific transport complexes called Enzyme II (EII), which confer substrate specificity (Postma *et al.*, 1993, Saier & Reizer, 1992). As a result, the PTS typically plays a key role in glucose repression (Brückner & Titgemeyer, 2002, Gorke & Stülke, 2008, Gunnewijk *et al.*, 2001). However, in *Streptomyces* species, deletion of either of the genes *ptsH*, *ptsI* or *crr* for HPr, EI and EIIA, respectively, has no influence on CCR, but instead leads to a block in morphological differentiation, with mutants failing to produce aerial hyphae and/or spores on a reference medium such as R2YE agar (Nothaft *et al.*, 2003, Rigali *et al.*, 2006). This sporulation defect is surprising and may be associated with lack of iron and/or copper in this medium, accompanied by a reduced production of the siderophore, desferrioxamine (Lambert *et al.*, 2014, Traxler *et al.*, 2012, Yamanaka *et al.*, 2005). This link between carbon availability, iron homeostasis and morphological differentiation has not yet been resolved. The limited role of the PTS in CCR may be explained by the fact that in streptomycetes, glucose is internalized via the GlcP permease, which belongs to the major facilitator

subfamily of transporters (Perez-Redondo *et al.*, 2010, Romero *et al.*, 2015, van Wezel *et al.*, 2005). For a summary of central carbon metabolism and CCR, see Figure 2.

It was recognized many decades ago that randomly generated mutants lacking CCR are invariably mutated in the gene *glkA,* which encodes a glucose kinase (Hodgson, 1982, Seno & Chater, 1983). Indeed, a targeted deletion of *glkA* in a clean genetic background was pleiotropically defective for CCR (Angell *et al.*, 1994, Angell *et al.*, 1992, Kwakman & Postma, 1994). The activity of Glk is mediated by as of yet unknown mechanism (van Wezel *et al.*, 2007). Its role in catabolite repression may be co-ordinately controlled with a number of other proteins. These include SCO2127, a protein of unknown function, which is encoded by the gene upstream of *glkA* (Chavez *et al.*, 2011, Guzman *et al.*, 2005) and regulatory proteins that control the transcriptional network of genes that mediate CCR, such as the global regulators Rok7B7 and DasR (see below). Another interesting protein is the phosphoinositide phosphatase, SblA (Gagnat *et al.*, 1999). Deletion of *sblA* in *S. lividans* leads to relief of CCR, with accelerated growth and development in the presence of glucose on some media (Chouayekh *et al.*, 2007). These phenotypes correlated with reduced glucose uptake by the mutant and may therefore affect the activity of GlcP. The cleavage of phosphoinositides by SblA is apparently required to resume growth in transition phase, although the mechanism has not been elucidated (Chouayekh *et al.*, 2007).

Studies with *S. peucetius* suggested the existence of an integral regulatory system that responds to glucose transport and metabolism, which probably elicits CCR (Sanchez *et al.*, 2010). Indeed, addition to growth media of either of the glycolytic intermediates fructose 1,6-biphosphate and phosphoenolpyruvate results in glucose repression of daunorubicin and doxorubicin biosynthesis in *S. peucetius* (Ramos *et al.*, 2004). This connects to observations that the activity of GlkA depends on interaction with the glucose permease GlcP in *S. coelicolor* (van Wezel *et al.*, 2007).

Many antibiotics show growth phase-dependent control. As a consequence, developmental mutants that are blocked in an early phase of the life cycle - in particular *bld* mutants - typically fail to produce antibiotics. A well-studied case is represented by mutants that lack the developmental gene, *bldB*, as these

are not only disturbed in development and antibiotic production, but are also defective in CCR (Pope *et al.*, 1998, Pope *et al.*, 1996). This links the pathways that regulate carbon utilization and morphological differentiation. BldB is a member of a family of DNA-binding proteins that are only found in Actinobacteria. The family is widespread in streptomycetes, with several paralogues in *S. coelicolor*, including AbaA and WhiJ, which play a role in the control of antibiotic synthesis and development, respectively (Eccleston *et al.*, 2002). Identification of the BldB regulon and the way its activity is modulated will likely offer important new insights into the growth phase-dependent control of antibiotic production and the role of CCR in this process.

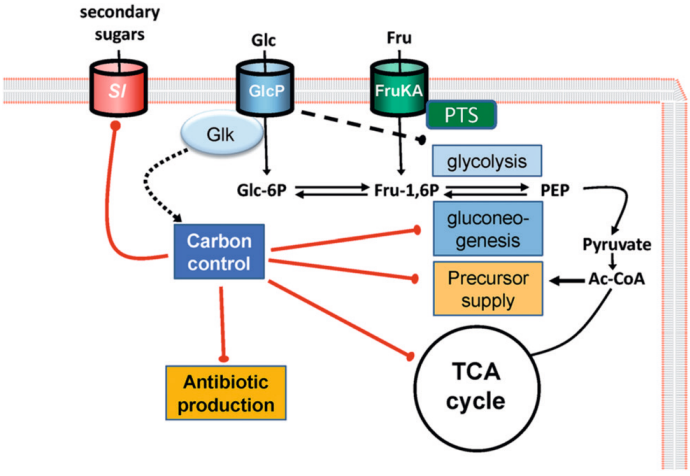


Figure 2 CCR and the control of antibiotic production. Glucose repression is shown for primary and secondary metabolism.
Black arrows indicate activation, red lines repression. Glucose kinase (Glk) is activated post-translationally in a glucose transport-dependent manner (van Wezel *et al.*, 2007). Glc, glucose; Fru, fructose, secondary sugars (energetically less favorable sugars, such as lactose, mannitol and glycerol). SI, substrate induction. Note that glucose is transported by an MFS transporter and not by the PTS in *Streptomyces*.

New insights into the nutrient-sensory DasR system

In streptomycetes, the PTS plays a major role as the first step in a global antibiotic sensory system revolving around the nutrient sensory protein, DasR, which is conserved in streptomycetes and many other actinobacteria. DasR is a GntR-family repressor with a pleiotropic role in the regulation of primary and secondary metabolism and of development. For details, we refer to reviews elsewhere (Urem *et al.*, 2016a, van Wezel & McDowall, 2011) Here we summarise the key elements of

the regulon and highlight recent insights (Figure 3). The core regulon of DasR in all Gram-positive bacteria revolves around the genes for aminosugar transport (*pts*) and metabolism (*nag*) and in streptomycetes also the genes for the chitinolytic system (*chi*). Originally identified as the repressor of the chitobiose transporter DasABC (Colson *et al.*, 2008, Seo *et al.*, 2002), it was soon recognized that DasR also controls many genes involved in antibiotic production. Comprehensive analysis of the DasR regulon of *S. coelicolor* showed that it acts as a direct and very global transcriptional repressor of antibiotic production by binding to the promoter regions of the CSRs for all known chromosomally located antibiotic BCGs in *S. coelicolor* (Nazari *et al.*, 2012, Rigali *et al.*, 2006, Rigali *et al.*, 2008, Swiatek-Polatynska *et al.*, 2015). DasR also represses siderophore biosynthesis via control of the iron-homeostasis regulator *dmdR1* (Lambert *et al.*, 2014, Craig *et al.*, 2012). A similar pleiotropic role of DasR has also been reported in the erythromycin producer *S. erythraea* (Liao *et al.*, 2014b, Liao *et al.*, 2015a), but is not typical of all streptomycetes.

The DNA-binding activity of DasR is modulated by ligands derived from GlcNAc or glucosamine (GlcN), in particular GlcNAc-6P and GlcN-6P, and the crystal structure of DasR and its orthologue NagR of *Bacillus subtilis* in complex with these ligands have been elucidated (Fillenberg *et al.*, 2016, Fillenberg *et al.*, 2015). GlcN-6P stands at the cross-roads of carbon and nitrogen metabolism and cell-wall synthesis, and by acting as an effector of the DasR-dependent antibiotic control system, it plays a major role in the connection between primary and secondary metabolism (Figure 3). The DNA-binding activity of DasR depends on environmental conditions. High concentrations of GlcNAc under famine conditions (*e.g.* on minimal media) result in inactivation of DasR, and thus derepression of its targets, leading to enhanced antibiotic production and development. Conversely, on rich media, GlcNAc represses antibiotic and development, leading to a complete developmental block (Rigali *et al.*, 2006, Rigali *et al.*, 2008, van Wezel *et al.*, 2009). This phenomenon is known as *feast* or *famine*; under conditions of nutritional richness, aminosugars are perceived as derived from chitin, signalling plenty of nutrients, while under poor growth conditions (famine) it is perceived as coming from autolytic degradation of the cell wall and hence cell death. The latter elicits development and antibiotic production. Besides the phosphorylated aminosugars GlcN-6P and GlcNAc-6P, other metabolites may also modulate the DNA-binding activity of DasR. These include high concentrations of phosphate (organic or inorganic), which were shown

to enhance the binding of DasR to its recognition sites (Swiatek-Polatynska *et al.*, 2015, Tenconi *et al.*, 2015). Thus, the affinity of DasR for its recognition sites (and with that the expression of its regulon, including many BGCs for natural products) depends on the metabolic status of the cell. Interestingly, high concentrations of phosphate (either organic or inorganic) enhance binding of DasR to its recognition site *in vitro*, which reinforces the PhoP-mediated repression of antibiotic production by phosphate (Swiatek-Polatynska *et al.*, 2015, Tenconi *et al.*, 2015).

Full genome-scale identification of the DasR binding sites *in vivo* using ChIP-chip analysis corroborated the identity of canonical DasR binding sites or *dre* (DasR-responsive elements), but also revealed so-called class II sites, which do not conform to the known consensus sequence (Swiatek-Polatynska *et al.*, 2015). These sites are not found by the regulon prediction algorithm PREDetector (Hiard *et al.*, 2007). Binding of DasR to class II sites may require a co-repressor, which has not yet been identified. The ChIP-Chip analysis also showed that the binding profile of DasR changes dramatically over time, with only small overlap in the binding profiles between 24 (vegetative growth) and 54 hours (morphological differentiation and antibiotic production). Thus, the DasR regulon is a highly complex system, which is influenced by metabolic status and most likely also by other regulatory proteins. Taken together, the metabolic status of the cell determines the selectivity of DasR for its recognition sites and thus the expression of its regulon, which includes many secondary metabolite BGCs.

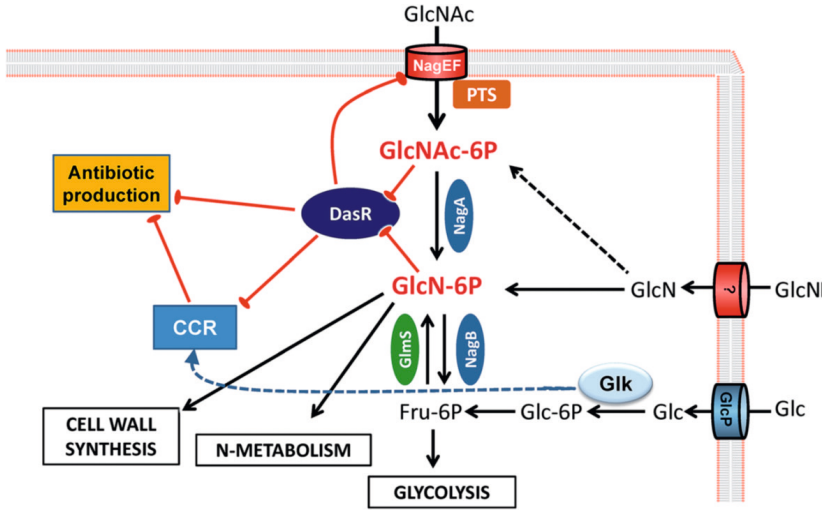


Figure 3 The DasR regulatory network.
The primary metabolism of *S. coelicolor* is shown for N-acetylglucosamine (GlcNAc), glucose (Glc) and glucosamine (GlcN). Glucosamine 6-phosphate (GlcN-6P) is a central metabolite that stands at the crossroads of aminosugar metabolism, glycolysis, nitrogen metabolism and cell wall synthesis. GlcN-6P and GlcNAc-6P are ligands that modulate the DNA-binding activity of DasR. DasR is a global repressor of specialised metabolism. Internalised glucose is phosphorylated by glucose kinase (Glk), which is key to carbon catabolite repression in *S. coelicolor*. In turn, DasR suppresses CCR by downregulating Glk expression. The broken lines represent known routes that have not yet been fully characterised.

Competition between AtrA, Rok7B7 and DasR and connections to CCR

Until the discovery of DasR, it was unclear how global carbon control was related to the control of specific carbon utilization regulons and antibiotic biosynthetic genes. Deletion of the genes for either GylR or MalR relieves both CCR and substrate induction of glycerol and maltose utilization, respectively, and hence gives constitutive expression even in the absence of inducer, while over-expression results in hyperrepression (Hindle & Smith, 1994, van Wezel *et al.*, 1997). This suggests that a global regulatory system for carbon utilization does not exist in *S. coelicolor*. In most bacteria, global carbon control depends on the cAMP receptor protein (CRP). Streptomycetes do have a cAMP receptor protein, but in contrast to other bacteria, it does not seem to play a role in CCR. Instead, CRP plays a role in the control of germination, and *crp* null mutants show prolongued dormancy (Piette *et al.*, 2005). Importantly, genome-wide DNA binding studies and transcriptional analysis revealed that CRP also globally controls antibiotic BGCs in *S. coelicolor* ((Gao *et al.*, 2012); see also section 6).

There is also growing evidence that besides DasR, the TetR-family regulator AtrA plays a role in carbon utilization (Figure 4). Very recent ChIP-seq experiments (McDowall *et al.*, unpubl. data) have confirmed that AtrA binds upstream of *nagE2*, which encodes a known permease for the uptake of GlcNAc (Nothaft *et al.*, 2010). Similar to what was found for *act*II-ORF4, this binding appears to activate transcription as disruption of *atrA* results in reduced levels of *nagE2* transcript (Nothaft *et al*, 2010). This led to the suggestion that AtrA may increase Act production indirectly through enhanced GlcNAc-induced inactivation of DasR as well as directly through activation of *act*II-ORF4 transcription (Nothaft *et al.*, 2010). The control of DasR activity by AtrA via cellular levels of GlcNAc may extend beyond *nagE2* as recent ChIP-seq also identified AtrA binding to recognisable motifs upstream of SCO0481, which encodes a protein that binds chitin (a rich source of GlcNAc), and *crr* (SCO1390), for the global PTS component EIIA, that is required for GlcNAc transport. The role of AtrA in carbon utilisation almost certainly extends beyond GlcNAc metabolism (Figure 4). ChIP-seq also identified AtrA binding to sites upstream of *gylR* (SCO1658) and *glpk2* (SCO0509), which encodes a glycerol kinase outside the *gyl* operon. Control of morphological differentiation by AtrA is explained at least in part by transcriptional control of *ssgR* (Figure 4) (Kim *et al.*, 2015b), the transcriptional activator of the gene encoding SsgA, which is involved in cell division and sporulation (Traag *et al.*, 2004, Traag & van Wezel, 2008). Disruption of *atrA* suggests it activates transcription of *ssgR* (Kim *et al.*, 2015b), and direct binding of AtrA within the upstream regulatory region of *ssgR* was confirmed by ChIP-seq (McDowall *et al*, unpubl. data).

The ROK-family protein, Rok7B7 takes up an interesting position in the regulatory network as it connects the control of antibiotic production and carbon catabolite repression (Swiatek *et al.*, 2013). Mutants lacking *rok7B7* are delayed in their developmental programme and are pleiotropically disturbed in terms of antibiotic production, perhaps as a consequence of a yet unexplained change in CCR. Rok7B7 activates the transcription of *act*II-ORF4 (and hence Act production) and represses the biosynthesis of Red and Cda, although its binding site has so far not been identified (Park *et al.*, 2009, Swiatek *et al.*, 2013). Aside from *act*II-ORF4, Rok7B7 also activates the GlcNAc *pts* gene, *nagE2,* which means it counteracts the activity of DasR in a manner very similar to AtrA.

The signals that are required for activation of AtrA and Rok7B7 are unknown. Since AtrA is a TetR-regulator it is suggested that this protein is regulated in an allosteric manner by a ligand to exert its effect on secondary metabolism. In *S. globisporus*, AtrA is inhibited by the binding of heptaene, a biosynthetic intermediate of lidamycin whose biosynthesis is controlled by AtrA via activation of its CSR (Li *et al.*, 2015). As part of this work, it was also reported that the DNA-binding activity of *S. coelicolor* AtrA is regulated by Act (Li *et al.*, 2015). Whilst this finding was shown with different preparations of Act, the specificity of this effect needs to be evaluated further. To our knowledge, in all streptomycetes *atrA* is co-located with a divergent AtrA-target gene (SCO4119 in *S. coelicolor*) that encodes NADH dehydrogenase (Ahn *et al.*, 2012). There is interest in identifying the substrate of SCO4119 as at least some members of the TetR family interact with ligands that are structurally identical or related to the substrates of proteins encoded by genes divergent to their own (Cuthbertson & Nodwell, 2013). As ChIP-chip experiments failed to show binding of ROK7B7 to genomic DNA under standard growth conditions on minimal media, it was proposed that the regulator requires a co-factor or ligand to facilitate its DNA binding activity. The control of - and gene synteny with - the xylose transport operon x*ylEFG* by Rok7B7 hints at C5-sugars as candidate ligands for this regulator(Swiatek *et al.*, 2013).

Interestingly, there is an intricate link between Rok7B7, DasR and CCR, which in turn has important implications for the control of antibiotic production. Proteomic comparison of *S. coelicolor* and a *glkA* null mutant showed that glucose activates the expression of Rok7B7 in a Glk-independent manner (Gubbens *et al.*, 2012), which was later confirmed by transcriptomic analysis (Romero-Rodríguez *et al.*, 2016). In turn, DasR and Rok7B7 repress the expression of *glkA* and thus CCR (Swiatek *et al.*, 2013, Swiatek-Polatynska *et al.*, 2015), while conversely, Glk represses Rok7B7 (Gubbens *et al.*, 2012). Deletion of *rok7B7* results in a loss of CCR, which directly implicates Rok7B7 in CCR (Gubbens *et al.*, 2012, Romero-Rodríguez *et al.*, 2016). It is unlikely however that *glkA* is a member of the *rok7B7* regulon, as *glkA* transcription is constitutive, and its activity is post-translationally controlled (van Wezel *et al.*, 2007, Romero-Rodríguez *et al.*, 2016).

In summary for this chapter, there are multiple regulatory networks that connect carbon control to the control of antibiotic production. Understanding carbon source-dependent control of antibiotic production is important from the

perspective of both the design of growth media for yield optimization and for screening of new bioactive molecules. Despite the wealth of literature, it is still unclear how Glk exerts CCR, and we expect that more regulatory proteins that play a role in this important process will be discovered. It is becoming clear that there is a strong connection to the regulons of DasR, Rok7B7 and AtrA. Future research will need to elucidate precisely how this multi-layer control network is governed. Finding the ligands for AtrA and Rok7B7 would be one of the major steps to take.

CONNECTIONS BETWEEN PHOSPHATE, NITROGEN AND CARBON METABOLISM

Carbon, nitrogen and phosphate are essential components for the basic building blocks of all cellular life. It is reasonable to assume that acquisition and utilization of these elements would be coordinately controlled. Although widely accepted, molecular characterization of this interconnectivity has only emerged recently, with the important discovery that GlnR, DasR and CRP jointly regulate three genes for citrate synthesis in the erythromycin producer *S. erythraea* (Liao *et al.*, 2014a). CRP controls early processes during growth in *Streptomyces* species (Derouaux *et al.*, 2004, Piette *et al.*, 2005) and acts as a global regulator of Act, Cda and Red production, perhaps by coordinating precursor flux (Gao *et al.*, 2012). Indeed, 8 out of 22 secondary metabolic clusters on within the *S. coelicolor* genome harbour Crp binding sites, suggesting a pleiotropic role in control of antibiotic production. Further evidence for the connection between C- and N-metabolism via GlnR came from elegant experiments showing that several ABC transporter systems are under direct control of GlnR in *S. erythraea*, affecting growth on maltose, mannitol, mannose, sorbitol and trehalose (Liao *et al.*, 2015b). Recent data show that in *S. coelicolor,* GlnR is activated by glucose (Romero-Rodríguez *et al.*, 2016), while GlnR directly activates transcription of a putative carbohydrate transport operon *agl3EFG* (Cen *et al.*, 2016). Taken together, these data suggest direct linkage between carbon and nitrogen metabolism, albeit perhaps only when certain carbon sources are available.

The understanding of links between nitrogen and phosphate metabolism in *S. coelicolor* is better developed. PhoP and GlnR control antibiotic production in response to the availability of phosphate and nitrogen sources, respectively (Santos-Beneit *et al.*, 2009, Santos-Beneit *et al.*, 2012). Similar to the competitive

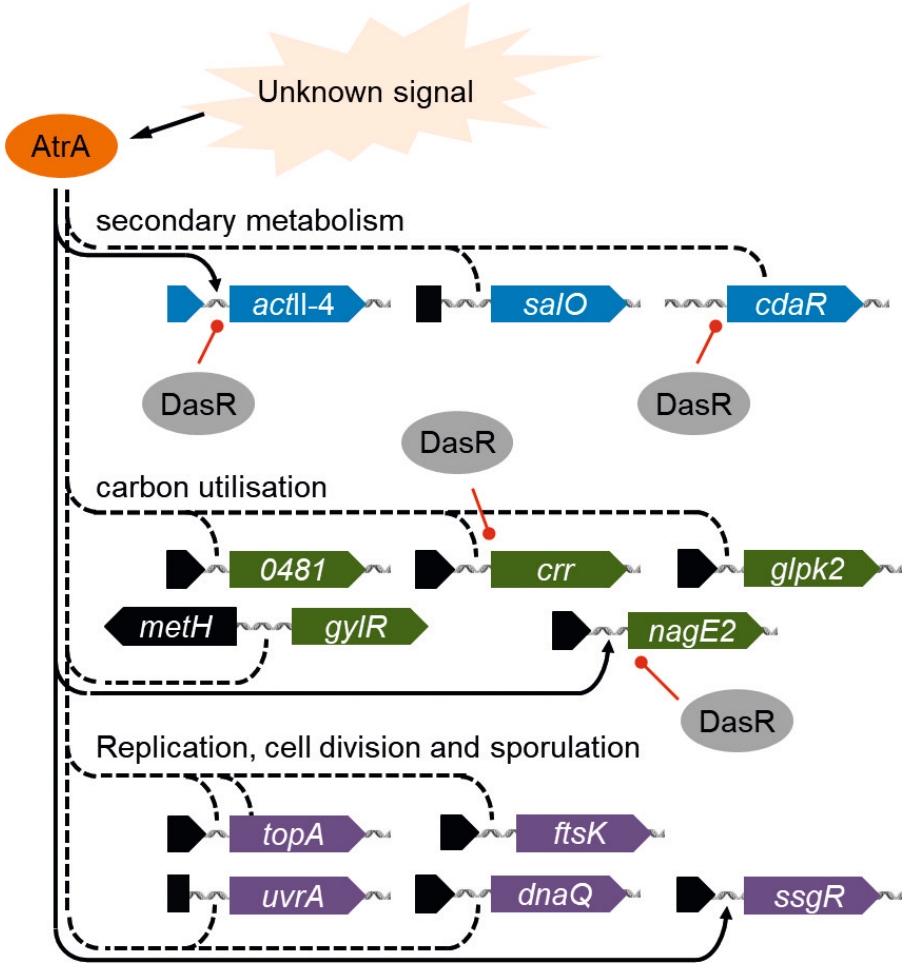


Figure 4 Schematic illustration of a selection of genes corresponding to sites of AtrA binding in S. *coelicolor*.

Black and red solid black lines with arrow heads represent previously described interactions associated with activation by AtrA and repression by DasR, respectively. The broken lines represent interactions identified by chromatin immunoprecipitation but not yet characterized AtrA binds to upstream regions of genes encoding CSRs (*act*II-ORF4, cdaR of S. *coelicolor* and *salO* of S. *albus*; the latter encodes the CSR for salinomycin biosynthesis). The activator AtrA and the repressor DasR compete for binding to the upstream regions of *act*II-ORF4 and *cdaR* and upstream regions of genes that are involved in the uptake of GlcNAc (*crr* and *nagE2*). In addition, AtrA binds to an upstream region of SCO0481, which encodes a protein that binds chitin, a rich source of GlcNac. The positive control of AtrA on GlcNac uptake suggest that AtrA increases Act production indirectly through enhanced GlcNAc-induced inactivation of DasR as well as directly through activation of *act*II-ORF4 transcription. AtrA also binds to upstream regions of genes involved in glycerol catabolism (*gylR* and *glpk2* (SCO1658)). The binding of AtrA to the upstream region of genes involved in DNA replication (*topA*, DNA topoisomerase 1, *uvrA*, *dnaQ*) cell division and sporulation (*ssgR* and *ftsK*) explains the role of AtrA in the control of morphological development.

activation of *afsS* by AfsR and PhoP described in section 3, these two regulators bind to overlapping regions within the *glnR* promoter, but unlike the *afsS* story, PhoP represses *glnR* transcription while only AfsR promotes it (Santos-Beneit *et al.*, 2011) (Figure 1). When phosphate is plentiful, PhoP is inactive and thus AfsR (dependent on the growth phase) activates transcription of *glnR*, but when phosphate is in short supply, PhoP is phosphorylated by PhoR and represses the expression of *glnR* (Figure 1) (Santos-Beneit *et al.*, 2012). In addition, PhoP also directly represses transcription of genes within the GlnR regulon, namely two glutamine synthetases (*glnA* and *glnII*) and the promoter for the *amtB-glnK-glnD* operon, which encodes an ammonium transporter and putative nitrogen sensing/regulatory proteins (Sola-Landa *et al.*, 2013). Uptake/utilization of nitrogen is presumably superfluous if insufficient phosphate is available, hence the PhoP-mediated repression of genes involved in these processes. Thus, PhoP-mediated control of nitrogen metabolism may help balancing the cellular P/N equilibrium.

Connection between phosphate and carbon metabolism is less well studied, but one link may be governed via the PhoP-controlled enzyme PPK (polyphosphate kinase), which affects antibiotic production in response to the level of inorganic phosphate (Pi) (Chouayekh & Virolle, 2002, Ghorbel *et al.*, 2006). PPK is involved in maintaining the cellular energy balance by regenerating ATP from ADP and polyphosphates and *ppk* mutants show enhanced Act production under Pi-limited growth conditions (Chouayekh & Virolle, 2002). This was recently explained by increased degradation of triacylglycerols (TAGs), resulting in accumulation of the polyketide precursor acetyl-CoA (Le Marechal *et al.*, 2013). Additionally, phospho-sugars inhibit antibiotic production in streptomycetes. This effect is mediated by the phosphate- rather than of the glyco-moiety, as the inactivation of *phoP* or *ppk* prevents or enhances, respectively, their utilization as nutrient sources and their inhibitory effect on antibiotic production(Tenconi *et al.*, 2012).

Thus, it is becoming evident that the conventional understanding of the PhoRP, AfsR and GlnR as the elements of the linear transduction systems regulating primary and secondary metabolism have been revised significantly over the last several years. Recent discoveries made it possible to understand, at least partially, the cross-talk occurring between regulators for phosphate and nitrogen metabolism, and to a lesser extent carbon metabolism in streptomycetes. It is a reasonable expectation to predict that established methods for assessing DNA

binding *in vivo* (i.e. ChIP-seq (Robertson *et al.*, 2007) in combination with new strategies for robustly mutagenizing and identifying mutants (i.e. Tn-Seq (van Opijnen *et al.*, 2009) will enhance the ability to probe these regulons and their cross regulation.

THE IMPACT OF METALS ON SECONDARY METABOLISM

Iron is an essential metal that plays important roles in DNA replication, protein synthesis and respiration. Iron is relatively unavailable in the soil due to the low solubility of the Fe^{3+} ion under aerobic conditions at neutral pH. Production of iron-chelating compounds called siderophores is the most common way that bacteria circumvent this problem (Guerinot, 1994). Moreover, some bacteria have developed systems that allow them to utilize siderophores synthesised by neighbouring microorganisms (Arias *et al.*, 2015, Galet *et al.*, 2015, Traxler *et al.*, 2012). The primary impact of iron deficiency in *Streptomyces* and other bacteria, is the stimulation of siderophore production. All *Streptomyces* species examined thus far appear to harbour a BGC for desferrioxamine, which has been proposed to be part of the 'core' secondary metabolome of the genus (Seipke, 2015), while other streptomycetes produce additional siderophores; *S. coelicolor* and *S. scabies* produce coelichelin and pyochelin, for example (Lautru *et al.*, 2005, Seipke *et al.*, 2011). Production of desferrioxamine is normally repressed by the DmdR1 protein, which becomes derepressed in the absence of iron (Flores *et al.*, 2005, Flores & Martin, 2004, Tunca *et al.*, 2007). The *dmdR1* gene is unusual in that its DNA sequence encodes a second gene (*adm*) using the anti-sense strand of DNA (Tunca *et al.*, 2009). Deletion of the *dmdR1-amd* locus in *S. coelicolor* abolished sporulation and the production of Act and Red (Flores *et al.*, 2005). Subsequent experimentation whereby either *dmdR1* or *amd* were individually mutated by a point mutation revealed that inactivation of *dmdR1* had no impact on Act and Red production where as these compounds were overproduced when only *amd* was mutated (Tunca *et al.*, 2009). Another link between iron availability and secondary metabolism in *S. coelicolor* is that iron de-represses the pleiotropic TCS, AbrA1/A2, which negatively regulates Act and Red production, although the mechanism has not yet been resolved (Rico *et al.*, 2014b).

Zinc is an important transition metal required as a cofactor for many enzymes and regulatory proteins important for normal bacteria physiology. However, the

intracellular free level of this element should be maintained within a narrow range due to its potential toxicity (Finney & O'Halloran, 2003, Reyes-Caballero *et al.*, 2011). Its uptake in streptomycetes as well as in other bacteria is regulated by Zur, a zinc-responsive transcriptional regulator (Panina *et al.*, 2003, Shin *et al.*, 2007). Interestingly, there is a Zur-binding site within the BGC for the metal chelator, coelibactin and adjacent to this is a binding site for another zinc-sensitive regulator, AbsC; together these regulators repress coelibactin biosynthesis (Hesketh *et al.*, 2009). Interestingly, AbsC also seems to be required for the production of Act and Red when *S. coelicolor* is cultivated under the specific conditions of zinc limitation and inactivation of *zur* and *absC* genes block sporulation. Binding of AtrA upstream of the promoter for *zur* (Romero *et al.*, 2014) has been identified both biochemically and by ChIP-seq (McDowall *et al*, unpubl. data) suggesting yet another layer of regulation that potential facilitates integration with primary metabolism as well as secondary metabolism and morphological development. More detailed study of these regulators is necessary in order to fully illuminate their regulons and the nature in which they overlap and interconnect with other metal acquisition systems. *Amycolatopsis japonicum* produces the biodegradable ethylenediamine-tetra acetate (EDTA) isomer [S,S]-EDDS, whose gene cluster was elucidated (Spohn *et al.*, 2016). Trace amounts of zinc in the culture media inhibit the production of [S,S]-EDDS, which led to the proposal that the molecule is required for zinc uptake. The synthesis of the zincophore is repressed by the zinc regulator Zur (Spohn *et al.*, 2016).

Recently, the impact of rare earth elements (REEs) on secondary metabolism was explored. Supplementation of culture medium with scandium or lanthanum stimulated the production Act by *S. coelicolor*, Str by *S. griseus* and actinomycin by *S. antibioticus* (Kawai *et al.*, 2007). Although precise mechanistic detail is lacking, scandium stimulation of Act production is dependent on the ppGpp synthetase, RelA and is mediated by upregulation of *act*II-ORF4 (Kawai *et al.*, 2007). Interestingly, scandium was also able to rescue the ability of *S. lividans* to produce Act, a compound that the species does not normally produce despite harbouring a nearly identical gene cluster (Kawai *et al.*, 2007). Quantitative RT-PCR and HPLC analyses showed that in addition to Act, scandium supplementation stimulated the expression of eight other BGCs in *S. coelicolor* (Tanaka *et al.*, 2010). Stimulation of secondary metabolism by REEs is not restricted to actinobacteria – scandium was recently shown to elicit the production of amylase and bacilysin in *B. subtilis*

(Inaoka & Ochi, 2011). Thus, REEs represent a relatively unexplored method for activating the expression of silent or weakly expressed BGCs and future studies should be aimed at understanding the molecular mechanism(s) by which this occurs.

MORPHOLOGICAL DEVELOPMENTAL CONTROL OF ANTIBIOTIC PRODUCTION

As mentioned in the introduction to this review, the production of antibiotics (and other secondary metabolites) is temporally correlated to the onset of development of *Streptomyces* colonies (Bibb, 2005, van Wezel & McDowall, 2011). A model of the linkage between the control of antibiotic production and development is presented in Figure 5. A likely explanation is that the colony is particularly vulnerable to competitors when it is undergoing programmed cell death (PCD), and antibiotics are produced to protect the colony and the nutrients released during PCD. Until recently, the occurrence of PCD in bacteria has been a subject to major debate, but it is becoming increasingly clear that PCD plays a major role the life cycle of multicellular bacteria (Bayles, 2014, Claessen *et al.*, 2014, Rice & Bayles, 2003, Rosenberg, 2009), and in that of streptomycetes in particular (Manteca *et al.*, 2005, Miguelez *et al.*, 2000). A direct link between PCD and antibiotic production was demonstrated with the discovery that GlcNAc, which together with N-acetylmuramic acid forms the peptidoglycan strands, acts as an elicitor of antibiotic production via metabolic inactivation of the global antibiotic repressor DasR (Rigali *et al.*, 2006, Rigali *et al.*, 2008). For details we refer to section 5. Interestingly, production of prodiginines, which have anticancer activity by degrading the DNA, may play a direct role in triggering PCD in *S. coelicolor*, and mutants that fail to produce prodiginines have strongly reduced PCD, whereby vegetative growth is prolongued (Tenconi *et al.*, 2020).

As a consequence of the growth phase-dependent control of antibiotic production, developmental mutants that are blocked in an early phase of the life cycle - in particular *bld* mutants - typically fail to produce antibiotics. As mentioned in Section 5.1, mutants of the developmental gene *bldB* are not only disturbed in development and antibiotic production, but are also defective in CCR (Pope *et al.*, 1998, Pope *et al.*, 1996). This links the pathways that regulate carbon utilization and morphological differentiation. BldB is a member of a family of DNA-binding proteins

that are only found in Actinobacteria. The family is widespread in streptomycetes, with several paralogues in *S. coelicolor*, including AbaA and WhiJ, which play a role in the control of antibiotic synthesis and development, respectively (Eccleston *et al.*, 2002). Identification of the BldB regulon and the way its activity is modulated will likely offer important new insights into the growth phase-dependent control of antibiotic production and the role of CCR in this process.

BldD is a small DNA-binding protein that is required for development and antibiotic production (Figure 5) (Elliot *et al.*, 1998). BldD is related to SinR, a master regulator of the transition from the motile to a sessile state in *Bacillus subtilis*, and hence associated with the control of biofilm formation (Gaur *et al.*, 1991, Kearns *et al.*, 2005). The BldD regulon encompasses over 150 transcriptional units, many of which are involved in the control of development (den Hengst *et al.*, 2010). One of its targets is *bldA*, which at least in part explains the requirement of BldD for antibiotic production. BldD binds to DNA as a homodimer, and dimerization is dependent on the binding of a tetramer of the signalling molecule cyclic-di-GMP (Tschowri *et al.*, 2014). This is another interesting example of small molecule-based control of antibiotic production in *Streptomyces*.

Other *bld* mutants also fail to produce antibiotics, but the phenotype of these mutants is not independent of the growth medium (Figure 5). In fact, *bldA, bldC. bldG, bldH* (*adpA*)*, bldJ* and *bldK* mutants produce spores on non-repressing carbon sources such as mannitol or glycerol, but not on media containing glucose. Interestingly, mutation of *glkA* restores antibiotic production and morphological development to *bldA* mutants(van Wezel & McDowall, 2011), while *bldJ* and *bldK* mutants are rescued by supplementing the colonies with iron. The latter is due to their failure to produce the siderophore desferrioxamine (Lambert *et al.*, 2014). In fact, most *bld* mutants are affected in desferrioxamine biosynthesis, with strongly reduced production of the siderophore in *bldA, bldJ,* and *ptsH* mutants, and overproduction in *bldF, bldK, crr* and *ptsI* mutants (Lambert *et al.*, 2014).

An infamous example of translational control of development and antibiotic production is BldA, the tRNA that recognizes the rare UUA codon for leucine. Mutants of *S. coelicolor* defective in *bldA* have a bald phenotype and fail to produce antibiotics (Lawlor *et al.*, 1987, Leskiw *et al.*, 1991). The latter is a direct consequence of the presence of UUA codons in the mRNA of the genes for ActII-

ORF4 and RedZ (Guthrie *et al.*, 1998, White & Bibb, 1997). The presence of TTA codons in BGCs for specialized metabolites - and in particular in genes encoding CSRs - is more a rule than an exception, which provides strong phylogenetic evidence for the fact that control of antibiotic production by BldA has evolved with a purpose (Chandra & Chater, 2008).

Mutants that are blocked in sporulation (so-called *whi* mutants) generally are not affected in antibiotic production. This is most likely because the decisions to switch on secondary metabolism made at an earlier stage in the life cycle. The exception is *ssgA*, whose transcription does not depend on any of the 'classical' *whi* genes (Traag *et al.*, 2004). SsgA activates sporulation-specific cell division by controlling the localization of its paralogue SsgB, which in turn recruits FtsZ to initiate sporulation-specific cell division (Figure 5) (Willemse *et al.*, 2011). In contrast to most developmental control proteins, SsgA and SsgB lack DNA-binding domains. The SsgA-like proteins are unique to sporulating actinobacteria, and most likely function as chaperones that recruit multi-component complexes (Girard *et al.*, 2013, Noens *et al.*, 2007). Over-expression of *ssgA* results in overproduction of prodiginines (Red), while Act production is blocked (van Wezel *et al.*, 2006, van Wezel *et al.*, 2000a). The most likely explanation is that SsgA blocks *S. coelicolor* development at a stage corresponding to early aerial growth, where Red production has been switched on, while Act production has not yet been initiated. SsgA and SsgB probably represent another important link in the coordination of secondary metabolite production with vegetative growth (van Dissel *et al.*, 2014).

WblA is a member of the WhiB-like proteins, and 11 paralogues are encoded by the *S. coelicolor* chromosome (Fowler-Goldsworthy *et al.*, 2011). The Wbl proteins are small iron-sulphur proteins that are unique to actinobacteria. Disruption of *wblA* has a highly pleiotropic effect on overall gene expression in *S. coelicolor* and prevents development while strongly increasing antibiotic production in this organism (Fowler-Goldsworthy *et al.*, 2011). Conversely, overproduction of WblA pleiotropically represses the biosynthesis of Act, Red and Cda in *S. coelicolor* and of anthracyclines in *S. peucetius* (Kang *et al.*, 2007). Deleting *wblA* also results in enhanced production of specialized metabolites in other streptomycetes, such as *Streptomyces ansochromogenes, Streptomyces glaucescens, Streptomyces roseosporus* and *Streptomyces sp.* C4412 as well as in *Pseudonocardia* (Huang *et al.*, 2016, Huang *et al.*, 2017, Kim *et al.*, 2014, Nah *et al.*, 2012, Noh *et al.*, 2010,

Rabyk *et al.*, 2011), and should therefore be considered as a general approach to achieve enhanced production of cryptic antibiotics in a given strain. It is yet unclear how WblA controls antibiotic production.

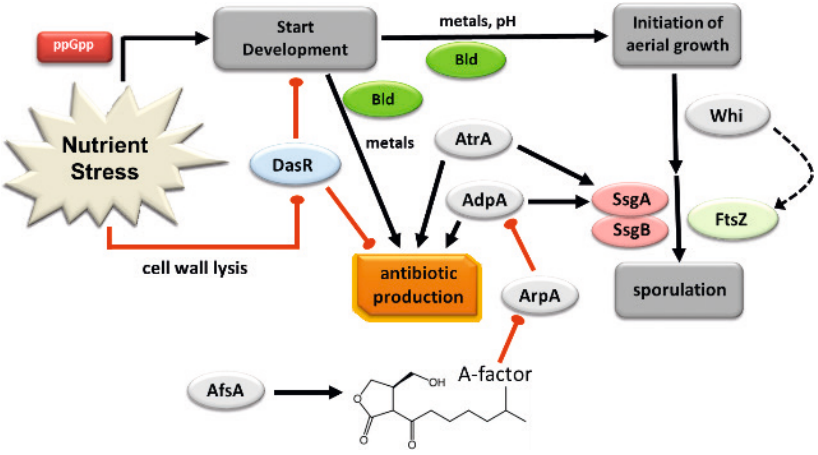


Figure 5 Initiation of development and antibiotic production.
The developmental programme starts with nutrient stress and growth cessation, followed by the accumulation of ppGpp. The autolytic dismantling of the cell wall (PCD) releases cell wall-derived metabolites that inhibit the activity of the nutrient sensory DasR. The onset of antibiotic production correlates temporally to the transition from vegetative to aerial growth, and is controlled by multiple pathway-specific and global regulators. Shown here are three key pleiotropic regulators, namely the antibiotic repressor DasR which responds to phosphorylated aminosugars likely derived from PCD, the activator AtrA (signal unknown) and AdpA, which responds to the accumulation of A-factor (synthesized by AfsA). Bld proteins and environmental signals control the procession towards aerial growth and antibiotic production. Whi proteins control aerial growth. Eventually, FtsZ accumulates and localizes to septum sites in an SsgAB dependent manner. Solid black arrows represent major transitions in development. The arrow indicates the FtsZ accumulation checkpoint controlled by the Whi proteins. Red lines indicate repression.

AUTOREGULATORS AND THE CONTROL OF ANTIBIOTIC PRODUCTION

Bacteria communicate with each other through production of small extracellular molecules, called bacterial hormones or autoregulators. After the discovery of the gamma-butyrolactone A-factor (2-isocapryloyl-3*R*-hydroxymethyl-γ-butyrolactone), produced by *S. griseus*, many more bacterial hormones have been identified, such as GBLs similar to A-factor, furans, gamma-butenolides and PI-factor. In general, these signalling molecules are active in nanomolar concentrations and diffuse readily from one actinomycete to another, thereby

affecting development and antibiotic production. GBL production is most likely not species-specific, as different species can produce the same GBL, suggesting extensive interspecies communication and 'eavesdropping'. Antibiotics may also function as signalling molecules, thereby induce antibiotic activity and/or resistance, and again in a more general fashion, affecting a broad range of hosts. Thus, the usage of bacterial hormones or antibiotics is an important factor in the discovery of novel antibiotics, as well as co-culturing micro-organisms (recently reviewed in (Niu *et al.*, 2016)).

The gamma-butyrolactone regulatory system in *S. coelicolor* and *S. avermitilis*

Enzymes responsible for the synthesis of gamma-butyrolactones (GBLs) in streptomycetes are identifiable through their homology to the A-factor synthetase AfsA of *S. griseus* (Kato *et al.*, 2007). The orthologue of AfsA is encoded by *scbA* (SCO6266) within the *cpk* gene cluster responsible for the production of the yellow compound coelimycin P1 (Xu *et al.*, 2010). ScbA is required for the production of the GBLs of *S. coelicolor.* This strain produces 8 different GBLs (SCB1-8). The structure of these molecules have recently been solved after they were overproduced in the super host M1152 (Sidda *et al.*, 2016). Deletion of *scbA* resulted in the overproduction of Act and Red biosynthesis and reduced *cpk* expression (D'Alia *et al.*, 2011) Divergent to *scbA* lies *scbR* (SCO6265), which encodes a transcription factor that appears to activate transcription of *scbA* as well as a repressor of its own transcription and that of *cpkO* (*kasO*), which encodes the CSR of the coelimycin BGC cluster, provided GBL is not bound by ScbR (Takano *et al.*, 2001), (Takano *et al.*, 2005). It also positively regulates CdaR, the CSR of the Cda BGC. Deletion of *scbR* resulted in reduced Act, Red and Cda production and increased coelimycin P1 production (Yang *et al.*, 2015). The regulation of *scbA* is complex, with no fewer than five *scbR* paralogues in *S. coelicolor* (Niu *et al.*, 2016), one of which *scbR2* (SCO6286) is also encoded within the coelimycin BGC (Gottelt *et al.*, 2010). The reader is referred to our previous review for more details (van Wezel & McDowall, 2011).

ScbR2 is highly similar to ScbR, but unlike ScbR it is not able to bind GBLs, and is hence considered a pseudo gamma-butyrolactone receptor (Xu *et al.*, 2010, Wang *et al.*, 2011). Instead, it binds the endogenous antibiotics Act and Red and the exogenous antibiotic jadomycin B and related angucyclines (Xu *et al.*, 2010,

Wang *et al.*, 2014b). Interestingly, addition of non-endogenous jadomycin B from *S. venezuelae* releases ScbR2 from the promoters of *redD* and *adpA* in *S. coelicolor*, leading to accelerated Red production and morphological differentiation. ScbR2 probably has a greater effect on secondary metabolism than ScbR. Deletion of *scbR2* abolishes Act, Red and Cda production and induced coelimycin production (Gottelt *et al.*, 2010, Wang *et al.*, 2014b). Like ScbR, ScbR2 directly represses *cpkO* (Xu *et al.*, 2010). ScbR2 is also a repressor of *scbA*, and acts both directly and indirectly on antibiotic production (Wang *et al.*, 2011). ChIP-seq showed that ScbR and ScbR2 have many shared targets genes related to primary and secondary metabolism (Yang *et al.*, 2015, Li *et al.*, 2016). Both directly act on *afsK* and on genes involved in malonyl-CoA synthesis and hence precursor supply for polyketide natural products. Interestingly, the TetR-like proteins ScbR and ScbR2 can also bind as heterodimers, and co-immunoprecipitation of ScbR2 and ScbR revealed that only the ScbR-ScbR2 heterodimer can control SCO5158, which encodes an uncharacterized protein (Li *et al.*, 2017b). Such heterodimer formation is not unique, and was previously proposed for the gene products of *mmfR* and *mmyR* of the methylenomycin BGC (O'Rourke *et al.*, 2009).

S. avermitilis contains three GBL-like receptors encoded by genes that are located in a single locus, namely *aveR1*, *aveR2* and *aveR3*. This locus also contains the genes *aco* and *cyp17* required for avenolide biosynthesis. The bacterial hormone avenolide increases avermectin production in a dose-dependent manner when added in nanomolar concentrations to an *aco* deletion mutant (Kitani *et al.*, 2011). The AveR1 protein was identified as its cognate receptor (Wang *et al.*, 2014a). Deletion of *aveR1* or addition of avenolide did not influence avermectin production, but increased avenolide production. An explanation for the latter might be that the threshold that is required for avermectin production has already been reached at the start of growth. This led to the suggestion that AveR1 acts as a repressor in the early stages of growth (Sultan *et al.*, 2016). AverR1 represses its own transcription and that of *aco* (Sultan *et al.*, 2016).

AveR2 is a pseudo GBL-receptor that represses the transcription of *aveR*, encoding the positive CSR of the *ave* cluster (Zhu *et al.*, 2016). Additionally, AveR2 represses *aco* and *cyp17*, and controls genes involved in primary metabolism, ribosomal protein synthesis and stress responses. Such an extended regulon is reminiscent of ScbR2 (see above), and it is important to note that both regulators can bind

endogenous and exogenous antibiotics. Indeed, the affinity of AveR2 for DNA is influenced by avermectins and also by the exogenous antibiotics jadomycin B and by aminoglycosides. Thus, we note that such pseudo-GBL receptors should be considered as important pleiotropic regulators (Zhu *et al.*, 2016).

AveR3 shows similarity to autoregulator receptors and activates *aveR* transcription of the avermectin BGC, and indirectly also filipin biosynthesis (Suroto *et al.*, 2017, Miyamoto *et al.*, 2011). Interestingly, deletion of *aveR3* resulted in the discovery of the cryptic natural product, phthoxazolin A, a cellulose synthesis inhibitor that shows activity against plant pathogenic oomycetes. The fact that GBL-mediated regulatory systems control cryptic genes in both *S. coelicolor* and *S. avermitilis* makes them candidate targets for drug discovery.

GBL-receptors and antibiotic production in other streptomycetes

The examples of *S. coelicolor* and *S. avermitilis* suggest that the presence of genes for GBLs and their receptor proteins may serve as beacons for cryptic BCGs. Similarly, the BGCs for the angucyclines jadomycin B (from *S. venezuelae*) and auricin (from *S. aureofaciens*) and also contain genes for GBL synthases and their cognate receptors (Mingyar *et al.*, 2015), (Zou *et al.*, 2014). The gene *jadR3* harboured within the jadomycin B BGC encodes a putative GBL receptor located upstream of the GBL synthase genes *jadW123*. The product of this GBL synthase system is SVB1, which is identical to the GBL SCB3, produced by *S. coelicolor*. *In* S. venezuelae, *only* JadW2 is required for jadomycin production (Zou *et al.*, 2014). Nevertheless, deletion of *jadW1* abolishes both jadomycin B and chloramphenicol production under conditions that are known to be favourable for production of these antibiotics (Wang & Vining, 2003). JadR3 is an autorepressor and also represses *jadW1* transcription, and thereby represses jadomycin B production (Zou *et al.*, 2014). The auricin BGC of *S. aureofaciens* is controlled by the GBL synthase SagA and its cognate receptor SagR, and again the genes encoding these proteins are located directly next to the biosynthetic genes. Deletion of *sagR* results in early but reduced auricin production, while deletion of *sagA* abolishes auricin production, establishing their key role in controlling auricin biosynthesis. In contrast to other GBL receptor proteins, SagR does not auto-regulate its own transcription, but instead *sagR* and *sagA* are repressed by the CSR Aur1R (Mingyar *et al.*, 2015).

Further on the theme, the production of indigoidine (a blue-pigmented compound), of nucleoside antibiotics (showdomycin and minimycin) and of D-cycloserine by *S. lavendulae* FRI-5 is controlled by the bacterial hormone IM-2 and its cognate receptor FarA (Kitani *et al.*, 2010, Kitani *et al.*, 2001). Supplementation of culture media with IM-2 enhances production of indigoidine, but abolishes production of D-cycloserine (Kitani *et al.*, 2010). FarA inhibits its own expression and activates the expression of FarX, the protein required for IM-2 biosynthesis. The genes encoding FarA and FarX are located on a regulatory island spanning 12.1 kb (Kitani *et al.*, 2008). This island contains the genes *farA-E*, *farR1-5* and *farX* (Kitani *et al.*, 2008). FarA negatively regulates its own expression and the expression of *farR1* (which encodes an orphan response regulator), *farR2* (for a pseudo-GBL receptor), *farR4* (for a SARP regulator) (Kurniawan *et al.*, 2014), *farB* (for a structural protein) (Kitani *et al.*, 2008). Since *farR3* and *farR4* can be transcribed both as monocistronic and bicistronic mRNA, it appears that *farR3* is also a target of FarA (Kurniawan *et al.*, 2014). FarR2 is a pseudo-GBL receptor that positively regulates the production of indigoidine, but negatively regulates the expression of the *far* regulatory genes in the regulatory island, including the expression of *farX* (Kurniawan *et al.*, 2016). Similarly, FarR3 positively regulates the production of indigoidine (Kurniawan *et al.*, 2014), but in both cases the control is most likely indirect (Pait *et al.*, 2017, Kurniawan *et al.*, 2016). The SARP regulator FarR4 represses IM2 biosynthesis (Kurniawan *et al.*, 2014). which offers a unique example of a SARP regulator that acts at the front instead of the end of a regulatory cascade (Kurniawan *et al.*, 2014).

The complex regulatory network of the "pristinamycin supercluster" of *S. pristinaespiralis* is also under the control of a GBL-receptor. Pristinamycin is a mixture of two compounds, including the cyclohexanedepsipeptide pristinamycin I (PI) and the poly-unsaturated macrolactone pristinamycin II (PII) that are produced in a 30:70 ratio. The mixture of pristinamycin is significantly more active against pathogenic bacteria than PI and PII separately (Mast & Wohlleben, 2014). PI is synthesized by non-ribosomal peptide synthetases (NRPS) and PII by hybrid polyketide synthases (PKS)/NRPS (Mast *et al.*, 2011). The genes required for PI and PII production are not arranged in a single BGC, but are heterogeneously divided over a 210 kb genomic region whereby the biosynthetic genes are interspersed by a cryptic BGC (Mast *et al.*, 2011). These characteristics of the BGC and the fact that the cluster contains seven genes encoding CSRs makes the regulation of pristinamycin biosynthesis very complex (Mast *et al.*, 2015). These CSRs include

the GBL-receptor SpbR, two TetR-like regulators (PapR3 and PapR5), three SARP regulators (PapRI, PapR2, PapR4) and a response regulator (PapR6) (Mast *et al.*, 2015, Mast *et al.*, 2011). The regulatory cascade starts with the release of SpbR from the DNA when its ligand reaches a critical concentration (Mast *et al.*, 2015). The pristinamycin BGC is under the direct control of the SARP regulators PapR1, PapR2 and the response regulator PapR6 (Mast *et al.*, 2015). PapR2 is most likely the master regulator of the pristinamycin BGC, as this is the only regulator that is fully required for pristinamycin biosynthesis (Mast *et al.*, 2015). The regulatory genes that directly control the pristinamycin BGC are repressed by the TetR-regulator PapR5 (Mast *et al.*, 2015),(Dun *et al.*, 2015). PapR5 shows similarity to pseudo-GBL receptors, suggesting that perhaps pristinamycin and/or biosynthetic intermediates act as ligands for PapR5 and may thereby control the level of pristinamycin (Mast *et al.*, 2015). Similar as to other regulatory networks, the GBL-receptor is not the first regulator in the regulatory cascade, since SpbR is positively regulated by an AtrA (SSDG_00466) regulator outside the BGC. AtrA in turn positively controls the transcription of PapR5 (Dun *et al.*, 2015). Thus, the pristinamycin BGC is subject to complex and multi-level control, several elements of which deserve further investigation, so as to unravel the full regulatory network.

EMERGING THEMES IN THE CONTROL OF ANTIBIOTIC PRODUCTION IN ACTINOBACTERIA

Besides the usual suspects, less well-studied genera of Actinobacteria (often referred to as rare Actinobacteria) also produce a wide range of natural products, and insights into their molecular regulation is important from the perspective of drug discovery and production improvement. Culture collections housed by biotechnology companies and research institutes possess several rare Actinobacteria, including *Micromonosporaceae, Streptosporangiae*, *Pseudonocardiaceae*, *Nocardiaceae*, and *Thermomonosporaceae,* and many other rare and unclassified species that have yet to be explored (Genilloud *et al.*, 2011, Monciardini *et al.*, 2014, Yan *et al.*, 2016, Fenical & Jensen, 2006). In recent years, interest in strains isolated from marine environments and other ecological niches such as plants and insects has grown because they offer a rich new microbial source for NP discovery (Freel *et al.*, 2012, van der Meij *et al.*, 2017, Kamjam *et al.*, 2017). The regulation of natural product biosynthesis by rare Actinobacteria is poorly characterised, because many of them are genetically intractable and limited

genetic tools are available. As the cell wall structure between Actinobacteria often varies and is different from that of streptomycetes, preparation of protoplasts (and regeneration) typically requires different methods (Marcone *et al.*, 2010). A protocol to prepare protoplasts of *Planobispora rosea*, the producer of the thiazolyl peptide antibiotic GE2270 that targets elongation factor EF-Tu (Anborgh & Parmeggiani, 1991) was applied to different rare Actinobacteria(Marcone *et al.*, 2010). This protocol demonstrated the applicability of both lysozyme and mutanolysin (from *S. globisporus*) to produce protoplasts from these industrially important strains (Marcone *et al.*, 2010). Other issues that need to be solved for genetic manipulation of rare actinobacteria include identification of suitable origins of replication for plasmids (Dairi *et al.*, 1999), the methylation pattern of the DNA (Suzuki *et al.*, 2011, Flett *et al.*, 1997) and the use of specific promoters for expression (Horbal *et al.*, 2013, Bai *et al.*, 2015). Many of these technical difficulties can in principle be circumvented by the use of expression of a BGC in a heterologous host. Expression of the BGC for GE2270 of *P. rosea* in *S. coelicolor* M1146 allowed the study of its regulation (Flinspach *et al.*, 2014). Deletion of *pbtR* for a TetR-family regulator abolished the production of GE2270. Similarly, the BGC for taromycin A from *Saccharomonospora sp.* CNQ490 was also expressed in *S. coelicolor* M1146 to allow its genetic manipulation. Deletion of *tar20*, encoding a LuxR regulator of the taromycin BGC, increased the production of the compound in the heterologous strain (Yamanaka *et al.*, 2014). Heterologous expression of a BGC may often be suitable to study the function of CSRs within a BGC, but for understanding of the global regulatory network and the ecological responses that control the BGC of interest, it is necessary to study the BGC in its natural host. In a number of Actinobacteria, the molecular regulation of antibiotic production has been studied. Especially in strains that produce clinically important antibiotics, such as glycopeptide producers. It appears that the rare Actinobacteria that have been studied indeed contain similar regulators as *Streptomyces* and therefore we expect that most of the control mechanisms of antibiotic production are similar. Below the control of antibiotic production in a number of Actinobacteria is discussed and compared to that of *Streptomyces*.

Control of glycopeptide biosynthesis

The glycopeptide antibiotics vancomycin and teicoplanin are important last line of defence antibiotics that are used to treat infections associated with multi-drug resistant Gram-positive bacteria (Wood, 1996) (Sosio & Donadio, 2006).

Their target is the peptidoglycan precursor lipid II, thereby inhibiting synthesis of the bacterial cell wall (Barna & Williams, 1984). Vancomycin is produced by *Amycolatopsis orientalis* and teicoplanin by *Actinoplanes teichomyceticus* (Sosio *et al.*, 2004, Li *et al.*, 2004). Other well-studied members include the precursor of dalbavancin, A40926 produced by *Nonomuraea sp.* ATCC39727 (Sosio *et al.*, 2003), balhimycin produced by *Amycolatopsis balhimycina* (Pelzer *et al.*, 1999), and the sugarless glycopeptide A47934 produced by *S. toyocaensis* (Pootoolal *et al.*, 2002). A comparison of the BGCs for these compounds (*tei* for teicoplanin, *bal* for balhimycin and *dbv* for A40926) and their control is presented in Figure 6. Members of the glycopeptides share a heptapeptide core, which is synthesized by non-ribosomal peptide synthetases (NRPS), with further modifications such as cross-linking, methylation, halogenation glycosylation or attachment of sulphur groups (Donadio *et al.*, 2005, Sosio & Donadio, 2006). Glycopeptides bind to the D-alanyl-D-alanine(D-ala-D-ala) terminus of the growing lipid attached peptidoglycan chain on the outside of the cytoplasmic membrane and thereby prevent the binding of transpeptidases that create the cross-links between the polysaccharides, required for cell wall integrity (Barna & Williams, 1984).

The BGCs of these antibiotics are typically controlled by CSRs of the StrR and LuxR families (Lo Grasso *et al.*, 2015, Shawky *et al.*, 2007, Horbal *et al.*, 2014). The teicoplanin BGC spans 89 kb and includes five regulatory genes, *tei2*, *tei3*, *tei15*, tei16** and *tei31** (Li *et al.*, 2004, Sosio *et al.*, 2004). Tei2 and Tei3 show high homology with the VanR/VanS system of *S. coelicolor* (Hong *et al.*, 2004, Hutchings *et al.*, 2006) and are involved in the control of teicoplanin resistance. The genes *tei15** and *tei16** encode members of the StrR and LuxR family regulators, respectively. Overexpression of Tei15* results in 30-40-fold increase in teicoplanin biosynthesis (Horbal *et al.*, 2014, Horbal *et al.*, 2012). Tei15* is the primary CSR, and directly controls the transcription of the regulatory genes *teiA* for the NRPS module, *tei2** (which encodes a deacetylase), *tei16**, *tei17** involved in Dpg synthesis and *tei27** (for an unknown protein). Tei15* also controls the expression of the LuxR family regulator Tei16* and the SARP family regulator Tei31*. The targets of Tei16* and Tei31* in the teicoplanin cluster remain unknown, although Tei16* does positively control teicoplanin production (Horbal *et al.*, 2014). Tei15* does not show autoregulation, in contrast to its orthologue BbR in the balhimycin BGC (Horbal *et al.*, 2014, Shawky *et al.*, 2007). See Figure 6.

The dalbavancin BGC of *Nonomuraea* sp. ATCC39727 contains four regulatory genes, namely *dbv3*, *dbv4*, and the TCS *dbv6* and *dbv22* for the control of resistance (Figure 6). Dbv4 (similar to StrR and Tei15*) is the likely CSR, and is expressed under phosphate-limiting conditions, while Dbv3 is a LuxR-type regulator similar to Tei16*. Both Dbv3 and Dbv4 are required for A40926 production (Lo Grasso *et al.*, 2015). Dbv3 controls the transcription of *dbv4*, as well as genes for the biosynthesis of 4-hydroxyphenylglycine, the heptapeptide backbone, and for glycosylation and export. However, similar to the situation for Tei16* in the teicoplanin BGC, no common regulatory elements were identified in the promoter regions of the Dbv3-controlled genes, and control could therefore be indirect (Lo Grasso *et al.*, 2015). Dbv4 is directly involved in the regulation of genes involved in 3,5-dihydroxyphenylglycine, cross-linking, halogenation, glycosylation and acylation (Lo Grasso *et al.*, 2015). Dbv4 and the Dbv4 regulon are repressed by phosphate, whereas Dbv3 and its regulon are not. No Pho-boxes were identified upstream of the *dbv4* genes, suggesting the phosphate repression is indirect (Alduina *et al.*, 2007).

The glycopeptide balhimycin is produced by *Amycolatopsis balhimycina* (formerly *Amycolatopsis mediterranei*). The balhimycin BGC has a simpler control system with three regulatory genes, namely the VanR/VanS TCS for resistance and the StrR-like regulator Bbr (Figure 6). Bbr binds to a consensus sequence ($GTCCAR(N)_{17}TTGGAC$) that is found within the promoter for its own transcription, the putative ABC transporter gene *tba*, *oxyA* for a P450 monooxygenase, *dvaA* involved in dehydrovancosamine synthesis and the putative sodium proton antiporter gene orf7 (Shawky *et al.*, 2007). In the three glycopeptide BGCs the StrR CSR binds to the consensus sequence that is conserved in the intergenic regions of the glycopeptide BGCs, although the target sequence may vary and deviate from the consensus (Alduina *et al.*, 2007, Donadio *et al.*, 2005, Horbal *et al.*, 2014, Shawky *et al.*, 2007). Although these three BGCs are organised in a similar manner and contain regulatory genes, the mechanism of regulation differs between them, and therefore making assumptions about the regulatory network based on bioinformatics alone is not sufficient (Lo Grasso *et al.*, 2015). In *S. griseus,* StrR is positively controlled by the pleiotropic regulator AdpA. However, overexpression of the putative *adpA* gene of *A. balhimycina* did not induce antibiotic production, although heterologous expression of this regulator in *S. coelicolor*, *S. ghanaensis* and several soil Actinobacteria was successful (Ostash *et al.*, 2015). Vancomycin

biosynthesis and its control are well understood, but the role of StrR regulator in the BGC (AORI_1475) has not been elucidated.

Since most glycopeptide BGCs contain a StrR-like positive regulator, over-expression of the corresponding gene is a logical generic strategy to induce the expression of (cryptic) glycopeptide BGCs. A good example is the production of ristomycin A in *Amycolatopsis japonicum*. This strain is known for the production of (*S*,*S*)-ethylenediamine disuccinic acid [(*S*,*S*)-EDDS], the biodegradable isoform of EDTA (section 7). Under standard laboratory conditions this strain does not produce antibiotics, but over-expression of the StrR orthologue in *A. japonicum* induced the production of ristomycin A, which is used for the diagnosis of von Willebrand disease and Bernard-Soulier syndrome (Spohn *et al.*, 2014).

Control of glycopeptide resistance

Bacteria that are resistant against glycopeptide antibiotics replace the D-alanine for D-lactate as the terminal residue of the peptide chain of the peptidoglycan. As the affinity of the glycopeptide for the latter is a lot lower than for D-ala-D-ala, binding of the glycopeptide is prevented (Arthur *et al.*, 1996, Bugg *et al.*, 1991). The glycopeptide BGCs contain genes that encode homologues of the VanR/VanS TCS that governs glycopeptide resistance.

S. coelicolor is resistant against vancomycin and this resistance is conferred by genes that are similar to the ones present in vancomycin resistant enterococci (Hong *et al.*, 2004, Hutchings *et al.*, 2006). The resistance cluster of *S. coelicolor* is organized in four transcription units, namely *vanRS*, *vanJ*, *vanK* and *vanHAX*. The latter encode the enzymes required for biosynthesis and incorporation of D-lac in the peptide moiety of the PG. All transcription units are regulated by VanRS (Hong *et al.*, 2004). Binding of vancomycin by the N-terminal part of VanS leads to its autophosphorylation, and this phosphate is then transferred to the N-terminal receiver domain of VanR, thereby activating its C-terminal DNA binding effector domain. This results in expression of the resistance genes. In the absence of vancomycin VanS acts a phosphatase that dephosphorylates VanR, and hence *vanS* mutants show constitutive expression of vancomycin resistance (Hutchings *et al.*, 2006, van der Aart *et al.*, 2016). In contrast, deletion of *vanS* in *S. toyocaensis* results in sensitivity to A47934, and it was suggested that VanR of *S. coelicolor* is phosphorylated by other proteins while that of *S. toyocaensis* is not

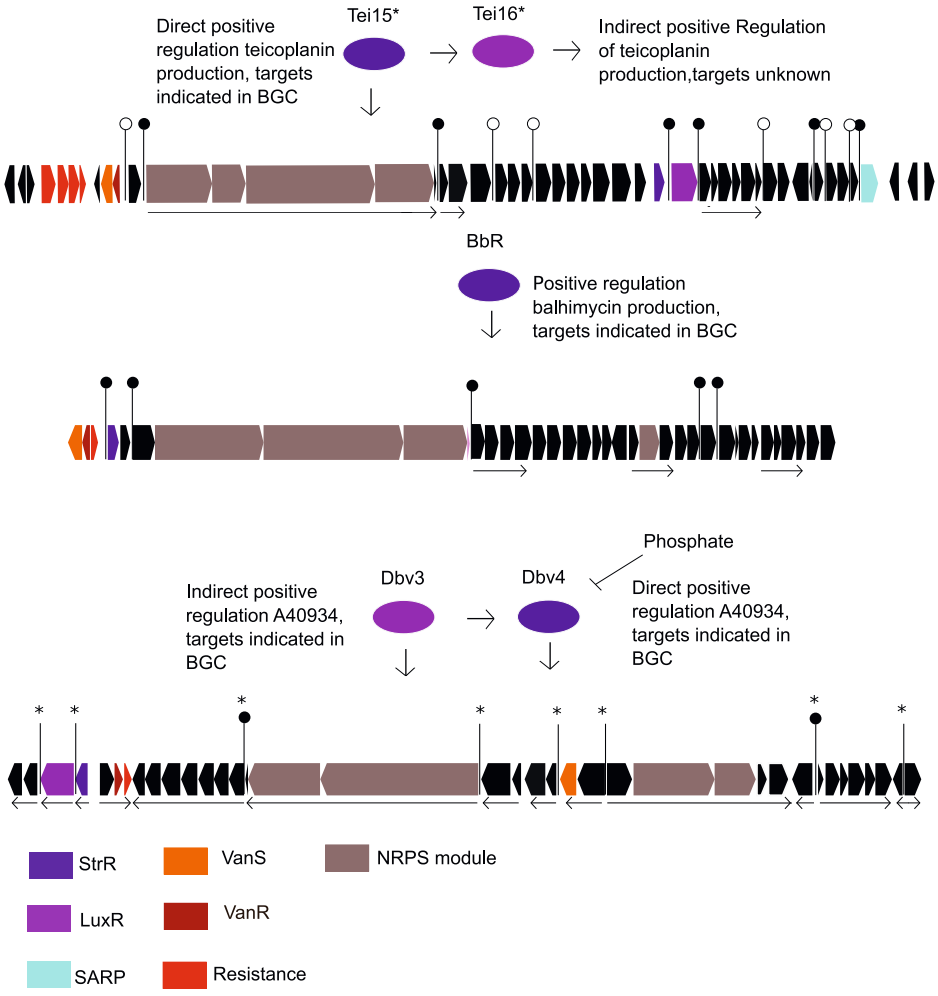


Figure 6 Regulation of glycopeptide biosynthetic gene clusters.
Shown are the BGCs for teicoplanin (*tei*), balhimycin (*bal*) and A40926 (*dbv*). Known and putative binding sites for StrR (purple) are indicated in the clusters with closed and open circles, respectively. The consensus sequence for the StrR binding sites GTCCAR(N)17TTGGAC is shared between all three BGCs. Genes regulated by LuxR (magenta) are indicated with an asterisk. Experimentally confirmed operons are indicated with an arrow. The primary CSR of the teicoplanin BGC is Tei15*, which positively regulates the expression of LuxR-family regulator Tei16* and of the SARP-family regulator Tei31*, with both regulators having unknown targets. The *bal* cluster is regulated by the CSR BbR, and lacks a gene for a LuxR regulator. The primary CSR of the dbv cluster is the LuxR regulator Dbv3, which positively regulates the expression of StrR regulator Dbv4, most likely indirectly. For details see the text. BGCs adapted from the MIBiG database.

(Novotna *et al.*, 2015). Interestingly, the VanRS TCS is an important determinant of the species-specific glycopeptide resistance profile. *S. coelicolor* is resistant against vancomycin and A47934, but sensitive to teicoplanin, while *S. toyocaensis* is only resistant against A47934 (Abrudan *et al.*, 2015). Exchanging the VanRS TCSs between the two *Streptomyces* strains is sufficient to switch the resistance profile (Abrudan *et al.*, 2015). Surprisingly, expression of the VanR orthologue of *A. balhimycina* (VnlR) in *S. coelicolo*r even governed resistance to teicoplanin, and led to increased actinorhodin biosynthesis (Kilian *et al.*, 2016). VnlR controls *vanHAX* in *S. coelicolor*, despite the fact that it does not control *vanHAX* in *A. balhimycina* itself (Kilian *et al.*, 2016).

σ-factor/anti-σ-factor systems and the control of antibiotic biosynthesis

An important new element of antibiotic control that was discovered in recent years is the control by σ-factors, the subunits of the RNA polymerase responsible for promoter recognition. An important example is that of the control of lantibiotics. Lantibiotics are ribosomally synthesized, post translationally modified peptide antibiotics (RiPPs; (Arnison *et al.*, 2013)). The best known lantibiotic is the food-preservative nisin, produced by *Lactococcus lactis* and discovered as early as 1928 (Rogers & Whittier, 1928). Lantibiotics are synthesized as a prepropeptide encoded by a precursor gene generally referred to as *lanA*. This propeptide is post-translationally modified via intramolecular lanthionine bridges that are formed between unusual amino acids to yield the mature peptide (Willey & van der Donk, 2007). Nisin and several other lantibiotics target the pyrophosphate linkage component of the cell-wall precursor lipid II. As this target is different from that of the clinically used antibiotic vancomycin, there is no cross-resistance with glycopeptides, making them interesting new antibiotics for the treatment of methicillin resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant enterococci (VRE) (Munch *et al.*, 2014). Screening a library of 120,000 chemical extracts derived from 40,000 Actinobacteria for activity against cell-wall biosynthesis by Vicuron Pharmaceuticals identified five novel lantibiotics, including microbisporicin (also known and NAI-107) and planosporicin, produced by *Microbispora corallina and Planobispora alba,* respectively (Castiglione *et al.*, 2007, Castiglione *et al.*, 2008). The control of the BGCs for microbisporicin (*mib* in *M. corallina* and *mlb* in *M.* ATCC-PTA-5024) and for planosporicin (*psp*) have been studied in detail (Fernandez-Martinez *et al.*, 2015, Sherwood & Bibb, 2013, Foulston & Bibb, 2011). The BGCs have a gene for an extracytoplasmic function

(ECF) σ-factor /anti- σ-factor complex (MibX/MibW for microbisporicin and PspX/PspW for planosporicin). ECF σ factors mediate responses to extracellular signals and stress or steps in morphological differentiation (Helmann, 2002, Paget *et al.*, 2002), but their involvement in the control of antibiotic production was only recognized recently. The microbisporicin and planosporicin BGCs also contain a gene for a regulator with a LuxR-like C-terminal domain. Herein, we use microbisporicin biosynthesis as the example for both BGCs, see Figure 7 for an overview of its control. The BGC is controlled by its own production by a feed-forward mechanism: deletion of *mibA* results in decreased transcription of the other *mib* genes, while growth of *mibA* mutant colonies adjacent to wild-type microbisporicin-producing colonies restored *mib* transcription (Foulston & Bibb, 2011, Foulston & Bibb, 2010, Sherwood & Bibb, 2013). This effect is specific, since microbisporicin cannot induce the production of planosporicin by *Planobispora alba* (Sherwood & Bibb, 2013). The *mib* cluster includes six transcription units, for synthesis, modification, proteolysis, export, immunity and regulation, and all except the *mibA* structural gene contain the ECF σ-factor promoter motif (GACC-N15-GCTAC) that is recognized by MibX (Foulston & Bibb, 2010, Fernandez-Martinez *et al.*, 2015, Foulston & Bibb, 2011) (Figure 7). The promoter of *mibA* is controlled by MibR; in turn, transcription of *mibR* depends on MibX and is enhanced by the stringent response. Indeed, deletion of *relA* in *M. corallina* abolishes microbisporicin production. Thus, a complex regulatory network ensures the correct timing of microbisporic biosynthesis, which is induced by both nitrogen starvation and the ensuing stringent response, which activates MibR expression and hence the expression of the (non-toxic) precursor peptide. This precursor is then exported and processed to yield the active antibiotic (Fernandez-Martinez *et al.*, 2015). Under repressing conditions, MibX is recruited by the membrane bound anti sigma factor MibW, thereby shutting down the biosynthetic pathway. Microbisporicin production also directly depends on the developmental programme, with reduced expression in *bld* mutants, similarly to the biosynthesis of the morphogenic lantibiotic-like morphogen SapB in *S. coelicolor* (Gallo *et al.*, 2016). For a detailed overview on the regulation of RiPPs in Actinobacteria and other bacterial genera, we refer the reader a recent review (Bartholomae *et al.*, 2017).

Involvement of *σ* factors in the control of antibiotic production is not exclusive to lantibiotic BGCs. SigT regulates Act production in *S. coelicolor* via *relA* in response to nitrogen starvation, which links nitrogen stress to secondary

metabolism (Feng *et al.*, 2011). In *S. albus*, the ECF σ^{AntA} controls the synthesis of the antimycin precursor, 3-formamidosalicylate (Seipke *et al.*, 2014, Seipke & Hutchings, 2013), and σ^{25} differentially controls the biosynthesis of oligomycin and of the important anti-helminthic drug avermectin in *S. avermitilis* (Luo *et al.*, 2014). Antimycin is a mitochondrial cytochrome c reductase inhibitor produced by diverse actinobacteria. σ^{AntA} was the first example of regulation of antibiotic production by a cluster-situated ECF σ factor in *S.* species and it was recently shown that this is likely to be a conserved strategy of regulation for more than 70 antimycin BGCs (Joynt & Seipke, 2018). Unlike other ECFs, which are controlled by an anti-σ factor that is unable to maintain an inactive complex in the presence of cognate stimulus, σ^{AntA} is an orphan and is not controlled by such a factor. Instead, evidence to date suggests that σ^{AntA} is controlled by Clp proteolysis (Seipke *et al.*, 2014). The involvement of σ-factor genes in the control of antibiotic production is a new concept, and in particular the presence of σ factor genes within BGCs may function as beacons to identify BGCs in genome mining.

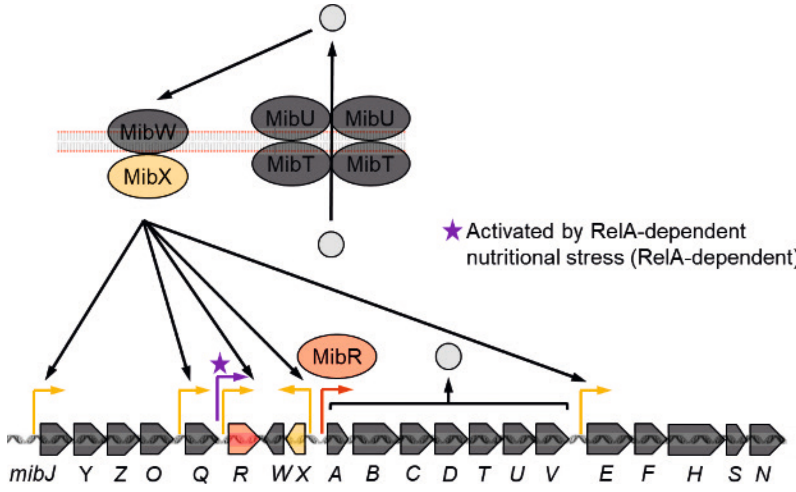


Figure 7 The regulation of microbisporicin production by *Microbispora corallina*. Nutritional stress leads to the RelA-dependent production of ppGpp which results in the expression of the LuxR-family regulator MibR. MibR activates the expression of *mibABCDTUV*, which results in the production of an immature and less active form of microbisporicin (grey circle) and the means for its export. A basal level of expression of the genes encoding an ECF s-factor (MibX)/anti-s-factor (MibW) system enables a feed-forward regulatory mechanism. The immature compound itself or possibly interaction with its lipid II to be sensed by MibW, at which point the ECF s-factor, MibX is released. MibX then in turn activates its own expression and that of *mibR* as well as the remaining genes in the BGC.

Regulation of antibiotic production in *Salinispora*

Recently, studies have also been dedicated to the regulatory network of natural product biosynthesis in the marine actinomycete *Salinispora*. *Salinispora* is an obligate marine actinomycete and most of the isolates are derived from marine sediments. The genus knows three different species, under which *S. pacifica*, *S. tropica* and *S. arenicola*. The compounds that were discovered from this genus are predominantly new and therefore this genus is a good example of the concept that new genera derived from remote areas are a good source for the discovery of novel natural products (Jensen *et al.*, 2015). One of these studies reveals that in *S. tropica* CNB-440, a LuxR-type regulator positively regulates the biosynthesis of the important natural product salinisporamide A, a proteasome inhibitor that is in stage 1 of clinical trials of anti-cancer treatment. This regulator controls the genes involved in the biosynthesis of the salinisporamide A precursor chloroethylmaloyl-CoA, and thereby specifically regulates the production of salinisporamide A and not of other salinosporamides that are produced by *S. tropica CNB-440* (Lechner *et al.*, 2011).

In the genus *Salinispora* an important concept for the study of cryptic gene clusters was revealed (Amos *et al.*, 2017). Transcriptomic comparison of the *Salinispora* strains *S. pacifica* CNT-150, *S. tropica* CNB-440, *S. arenicola* CNS-205 and *S. arenicola* CNS-991 revealed that BGCs common between different strains are not necessarily controlled in the same way and could be active in one while silent in another. Such strain-specific silencing of a BGC was explained by mutation of regulatory genes. Indeed, an orphan BGC in *S. pacifica* (STPKS1) was expressed normally, while its counterpart in *S. tropica* was silent due to the lack of the AraC-family CSR, which was replaced by a transposase. Interestingly, this silent gene cluster is conserved throughout the *S. tropica* clade, which suggests that either this BGC is permanently silenced or that another regulator is involved in the control of the BGC. The BGC for the enediyene PKS1A was silent in CNS-991 and expressed in CNS-205. Comparative genomics and transcriptomic data revealed that a σ factor upstream of the BGC was expressed in *S. arenicola* CNS205, but not in CNS991. Differential expression of this σ factor was proposed be a consequence of its different chromosomal location in the two strains. The BGC for the black spore pigment was present in all four *Salinispora* strains, but the full BGC was only expressed by *S. tropica* CNB-440 and *S. pacifica* CNT-150, whereas only a subset of the genes within the gene cluster was expressed in the

two *S. arenicola* strains. The spore pigment BGCs that were entirely expressed contained one or two *luxR* genes, whereas the partially expressed BGC contained small genes encoding hypothetical proteins of unknown function. The *sta* gene cluster for staurosporine was also differentially expressed between the four *Salinispora* strains, but all strains contained the *malT* gene for the CSR. Finally, the fact that a BGC (NRPS4) was expressed in *S. arenicola* and *S. pacifica,* but not in *S. tropica* was explained by the lack of a xenobiotic response element in *S. tropica* (Amos *et al.*, 2017). Further genetic analysis of these interesting examples is required to fully understand the regulatory mechanisms for these BGCs. The differential expression of gene clusters between different species suggests that one feasible approach to the problem of silent gene clusters may be to look for the same (or highly similar) gene cluster in related actinobacteria, and see if the cluster is expressed there. With the ever-growing genome sequence information, this approach is becoming increasingly feasible, and is particularly attractive in strains that are not genetically tractable.

Regulation of rifamycin biosynthesis in *Amycolatopsis mediterranei*

Recently, the molecular regulation of the rifamycin BGC was studied in *Amycolatopsis mediterranei.* Although rifamycin and its derivatives are the first-line anti-tuberculosis drugs, the regulation of the rifamycin BGC was only studied recently. Deletion of *glnR* influences the biosynthesis of rifamycin, although this control is indirect (Yu *et al.*, 2007). The LuxR-type regulator RifZ, encoded by the last gene in the gene cluster, positively controls all of the operons in the rifamycin BGC (Li *et al.*, 2017a). The rifamycin BGC also encodes a TetR-family repressor (RifQ), which represses rifamycin biosynthesis and efflux. Deletion of *rifQ* resulted in increased production of rifamycin, while accumulation of rifamycin B lowered the affinity of RifQ for its target sequences (Lei *et al.*, 2018). This system is consistent with what is known for other TetR-family regulators that control natural product biosynthesis.

GBL-receptors and antibiotic production in Actinobacteria other than *Streptomyces*

GBL-like molecules are produced by many actinobacteria, including the industrial important strains *A. teichomyceticus* (producer of teicoplanin), *A. mediterranei* (produces rifamycin), and *Micromonospora echinospora* (produces gentamicin) (Choi *et al.*, 2003). The exact structures of the GBL molecules produced

by these strains are unknown, but the type of GBL that is produced could be determined using binding assays with tritium-labeled GBL molecules as ligands (Choi *et al.*, 2003) (Polkade *et al.*, 2016). These binding assays confirmed that *A. teichomyceticus* produces a GBL similar to virginiae butenolide (VB) derived from *S. viginiae*. The strains *A. mediterranei* and *M. echinospora* produce a GBL similar to IM-2, derived from *S. lavendulae* (see section 9.2) (Choi *et al.*, 2003). In the rifamycin producer *A. mediterranei,* four genes that encode GBL-receptor paralogues are present, namely *bamA1-bamA4* (Aroonsri *et al.*, 2008). All four receptor proteins can bind GBLs derived from *Streptomyces*, including VB from *S. virginiae* and SCB1 from *S. coelicolor*. Only BamA1 was shown to bind the IM-2 GBL , an autoregulator produced by *A. mediterranei* itself (Aroonsri *et al.*, 2008, Choi *et al.*, 2003).

Kitasatospora setae, a member of a genus closely related to *Streptomyces*, harbours several GBL-receptors (Girard *et al.*, 2013, Choi *et al.*, 2004). *K. setae* produces bafilomycins A1 and B1. These macrolides specifically inhibit vacuolar H^+ -ATPases and are used in studies of molecular transport in eukaryotes. The genome of *K. setae* contains three genes that are similar to GBL-receptors, namely *ksbA*, *ksbB* and *ksbC* (Aroonsri *et al.*, 2012). KsbA binds ^{3}H-labeled SCB1, and deletion of *ksbA* increases bafilomycin biosynthesis (Choi *et al.*, 2004). Conversely, KsbC indirectly represses bafilomycin biosynthesis, perhaps via the activation of the gene for the autoregulator KsbS4 (Aroonsri *et al.*, 2012). KsbC also indirectly activates the production of kitasetaline, a β-carboline alkaloid, and of the kitasetaline derivative JBIR-133 (Aroonsri *et al.*, 2012).

Interestingly, *Rhodococcus jostii*, a genus of the *Nocardiaceae* produces the GBL (called RJB) that is structurally identical to a precursor of SCB2 (6-dehydro SCB2) produced by *S. coelicolor*, and can bind to the *S. coelicolor* GBL receptor ScbR (Ceniceros *et al.*, 2017). This suggests cross-family communication mediated by GBLs in the natural environment. The gene for GBL biosynthesis, *gblA*, is located in a GBL BGC that is conserved between different *Rhodoccocus* species. This GBL BGC also encodes a GBL-receptor protein GblR and the biosynthesis enzyme GblE, which is an NAD-epimerase/dehydratase. Genome sequencing of *R. jostii* RHA1 indicated that the strain potentially has a rich NP biosynthetic repertoire. The precise role of GBLs in the regulation of natural product biosynthesis in

Rhodococcus, and the value of the NPs these Actinobacteria can produce, merit further investigation.

OUTLOOK

Over the last decade it has become increasingly clear that *Streptomyces* species and other antibiotic-producing Actinobacteria produce only a small percentage of their secondary metabolome under laboratory conditions. Accessing the chemistry specified by this 'silent majority' - also referred to as dark matter - without a doubt holds potential for drug discovery. This untapped resource can be harnessed by both genetic and non-genetic methods which been reviewed recently (Zarins-Tutt *et al.*, 2016). The proverbial 'holy grail' in this respect is development of small molecules that can simply be added to culture media to elicit the production of all or ideally only a subset of compounds. Progress has been achieved in this area (i.e. sugar-responsive antibiotic repressors, REEs, GBLs and manipulation of C, N and P concentrations, discussed above); the molecular insights that is reviewed above can be harnessed to develop strategies to activate antibiotic production. Clearly, more work is required with the identification of other small molecules. Reporter-based methods have therefore been developed to aid detection of activated or de-repressed gene clusters (Guo *et al.*, 2015, Sun *et al.*, 2017), and screening using small molecule libraries forms an attractive black box alternative to rational approaches that are based on molecular insights (Craney *et al.*, 2012, Xu *et al.*, 2017). For details on molecular, environmental and HT screening approaches to find elicitors we refer the reader to recent reviews (Okada & Seyedsayamdost, 2017, van der Meij *et al.*, 2017). Elicitors are also instrumental in unsupervised metabolomics approaches, required to identify compounds in the complex metabolic matrix of microbial cultures (Bingol *et al.*, 2016). Here, significant fluctuation of the secondary metabolome needs to be achieved, allowing statistical coupling of a yet unknown bioactivity of interest to a specific metabolite and/or a BGC. NMR- or MS-based metabolomics then facilitate the identification of the sought-after bioactive molecules (Gaudêncio & Pereira, 2015, Wu *et al.*, 2015b). Ultimately, the productivity of any given biosynthetic pathway is dictated by one or more CSRs. The examples provided by among others *Salinispora* show that BGCs may be silent in one species of a given genus, and active in another. Thus, with the growing wealth of genome sequence information, a promising strategy is to look for related bacteria that harbour a close relative of the gene cluster of

interest. Indeed, it is not illogical to assume that over the hundreds of millions of years of evolution, the natural products specified by the BGCs have remained structurally the same or highly similar, but are expressed under different growth conditions or in response to different environmental stimuli. The functionality of most putative CSRs can be deduced bioinformatically (i.e. as a repressor or an activator). Therefore, an obvious strategy and one that is commonly employed for elicitation of poorly expressed BGCs is augmentation of endogenous regulatory system(s). For example, by deleting genes encoding repressors or over-expressing those encoding activators (Olano *et al.*, 2014, Seipke *et al.*, 2011). This strategy depends upon the genetic tractability of the organism, but this is becoming less and less of a requirement as the cloning of large genomic fragments and their *de novo* synthesis becomes more feasible, which enables their tractability and heterologous expression in a panel of potential hosts (Komatsu *et al.*, 2013, Gomez-Escribano & Bibb, 2011, Nah *et al.*, 2017). Indeed, it is now possible to completely refactor the regulation of a biosynthetic pathway by replacing native promoters with those that are constitutively expressed to increase production titres using CRISPR-Cas9 technology (Zhang *et al.*, 2017a). Longer term, improved understanding of how secondary metabolism is controlled and the development of approaches to exploit this and/or efficient synthetic biology strategies to activate biosynthetic pathways are required in order to capitalise on the treasures beneath our feet.

Conflicts of interests

The authors have no conflicts to declare.

Acknowledgements

The work was supported by grants 731.014.206 and 14221 from the Netherlands Organization for Scientific Research to GPvW.

Phylogenetic and transcriptional analysis of the biosynthetic gene cluster for angucyclines and lugdunomycin in *Streptomyces* sp. QL37

Helga U. van der Heul
Changsheng Wu
Guillermo Guerrero Egido
Somayah S. Elsayed
Victor J. Carrión Bravo
Gilles P. Van Wezel

P art of this chapter is published as:
Wu, C., van der Heul, H.U., Melnik, A.V., Lubben, J., Dorrestein, P.C., Minnaard, A.J., Choi, Y. H., Wezel, G.P. (2019) Lugdunomycin, an angucycline-derived molecule with unprecedented chemical architecture. *Angew Chem Int Ed Engl* **58**: 2809–2814

ABSTRACT

Angucyclines form the largest group of polycyclic aromatic polyketides. These extensively studied compounds are well known for their anti-tumor and antibacterial activities. Previously we discovered the extensively rearranged angucycline, lugdunomycin, which is produced by *Streptomyces* sp. QL37. Despite high production of the angucyclines, lugdunomycin is produced in low amounts under all studied growth conditions. To obtain more insight into the mode of action of lugdunomycin, obtaining higher titers is important. Here, we study the polyketide type II gene cluster for angucyclines (designated *lug*) in *Streptomyces* sp. QL37, using targeted mutagenesis, RNA sequencing and phylogeny. This revealed that the *lug* gene cluster drives the production of non-rearranged and C-ring rearranged angucyclines and is also required for lugdunomycin biosynthesis. Transcriptomics in combination with bioinformatic comparisons of the *lug* gene cluster with other angucycline BGCs indicate that the *lug* gene cluster likely comprises 28 genes. Two oxygenase genes (*lugOIII* and *lugOIV*) and one gene with an unknown product (*lugX*) were not detected in BGCs directing the production of non-rearranged and A- and B-ring rearranged angucyclines, suggesting possible novel enzymology. Orthologues of *lugF–lugE*, genes required for the production of the angucycline backbone were found in some 25% out of 1020 searched Actinobacterial genomes, suggesting that angucycline production is widely spread. Angucycline BGCs with similar architecture as *lug* were identified in at least 36 genomes; these strains may be a good source for the discovery of similar novel rearranged angucyclines.

INTRODUCTION

Polyketides are one of the largest family of natural products, and many compounds or semisynthetic derivatives thereof are clinically relevant as antibiotics (oxytetracycline, erythromycin, monensin), anticancer agents (doxorubicin, mithramycin), immunosuppressants (FK506), antifungals (amphotericin B), cholesterol lowering agents (lovastatin) and antiparasitic agents (Ivermectin) (Tibrewal & Tang, 2014, Zhang *et al.*, 2006). Though polyketides are synthesised by multiple kingdoms of life, the above mentioned molecules are mostly produced by *Streptomyces*, a genus of Gram-positive bacteria, characterised by their filamentous growth and spore formation (Risdian *et al.*, 2019). Streptomycetes harbour different types of polyketide synthases (PKS I, PKS II and PKS III) (Hertweck *et al.*, 2007) that form distinct core structures, using small organic acids as substrates. The generated framework is further modified by post-PKS tailoring enzymes, resulting in final products with a remarkable chemical diversity (Patrikainen, 2015).

In our search for novel natural products, we previously discovered the extensively rearranged polyketide lugdunomycin produced by *Streptomyces* sp. QL37, which was isolated from the Qinling mountains in China (Wu *et al.*, 2019). Besides lugdunomycin, a variety of angucyclines is produced by this strain (Wu *et al.*, 2019), these include non-rearranged angucyclines (**2**–**6**) and rearranged angucyclines, such as the limamycins (**8**–**9**). Recently the structure of one of the limamycin derivatives with $[M+H]^+ = 334.11$ was revised to the structure of pratensilin A (**7**) (Figure 1)(Mikhaylov *et al.*, 2021, Wu *et al.*, 2019). The postulated biosynthesis pathway of lugdunomycin includes C-ring cleavage of a non-rearranged angucycline by Baeyer–Villiger oxidation, followed by a structural rearrangement, and a proposed Diels-Alder reaction to mediate the reaction between the modified angucycline as a diene and *iso*-maleimycin as a dienophile (Wu *et al.*, 2019, Uiterweerd, 2020). The new structural features of lugdunomycin not only make the molecule interesting in terms of drug discovery, but also imply potentially novel enzymology involved in its biosynthesis (Wu *et al.*, 2019). This may be used for the modification of other molecules. However, lugdunomycin is produced in extremely low amounts, and 7.5 L of agar was required to obtain only 0.5 mg of the compound (Wu *et al.*, 2019). Accordingly, it is vital to increase the production of lugdunomycin, to obtain more knowledge on its biological activity.

Angucyclines exhibit a variety of biological activities, including antitumor, antibacterial and enzyme inhibitory activities (Kharel *et al.*, 2012, Korynevska *et al.*, 2007, Yixizhuoma *et al.*, 2017, Fakhruzzaman *et al.*, 2015). These molecules are polycyclic aromatic polyketides, synthesised by type II polyketide synthases (Kharel *et al.*, 2012). The polyketide carbon chain is produced by the so-called minimal PKS, consisting of an heterodimer of ketosynthases, KSα and KSβ, and an acyl carrier protein (ACP) which catalyse consecutive Claisen condensations of acetyl-CoA and malonyl-CoA into a long polyketide chain, which is further modified by ketoreductases and cyclases, resulting in a framework featuring a four ring system (Risdian *et al.*, 2019). For the production of angucyclines, a distinct cyclase is required that folds the fourth ring in an angular position (Sharif & O'Doherty, 2012). This cyclase differentiates the angucyclines, with a typical polycyclic aromatic benz[a]anthracene structure, from the anthracyclines (e.g. doxorubicin) (Metsa-Ketela *et al.*, 2003). The genes encoding the minimal PKS are situated in a biosynthetic gene cluster (BGC) that further contains modification-, transporter- and regulatory genes, essential for structural diversity, resistance and the control of gene expression in response to environmental and intracellular signals (Yushchuk *et al.*, 2019, Zou *et al.*, 2014).

Biosynthesis pathways of angucyclines have been widely studied *in vivo* as well as *in vitro*; especially the early post-PKS modification steps are very well known (Kharel & Rohr, 2012). Oxygenases are crucial modification enzymes and differentiate angucyclines in non-rearranged (typical) and rearranged (atypical) angucyclines (Fan & Zhang, 2018). Non-rearranged angucyclines are exemplified by urdamycins, landomycins and the angucyclines **2**–**6**, and are specified by the typical benz[*a*] anthracene skeleton. Rearranged (atypical) angucyclines are biosynthesised via oxidative ring opening (cleavage) of one of the angucycline rings, followed by rearrangement of the intermediates (Fan & Zhang, 2018). This group of molecules is exemplified by the jadomycins and kinamycins, biosynthesised via oxidative B-ring opening of dehydrorabelomycin; pratensilins (**7**), limamycins (**8**,**9**), and lugdunomycin (**1**), derived from a postulated C- ring opening and rearrangement (Fan & Zhang, 2018) (Fan *et al.*, 2012a, Pan *et al.*, 2017, Zhang *et al.*, 2017b, Wu *et al.*, 2019, Yixizhuoma *et al.*, 2017). In addition, the A-ring can be extensively modified and cleaved resulting in the compounds BE-7585A, grincamycins, fridamycins and urdamycin L (Lai *et al.*, 2018, Sasaki *et al.*, 2010, Yoon *et al.*, 2019, Maskey *et al.*, 2003, Rix *et al.*, 2003). It is postulated that the A-, B- and C-ring opening

involve a Baeyer–Villiger oxidation, due to the discovery of lactone intermediates (Figure 2)(Rix *et al.*, 2003, Tibrewal *et al.*, 2012, Wu *et al.*, 2019). For discussion on oxidative ring opening and the possible oxygenases involved in this reaction, I refer to Chapter 5.

Besides lugdunomycin many other C-ring cleaved angucyclines have been isolated from various *Streptomyces* species and this group of molecules has grown into a distinct subclass of natural products (Mikhaylov *et al.*, 2021). However, not much is known about their cognate BGC, and the reaction mechanism involved in C-ring cleavage (Mikhaylov *et al.*, 2021, Cao *et al.*, 2021).

In this study we identified the BGC required for the production of angucyclines, limamycins, pratensilin A and lugdunomycin in *Streptomyces* sp. QL37, which we designated *lug*. The role of this BGC was confirmed with targeted gene deletion. The borders of the *lug* gene cluster were predicted using RNA-seq and by comparison with angucycline BGCs from other *Streptomyces* spp. These comparisons also revealed genes specific to the *lug* gene cluster and other potential producers of C-ring rearranged angucyclines. Phylogenetic analysis of various *Streptomyces* strains and *Kitasatospora* strains revealed that the presence of the *lug* gene cluster and angucycline production was not limited to a specific phylogenetic clade but were distributed all over the phylogenetic tree.

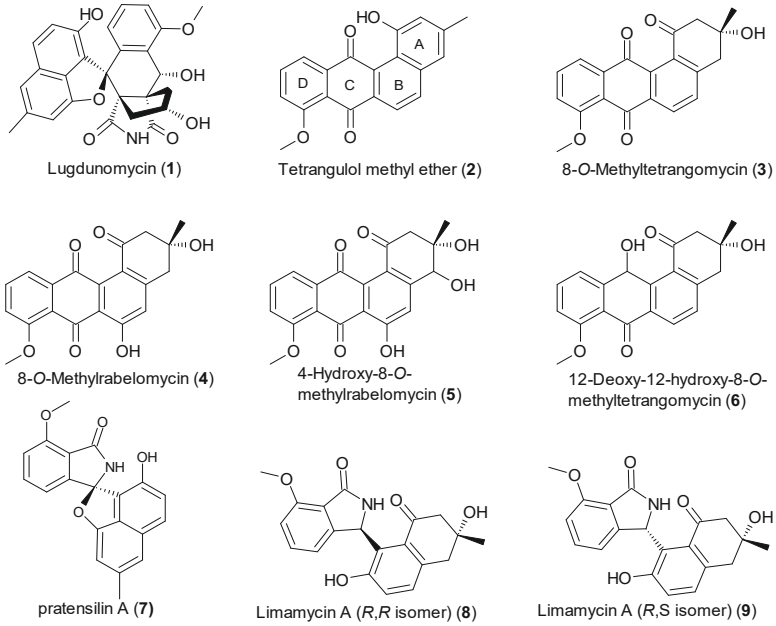


Figure 1 Structures of the compounds discussed in this study.
Lugdunomycin (**1**), non-rearranged angucyclines (**2**, **3**, **4**, **5**, **6**) and limamycins (**8**, **9**) were previously isolated from *Streptomyces* sp. QL37 (Wu *et al.*, 2019). The structure of compound **7** was later revised to be pratensilin A (Mikhaylov *et al.*, 2021).

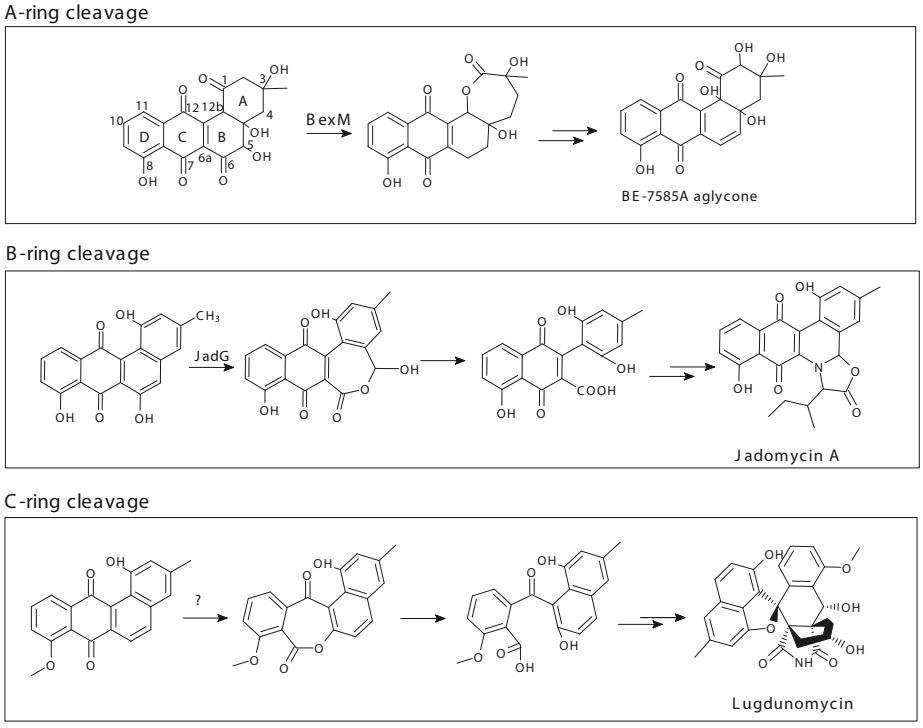


Figure 2 (Proposed) biosynthesis pathways of the BE-7585A aglycone, jadomycin A and lugdunomycin.

Each biosynthesis pathway starts with the non-rearranged or typical angucycline structure (Fan & Zhang, 2018). The ring opening at each ring is postulated to be catalysed by a Baeyer–Villiger mechanism, whereby an oxygen is incorporated in one of the rings, resulting in a lactone intermediate. BexM, an oxygenase-reductase is proposed to introduce an oxygen at position C12b–C1 of the angucycline A-ring, followed by hydrolysis of the lactone ring (Sasaki *et al.*, 2010). JadG introduces an oxygen at position C5–C6 of the angucycline B-ring followed by hydrolysis of the formed lactone and incorporation of *iso*-leucine, forming jadomycin A (Fan *et al.*, 2012a) . The activity of JadG was experimentally confirmed (Fan *et al.*, 2012a). For the production of the C-ring rearranged angucyclines, a similar mechanism is proposed as for the opening of the A-and B-rings, whereby an oxygen is incorporated in the C6a–C7 bond (Wu *et al.*, 2019). However, the question is which enzymes are involved in the opening of the C-ring (for further discussion I refer to below and Chapter 5).

RESULTS

Targeted deletion of the minimal PKS genes in the *lug* cluster

Lugdunomycin is derived from angucyclines (Wu *et al.*, 2019), which are generated by type II polyketide synthases (PKS) (Kharel *et al.*, 2012). Genome analysis of *Streptomyces* sp. QL37 using the bioinformatic tool antiSMASH 5.0 revealed 35 predicted BGCs, including a type II PKS cluster which is designated as *lug* (Figure S1) (Medema *et al.*, 2011). The by antiSMASH proposed region (BGC12) is 69,871 nt long (Figure S2 and Table S1). In this Chapter we propose a region of 29,183 kb as designated in Figure 3A.

To verify whether the *lug* gene cluster is indeed required for the biosynthesis of angucyclines, limamycins and lugdunomycin, genes encoding for minimal PKS enzymes were deleted in the wild-type strain via homologous recombination. LugA–C are the enzymes required for the biosynthesis of the angucycline backbone (Kharel *et al.*, 2012). The *lugA–OII* genes were replaced by double recombination with an apramycin resistance cassette (*aacC4*) using the unstable multi-copy plasmid pWHM3-*oriT* (Wu *et al.*, 2019, Vara *et al.*, 1989). The created mutant is further indicated as *lug-pks* mutant. Primers used for the generation of the knock-out construct are indicated in Table 3. We previously isolated lugdunomycin from cultures grown on minimal medium (MM) supplemented with 0.5% mannitol and 1% glycerol, while pratensilin A (**7**) and the limamycins (**8**, **9**) were isolated from cultures on R5 medium supplemented with 1% mannitol and 0.8% peptone, and angucyclines from both media (Wu *et al.*, 2019). Therefore, the wild-type strain and its *lug-pks* mutant were grown on both MM and R5 agar plates for seven days at 30 °C, and metabolites were extracted using ethyl acetate. The extracts were subsequently dried, re-dissolved in methanol and analysed using liquid chromatography-mass spectrometry (LC-MS). The *lug-pks* deletion mutant failed to produce any of the non-rearranged or rearranged angucyclines, including pratensilin A, the limamycins and lugdunomycin, under all conditions (Figure 3). This confirms that the *lug* gene cluster is indeed required and essential for (non)-rearranged angucyclines and lugdunomycin biosynthesis.

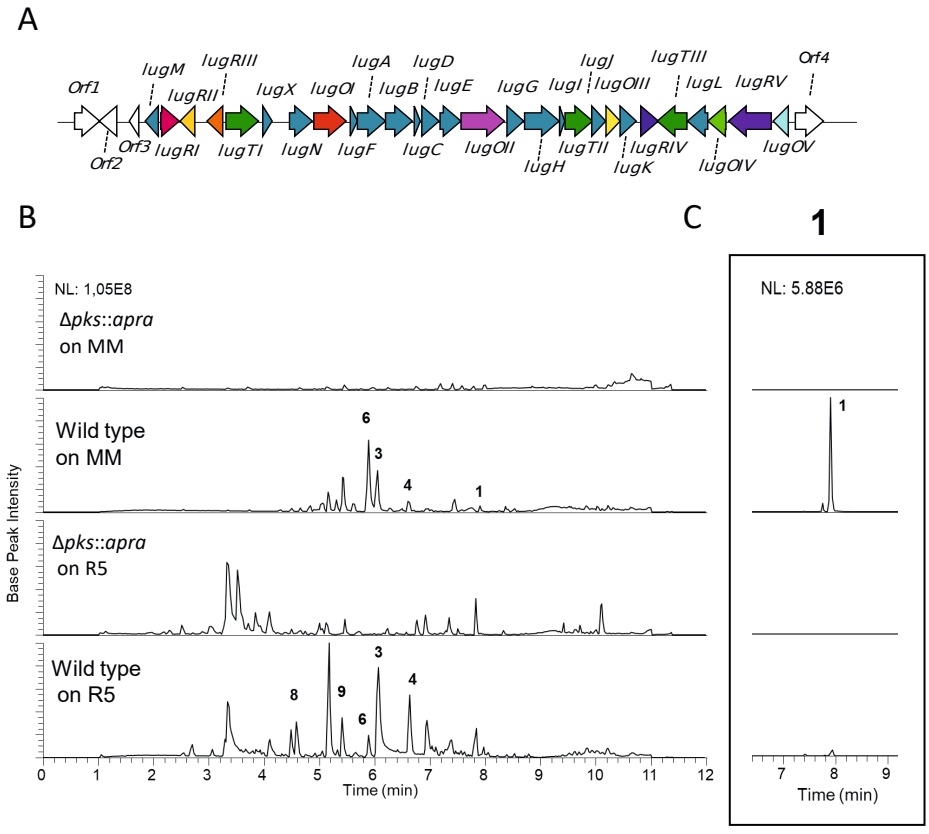


Figure 3 The *lug* gene cluster is required for the biosynthesis of angucyclines, pratensilin A, limamycins and lugdunomycin.
A) Genetic organization of the *lug* gene cluster. More information on the annotation of the genes can be found in Table S2. The marine blue genes indicate genes that are not annotated as either regulatory or oxygenase gene, the white genes are not part of the *lug* gene cluster B) Base peak LC-MS chromatogram of the extracts derived from *Streptomyces* sp. QL37 wild-type and its *lug-pks* mutant grown on MM and R5. Compound numbers are given over their corresponding peaks (Figure 1). The wild-type strain produced angucyclines (**6**,**2**,**3**,**4**) pratensilin A (**7**) and limamycins (**8**,**9**) and lugdunomycin (**1**), whereas the *pks* mutant was unable to produce these compounds (**2** and **7** could only be visualised in an extracted ion chromatogram). C) Extracted ion chromatogram of the water adduct ion of lugdunomycin (**1**, m/z 456.14) in the corresponding extracts shown in B. Lugdunomycin is produced by the wild type grown on MM, whereas the *pks* mutant is incapable of producing the molecule.

Prediction of the extent of the *lug* gene cluster using time-course RNA-seq on the transcriptome of *Streptomyces* sp. QL37

To obtain more insights in the extent of the *lug* gene cluster (described in this Chapter) and whether other BGCs are co-expressed with the cluster (described in Chapter 4), RNA-sequencing (RNA-seq) was performed. For this, *Streptomyces* sp. QL37 was grown on MM agar at 30 °C, and mycelia were harvested after 24 h (T_{24h}; vegetative growth), 36 h (T_{36h}; aerial growth), 48 h (T_{48h}; onset of sporulation) and 60 h (T_{60h}; sporulation). Four biological replicates were collected for all time points. From these samples, RNA was isolated from the mycelium, and secondary metabolites were simultaneously extracted from both the medium and the mycelium. The RNA sequencing (RNA-seq) data were normalised using DeSeq2 (Love *et al.*, 2014).

Principal component analysis (PCA) of the transcriptome data was done to evaluate the different replicates and treatments (Figure S3) This revealed that the replicates of the transcriptome of *Streptomyces* sp. QL37 at the different time points clustered into separate groups. The variation among the replicates of the 24 h samples was small, while that among the 36 h samples was relatively large.

At first the expression of the *lug* gene cluster was analysed (Figure 4) and thereby an indication could be made on the extent of the cluster; a similar gene expression pattern can indicate that these genes are co-regulated and are thus part of one gene cluster. To do so, the expression of all genes in region 12 (68,871 nt) predicted by antiSMASH was studied (Figure S2 and Table S1). Figure 3 depicts the average expression of each gene in region 12 at each of the four time points. In certain regions (QL37_12745–QL37_128760, DNA replication), (*orf1–orf2*, hypothetical protein, GntR-regulator), (*lugRI–lugOV*) the gene expression was relatively high (Figure 4). Comparing the normalised gene expression counts of each gene at T_{36h}, T_{48h} and T_{60h} with the expression counts at T_{24h}, confirmed that the region surrounding the *lugA–C* genes was significantly increased at the time points T_{36h}, T_{48h} and T_{60h}. On the left flank of *lugA–C* this expression trend extended until *lugRI* and on the right border of these genes until *lugOV*, indicating that *lugRI* and *lugOV* may define the borders of the *lug* gene cluster. On the right flank of the gene cluster *orf10* also showed an expression trend similar to that of the *lug* gene cluster, this gene is possibly involved in sugar transport (Table S1).

Most of the *lug* genes showed a similar expression pattern. The lowest expression was seen at T_{24h} and the highest expression was seen at T_{36h}. From T_{36h} onwards, the expression of the genes decreased. The genes *lugRI*, *lugRII*, *lugRIII*, *lugTI* and *lugX* had an expression pattern different from the other *lug* genes. The increase in *lug* gene expression at T_{36h} corresponded well with the appearance of angucyclines at T_{48h} and T_{60h} (Figure S4). These are also the time points were lugdunomycin was observed at T_{60h} and possibly at T_{48h}.

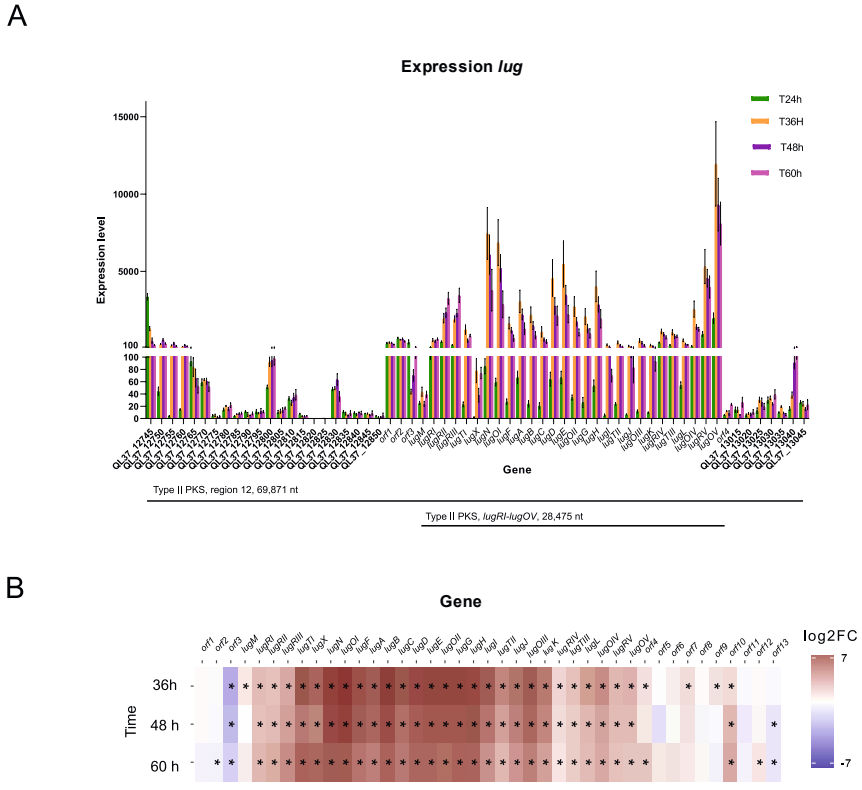


Figure 4 Expression level of the *lug* genes at T_{24h}, T_{36h},T_{48h} and T_{60h}
A) Barplot indicates the absolute normalised expression level of each gene in the predicted *lug* gene cluster. Error bars indicate the standard error of the mean (SEM). B) The heatmap indicates the log2FC of the expression of each gene in or around the *lug* cluster at $T_{36h,}$ T_{48h} and T_{60h} compared to the expression of each gene at T_{24h}. Data with *p*-value <0.05 are indicated with (*) . The expression level of each gene is increased from T_{36h}, which is in line with the production of angucyclines at T_{48h} (Figure S4). The heatmap suggests that the *lug* gene cluster includes the genes from *lugRI–lugOV*.

Prediction of the functional *lug* gene cluster using bioinformatic comparison

The RNA-seq data suggested that the *lug* gene cluster runs from *lugRI* to *lugOV* (Figure 4). We then analysed which of these genes are conserved in other *Streptomyces* spp. This should facilitate finding other *Streptomyces* strains that may produce C-ring rearranged angucyclines and perhaps lugdunomycin. Despite its differential regulation as shown by the RNAseq, we also included *lugM* (encoding an NADPH-dependent FMN reductase) in the bioinformatic comparison.

For the gene comparisons a database was generated of 1020 fully sequenced genomes of *Streptomyces* and *Kitasatospora* spp., obtained from the NCBI database and from our MBT strain collection (Zhu *et al.*, 2014b). As a cut-off for the presence of a gene, we used a similarity threshold of 40% aa identity between the predicted gene products. The core feature of the *lug* gene cluster is that it is an angucycline gene cluster and thus the *lug* type gene clusters should contain the genes *lugA–C* typical of type II PKS BGCs, encoding the ketosynthase α (KSα) and β (KSβ) subunits and an acyl carrier protein (ACP), respectively, which form the minimal PKS (Kharel *et al.*, 2012). These genes must be flanked by genes encoding cyclases (*lugF* and *lugE*) and a ketoreductase (*lugD*). The cyclase *lugF* is required for the cyclisation of the fourth angular ring (ring A), which leads to the typical angucycline framework (Kharel *et al.*, 2012, Kulowski *et al.*, 1999). Thus, first of all we filtered the collection for strains containing these core genes. In the searched genomes 262 strains contained genes similar to *lugF–lugE*, suggesting that these strains may all produce angucyclines. In other words, some 25% of all *Streptomyces* and *Kitasatospora* strains produce angucyclines.

Of the 262 strains containing the angucycline core genes, 36 strains contained at least 18 genes with significant similarity to those in the *lug* cluster. The 36 putative *lug*-like gene clusters were analysed using antiSMASH and compared using the software tool clinker using the complete region 12 from the genome of *Streptomyces* sp. QL37 (Figure 5 and Figure S5) (Gilchrist & Chooi, 2021, Blin *et al.*, 2019). This showed that *orf3* and the left flanking region of this gene are absent in all of the 36 strains and it showed that the gene *orf4* and the right flanking region of this gene are not found in the neighbourhood of the *lug*-type gene clusters of the 36 strains, using a 40% similarity cut-off. Hierarchical clustering using Ward's method mapped to the generated table showed a distribution of BGCs in different groups (or clusters) based on the presence or absence of a gene in the *lug*-type BGCs.

The genes from *lugM* to *lugOV* are frequently shared with the genes in the *lug*-like clusters, which indicates that *lugM* and *lugOV* most likely define the boundaries of the *lug* gene cluster. The region *lugOI –lugOII* was conserved between all BGCs. Some BGCs contained the full-length *lugOII*, while in other strains *lugOII* was split so as to encode a separate oxygenase and reductase (Xiao *et al.*, 2020). Only *lugX* and *lugK* were not found near any of the 36 *lug*-like clusters. Homologues of *lugK* were identified elsewhere in other regions of the chromosome in 28 of the 36 strains. LugK a phosphopantetheinyl transferase (PPTase), which can be used in multiple pathways (Bunet *et al.*, 2014). LugX is a hypothetical protein, and it is not only absent from BGCs, but in fact no homologues of this protein are found in the NCBI database. The role of this putative protein in lugdunomycin biosynthesis (if any) still has to be established. The gene *lugRI* was shared with only one *lug*-type gene cluster; that of *Streptomyces* sp. 94 (Hulcr *et al.*, 2011). However, there was a distance of 23 genes between the *lugRI* homologue and the rest of the *lug*-type gene cluster of *Streptomyces* sp. 94, including many mobile genetic elements.

The BGCs with highest similarity to the *lug* gene cluster were those from *S. scopuliridis* RB72 strain NRRL B-24574, *Streptomyces* spp. CB00072, b84, w007 (Figure 6). Remarkably, the *lug*-like BGCs in the strains *S. scopuliridis*, *Streptomyces* spp. CB00072, b84 and w007 were very similar to each other and to the *lug* cluster in terms of gene order and orientation. Additionally, the four clusters lacked the genes which are homologous to *lugM*, *lugRI*, *lugX*, *lugJ*, *lugK*, and *lugTIII*. *Streptomyces* sp. w007 and *Streptomyces* sp. CB00072 are known angucycline producers (Zhang *et al.*, 2012, Zhang *et al.*, 2015a, Cao *et al.*, 2021). They also produce C-fragmented angucyclines, suggesting that the genes that are absent in these clusters (*lugM*, *lugRI*, *lugX*, *lugJ*, *lugK*, and *lugTIII*) are not required for the production or transport of the C-ring rearranged angucycline that is required for lugdunomycin biosynthesis.

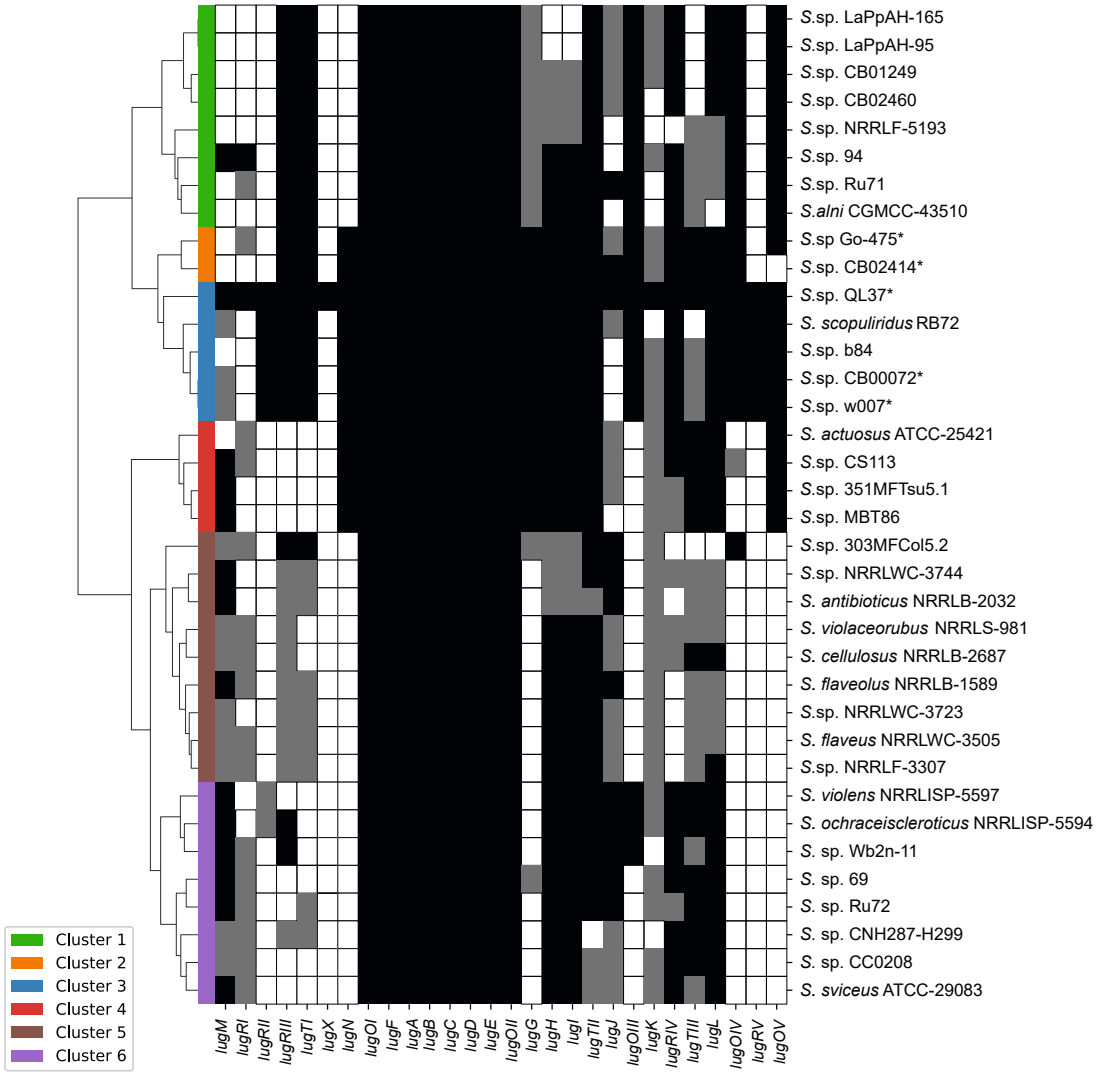


Figure 5 Absence or presence of the *lug* genes in other Streptomyces strains. Hierarchical clustering analysis using Ward's method mapped against a table indicating if a gene is absent (white) present in a polyketide type II BGC (black) or present outside a polyketide type II BGC (grey) in the genome of *Streptomyces* strains that contain at least 18 genes of the *lug* gene cluster. More information about the annotation of each gene can be found in the Table S2.

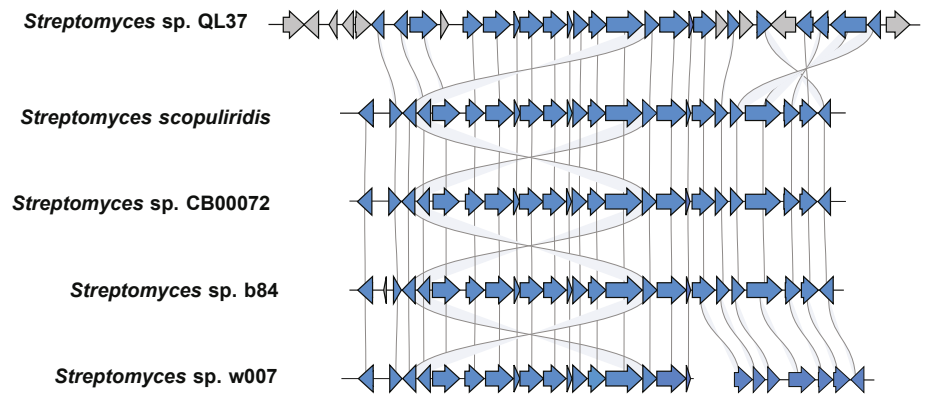


Figure 6 Gene by gene comparison of the *lug* gene cluster with the highest similar ones in other *Streptomyces* strains.
The gene clusters were compared using Clinker (Gilchrist & Chooi, 2021, Berdy, 2005). The strokes between the BGCs link genes that share at least 40% identity. Grey coloured genes, like ORF1, ORF2, ORF3, *lugM*, *lugRI*, *lugX*, *lugK* and ORF4 have no homologues in any of the compared gene clusters.

Prediction of genes that are specific to *lug*-type gene clusters

To obtain a further indication on which genes are specific to the *lug*-type gene clusters, the *lug* gene cluster was compared with 27 BGCs directing non-rearranged angucyclines and A- and B-ring rearranged angucyclines using the software tool Clinker (Table 1)**.** (Gilchrist & Chooi, 2021). Most of these BGCs were derived from the Minimum Information about a Biosynthetic Gene cluster (MIBiG) database, as the organisms containing these BGCs are not fully sequenced, or are detected in species other than *Streptomyces* or *Kitasatospora* (Kautsar *et al.*, 2020). Furthermore, the region of *lugOI–lugE* was highly conserved in all angucycline BGCs. The oxidoreductase gene *lugOII* was not found in all BGCs, but those lacking *lugOII* contained a second paralogue of *lugOI*, a gene similar to *jadF* from the jadomycin BGC (Xiao *et al.*, 2020, Kallio *et al.*, 2008a, Kharel & Rohr, 2012, Fan & Zhang, 2018, Patrikainen *et al.*, 2012). The *lugM* gene was frequently found in BGCs known to produce angucyclines, further supporting the idea that this gene is a *bonafide* member of the *lug* gene cluster. Interestingly, the entire region spanning from *lugRI* to *lugX*, as well as *lugOIII*, *lugOIV* and *lugRV* were absent in all 27 BGCs. This suggests that all of these genes may be specific for biosynthetic pathways involved in the production of C-ring cleaved angucyclines. Again, *lugK* was rarely found; only the simocyclinone BGC of *S. antibioticus* Tü 6040 has an orthologue of the gene, suggesting it plays a minor role in angucycline biosynthesis. Finally,

lugN and *lugOV* were only detected in the hatomarubigin BGC of *Streptomyces* sp. 2238 SVT-4 (Izawa *et al*., 2014)

Table 1 Overview of *lug* genes detected in BGC specifying the production of known non-rearranged angucyclines and A- and B-ring rearranged angucyclines*.

Strain	BGC	*lugM*	*lugRI*	*lugRII*	*lugRIII*	*lugTI*	*lugX*	*lugN*	*lugOI*	*lugF*	*lugA*	*lugB*	*lugC*	*lugD*	*lugE*	*lugOII*	*lugG*	*lugH*	*lugI*	*lugTII*	*lugJ*	*lugOIII*	*lugK*	*lugRIV*	*lugTIII*	*lugL*	*lugOIV*	*lugRV*	*lugOV*		
Streptomyces sp. 120454	Mayamycin																												Non rearragend (typical)	Angucycline backbone
S. antibioticus ATCC 11891	Oviedomycin																													
S. globisporus 1912	Landomycin																													
S. cyanogenus S136	Landomycin																													
S. griseus NTK97	Frigocyclinone																													
Streptomyces sp. 2238-SVT4	Hatomarubigin																													
Streptomyces sp. PGA64	Gaudimycin																													
Streptomyces sp. CS057	Warkmycin																													
Streptomyces sp. TK08046	Saprolmycin																													
S. antibioticus Tü 6040	Simocyclinone																													
Streptomyces sp. NRRL B24484	Simocyclinone																													
Streptomyces sp. KY 40-1.	Saquayamycin																													
Streptomyces sp. MBT86	Saquayamycin																													
Streptomyces sp. SCC2136	Sch-47554 and Sch-47555																													
Kibdelosporangium sp. MJ126-NF4	Azicemycin																													
Streptomyces sp. CNZ748	Grincamycin																												A-ring cleaved	
Amycolatopsis orientalis subsp. vinearia BA-07585	BE-7585A																													
S. griseoflavus Gö 3592	Gilvocarcin																												B-ring cleaved	
S. albaduncus C38291	Chrysomycin																													
S. ravidus NRRL 11300	Ravidomycin																													
S. venezuelae ISP5230	Jadomycin																													
S. murayamaensis sp. nov. Hata et Ohtani	Kinamycin																													
S. ambofaciens ATCC 23877	Kinamycin																													
S. albus DSM 41398	Fluostatin																													
Salinispora pacifica DPJ0016	Lomaiviticin																													
Micromonospora echinospora SCSIO 0408	Nenestatin																													
S. aureofaciens CCM3239/*S. lavenduleae* subsp. *lavendulae* CCM 3239	auricin																												Pyranonaphthoquinone (**)	

*As cut-off for the presence of a gene we used a threshold of 40% aa identity between the predicted gene product. Gene products with less than 40% aa identity are indicated in white. Gene products with an identity of 40≤aa<60% identity are indicated in light grey; gene products with an identity of 60≤aa<80% are indicated in dark grey and gene products with an identity of 80≤aa≤100% are indicated in black.

** Though auricin has been characterised as part of the angucycline family, due to the typical angucycline cyclase gene in its BGC, recent developments in the characterisation of the auricin structure and extensive research on its BGC, revealed it possibly belongs to the pyranonaphtoquinones instead (Matulova *et al.*, 2019).

Phylogenetic analysis of *Streptomyces* sp. QL37 and its relatives

A phylogenetic tree of the strains analysed in this Chapter was generated using the PhyloPhlAn algorithm (Figure 7A) (Asnicar *et al.*, 2020). The phylogenetic analysis revealed that *Streptomyces* sp. QL37 is closely related to *Streptomyces* sp. LamerLS-316 (Book *et al.*, 2016). It is important to note that *Streptomyces* sp. LamerLS-316 lacks an angucycline BGC (Figure 7B). In fact, strains that contain *lug*-like gene clusters are scattered over the phylogenetic tree, indicating that the presence of the *lug* cluster is not correlated to overall phylogeny of the strains. Still, some of the strains that contain a *lug*-like gene cluster are closely related to each other, for example *Streptomyces* spp. w007, CB00072 and b84.

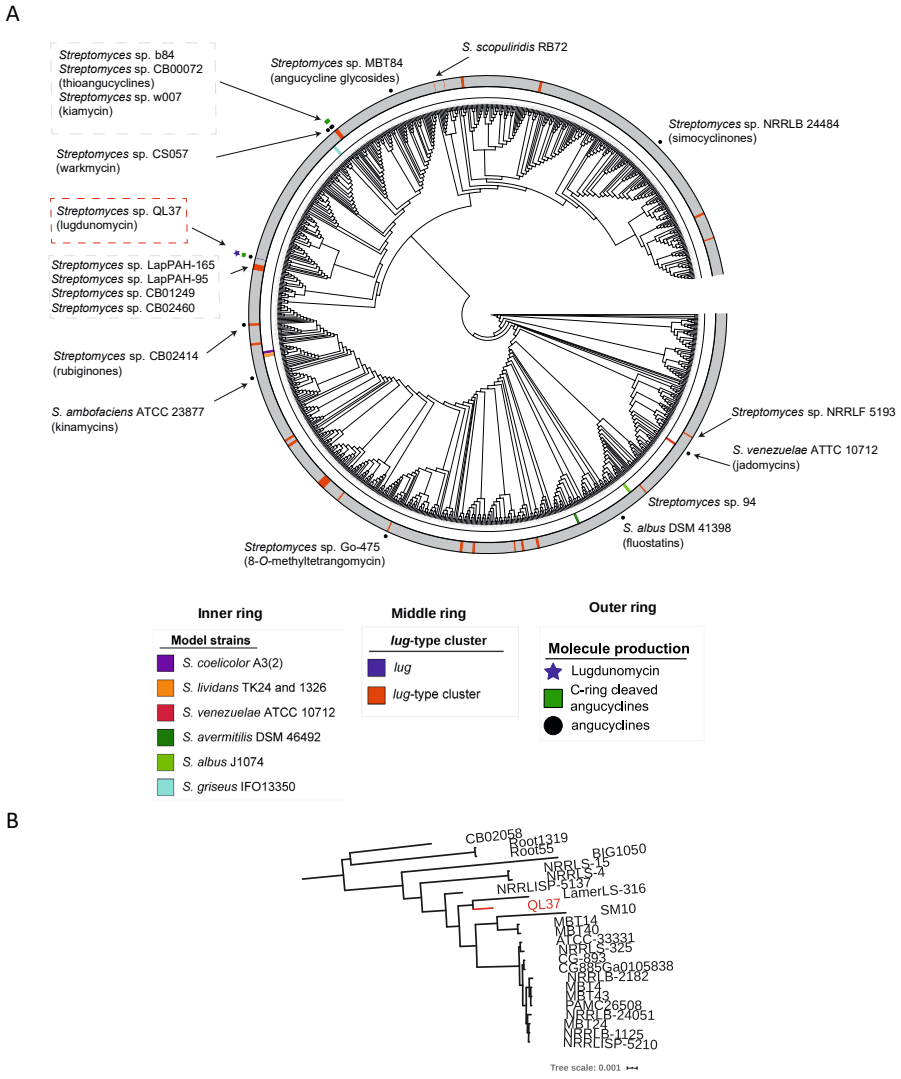


Figure 7 Phylogenetic tree of 1020 *Streptomyces* and *Kitasatospora* strains from the NCBI database and the Leiden MBT collection.
A) The phylogenetic tree was generated using the PhyloPhlAn algorithm (Asnicar *et al.*, 2020) based on four housekeeping genes (*atpD*, *recA*, *trpB* and *gyrB*). The colours of the inner ring represent different well-known *Streptomyces* spp. In the second ring, the colours indicate strains containing different *lug*-type gene clusters. The outer ring shows the *Streptomyces* spp. that are known to be producers of angucyclines, C-ring rearranged angucyclines or lugdunomycin. B) Clade of the phylogenetic tree, showing *Streptomyces* sp. QL37 and its close relatives

DISCUSSION

Streptomyces sp. QL37 produces the highly complex angucycline derivative lugdunomycin (Wu *et al.*, 2019). A biosynthesis pathway was proposed wherein the C-ring of the angucycline backbone is cleaved and subsequently rearranged. The rearranged angucycline was proposed to subsequently react with *iso*-maleimycin through a Diels-Alder [4+2] cycloaddition to form lugdunomycin (Uiterweerd, 2020, Wu *et al.*, 2019). This new study provides clues as to which genes might be involved in lugdunomycin biosynthesis. In this study an angucycline BGC (*lug*) was characterised and analysed using deletion experiments, RNA-seq data analysis and bioinformatic comparisons.

The *lug* gene cluster is a type II PKS gene cluster that shows high similarity to angucycline BGCs. Deletion of *lugA–OII*, of which *lugA–C* encode the minimal PKS enzymes KSα, KSβ and ACP, yielded a strain that did not produce any angucyclines, limamycins or lugdunomycin. Indeed, deletion of the minimal PKS of BGC12 abolished production of all angucycline-related metabolites, including lugdunomycin. AntiSMASH predicted a 70 kb genomic region which includes a complete angucycline BGC and flanking regions, and we performed a more in depth analysis to determine the likely borders of the BGC. Bioinformatic comparison of 1020 *Streptomyces* spp. identified that some 25% of all strains contained angucycline BGCs, a huge number that again underlines how abundant angucyclines are in nature. Of these, 36 strains had a BGC that shared at least 18 genes with the *lug* gene cluster, which were therefore called *lug*-like gene clusters. These include the known angucycline producers *Streptomyces* sp. w007, *Streptomyces* sp. CB00072, *Streptomyces* sp. CB02414 and *Streptomyces* sp. Go-475 (Zhang *et al.*, 2012, Zhang *et al.*, 2015a, Cao *et al.*, 2021, Kibret *et al.*, 2018). We did not find a clear correlation between strain phylogeny and the presence or absence of angucycline BGCs, except for a few strains sharing a highly similar BGC. Based on our comparative analysis, the *lug* gene cluster spans from *lugM* to *lugOV*, considering that the region left of *lugM* and right of *lugOV* were not conserved in any of the 36 *lug*-type gene clusters, nor in the 27 angucycline BGCs directing the production of non-rearranged and A- and B-ring rearranged angucyclines. Time course RNA-seq of the transcriptome of *Streptomyces* sp. QL37 under conditions that lugdunomycin is produced (MM agar) was performed, which showed that the genes from *lugRI–lugOV* had a significantly higher transcription than the

flanking genes. Although, the expression of *lugM* was relatively low compared to the expression of *lugRI–lugOV*, the gene is part of the *lug* gene cluster according to the bioinformatic comparisons. Possibly the expression of *lugM* is regulated differently compared to the expression of *lugRI–lugOV*. Thus, the RNA-seq data combined with the bioinformatic predictions support that the *lug* gene cluster spans from *lugM* to *lugOV*, with in total 28 genes, comprising 29,183 nt of DNA.

The *lug* gene cluster includes among others five oxygenase genes (Chapter 5), five regulatory genes (Chapter 6), and one gene of unknown function (*lugX*) that was exclusively found in *Streptomyces* sp. QL37. Although all 36 strains with *lug*-like gene clusters shared the core set of minimal PKS genes, the post-PKS tailoring genes varied between strains. Accordingly, they could be a good source for the discovery of novel angucycline derivatives, and thus to help understand the evolutionary relationships in the biosynthesis of angucyclines. Of these 36 strains, only *Streptomyces* sp. 94 contains an orthologue of *lugRI* (Hulcr *et al.*, 2011), while *lugK* was not found in any of the angucycline BGCs, but instead was found elsewhere in the genome for 28 out of the 36 strains; *lugK* encodes a phosphopantetheinyl transferase (PPTase), which can be used in multiple pathways (Bunet *et al.*, 2014), and it remains to be seen whether the orthologues play a role in angucycline biosynthesis. Further comparison of the *lug* gene cluster with the 27 previously studied angucycline BGCs that specify the biosynthesis of non-rearranged and A- and B-ring rearranged angucyclines showed that besides *lugX*, seven other *lug* genes were absent in any of these clusters, suggesting they are specific to *lug*-type gene clusters. This included the regulatory genes *lugRI–lugRIII* and *lugRV*, a transporter gene (*lugTI*), and the oxygenase genes *lugOIII* and *lugOIV*. Orthologues of *lugRIV*, the only regulatory genes found in most of the BGCs, typically activate the angucycline BGC they belong to, including *jadR1* and *aur1P* from the jadomycin and auricin BGC in *S. venezuelae* ISP5230 and *S. aureofaciens* CCM3239, respectively (Wang *et al.*, 2009, Novakova *et al.*, 2005). Of the 27 angucycline BGCs, only the hatomarubigin BGC harboured copies of *lugN* and *lugOV*. it is important to note that the *lug* gene cluster and the hatomarubigin BGC both direct the production of 8-*O*-methylated angucyclines, which may require the methyltransferase encoded by *lugN* (Izawa *et al.*, 2014). HrbF shares homology with LugOV (48 % amino acid identity, 84% coverage) and is an oxygenase that regulates the regiospecificity of oxygenase enzymes during hatomarubigin biosynthesis in *Streptomyces* sp. 2238 SVT-4 (Izawa *et al.*, 2014). No orthologues

of *lugOIII* and *lugOIV* could be found in any of the 27 previously studied angucycline BGCs; however, they were conserved in BGCs that likely direct the biosynthesis of C-ring cleaved angucyclines, suggesting a direct correlation to opening of the C-ring. The Lug oxygenases are described in detail in Chapter 5.

In summary, the *lug* gene cluster drives the production of C-ring-rearranged angucyclines and is required for the complex angucycline derivative lugdunomycin. Comparative analysis could clearly separate BGCs for C-ring rearranged angucycline from those for B-ring cleaved angucyclines. In fact, the entire region spanning from *lugRI* to *lugX*, as well as *lugOIII*, *lugOIV* and *lugRV* were absent in all 27 previously studied BGCs for non-rearranged and A- and B-ring rearranged angucyclines. This provides important leads for the understanding of these distinct biosynthetic pathways of angucycline biosynthesis. Further studies should reveal which genes drive the production of lugdunomycin itself.

Acknowledgments
We thank Du Chao for his assistance in creating Figure 5. The work was supported by NACTAR grant 16439 from the Netherlands Organization for Scientific research (NWO).

MATERIALS AND METHODS

Bacterial strains and growth conditions

Streptomyces sp. QL37 was obtained from the MBT strain collection (Zhu *et al.*, 2014b). The strain was isolated from soil in the Qinling mountains (P. R. China) as described previously and was deposited to the collection of the Centraal Bureau voor Schimmelcultures (CBS) in Utrecht, the Netherlands, under deposit number 138593 (Wu *et al.*, 2019). *Streptomyces* sp. QL37 was grown on soya flour mannitol (SFM) for seven days at 30 °C, then spores were collected and stored in 20% glycerol at -20 °C (Kieser *et al.*, 2000). For conjugation *Streptomyces* sp. QL37 was grown on SFM containing 60 mM $MgCl_2$ and 60 mM $CaCl_2$ (Wang & Jin, 2014). Ex-conjugants were selected using thiostrepton (20 µg/ml) and apramycin (50 µg/ml) (Wu *et al.*, 2019). For routine cloning, *Escherichia coli* JM109 was used and grown on Luria Broth (LB), where needed supplemented with antibiotics as selection markers. *E. coli* ET12567/pUZ8002 was used for conjugation of plasmids towards *Streptomyces sp.* QL37. Strains containing plasmids were selected on

LB containing ampicillin (100 µg/ml), apramycin (50 µg/ml), chloramphenicol (25 µg/ml) and kanamycin (50 µg/ml). All strains and plasmids used in this study are listed in Table 2.

Construction of knock-out mutants

Targeted deletion of *lugA–OII* was executed as previously described (Wu *et al.*, 2019, Vara *et al.*, 1989). To obtain a construct for gene disruption, approximately 1.5 kb regions up- and downstream of *lugA–OII* were amplified from the chromosome of *Streptomyces* sp. QL37 using primer pairs 1+2 and 3+4, respectively (Table 3) and cloned as EcoRI-XbaI and XbaI-BamHI fragments into a derivative of the unstable multi-copy plasmid pWHM3 [14] that harbours *oriT* in the PstI site to allow its conjugative transfer. The apramycin resistance cassette *aac(3)IV* flanked by *loxP* sites was then inserted in-between using an engineered XbaI site. The correct knock-out construct was transformed to the methylase-deficient strain *E. coli* ET12567/pUZ8002 [15] and subsequently further introduced into *Streptomyces* sp. QL37 by conjugation, following the protocol as described previously [8]. The correct mutant was selected by resistance to apramycin (50 µg/ml) and sensitivity to thiostrepton (10 µg/ml) [16]. The introduction of the apramycin cassette was verified by PCR with the primers 5-8 (Table 3) and Sanger sequencing of the PCR products (Baseclear).

Table 2 Bacterial strains used in this study

Strains		
Name	Description	Reference
***Streptomyces sp.* QL37**	Wild type, CBS 138593	(Wu *et al.*, 2019), (Zhu *et al.*, 2014b)
Streptomyces sp.* QL37Δ *lug-pks	Deletion mutant of the *lugA–OII* genes (the minimal PKS), Apra^r.	(Wu *et al.*, 2019)
***Escherichia coli* ET12567/ pUZ8002**	*E. coli* ET12567 containing pUZ8002. The strain was used for conjugal transfer. pUZ8002 is a derivative of RK2 with a mutation in *oriT* and containing the *tra* gene; Cm^{r}Km^r.	(MacNeil *et al.*, 1992), (Paget *et al.*, 1999)
***Escherichia coli ET12567/*pUWL-Cre**	*E. coli* ET12567 harbouring pUWL-Cre, a pUWL-*oriT* derivative containing the *cre* gene under the control of the *ermE* promoter. Cre recombinase recognises the *loxP* site, used for removal of *aac (3)IV* (Apra^R); Cm^r, Thio^r.	(Fedoryshyn *et al.*, 2008)
***E. coli* JM109**	*end*A1, *rec*A1, *gyr*A96, *thi, hsd*R17 (r_k^-, m_k^+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F´ *tra*D36, *pro*AB, *laq*I^{q}ZΔM15 Used for general cloning	(Yanisch-Perron *et al.*, 1985)
Plasmids		
Name	Description	Reference
pWHM3-*oriT*	High copy number *E. coli-Streptomyces* shuttle vector. The *oriT* from pSET152 was inserted in the PstI site of pWHM3.	(Wu *et al.*, 2019, Garg & Parry, 2010, Vara *et al.*, 1989)

Table 3 PCR primers used in this study

No.	Name	Sequence 5' --> 3'^
1	MinPKS_LF_Fw	ctagGAATTCCCGCCACCACCGAGCTCTTC
2	MinPKS_LF_RV	GAAGTTATCCATCACCTCTAGAGATACCGGTGATGACGACCC
3	MinPKS_RF_Fw	GAAGTTATCGCGCATCTCTAGAGCCGAGCAGCTCGACCGTTAC
4	MinPKS_RF_Rv	ctagGGATCCCTGCCCTTGTCGAGAAGCAGTG
5	MinPKS_Apra_ Fw	GGATCAGAGATGATCTGCTCTGCCT
6	MinPKS_Apra_Rv	GTCGCCCGTGTCCATGAACTC
7	MinPKS_LF_ Check_Fw	GGTAGCCGCACCCACCGGAC
8	MinPKS_RF_ Check_Rv	TCCTTCGGTCGGGAACCGGG

^ restriction sites underlined. GAATTC, EcoRI; GGATCC, BamHI; TCTAGA, XbaI

Metabolite extraction

For angucycline and lugdunomycin production *Streptomyces* sp. QL37 was grown on minimal media agar plates (MM) supplemented with 0.5% mannitol and 1% glycerol (w/v) as the carbon source. The strains were grown on R5 agar medium plates supplemented with 0.8% peptone and 1% mannitol (w/v) for angucycline and limamycin production (Wu *et al.*, 2019, Kieser *et al.*, 2000). For MM we used Iberian agar (TM Duche & Sons Ltd, batch from 2017) and for R5 Bacto agar (Brunswig Chemie). After seven days of growth at 30 °C the agar plates were cut into small pieces and soaked in 25 ml of ethyl acetate for 12 hours. Subsequently the ethyl acetate was decanted and evaporated at room temperature. This process was repeated two times. The dried extract was re-dissolved in methanol (MeOH) and centrifugated to remove any undissolved matters. Subsequently the MeOH solutions were transferred to new pre-weighed glass vials, where it was dried under nitrogen. The crude extracts were weighed and dissolved in methanol to a fixed concentration for LC-MS analysis. The prepared solutions were centrifuged again for 20 min at 4 °C to remove any suspended matters.

Method LC-MS runs on extracts from the *lug-pks* mutant

LC-DAD-HRESIMS spectra were obtained using a Waters Acquity UPLC system, equipped with Waters Acquity PDA, and coupled to a Thermo Instruments MS

system (LTQ XL/LTQ Orbitrap XL). The UPLC system was run using Acquity UPLC HSS T3 C_{18} column (1.8 µm, 100 Å, 2.1 × 100 mm). Solvent A was 0.1% formic acid, 95% H_2O and 5% ACN. Solvent B was 0.1% formic acid, 95% ACN and 5% H_2O. The gradient used was 2% B for 0.5 min, 2-40% for 5.5 min, 40-100% for 2 min, and 100% for 3 min. The flow rate used was 0.5 ml/min. The MS conditions used were: capillary voltage 5 V, capillary temperature 300 °C, auxiliary gas flow rate 5 arbitrary units, sheath gas flow rate 50 arbitrary units, spray voltage 3.5 kV, mass range 100-2000 amu, FT resolution 30000. Spectra were analysed using Thermo Scientific Xcalibur.

RNA isolation

To collect biomass for RNA isolation, 6×10^6 spores of *Streptomyces* sp. QL37 were confluently spread over an MM agar plate supplemented with 0.5 % mannitol and 1 % glycerol, and samples were taken at 24, 48, 36 and 60 hours at 30 °C. For each time point four replica plates were incubated. At each time point 12 small (1.5 × 0.5 cm) agar pieces were cut from the agar from each of the four replicas. The excess agar below the mycelium was sliced off. It was ensured that the vegetative mycelium was still present in the agar. Two pieces of agar were transferred to a 2 ml tube containing 1 × 4 mm and 2 × 3 mm metal beads and were immediately frozen in liquid nitrogen. This led to 6 samples per plate. In total 96 samples were required for the experiment. The samples were stored at -80 °C until RNA isolation. The following protocol was followed-up in four separate rounds for each replica plate of each time point. Each round of isolation included 6 tubes (6 × 2 pieces of agar) per time point. In the final steps of the protocol, the RNA derived from the 6 samples was pooled. After removal of the samples from -80 °C, the samples were homogenized using a TissueLyserII (Qiagen, Hilden Germany). In order to do so, the tubes were transferred to a pre-cooled TissueLyser block. The samples were shaken at 30 Hz for 15 s and this cycle was repeated four times. Between each cycle the machine was stopped for 5 s. When no large pieces were detected in the samples, the tubes were transferred to liquid nitrogen.

To each sample 0.9 ml of pre-warmed (40 °C) lysis buffer (5% (v/v) Triton X-100 and 10 mM EDTA pH 8.0) was added. The samples were vortexed for 30 s. Subsequently 0.9 ml phenol:chloroform (50:50, pH 4.9, VWR) was added. The mixture was vortexed for 1 min, followed by a centrifugation step (15 min, 12,000 × g). The aqueous phase was transferred to a new tube. The phenol:chloroform RNA

extraction was repeated for another two cycles, in order to make sure the sample was clean. After the final phenol:chloroform cleaning step, the aqueous phase was divided over two tubes. The nucleic acids were precipitated by addition of 62.5 µL 1 M Tris-HCL pH 8.0, 25 µL 5 M NaCl and 0.75 ml cold absolute ethanol. After overnight incubation at -20 °C, the samples were centrifuged (5 min at 7,500 × g) and the supernatant was removed. The pellet was washed with 1 ml 80% ethanol. The ethanol was removed, and the pellet was dried at room temperature for 10 min. The pellets derived from the samples that were earlier divided were pooled by dissolving in 87.5 µL nuclease-free water (Ambion). To that, 10 µL DNase buffer and 5 units of DNase I (New England Biolabs) were added, and samples were incubated for 45 min at 37 °C. This was followed by another cleaning step using phenol: chloroform. To do so, nuclease free water was added up to a volume of 200 µL. Then 200 µL phenol:chloroform were added to each sample and vortexed for 1 min. The samples were centrifuged (10 min, 7500 × g) and one other round of phenol:chloroform extraction was executed. The aqueous phase was transferred to a new tube and precipitated by addition of 25 µL 1M Tris-HCl pH 8.0, 10 µL 5M NaCl and 300 µL cold ethanol absolute. This was incubated for at least 10 min at -20 °C. Again, the samples were centrifuged (5 min, 7,500 × g) and the supernatant was removed. The pellet was washed with 200 µL 80% ethanol, dried and subsequently dissolved in 50 µL nuclease free water. The six samples of each different time point were pooled, which led to a final volume of 300 µL RNA. The purity of the RNA was tested using the A260/A280 and A260/A230 ratio using the Nanodrop ND-1000 Spectrophotometer (PEQLAB) and the integrity of the RNA was assessed on a 1.2% 1x TAE gel before sending to Baseclear. The RNA integrity number was estimated using the Bioanalyzer 2100 (Agilent).

Library preparation and RNA-sequencing

rRNA depletion and RNA sequencing were performed by BaseClear. RNA samples were subjected to the rRNA depletion kit using the siTOOLS riboPOOLs Actinobacteria kit. Single-end or paired-end sequence reads were generated using the Illumina NovaSeq 6000 or MiSeq system at a depth of 6–9 million reads each. The sequences generated with the MiSeq system were performed under accreditation according to the scope of BaseClearB.V. (L457; NEN-EN-ISO/IEC 17025).

Quality assessment and analysis of the RNA-seq data

Illumina sequences were trimmed and filtered with FastQC using a threshold of 25 (quality value [Q] > 25). Reads were mapped to the *Streptomyces* sp. QL37 reference genes using the software Bowtie 2 v.2.1.0 (Langmead *et al.*, 2009). The Bioconductor package DESeq2 (Love *et al.*, 2014) was used for normalization and differential expression analyses. The *p*-value was obtained from the differential gene expression test, and False discovery rate (FDR) was used to correct for multiple testing. Accordingly, differentially expressed genes (DEGs) were regarded significant using an FDR adjusted *p*-value of <0.05.

Bioinformatics

For the prediction of the BGCs in the genome sequence of *Streptomyces* sp. QL37 (NZ_PTJS00000000.1) and other *Streptomyces* strains, the bioinformatic tool antiSMASH 5.0 was used (Blin *et al.*, 2019). The bioinformatic tool Clinker was applied to compare the *lug*-like and angucycline and betalactone gene clusters (Gilchrist & Chooi, 2021). As cut-off for the presence of a gene we used a threshold of 40% aa identity between the predicted gene products.

Phylogenetic analysis

A set of 1020 genomes, including 1004 *Streptomyces* and 16 *Kitasatospora* genomes, was downloaded from the NCBI database by querying the fasta files in combination with the taxonomic identifier. 116 unpublished draft genome sequences of an in-house collection of actinomycetes (Zhu *et al.*, 2014b) were added to the set. A maximum-composite likelihood phylogeny was constructed using PhyloPhlAn (Asnicar *et al.*, 2020). Ortholog identification and alignment was performed in Phylophlan using the "-u" command. iTOL (70) was used for the visualization of the phylogenetic tree.

SUPPLEMENTARY TABLES

Table S1 Annotation of the genes in region 12 in the genome of *Streptomyces* sp. QL37 as predicted by antiSMASH (see also Figure S2) (Blin *et al.*, 2019)

Gene	Gene name	Prokka annotation	Gene	gene name	Prokka annotation (Seemann, 2014)
QL37_12745	-	hypothetical protein	QL37_12900	*lugN*	3-hydroxy-5-methyl-1-naphthoate 3-O-methyltransferase
QL37_12750	-	Single-stranded DNA-binding protein 1	QL37_12905	*lugOI*	Anhydrotetracycline monooxygenase
QL37_12755	-	GTPase Era	QL37_12910	*lugF*	Tetracenomycin F2 cyclase
QL37_12760	-	GTPase Era	QL37_12915	*lugA*	Actinorhodin polyketide putative beta-ketoacyl synthase 1
QL37_12765	-	NA	QL37_12920	*lugB*	Actinorhodin polyketide putative beta-ketoacyl synthase 2
QL37_12770	-	Tyrosine recombinase XerC	QL37_12925	*lugC*	Oxytetracycline polyketide synthase acyl carrier protein
QL37_12775	-	hypothetical protein	QL37_12930	*lugD*	Putative ketoacyl reductase
QL37_12780	-	hypothetical protein	QL37_12935	*lugE*	Putative polyketide cyclase
QL37_12785	-	hypothetical protein	QL37_12940	*lugOII*	Anhydrotetracycline monooxygenase
QL37_12790	-	hypothetical protein	QL37_12945	*lugG*	3-oxoacyl-[acyl-carrier-protein] reductase FabG
QL37_12795	-	hypothetical protein	QL37_12950	*lugH*	Methylmalonyl-CoA carboxyltransferase 12S subunit
QL37_12800	-	hypothetical protein	QL37_12955	*lugI*	hypothetical protein
QL37_12805	-	hypothetical protein	QL37_12960	*lugTII*	putative MFS-type transporter EfpA

Table S1 Annotation of the genes in region 12 in the genome of *Streptomyces* sp. QL37 as predicted by antiSMASH (see also Figure S2) (Blin *et al.*, 2019) *(continued)*

Gene	Gene name	Prokka annotation	Gene	gene name	Prokka annotation (Seemann, 2014)
QL37_12810	-	hypothetical protein	QL37_12965	*lugJ*	Chromate reductase
QL37_12815	-	hypothetical protein	QL37_12970	*lugOIII*	Tetracenomycin-F1 monooxygenase
QL37_12820	-	hypothetical protein	QL37_12975	*lugK*	4'-phosphopantetheinyl transferase Npt
QL37_12825	-	hypothetical protein	QL37_12980	*lugRIV*	Transcriptional regulatory protein BaeR
QL37_12830	-	hypothetical protein	QL37_12985	*lugTIII*	Enterobactin exporter EntS
QL37_12835	-	hypothetical protein	QL37_12990	*lugL*	2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase
QL37_12840	-	Agglutinin receptor	QL37_12995	*lugOIV*	putative oxidoreductase YciK
QL37_12845	-	hypothetical protein	QL37_13000	*lugRV*	Regulatory protein AfsR
QL37_12850	-	hypothetical protein	QL37_13005	*lugOV*	Tetracenomycin-F1 monooxygenase
QL37_12855	-	hypothetical protein	QL37_13010	-	Putative phosphoribosyl transferase
QL37_12860	-	Methanol dehydrogenase activator	QL37_13015	-	hypothetical protein
QL37_12865	-	hypothetical protein	QL37_13020	-	hypothetical protein
QL37_12870	*lugM*	Chromate reductase	QL37_13025	-	hypothetical protein
QL37_12875	*lugRI*	HTH-type transcriptional regulator DdrOP3	QL37_13030		hypothetical protein

Table S1 Annotation of the genes in region 12 in the genome of *Streptomyces* sp. QL37 as predicted by antiSMASH (see also Figure S2) (Blin *et al.*, 2019) *(continued)*

Gene	Gene name	Prokka annotation	Gene	gene name	Prokka annotation (Seemann, 2014)
QL37_12880	*lugRII*	Transcriptional regulatory protein LiaR	QL37_13035	-	hypothetical protein
QL37_12885	*lugRIII*	Fatty acid metabolism regulator protein	QL37_13040	-	Ribose import ATP-binding protein RbsA
QL37_12890	*lugTII*	putative transport protein HsrA	QL37_13045	-	Ribose transport system permease protein RbsC
QL37_12895	*lugX*	hypothetical protein		-	

Table S2 Functional annotation of the gene products of the *lug* gene cluster in *Streptomyces* sp. QL37

Protein	Gene Annotation	Aa	Putative function	Nearest homologue	Identity (%)	Coverage (%)	Accession
Orf1	WP_187355741.1	378	Hypothetical protein	*Streptomyces mediolani*	97,4	100	WP_030801858.1
Orf2	WP_104791672.1	264	GntR-like regulator, Nudix hydrolase	*Streptomyces pratensis*	98,5	100	MBD2831978.1
Orf3	WP_203186292.1	143	Hypothetical protein	*Streptomyces pratensis*	84,6	100	MBD2831977.1
LugM	WP_104785779.1	199	NADPH-dependent FMN reductase	*Streptomyces* sp. M3	91.4	100	WP_129263039.1
LugRI	WP_104785781.1	280	XRE family transcriptional regulator	*Streptomyces* sp. M3	88.2	100	WP_129263041.1
LugRII	WP_104785782.1	223	LuxR family transcriptional regulator	*Streptomyces* sp. W007	52.1	97	EHM27498.1
LugRIII	WP_104785784.1	239	TetR family transcriptional regulator	*Streptomyces* sp. L-9-10	63.1	96	RYJ29277.1
LugTI	WP_104785785.1	493	Transporter	*Streptomyces* sp. W007	76	98	WP_032792571.1
LugX	WP_104785787.1	144	Hypothetical protein	---	---	---	---
LugN	WP_104791674.1	346	O-methyltransferase	*Streptomyces* sp. CB00072	71.2	99	WP_073867019.1
LugOI	WP_104785789.1	490	FAD-dependent Monooxygenase	*Streptomyces* scopuliridis	81.4	100	WP_030349271.1
LugF	WP_104785790.1	109	Cyclase	*Streptomyces scopuliridis*	83.	98	WP_030349270.1
LugA	WP_104785792.1	427	Polyketide-α-ketoacyl synthase II	*Streptomyces scopuliridis*	87.3	100	WP_030349269.1
LugB	WP_104785793.1	407	Polyketide-β-ketoacyl synthase	*Streptomyces Scopuliridis*	80.3	99	WP_030349268.1

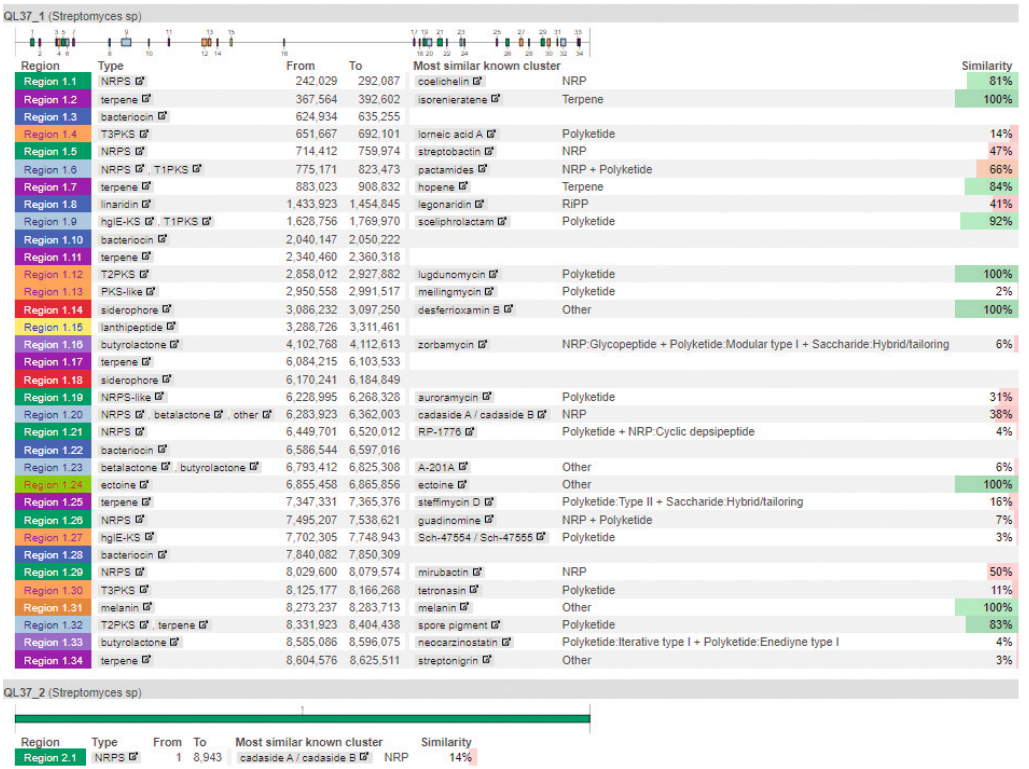
Table S2 Functional annotation of the gene products of the *lug* gene cluster in *Streptomyces* sp. QL37 *(continued)*

Protein	Gene Annotation	Aa	Putative function	Nearest homologue	Identity (%)	Coverage (%)	Accession
LugC	WP_104785795.1	91	Acyl carrier protein	*Streptomyces scopuliridis*	70.8	98	WP_030349267.1
LugD	WP_104785796.1	262	Ketoacyl reductase	*Streptomyces scopuliridis*	85.1	100	WP_030349266.1
LugE	WP_104785798.1	527	Cylclase	*Streptomyces scopuliridis*	81.3	100	WP_030349265.1
LugOII	WP_104785800.1	656	FAD dependent Monooxygenase-SDR family reductase	*Streptomyces* sp. CB00072	76.4	100	WP_073867012.1
LugG	WP_107528480.1	262	Short chain dehydrogenase (SDR)/reductase	*Streptomyces* sp. b84	77.	98	WP_097910307.1
LugH	WP_104791675.1	527	Acyl-coA carboxylase subunit beta	*Streptomyces* sp. w007	87.1	97	WP_050987785.1
LugI	WP_104785801.1	79	Acyl-CoA carboxylase subunit epsilon	*Streptomyces* sp. NBRC 110028	53.8	94	WP_055547361.1
LugTII	WP_104785803.1	417	Transporter	*Streptomyces scopuliridis*	67.7	96	WP_051745235.1
LugJ	WP_104785804.1	211	NAD(P)H dependent FMN reductase	*Streptomyces* sp. Ru71	63.9	90	WP_103781763.1
LugOIII	WP_104785806.1	214	Antibiotic biosynthesis monooxygenase	*Streptomyces* sp. CB00072	62.3	99	WP_073867008.1
LugK	WP_104785807.1	245	4'-phosphopantetheinyl transferase superfamily protein	*Goodfellowiella* sp. AN110305	51.1	94	WP_104785807.1
LugRIV	WP_104785809.1	267	Response regulator transcription factor	*Streptomyces violens*	62	91	WP_078601061.1

Table S2 Functional annotation of the gene products of the *lug* gene cluster in *Streptomyces* sp. QL37 *(continued)*

Protein	Gene Annotation	Aa	Putative function	Nearest homologue	Identity (%)	Coverage (%)	Accession
LugTIII	WP_104785811.1	458	Transporter	*Streptomyces violens*	75.5	84	WP_078601068.1
LugL	WP_104785812.1	307	Alpha/beta hydrolase	*Streptomyces scopuliridis*	69	94	WP_030349255.1
LugOIV	WP_104785814.1	275	SDR family NAD(P)-dependent oxidoreductase	*Streptomyces* sp. TSRI0281	75.6	100	WP_107468115.1
LugRV	WP_104785815.1	646	AfsR/SARP family transcriptional regulator	*Streptomyces scopuliridis*	60.8	93	WP_078490261.1
LugOV	WP_146111254.1	229	Hypothetical protein	*Streptomyces zhaozhouensis*	57.5	90	WP_141514657.1
Orf 4	WP_104785818.1	425	Phosphoribosyltransferase	*Streptomyces* sp. ADI93-02]	93	100	WP_124276433.1
Orf 5	WP_104791676.1	989	Alpha_L-rhamnosidase	*Streptomyces* sp. wa22	93	96	WP_147960798.1
Orf 6	WP_104785820.1	868	glycoside hydrolase family 127	*Streptomyces* sp. TRM S81-3]	96	76.1	WP_188185628.1
Orf 7	WP_104785821.1	1133	glycoside hydrolase family 106 (alpha L-rhamnosidase)	*Streptomyces* sp. TRM S81-3	100	74.6	WP_188185460.1
Orf 8	WP_104785823.1	1092	glycoside hydrolase family 78	*Streptomyces* sp. TRM S81-3	99	77.7	WP_188185461.1
Orf 9	WP_104785824.1	184	GNAT-family acetyltransferase	*Streptomyces* sp. For3	98	81.8	WP_202079774.1
Orf 10	WP_104785825.1	507	Sugar ABC-transport system	*Streptomyces* sp. wa22	100	97.8	WP_147960792.1

SUPPLEMENTARY FIGURES



QL37_1 (Streptomyces sp)

Region	Type	From	To	Most similar known cluster		Similarity
Region 1.1	NRPS	242,029	292,087	coelichelin	NRP	81%
Region 1.2	terpene	367,564	392,602	isorenieratene	Terpene	100%
Region 1.3	bacteriocin	624,934	635,255			
Region 1.4	T3PKS	651,667	692,101	lorneic acid A	Polyketide	14%
Region 1.5	NRPS	714,412	759,974	streptobactin	NRP	47%
Region 1.6	NRPS, T1PKS	775,171	823,473	pactamides	NRP + Polyketide	66%
Region 1.7	terpene	883,023	908,832	hopene	Terpene	84%
Region 1.8	linaridin	1,433,923	1,454,845	legonaridin	RiPP	41%
Region 1.9	hglE-KS, T1PKS	1,628,756	1,769,970	scelphrolactam	Polyketide	92%
Region 1.10	bacteriocin	2,040,147	2,050,222			
Region 1.11	terpene	2,340,460	2,360,318			
Region 1.12	T2PKS	2,858,012	2,927,882	lugdunomycin	Polyketide	100%
Region 1.13	PKS-like	2,950,558	2,991,517	mellingmycin	Polyketide	2%
Region 1.14	siderophore	3,086,232	3,097,250	desferrioxamin B	Other	100%
Region 1.15	lanthipeptide	3,288,726	3,311,461			
Region 1.16	butyrolactone	4,102,768	4,112,613	zorbamycin	NRP:Glycopeptide + Polyketide:Modular type I + Saccharide:Hybrid/tailoring	6%
Region 1.17	terpene	6,084,215	6,103,533			
Region 1.18	siderophore	6,170,241	6,184,849			
Region 1.19	NRPS-like	6,228,995	6,268,328	auroramycin	Polyketide	31%
Region 1.20	NRPS, betalactone, other	6,283,923	6,362,003	cadaside A / cadaside B	NRP	38%
Region 1.21	NRPS	6,449,701	6,520,012	RP-1776	Polyketide + NRP:Cyclic depsipeptide	4%
Region 1.22	bacteriocin	6,586,544	6,597,016			
Region 1.23	betalactone, butyrolactone	6,793,412	6,825,308	A-201A	Other	6%
Region 1.24	ectoine	6,855,458	6,865,856	ectoine	Other	100%
Region 1.25	terpene	7,347,331	7,365,376	steffimycin D	Polyketide:Type II + Saccharide:Hybrid/tailoring	16%
Region 1.26	NRPS	7,495,207	7,538,621	guadinomine	NRP + Polyketide	7%
Region 1.27	hglE-KS	7,702,305	7,748,943	Sch-47554 / Sch-47555	Polyketide	3%
Region 1.28	bacteriocin	7,840,082	7,850,309			
Region 1.29	NRPS	8,029,600	8,079,574	mirubactin	NRP	50%
Region 1.30	T3PKS	8,125,177	8,166,268	tetronasin	Polyketide	11%
Region 1.31	melanin	8,273,237	8,283,713	melanin	Other	100%
Region 1.32	T2PKS, terpene	8,331,923	8,404,438	spore pigment	Polyketide	83%
Region 1.33	butyrolactone	8,585,086	8,596,075	neocarzinostatin	Polyketide:Iterative type I + Polyketide:Enediyne type I	4%
Region 1.34	terpene	8,604,576	8,625,511	streptonigrin	Other	3%

QL37_2 (Streptomyces sp)

Region	Type	From	To	Most similar known cluster		Similarity
Region 2.1	NRPS	1	8,943	cadaside A / cadaside B	NRP	14%

Figure S1 Overview of antiSMASH 5.0 secondary metabolites predicted from the genome of *Streptomyces* sp. QL37. Region 12 contains 100% of the genes of the lugdunomycin BGC, since this cluster is already deposited in the Minimum Information about a Biosynthetic Gene cluster (MIBiG) database(Kautsar *et al.*, 2020).

Figure S2 Region 12 in the genome of *Streptomyces* sp. QL37 contains a type II polyketide gene cluster as predicted by antiSMASH
For further details on this region, see Table S1.

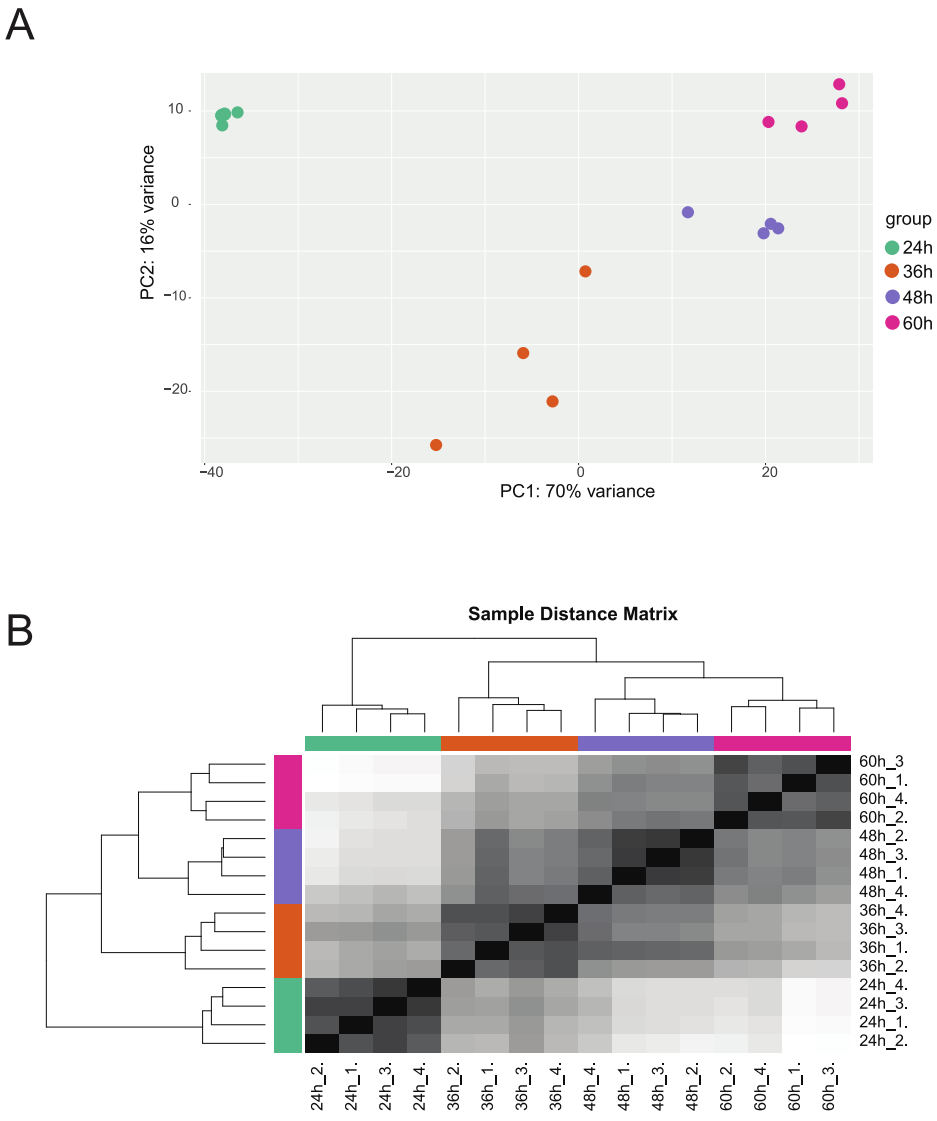


Figure S3 Global analysis of the RNA-seq data of the transcriptome of *Streptomyces* sp. QL37.
RNA was isolated from *Streptomyces* sp. QL37 after 24 hours, 36 hours, 48 hours and 60 hours of growth on MM at 30 °C. A) Principal component analysis (PCA) and B) sample distance matrix of the normalised expression counts of the transcriptome of *Streptomyces* sp. QL37. The samples distance matrix shows the levels of similarity between samples. The square at the intersection of a certain row and a certain column is shaded according to the similarity between the samples. Black indicates a high degree of similarity between two samples, and lighter shades of gray indicate lower levels of similarity (white, no relevant similarity).

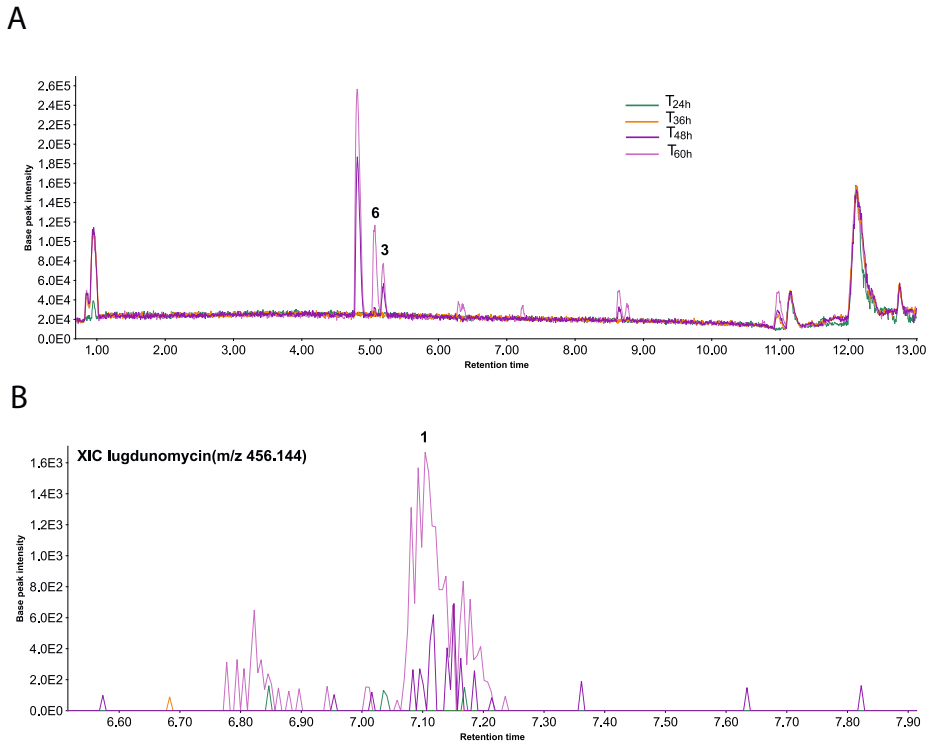


Figure S4 LC-MS chromatogram of the extracts derived from *Streptomyces* sp. QL37 grown on MM for four different time points.
The strain was grown for 24 h, 36 h, 48 h or 60 h on MM agar supplemented with 0.5% mannitol and 1% glycerol. Angucyclines were produced from 48 h (A) and a minute amount of lugdunomycin was identified after 48 and 60 h (B).

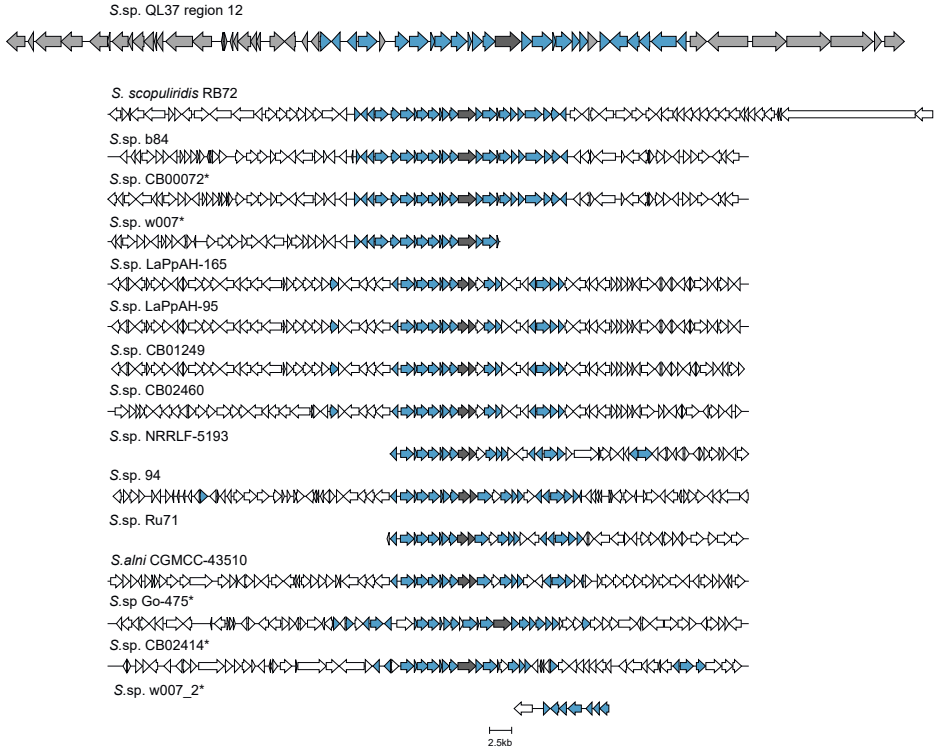


Figure S5 Comparison of region 12 from *Streptomyces* sp. QL37 with *lug*-type gene clusters in other *Streptomyces* strains.
Blue genes are detected in other BGCs, besides the *lug* gene cluster; light grey genes are only detected in the *lug* gene cluster. Dark grey genes indicates either the full length *lugOII*, or the separate oxygenase and reductase encoding genes of which the products show similarity with the oxygenase and reductase domain of LugOII. White genes are genes unrelated tot he *lug* gene cluster. The other 22 *lug*-type gene clusters, indicated in Figure 5 also did not contain the light grey genes (not shown) flanking *lugM* and *lugOV*, indicating these genes mark the borders of the *lug* gene cluster. * indicates known angucycline producers.

The *lug* gene cluster is sufficient for the biosynthesis of angucyclines and limamycins but not lugdunomycin

Helga U. van der Heul
Michiel Uiterweerd
Isabel Nuñez-Santiago
Guillermo Guerrero Egido
Changsheng Wu
Adriaan J. Minnaard
Somayah S. Elsayed
Gilles P. van Wezel

ABSTRACT

Oxidative ring cleavage of the angucycline C-ring expands the chemical diversity of the angucycline family. An exciting example is the angucycline derivative, lugdunomycin, produced by *Streptomyces* sp. QL37. This molecule is characterised by its unusual structural features, including a heptacyclic ring system, a spiroatom, two all-carbon stereocenters, and a benzaza[4,3,3]propellane motif. The final step in the biosynthesis of lugdunomycin was previously proposed to be catalysed by a Diels-Alder reaction between a C-ring cleaved angucycline and *iso*-maleimycin. Heterologous expression of the angucycline biosynthetic gene cluster (BGC; designated *lug*) of *Streptomyces* sp. QL37 resulted in the production of C-ring rearranged angucyclines and limamycins, but not lugdunomycin. In this study we applied a paired-omics strategy to find the missing factor in the biosynthesis of lugdunomycin, using genomics, metabolomics, and temporal RNA-seq profiling. Molecular networking of the metabolome of the wild-type strain and its *pks* mutant identified all molecules that are not derived from the *lug* cluster. Interestingly, *iso*-maleimycin was detected in a *pks* mutant of the *lug* gene cluster, which implies that the molecule is derived from a BGC different from that for the angucyclines. Transcriptome analysis revealed that transcription of the *lug* gene cluster is coordinated with a BGC homologous to the one for maleimycin biosynthesis in *Streptomyces showdoensis* ATCC 15227, and this BGC is thus the major candidate to produce the dienophile for the final step in lugdunomycin biosynthesis.

INTRODUCTION

Streptomyces species are one of the most important resources for bioactive molecules with high chemical diversity (Barka *et al.*, 2016). One of these chemically diverse structures is the angucycline group, the largest molecular class synthesised by polyketide type II synthases (Kharel *et al.*, 2012, Mikhaylov *et al.*, 2021). They exhibit a wide range of antimicrobial and anti-cancer activities possibly due to their variation in chemical modifications of their typical benz[*a*]anthracene backbone, such as glycosylation, epoxidation and hydroxylation (Kharel *et al.*, 2012, Buttner *et al.*, 2018, Schafer *et al.*, 2015, Lehka *et al.*, 2015, Yoon *et al.*, 2019, Fakhruzzaman *et al.*, 2015). The typical angucycline backbone can be drastically modified by oxidative cleavage of one of the aromatic rings which may be followed by incorporation of various other molecules (Fan & Zhang, 2018, Wang *et al.*, 2015, Rix *et al.*, 2003, Tibrewal *et al.*, 2012, Fan *et al.*, 2012b). Previously we discovered the complex angucycline derived molecule, lugdunomycin, characterised by its benzaza[4,3,3]propellane-6-spiro-2'-2H-naphtho[1,8-bc]furan backbone (Wu *et al.*, 2019). The molecule was discovered by screening *Streptomyces* strains derived from remote areas, and is produced by *Streptomyces* sp. QL37 isolated from soil from the Qinling mountains (Figure 1) (Zhu *et al.*, 2014b). However, lugdunomycin is produced in very low amount under all growth conditions, whereas the angucyclines are typically produced in rather high amounts (Wu *et al.*, 2019). It is crucial to increase the production of the molecule to obtain more knowledge on its bioactivity and functional role. To do so, we use a fundamental approach by obtaining insights in the steps required for the biosynthesis of lugdunomycin. Previously we hypothesised a biosynthesis pathway wherein a typical non-rearranged angucycline is oxygenated by a Baeyer-Villiger oxygenase, leading to a C-ring rearranged angucycline (Wu *et al.*, 2019). Subsequently the C-ring rearranged angucycline (diene) reacts with *iso*-maleimycin (dienophile) in a Diels-Alder reaction. Indeed *iso*-maleimycin was also detected in the extracts of *Streptomyces* sp. QL37 (Uiterweerd, 2020). We hypothesised that *iso*-maleimycin is derived from the limamycins via a highly complex biosynthetic route. In chapter 3 we proposed a biosynthetic gene cluster (BGC) that is required for the production of angucyclines, pratensilin A, limamycin and lugdunomycin, *lug*. The functional role of this BGC was confirmed by deletion of the genes involved in the production of the typical angucycline backbone (*lug-pks* mutant). This polyketide type II BGC contains amongst others five genes for oxygenases, which

likely mediate C-ring cleavage (Chapter 5). One approach to determine whether the complete set of genes required for the biosynthesis of a molecule are present in a BGC is heterologous expression in an optimised host such as *S. coelicolor* A3(2) M1152 (Liu *et al.*, 2018, Gomez-Escribano & Bibb, 2011, Izawa *et al.*, 2014). In this strain the BGCs specifying the production of actinorhodin (Act), undecylprodigiosin (Red), Calcium-dependent antibiotic (CDA) and coelimycin P1 (CPK) have been deleted. This removes the possibility that that energy and nutrients are used by other pathways. In this study we introduced a large clone harbouring the *lug* gene cluster from the *Streptomyces* sp. QL37 genome into *S. coelicolor* M1152, producing angucyclines and C-ring cleaved derivatives, as well as limamycins, but not lugdunomycin. Paired-omics using metabolomics, genomics and temporal RNA-seq, suggests that a second BGC is required to produce the dienophile *iso*-maleimycin required for the final step in lugdunomycin biosynthesis.

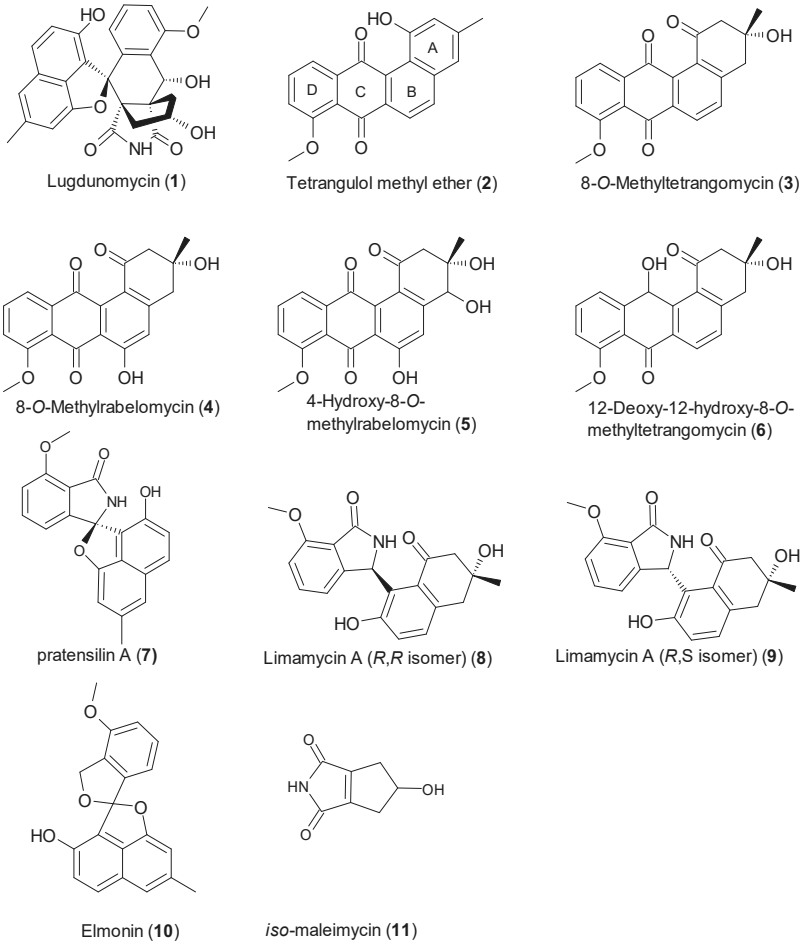


Figure 1 Structures of the compounds discussed in this study.
Lugdunomycin (**1**), the non-rearranged angucyclines (**2**, **3**, **4**, **5**, **6**) and the limamycins (**8**, **9**) were previously isolated from *Streptomyces* sp. QL37 (Wu *et al*., 2019). The structure of compound **7** was later revised to be pratensilin A (Mikhaylov *et al*., 2021). Elmonin (**10**) and *iso*-maleimycin (**11**) were detected in the extracts of *Streptomyces* sp. QL37 by comparison with a standard (Figure S3) (Uiterweerd, 2020).

RESULTS

Heterologous expression of the *lug* gene cluster in *Streptomyces coelicolor* M1152

The *lug* gene cluster of *Streptomyces* sp. QL37 is similar to other angucycline BGCs, but so far, no strain has been reported that produces lugdunomycin and study of several angucycline producers in our laboratory failed to detect the molecule (our unpublished data). A question we therefore sought to resolve is whether the cluster is sufficient to produce lugdunomycin, and if so, if we could enhance its expression. To this end, the cluster was expressed in the angucycline non-producer *S. coelicolor* A3(2) M1152, a derivative of M145 that has been optimised as chassis strain for secondary metabolite production (Gomez-Escribano & Bibb, 2011). In order to do so, a P1-derived artificial chromosome (PAC) library of the genome from *Streptomyces* sp. QL37 was constructed using pESAC-13-A, an *E. coli–Streptomyces* artificial chromosome (Tocchetti *et al.*, 2018, Jones *et al.*, 2013). Three sets of primers were used to define the start (primers 1 and 2), middle (primers 3 and 4) and end (primers 5 and 6) of the BGC, respectively (Table 2). Screening of the library revealed two clones that contained the *lug* gene cluster, one included a 120 kb fragment and the other included a 190 kb fragment of the *Streptomyces* sp. QL37 genome. Both were introduced into *S. coelicolor* M1152. This yielded the strains M1152::*lug*11O and M1152::*lug*3G, respectively. Lugdunomycin was previously isolated from cultures grown on minimal medium (MM) supplemented with 0.5% mannitol and 1% glycerol; pratensilin A (**7**) and the limamycins (**8**, **9**) were isolated from cultures on R5 medium supplemented with 1% mannitol and 0.8% peptone, and angucyclines from both media (Wu *et al.*, 2019). Therefore, the wild-type strain, *S. coelicolor* M1152, strains M1152::*lug*11O and M1152::*lug*3G were grown on both MM and R5 agar plates for seven days at 30 °C.

S. coelicolor M1152::*lug* ex-conjugants produced a brown-yellow pigment on MM and R5 agar plates, whereas *S. coelicolor* M1152 without the cluster did not produce this pigment (Figure 2A). The brown-yellow pigment production by the ex-conjugants containing the *lug* gene cluster suggests angucycline production. Therefore, extracts were prepared from these cultures, followed by LC-MS analysis, where angucyclines production was confirmed in both media (Figure 2B and Figure S1) On R5 agar, angucyclines (**2**, **3**, **4**, **6**), pratensilin A (**7**) and limamycins (**8**,**9**)

were identified in the heterologous host. This indicates that the introduced clone contained all the genes required for angucycline, pratensilin A and limamycin production. Angucyclines were also produced in MM-grown cultures, while lugdunomycin could not be detected under either growth condition. A peak was observed with a monoisotopic mass of 456.14, corresponding to the $[M+H-H_2O]^+$ ion for lugdunomycin, and with the same retention time. However, its MS/MS spectrum did not match that of lugdunomycin, and the mass of the $[M+H]^+$ ion could not be simultaneously observed (data not shown). We therefore conclude that the *lug* gene cluster alone is not sufficient to enable lugdunomycin production by *S. coelicolor* M1152.

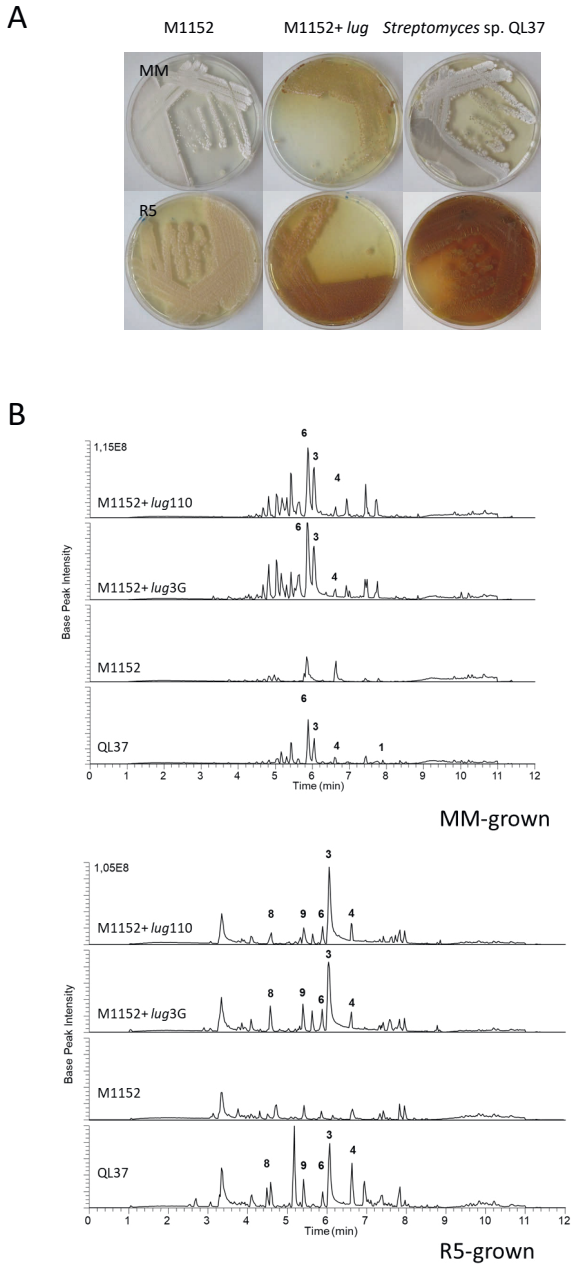


Figure 2 Heterologous expression of the *lug* gene cluster in *S. coelicolor* M1152.
A) *S. coelicolor* M1152 (M1152), *S. coelicolor* M1152::*lug*11O (M1152+*lug*11O), *S. coelicolor* M1152::*lug*3G (M1152+*lug*3G), and wild-type *Streptomyces* sp. QL37 (QL37) grown on MM (top) and R5 (bottom) agar plates. B) LC-MS chromatograms of the crude extracts of the cultures shown in A. Compound numbers are given over their corresponding peaks (Figure 1) As expected M1152 itself did not produce any of the angucycline derived

compounds. *Streptomyces* sp. QL37 produced the non-rearranged angucyclines (**2**,**3**,**4**,**6**), the rearranged angucyclines (**7**,**8**,**9**) and lugdunomycin (**1**). The heterologous expression strains M1152+*lug*11O and M1152+*lug*3G produced the (non)-rearranged angucyclines (**2**–**4**, **6**, **7**–**9**), but did not produce lugdunomycin (**1**) (see also Figure S1).

Metabolome analysis for *Streptomyces* sp. QL37 and its *lug-pks* mutant

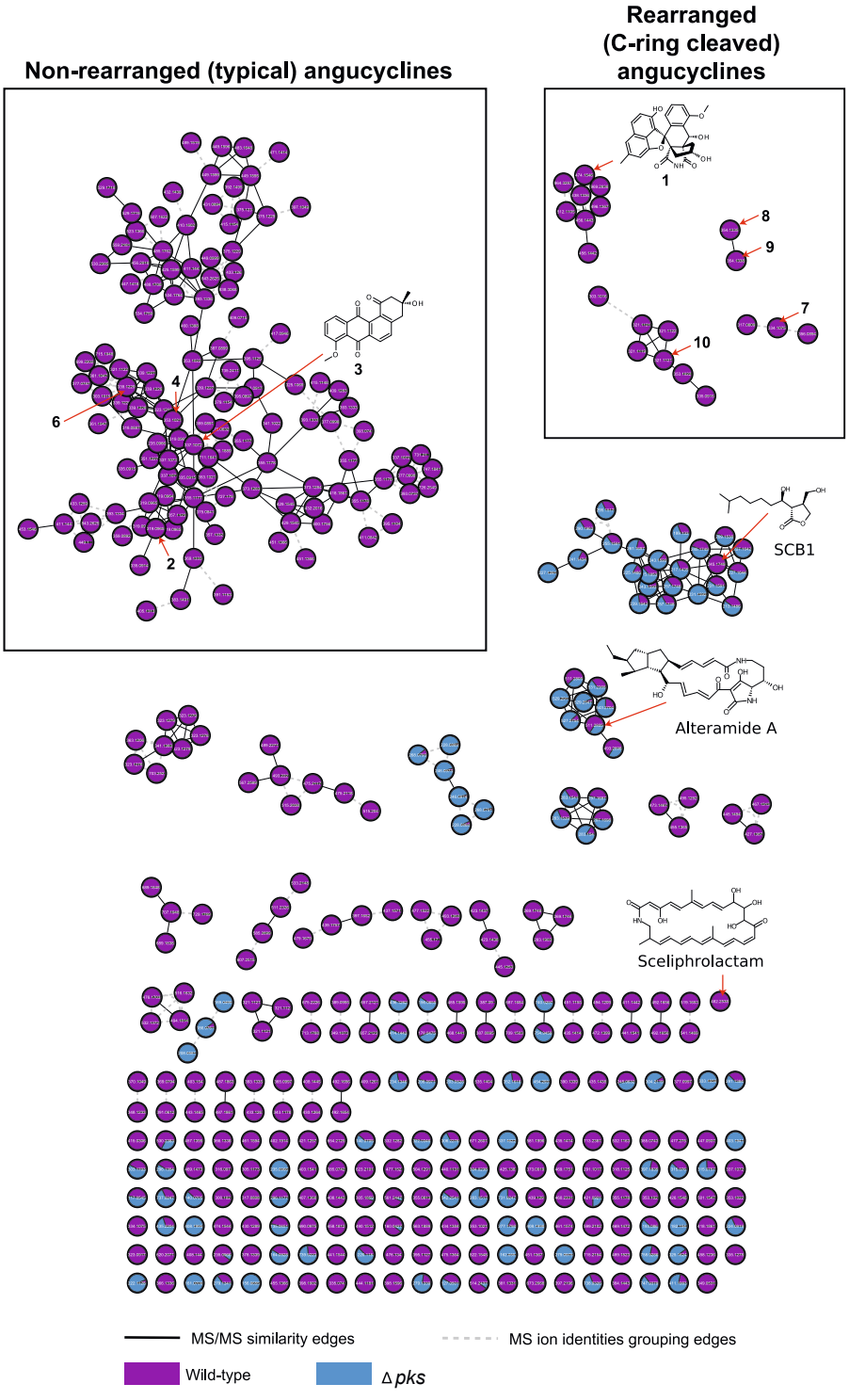
We then wanted to see which metabolites are produced in a *lug-pks* null mutant, which cannot produce angucycline-related molecules (Chapter 3). Molecular networking was employed to analyse the LC-MS data generated from the extracts of the wild-type strain and its *lug-pks* mutant cultures grown for seven days on MM agar plates (Figure 3) This technique enables clustering of structurally related metabolites based on similarities in their MS/MS spectra (Nothias *et al.*, 2020, Schmid, 2020). A molecular network was generated using the Global Natural Products Social Molecular Networking (GNPS) platform, and was visualised using Cytoscape (Kohl *et al.*, 2011, Wang *et al.*, 2016, Nothias *et al.*, 2020, Schmid, 2020). In the wild-type strain, angucyclines (**2**–**4**,**6**), lugdunomycin, pratensilin A (**7**), limamycins (**8**,**9**) and lugdunomycin (**1**) were identified, which clustered as separate molecular families (Figure 3) Additionally, a structurally related isomer of lugdunomycin, with the same mass but different retention time, could be identified for the first time (Figure S2). Possibly *iso*-maleimycin can react in different manners with the angucycline diene, leading to a variety of lugdunomycin isomers (M. Uiterweerd and A.J. Minnaard, unpublished data). We also identified elmonin or its enantiomer oleaceran, a rearranged angucycline, through matching one of the nodes with the chemically synthesised elmonin (**10**) (Figure S3). This molecule is a C-ring cleaved angucycline, which also clustered separately from the non-rearranged angucyclines (Yixizhuoma *et al.*, 2017, Raju *et al.*, 2013). Elmonin may be the direct precursor of the diene that reacts with *iso*-maleimycin in the final Diels-Alder reaction. All the ions in these molecular families were exclusively present in the wild-type strain.

Besides angucycline-related molecules, we detected a molecular family that is related to γ-butyrolactones (GBLs) and one that is related to tetramate macrolactams. The annotation of the latter two was based on manual comparison of measured MS and MS/MS spectra with those previously reported for SCB1 and alteramide A respectively (Figure S4 and Figure S5) (Yang *et al.*, 2005, Shaikh *et al.*, 2021, Moree *et al.*, 2014). Additionally, the genome of *Streptomyces* sp. QL37 contains three predicted BGCs for GBLs (BGC16, -23 and -33) and a predicted

tetramate macrolactams one (BGC6), in agreement with the annotation. The monoisotopic mass of sceliphrolactam was also detected (Figure S6), supported by the presence of a sceliphrolactam BGC (BGC9). For details on the predicted BGCs in *Streptomyces* sp. QL37 I refer to Chapter 3. All the non-angucycline metabolite families were also identified in the *lug-pks* null mutant, with a general tendency of being upregulated in the mutant. This could be due to the enhanced availability of precursors for the other biosynthetic pathways given that the highly abundant angucyclines are not produced anymore.

The production of lugdunomycin was proposed to proceed through a Diels-Alder [4+2] cycloaddition involving a dienophile which is called *iso*-maleimycin (Wu *et al.*, 2019, Uiterweerd, 2020). It was not possible to detect the *iso*-maleimycin moiety using LC-MS, however it could be readily detected using GC-MS (Uiterweerd, 2020). Accordingly, the extracts of *Streptomyces* sp. QL37 wild-type strain and its *lug-pks* mutant were additionally analysed using GC-MS. Interestingly, *iso*-maleimycin could be detected in the extracts of both the wild-type strain (Uiterweerd, 2020) and its *pks* mutant (Figure 4), suggesting that it is not derived from the *lug* gene cluster. This contrasts with the previous hypothesis that *iso*-maleimycin may be derived from limamycins (Wu *et al.*, 2019).

Figure 3 ***(right)*** Molecular network of the ions detected in the extracts of *Streptomyces* sp. QL37 and its *lug*-pks mutant when grown on MM.
The nodes are labelled by the precursor mass of their ions and pie charts are mapped to the nodes to indicate the relative intensities of the ions in the different samples. The dashed edges in the network connect different ion species of the same molecule and the solid edges connect nodes based on their MS/MS similarity. Arrows highlight metabolites previously identified in *Streptomyces* sp. QL37 (Figure 1).


Non-rearranged (typical) angucyclines
Rearranged
(C-ring cleaved)
angucyclines
1
2
3
4
6
7
8
9
10
SCB1
Alteramide A
Sceliphrolactam
MS/MS similarity edges
MS ion identities grouping edges
Wild-type
Δpks

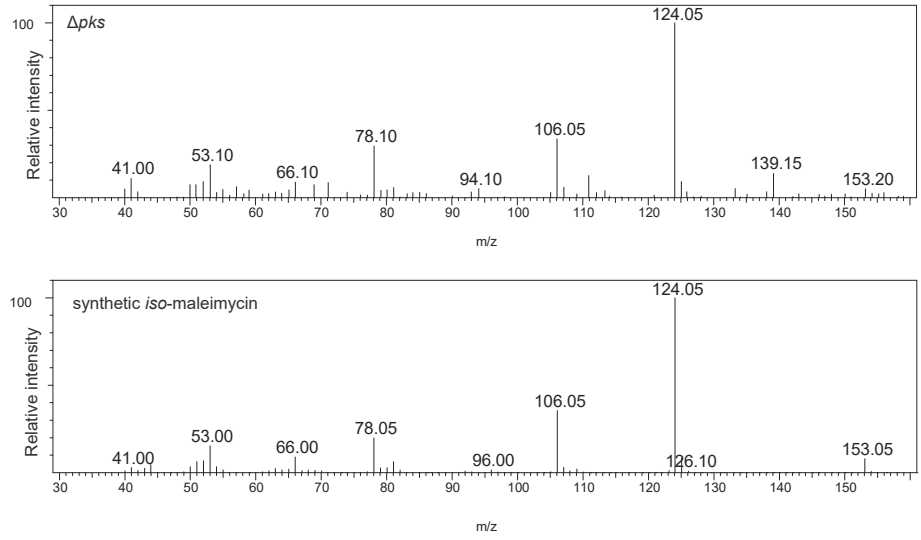


Figure 4 Comparison of the GC-MS spectrum of synthetic iso-maleimycin and the corresponding peak in the extract of the *lug*-pks mutant of *Streptomyces* sp. QL37.
The mutant was grown for seven days on MM at 30°C. Extracts were prepared using ethyl-acetate.

Transcriptional analysis of the BGCs of *Streptomyces* QL37

To study the expression of the different BGCs of *Streptomyces* sp. QL37, and their temporal correlation, RNA-sequencing (RNA-seq) was performed on RNA isolated from *Streptomyces* sp. QL37 grown for different incubation times (24 h, 36 h, 48 h and 60 h) on MM at 30°C. (Schorn *et al.*, 2021). For more details on the RNA-seq data see Chapter 3. The genome of *Streptomyces* sp. QL37 was analysed using antiSMASH, which predicted the presence of in total 35 BGCs. Four hybrid BGCs were recovered, including BGC 6 (Non-ribosomal peptide synthetase (NRPS), T1PKS), BGC9 (hlgE-KS and T1PKS), BGC20 (NRPS, β-lactone and other), BGC23 (β-lactone and γ-butyrolactone) and 32 (T2PKS and terpene). For details on the antiSMASH outcome, see Chapter 3, Figure S1. To define the expression level of each BGC at each time point, the average expression level of its core biosynthetic genes was calculated, such as NRPS, ketosynthase or terpene cyclase genes. Table S1A and S1B show the genes that were assigned as core genes. For some of the hybrid BGCs (BGC 9, BGC 23 and BGC 32) the different BGC types were clearly separated over two fragments of the hybrid BGC, suggesting they could each have a different expression pattern. Therefore, the average expression level

was calculated for each BGC type in these hybrid BGCs (with each type designated as region "a" an "b", Table S1A and B and Figure S7).

A heatmap with hierarchical clustering analysis using Ward's method was generated to analyse the similarities between the expression of each BGC based on the expression level and the expression pattern over time (Figure 5). A corresponding bar plot with error bars is indicated in Figure S8. BGCs that showed a similar expression pattern and level as the *lug* gene cluster (BGC12) were BGC32a (for a spore pigment), BGC6 (for a polycyclic tetramate macrolactam), and BGC23a (for a β-lactone). BGC32a is most likely responsible for the production of the black spore pigment. BGC6 produces polycyclic tetramate macrolactams and these were abundant in the extracts of the wild-type strain and its *pks* mutant (Figure 3).

Other BGCs that were highly expressed in the transcriptome of *Streptomyces* sp. QL37 were BGC7 (for a terpene), BGC9b (for sceliphrolactam), BGC21 (for a nonribosomal peptide), and BGC32b (for a terpene). BGC9b is predicted to produce sceliphrolactam, and was detected in the extracts of the wild-type strain. BGC21 is a likely candidate for the biosynthesis of a novel lipopeptide that was identified in the extracts of the *lugOI* mutant (Chapter 5). The GBLs that were identified in the crude extracts might be derived from BGC16 and/or BGC23b, both of which are predicted GBL BGCs and showed similar expression levels.

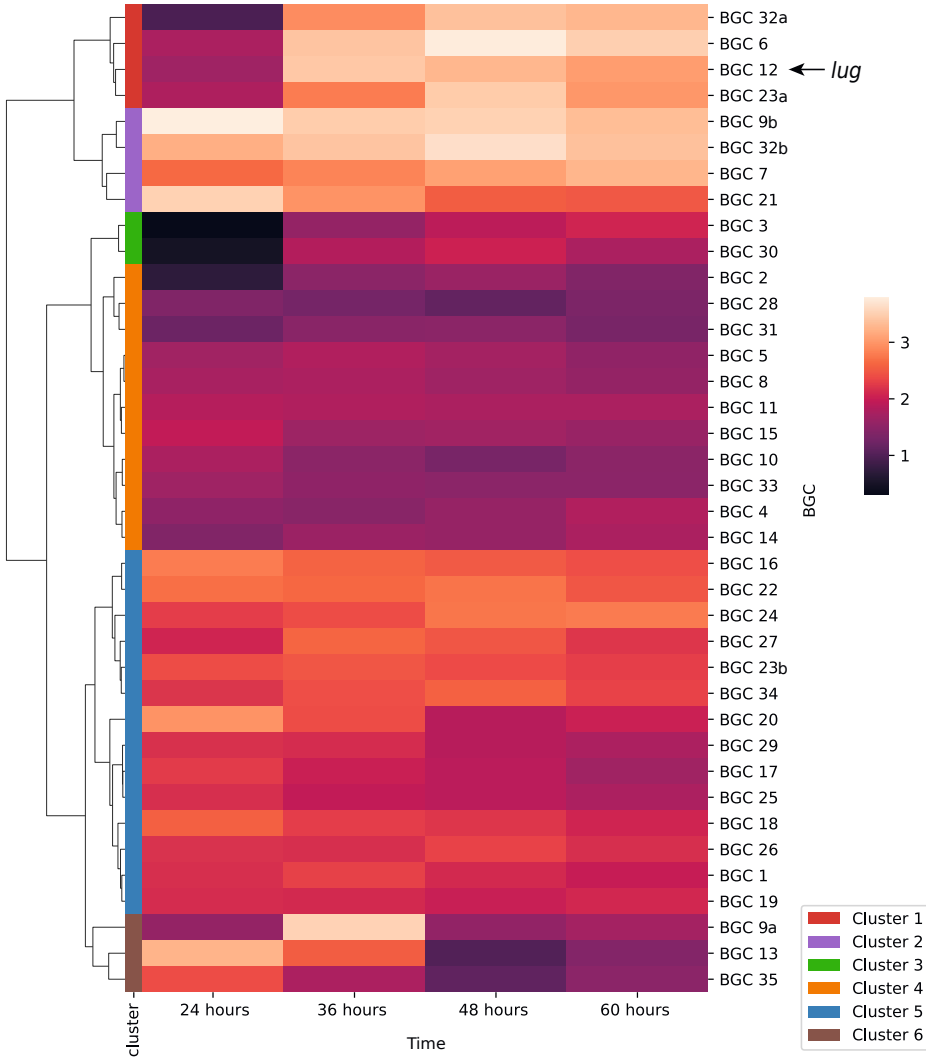


Figure 5 Hierarchical clustering analysis and heatmap of the average expression level of the core genes of the BGCs in the genome of *Streptomyces* sp. QL37.
Heatmap showing the log10 of the absolute expression levels of the core genes in each BGC predicted by antiSMASH at the different measured time points (T_{24h}, T_{36h}, T_{48h} and T_{60h}). The lug gene cluster (BGC12) is relatively highly expressed together with the *iso*-maleimycin gene cluster (BGC23a), the alteramide A BGC (BGC6) and the spore pigment BGC (BGC32a) (Refer to Table S1 for the core genes used in the calculation, and Figure S7 for the annotation of regions “a” and “b”).

Possible role of BGC 23 in lugdunomycin biosynthesis

One of the BGCs co-expressed with the *lug* gene cluster was BGC 23a (Figure 5). This BGC is annotated as a β-lactone BGC by antiSMASH, due to the presence of a gene encoding a 2-isopropylmalate synthase and a gene encoding an AMP-ligase, similar as to the β-lactone BGCs for belactosin and cystargolide ((Wolf *et al.*, 2017a) and Serina Robinson, pers. comm.). In addition, BGC 23 encodes an orthologue of LysW, a carrier protein required for the synthesis of lysine (Matsuda *et al.*, 2017). Recently it was shown that these proteins are also required for the biosynthesis of non-proteogenic building blocks of natural products, such as for vazabitide A (Hasebe *et al.*, 2016). One member of this family of BGCs is required for the biosynthesis of maleimycin ((Makato, 2012-2017, Prima *et al.*, 2017)), produced by *S. showdoensis* ATCC 15227 (Elstner *et al.*, 1973) . BGC23 is highly similar to this likely maleimycin BGC (Figure 6, Table S3). This is highly suggestive, considering that the closely related *iso*-maleimycin was proposed to be involved in lugdunomycin biosynthesis (Uiterweerd, 2020, Wu *et al.*, 2019). Preliminary mutational analysis revealed that neither lugdunomycin nor *iso*-maleimycin could be detected in mutants of BGC23a, whereas the angucyclines, pratensilin A, elmonin and the limamycins were still produced (I. Nuñez-Santiago, S.S. Elsayed and G. P. van Wezel, unpublished data). Taken together, our data show that the *lug* gene cluster is responsible for the biosynthesis of all angucyclines and derivatives as well as limamycins, which was further corroborated by heterologous gene expression. However, the final step for lugdunomycin biosynthesis likely requires a second BGC.

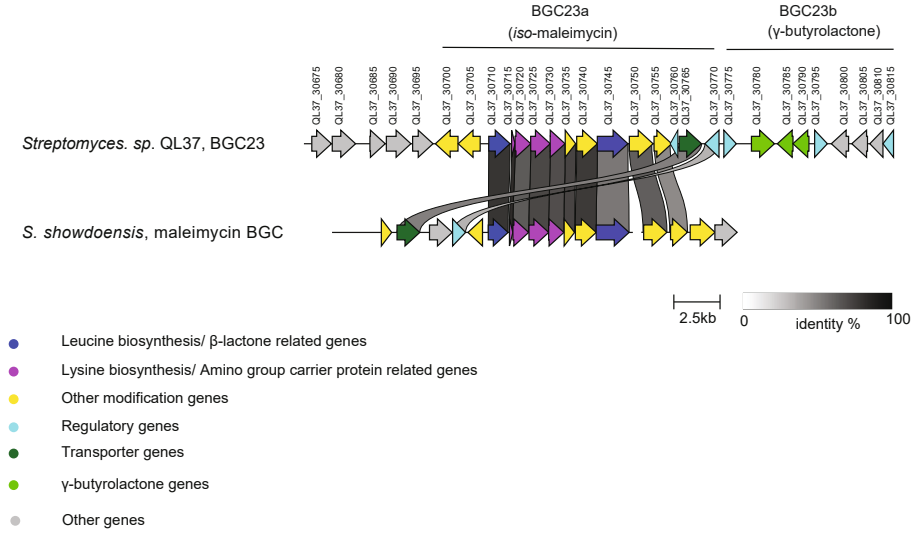


Figure 6 Comparison of BGC23 from *Streptomyces* sp. QL37 with the likely maleimycin BGC cluster in the genome of the maleimycin producer, S. *showdoensis* ATCC 15227. The strokes between the BGCs link genes that share at least 40% identity. Refer to Table S2 for information about the annotation of each gene. The orthologues of the genes from QL37-30750 and QL37-30755 of *Streptomyces showdoensis* were detected on a different scaffold.

DISCUSSION

Streptomyces sp. QL37 produces the highly rearranged angucycline derivative lugdunomycin (Wu *et al.*, 2019). A biosynthesis pathway was proposed wherein a C-ring rearranged angucycline acts as a diene and reacts with *iso*-maleimycin as a dienophile in a Diels-Alder [4+2] cycloaddition to form lugdunomycin. (Uiterweerd, 2020, Wu *et al.*, 2019). We previously hypothesised that *iso*-maleimycin is derived from the limamycins (Wu *et al.*, 2019). Our current study shows that in fact the proposed *lug* gene cluster is an angucycline BGC that only specifies angucycline, pratensilin A, limamycin production, but is not sufficient for lugdunomycin production. An additional BGC is likely required for the biosynthesis of *iso*-maleimycin.

Heterologous expression of the *lug* gene cluster and flanking regions in *S. coelicolor* M1152 led to the biosynthesis of angucyclines and their C-ring rearranged derivatives (pratensilin A and the limamycins), while no lugdunomycin was detected. Therefore, we thoroughly examined the metabolome of *Streptomyces*

sp. QL37 using molecular networking. To do so, we compared the metabolome of the wild type and its *pks* mutant, which does not produce any angucyclines, to deviate between angucycline-related and non-angucycline related molecules (Chapter 3). Besides angucycline-related molecules, *Streptomyces* sp. QL37 also produces alteramide A, sceliphrolactam and γ-butyrolactones (GBLs), which were mainly produced by the *pks* mutant. Interestingly, we found that *iso*-maleimycin is produced by a *lug-pks* mutant that lacks the minimal PKS genes of the lugdunomycin BGC, suggesting the molecule is not derived from the limamycins as previously anticipated, and perhaps from another BGC.

Temporal RNA-seq on the transcriptome of *Streptomyces* sp. QL37 under conditions in which lugdunomycin is produced revealed BGCs that were co-expressed with the *lug* gene cluster, including BGCs for alteramide A, sceliphrolactam and GBLs, in line with the metabolomics data. In addition, BGC23a showed very good co-expression with the *lug* gene cluster; this BGC contains amongst others β-lactone related genes, and AmCP-related genes. Additionally, BGC23 is very similar to a BGC for maleimycin biosynthesis in *S. showdoensis* (Makato, 2012-2017, Wolf *et al.*, 2017a, Prima *et al.*, 2017). This data, together with preliminary mutational studies (I. Nuñez-Santiago, S.S. Elsayed and G. P. van Wezel, unpublished data), makes it highly likely that BGC12 (for lugdunomycin) and BGC23 (for *iso*-maleimycin) are both required to produce lugdunomycin. It is important to note that *S. coelicolor* lacks a cluster homologues to BGC23, and this likely explains why the heterologous expression of the *lug* gene cluster in *S. coelicolor* M1152 did not result in lugdunomycin biosynthesis.

We propose that the final Diels-Alder reaction between *iso*-maleimycin and the C-ring rearranged angucyclines is spontaneous, based on the racemic mixture that is produced (Wu *et al.*, 2019). If our assumptions are correct, introduction of the *lug* gene cluster in an *iso*-maleimycin producer (or vice versa) could facilitate heterologous lugdunomycin production. This is currently being investigate in our laboratory.

We here argue that BGC12 and BGC23a collaborate to produce lugdunomycin, via the joining of metabolites from the angucycline and *iso*-maleimycin pathways. Such joining of two pathways is an exciting new theme, and is something that requires more attention. Recent examples discovered in our laboratory are

the production of endophenasides in *Kitasatospora* sp. MBT66, whereby the glycosylation of phenazines by a methyl-rhamnose moiety is encoded by a BGC for the macrolide leucanicidin, and the biosynthesis of actinomycin L in *Streptomyces* sp. MBT27, whereby an anthranilamide unit, derived from primary metabolism, is incorporated into the actinomycin core (Wu *et al.*, 2016b, Machushynets *et al.*, 2022). As shown in the current work, time- and cost-intensive synthetic biology or heterologous expression approaches on a single BGC will not be successful if the final product is in fact governed by more than one BGC, so this is a concept that needs more attention in the field of microbial drug discovery.

A question that remains to be answered is what the environmental signals are that activate lugdunomycin biosynthesis and how the expression of the BGCs of *Streptomyces* sp. QL37 is coordinated. Perhaps more importantly, it will be important to see whether lugdunomycin offers an ecological advantage to the producer and if so, under which circumstances, and what its bioactivity is. Answers to these questions will shed further light on the biosynthesis, the biological role, and the possible application of this intriguing molecule.

Acknowledgments

We thank Serina Robinson and Marnix Medema for discussions on *iso*-maleimycin biosynthesis. We thank Du Chao for his assistance in creating Figure 5. The work was supported by NACTAR grant 16439 from the Netherlands Organization for Scientific research (NWO).

MATERIALS AND METHODS

Bacterial strains and growth conditions

Streptomyces sp. QL37 was obtained from the MBT strain collection (Zhu *et al.*, 2014b). The strain was isolated from soil in the Qinling mountains (P. R. China) as described previously and was deposited to the collection of the Centraal Bureau voor Schimmelcultures (CBS) in Utrecht, the Netherlands, under deposit number 138593 (Wu *et al.*, 2019). *Streptomyces coelicolor* M1152 was obtained from the John Innes Centre. *S. coelicolor* M1152 lacks genes that are required for prodigiosin, coelimycin P1, calcium dependent antibiotic, and actinorhodin biosynthesis. In additions it carries mutations in the gene *rpoB* for the RNA polymerase β subunit,

which often promotes antibiotic production (Hu *et al.*, 2002, Gomez-Escribano & Bibb, 2011). Strains used in this study are described in Table 1.

Streptomyces were grown on soya flour mannitol (SFM) for seven days at 30 °C, then spores were collected and stored in 20% glycerol at -20 °C (Kieser *et al.*, 2000). For conjugation *Streptomyces* sp. QL37 was grown on SFM containing 60 mM $MgCl_2$ and 60 mM $CaCl_2$ (Wang & Jin, 2014). Ex-conjugants were selected using thiostrepton (20 µg/ml) and apramycin (50 µg/ml) (Wu *et al.*, 2019).

Heterologous expression of the *lug* gene cluster in *Streptomyces coelicolor* M1152

The pESAC13-A library was prepared by the company BioS&T (Montreal, Canada) as previously described (Jones *et al.*, 2013). For the preparation of the genomic library, genomics DNA of *Streptomyces* sp. QL37 was isolated and partially digested with BamHI. These fragments were cloned into the BamHI-digested and dephosphorylated pESAC13-A. This vector is a Phage P1 *E.coli-Streptomyces* Artificial Chromosome containing an *oriT* for conjugation, a phiC31 integrase and an *attP* site for integration into the *attB* recombination locus of the *Streptomyces* genome (Sosio *et al.*, 2000, Jones *et al.*, 2013). This vector is derived from the *Streptomyces* pPAC-S1 (Sosio *et al.*, 2000). The QL37 library contained 3840 clones with an average insert size of 112 kb. PCR primers (1-6) were developed for screening of the PAC library for positive clones containing the *lug* gene cluster (Table 2). PCR screening identified two clones that contain the *lug* cluster (11O of around 120 kb and 3G of around 190 kb).

DH10β cells containing the PAC clones were used to introduce the *lug* cluster into the heterologous strain *S. coelicolor* M1152. First the *lug* cluster was introduced in the methylase deficient strain ET12567 by tri-parental mating including the strains *E. coli* ET12567, ET12567 containing the driver plasmid pUB307 and DH10β containing *lug* (PAC11O and PAC3G) according to A.C. Jones and colleagues (Jones *et al.*, 2013). ET12567 transformants containing both the PAC construct and the driver plasmid pUB307 were selected on LB agar containing kanamycin (50 µg/ml), chloramphenicol (25 µg/ml) and apramycin (50 µg/ml). Colonies was selected and glycerol stocks were prepared of these strains. The three PCR primer pairs (1-6) were used to confirm that the strains contained the vector. These fragments

Table 1 Bacterial strains and plasmids used in this study

Strains		
Name	Description	Reference
***Streptomyces* sp. QL37**	Wild type, CBS 138593	(Wu *et al*., 2019), (Zhu *et al*., 2014b)
Streptomyces* sp. QL37Δ *lug-pks	Deletion mutant of the *lugA–OII* genes (for the minimal PKS), Apra^R.	(Wu *et al*., 2019), Chapter 3
***Streptomyces coelicolor* M1152**	*Streptomyces coelicolor* M145 Δ*act* Δ*red* Δ*cpk* Δ*cda rpoB*(C1298T), conferring resistance to rifampicin	(Gomez-Escribano & Bibb, 2011)
***S. coelicolor* M1152::lug11O**	*S. coelicolor* M1152 containing fragment of 120 kb from genomic DNA of *Streptomyces* sp. QL37 including the *lug* gene cluster; Apra^R, Thio^R.	This study
***S. coelicolor* M1152::lug3G**	*S. coelicolor* M1152 containing fragment of 190 kb from genomic DNA of *Streptomyces* sp. QL37 including the *lug* gene cluster; Apra^R, Thio^R	This study
***E. coli* ET12567**	Methylation deficient host (*dam*-13::Tn9 *dcm*-6 *hsdM hsdS hsdR* Cm^R).	(MacNeil *et al*., 1992)
***E. coli* ET12567/ pUB307**	*E. coli* ET12567 harbouring pUB307.The strain is used for conjugal transfer. It is a derivative of RP1, and contains *oriT*; Cm^R, Tet^R, Km^R.	(Jones *et al*., 2013)
Plasmids		
Name	Description	Reference
pESAC13-A	Used for the creation of *Streptomyces* genomic library by BioS&T. The vector is derived from pPAC-S1 and contains the *oriT* from RK2 for conjugative transfer to *E. coli* and the *phiC31 attP* sites for integration into the *Streptomyces* genome; Thio^{R}and Apra^R.	[34, 43]

were sequenced to verify that the PCR products correspond to the fragments of the *lug* gene cluster.

Conjugation of the PAC constructs was executed according to the protocol of T. Kieser and colleagues(Kieser *et al.*, 2000). After mixing *E. coli* and *Streptomyces* the strains were spread on an SFM plate supplemented with 10 mM $MgCl_2$ and incubated at 30 °C. After 20 hours the plates were overlayed with nalidixic acid (10 µg/ml) and apramycin (50 µg/ml) and incubated for another 6 days. Ex-conjugants were selected streaked onto new plates to remove *E. coli* and to obtain pure isolates. To obtain spore suspensions the strains were grown on SFM medium containing the selection marker apramycin. The mycelium was scraped off the plate and re-suspended into 20% glycerol. The integration was verified using the primers 1-6 (Table 2).

Table 2 Primers used in this study

Number	Name	Sequence 5' --> 3'
1	StartLug+1_Fw	ACGGACAACGCCTTCTCCGCTGAAC
2	StartLug+336_Rv	CTGCCTACGCCGACGTTGTCACCG
3	MiddleLug+1_Fw	CCGACCGTACCGAAGGCATGATCTC
4	MiddleLug+477_Rv	CGGTAGCCCTCGCAGTGGGTGATG
5	EndLug+1 _Fw2	TATCCGGGTCAAGCGGATCAAGCTG
6	EndLug+425_Rv2	TGGCGGCGTGTGCCTTTCGGAGAG

Natural products extraction

Streptomyces strains were confluently grown on minimal medium (MM) agar plates (25 ml) containing 0.5% mannitol + 1% glycerol and on R5 containing 1% mannitol and 0.8% peptone (Wu *et al.*, 2019). The same agars as described in Chapter 3 were used for the preparation of the media. The experiment was conducted in triplicate. After seven days of growth at 30 °C the agar plates were cut into small pieces and soaked in 25 ml of ethyl acetate for 12 hours. Subsequently the ethyl acetate was decanted and evaporated at room temperature. This process was repeated twice. The dried extract was re-dissolved in methanol (MeOH) and centrifugated to remove any undissolved material. Subsequently the MeOH solutions were transferred to new pre-weighed glass vials, and the solvent dried under ambient conditions (heterologous expression samples) or under nitrogen

(GNPS networking). The crude extracts were weighed and dissolved in methanol to a fixed concentration for LC-MS analysis. The prepared solutions were centrifuged again for 20 min at 4 °C to remove any suspended matters. For the detection of *iso*-maleimycin using GC-MS in the metabolome of the *pks* mutant, extracts were prepared in a similar manner as for the detection of angucyclines, limamycins and lugdunomycin (Uiterweerd, 2020). However, extracts were not re-dissolved in methanol, because *iso*-maleimycin degrades in this solvent.

LC-MS methods

Extracts from S. coelicolor ***(heterologous gene expression)***
LC-DAD-HRESIMS spectra were obtained using a Waters Acquity UPLC system, equipped with Waters Acquity PDA, and coupled to a Thermo Instruments MS system (LTQ XL/LTQ Orbitrap XL). The UPLC system was run using Acquity UPLC HSS T3 C_{18} column (1.8 µm, 100 Å, 2.1 × 100 mm). Solvent A was 0.1% formic acid, 95% H_2O and 5% ACN. Solvent B was 0.1% formic acid, 95% ACN and 5% H_2O. The gradient used was 2% B for 0.5 min, 2-40% for 5.5 min, 40-100% for 2 min, and 100% for 3 min. The flow rate used was 0.5 ml/min. The MS conditions used were: capillary voltage 5 V, capillary temperature 300 °C, auxiliary gas flow rate 5 arbitrary units, sheath gas flow rate 50 arbitrary units, spray voltage 3.5 kV, mass range 100-2000 amu, FT resolution 30000. Spectra were analysed using Thermo Scientific Xcalibur.

Extracts for molecular networking
LC-MS/MS acquisition was performed using Shimadzu Nexera X2 UHPLC system, with attached PDA, coupled to Shimadzu 9030 QTOF mass spectrometer, equipped with a standard ESI source unit, in which a calibrant delivery system (CDS) is installed. The dry extracts were dissolved in MeOH to a final concentration of 1 mg/ml or 0.5 mg/ml, and 2 µL were injected into a Waters Acquity HSS C_{18} column (1.8 µm, 100 Å, 2.1 × 100 mm). The column was maintained at 30 °C, and run at a flow rate of 0.5 ml/min, using 0.1% formic acid in H_2O as solvent A, and 0.1% formic acid in acetonitrile as solvent B. A gradient was employed for chromatographic separation starting at 5% B for 1 min, then 5–85% B for 9 min, 85–100% B for 1 min, and finally held at 100% B for 3 min. The column was re-equilibrated to 5% B for 3 min before the next run was started.

All samples were analysed in positive polarity, using data dependent acquisition mode. In this regard, full scan MS spectra (*m/z* 100–1700, scan rate 10 Hz, ID enabled) were followed by two data dependent MS/MS spectra (*m/z* 100–1700, scan rate 10 Hz, ID disabled) for the two most intense ions per scan. The ions were selected when they reach an intensity threshold of 1500, isolated at the tuning file Q1 resolution, fragmented using collision induced dissociation (CID) with fixed collision energy (CE 20 eV), and excluded for 1 s before being re-selected for fragmentation. The parameters used for the ESI source were: interface voltage 4 kV, interface temperature 300 °C, nebulizing gas flow 3 L/min, and drying gas flow 10 L/min.

GNPS molecular networking

For molecular networking using GNPS, the Feature Based Molecular Networking (FBMN) platform was used in combination with ion identity networking (IIMN), where the data were first processed using MZmine version 2.53 and subsequently further processed in MZmine version 2.37.1.corr17.7 (Nothias *et al.*, 2020, Schmid, 2020). LC-MS data were processed similarly as described by D. A. van Bergeijk and colleagues (van Bergeijk *et al.*, 2022).

Pre-processing LC-MS data in MZmine 2.53 for feature -based molecular networking

Raw data obtained from LC-MS analysis were converted to mzXML centroid files using Shimadzu LabSolutions Postrun analysis. The files were imported into MZmine 2.53 for data processing (Pluskal *et al.*, 2010). Unless stated otherwise, m/z tolerance was set to 0.002 m/z or 10.0 ppm, RT tolerance was set to 0.05, noise level was set to 2.0E2 and the minimum absolute intensity to 4.0E2. Raw data were cropped to RT 0.5-10.0min. Mass ion peaks were detected (positive polarity, mass detector: centroid). Additionally mass ion peaks at an MS level 2 were detected using the centroid algorithm with noise level of 0. Their chromatograms were built using ADAP chromatogram builder (minimum group size of scans: 10, group intensity threshold: 2.0E2). To enhance the deconvolution step, the chromatograms were smoothed (filter width: 9), and subsequently deconvoluted (algorithm: local minimum search; chromatographic threshold: 90%; search minimum in RT range: 0.05; minimum relative height: 1%, minimum ratio of peak top/edge: 2, peak duration range: 0.03–3.00) For chromatogram deconvolution, additional MS2 scan pairing was performed with *m/z* range of 0.02 Da. The detected peaks were

deisotoped (monoisotopic shape; maximum charge: 2; representative isotope: most intense). The peak lists of each extract were aligned (weight for RT= weight for m/z =20); compare isotopic pattern with a minimum score of 50%). Missing peaks that were detected in at least one of the samples were filled with gap-filling algorithm.

Identification of complexes, adducts, and fragments in the aligned peak list was done using the identification options complex search (Ionization method: $[M+H]^+$, Maximum complex peak height: 50%), Adduct Search (Adducts: $[M+Na]^+$, $[M+K]^+$ and $[M+NH_3]^+$; Maximum relative adduct peak height: 3000%), and fragment search (Maximum fragment peak height: 50%, Minimum MS2 peak height: 0.0E0). After identification of the mass features and filtering to remove duplicate peak row and removal of mass features due to detected ringing (*m/z* tolerance: 1.0 *m/z* or 1000.0 ppm), only the mass features with associated MS/MS spectra were selected with peak list rows filter.

Pre-processing LC-MS data in MZmine 2.37.1 corr17.7 for ion identity networking
The data were further processed in in MZmine version 2.37.1.corr17.7 to detect the ion species and to generate an ion identity family (Schmid, 2020). MetaCorrelate module was used, an algorithm that groups ions derived from the same molecule, based on the retention time and the peak shape (RT tolerance 0.1 min, min height 4.0E2, noise level 2.0E2, min sample filter (default setting) min samples in all: abs=1, rel= 0.0%; min samples in group: abs= 0; rel = 0.0%, Min % intensity overlap: 60%, exclude estimated features (gap-filled)), correlation grouping (min data points: 5; min data points on edge: 2; measure: Pearson; min feature shape correlation: 5), Feature height correlation ((default settings) min data points: 2; measure: Pearson; min Correlation: 70%)). Next, ion identity networking was applied to annotate the different ion adducts (check all features, min height 4.0E2, adducts were selected in the ion library ($[M+H]^+$, $[M+Na]^+$, $[M+K]^+$ and $[M+NH_4]^+$ and modifications were selected ($[M-H_2O]$, $[M-2H_2O]$). The MS mode was set on positive; the maximum charge was 2 and the maximum molecules/clusters on 3. The data were subsequently exported for GNPS (select spectra to merge across samples; *m/z* merge mode: most intensive; intensity merge mode: sum intensities; cosine threshold is 70%, peak count threshold 20% and isolation window offset *m/z* 0.00; isolation window width *m/z* 3.0.

Data filtering

The mass feature was kept when its peak aera was higher than 3000 in two out of three replicates of each group. All other mass features were considered as noise and removed from the dataset. Afterwards the mass features that were also present in the media blanks were removed by a filtering step: the mean peak area of a mass feature in each group should be higher than 50 times the peak area of the corresponding mass feature in the blank.

Generation feature-based ion identity molecular network

The data were submitted to the GNPS web tool and a network was generated using feature networking. To generate the network the data were filtered, where MS/MS fragment ions within +/-17 Da of the precursor *m/z* were removed. Only the top 6 fragment ions in the +/- 50 Da were selected for window filtering of the MS/MS spectra. Both the precursor ion mass tolerance and the MS/MS ion mass tolerance were set to 0.02 Da. A molecular network was generated where the edges were filtered to have a cosine score above 0.7 and more than 7 matched peaks. Spectral library search was performed using the default settings. The molecular network was visualised using the software Cytoscape (Kohl *et al.*, 2011).

GC-MS analysis for the identification of *iso*-maleimycin

GC-MS analysis was executed as previously described (Uiterweerd, 2020). For analysis and identification, synthetic samples of *iso*-maleimycin (0.23 mg/ml), and a sample of the extract (1 mg/ml) were injected on a Shimadzu GC-2010 gas chromatograph coupled to a Shimadzu GC-MS-QP2010 gas chromatograph mass spectrometer. For the synthetic samples, split ratio 20.0, for the extract, split ratio 10.0. The injection temperature was 200 °C, detector temperature 250 °C. Separation was carried out on an HP-1 column, 100 % dimethylsiloxane, (30 m L × 0.25 mm Ø × 0.25 µm thickness). The temperature program was set as follows: 150 °C hold for 5 min, then increase 150 °C to 200 °C with 20 °C/min, finally 3 minutes hold at 200 °C. The carrier gas was He, with a column flow of 0.9 ml/min, pressure 82.7 kPa. Ionization by means of EI, ionization energy 70 eV, mass range from m/z 50–225. The identification was done by comparing retention times and mass fragmentation patterns of the synthetic standards with those obtained from the extracts.

Bioinformatics

For the prediction of the BGCs in the genome sequence of *Streptomyces* sp. QL37 (NZ_PTJS00000000.1) and other *Streptomyces* strains, the bioinformatic tool antiSMASH 5.0 was used (Blin *et al.*, 2019). BGC23 was compared with the maleimycin BGCs using the genome sequence of *S. showdoensis* ATCC 15227 (NZ_LAQS01000000).

Statistics

Heatmaps were generated using Matplotlib package (version 3.5.0)(Hunter, 2007). Ward's hierarchical clustering was applied using "ward" method in Scipy package (version 1.8.0) (Virtanen *et al.*, 2020).

SUPPLEMENTARY TABLES

Table S1A Genes used for the calculation of the average expression of a BGC shown in Figure 5

BGC Region	Class	Gene	locus_tag (NCBI, NZ_PTJS01000000)	Prokka annotation
BGC 1	NRPS	QL37_01200	C5F59_RS01200	Linear gramicidin synthase subunit D
BGC 2	Terpene	QL37_01665	C5F59_RS01660	tRNA 5-methylaminomethyl-2-thiouridine biosynthesis bifunctional protein MnmC
		QL37_01670	C5F59_RS01665	Demethylmenaquinone methyltransferase
		QL37_01675	C5F59_RS01670	15-cis-phytoene desaturase
		QL37_01680	C5F59_RS01675	Aminopyrrolnitrin oxygenase PrnD
		QL37_01685	C5F59_RS01680	All-trans-phytoene synthase
BGC 3	Bacteriocin	QL37_02695	C5F59_RS02730	Linocin-M18
BGC 4	T3PKS	QL37_02885	C5F59_RS02920	type III polyketide synthase
BGC 5	NRPS	QL37_03200	C5F59_RS03230	Dimodular nonribosomal peptide synthase
BGC 6	NRPS,T1PKS	QL37_03460	C5F59_RS03495	Gramicidin S synthase 2
BGC 7	Terpene	QL37_03860	C5F59_RS03900	Squalene--hopene cyclase
BGC 8	Linaridin	QL37_06275	C5F59_RS06355	hypothetical protein
		QL37_06280	C5F59_RS06360	hypothetical protein

Table S1A Genes used for the calculation of the average expression of a BGC shown in Figure 5 *(continued)*

BGC Region	Class	Gene	locus_tag (NCBI, NZ_PTJS01000000)	Prokka annotation
BGC 9a	T1PKS,hglE-KS	QL37_07220	C5F59_RS07315	Phthiocerol/phenolphthiocerol synthesis polyketide synthase type I PpsA
		QL37_07225	C5F59_RS07320	Phenolphthiocerol synthesis polyketide synthase type I Pks15/1
BGC 9b	T1PKS	QL37_07410	C5F59_RS07500	Phenolphthiocerol synthesis polyketide synthase type I Pks15/1
		QL37_07420	C5F59_RS07505	Erythronolide synthase, modules 3 and 4
		QL37_07440	C5F59_RS07510	Erythronolide synthase, modules 1 and 2
		QL37_07450	C5F59_RS07515	Phenolphthiocerol synthesis polyketide synthase type I Pks15/1
		QL37_07460	C5F59_RS07520	Oxygen regulatory protein NreC
		QL37_07465	C5F59_RS07525	Erythronolide synthase, modules 3 and 4
		QL37_07480	C5F59_RS07530	Phenolphthiocerol synthesis polyketide synthase type I Pks15/1
BGC 10	Bacteriocin	QL37_08865	C5F59_RS08915	hypothetical protein
BGC 11	Terpene	QL37_10150	C5F59_RS10245	Pentalenene synthase
BGC 12	T2PKS	QL37_12915	C5F59_RS12910	Actinorhodin polyketide putative beta-ketoacyl synthase 1
		QL37_12920	C5F59_RS12915	Actinorhodin polyketide putative beta-ketoacyl synthase 2

Table S1A Genes used for the calculation of the average expression of a BGC shown in Figure 5 *(continued)*

BGC Region	Class	Gene	locus_tag (NCBI, NZ_PTJS01000000)	Prokka annotation
BGC 13	PKS-like	QL37_13225	C5F59_RS13225	3-oxoacyl-[acyl-carrier-protein] synthase 3 protein 1
BGC 14	Siderophore	QL37_13755	C5F59_RS13750	Aerobactin synthase
BGC 15	Lanthipeptide-classIII	QL37_14765	C5F59_RS14775	Serine/threonine-protein kinase D
BGC 16	Butyrolactone	QL37_18390	C5F59_RS18460	hypothetical protein
BGC 17	Terpene	QL37_27840	C5F59_RS27970	(+)-T-muurolol synthase
BGC 18	Siderophore	QL37_28170	C5F59_RS28300	N(2)-citryl-N(6)-acetyl-N(6)-hydroxylysine synthase
		QL37_28175	C5F59_RS28305	N(6)-hydroxylysine O-acetyltransferase

Table S1B Genes used for the calculation of the average expression of a BGC shown in Figure 5

BGC Region	Class	Gene	locus_tag (NCBI, NZ_PTJS01000000)	Prokka annotation
BGC 19	NRPS-like	QL37_28445	C5F59_RS28575	Tyrocidine synthase 3
BGC 20	NRPS,Betalactone, other	QL37_28675	C5F59_RS28800	2-isopropylmalate synthase
		QL37_28680	C5F59_RS28805	Succinate-semialdehyde dehydrogenase (acetylating)
		QL37_28685	Not identified	hypothetical protein
		QL37_28690	C5F59_RS28815	Linear gramicidin synthase subunit B
		QL37_28695	C5F59_RS28820	Dimodular nonribosomal peptide synthase
		QL37_28700	C5F59_RS28825	Linear gramicidin synthase subunit B
		QL37_28705	C5F59_RS28830	Linear gramicidin synthase subunit D
BGC 21	NRPS	QL37_29280	C5F59_RS29410	Tyrocidine synthase 3
		QL37_29285	C5F59_RS29415	Linear gramicidin synthase subunit D
BGC 22	RiPP-like	QL37_29740	C5F59_RS29875	hypothetical protein

Table S1B Genes used for the calculation of the average expression of a BGC shown in Figure 5 *(continued)*

BGC Region	Class	Gene	locus_tag (NCBI, NZ_PTJS01000000)	Prokka annotation
BGC 23a	Betalactone	QL37_30710	C5F59_RS30845	2-isopropylmalate synthase
		QL37_30715	C5F59_RS30850	hypothetical protein
		QL37_30720	C5F59_RS30855	Alpha-aminoadipate--LysW ligase LysX
		QL37_30725	C5F59_RS30860	N-acetyl-gamma-glutamyl-phosphate reductase
		QL37_30730	C5F59_RS30865	Acetylglutamate kinase
		QL37_30735	C5F59_RS30870	GMP synthase [glutamine-hydrolyzing]
		QL37_30740	C5F59_RS30875	Catabolic NAD-specific glutamate dehydrogenase RocG
		QL37_30745	C5F59_RS30880	Short-chain-fatty-acid--CoA ligase
BGC 23b	Butyrolactone	QL37_30790	C5F59_RS30925	hypothetical protein
BGC 24	Ectoine	QL37_30990	C5F59_RS31125	L-ectoine synthase
BGC 25	Terpene	QL37_33360	C5F59_RS33500	Germacrene A synthase
BGC 26	NRPS	QL37_34050	C5F59_RS34210	Dimodular nonribosomal peptide synthase
BGC 27	hglE-KS	QL37_34940	C5F59_RS35105	Mycocerosic acid synthase
		QL37_34950	C5F59_RS35115	Polyketide biosynthesis protein PksE

Table S1B Genes used for the calculation of the average expression of a BGC shown in Figure 5 *(continued)*

BGC Region	Class	Gene	locus_tag (NCBI, NZ_PTJS01000000)	Prokka annotation
BGC 28	Bacteriocin	QL37_35390	C5F59_RS35545	hypothetical protein
BGC 29	NRPS	QL37_36325	C5F59_RS36490	Linear gramicidin synthase subunit D
BGC 30	T3PKS	QL37_36750	C5F59_RS36920	1,3,6,8-tetrahydroxy-naphthalene synthase
BGC 31	Melanin	QL37_37320	C5F59_RS37485	hypothetical protein
BGC 32a	T2PKS	QL37_37710	C5F59_RS37875	Actinorhodin polyketide putative beta-ketoacyl synthase 1
		QL37_37715	C5F59_RS37880	Actinorhodin polyketide putative beta-ketoacyl synthase 2
BGC 32b	Terpene	QL37_37755	C5F59_RS37920	Germacradienol/geosmin synthase
BGC 33	butyrolactone	QL37_38805	C5F59_RS38975	hypothetical protein
BGC 34	Terpene	QL37_38925	C5F59_RS39090	Pentalenene synthase
C2 BGC 1	NRPS	QL37_39520	C5F59_RS39750	hypothetical protein
		QL37_39525	C5F59_RS39755	Dimodular nonribosomal peptide synthase
		QL37_39530	C5F59_RS39760	Chondramide synthase cmdD
		QL37_39535	C5F59_RS39765	hypothetical protein
		QL37_39540	C5F59_RS39770	hypothetical protein

Table S2 Organization of BGC 23 and its comparison with the maleimycin cluster in *S. showdoensis* ATCC 15227

BGC 23		
Gene *S. sp.* QL37	**Protein**	**aa identity (%)with Mal protein in *S. showdoensis***
QL37-30710	Isopropylmalate synthase	88.7
QL37-30715	LysW	90,2
QL37-30720	RimK /LysX	70.6
QL37-30725	Semialdehyde dehydrogenase/LysY/ArgC	77.1
QL37-30730	Acetylglutamate kinase LysZ/ArgB	78.6
QL37-30735	Glutamine aminotransferase class I	77.5
QL37-30740	Glutamine /leucine/phenylalanine/valine dehydrogenase	83.4
QL37-30745	AMP-ligase	67.7
QL37-30750	Aminotransferase class I and II, Pyridoxal phosphate dependent	70.9
QL37-30755	Fumyrylacetoacetate hydrolase	58.1
QL37-30760	MerR family transcriptional regulator	Not in BGC
QL37-30765	MFS transporter	59.8
QL37-30770	DNA binding response regulator	51.9
Biosynthetic genes detected in *mal* gene cluster *S. showdoensis* and absent in BGC 23 of *S.* sp. QL37		
Gene in *S. showdoensis*	**Protein**	**aa identity (%) with protein in *S.*sp. QL37 outside BGC 23**
VO63_RS35620	Short-chain dehydrogenase	41.5
VO63_RS35610	Transposase	34.7
VO63_RS35600	Arylmalonate decarboxylase	No BLAST hit in genome
VO63_RS31285	Flavin dependent oxidoreductase	65.7
VO63_RS31290	Flavin dependent oxidoreductase	43

Table S2 Organization of BGC 23 and its comparison with the maleimycin cluster in *S. showdoensis* ATCC 15227 *(continued)*

BGC 23		
Biosynthetic genes detected in BGC 23 *S.*sp. QL37 absent in maleimycin BGC of *S. showdoensis*		
Gene *S. sp.* QL37	**Protein**	**aa identity with protein in *S.showdoensis***
QL37-30705	Acyl-CoA dehydrogenase	27.1
QL37-30700	Acyl-CoA dehydrogenase	28.6

SUPPLEMENTARY FIGURES

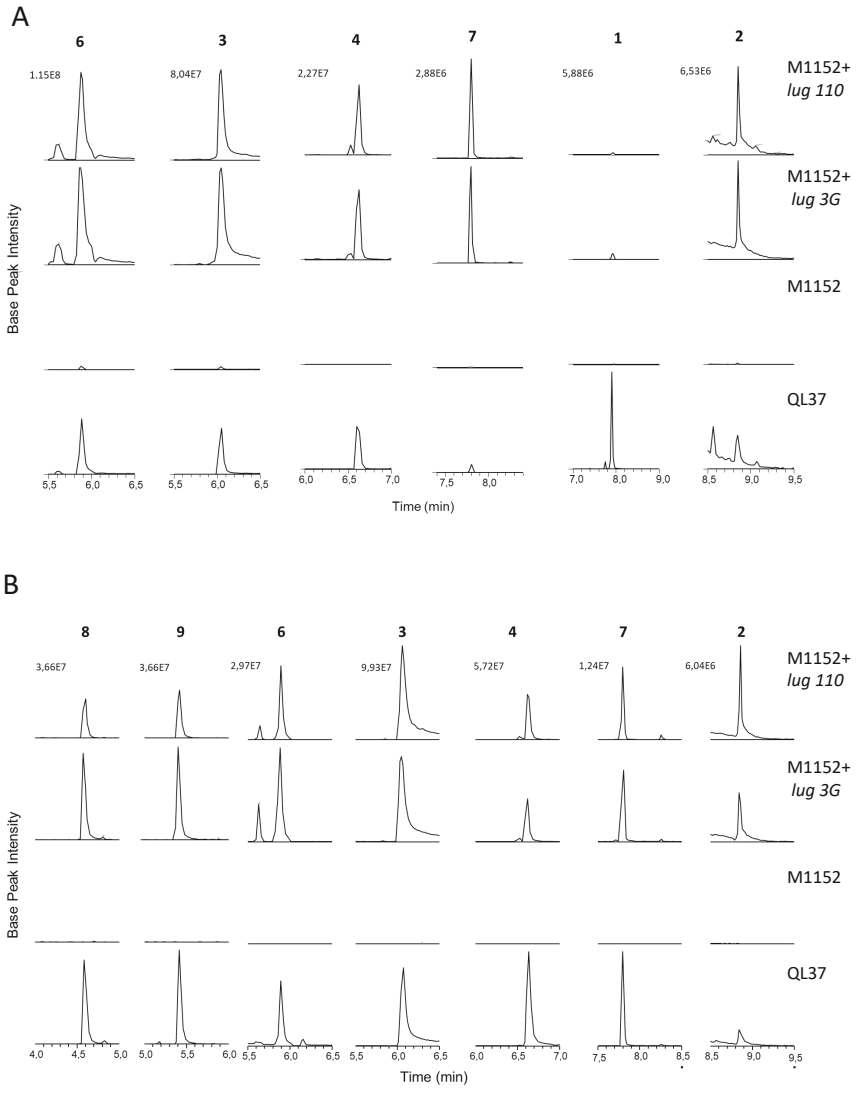


Figure S1 Extracted ion chromatogram of compounds **1–9** in the extracts of *Streptomyces* sp. QL37 and S. *coelicolor* carrying the *lug* gene cluster.
Extracted chromatogram (XIC) of the angucyclines (**2**,**3**,**4**,**6**) pratensilin A (**7**) limamycins (**8**,**9**) and lugdunomycin (**1**) in extracts derived from *S coelicolor.*M1152 carrying the *lug* gene cluster (MM152+lug 11O and MM152+lug 3G), *S. coelicolor* M1152 (MM152) and *Streptomyces* sp. QL37 (QL37). The strains were grown on MM (A) and R5 (B)

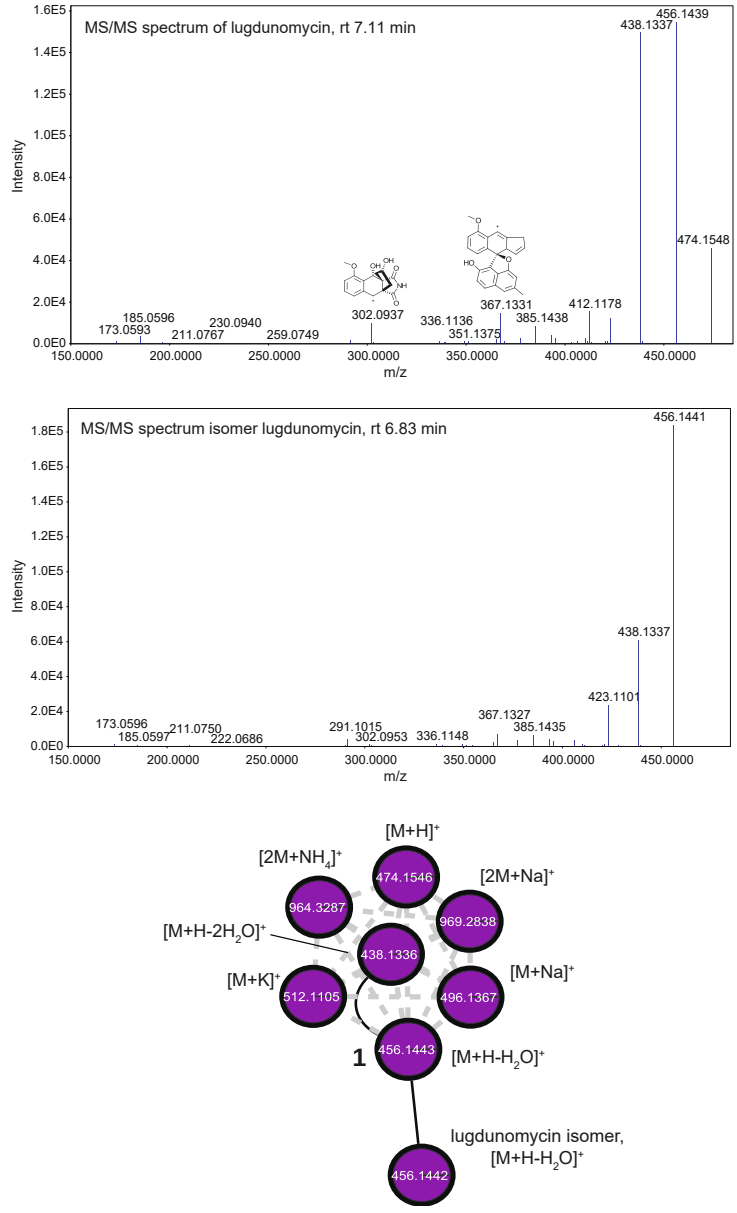


Figure S2 Subnetwork for lugdunomycin reveals lugdunomycin isomer.
Comparison of the MS/MS spectra of the different lugdunomycin isomers. These different molecules were detected at different retention times, but share the characteristic fragments of lugdunomycin, and thus could be analogues structures.

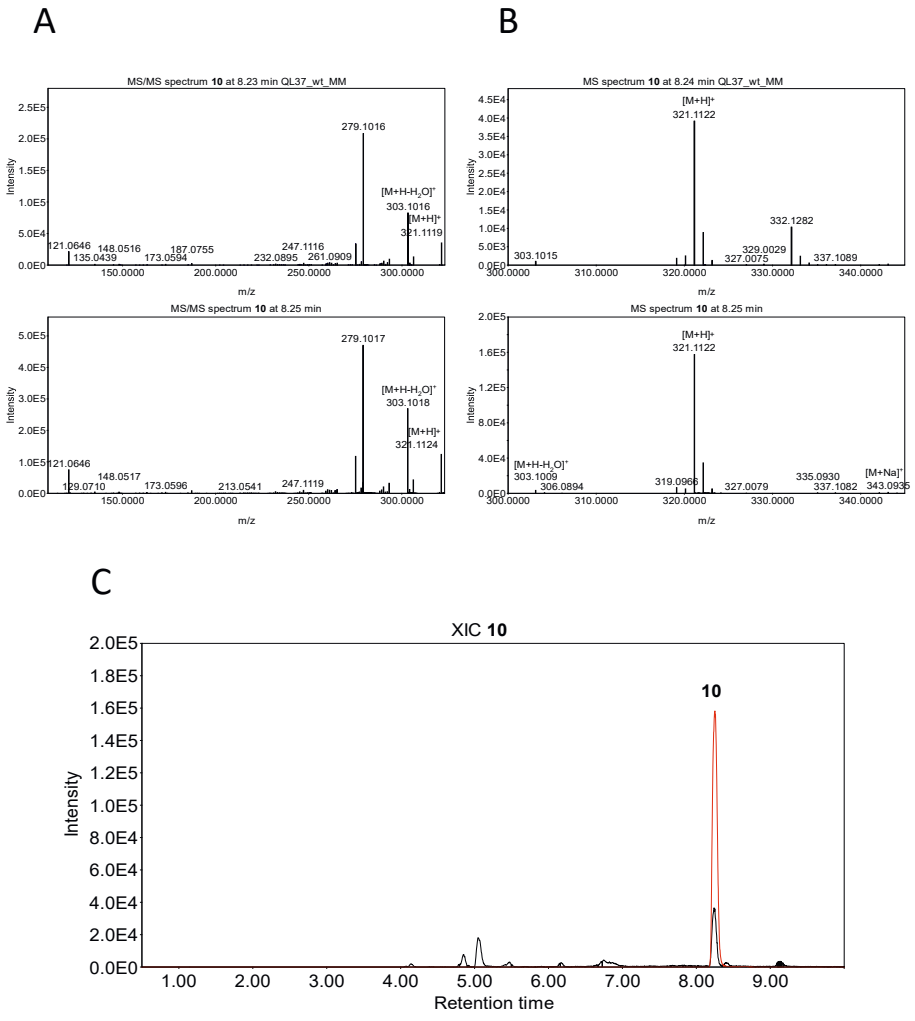


Figure S3 Identification of compound 10 in the extract of *Streptomyces* sp. QL37. Comparison of the MS/MS (A) and MS spectra (B) of compound **10** and its corresponding peak in the crude extract of *Streptomyces* sp. QL37. The extracted ion chromatograms of the two peaks are shown in (C) (Elmonin was a kind gift from Michiel Uiterweerd and Prof. Dr. A. Minnaard (university of Groningen).

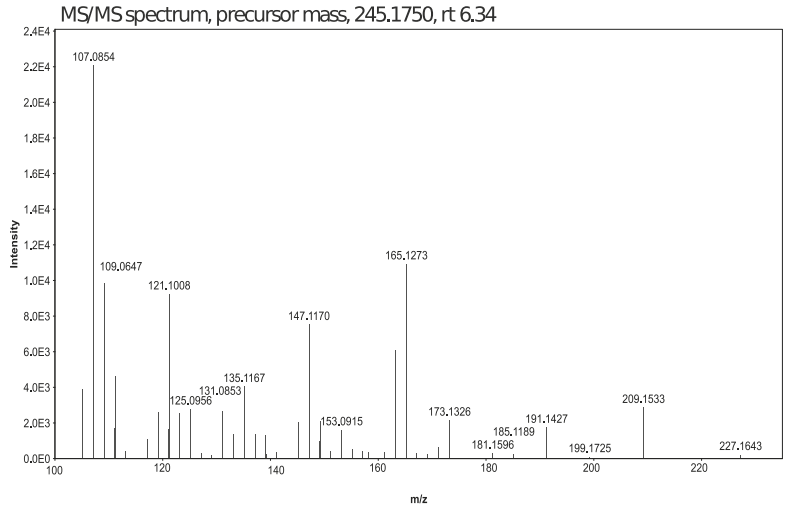


Figure S4 MS/MS spectrum of the precursor mass related to SCB1.
MS/MS spectrum of a node in one of the molecular families in the GNPS network indicated in Figure 3. This node showed a fragmentation pattern consistent with that of SCB1 as previously described (Yang *et al.*, 2005).

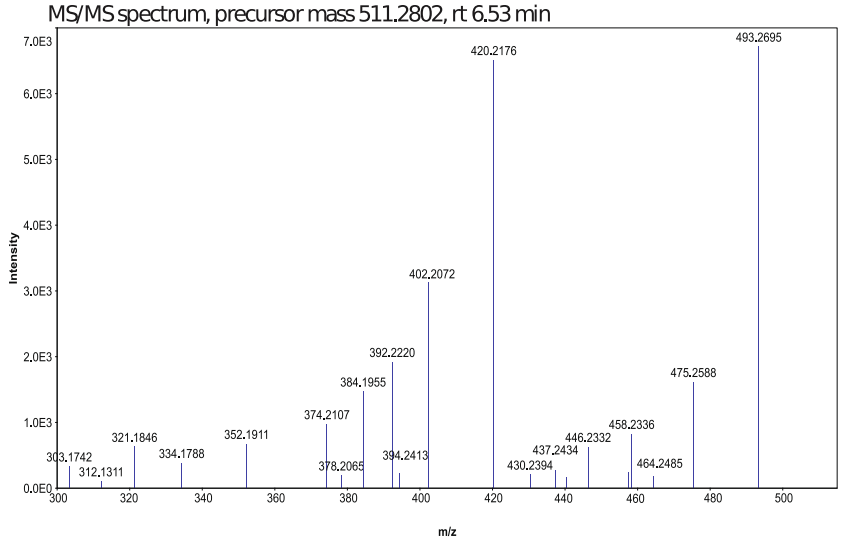


Figure S5 MS/MS spectrum of the precursor mass related to alteramide A.
MS/MS spectrum of a node in one of the molecular families in the GNPS network indicated in Figure 3. This node showed a fragmentation pattern consistent with that of alteramide A as previously described (Shaikh *et al.*, 2021, Moree *et al.*, 2014).

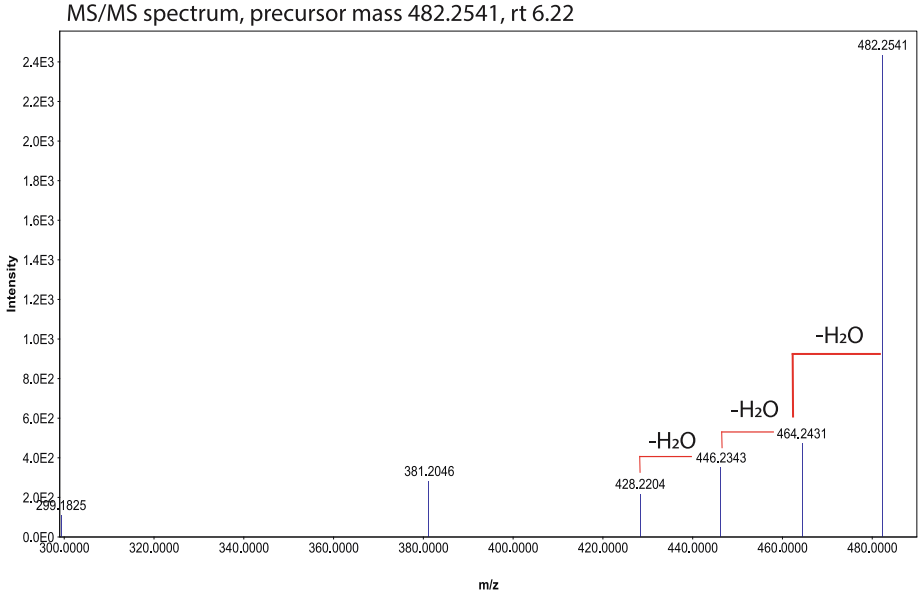


Figure S6 MS/MS spectrum of the precursor mass related to sceliphrolactam.
Detection of the monoisotopic mass of sceliphrolactam and the detected MS/MS spectrum in the extracts of *Streptomyces* sp. QL37 grown on MM. However, the MS/MS of sceliphrolactam has never been reported (Low *et al.*, 2018).

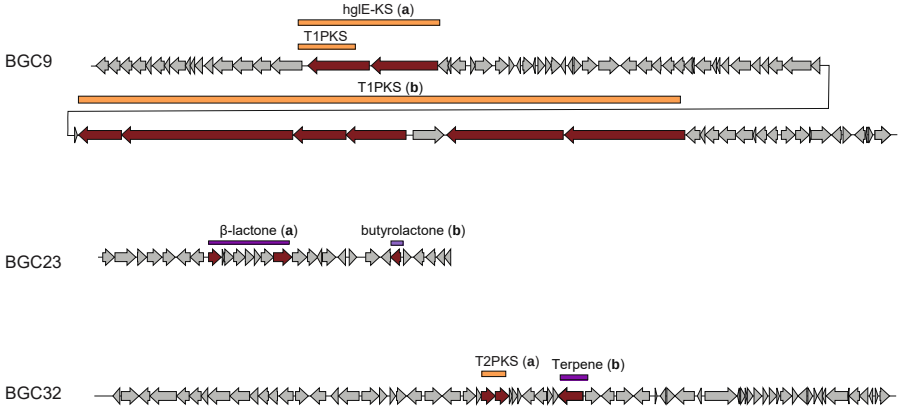


Figure S7 Organization of BGC9, BGC23 and BGC32 and annotation of BGC types.
Illustration was adapted from antiSMASH 5.0 (Blin *et al.*, 2019). Bars above the BGCs indicate the core regions (protocluster) that were assigned to each BGC and the BGC type, based on the predicted core genes (dark red genes). The expression of the genes included in the core regions was used to predict the average gene expression of the BGC. The different BGC types are indicated as “a” and “b” in Figure 5.

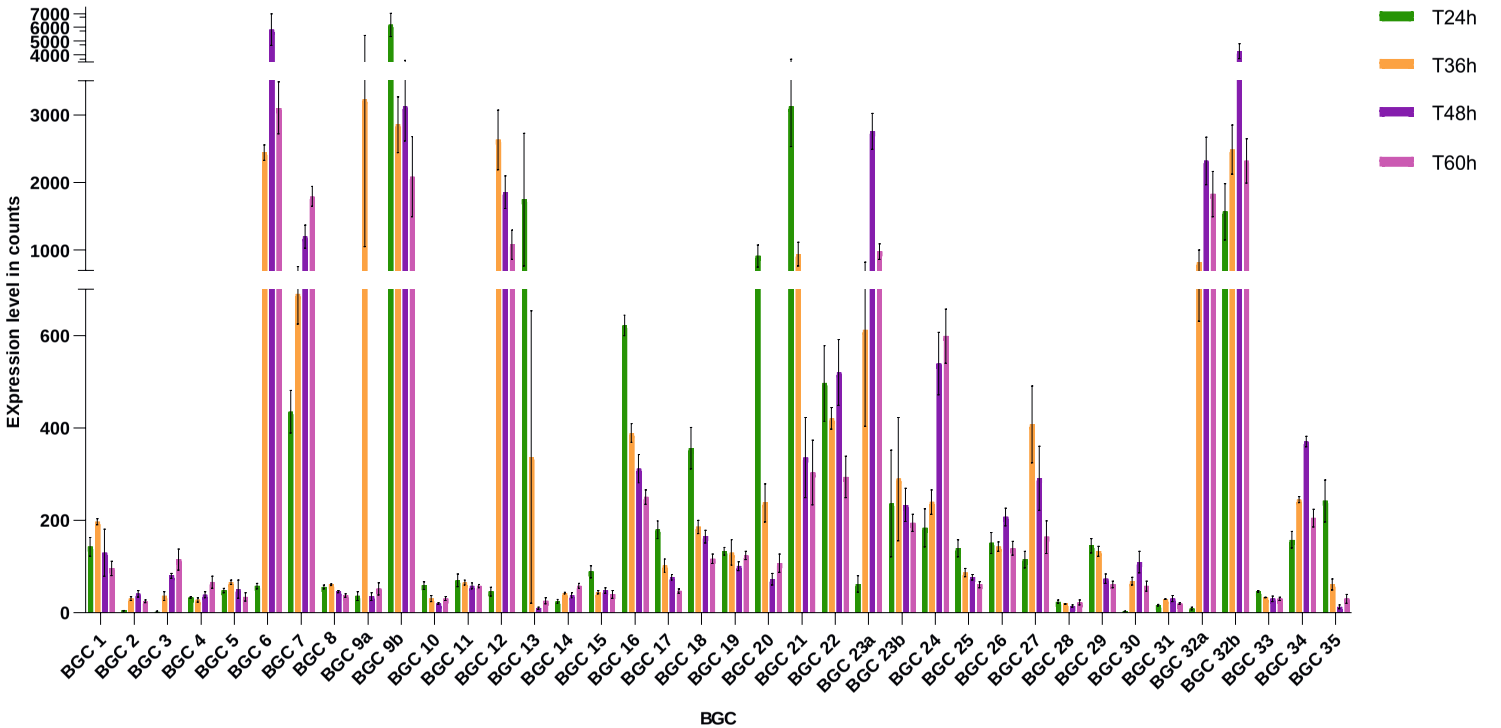


Figure S8 Average expression level of the core genes of the BGCs predicted in the genome of *Streptomyces* sp. QL37.
The graph corresponds to the heatmap indicated in Figure 5. The expression level is indicated in normalised counts. The RNA-seq data was normalized using DESeq2(Love *et al.*, 2014). The core biosynthetic genes encoding for a nonribosomal peptide synthase (NRPS), polyketide synthase (PKS), terpene cyclase or ribosomally-synthesised and post-translationally modified peptides (RiPP) were predicted by antiSMASH. In case the BGC contains multiple core genes, the average gene expression of multiple core genes was calculated. The error bars indicate the SEM.

Functional and metabolomic study of the LugO oxygenases of the lugdunomycin biosynthetic pathway reveals their role in angucycline C-ring cleavage and elicits hidden biosynthetic pathways

Helga U. van der Heul
Xiansha Xiao
Changsheng Wu
Somayah S. Elsayed
Gilles P. van Wezel

Part of this chapter is published as:
Xiao, X., Elsayed, S.S., Wu, C., van der Heul, H.U., Metsä-Ketelä, Du, C., Prota, A.E., Chen, C.-C., Liu, W., Guo, R.-T., Abrahams, J.P., van Wezel, G.P. (2020) Functional and structural insights into a novel promiscuous ketoreductase of the lugdunomycin biosynthetic pathway. *ACS Chem Biol.* **15**: 2529-2538

ABSTRACT

Lugdunomycin is a highly rearranged angucycline produced by *Streptomyces* sp. QL37. More insights into this biosynthetic pathway are required for metabolomic engineering with the aim to steer angucycline biosynthesis towards lugdunomycin biosynthesis and to increase the angucycline chemical space. Here we present a bioinformatic, functional and metabolomics study of the role of the five LugO oxygenases encoded by the lugdunomycin biosynthetic gene cluster (BGC). These studies showed that the flavoproteins LugOI and LugOII are part of the minimal PKS and hence required for angucycline biosynthesis, while the monooxygenases LugOIII and LugOV are required for the production of C-ring rearranged angucyclines, including pratensilin A, limamycins, elmonin and lugdunomycin. Therefore, these enzymes likely catalyse the C–C bond cleavage in the angucycline C-ring. LugOIV is apparently redundant for angucycline biosynthesis. Importantly, GNPS mass-spectral networking revealed major changes in the metabolite profiles of *lugO* null mutants, with many previously unknown angucyclines and a molecular family of N-acyl amino acids, amino alcohols as well as a novel lipopeptide. This provides an example of how interference with major biosynthetic pathways, in this case for angucyclines, may give way to the production of potentially exciting novel compounds. The *lugOV* mutant accumulated large amounts of oxygenated angucyclines, including putative epoxide angucyclines. Further analysis suggested that LugOIII acts as an epoxidase, and LugOV then acts as a Baeyer–Villiger oxygenase that cleaves between the C6a–C7 to produce a lactone C-ring. Taken together, our study provides new clues into the roles of the LugO oxygenases in the biosynthesis of angucyclines, limamycins and lugdunomycin.

INTRODUCTION

Lugdunomycin (**1**) is a highly rearranged angucycline with an unprecedented biosynthetic complexity, which is produced by *Streptomyces* sp. QL37 (Figure 1) (Wu *et al.*, 2019, Wu, 2016). It has a benzaza[4,3,3]propellane-6-spiro-2'-2*H*-naphtho[1,8-*bc*]furan backbone and is derived from angucyclines, one of the largest group of type II polyketides (Kharel *et al.*, 2012, Wu *et al.*, 2019). Angucyclines are characterised by their benz[a]anthracene ring, which consists of three fused benzene rings (designated as B–D-rings), to which a fourth benzene ring is fused to the B-ring at an angle (designated as A-ring) (Figure 1). The most well-known angucyclines with this typical structure are the urdamycins and the landomycins (Kharel *et al.*, 2012). Besides metabolites with this angucycline backbone, many rearranged, atypical angucyclines have been identified. These include angucyclines that are cleaved in the A-ring (fridamycins, BE-7585A and urdamycin L), the B-ring (kinamycins, jadomycin, lomaiviticin, gilvocarcin and fluostatins) and the C-ring (emycins, elmenols, pratensilins, limamycins, and lugdunomycin) (Chen *et al.*, 2011, Maskey *et al.*, 2003, Yoon *et al.*, 2019, Tibrewal *et al.*, 2012, Wang *et al.*, 2015, Fan & Zhang, 2018, Pan *et al.*, 2017, Walker *et al.*, 1999, Ma *et al.*, 2015, Wang *et al.*, 2019, Yixizhuoma *et al.*, 2017, Fotso *et al.*, 2008, Raju *et al.*, 2013, Zhang *et al.*, 2017b, Sasaki *et al.*, 2010). Especially the rearrangement of the B-ring has been studied extensively (Fan & Zhang, 2018). The oxygenases that catalyse these reactions have been biochemically characterised. These include GilOII (gilvocarcin), JadG (jadomycin) and AlpJ (kinamycins) (Pan *et al.*, 2017, Fan *et al.*, 2012a, Tibrewal *et al.*, 2012). These enzymes were proposed to catalyse a C–C cleavage through a Baeyer–Villiger oxidation (BVO) reaction, whereby an oxygen is incorporated between C5 and C6 in the ketone-containing B-ring of the angucycline, leading to the formation of a lactone or an aldehyde/acid intermediate (Tibrewal *et al.*, 2012, Tolmie *et al.*, 2019, Fan *et al.*, 2012a). For the catalysis of this reaction GilOII, JadG and AlpJ require a separate NADPH-dependent flavin reductase. The proteins show weak sequence similarity with the anthrone oxygenases TcmH and ActVA-Orf6, which catalyse the incorporation of an oxygen leading to the formation of a quinone ring (Hertweck *et al.*, 2007). An example of other polyketides that are derived from BVO are mithramycin DK and xantholipin, which are products from the Baeyer–Villiger monooxygenases (BVMOs) MtmOIV and XanO4, respectively (Kong *et al.*, 2016, Beam *et al.*, 2009). These proteins are similar to flavoproteins, that are mainly involved in hydroxylation reactions.

In addition to lugdunomycin (**1**) and the non-rearranged angucyclines (**2**–**6**), *Streptomyces* sp. QL37 also produces a limamycin derivative of which the structure was later revised to pratensilin A (**7**) and the limamycins (**8**, **9**) (Figure 1) (Wu *et al.*, 2019, Mikhaylov *et al.*, 2021). Pratensilin A was previously isolated from the marine-derived *Streptomyces pratensis* KCB-132, together with the typical angucyclines **2** and **3** (Guo *et al.*, 2020). The limamycins were previously identified in the extracts of *Streptomyces* sp. ICBB8309 and ICBB8415, and were designated as limamycin A (**8**, **9**) and limamycin B (Fotso *et al.*, 2008). These strains also produce the angucyclines **2**, **3**, and **4**; emycin; and angucyclinone C (Fotso *et al.*, 2008). Lugdunomycin, pratensilin A and the limamycins are derived from angucyclines that undergo a C-ring cleavage and nitrogen incorporation (Fotso *et al.*, 2008, Wu *et al.*, 2019, Guo *et al.*, 2020).

So far it was unclear which enzymes are involved in the opening of the angucycline C-ring, or which proteins synthesise limamycins and lugdunomycin (Wu *et al.*, 2019).

The sequence of the biosynthetic gene cluster (BGC; named *lug*) that specifies angucycline biosynthesis in *Streptomyces* sp. QL37 was elucidated, and based on the annotation and the chemical structures identified, a biosynthetic pathway was proposed (Wu *et al.*, 2019). The biosynthesis of the angucyclines starts with malonyl and acetyl-CoA (Kharel *et al.*, 2012). The enzymes of the minimal PKS (KSα, KSβ, and ACP), encoded by the genes *lugA–C*, construct a long carbon chain. This chain is folded into an angucycline backbone by ketoreductases (LugD) and cyclases (LugE and LugF). The first stable angucycline intermediate is UWM6, which is converted to the proposed lugdunomycin precursors **2** and **3**. Subsequently, the angucycline backbone is cleaved at the C6a–C7 bond of the C-ring by an oxygenase, and further rearranged. See also Chapter 3. This can lead to the synthesis of pratensilin A (**7**) the limamycins (**8** and **9**), elmonin (**10**) or a diene that acts as a substrate for a [4+2] Diels-Alder reaction with the dienophile *iso*-maleimycin resulting in lugdunomycin (**1**) (Wu *et al.*, 2019, Uiterweerd, 2020). See also Chapter 4.

The purpose of this study was to define the role of the oxygenases that are encoded by the *lug* gene cluster, which we designated LugOI–LugOV. To determine the function of each oxygenase in the lugdunomycin biosynthetic pathway, the

respective genes were deleted, and the metabolome of the mutants was analysed using metabolomics. Additionally, mass-spectral molecular networking using GNPS was used to unveil the differential production of the key metabolites, together with their structural relatedness. This study uncovers clues on the functionality of the oxygenases of the lugdunomycin gene cluster in the biosynthesis of the typical angucyclines towards limamycins and lugdunomycin.

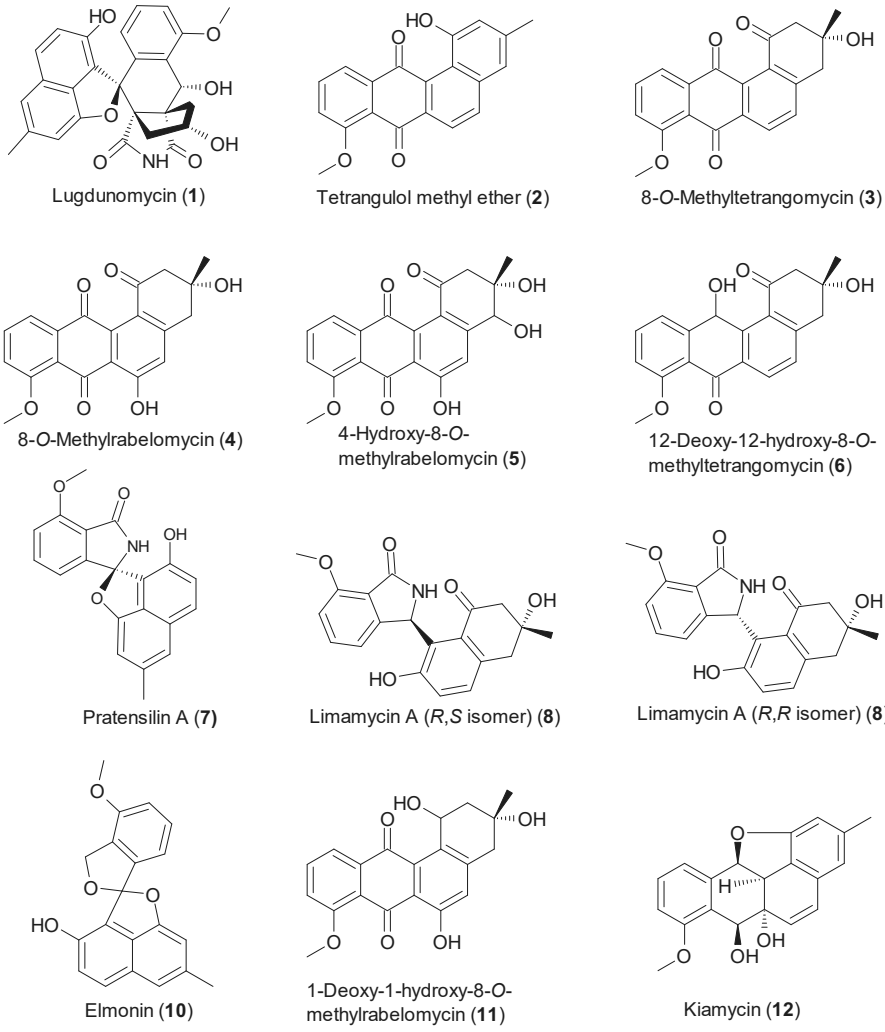


Figure 1 Structures of molecules discussed in this study.
Lugdunomycin (**1**), the non-rearranged angucyclines (**2**, **3**, **4**, **5** and **6**), pratensilin A (**7**) and the limamycins (**8**, **9**) were previously isolated from *Streptomyces* sp. QL37 (Wu *et al.*, 2019). Elmonin (**10**) was detected in the extracts of *Streptomyces* sp. QL37 by comparison with a standard (Chapter 4), 1- deoxy-1 hydroxy-8-*O*-methylrabelomycin (**11**) was characterised as a product of LugOII and kiamycin (**12**) was isolated from *Streptomyces* sp. w007 (Xiao *et al.*, 2020, Zhang *et al.*, 2012)

RESULTS

Bioinformatic analysis of the LugO enzymes

The *lug* gene cluster encodes five putative oxygenases (LugOI–OV). Orthologues of the five oxygenases are also encoded by other *lug*-type gene clusters, amongst others in *Streptomyces* sp. w007*, Streptomyces sp.* CB00072 and *Streptomyces* sp. Go-475. See also Chapter 3 for the relationship between BGC and strain phylogeny. All three gene clusters govern the production of the non-rearranged angucyclines tetrangulol methyl ether (**2**) and methyltetrangomycin (**3**) (Zhang *et al.*, 2015a, Zhang *et al.*, 2012, Cao *et al.*, 2021, Kibret *et al.*, 2018). The *lug*-type cluster of *Streptomyces* sp. w007 also directs the production of kiamycin (**12**) and an isobenzofuran derivative. The *lug*-type cluster of *Streptomyces* sp. CB00072 directs the production of thioangucyclines (Figure 2A) (Zhang *et al.*, 2012, Zhang *et al.*, 2015a, Cao *et al.*, 2021).

To get a first indication of the functionality of the various LugO enzymes, the enzymes were subjected to phylogenetic analysis (Figure 2B). The phylogenetic tree includes known angucycline oxygenases, together with anthrone oxygenases like TcmH and ActVA-ORF6 (Sciara *et al.*, 2003) and Baeyer–Villiger monooxygenases (BVMOs) involved in the biosynthesis of other polyketides, such as MtmOIV, XanO4, UrdM, BexM, GilOII, JadG and AlpJ (Tibrewal *et al.*, 2012, Fan *et al.*, 2012a, Pan *et al.*, 2017, Kong *et al.*, 2016, Sasaki *et al.*, 2010). LugOI and LugOII could be separated into different branches, and share around 45% aa identity. Both enzymes clustered with flavoproteins and have orthologues in many different angucycline pathways, suggesting their involvement in the biosynthesis of the angucycline backbone (Fan & Zhang, 2018). Notably, both enzymes clustered with a flavoprotein involved in the biosynthesis of hatomarubigin in *Streptomyces* sp. 2238-SVT4 (Izawa *et al.*, 2014). Additionally, LugOII clustered with amongst others the type 'O' BVMOs UrdM, PgaM and BexM involved in the biosynthesis of urdamycin, gaudimycin, BE-7585A in *S. fradiae* Tü 2717, *Streptomyces* sp. PGA64, and *Amycolatopsis orientalis* subsp. *vinearia* BA-07585 (Tolmie *et al.*, 2019, Sasaki *et al.*, 2010). See also Chapter 3.

LugOIII and LugOV are predicted antibiotic monooxygenases, although the putative antibiotic monooxygenase pfam domain LugOV has only low similarity with the consensus (e-value 0.00038). LugOIII clustered together with May19, an oxygenases protein involved in mayamycin biosynthesis in *Streptomyces* sp.

120454 (Bo *et al.*, 2018). However, no function was proposed for this enzyme. LugOV clustered into one clade with HrbF and shows high similarity with this protein (48% aa identity) which was reported to regulate the regiospecificity of oxygenation enzymes in hatomarubigin biosynthesis by *Streptomyces* sp. 2238-SVT4 (Izawa *et al.*, 2014). Interestingly, the homologues of *lugOIII* and *lugOV* lie directly adjacent to one another and are transcriptionally linked in among others the kiamycin BGC of *Streptomyces* sp. w007 and the *lug*-like gene clusters of *Streptomyces* sp. CB00072, b84 and *S. scopuliridis* (Fig. 2A). This strongly suggests linkage in terms of timing of their expression and function Figure 2A. The oxygenase LugOIV contains a short-chain dehydrogenase (SDR) domain and only clustered with LugOIV from other *lug*-type gene clusters.

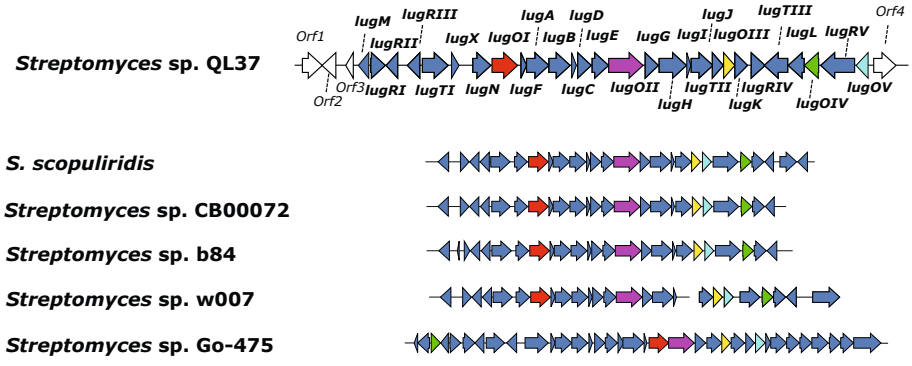


Figure 2A Phylogenetic analysis of oxygenases encoded by angucycline BGCs. Comparison of the *lug* gene cluster with related clusters from other *Streptomyces* strains.

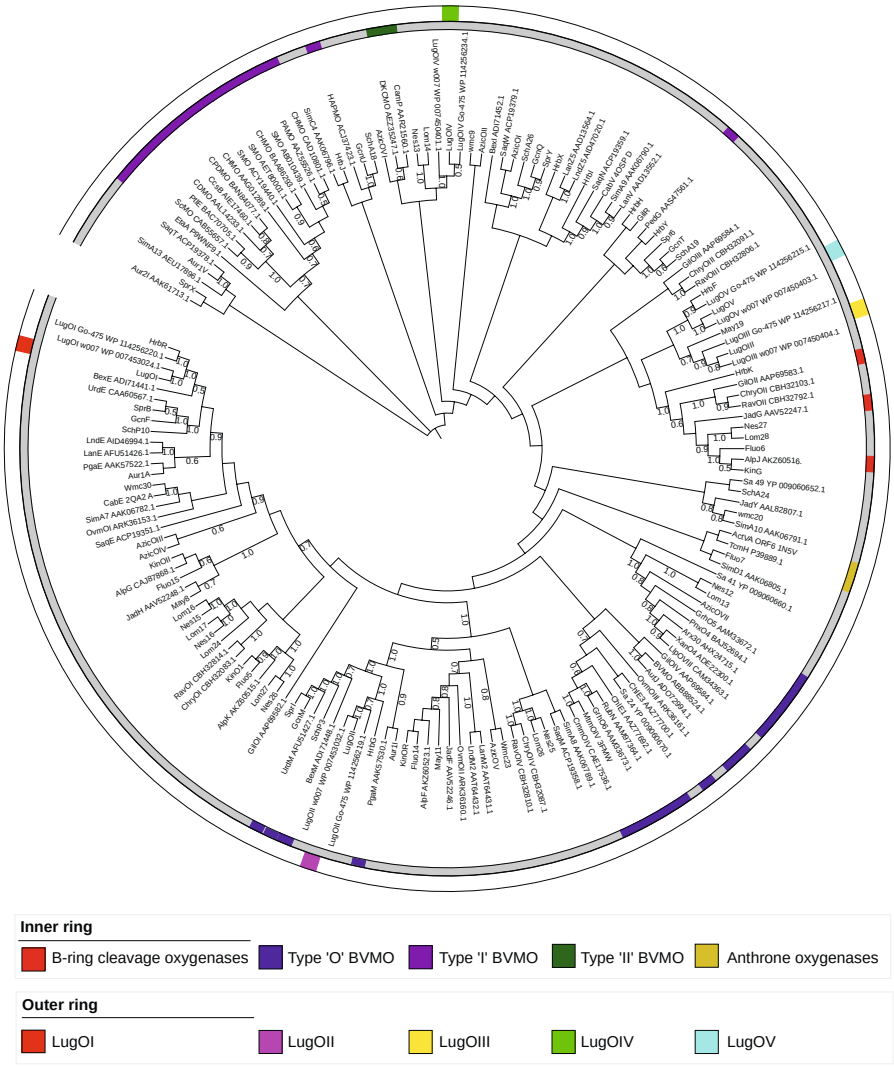


Figure 2B Phylogenetic analysis of oxygenases encoded by angucycline BGCs. Maximum-likelihood tree of oxygenases encoded by all known angucycline BGCs and BVMOs known to be involved in ring opening of aromatic compounds. The analysis involved 169 amino acid sequences and in total 1480 positions. The percentage of replicate trees in which the associated proteins clustered together in the bootstrap test (500 replicates) are indicated next to the branches (in a scale from 0-1). Only bootstrap values of >0.5 are shown at nodes. The tree was generated using MEGA-X (Kumar *et al.*, 2018) .

Mutational analysis of the *lugO* genes

To obtain more insights into the predicted roles of the oxygenases in lugdunomycin biosynthesis, *lugOI–OV* null mutants were created in *Streptomyces* sp. QL37. For each *lugO* gene, deletion mutants were created using a method published previously (Swiatek *et al.*, 2012). Briefly, for generation of the knock-out construct, around 1.5 kb of the upstream and downstream regions of the genes were amplified by PCR from the genomic DNA from *Streptomyces* sp. QL37 with the primer pairs listed in Table 2. The DNA fragments were cloned into the unstable multicopy vector pWHM3-*oriT*, and the engineered XbaI site was used for insertion of the *aac*(3)*IV* apramycin resistance cassette flanked by *loxP* sites between the flanking regions. The construct was conjugated to *Streptomyces* sp. QL37 using the methylase deficient strain *Escherichia coli* ET12567/pUZ8002 (MacNeil, 1988, Kieser *et al.*, 2000). The presence of the *loxP* recognition sites allowed the efficient removal of the apramycin resistance cassette after the introduction of plasmid pUWLcre expressing Cre recombinase(Khodakaramian *et al.*, 2006).

To assess the effect of the deletion of the *lugO* genes on morphology and angucycline production, the respective mutants were grown on different media for seven days at 30°C (Figure 3). These included Minimal Media (MM) agar plates supplemented with 0.5% mannitol and 1% glycerol, R5 agar plates supplemented with 0.8% peptone and 1% mannitol and Soya Flour Mannitol (SFM) agar plates (Wu *et al.*, 2019). These media were selected as lugdunomycin was previously isolated from cultures grown on MM, whereas pratensilin A (**7**) and the limamycins (**8**, **9**) were isolated from cultures on R5 medium. The angucyclines were isolated from both media (Wu *et al.*, 2019). SFM was included as it promotes sporulation of *Streptomyces* (Kieser *et al.*, 2000). The wild-type strain and the *lug-pks* mutant were used as the controls. The *lug-pks* deletion mutant does not produce any angucyclines, pratensilin A, limamycins or lugdunomycin (Wu *et al.*, 2019), (for the description of the mutant I refer to Chapter 3). Angucyclines are mostly yellow/brownish diffusible metabolites and their production is therefore readily assessed visually (Wu *et al.*, 2019). The degree of pigmentation depends on the type and amount of angucycline intermediates that are produced. As lugdunomycin is colourless, it can only be identified based on its spectral properties (Wu *et al.*, 2019).

All strains developed well on MM and SFM agar. On R5 agar, the wild-type *Streptomyces* sp. QL37 had a non-sporulating phenotype with sparse aerial hyphae, coinciding with the production of brown-pigmented compounds. Conversely, its

lug-pks null mutant sporulated well and the media were less pigmented, although some yellow pigmentation was still observed. Since the *lug-pks* mutant does not produce angucyclines (Wu *et al*., 2019), this shows that angucyclines are not the only yellow/brown-pigmented metabolites produced by *Streptomyces* sp. QL37. The *lugOI* and *lugOII* mutants produced aerial hyphae and spores, while the *lugOIII–OV* mutants failed to develop. Many angucyclines have anticancer and/or antibacterial activity, and we therefore propose that the observed inhibition of development is due to their enhanced production. On all media, pigmentation increased in the *lugOI* mutant, decreased in *lugOII* and *lugOIII* mutants, while no change was seen for the *lugOIV* mutant. Colonies of the *lugOV* mutant were much smaller than those of all other strains, and their agar plates turned very dark when grown on R5 and SFM. Taken together, major changes were seen in development and pigmentation in *lugOI, lugOII, lugOIII* and lug*OV* mutants, suggesting major changes in angucycline production, while *lugOIV* mutants had a very similar phenotype as the parental strain.

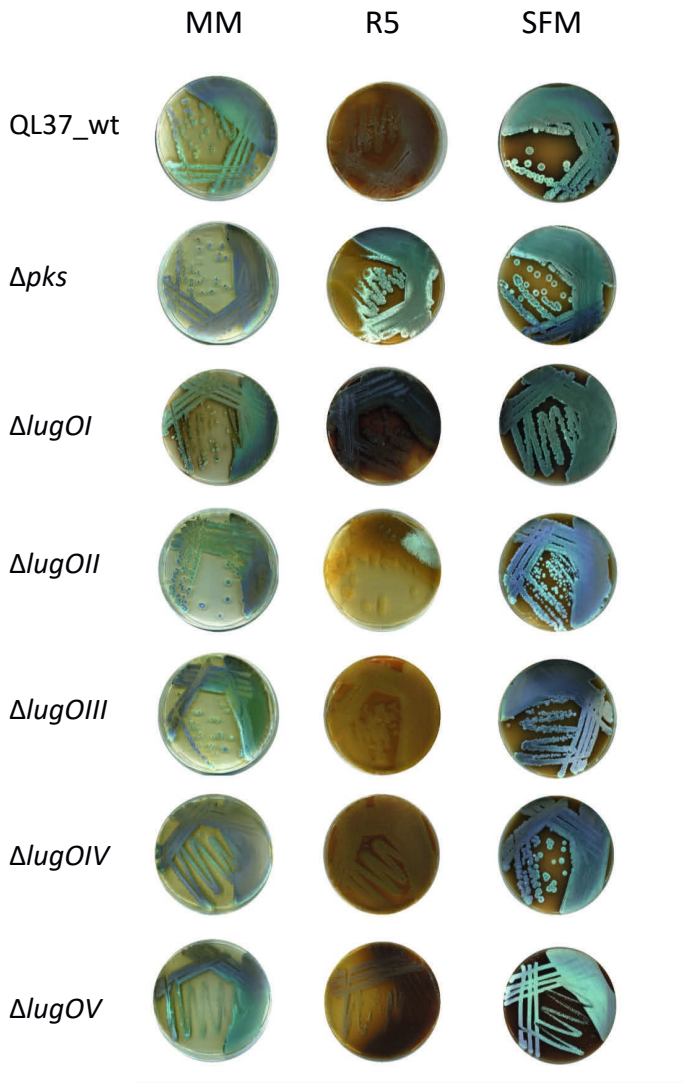


Figure 3 *Streptomyces* sp. QL37, and its *lug-pks* and *lugOI–OV* mutants on different media show differential pigmentation, suggesting differential angucycline production.
The strains were grown on MM (Minimal medium with 0.5% mannitol and 1% glycerol), on R5 (R5 supplemented with 0.8% peptone and 1% mannitol) and on SFM (Soya Flour Mannitol) for seven days at 30 °C.

Analysis of the metabolome of *lugO* null mutants using LC-MS based metabolomics

To assess the effect of the deletion of the *lugO* genes on secondary metabolism, and in particular the production of angucyclines, all strains were grown in triplicate on MM and on R5 agar plates. *Streptomyces* sp. QL37 produces lugdunomycin on MM but not on R5 (Wu *et al.*, 2019). Subsequently, the metabolites in the agar and the mycelium were extracted with ethyl acetate and analysed using liquid chromatography coupled to mass spectrometry (LC-MS). The wild-type strain and its *lug-pks* mutant were included, so as to relate the metabolites to the *lug* gene cluster. The LC-MS data obtained were processed using MZmine, resulting in a list containing all the mass features and their peak areas detected in any of the extracts. Metabolomics analysis on the processed data was performed using MetaboAnalyst (Chong *et al.*, 2019, Pluskal *et al.*, 2010).

Principle component analysis (PCA) was applied to the peak list to assess the reproducibility of the replicates and to analyse the similarity of the metabolomic profiles of the mutants relative to those of the wild-type strain or its *lug-pks* mutant (Figure 4A and B). Heatmaps with added hierarchical clustering analysis were additionally generated to better visualise the metabolite diversity and the similarities among the samples (Figure 4C and D). Both analyses of the samples obtained from MM- or R5-grown cultures revealed clear separation of all the metabolomic profiles (Figure 4A and B). Reproducibility of the analysis was supported by the observed clustering of the replicates.

The metabolomic profile of the *lugOIV* mutant was mostly similar and clustered with that of the wild-type strain on both media, indicating that LugOIV only plays a minor role in the biosynthesis of angucycline-related molecules. Conversely, the metabolomic profile of extracts obtained from the *lugOI*, *lugOII*, *lugOIII* and *lugOV* mutants was significantly different from that of the parent on both media. This differential production was observed for mass features that were highly produced in the wild -type strain and not produced in the *lug-pks* mutant, suggesting these LugO enzymes play a crucial role in the biosynthesis of angucycline-related molecules. Interestingly, mass features that were not directly linked to the *lug* gene cluster also showed a differential production profile amongst the different strains. Most obvious is the *lugOV* mutant, which produced a distinct group of mass features with relative high intensities, which were not produced by the *lug-pks* mutant and most of them were also absent or very low in any of the other strains.

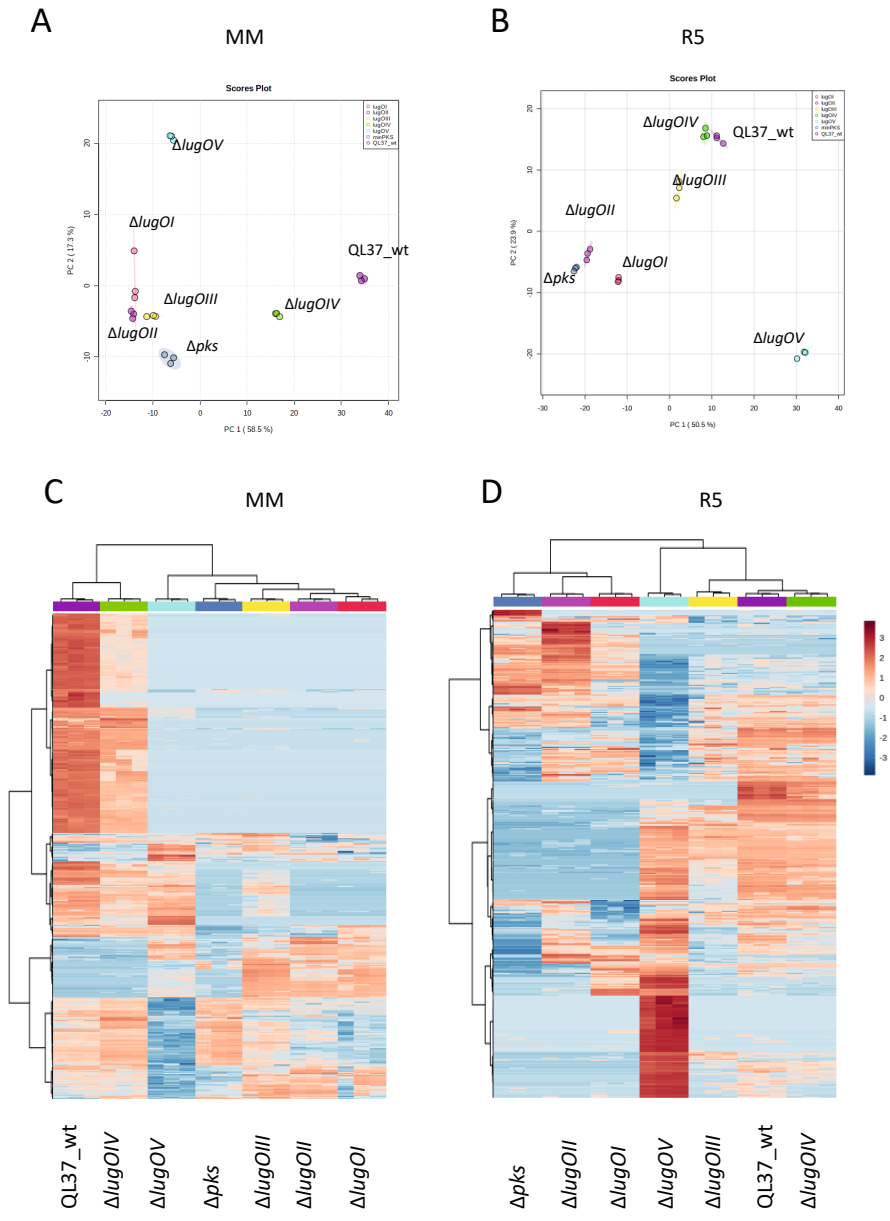


Figure 4 Exploratory statistical analysis of the metabolomic profiles of *Streptomyces* sp. QL37, and its *lugOI–OV* and *pks* mutants grown on MM and R5 agar.
Principle component analysis (PCA) score plots of the extracts from the various strains grown on MM (A) and R5 (B). Heatmaps with additional hierarchical clustering of the metabolomic profile of the extracts from the strains grown on MM (C) and R5 (D). Mass features with relatively high peak areas are displayed in red and mass features with relatively low peak areas in blue.

Analysis of the metabolome of *lugO* null mutants using molecular networking

To gain more insights into the chemical nature and structural relatedness of the metabolites whose production was altered in the *lugO* mutants, mass spectral networking was applied to the metabolites of *Streptomyces* sp. QL37, using Global Natural Products Social Molecular Networking (GNPS) (Wang *et al.*, 2016). This generates networks of parent masses that share similar MS/MS fragmentation patterns. As similar chemical structures share similar MS/MS patterns, the generated networks are thus a predictor of which mass features are having analogous structures. The molecular networks were generated using the GNPS Feature-Based Molecular Networking (FBMN) platform (Nothias *et al.*, 2020). This enables the identification of isomers with different retention times, together with the relative quantification of the different mass features using their peak areas. Additionally, Ion Identity Molecular Networking (IIMN) was applied to connect ion adducts of the same molecule based on several criteria, such as the mass difference, peak shape, and retention time (Schmid, 2020). In the end, two networks were generated representing the ions detected in the extracts from the cultures of *Streptomyces* sp. QL37 and its *lugO* mutants on either MM- or R5-grown cultures. The ions detected in the media blanks and the *lug-pks* mutant were removed before generation of the networks (Figure 5 and Figure 6). The dashed edges in the network connect different ion species of the same molecule and the solid edges connect nodes based on their MS/MS similarity. The relative intensities of the ion in each strain was mapped as a pie chart on the node representing the ion (For the absence and presence of each molecule in the extract of each strain I refer to Table S1 and Table S2 and for their corresponding extracted chromatograms and peak areas I refer to Figure S6–Figure S9).

In the molecular network generated from the extracts of MM-grown cultures (Figure 5), the ions detected in extracts of the wild-type strain represented 78% of these nodes, most of which had not previously been identified. Based on the known metabolites, we identified all molecular families for typical (non-rearranged) angucyclines (**2**–**6**), pratensilin A (**7**), and the limamycins (**8**,**9**), as well as lugdunomycin and a structurally related isomer of lugdunomycin with the same mass but different retention times (Figure 2). This may be explained by the possibility that *iso*-maleimycin can react in various ways with the diene (M. Uiterweerd and A.J. Minnaard, unpublished data). We also identified a molecular family that relates to elmonin (**10**), which candidates as the direct precursor of

the diene in the Diels-Alder reaction required for the production of lugdunomycin (Chapter 4 and M. Uiterweerd and A.J. Minnaard, unpublished data). Two additional molecular families could be related to the elmonin skeleton, since their parent masses show a fragment ion with the same mass as **10**, which further fragments in a similar way (Figure S10). All these angucycline-related molecular families were produced by the wild-type strain and its *lugOIV* mutant, although the production levels were generally reduced in the *lugOIV* mutant, indicating that LugOIV only plays a minor role in the production of the angucyclines. The deletion of *lugOI* and *lugOII* had a major impact on the metabolome, as these mutants only produced compound **4** of the previously isolated angucycline-related compounds (see also Table S1). Some angucyclines were (over)produced by the *lugOI* mutant; including a compound that could be dereplicated as rabelomycin ($[M+H]^+$ 339.0863), based on the detection of its monoisotopic mass in the GNPS network (Figure 5 and Figure 7). These data suggest that LugOI and LugOII are required in the early biosynthesis steps of lugdunomycin and for the production of the proposed substrates for oxidative C-ring cleavage; **2** and **3** (Wu *et al.*, 2019). Interestingly, the *lugOIII* and *lugOV* mutants did not produce any rearranged angucyclines, such as lugdunomycin (**1**), pratensilin A (**7**), limamycins (**8**, **9**) or elmonin (**10**), while the mutants could produce the non-rearranged angucyclines **2**, **3**, **4** and **6** (see also Table S1). Thus, LugOIII and LugOV are required for the C-ring opening and expansion of the angucycline family. It should be noted that the *lugOIII* mutant produced the non-rearranged angucyclines in lower amounts. Conversely, the *lugOV* mutant overproduced non-rearranged angucyclines, some of which were exclusively detected in this mutant. These molecules may provide clues on the substrate used by LugOV—this knowledge is crucial to elucidate the lugdunomycin biosynthesis pathway.

Interestingly, three molecular families that could be dereplicated as lipid-derived metabolites were produced by the *lugOI–III* mutants, and to a lesser extent by the *lugOV* mutant. Two of the molecular families were linear amino alcohols, comprised of an unsaturated aliphatic chain of varying lengths, possessing one primary amino group and either one or two hydroxy groups (Figure S11 and Figure S12). The other molecular family comprised *N*-acyl amino acids, with a glutamine conjugated to an unsaturated fatty acid chain of varying lengths (Figure S13). Why they were produced only in the mutants and not in the *lug-pks* mutant is yet unclear.

Also in the molecular network of the extracts of R5 grown cultures (Figure 6), molecular families were detected of the rearranged angucyclines pratensilin A (**7**) and the limamycins (**8**,**9**). The ones for elmonin and lugdunomycin were absent (see also Table S2). Thus, the growth conditions considerably influence the metabolomic profile, as was already observed in the statistical analyses (Figure 4). The major difference was the increased number of nodes in the canonical angucyclines that were upregulated in comparison to what was observed in the molecular network generated from the MM extracts, especially in *lugOI* and *lugOV* mutants. This resulted in the clustering of non-rearranged angucyclines in multiple molecular families, because they exceeded the maximum number of connected nodes which are set in the networking parameters. Therefore, both mutants are suited for the isolation of specific angucycline derivatives. Similar to the observations made on MM, when grown on R5 agar the *lugOI* and *lugOII* mutants only produced a few of the previously isolated angucyclines, including **4** and **5** (see also Table S2). The *lugOI* mutant could also still produce compound **11**. In addition, nodes with the monoisotopic mass of rabelomycin ($[M+H]^+$=339.0865), prejadomycin ($[M+H]^+$=325.1071), UWM6 ($[M+H]^+$=343.1178) and a molecular family of angucycline dimers had accumulated in the *lugOI* mutant (Figure 7 and Figure S14). Taken together, LugOI and LugOII play a role in the early post-PKS modification steps and the lugdunomycin biosynthesis pathway shares the two common key intermediates UWM6 and prejadomycin with other angucycline biosynthesis pathways (Fan & Zhang, 2018). The *lugOIII* and *lugOV* mutants were able to produce the unrearranged angucyclines (**2**–**6** and **11**) and unable to produce the rearranged angucyclines (**7**, **8** and **9**), again underlining the importance of LugOIII and LugOV in C-ring cleavage of the angucycline backbone. In addition, the *lugOIV* mutant had the same metabolomic profile as the wild-type strain, showing it does not play a major role in lugdunomycin biosynthesis under these conditions (see also Table S2).

Another interesting observation is that based on exact mass and fragmentation pattern, a family of lipopeptide metabolites were upregulated in the *lugOI* mutant, when grown on R5. These compounds consist of a pentapeptide that is linked to a fatty acyl chain of varying lengths (Figure S15A). These lipopeptides were not identified previously from any natural product source, and it shows again the exciting principle that the expression of small molecules can be elicited by the deletion of highly expressed abundant compound families (Culp *et al.*, 2019).

Analysis of the *Streptomyces* sp. QL37 genome using antiSMASH identified a BGC that likely specifies lipopeptides (Figure S15B). This BGC was also highly expressed together with the *lug* gene cluster (Chapter 4). The nonribosomal peptide synthetase (NRPS) gene in this cluster is composed of six modules. Five out of the six amino acids predicted to be incorporated by these modules, matched the pentapeptide sequence observed in the MS/MS spectra (Figure S15). The absence of the sixth residue can likely be explained by skipping of the last module during biosynthesis, which may for example be due to an inactive A domain (Wenzel *et al.*, 2006).

Altogether, the metabolomic profiles of the mutants revealed that LugOIV plays a minor role in lugdunomycin production (Table S1 and Table S2). The *lugOI* and *lugOII* mutants hardly produced any angucyclines, which implies that the gene products are required for the early steps in angucycline biosynthesis. The *lugOIII* and *lugOV* mutants only produced non-rearranged angucyclines on both media, indicating these genes are required for oxidative C-ring cleavage. Remarkably, many non-rearranged angucyclines were produced in much higher levels in the *lugOI* and the *lugOV* mutants, especially when the strains were grown on R5 agar. The *lugOI* mutant is ideally suited for the isolation of early angucycline intermediates, whereas the *lugOV* mutant is ideally suited for the isolation of late-intermediates, and structural elucidation of more angucyclines should give more insights into the biosynthesis of lugdunomycin and other C-ring rearranged angucyclines.

Interestingly on both media the mutants accumulated molecules unrelated to angucyclines. On MM amino alcohols and N-acetyl amino acids were observed in the extracts of the *lugOI–OIII* mutants; on R5 medium an unprecedented lipopeptide was produced by the *lugOI* mutant. Thus, multiple biosynthesis pathways of various compound families can be stimulated by alterations in the production profile of one abundantly produced molecular family.

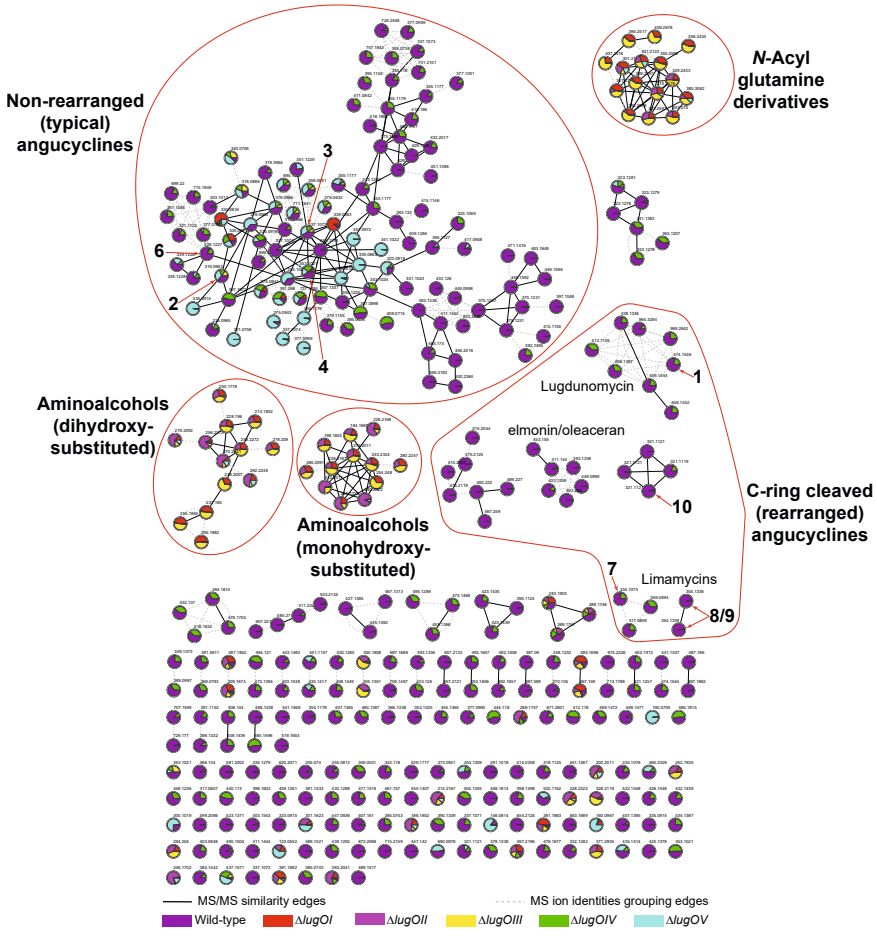


Figure 5 Molecular network of the ions detected in the extracts of *Streptomyces*. sp. QL37 and its *lugOI–lugOV* mutants grown on MM agar.
The nodes are labelled by the precursor mass of their ions and pie charts are mapped to the nodes to indicate the relative intensities of the ions in the different samples. Arrows are used to highlight the metabolites which were previously identified in *Streptomyces* sp. QL37. The molecular families, which were assigned based on the previously known metabolites or based on analysing their MS/MS spectra, are highlighted.

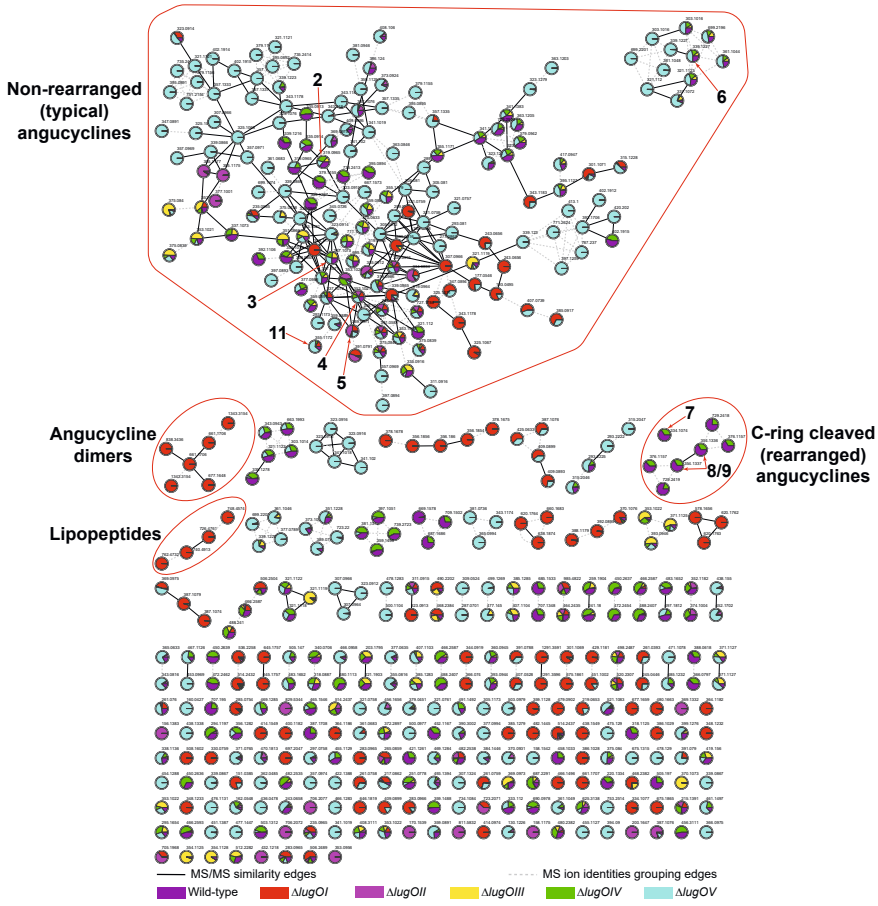


Figure 6 Molecular network of the ions detected in the extracts of *Streptomyces*. sp. QL37 and its *lugOI–lugOV* mutants grown on R5 agar.
The nodes are labelled by the precursor mass of their ions and pie charts are mapped to the nodes to indicate the relative intensities of the ions in the different samples. Arrows are used to highlight the metabolites which were previously identified in *Streptomyces* sp. QL37. The molecular families, which were assigned based on the previously known metabolites or based on analysing their MS/MS spectra, are highlighted.

DISCUSSION

Angucyclines are among the largest group of type II polyketides isolated so far (Kharel *et al.*, 2012). Despite the many dedicated studies to unravel their special biosynthetic mechanisms leading to structural rearrangements and generation of unusual chemical scaffolds, there is still more to be learned from this group of natural products. The structure of lugdunomycin, together with the additionally

isolated nitrogen containing pratensilin A and limamycins, implies C-ring cleavage and rearrangements (Wu *et al.*, 2019). While the enzymes involved in the B-ring cleavage of angucyclines have been previously studied, little is known on the enzymes and mechanisms underlying the C-ring cleavage. An approach involving targeted deletion of the different oxygenase genes of the *lug* gene cluster followed by metabolomics and molecular networking analysis revealed that LugOI and LugOII are required for the early-post PKS tailoring steps, while LugOIII and LugOV are required for the production of the C-ring rearranged angucyclines, including lugdunomycin. LugOIV may play a role in promoting angucycline biosynthesis, especially under nutrient limiting conditions, but it is not required for their production, as the metabolomic profile of the mutant was highly similar to that of the wild-type strain and all (non)-rearranged angucyclines were produced. Based on the findings a biosynthesis pathway was proposed including the functional role of LugOI, lugOII, LugOIII and LugOV (Figure 7).

LugOI and LugOII are involved in the early post-PKS modifications of angucyclines

Metabolomics and molecular networking analysis showed that the deletion of either *lugOI* or *lugOII* abolished the production of almost all the non-rearranged angucyclines (**2**, **3**, **6**), together with the elmonin (**10**), pratensilin A (**7**), the limamycins (**8**, **9**) and lugdunomycin (**1**) molecular families. Thus, *lugOI* and *lugOII* are required for the production of the precursors of the C-ring rearranged angucyclines. Several masses accumulated in the *lugOI* mutant, like rabelomycin, a rabelomycin dimer, 8-*O*-methylrabelomycin (**4**), and its 4-hydroxy derivative (**5**). This can be explained by the spontaneous oxidation of UWM6 and prejadomycin, the first stable intermediates in angucycline biosynthesis pathways (Chen *et al.*, 2005, Kulowski *et al.*, 1999). Masses corresponding to these intermediates were also observed in the molecular network. The angucycline dimers can be formed by a spontaneous non-enzymatic reaction (Huang *et al.*, 2018). Thus, the production profile of the *lugOI* mutant aligns with the function of the LugOI homologues PgaE and UrdE, which are flavoprotein monooxygenases that are required for the C12 hydroxylation of UWM6 or prejadomycin in the biosynthesis of gaudimycin and urdamycin (Kallio *et al.*, 2008b, Kallio *et al.*, 2013, Patrikainen *et al.*, 2012). The increased levels of the metabolites due to spontaneous oxidation/dehydration reactions in the *lugOI* mutant is the likely explanation of the darker colour of its

agar plates (Figure 3). This is because such reactions result in metabolites with extended double bond conjugation and accordingly more intense colour.

The *lugOII* mutant still produced **4** and **5**. The introduction of *lugOII* under the control of the constitutive *ermE** promoter in the *lugOII* mutant restored the production of **2**, **3**, and **6** (Xiao *et al.*, 2020). These observations imply that LugOII has C6 ketoreduction activity, which is in line with the predicted activity of the reductase domain of LugOII based on homology with UrdM and PgaM. From *in vitro* assays it was concluded that LugOII is a promiscuous enzyme, as it catalyses a C1 ketoreduction using **3** or **4** as substrates, leading to either SM-196B or **11** (Xiao *et al.*, 2020). This reaction has never been reported to be catalysed by LugOII homologues. It was shown in a previous research that the homologue of LugOI, PgaE, is required for the C6 ketoreduction activity of the homologue of LugOII, PgaM (Kallio *et al.*, 2008b). The same clearly applies to LugOII. Notably, the recently identified C1 ketoreduction activity of LugOII does not seem to require LugOI, as the previously isolated compound **11**, could still be produced in the *lugOI* mutant. LugOII was previously crystallised with compounds **3** and **4** and a mechanism was proposed for the catalysis of the C1-and C6 ketoreduction (Figure S19 and Figure S20) (Xiao, 2020).

The metabolites produced by the *lugOII* mutant are pretty much the same as those produced by the *lugOI* mutant, since the product of LugOI can convert to rabelomycin through keto-enol tautomerism (Figure 7). On the other hand, the presence of an intact LugOI in the *lugOII* mutant seemed to reduce the number, and sometimes the amount, of metabolites produced through the spontaneous reaction pathway.

LugOIII and LugOV are required for pratensilin A, limamycin and lugdunomycin biosynthesis

The molecular networks generated for the different mutants, together with the wild-type strain, showed that both *lugOIII* and *lugOV* mutants failed to produce C-ring rearranged angucyclines (**1**, **7**, **8**, **9** and **10**), while they were still able to produce the non-rearranged angucyclines (Figure 5 and Figure 6), indicating that both LugOIII and LugOV are involved in the catalysis of the C-ring cleavage. To gain more insight into the likely functions of LugOIII and LugOV, we interrogated the molecular networks at the nodes representing compounds **2** and **3**. This is

because both compounds are the products of the early post-PKS tailoring steps performed by LugOI and LugOII. Additionally, the structures of the C-ring rearranged angucyclines identified so far imply that **2** and **3** are the likely substrates for their production (Figure 7).

The networks in Figures 5 and 6 show **2** and **3** and their 8-hydroxy analogues to be connected to nodes upregulated in the *lugOV* mutant, which represent metabolites containing an additional oxygen atom (Figure S16 and Figure S17). We propose that these metabolites contain an epoxide moiety, because we also observed nodes that implied the oxidised metabolites were further hydrolysed or reduced (Figure S17). The hydrolysed or reduced products upregulated in the *lugOV* mutant, were absent in the *lugOIII* mutant. The angucyclines that were proposed to contain an epoxide moiety were also detected in the *lugOIII* mutant, but at relatively lower levels.

Based on these data, we conclude that LugOIII likely acts as an epoxidase, which introduces an epoxide group between C6a and C12a on the product of LugOII, while LugOV acts as a BVMO, which acts on the product of LugOIII, resulting in a lactone C-ring that may undergo further reactions like reduction, dehydration, and hydrolysis, to finally produce rearranged angucyclines, such as elmonin, pratensilin A, limamycins, and lugdunomycin (Figure 7). A recent study of the *lug*-type gene cluster from *Streptomyces* sp. CB00072 (*tac*), showed that the orthologues of *lugOIII* (*tacS*) and *lugOV* (*tacT*) are required for the production of epoxide- and C-ring cleaved angucyclines (Cao *et al.*, 2021), which is well in line with the results obtained in our study. The previously solved crystal structure of LugOIII showed resemblance with the anthrone oxygenase ActVA-ORF6 from the actinorhodin biosynthesis pathway in *S. coelicolor* and the B-ring cleavage enzyme AlpJ from the kinamycin pathway from *S. ambofaciens* (Figure S21 and Figure S22) (Xiao, 2020) . The structure suggested that indeed LugOIII contains all the key residues required to function as a monooxygenase (Xiao, 2020).

The suggested epoxidation step likely facilitates the subsequent Baeyer–Villiger oxidation of the stable benzanthraquinone skeleton of **2** and **3**, respectively. A similar mechanism was proposed in the biosynthesis of xantholipin, where an anthraquinone epoxide intermediate is initially produced, followed by a Baeyer–Villiger oxidation, which eventually leads to a structural rearrangement and

formation of a xanthone ring (Kong *et al.*, 2016). Additionally, the B-ring cleavage required for the biosynthesis of the angucycline gilvocarcin was shown to proceed through an initial oxidation of C5, followed by Baeyer–Villiger oxidation of the C6 ketone group (Tibrewal *et al.*, 2012).

The proposed roles of LugOIII and LugOV, together with their proposed sequential activities, explain the accumulation of masses resulting from oxidations and further hydrolysis in the *lugOV* mutant. The masses which showed additional reductions are likely due to LugOII or other reductases (*lugM*, *lugG* or *lugJ*) (Xiao *et al.*, 2020). The accumulation of 8-hydroxy-containing angucyclines in the *lugOV* mutant might be explained by LugOV acting preferentially on the 8-methoxy angucyclines. In the presence of LugOV, the biosynthetic pathway is subsequently directed towards producing the 8-methoxy analogues as they are continually consumed. Deletion of *lugOV* would thus result in the accumulation of the 8-hydroxy analogues, which are also substrates for LugOIII. Finally, the oxidised products produced in the *lugOIII* mutant might be the result of a side reaction of LugOV, since BVMOs can catalyse a variety of promiscuous oxidation reactions including epoxidation (Fürst, 2019). Morphologically, the increased accumulation of many angucyclines in the *lugOV* mutant, especially on R5, might be the cause of the darker discolouration of their agar plates (Figure 3). Future identification of the angucyclines which were upregulated in the *lugOV* mutant might be useful for confirming the proposed roles of LugOIII and LugOV.

Differential production of molecules unrelated to angucyclines in the *lugO* mutants

Besides the production of angucyclines, other metabolites were also affected by the deletion of *lugO* genes. On R5 medium the production of a lipopeptide was induced in the *lugOI* mutant (Figure 6). On MM medium molecular families of linear amino alcohols and N-acyl glutamine were observed in the metabolome of the *lugOI* mutant (Figure 5). These lipid-based metabolites were also detected in the *lugOII*, *lugOIII*, and to a lesser extent *lugOV* mutants grown on MM, all of which showed reduced angucyclines production as compared to the wild-type strain and its *lugOIV* mutant. The exact structures and functions of the lipid metabolites need to be investigated. Structurally similar metabolites were observed widely in eukaryotes and in some bacterial species, where some of these metabolites have been shown to function as signalling molecules, especially in response to stress

conditions (Battista *et al.*, 2019). However, to the best of our knowledge, similar lipid metabolites were not reported before in Actinobacteria. Previously it has been reported that inactivation of the biosynthesis of streptomycin or streptothricin, antibiotics commonly produced by actinomycetes, unlocked the production of hidden metabolites and novel chemistry (Culp *et al.*, 2019). We observed that blocking the biosynthesis pathway of angucycline-related molecules induces the production of previously unseen molecules that are unrelated to angucyclines. The same phenomenon was previously described for the pseudouridimycin (PUM) pathway (Iorio *et al.*, 2021), suggesting that mutation of a single modification gene in one BGC could also be applied to elicit novel chemistry, and not merely to elucidate the steps of a single biosynthesis pathway.

CONCLUSION

One of the key reactions in lugdunomycin biosynthesis is the C-ring cleavage of the non-rearranged angucyclines. The group of discovered C-ring modified angucyclines is expanding and the question remains as to how this key reaction is catalysed. Our work shows that LugOIII and LugOV are likely the major players in catalysing the C-ring cleavage, needed for the production of the rearranged pratensilin A, limamycins and lugdunomycin. Molecular networking allowed identifying the structure relatedness of the angucyclines that were affected by the absence of either enzymes, and accordingly deducing their functions. We propose that C-ring cleavage proceeds through the sequential action of LugOIII, which serves as an epoxidase, followed by LugOV, a predicted Baeyer–Villiger monooxygenase. 8-methoxy substitution on angucyclines likely promotes substrate binding to LugOV. As with other BVMOs, LugOV can additionally perform a promiscuous epoxidation reaction. However, it is not efficient enough to direct the biosynthetic pathway towards the production of detectable amounts of rearranged angucyclines.

Taken together, we provide new insights into the delineation of the lugdunomycin biosynthetic pathway, including a novel enzymatic mechanism in angucyclines biosynthesis, and at the same time awakened cryptic BGCs that are very suitable for further study. Important clues were obtained for designing future effective *in vitro* enzymatic reactions using LugOIII and LugOV, aiming at enhancing the production of lugdunomycin and increasing its structural diversity.

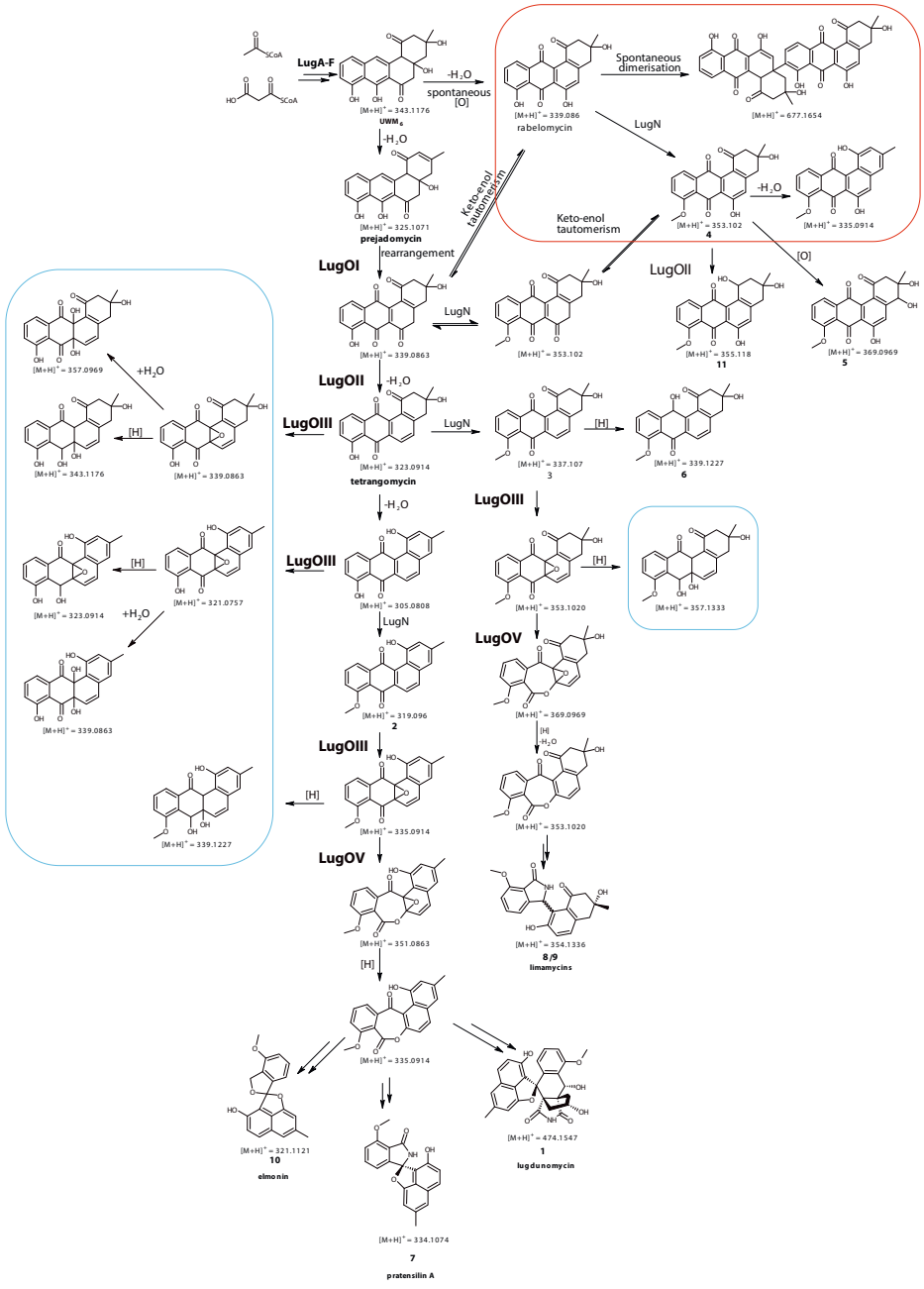


Figure 7 Proposed role of the oxygenases LugOI, LugOII, LugOIII and LugOV in lugdunomycin biosynthesis based on the metabolomics data (this study), and *in vitro* assays.
The calculated masses of the $[M+H]^+$ ions, which correspond to the precursor masses observed in the molecular networks, are given below the structures. The metabolites proposed to be accumulating in the *lugOI* or *lugOV* mutants are encircled with red and blue respectively.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Bacterial strains used in this study are indicated in Table 1. *Streptomyces* sp. QL37 was obtained from the soil in the Qinling mountains (P. R. China) (Zhu *et al.*, 2014b). The strain is deposited to the collection of the Centraal Bureau voor Schimmelcultures (CBS) in Utrecht, The Netherlands, under deposit number 138593. The metabolite production profile of each *Streptomyces* strain used in this study is indicated in Table S1 and Table S2. For general cloning *Escherichia coli* JM109 was used and grown on Luria Broth without antibiotics. *E. coli* ET12567/ pUZ8002 was used for conjugation of plasmids towards *Streptomyces* sp. QL37 (Wang & Jin, 2014, MacNeil *et al.*, 1992, MacNeil, 1988). Strains containing plasmids were selected on Luria Broth containing ampicillin (final 100 µg/mL), apramycin (50 µg/mL), chloramphenicol (25 µg/mL) and kanamycin (50 µg/mL). For the preparation of spore stocks, *Streptomyces* sp. QL37 was grown on SFM for seven days at 30°C and spores were collected as described by T. Kieser and colleagues (Kieser *et al.*, 2000). Spore stocks were stored in 20% glycerol at -20 °C.

Construction of knock-out mutants

For construction of the knock-out mutants the plasmid pWHM3-*oriT* was used, a derivative of the plasmid pWHM3, which harbours the *oriT* in the NdeI site, allowing its conjugative transfer (Vara *et al.*, 1989, Wu *et al.*, 2019, Garg & Parry, 2010). In this pWHM3-*oriT* variant the MCS is totally intact and all originally present restriction sites can be used for follow-up cloning. This is in contrast to the MCS of the pWHM3-*oriT* used in Chapter 3, where *oriT* was cloned within the MCS. The in-frame deletion mutants were created according to the method previously described (Swiatek *et al.*, 2012). Therefore, a knock-out construct was generated containing an apramycin cassette that is flanked by an up-and downstream region of the targeted gene. The ~1.5 kb upstream and downstream region of the targeted gene were amplified from *Streptomyces* sp. QL37 genomic DNA with primers that are annotated in Table 2. These were subsequently cloned in pWHM3-*oriT* using the correct restriction enzymes. An apramycin resistance cassette flanked by two *loxP* sites was cloned in between the fragments using XbaI(Khodakaramian *et al.*, 2006). The integrity of the construct was verified using Sanger sequencing and restriction enzyme analysis. The plasmid was transformed in the methylase deficient strain ET12567/ pUZ8002 that allows conjugation of the plasmid to

Streptomyces sp. QL37(MacNeil, 1988). Conjugation was executed as described by T. Kieser and colleagues (Kieser *et al.*, 2000). However, for conjugation *Streptomyces* sp. QL37 was grown on SFM containing 60 mM $MgCl_2$ and 60 mM $CaCl_2$ (Wang & Jin, 2014). The correct mutants were selected by their resistance against apramycin (50 µg/mL) and sensitivity to thiostrepton (10 µg/mL). In order to remove the apramycin cassette the pUWL-Cre construct was conjugated to the apramycin cassette containing strains (Khodakaramian *et al.*, 2006). The removal of the apramycin cassette was verified by PCR with the primers annotated in Table 2 and Sanger sequencing of the PCR products (Swiatek *et al.*, 2012).

Table 1 Strains used in this study

Bacterial strains	Characteristics	Reference
***Escherichia coli* JM109**	*end*A1, *rec*A1, *gyr*A96, *thi, hsd*R17 (rk–, mk+), *rel*A1, *sup*E44, λ–, Δ(*lac-pro*AB), [F´, *tra*D36, *pro*AB, *lac*IqZΔM15], IDE3	(Sambrook J., 1989)
***Escherichia coli* ET12567 / PUZ8002**	Methylation deficient strain with deletion in *dam*, *dcm* an *hdsM* genes with chloramphenicol resistance marker. The strain harbours a non-transmittable plasmid pUZ8002 that contains the *tra* genes, required for conjugation.	(MacNeil *et al.*, 1992)
***Streptomyces* sp. QL37**	lugdunomycin and angucycline producer	(Wu *et al.*, 2019, Zhu *et al.*, 2014b)
Streptomyces* sp. QL37 Δ*lugOI	QL37 with in-frame deletion of *lugOI*; the strain fails to produce lugdunomycin.	This study
***Streptomyces* sp. QL37 Δ*lugOI*I**	QL37 with in-frame deletion of *lugOII*; the strain fails to produce lugdunomycin.	(Wu, 2016)
Streptomyces* sp. QL37 Δ*lugOIII	QL37 with in-frame deletion of *lugOIII*; the strain fails to produce lugdunomycin.	This study
Streptomyces* sp. QL37 Δ*lugOIV	QL37 with in-frame deletion of *lugOIV*; the strain produces lugdunomycin.	(Wu, 2016)
Streptomyces* sp. QL37 Δ*lugOV	QL37 with in-frame deletion of *lugOV*; the strain fails to produce lugdunomycin.	This study
Streptomyces* sp. QL37 Δ*lug-pks	QL37 deleted for *lugA–OII*; the strain fails to produce angucyclines and lugdunomycin.	(Wu *et al.*, 2019), Chapter 3

Table 2 List of primers used in this study

Gene	Application	Sequence (5' -> 3') ^	Position #	Length Gene(nt)
lugOI	Left flank	CGATAAGCTTAGATCTTCCATCCCGCCTTCTGAAGAC	-3/+1467	1470
		CGATTCTAGAGACTACGACTGATGCGTCCATGTG		
	Right flank	CGATTCTAGAGGACCGGCCAGGTGAGATTC		
		CGATCTGCAGGCCGGTCGTTCTGCTTGGTC		
	Verification mutant	CGAGTGCCATGCCCTGATGAAG		
		GCGACGATCAGAGTGCTTGG		
lugOII	Left flank	CGATAAGCTTGGTACCGAGCTGTGGCTGGACGTGATCAAC	+9/+1967	1968
		CGATTCTAGACGTGGTGCCCACGGGTCAGC		
	Right flank	CGATTCTAGAGGCGGGACGCTCCTCGGATG		
		CGATGAATTCGGTGAGGGCCGGCGAGTAGG		
	Verification mutant	TCGACCCACACCACGGAATC		
		TTCCTCGCCGATGTGCTTGG		
lugOIII	Left flank	CGATAAGCTTGGTACCCGGGTTCGCGGAAAGTGAAG	+24/+624	642
		CGATTCTAGACGCGAGAGAGATCTGACGAG		
	Right flank	CGATTCTAGACTGCACGGCATCCGGTGATCG		
		CGATGAATTCGTCGAACTCCTTGCGGGTGAGC		
	Verification mutant	TGGTGACGCCCGAGTACAAC		
		GACATGACCTTGCGGAGCAC		

Table 2 List of primers used in this study *(continued)*

Gene	Application	Sequence (5' -> 3') ^	Position #	Length Gene(nt)
lugOIV	Left flank	CAGTGAATTCCACCAGGGTGCCGAGGTAGC	-32/+811	825
		CTAGTCTAGATCGACGAGCTGACCCGAGAAC		
	Right flank	CTAGTCTAGATTCGCACGCGGTCGGATGAG		
		CTAGGGATCCGGCCGCCCAGATGAAGGGTG		
	Verification mutant	GTATCCGGCCCGCATGCCTC		
		CCCTCGCCGACGCACACTTC		
lugOV	Left flank	CTAGAAGCTTAGCAGGTCGGCCAAGGGCAG	-24/+764	687
		CTAGTCTAGATGCGCGTCCTCACGACCAAC		
	Right flank	CTAGTCTAGATGCGCGACTCCCATGAATGGATACG		
		CTAGGAATTCGCCCGGCTCCTCGAACAGGTG		
	Verification mutant	GATAGGCGGGCCGCATGTGG		
		CCCAGCTCCGGGTGGTAGGG		

position relative to the translational start site (+1) of the respective genes.
^ restriction sites underlined. GAATTC, EcoRI; AAGCTT, HindIII; TCTAGA, XbaI; CTGCAG, PstI; GGATCC, BamHI ; AGATCT, BglII

Extraction of natural products

Streptomyces sp. QL37 spores were confluently grown on Minimal Medium (MM) agar plates (25 mL) containing 0.5% mannitol and 1% glycerol and on R5 Difco containing 1% mannitol and 0.8% peptone (Wu *et al.*, 2019). The same agars as described in Chapter 3 were used for the preparation of the media. The experiment was conducted in triplicate. After seven days of growth at 30 °C the agar plates were cut into small pieces and soaked in 25 mL of ethyl acetate for 12 hours. Subsequently the ethyl acetate was decanted and evaporated at room temperature. This process was repeated two times. The dried extract was re-dissolved in methanol (MeOH) and centrifugated in order to remove any undissolved matters. Subsequently the MeOH solutions were transferred to new pre-weighed glass vials, where it was dried under nitrogen. The crude extracts were weighed and dissolved in methanol to a final concentration of either 1 mg/mL (extracts derived from MM medium) or 0.5 mg/mL (extracts derived from R5 medium). The prepared solutions were centrifuged again for 20 min at 4 °C in order to remove any suspended matters.

Method LC-MS/MS runs

LC-MS/MS acquisition was performed using Shimadzu Nexera X2 UHPLC system, with attached PDA, coupled to Shimadzu 9030 QTOF mass spectrometer, equipped with a standard ESI source unit, in which a calibrant delivery system (CDS) is installed. The dry extracts were dissolved in MeOH to a final concentration of 1 mg/mL or 0.5 mg/mL, and 2 µL were injected into a Waters Acquity HSS C_{18} column (1.8 µm, 100 Å, 2.1 × 100 mm). The column was maintained at 30 °C, and run at a flow rate of 0.5 mL/min, using 0.1% formic acid in H_2O as solvent A, and 0.1% formic acid in acetonitrile as solvent B. A gradient was employed for chromatographic separation starting at 5% B for 1 min, then 5–85% B for 9 min, 85–100% B for 1 min, and finally held at 100% B for 3 min. The column was re-equilibrated to 5% B for 3 min before the next run was started.

All the samples were analysed in positive polarity, using data dependent acquisition mode. In this regard, full scan MS spectra (*m/z* 100–1700, scan rate 10 Hz, ID enabled) were followed by two data dependent MS/MS spectra (*m/z* 100–1700, scan rate 10 Hz, ID disabled) for the two most intense ions per scan. The ions were selected when they reach an intensity threshold of 1500, isolated at the tuning file Q1 resolution, fragmented using collision induced dissociation (CID) with fixed

collision energy (CE 20 eV), and excluded for 1 s before being re-selected for fragmentation. The parameters used for the ESI source were: interface voltage 4 kV, interface temperature 300 °C, nebulizing gas flow 3 L/min, and drying gas flow 10 L/min.

Angucycline, limamycin and elmonin standards

Compounds described in this study have been isolated and characterised previously from the wild-type *Streptomyces* sp. QL37 (Wu *et al.*, 2019). Chemically synthesised elmonin was kindly provided by Michiel Uiterweerd and Prof. dr. Adriaan J. Minnaard (University of Groningen).

Comparative Metabolomics

Before statistical analysis the data obtained from the LC-MS-runs were processed using MZmine version 2.53 (Pluskal *et al.*, 2010). The data derived from MM-grown cultures and R5- grown cultures were processed separately. Data processing in MZmine for statistical analysis in MetaboAnalyst was executed as in Chapter 4 (See Pre-processing LC-MS data in MZmine 2.53 in Chapter 4) with a few exceptions (Chong *et al.*, 2019). With the detection of the mass ion peaks only mass ion peaks at an MS level 1 were detected. Thus, in the chromatogram deconvolution step no additional MS2 scan pairing was performed. Furthermore, within the modules isotopic peak grouper, join aligner and gap-filling an RT tolerance of 0.1 min was applied, instead of 0.05 min. The final pre-processing step was removing mass features detected before 1 min and after 10 min with peak list row filter (RT 1.0–10.0). Peaks with and without MS/MS data were kept. The data were exported to MetaboAnalyst file and were filtered as described in Chapter 4.

For statistics the generated peak lists were uploaded to MetaboAnalyst V4.0 (Chong *et al.*, 2019). In order to generate the heatmap, the empty values in the dataset were replaced by the minimum nonzero value in the dataset divided by 5. Then the values were log2 transformed. Subsequently pareto scaling was applied to the dataset. To create the dendrograms, Euclidean distance measure and Ward's clustering algorithm was used. To generate the Principal Component Analysis Score plot (PCA), the PCA function in MetaboAnalyst was used. To determine whether the peak areas were significantly changed a two-sided unpaired *t*-test was performed, assuming equal variance. *p*-values below 0.05 were considered as significant. The peak area of each mass feature of a mutant was compared with

the peak area of the corresponding mass feature of the wild-type strain (Grace, 2016). For multiple testing corrections, the Benjamini-Hochberg adjustment was applied to the obtained *p*-values (Grace, 2016).

Molecular networking

Pre-processing LC-MS data in MZmine 2.53 and MZmine 2.37.1 corr17.7
For molecular networking using GNPS, the Feature Based Molecular Networking (FBMN) platform was used in combination with ion identity networking (IIMN) (Nothias *et al.*, 2020, Schmid, 2020). The data were first processed using MZmine version 2.53 and subsequently further processed in MZmine version 2.37.1.corr17.7 (Phelan, 2020, Xie *et al.*, 2020). Data processing in MZmine for statistical analysis in MetaboAnalyst was executed as in Chapter 4, with a few exceptions: the join alignment algorithm was executed using and RT range of 0.1 min, followed by gap-filling with an RT range of 0.05 min.

Data filtering
Data filtering was also executed as in Chapter 4, with one exception: the datasets created with ion identity networking (feature quantification table) were not only filtered by removal of mass features that were also present in the blanks, but also by removal of mass features that were detected in the *lug-pks* mutant. The latter was executed with the same method as for removal of mass features detected in the blanks.

Generation feature-based ion identity molecular network
The data were submitted to the GNPS web tool and a network was generated using FBMN. The same filtration method was used as described in Chapter 4, with one exception: a molecular network was generated where the edges were filtered to have more than 6 matched peaks, instead of 7. The molecular networks were visualised using the software Cytoscape (Kohl *et al.*, 2011).

Bioinformatics

The protein sequences of all the angucycline BGCs of known compounds were picked according to literature. Protein sequences were retrieved from the NCBI database or from MIBiG (Kautsar *et al.*, 2020). A phylogenetic tree was generated using MEGA-X (Kumar *et al.*, 2018). The sequences were aligned using the MUSCLE

algorithm. A Maximum Likelihood tree was generated with a bootstrap value of 500. For the visualisation an annotation the software tool iTol was used (Letunic & Bork, 2021). Protein similarity searches were executed using the Basic Local Alignment Search Tool against the NCBI database. GBK files from biosynthetic gene clusters were derived from antiSMASH (Blin *et al.*, 2021). The BGCs were drawn using Clinker (Gilchrist & Chooi, 2021).

SUPPLEMENTAL TABLES

Table S1 Production of metabolites produced by strains used in this study when grown on MM

	Strain						
Molecule	Δ*lugOI*	Δ*lugOII*	Δ*lugOIII*	Δ*lugOIV*	Δ*lugOV*	*Δlug-pks*	Wild type
1	-	-	-	+	-	-	+
2	-	-	+	+	+	-	+
3	-	-	+	+	+	-	+
4	+	+	+	+	+	-	+
5	-	-	-	-	-	-	-
6	-	-	+	+	+	-	+
7	-	-	-	+	-	-	+
8	-	-	-	+	-	-	+
9	-	-	-	+	-	-	+
10	-	-	-	+	-	-	+
11	-	-	-	-	-	-	-

Table S2 Production of metabolites produced by strains used in this study when grown on R5

	Strain						
Molecule	Δ*lugOI*	Δ*lugOII*	Δ*lugOIII*	Δ*lugOIV*	Δ*lugOV*	Δ*lug-pks*	Wild type
1	-	-	-	-	-	-	-
2	-	-	+	+	+	-	+
3	-	-	+	+	+	-	+
4	+	+	+	+	+	-	+
5	+	+	+	+	+	-	+
6	-	-	+	+	+	-	+
7	-	-	-	+	-	-	+
8	-	-	-	+	-	-	+
9	-	-	-	+	-	-	+
10	-	-	-	-	-	-	-
11	+	-	+	+	+	-	+

SUPPLEMENTAL FIGURES

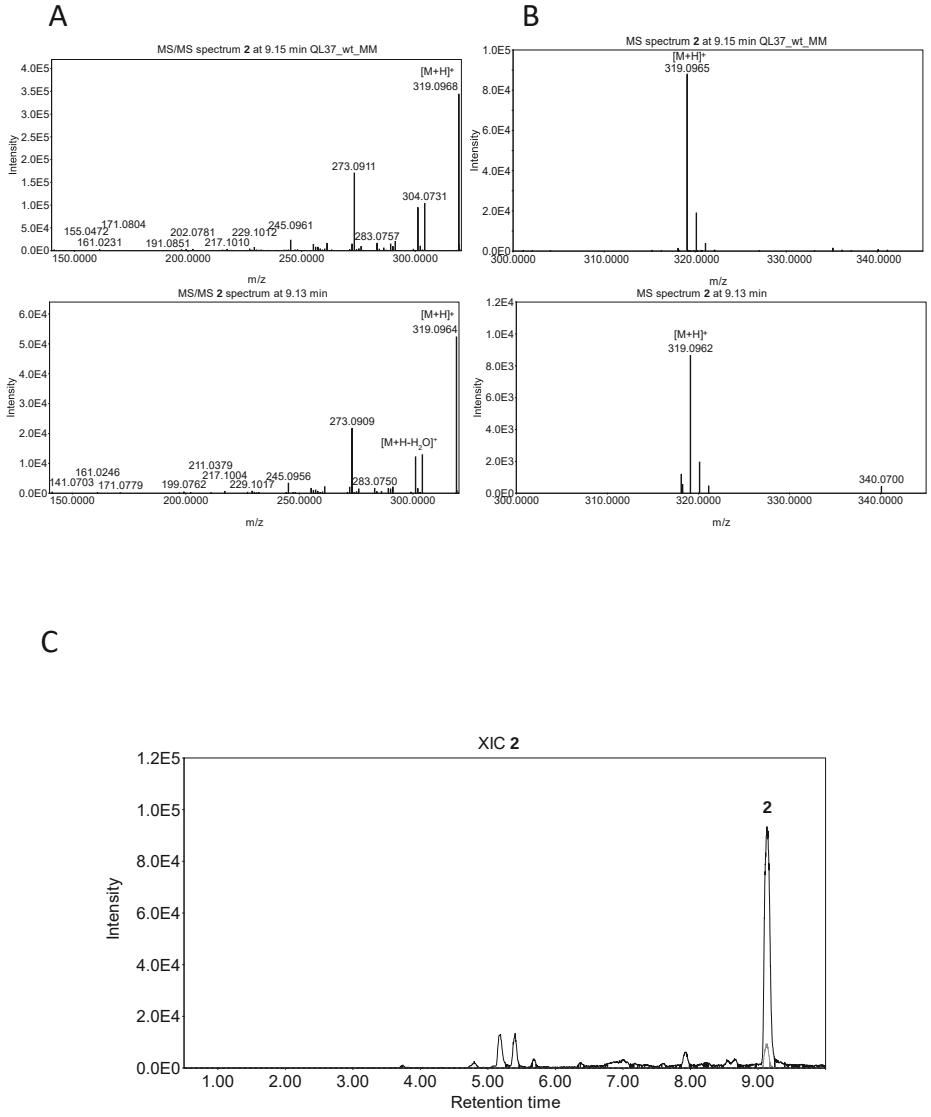


Figure S1 Identification of compound 2 in the extract of *Streptomyces* sp. QL37
Comparison of the MS/MS (A) and MS spectra (B) of the semi-pure compound **2** and its corresponding peak in the crude extract of wild-type *Streptomyces* sp. QL37. The extracted ion chromatograms of the two peaks are shown in (C).

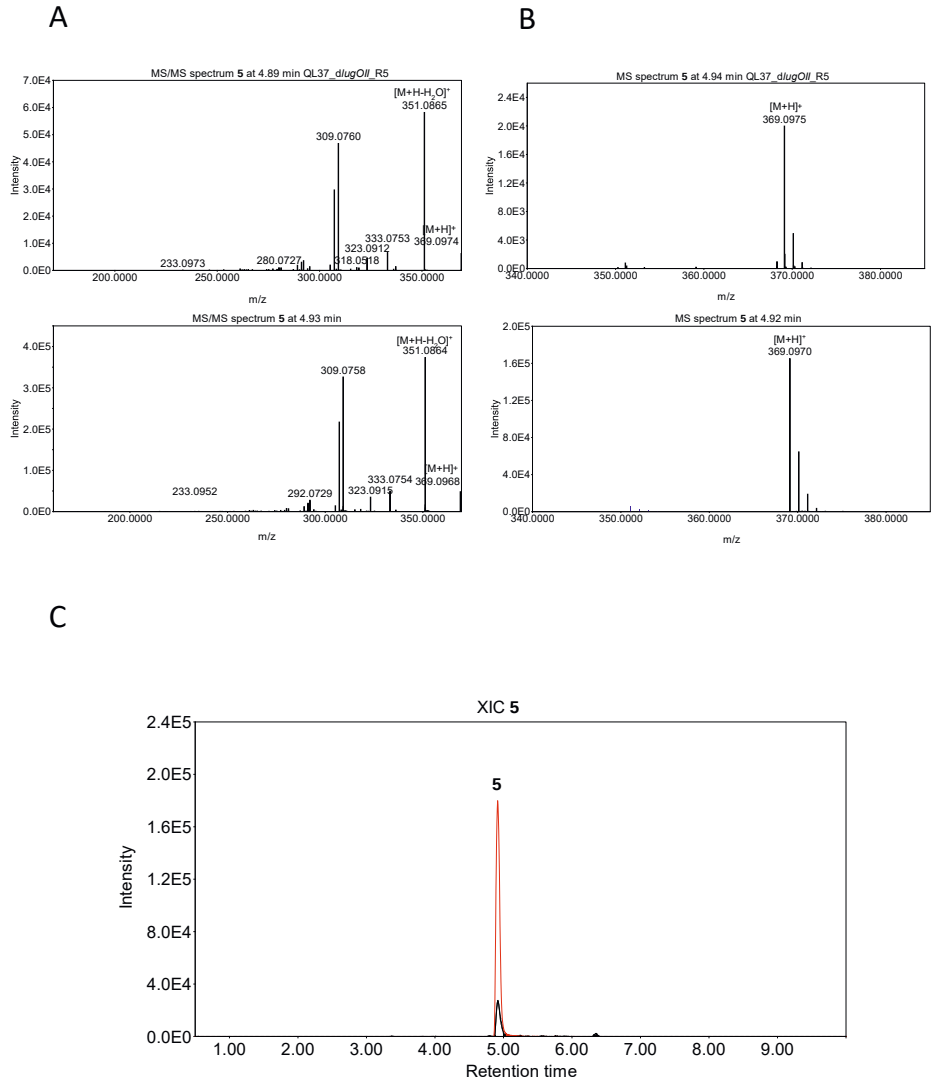


Figure S2 Identification of compound 5 in the extract of *Streptomyces* sp. QL37
Comparison of the MS/MS (A) and MS spectra (B) of the semi-pure compound **5** and its corresponding peak in the crude extract of *lugOII* mutant of *Streptomyces* sp. QL37. The extracted ion chromatograms of the two peaks are shown in (C).

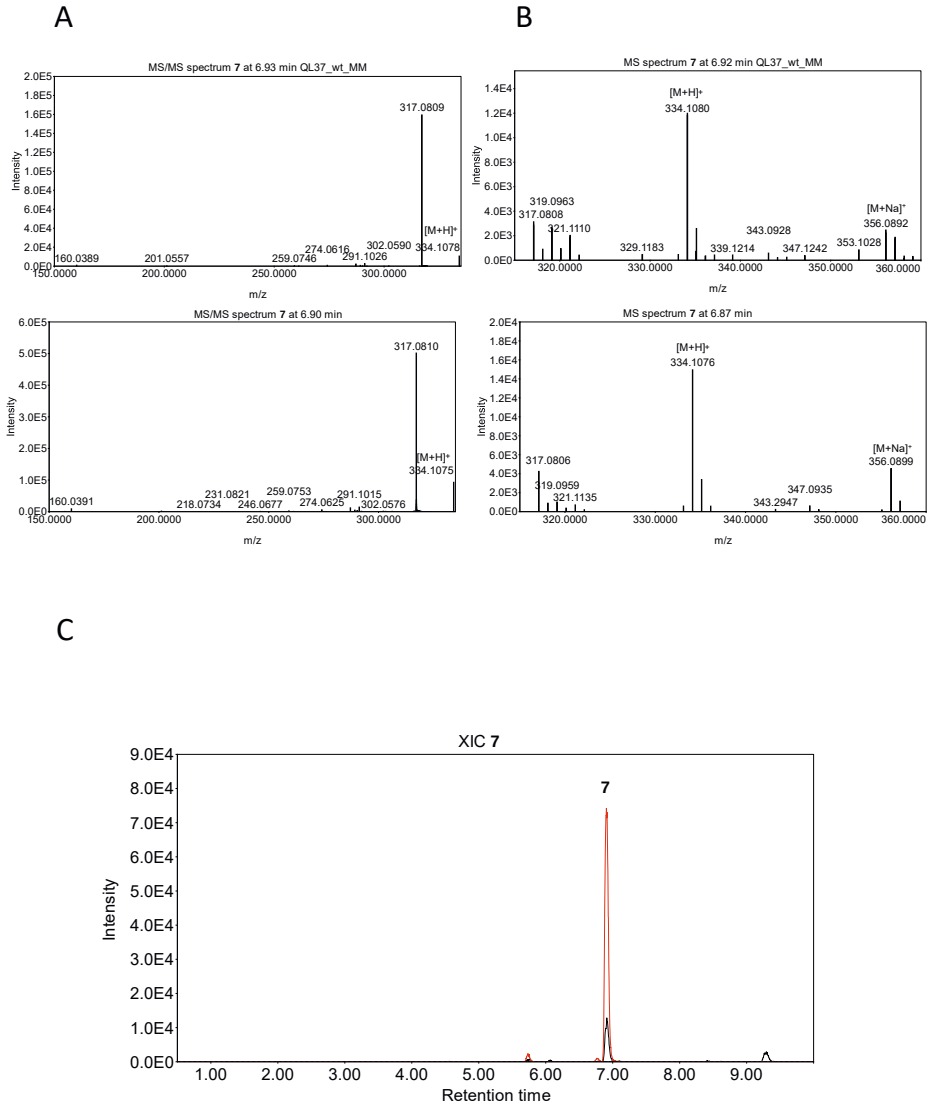


Figure S3 Identification of compound 7 in the extract of *Streptomyces* sp. QL37
Comparison of the MS/MS (A) and MS spectra (B) of the semi-pure compound **7** and its corresponding peak in the crude extract of *Streptomyces* sp. QL37. The extracted ion chromatograms of the two peaks are shown in (C)

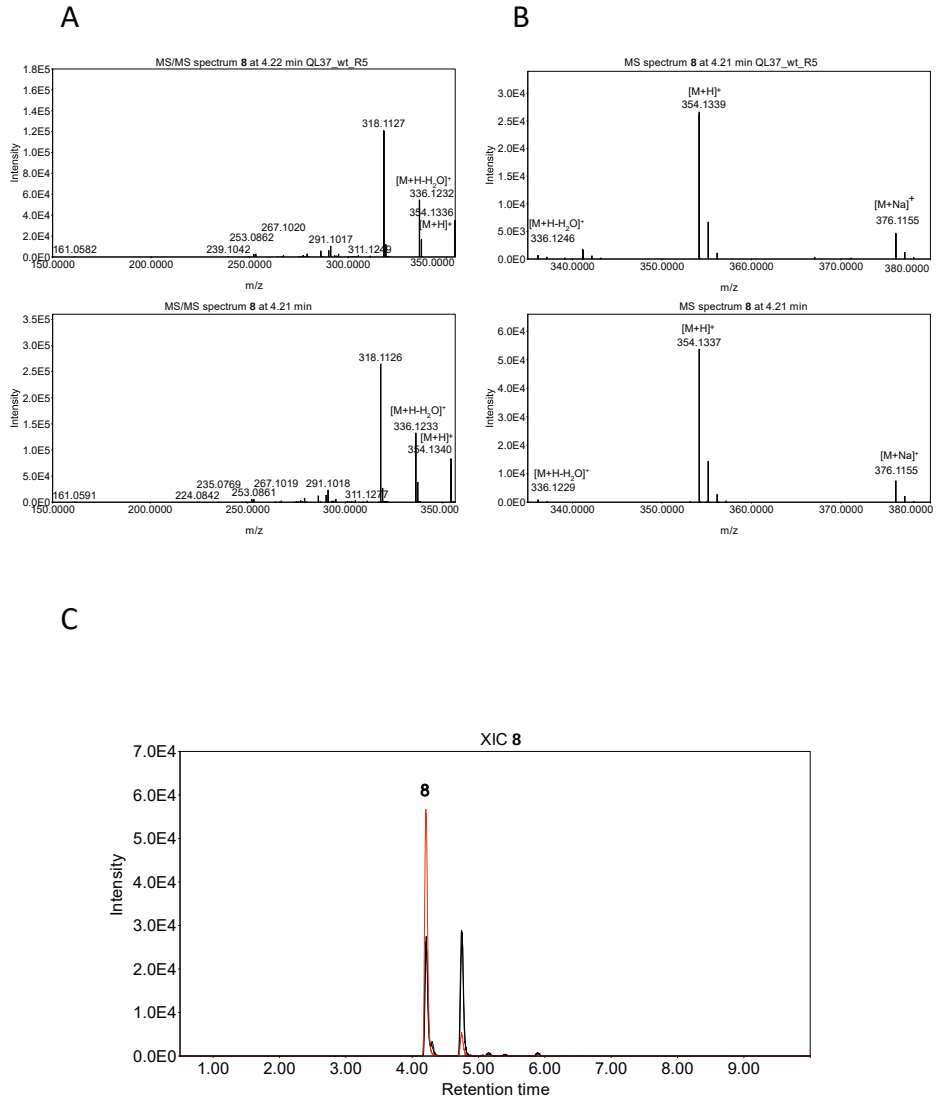


Figure S4 Identification of compound 8 in the extract of *Streptomyces* sp. QL37
Comparison of the MS/MS (A) and MS spectra (B) of the semi-pure compound **8** and its corresponding peak in the crude extract of *Streptomyces* sp. QL37. The extracted ion chromatograms of the two peaks are shown in (C).

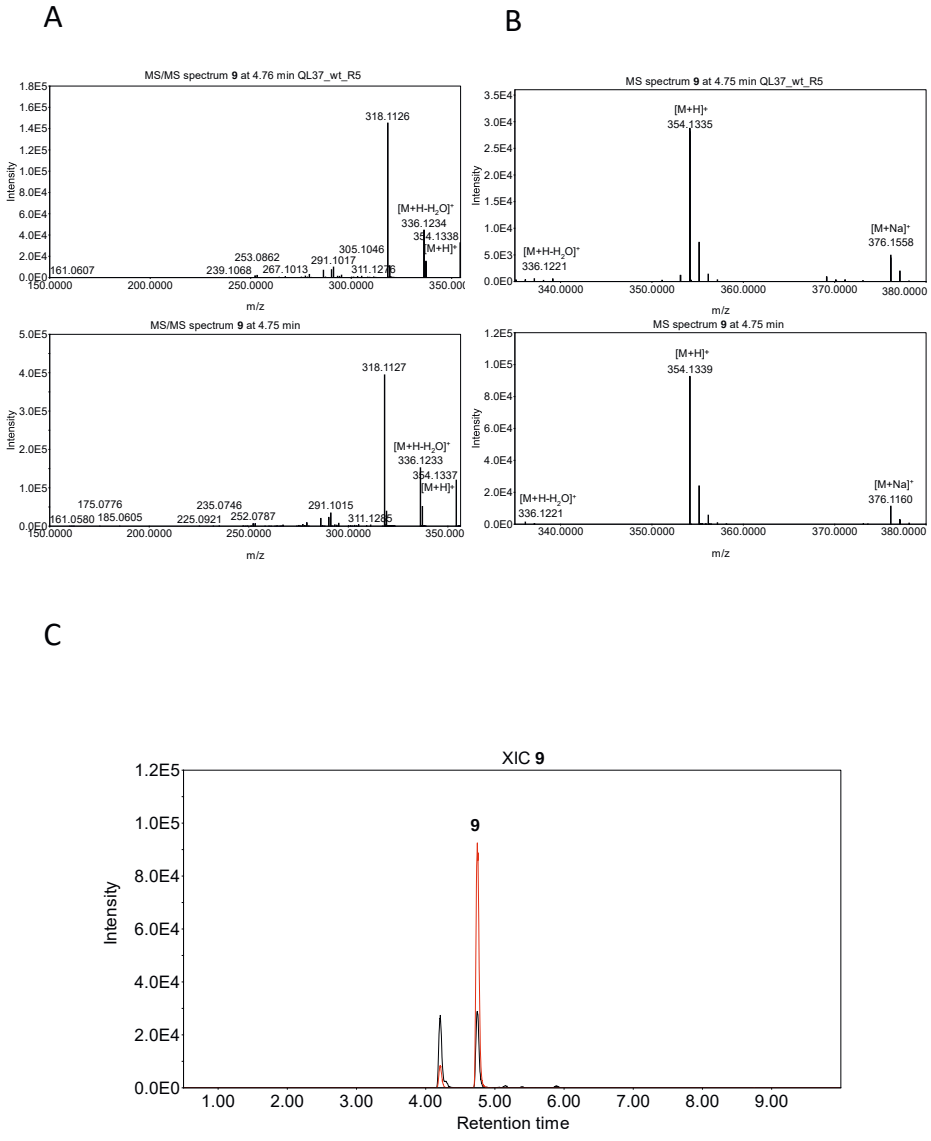


Figure S5 Identification of compound 9 in the extract of *Streptomyces* sp. QL37
Comparison of the MS/MS (A) and MS spectra (B) of the semi-pure compound **9** and its corresponding peak in the crude extract of *Streptomyces* sp. QL37. The extracted ion chromatograms of the two peaks are shown in (C).

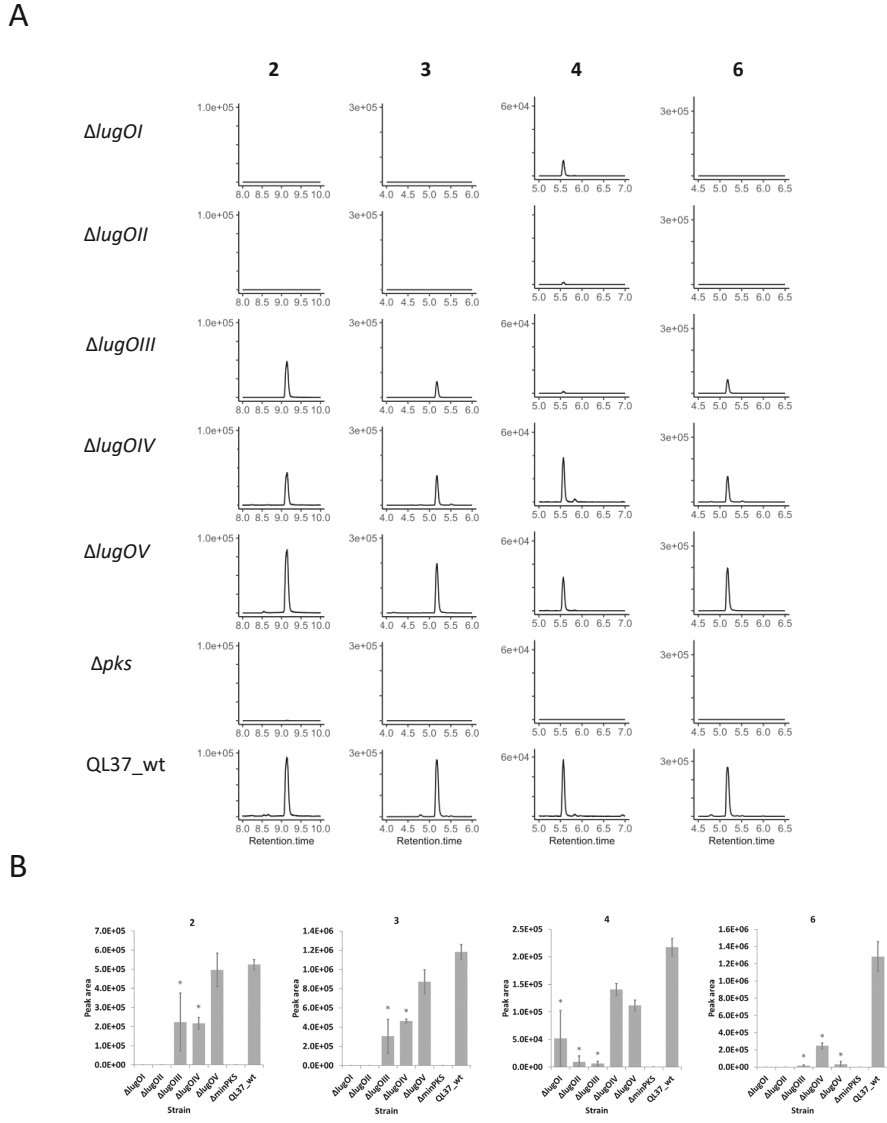


Figure S6 Production of non-rearranged angucyclines by *Streptomyces* sp. QL37, its lug-pks mutant and its *lugOI–lugOV* mutant, when grown on MM.
Extracted ion chromatograms of the non-rearranged angucyclines (**2**, **3, 4**, and **6**) in the extracts of *Streptomyces* sp. QL37 and its mutants of *lugOI-lugOV* and of the *lug-pks* mutant grown on MM (A). Bar plots of the peak areas of the previously isolated molecules from *Streptomyces* sp. QL37. (B). The error bars represent the 95% confidence intervals. The asterisk indicates the peak area of mass feature in the mutant is significantly different from the peak area of the corresponding mass feature in the wild type with a *p*-value of <=0.05.

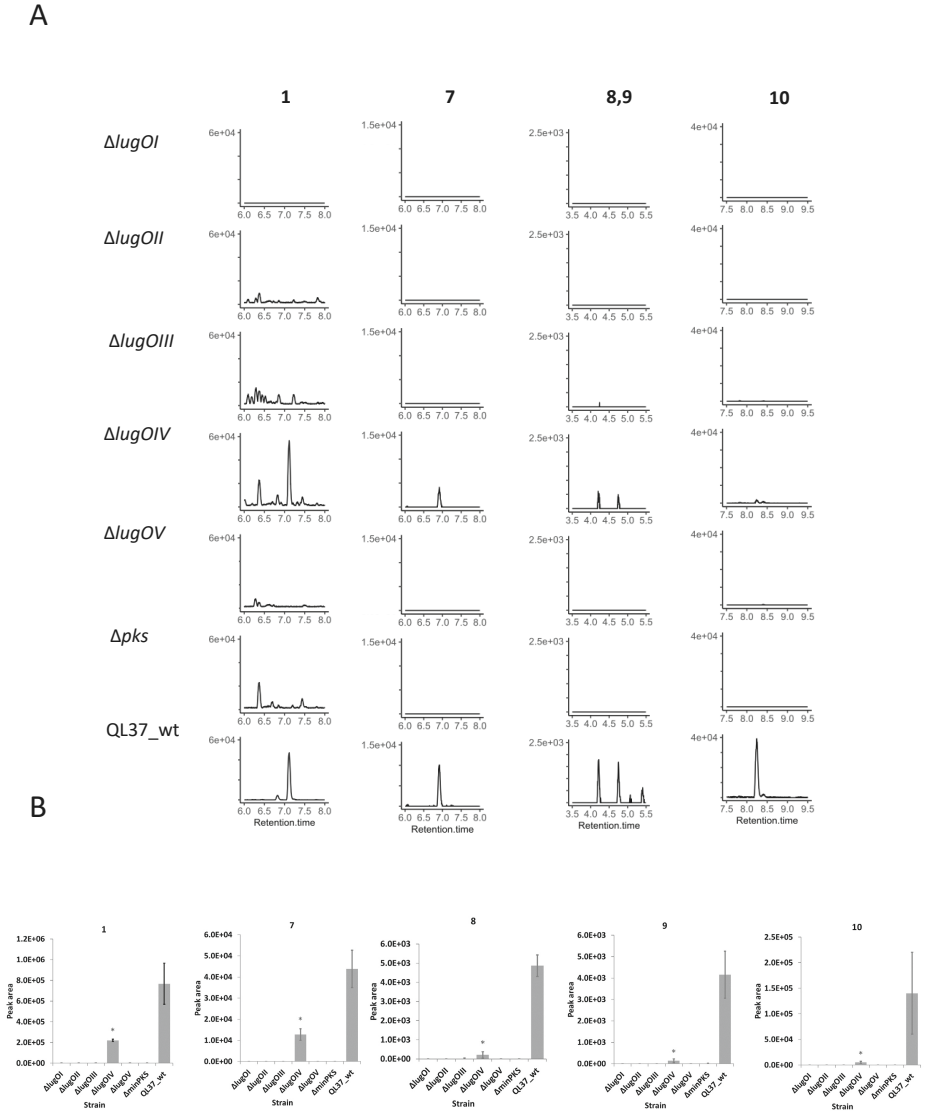


Figure S7 Production of C-ring-rearranged angucyclines by *Streptomyces* sp. QL37, its *lug-pks* mutant and its *lugOI–lugOV* mutant, when grown on MM.
Extracted ion chromatograms of the rearranged angucyclines (**1** $[M+H]^+$, **7** , **8** , **9** and **10**) in the extracts of *Streptomyces* sp. QL37 and its mutants of *lugOI–lugOV* and of the *lug-pks* mutant grown on MM (A). Bar plots of the peak areas of the previously isolated molecules from *Streptomyces* sp. QL37 (B). The error bars represent the 95% confidence intervals. The asterisk indicates the peak area of mass feature in the mutant is significantly different from the peak area of the corresponding mass feature in the wild type with a *p*-value of <=0.05.

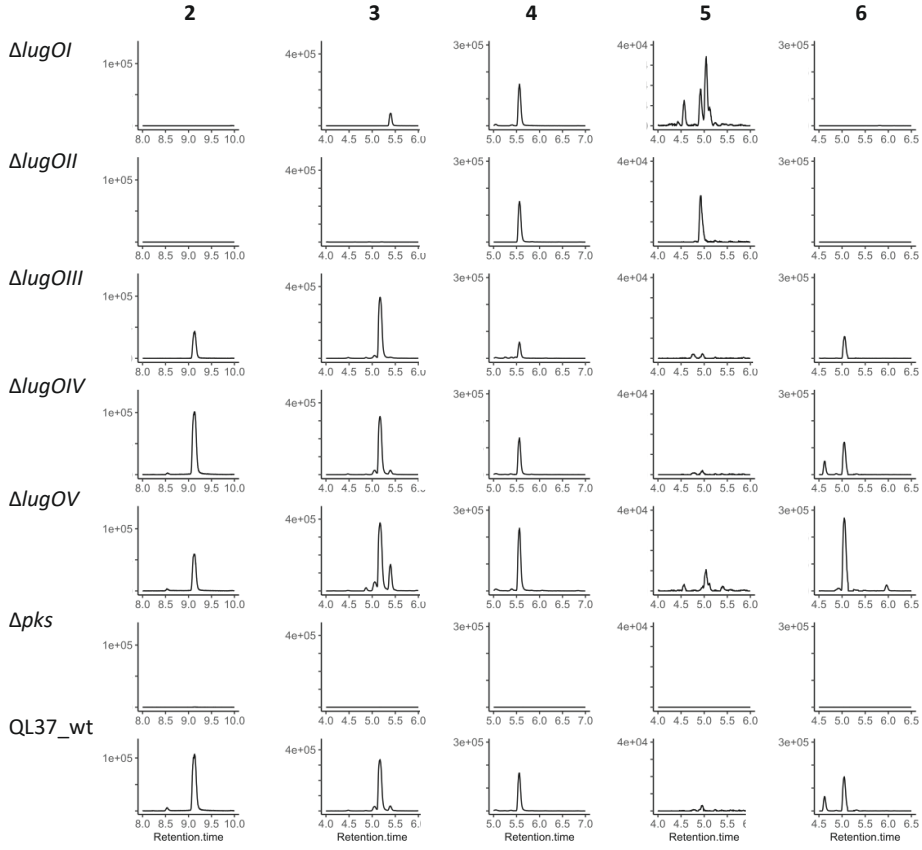


Figure S8A Production of non-rearranged angucyclines by *Streptomyces* sp. QL37, its lug-pks mutant and its *lugOI–lugOV* mutant, when grown on R5.
Extracted ion chromatograms of the typical angucyclines (**2**, **3**, **4**, **5**, and **6**) in the extracts of *Streptomyces* sp. QL37 and its mutants of *lugOI-lugOV* and of the *lug-pks* mutant grown on R5. Bar plots of the peak areas of the previously identified molecules involved in lugdunomycin biosynthesis (B). The error bars represent the 95% confidence intervals. The asterisk indicates the peak area of mass feature in the mutant is significantly different from the peak area of the corresponding mass feature in the wild type with a *p*-value of <0.05.

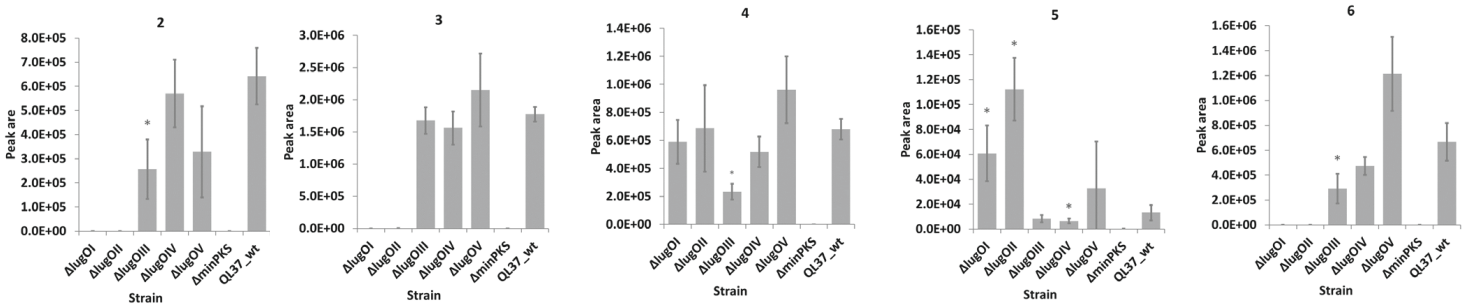


Figure S8B Production of non-rearranged angucyclines by *Streptomyces* sp. QL37, its lug-pks mutant and its *lugOI–lugOV* mutant, when grown on R5.

Bar plots of the peak areas of the previously identified molecules involved in lugdunomycin biosynthesis (B). The error bars represent the 95% confidence intervals. The asterisk indicates the peak area of mass feature in the mutant is significantly different from the peak area of the corresponding mass feature in the wild type with a *p*-value of <0.05.

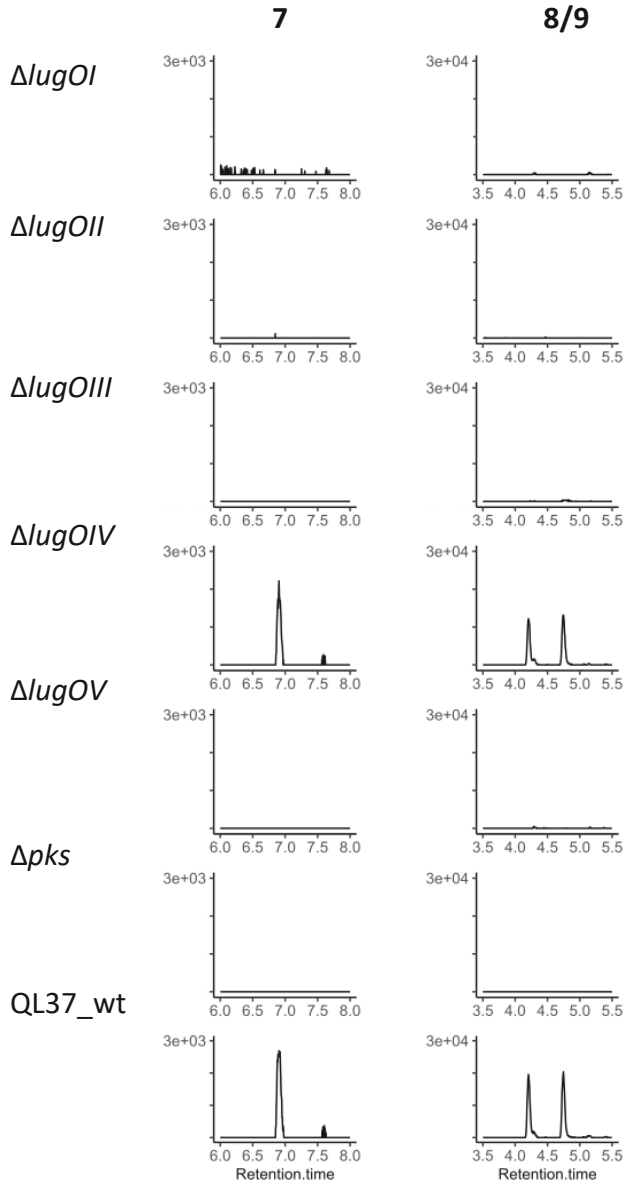


Figure S9A Production of C-ring-rearranged angucyclines by *Streptomyces* sp. QL37, its lug-pks mutant and its *lugOI–lugOV* mutant, when grown on R5.
Extracted ion chromatograms of the typical angucyclines (**7**, **8** and **9**) in the extracts of *Streptomyces* sp. QL37 and its mutants of *lugOI-lugOV* and of the *lug-pks* mutant grown on R5.

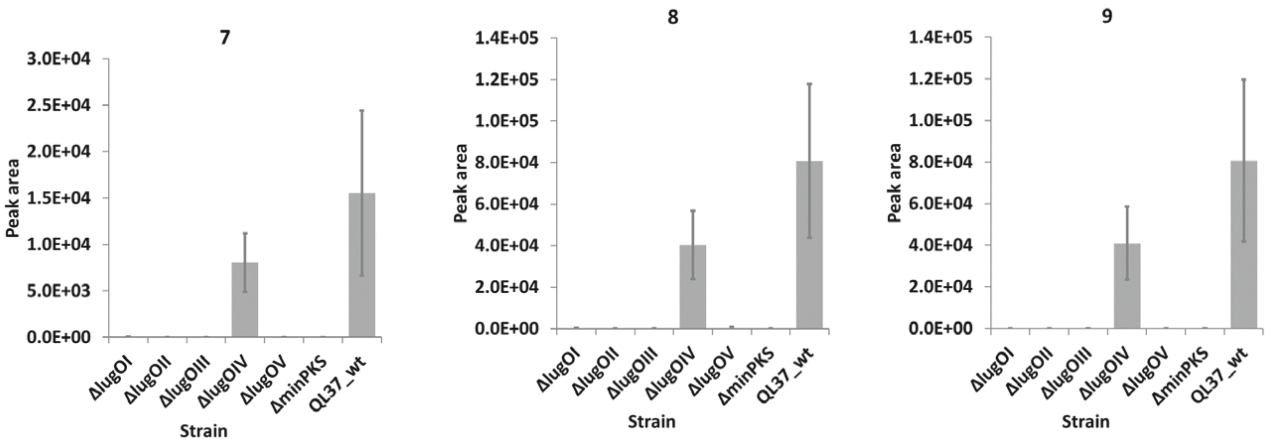


Figure S9B Production of C-ring-rearranged angucyclines by *Streptomyces* sp. QL37, its lug-pks mutant and its *lugOI–lugOV* mutant, when grown on R5.

Bar plots of the peak areas of the previously identified molecules involved in lugdunomycin biosynthesis (B). The error bars represent the 95% confidence intervals. The asterisk indicates the peak area of mass feature in the mutant is significantly different from the peak area of the corresponding mass feature in the wild type with a *p*-value of <0.05.

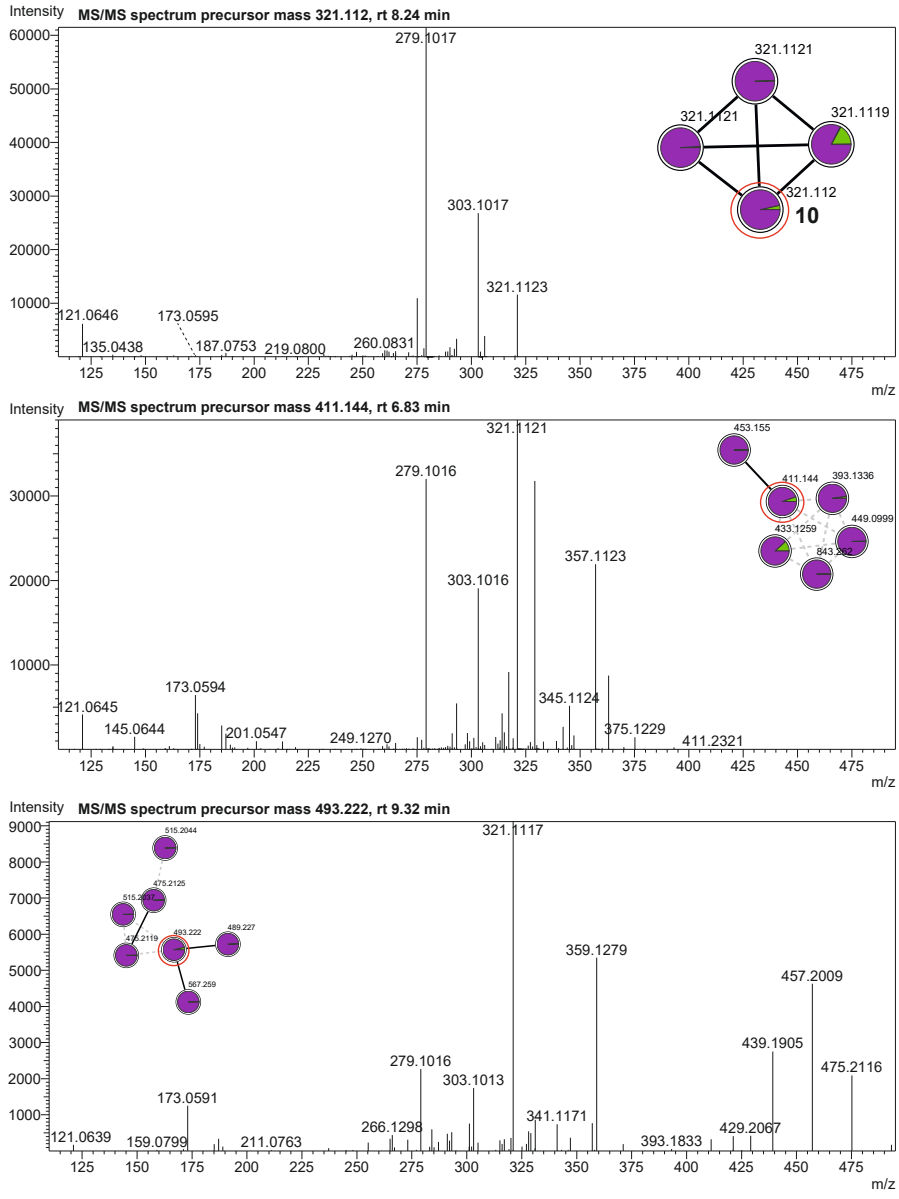


Figure S10 MS/MS spectrum of precursor masses related to elmonin and their molecular families.

Molecular families with similar MS/MS fragmentation as elmonin. MS/MS spectrum of the elmonin (**10**) (top) as compared to the MS/MS spectra of other mass features in different molecular families, which showed a fragment ion with the same mass as **10** that further fragments similarly. The nodes whose MS/MS spectra are given are highlighted.

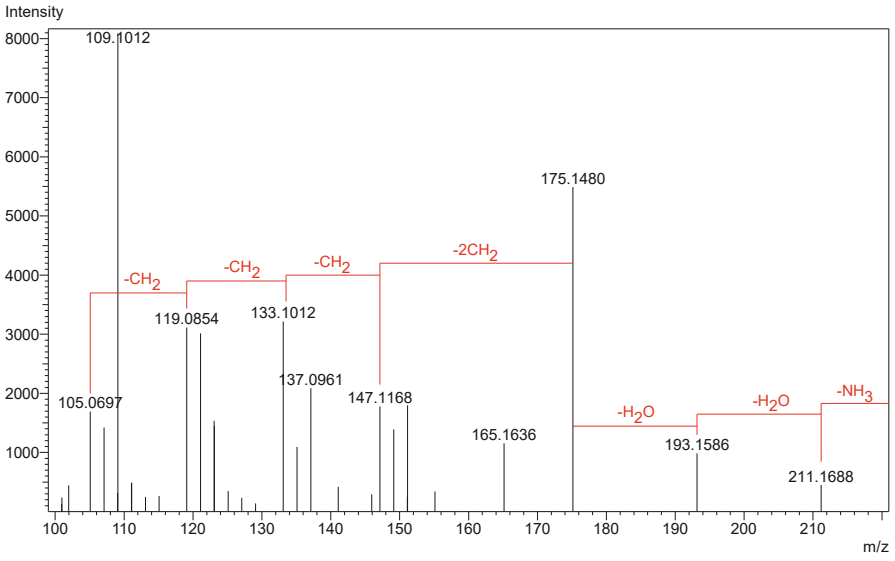


Figure S11 MS/MS spectrum dihydroxy amino alcohol.
MS/MS spectrum of a node in one of the molecular families of the MM extracts, which showed a fragmentation pattern consistent with an unsaturated aliphatic compound containing dihydroxy and a monoamino substitution. This is evident through the observation of a loss of an NH_3, followed by two molecules of H_2O, followed by a series of CH_2 groups.

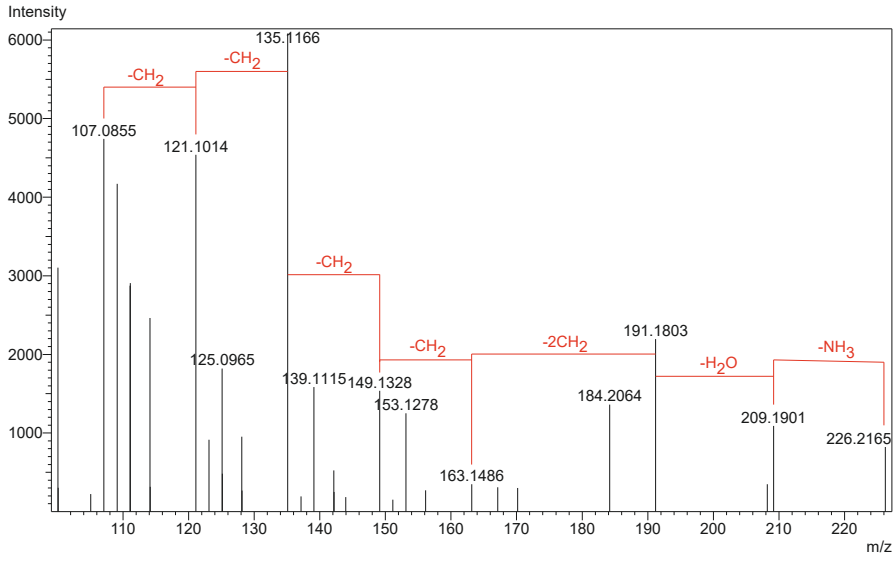


Figure S12 MS/MS spectrum monohydroxy amino alcohol.
MS/MS spectrum of a node in one of the molecular families of the MM extracts, which showed a fragmentation pattern consistent with an unsaturated aliphatic compound containing a monohydroxy and a monoamino substitution. This is evident through the observation of a loss of an NH_3, followed by an H_2O, followed by a series of CH_2 groups

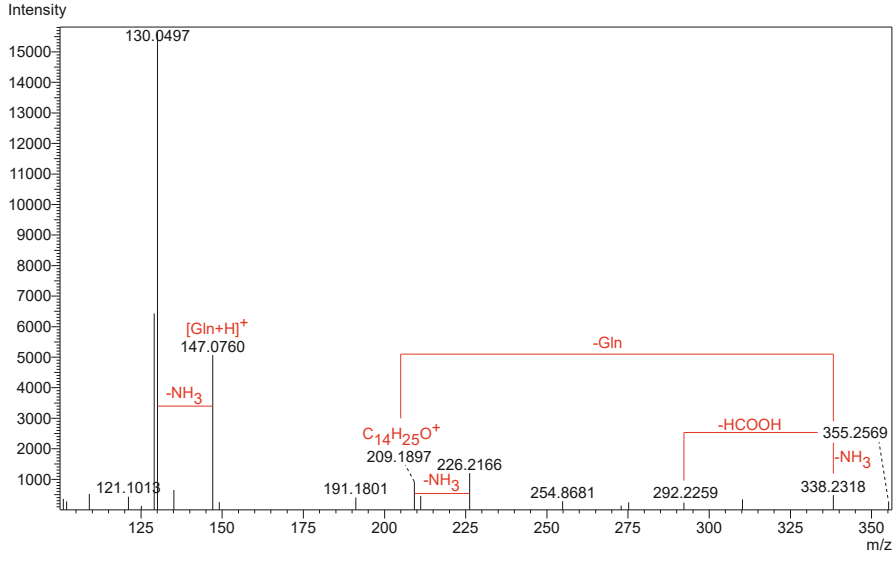


Figure S13 MS/MS spectrum acyl glutamine.
MS/MS spectrum of a node in one of the molecular families of the MM extracts, which showed a fragmentation pattern consistent with an unsaturated fatty acyl chain that is conjugated to glutamine amino acid.

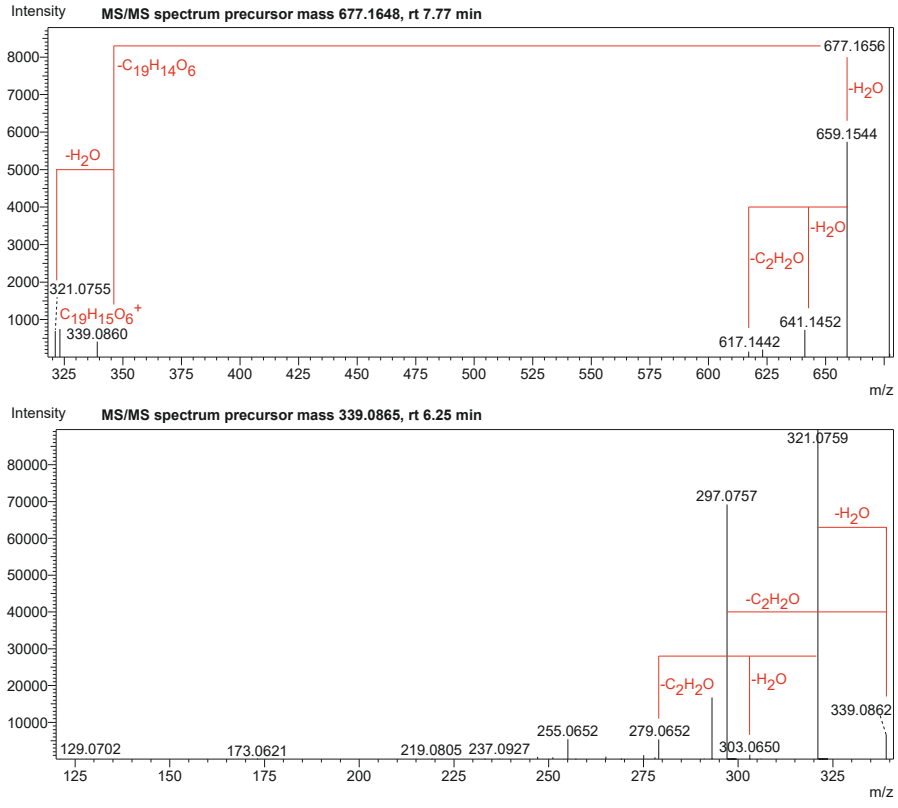
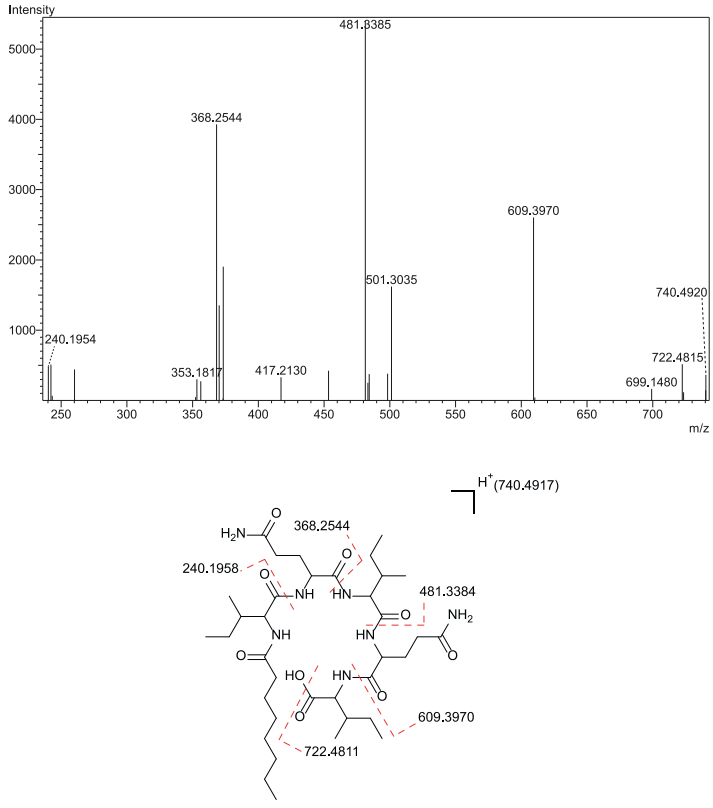


Figure S14 MS/MS spectrum of rabelomycin dimer.
MS/MS spectrum of a node in one of the molecular families of the R5 extracts representing a likely dimer of two rabelomycin molecules (top), as compared to the MS/MS spectrum of rabelomycin (bottom). The dimer shows in the MS/MS spectrum a loss of a rabelomycin molecule ($C_{19}H_{14}O_6$) resulting in the fragment ion of the second rabelomycin ($C_{19}H_{15}O_6^+$). Additionally, a similar fragmentation pattern could be observed for both the dimer and rabelomycin.

A



B

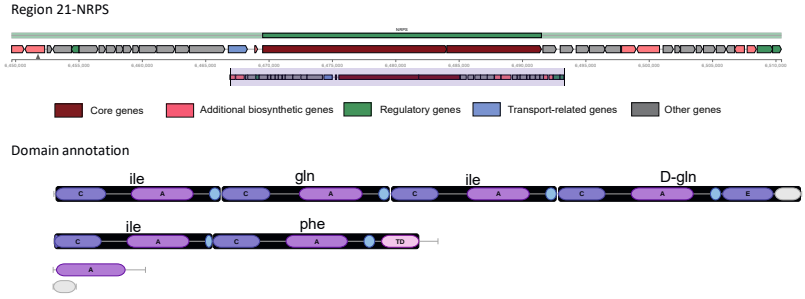


Figure S15 MS/MS spectrum of proposed lipopeptide, its structure and its putative BGC. MS/MS spectrum of a node in one of the molecular families of the R5 extracts, which showed a fragmentation pattern consistent with a lipopeptide metabolite. A proposed structure on which the major fragments are shown is given below the spectrum (A). Predicted nonribosomal peptide synthetase (NRPS) genes, identified by antiSMASH possibly involved in the synthesis of the proposed lipopeptide (B).

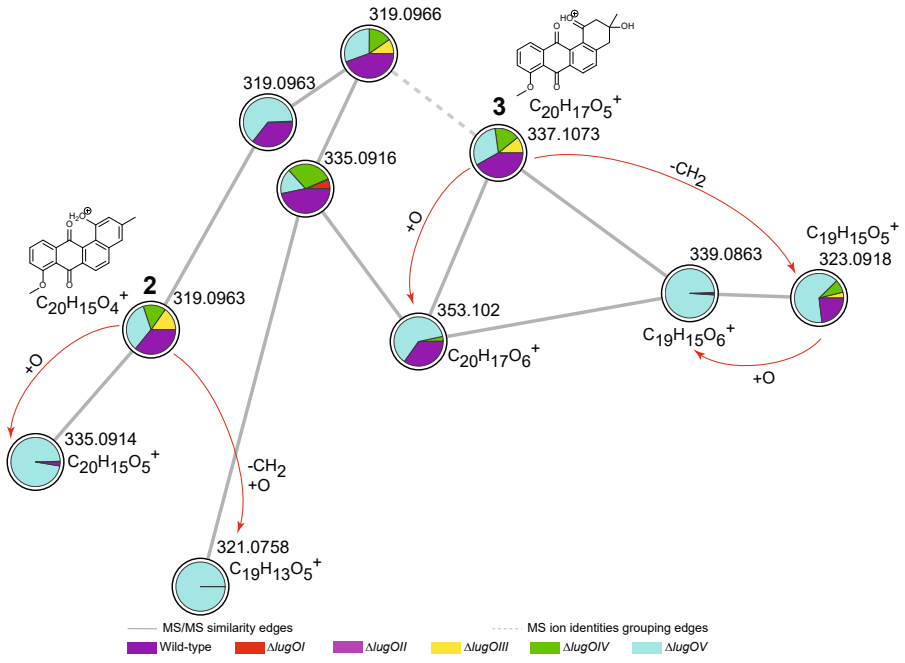


Figure S16 Relationship between nodes upregulated in the *lugOV* mutant grown on MM. Part of the non-rearranged angucycline molecular family in the MM extracts showing the likely structural relationship between the previously identified metabolites **2** and **3**, and the metabolites which were upregulated in the *lugOV* mutant.

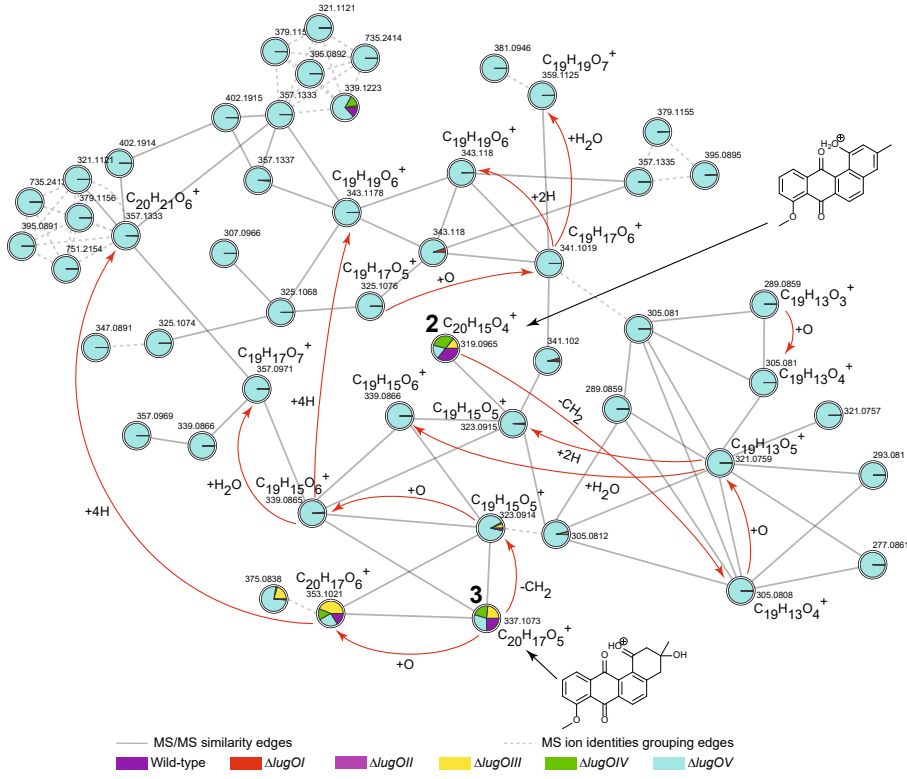


Figure S17 Relationship between nodes upregulated in the *lugOV* mutant grown on R5. Part of the non-rearranged angucycline molecular family in the R5 extracts showing the likely structural relationship between the previously identified metabolites **2** and **3**, and the metabolites which were upregulated in the *lugOV* mutant.

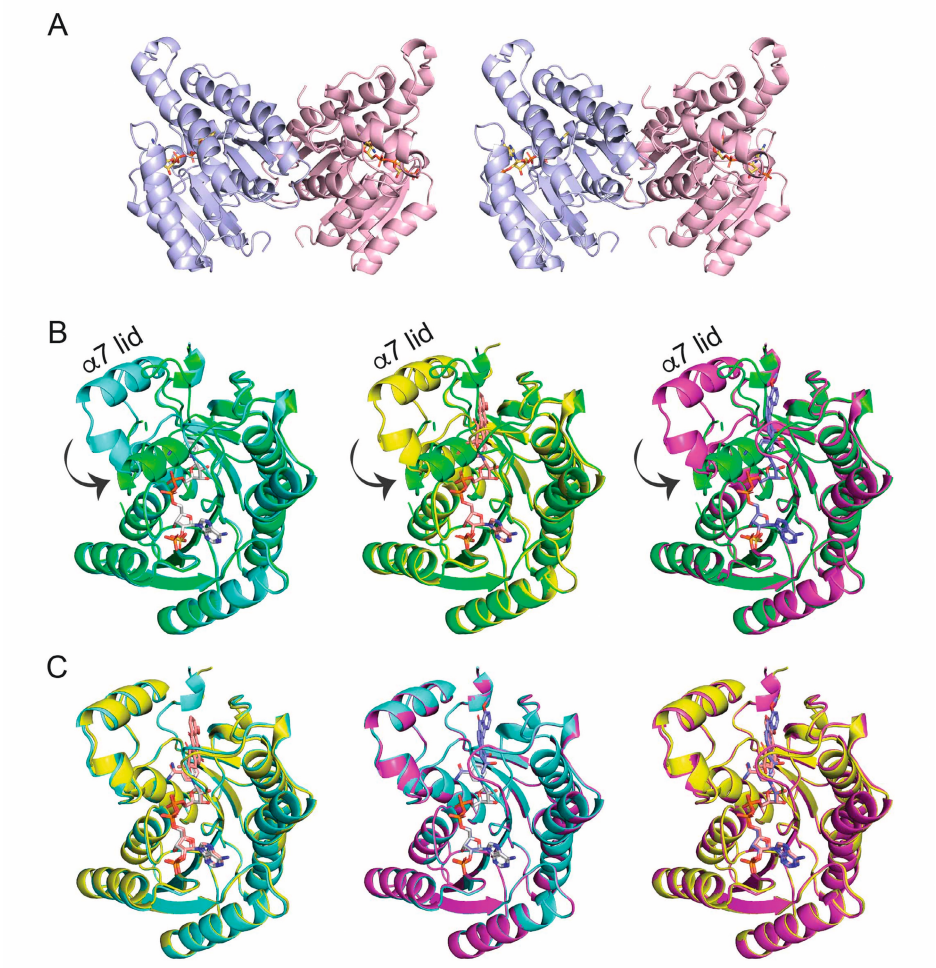


Figure S19 Arrangement of the dimeric structures of LugOIIred and the observed conformational changes between the LugOIIred structures.
A) Stereoscopic view of NADPH bond LugOIIred in dimeric form. The NADPH is displayed as sticks. B) Unliganded LugOIIred (green), LugOIIred/NADPH (cyan), LugOIIred/NADPH/**4** (yellow) and LugOIIred/NADPH/**3** (magenta) structures are superimposed. a6, a7 and the loop region between them serve as a lid, which turn around 180° and further rotate 90° towards the binding site of **3**. C) The alignment of LugOIIred/NADPH (cyan), LugOIIred/NADPH/**4** (yellow) and LugOIIred/NADPH/**3** (magenta) structures. NADPH, **3** and **4** are displayed in sticks (cited, (Xiao, 2020)).

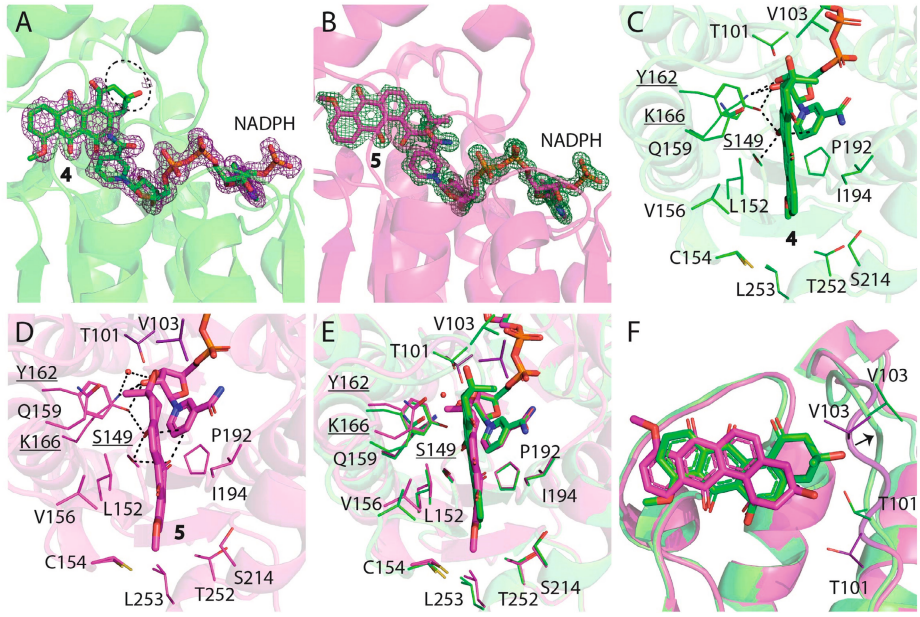


Figure S20 The LugOIIred active site.
The LugOIIred active site. A&B: 2Fo-Fc omit maps contoured at the 1s level corresponding to ligand **3**, **4** and cofactor NADPH. The missing density for **4** is highlighted with black circle. C&D: key residues that surround the binding site of **3** and **4**. Catalytic residues S149, Y162 and K166 are underlined and their distances (within 3.2 Å) to the ligand and cofactor dashed. E&F: Superposition of the two substrates **3** and **4** bounded to LugOIIred. E) Top view of the active pocket. F) Side view of the two aligned substrate structures. Major differences are found in the orientations of the two substrates and the movement of the a4-b4 loop that are close to the A-rings of the two substrates. (cited, (Xiao, 2020)).

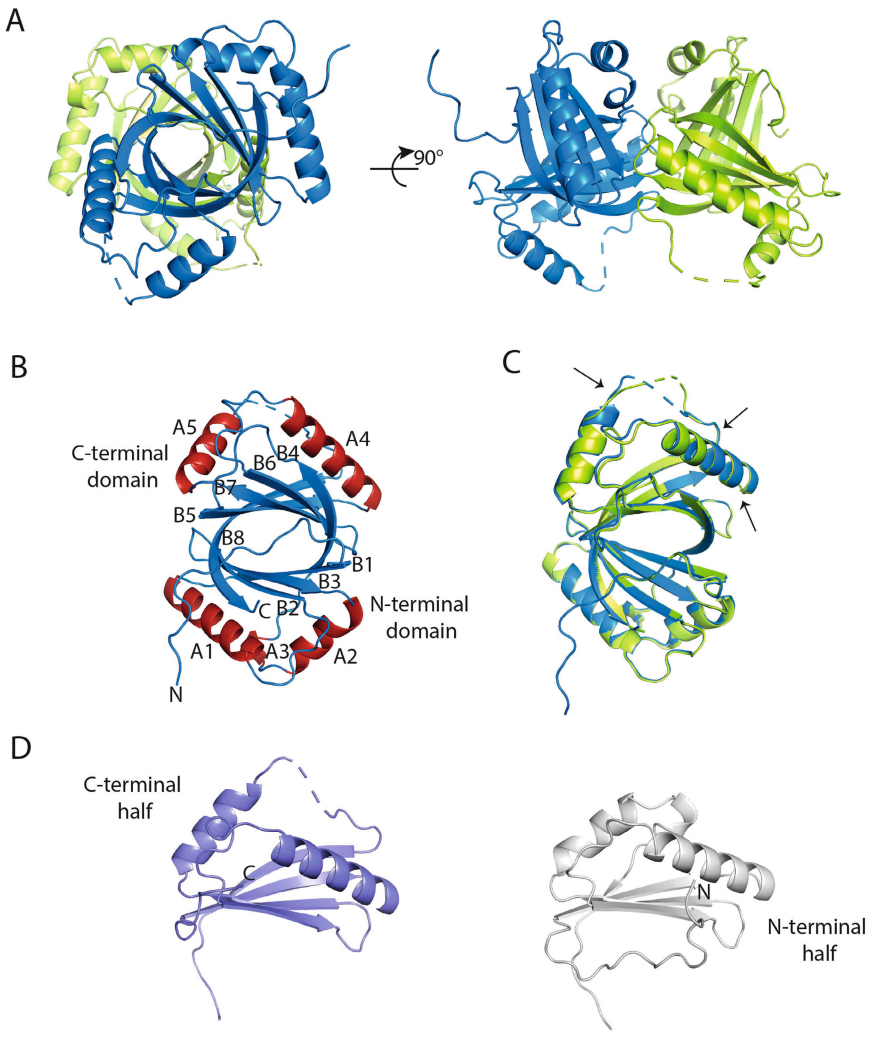


Figure S21 Dimeric arrangement of LugOIII structure.
A) Top and side views of LugOIII in dimeric form. B) LugOIII is consisted by eight b-strands and five a-helices. C) Structural alignment of the two chains of LugOIII structure. Three b-turns are highlighted with arrows. D) The alignment of the N-terminal half and C-terminal half of LugOIII structure (cited (Xiao, 2020)).

Analysis of the transcriptional regulatory genes of the biosynthetic gene cluster for angucyclines, limamycins and lugdunomycin in *Streptomyces* sp. QL37

Helga U. van der Heul
Chao Du
Lukas Kiefer
Somayah S. Elsayed
Gilles P. van Wezel

ABSTRACT

The *lug* gene cluster in *Streptomyces* sp. QL37 directs the production of non-rearranged angucyclines an C-ring cleaved angucyclines, such as the limamycins, pratensilin A and elmonin. In addition, the BGC is required for the production of lugdunomycin, a highly rearranged angucycline with an unprecedented scaffold. The extremely low production levels have so far prevented studies of mode of action and pharmacological properties of this molecule. To improve its production, more insight into the regulatory networks that controls its expression is needed. Here we present a mutational and functional analysis of four regulatory genes *lugRII-RV* of the angucycline biosynthetic gene cluster (*lug*) that is required for lugdunomycin biosynthesis, and report on the global changes in terms of development, metabolism, and protein expression. Our work shows that *lugRII*, *lugRIV* and *lugRV* are required for the expression of the BGC, and hence for production of angucyclines, limamycins and lugdunomycin. Deletion and overexpression experiments revealed that the orphan response regulator LugRIV and the SARP regulator LugRV are cluster-situated activators of the *lug* gene cluster, as was validated by metabolomics analysis, together with the upregulation of the *lug* proteins in the early stages of growth. Taken together this study shows that LugRIV and LugRV are the main positive regulators of the *lug* gene cluster and can be employed to increase angucycline production and for premature lugdunomycin production by *Streptomyces* sp. QL37.

INTRODUCTION

New antibiotics are desperately required, as the emergence of novel antimicrobial resistance pathogens increases (Lewis, 2013). *Streptomyces* are filamentous soil bacteria with a complex life cycle and are well known antibiotic producers. They have the potential to produce numerous natural products as is indicated by the number of biosynthetic gene clusters (BGCs) located on their genome. An average *Streptomyces* genome can contain more than 50 BGCs (Katz & Baltz, 2016). However, many of these are not or poorly expressed under laboratory conditions, also referred to as cryptic BGCs. Therefore various screening methods are applied to awaken and explore these BGCs (Rutledge & Challis, 2015). One of these is the "One strain-many compounds" strategy (OSMAC) (Romano *et al.*, 2018). Application of this method allowed the discovery of lugdunomycin (Wu *et al.*, 2019). The molecule is produced by *Streptomyces* sp. QL37, a strain derived from the Qinling mountains in China (Wu *et al.*, 2019). Besides lugdunomycin (**1**), also non-rearranged (typical) angucyclines (**2**–**6**) and the rearranged angucyclines pratensilin A (**7**), the limamycins (**8**, **9**) and elmonin (**10**) were detected in the extracts of *Streptomyces* sp. QL37 (Figure S1) (Wu *et al.*, 2019, Xiao *et al.*, 2020). A type II polyketide BGC (*lug*) was identified that encodes the biosynthetic enzymes for these molecules ((Wu *et al.*, 2019); Chapter 3). Rearranged angucyclines are derived from the angucycline backbone (**2**,**3**), following cleavage of the C6a–C7 bond of the canonical angucycline C-ring (Wu *et al.*, 2019). It was proposed that lugdunomycin is the product of a final Diels-Alder reaction between a rearranged angucycline and *iso*-maleimycin (**11**) (Wu *et al.*, 2019, Uiterweerd, 2020). While angucyclines are produced at high levels under various growth conditions, lugdunomycin is hardly produced. To obtain more insight into the activity of lugdunomycin, it is essential to improve its production.

One approach to enhance the expression of a BGC, and thereby induce the production of the respective natural products, is to manipulate the regulatory pathway that controls it (van der Heul *et al.*, 2018). In *Streptomyces* the control of antibiotic production is highly complex and is intertwined with the control of the morphological development. Environmental signals are often sensed by pleiotropic regulators that have a global impact on morphological development and specialised metabolism in *Streptomyces* (Liu *et al.*, 2013, van der Heul *et al.*, 2018). These pleiotropic regulators control the transcriptional activity of regulatory

genes within the BGC, which are referred to as cluster-situated regulators (CSRs). In general, the CSR controls the transcription of the specific BGC it is embedded in, though CSRs that control the activity of multiple BGCs have also been reported (van der Heul *et al.*, 2018, Liu *et al.*, 2013). Different CSR families are known; the best known belongs to the *Streptomyces* antibiotic regulatory protein (SARP) transcriptional regulatory family. These regulators activate the transcription of their target BGC. Famous examples are ActII-ORF4 (actinorhodin), RedD (undecylprodigiosin) and DnrI (daunorubicin) (Fujii *et al.*, 1996, Prija *et al.*, 2017, Hulst *et al.*, 2021). For details see Chapter 2.

The data presented in Chapter 4 suggest that the *lug* cluster specifies angucyclines and limamycins, but not lugdunomycin; indeed, a second BGCs is required for the production of lugdunomycin, namely BGC 23 that is most likely involved in the production of *iso*-maleimycin (**11**). Here, we analyse the transcriptional regulatory genes of the *lug* gene cluster. Comparison of the *lug* gene cluster with other *lug*-type clusters revealed that the regulatory genes *lugRII*, *lugRIII*, *lugRIV* and *lugRV* are typically shared among angucycline BGCs with high similarity to the *lug* gene cluster. We then studied the metabolic, proteomic and morphological changes resulting from the deletion and over-expression of the *lugR* genes. These studies demonstrated that LugRIV and LugRV act as the major transcriptional activators of the *lug* gene cluster.

RESULTS

Bioinformatic analysis of the putative regulators from the *lug* gene cluster

The *lug* gene cluster contains five regulatory genes designated as *lugRI*, *lugRII*, *lugRIII*, *lugRIV* and *lugRV*. The regulatory genes *lugRII–lugRV* are all located in BGCs that are highly similar to the *lug* gene cluster, suggesting that these are required for the control of the BGC (Figure 1, Table S1). Only the regulatory gene *lugRI* is absent in the highly related gene clusters, indicated in Figure 1. To predict the functional role of the *lugR* genes the pfam domains were predicted in the amino acid sequences of the regulatory proteins.

LugRI is a predicted 279 amino acid (aa) protein composed of a C-terminal MmyB-like transcription regulator ligand binding domain (pfam: MLTR-LBD) and an N-terminal helix-turn-helix (HTH) DNA-binding domain that belongs to the

XRE-family of transcriptional regulators (pfam: HTH-31). MmyB-like regulators occur in a wide range of bacteria and are associated with antibiotic and fatty acid metabolism in Actinobacteria (Xu *et al.*, 2012). Genes encoding an MmyB-regulator are often associated with genes for short-chain dehydrogenases/reductases, such as *lugM*, which implies a functional connection between *lugRI* and *lugM* (Xu *et al.*, 2012).

LugRII (222 aa) belongs to the family of NarL-type atypical orphan response regulators (ARRs). It contains a C-terminal LuxR-type HTH DNA-binding domain (pfam: GerE), which is found in NarL/FixJ family of response regulators. RedZ and DnrN, the activators of the undecylprodigiosin BGC in *S. coelicolor* and the daunorubicin BGC in *S. peucetius*, respectively, also belong to this regulatory family (Furuya & Hutchinson, 1996, Wang *et al.*, 2009, Guthrie *et al.*, 1998). RedZ and DnrN activate transcription of the pathway-specific activator genes of their cognate BGC, namely *redD* and *dnrI*, respectively (White & Bibb, 1997, Furuya & Hutchinson, 1996).

LugRIII (238 aa) contains a C-terminal AefR transcriptional regulatory protein domain (pfam: TetR_C_7) and an N-terminal TetR-like DNA-binding domain (pfam: TetR_N). Regulators belonging to the TetR family often repress the transcription of divergently oriented transporter genes (Ramos *et al.*, 2005). The repression is often released due to the binding of the end product of the BGC to the TetR regulator. Also in the *lug* gene cluster *lugRIII* lies next to *lugTI,* which encodes a putative transporter.

LugRIV (266 aa) belongs to the OmpR-type regulators. It contains a C-terminal OmpR-like DNA- binding domain (pfam: Trans_reg_C) and an N-terminal domain response regulator receiver domain (pfam: response _Reg). The protein shares significant homology with JadR1 (52% aa identity) and Aur1P (54% aa identity), which are activators of the jadomycin BGC in *S. venezuelae* ISP5230 and the auricin BGC in *S. aureofaciens* CCM 3239, respectively (Wang *et al.*, 2009, Novakova *et al.*, 2005).

Finally, LugRV (645 aa) is a member of the family of *Streptomyces* antibiotic regulatory proteins (SARPs). SARP regulators often act as cluster-specific activators of antibiotic production (van der Heul *et al.*, 2018, Wietzorrek & Bibb,

1997). They are characterised by an OmpR HTH DNA-binding domain (pfam: Trans_reg_C) and a bacterial transcription activator domain (pfam: BTAD domain) (Liu *et al.*, 2013). Similar to the pleotropic antibiotic regulator AfsR from *S. coelicolor* (Horinouchi *et al.*, 1990), LugRV contains a centrally located NB-ARC domain. SARPs are the most well studied regulators.

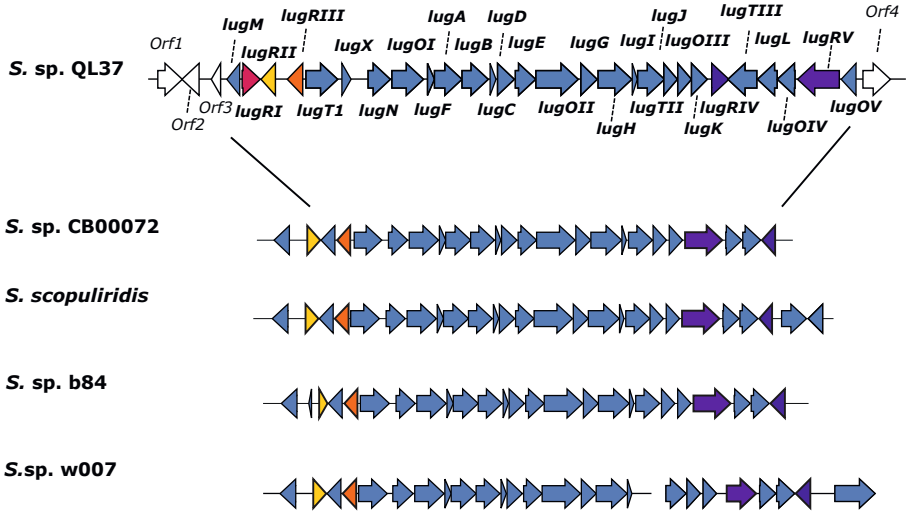


Figure 1 Location of the regulatory genes in the *lug* gene cluster and related type II polyketide gene clusters.
On top the *lug* gene cluster is shown. Below the *lug* gene cluster and the type II polyketide gene clusters are shown from *Streptomyces* sp. CB00072, *S. scopuliridis* RB72, *Streptomyces* sp. b84 and *Streptomyces* sp. w007 (Zhang *et al.*, 2012, Kibret *et al.*, 2018, Cao *et al.*, 2021). All gene clusters contain orthologues of *lugRII* (yellow), *lugRIII* (orange), *lugRIV* (blue) and *lugRV* (purple), but not of *lugRI* (red).

Mutational analysis of the *lugR* genes

The bioinformatic analysis presented above showed that genes highly similar to *lugRII–lugRV* are conserved in BGCs that are highly similar to the *lug* gene cluster. To evaluate the role of these genes in lugdunomycin biosynthesis, we sought to inactivate the genes independently in *Streptomyces* sp. QL37 via homologous recombination (Materials and Methods). To this end, knock-out constructs pΔ*lugRII*, pΔ*lugRIII*, pΔ*lugRIV*, pΔ*lugRV* were created, based on a method published previously (Swiatek *et al.*, 2012). The generated constructs were composed of the conjugatable pWHM3-*oriT* harbouring approximately 1.5 kb of the upstream and downstream regions of the respective *lugR* genes, with an apramycin resistance cassette (*aac*(3)*IV*) in-between (Vara *et al.*, 1989, Garg

& Parry, 2010). The constructs were conjugated to *Streptomyces* sp. QL37 using the methylation-deficient *E. coli* ET12567/pUZ8002 (MacNeil *et al.*, 1992). Through homologous recombination either one of the *lugR* regulatory genes was replaced by the apramycin cassette. The presence of *loxP* sites flanking the apramycin cassette allowed the efficient removal of the resistance cassette by introduction of the pUWLcre construct expressing Cre recombinase (Khodakaramian *et al.*, 2006, Swiatek *et al.*, 2012). The generated in-frame single gene deletion mutants for *lugRII, lugRIII, lugRIV* and *lugRV* were designated QL37Δ*lugRII*, QL37Δ*lugRIII*, QL37Δ*lugRIV*, and QL37Δ*lugRV*, respectively.

To assess the possible role of the *lugR* genes in the control of development and angucycline production, the respective mutants were grown on different media for seven days at 30 °C (Figure 2). Angucyclines are brown or yellowish diffusible pigments, and therefore their production can be seen by visual inspection (Wu *et al.*, 2019). However, lugdunomycin is colourless and its production can only be followed based on its spectral properties (Wu *et al.*, 2019). Wild-type *Streptomyces* sp. QL37 and its *lugR* mutants were grown on minimal media (MM) agar plates supplemented with 0.5% mannitol and 1% glycerol, on R5 agar supplemented with 0.8% peptone and 1% mannitol (R5), or on soya flour mannitol (SFM) agar. Strain QL37Δ*lug-pks* that lacks the minimal polyketide synthase (PKS) genes was included as a control for the absence of angucycline production (Wu *et al.*, 2019) (described in Chapters 3, 4 and 5). These media were selected because limamycins were previously isolated from R5, lugdunomycin from MM and angucyclines from both media. SFM was selected as it is routinely used for promoting sporulation. On MM agar plates, the wild-type *Streptomyces* sp. QL37 produced spores and a yellow pigment, on R5 agar it only produced a sparse aerial mycelium and a brown pigment, while on SFM agar it produced abundant spores and was also brown-pigmented. Small differences were seen between QL37Δ*lug-pks* and the wild-type strain on MM and SFM, but on R5 medium differences were significant; on this media, the wild-type strain hardly developed, while the mutant developed well.

On R5 media, QL37Δ*lugRIII* had a similar phenotype as the wild-type strain. Conversely, QL37Δ*lugRII*, QL37Δ*lugRIV*, and QL37Δ*lugRV* resembled the *lug-pks* mutant, as it was less pigmented and showed better development than the wild-type strain. On MM the phenotypes were similar to those of the wild-type strain, although QL37Δ*lugRII* sporulated better, while sporulation was reduced

in QL37Δ*lugRV*. Mutants QL37Δ*lugRII,* QL37Δ*lugRIV* and QL37Δ*lugRV* hardly produced brown pigments on SFM agar plates. We anticipate that the brown pigment reflects the level of angucycline production, and that lack of sporulation on R5 is due to high production of angucyclines. Following this logic, it is likely that angucyclines were not produced in significant quantities in cultures of the *lugRII, lugRIV* or *lugRV* mutants, while the *lugRIII* mutant still produced abundant angucyclines.

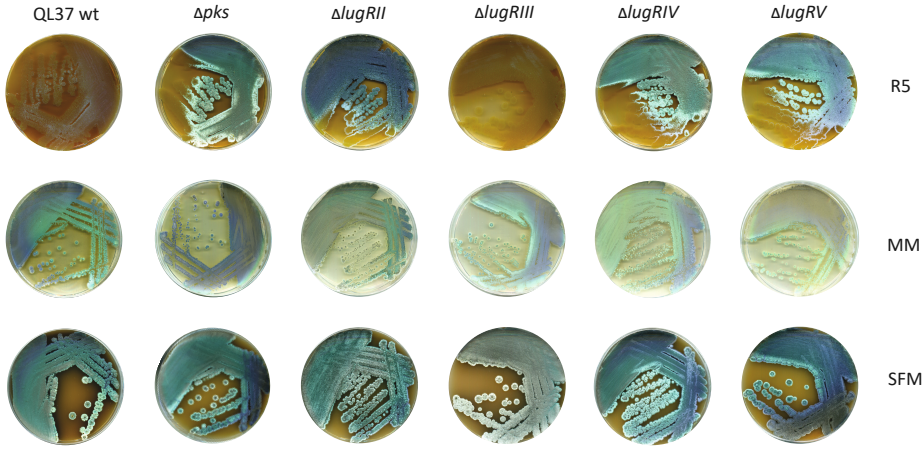


Figure 2 Phenotypes of *lugR* mutants of *Streptomyces* sp. QL37.
The strains were grown on MM, on R5 and on SFM agar plates for seven days at 30 °C. The phenotype of the *lugRII*, *lugRIV* and *lugRV* mutants is similar to the *pks* mutant, implying that these strains do not produce angucyclines, while the *lugRIII* mutant is similar to the wild type implying that this strain still produced angucyclines.

Metabolomic analysis of the *lugR* mutants

To analyse the effect of the deletion of the *lugR* genes on angucycline production in more detail, ethyl acetate extracts were prepared from the wild-type strain, the *lug-pks* mutant and the different *lugR* mutants and analysed using liquid chromatography coupled to mass spectrometry (LC-MS). For this, the strains were grown in triplicate on MM and R5 agar plates (Wu *et al.*, 2019). Natural products were extracted from both the mycelium and the agar. MZmine was used to process the LC-MS data. Principle component analysis (PCA) (Figure 3A and B), and a heatmap with hierarchical clustering analysis were generated based on the processed peak lists (Figure 3C and D). The PCA showed clear separation of the metabolome of the mutants QL37Δ*lugRII*, QL37Δ*lugRIV*, QL37Δ*lugRV* and QL37Δ*pks* from that of the wild-type strain and its *lugRIII* mutant. This is also

reflected in the heatmap. Analysis of the extracted ion chromatograms (Figure 4) and the respective peak areas of the molecules that were previously isolated confirmed that QL37Δ*lugRIII* still produced the typical angucyclines (**2**–**6**), the rearranged angucyclines (**7**–**10**), and lugdunomycin (**1**). On MM the production level of the molecules was lower in QL37Δ*lugRIII* as compared to the levels produced by the wild-type strain. No angucyclines or derivatives thereof were produced by QL37Δ*lugRII*, QL37Δ*lugRIV* and QL37Δ*lugRV*.

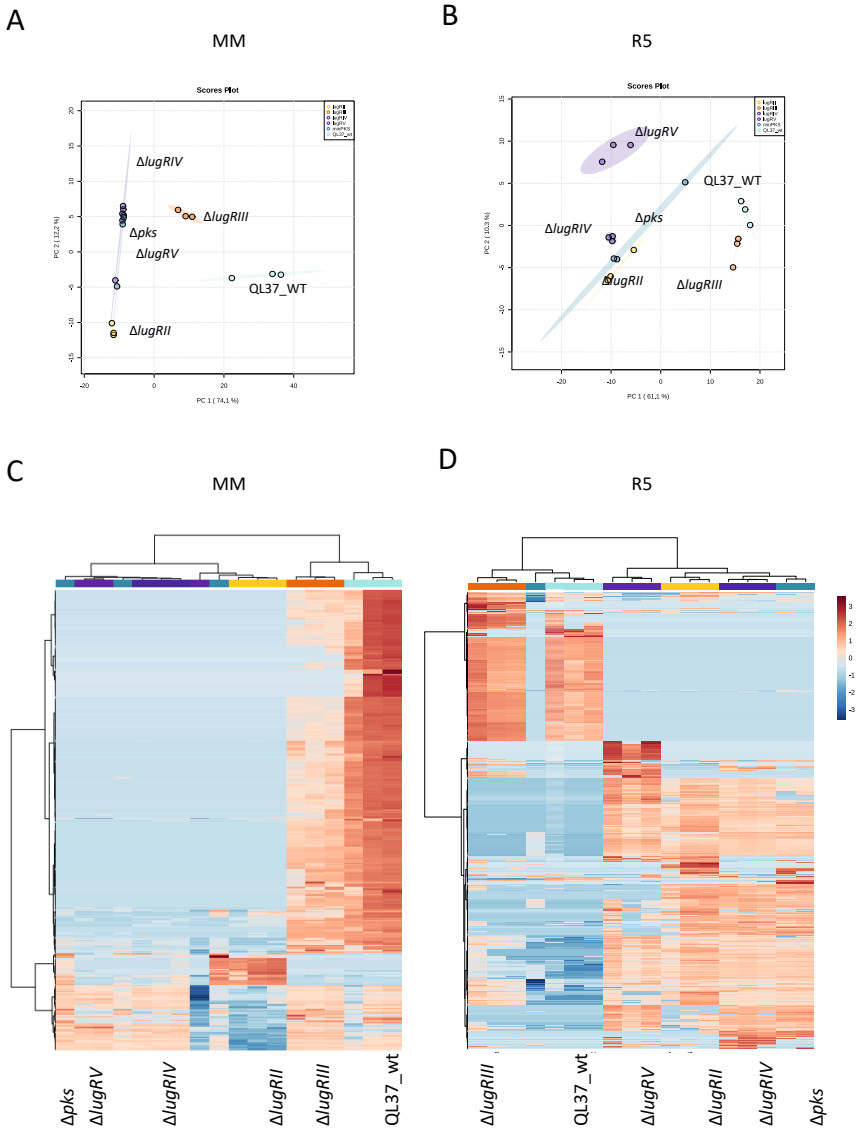


Figure 3 Exploratory statistical analysis of the metabolomic profiles of *Streptomyces* sp. QL37, its *lugRII–RV* and *lug-pks* mutants grown on MM and R5.
PCA score plots of the extracts from the strains grown on MM (A) and R5 (B). Heatmaps with additional hierarchical clustering of the metabolomic profile of the extracts from the analysed strains grown on MM (C) and R5 (D). Mass features that have relatively high peak aeras are displayed in red and mass features that have relatively low peak areas are displayed in blue. Note that mutants lacking *lugRII, lugRIV* or *lugRV* failed to produce angucyclines under these conditions.

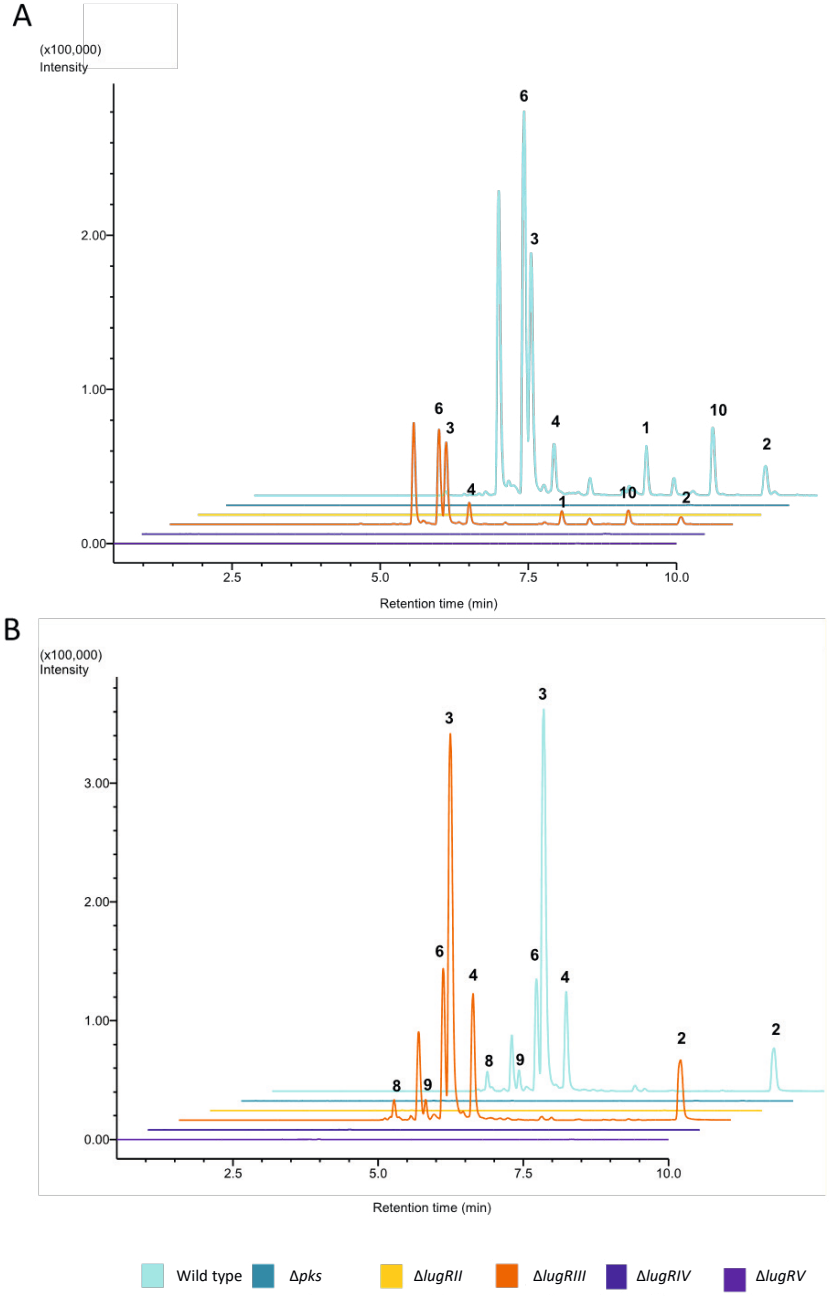


Figure 4 Overlay of extracted ion chromatograms of compounds 1–10 in the extracts from *Streptomyces* sp. QL37, its *lugR* mutants and its *pks* mutant grown on MM (A) or R5 (B). The wild-type strain and its *lugRIII* mutant produced angucyclines, rearranged angucyclines and lugdunomycin, while *lugRII*, *lugRIV* and *lugRV* mutants did not, similar to the *lug-pks* null mutant. See also Figure 3.

Complementation of the *lugRII*, *lugRIV* and *lugRV* mutants and hierarchical order of their regulatory activity

To ascertain that the observed effects of the deletion of *lugRII*, *lugRIV* and *lugRV* were indeed solely due to the deletion of these genes, and to assess whether the regulators are subject to a certain hierarchy in terms of activation of the *lug* gene cluster, the different *lugR* mutants were complemented by introduction of the overexpression constructs for the regulatory genes (Figure 5). For this, the *lugR* genes were each cloned behind the constitutive *ermE* promoter (*ermE**p) in the conjugative vector pSET152, which integrates at the øC31 attachment site in the genome (Motamedi *et al.*, 1995). This resulted in the overexpression constructs pO*lugRII*, pO*lugRIII*, pO*lugRIV*, and pO*lugRV*. Each construct was conjugated to all mutants that were unable to produce angucyclines, namely to QL37Δ*lugRII*, QL37Δ*lugRIV* and QL37Δ*lugRV*. The empty vector pSET152 was used as the control.

Introduction of the *lugRV* construct into the *lugRII*, *lugRIV* and *lugRV* mutants restored the production of the brown diffusible pigments and abolished development on R5 agar, indicative of restoration of angucycline production in all three mutants. However, overexpression of *lugRIV* restored angucycline production in only *lugRII* and *lugRIV* but not in *lugRV* mutants, indicating a hierarchical relationship between *lugRIV* and *lugRV*, with *lugRV* dominant over *lugRIV*. In turn, *lugRII* failed to complement any of the *lugR* mutants, including the *lugRII* mutant itself. Still, the observation that angucycline production was restored to the *lugRII* mutant by introduction of constructs expressing *lugRIV* or *lugRV* from the *ermE**p, strongly suggests that the *lugRII* mutant is not defective in angucycline biosynthesis due to second site mutations and thus that *lugRII* is indeed a positive regulator of the *lug* gene cluster. Why p*lugRII* could not complement the *lugRII* mutant remains to be elucidated. Introduction of *lugRIII* did not affect the production of angucyclines in the mutants *lugRII*, *lugRIV* and *lugRV*, which is in line with the observation that deletion of *lugRIII* only has a minor effect on angucycline production. As for cultures on MM and SFM agar, only introduction of the *lugRV* expression construct had an effect, in line with the minor phenotypic differences seen for the *lugR* mutants on these media.

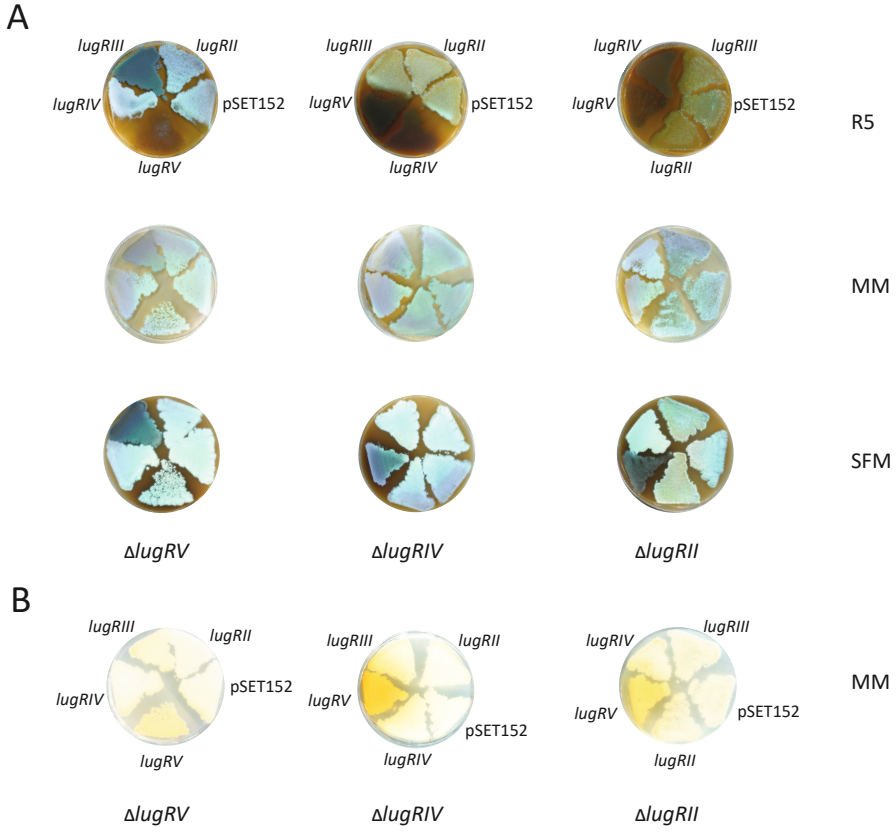


Figure 5 Complementation of *lugR* mutants via expression of the *lugR* genes from the *erm*E*p promoter.
Integrative plasmids expressing either *lugRII*, *lugRIII*, *lugRIV* or *lugRV* from the *erm*E* promoter were introduced into the respective mutants, and restoration of angucycline production assessed. As a control, transformants harbouring the empty vector (pSET152) were included in the experiment.

Effect of the enhanced expression of the *lugR* genes on development and angucycline production

We then wondered what the effect would be of the enhanced expression of the *lugRIV* and *lugRV* genes in the wild-type strain. These genes were chosen for more detailed analysis, considering their major impact on the control of the *lug* gene cluster. Overexpression strains O*lugRIV* and O*lugRV* and the control strain *Streptomyces* sp. QL37::pSET152 were grown on MM, R5 and SFM agar plates (Figure 6). When grown on SFM, the *Streptomyces* sp. QL37::pSET152 produced

aerial hyphae that were largely devoid of spores, indicating that the introduction of pSET152 already affected the morphological differentiation. Reduced sporulation was also observed for ex-conjugants O*lugRIV and* O*lugRV*. On this medium, O*lugRV* stood out from the other strains by producing large amounts of brown pigment, together with a bald phenotype with only sparse aerial mycelium that is most likely due to the DNA-targeting activity of angucyclines (Kharel *et al.*, 2012). On MM agar, the control strain and O*lugRIV* produced spores and yellow pigments, while O*lugRV* showed production of yellow pigment and slightly less sporulation than the control strain. Finally, when the strains were grown on R5 medium all the exconjugants produced abundant brown pigments, associated with reduced development. Strains O*lugRIV* and O*lugRV* produced more brown pigments than the control strain, which indicates that these strains produced more angucyclines.

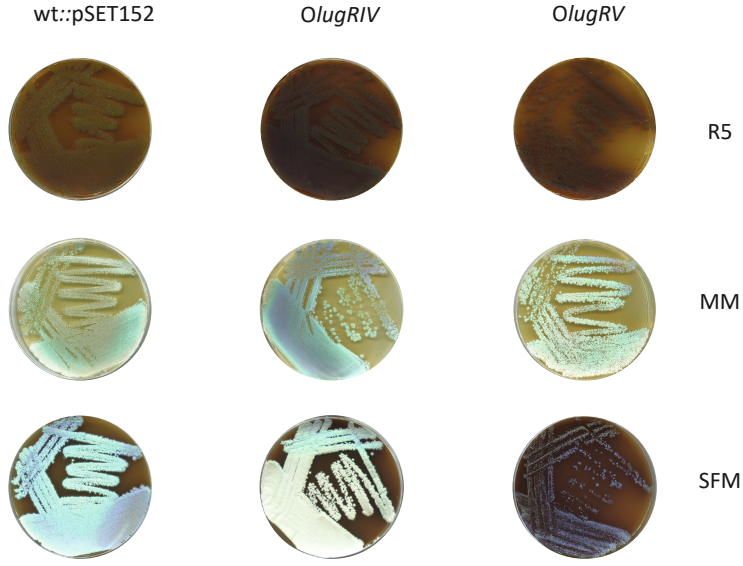


Figure 6 Effects of enhanced expression of the regulatory genes *lugRIV* and *lugRV* on morphogenesis and antibiotic production.
Strains O*lugRIV* and O*lugRV* produced more brown pigments on R5 than the control strain, while O*lugRV* was inhibited in sporulation on SFM after growth for seven days at 30 °C

We then aimed to verify whether the enhanced expression of the regulatory genes from the *lug* gene cluster affected the respective secondary metabolites and subsequently the expression of Lug proteins. To do so, we followed a paired omics strategy by combining metabolomics with quantitative proteomics (Schorn *et al.*, 2021). We analysed cultures of the overexpression strains O*lugRIV* and O*lugRV*,

which overexpress *lugRIV* and *lugRV*, respectively, from the strong *ermE** promoter. The plasmid pSET152 without insert was used as the control. To perform the combined metabolomics/proteomics experiment, all strains were grown in triplicate on MM and R5 covered with cellophane for two days. Mycelia and spores were harvested to prepare proteins extracts for quantitative proteomics, while the secondary metabolites were extracted from the remaining solid media, for analysis via LC-MS. In addition, the strains were grown for seven days on MM, the time point at which lugdunomycin is produced, and secondary metabolites were subsequently extracted from mycelium and agar.

In line with the observed changes in pigmentation and development, PCA and heatmap hierarchical clustering analysis revealed major changes in the metabolome of MM-grown O*lugRIV* and O*lugRV* (Figure 7A and C), while on R5 agar the differences were less significant (Figure 7B and D). The effect of the overexpression of the regulatory genes on the different structural classes produced by QL37 was then investigated in more detail using molecular networking. Using Feature-Based molecular networking (FBMN) in the Global Natural Product Social Networking (GNPS) platform, molecular networks of structurally related metabolites were generated in which isomers having different retention times can be identified, and the relative quantification of the mass features could be mapped (Nothias *et al.*, 2020). Ion Identity Molecular networking was then used to connect ion adducts of the same mass feature based on several criteria (Nothias *et al.*, 2020, Schmid, 2020). In this way, two molecular networks were generated. The ions detected in the media blanks were removed before generation of the networks (Figure 8 and Figure S2).

Analysis of the molecular network from MM-grown cultures obtained after 2 days of growth revealed different molecular families for the non-rearranged angucyclines, the limamycins pratensilin A and lugdunomycin (Figure 8). Elmonin (**10**) was identified as a single node in the network. In addition, a molecular family of γ-butyrolactones was detected (Chapter 4) (Yang *et al.*, 2005). Both the rearranged and non-rearranged angucycline molecular families were predominantly present in O*lugRIV* and O*lugRV*, while lugdunomycin was mainly detected in O*lugRV*. The vast majority of the mass features in the angucycline-related molecular families were detected only in O*lugRIV* and O*lugRV* and not or hardly in the parental strain; this could be due to the early extraction time point (two days), when angucyclines

are not yet or hardly produced in wild-type cells, while production is accelerated in the overexpression strains. The control strain produced more angucyclines and limamycins on R5 agar, albeit with lower concentrations as compared to the overexpression strains, particularly in comparison with O*lugRIV* (Figure S2).

When the metabolites were extracted from a seven days culture on MM (the time point which was previously used for the isolation of lugdunomycin (Wu *et al.*, 2019)), the angucycline-related metabolites could be detected in cultures of the control strain and of the overexpression strains (Figure S3). Again, concentrations were higher in the overexpression strains than in the control, although the difference was somewhat less pronounced as compared to the early time points. Interestingly, lugdunomycin production was not increased at this time point and was even slightly decreased compared to the control strain. This suggests uncoupling of lugdunomycin and angucycline biosynthesis, most likely because BGC 23 is required for the biosynthesis of lugdunomycin, but not for other angucyclines and derivatives.

Analysis of the proteome of the strains overexpressing the *lugR* genes

Quantitative proteomics was executed on the proteins isolated from the biomass of the O*lugRIV*, O*lugRV* and pSET152 control strains grown for two days on MM and R5 at 30 °C. First PCA analysis was executed to assess the reproducibility of the replicates and visualise the similarity of the proteome among the three recombinant strains (Figure S4). PCA analysis showed that the replicates of the proteome of the control and O*lugRIV* were separated in both MM and R5 media, with a relatively small variation within each set, while variation in O*lugRV* samples was relatively large in both cultures. Still, these replicates were also clearly separated from the control strain and to a lesser extent from O*lugRIV* in MM samples, which is in line with the previously observed difference in the metabolome of these strains. Bioinformatics analysis showed that the *lug* gene cluster most likely contains 28 genes (Chapter 3). Of the predicted gene products, 17 were detected in both the MM and R5-grown cultures by proteomics. To visualise the differential expression of the Lug proteins in each of the overexpression strains relative to the control, heatmaps and violin plots were generated (Figure 9A and B; Figure S5A and B). Violin plots were generated to show the distribution of the overall expression of the Lug proteins. When grown on MM, the Lug proteins were upregulated in O*lugRIV* and O*lugRV* compared to the control strain (Figure 9A and B). This correlated

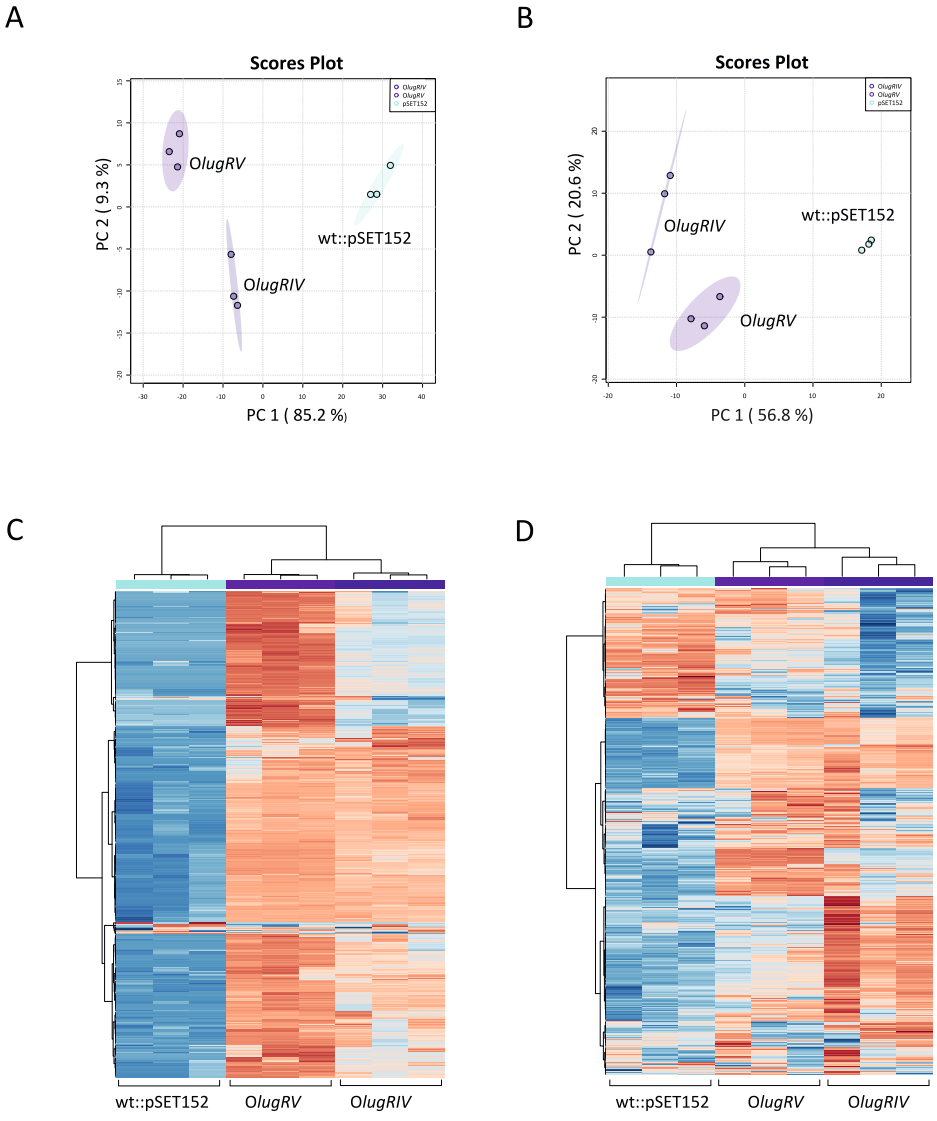


Figure 7 Statistical analysis of the metabolomic profiles of *Streptomyces* sp. QL37 over-expressing *lugRIV* and *lugRV*.
PCA score plots of the extracts from the analysed strains grown on MM (A) and R5 (B) for 48 hours. Heatmaps with additional hierarchical clustering of the metabolomic profile of the extracts from the analysed strains grown on MM (C) and R5 (D). Mass features that have relatively high peak areas are displayed in red and mass features that have relatively low peak areas are displayed in blue.

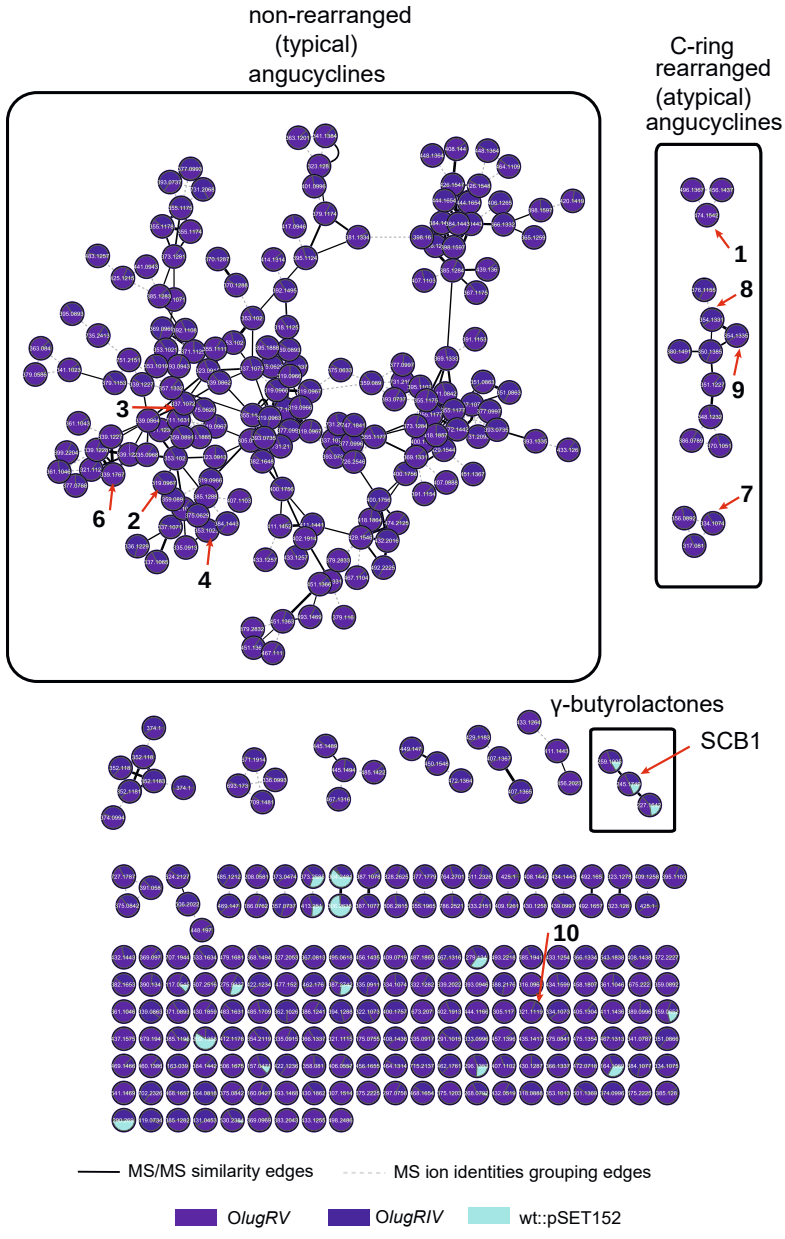


Figure 8 Molecular network of the ions detected in the extracts of *Streptomyces* sp. QL37 O*lugRIV*, O*lugRV* and the control strain grown on MM.
The nodes are labelled by the precursor mass of their ions and pie charts are mapped to the nodes to indicate the relative intensities of the ions in the different samples. Arrows are used to highlight the metabolites which were previously identified in *Streptomyces* sp. QL37. The molecular families, which were assigned based on the previously known metabolites or based on analysing their MS/MS spectra, are highlighted.

well with the enhanced or precocious production of the angucycline-related metabolites. The change in the Lug proteins was less significant on R5 medium, again in line with the metabolomics data (Figure S4). As lugdunomycin is derived from two BGCs, namely the angucycline BGC (*lug*; BGC 12) and the *iso*-maleimycin BGC (BGC 23) (Chapter 4), the proteins encoded by the *iso*-maleimycin BGC were also investigated. In samples from MM-grown cultures ten proteins encoded by the *iso*-maleimycin BGC were detected (Figure S6), whereas in samples from R5-grown cultures only three proteins were detected (Figure S5). Although the expression of the *iso*-maleimycin BGC appeared increased in the heatmaps in O*lugRV* as compared to the control strain, this increase was not statistically significant (*p*-value = 0.21733) according to the Mann-Whitney test (Figure 9 and Figure S6). This is in accordance with the observed failure to detect a significant increase in lugdunomycin production over time, despite the constitutive expression of the activators of the *lug* gene cluster (Figure S3).

DISCUSSION

The aim of this study was to obtain insight into the regulatory network of the *lug* gene cluster, as a prelude to improve the production of lugdunomycin. Based on bioinformatic predictions, the *lug* gene cluster encodes five regulatory proteins belonging to five different families, including an XRE-type regulator (LugRI), an atypical response regulator (ARR) of the LuxR-type (LugRII), a TetR-like regulator (LugRIII), an ARR of the OmpR-type (LugRIV) and a SARP regulator (LugRV). In this study the role of *lugRII–V* in angucyclines and lugdunomycin production was examined. These regulatory genes are conserved in *lug*-type gene clusters highly similar to *lug* (Chapter 4).

Targeted deletion combined with phenotypic and metabolomics analysis, showed that LugRII, LugRIV and LugRV are positive regulators of the *lug* gene cluster, while LugRIII played only a minor role under the conditions tested. The *lugRIV* and *lugRV* mutants could be complemented with the introduction of constructs expressing the respective genes from the *ermE* promoter, which showed that there were no polar effects caused by the generated mutations. However, the *lugRII* mutants could not be complemented by the introduction of *lugRII*; instead, angucycline production could be restored by expressing either *lugRIV* and *lugRV*. This suggests that the angucycline biosynthetic genes were not affected in the

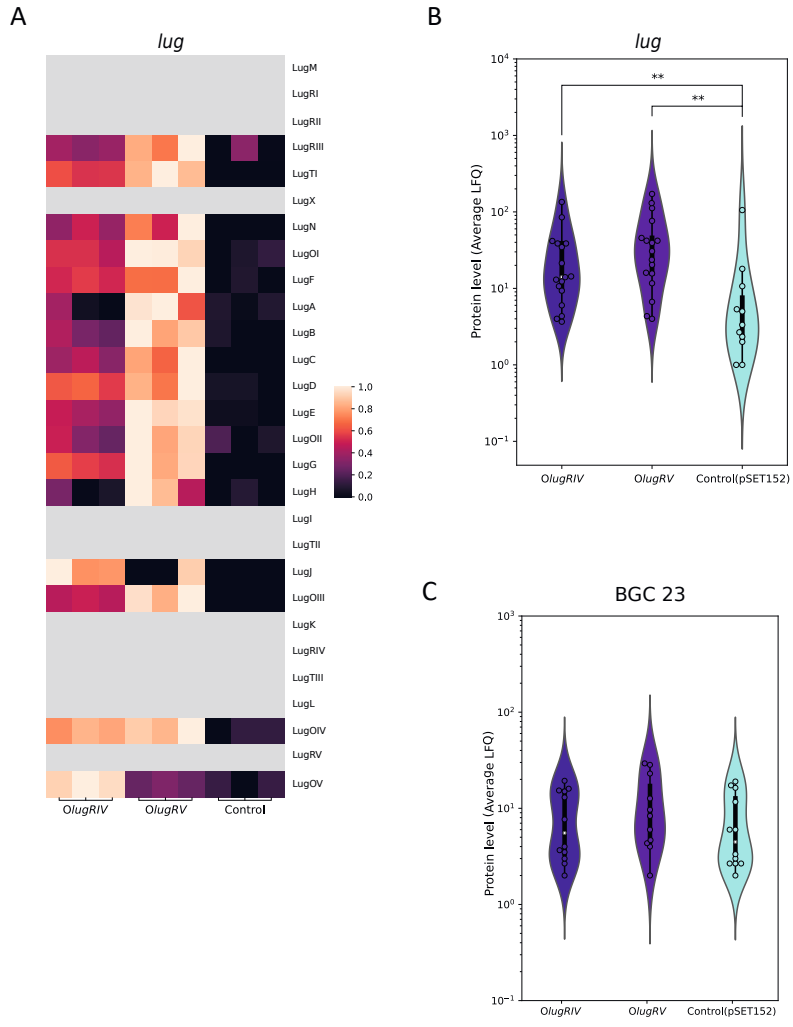


Figure 9 Statistical analysis of the proteins encoded by the gene clusters involved in lugdunomycin biosynthesis in O*lugRIV*, O*lugRV* and the control strain (QL37::pSET152) grown on MM for two days.

A) Heatmap showing the levels of the proteins encoded by the lug gene cluster in each replicate of O*lugRIV*, O*lugRV* and the control strain, normalised by row to the length of 1. The grey tiles indicate that the protein was not detected. B) The distribution of the overall expression level of the Lug proteins represented in a violin plot. The dots indicate the average expression level of each protein. ** indicates a p-value less than 0.01 (O*lugRIV* vs control: *p*-value =0.00901, O*lugRV* vs control: *p*-value =0.00190) C) The distribution of the overall expression level of the proteins from BGC 23. The expression of BGC 23 was not significantly different among the tested strains (OlugRIV vs control: *p*-value = 0.60230; O*lugRV* vs control: *p*-value = 0.21733). The p-values were calculated using the Mann-Whitney test.

lugRII mutant, and that perhaps insufficient expression of *lugRII* was achieved from plasmid O*lugRII*. The function of *lugRII* and its role in the control for angucycline biosynthesis requires further studies.

Cross-complementation experiments suggested a hierarchical order between the positive regulators LugRIV and LugRV, whereby the ARR LugRIV activates the SARP LugRV which in turn activates the structural genes of the *lug* gene cluster, resulting in the biosynthesis of the angucyclines. A similar hierarchical system wherein an ARR activates a SARP regulator that in turn activates the BGC, was observed for the undecylprodigiosin and daunorubicin BGCs in the angucycline non-producers *S. coelicolor* and *S. peucetius*, respectively (Hulst *et al.*, 2021, Fujii *et al.*, 1996, Guthrie *et al.*, 1998, Furuya & Hutchinson, 1996). The ARR-family regulators RedZ and DnrN belong to the NarL/FixJ superfamily of LuxR-type regulators, and show more similarity to LugRII, whereas LugRV belongs to the ARRs from the OmpR-type, such as JadR1, Aur1P and LanI (Wang *et al.*, 2009, Novakova *et al.*, 2005, Yushchuk *et al.*, 2019). The latter directly activate the expression of the structural genes of the angucycline BGCs governing the production of jadomycin, auricin and landomycin in *S. venezuelae*, *S. aureofaciens* and *S. cyanogenus*, respectively, without the involvement of a SARP regulator (Novakova *et al.*, 2005, Wang *et al.*, 2009, Yushchuk *et al.*, 2019). This highlights interesting regulatory differences between angucycline BGCs.

Angucyclines are toxic molecules and a mechanism to inhibit their production was expected. Various angucycline BGCs (kinamycin, jadomycin, auricin and landomycin) contain genes coding for activators and for repressors (Aigle *et al.*, 2005, Kormanec *et al.*, 2014, Zou *et al.*, 2014, Yushchuk *et al.*, 2019), whereas our experiments on the *lugR* genes only identified activatory roles. Similar results were observed in the study of the thioangucycline BGC of *Streptomyces* sp. CB00072 (Cao *et al.*, 2021). Repression of angucycline production may be executed by an autoregulatory mechanism involving the ARRs LugRII and LugRIV, since the binding of ARRs and thereby the expression of the target BGC is modulated by the binding of the intermediates and final products of the BGCs (Wang *et al.*, 2009). The function of LugRI in angucycline biosynthesis, if any, was not elucidated and needs more attention. Distant orthologues of *lugRI* (≤34% between the gene products) were found in the type II polyketide BGC of *Streptomyces* sp. CB02414, the producer of rubiginones, which belong to the angucyclines (Zhang *et al.*, 2021).

However, the role of these regulators in the control of the production of rubiginones has not been elucidated.

Overexpression of LugRIV and LugRV resulted in precocious angucycline and lugdunomycin production in the early stages of growth. A simultaneous increase in the expression of the Lug proteins in the O*lugRIV* and O*lugRV* strains relative to the control was also observed. However, prologues growth of the strains (seven days instead of two) did not result in higher levels of lugdunomycin in the overexpression mutants relative to the control strain, even though the angucyclines were still produced at a higher level. Lugdunomycin biosynthesis requires two substrates: a C-ring rearranged angucycline and *iso*-maleimycin (Wu *et al.*, 2019), which are probably produced from two different BGCs (Chapter 4). Investigation of the proteomics data revealed no statistically significant changes in the expression of the proteins encoded by the hypothesised *iso*-maleimycin BGC (BGC 23) between the control strains and either O*lugRIV* or O*lugRV*, which could explain the failure of O*lugRIV* and O*lugRV* to overproduce lugdunomycin. In MM-grown cultures more proteins encoded by the *iso*-maleimycin BGC were detected compared to R5-grown cultures, which may explain the absence of lugdunomycin in R5-grown cultures. Based on that, constitutive expression of the *iso*-maleimycin gene cluster can possibly help to increase lugdunomycin production.

In conclusion, this study revealed that the *lug* gene cluster that is required for angucycline, limamycin and lugdunomycin production is positively controlled by the three regulators LugRII, LugRIV and LugRV. A hierarchical relationship could be identified between LugRIV and LugRV, where LugRV is the ultimate cluster-situated regulator. This contrasts with the regulatory networks controlling other known angucycline BGCs, several of which lack a SARP regulator and instead fully depend on ARR-family regulators. Overexpression of these regulators led to increased angucyclines production, but not of lugdunomycin. This again underlines the complexity of the metabolic and regulatory pathways that together govern the biosynthesis of this cryptic molecule.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Bacterial strains used in this study are indicated in Table 1. *Streptomyces* sp. QL37 was obtained from the soil in the Qinling mountains (P. R. China) (Zhu *et al.*, 2014b). The strain is deposited to the collection of the Centraal Bureau voor Schimmelcultures (CBS) in Utrecht, The Netherlands, under deposit number 138593. The metabolite production profile of each *Streptomyces* strain used in this study is indicated in Table 1 and Table 2. *Escherichia coli* JM109 was used for general cloning and grown on Luria Broth without antibiotics. *E. coli* ET12567/pUZ8002 was used for conjugation of plasmids to *Streptomyces* sp. QL37 (Wang & Jin, 2014, MacNeil *et al.*, 1992, MacNeil, 1988). Strains containing plasmids were selected on LB containing ampicillin (100 µg/mL), apramycin (50 µg/mL), chloramphenicol (25 µg/mL) and kanamycin (50 µg/mL). Spore stocks were prepared according to (Kieser *et al.*, 2000) from cultures of *Streptomyces* sp. QL37 grown on SFM agar for seven days at 30°C. For conjugation *Streptomyces* sp. QL37 was grown on SFM containing 60 mM $MgCl_2$ and 60 mM $CaCl_2$ (Wang & Jin, 2014, Kieser *et al.*, 2000). Exconjugants were selected using thiostrepton (20 µg/mL) and apramycin (50 µg/mL).

Construction of knock-out mutants

The in-frame deletion mutants were created according to the method previously described (Swiatek *et al.*, 2012) and described in Chapter 5. The correct mutants were selected on SFM containing either apramycin (50 µg/mL) or thiostrepton (10 µg/mL), where they should only grow in the presence of apramycin. In order to remove the apramycin cassette by recombination of the *lox*P sites the pUWL-Cre construct was conjugated to the strains. The removal of the apramycin cassette was verified by PCR with the primers given in Table 2 and Sanger sequencing of the PCR products (Swiatek *et al.*, 2012).

Generation of strains for enhanced gene expression of the *lugR* genes

For constitutive expression of the *lugR* genes, the genes were amplified from the genome using the primers indicated in Table 3 and cloned downstream of the *ermE** promoter. For this, the genes *lugRIII* and *lugRV* were first sub-cloned in the pHM10a vector using the restriction enzymes NdeI and HindIIII. The constructs were then digested with BglII and EcoRI, resulting in a fragment that contained the

*ermE** promoter, the gene, and a part of the pHM10a plasmid. This fragment was subsequently cloned into pSET152 digested with BamHI and EcoRI. A three-way ligation including the *ermE** promoter, the plasmid pSET152, and the gene resulted in constructs with *lugRII* or *lugRIV* behind the *ermE** promoter in pSET152. The integrity of the constructs was verified using Sanger sequencing and restrictions enzyme analysis. The constructs were isolated from JM109 and transformed in the methylase deficient strain *E. coli* ET12567/pUZ8002 for conjugation to *Streptomyces* sp. QL37 (MacNeil *et al.*, 1992).

Table 1 Bacterial strains used in this work

Strain	Genotype	Reference
***Streptomyces* sp. QL37**	Wild type lugdunomycin producer	(Zhu *et al.*, 2014b, Wu *et al.*, 2019, Wu, 2016)
***E. coli* JM109**	*end*A1, *rec*A1, *gyr*A96, *thi, hsd*R17 (r_k^-, m_k^+), *rel*A1, *sup*E44, λ–, Δ(*lac-pro*AB), [F´, *tra*D36, *pro*AB, *lac*I^qZΔM15], IDE3.	(Sambrook J., 1989)
***E. coli* ET12567::pUZ8002**	Methylation-deficient host (*dam*$^-$ *dcm*$^-$*hsdM*$^-$) for intergeneric conjugation. Contains the nontransmissible, *oriT*$^-$ mobilizing plasmid pUZ8002.	(MacNeil *et al.*, 1992)
QL37Δ*lug-pks*	*Streptomyces* sp. QL37 Δ*pks*::apra^R	(Wu *et al.*, 2019)
QL37Δ*lugRII*	*Streptomyces* sp. QL37 Δ*lugRII*	This work
QL37Δ*lugRIII*	*Streptomyces* sp. QL37 Δ*lugRIII*	This work
QL37Δ*lugRIV*	*Streptomyces* sp. QL37Δ*lugRIV*	This work
QL37Δ*lugRV*	*Streptomyces* sp. QL37Δ*lugRV*	This work
WT::pSET152	*Streptomyces* sp. QL37 carrying pSET152	
OE-*lugRII*	*Streptomyces* sp. QL37::pSET152-*ermE* -lugRII*	This work
OE-*lugRIII*	*Streptomyces* sp. QL37::pSET152-*ermE* -lugRIII*	This work
OE-*lugRIV*	*Streptomyces* sp. QL37::pSET152-*ermE* -lugRIV*	This work
OE-*lugRV*	*Streptomyces* sp. QL37::pSET152-*ermE*-lugRV*	This work

Table 2 Oligonucleotides used in this study for the creation of knock-out constructs

Gene	Application	Sequence (5' -> 3') ^	Position #	Gene length (nt)
lugRII	Left flank	GTCAGAATTCCTTCAGCGATGCCGGATAAC	+19/-694	669
		GAAGTTATCCATCACCTCTAGACAGGGATGCTCTACGAGATG		
	Right flank	GAAGTTATCGCGCATCTCTAGAACCGTTGTCGACCTGCATGAATC		
		GTCAAAGCTTGGCGGTAGGATGAGTTGGTGACTG		
	Mutant verification	TGACCGAGGTTGCCGAAACAGG		
		CGACTTCACGGGCCGCCTGAAC		
lugRIII	Left flank	GTCAGAATTCGCGTAGGTGGCGAGTCCGACAAG	+63/-794	717
		GTCATCTAGAGTTGCACAACGCCGAGCGAGG		
	Right flank	GTCATCTAGAGTCCGTCGAGACGGCACTGAC		
		GTCAAAGCTTCGGAGTGCAGGTGCCGTTTCAC		
	Mutant verification	AGCAGGACCACACCGATCAG		
		TGACCGGCTCGCCAAGATTTC		

Table 2 Oligonucleotides used in this study for the creation of knock-out constructs *(continued)*

Gene	Application	Sequence (5' -> 3') ^	Position #	Gene length (nt)
lugRIV	Left flank	GTCAGAATTCGTCCTCAACTACGCGGAGTG	+24/-698	801
		GTCATCTAGAATGGGTATGGGCAGGCCAATC		
	Right flank	GTCATCTAGATAAGCTCGGTTCGAGCCAGTG		
		GTCAAAGCTTGCTACGACAAGCGCAAGCTG		
	Mutant verification	GCCGTCTACAAGACCTGGTTCC		
		CCGATCCGTACGACGGACTGC		
lugRV	Left flank	GTCAGAATTCACCTTGCCGCCCAGTGAATG	-18/-1917	1939
		GAAGTTATCCATCACCTCTAGACCGTGCGGTTCATTCCAGTAG		
	Right flank	GAAGTTATCGCGCATCTCTAGAGAGAACGGCGAAGTGCATGAG		
		GTCAAAGCTTGGACGATCTCGTCGCTGATG		
	Mutant verification	ACTTCCATGTCAGCCTGGACG		
		GATCTCCTCGACGACCCGGTC		

position relative to the translational start site (+1) of the respective genes.
^ restriction sites underlined. GAATTC, EcoRI; AAGCTT, HindIII; TCTAGA, XbaI;

Table 3 Oligonucleotides used in this study for the creation of overexpression constructs

Gene	Sequence (5' -> 3') ^	Position #	Gene length (nt)
lugRII	CGATAAGCTTTCCCTGGCCACGCCTTACGC	+1/-700	669
	GCGCCCATCCATATGGTTTTCGCAG		
lugRIII	CGATAAGCTTTCGTCCGGGTGGCGGTCATC	+1/-729	717
	CGATGAATTCATATGACTGAAGCCAGAACGTCCGAGG		
lugRIV	CGATTCTAGAGAGGTCAGGGACAGCGCATGG	+1/-805	801
	CGATGGATCCATATGTCGCACGACCTCGATCTG		
lugRV	CGATAAGCTTGGATGAGCCGGCATGAGATGAACCG	+1/-2009	1938
	CGATGAATTCATATGCACTTCGCCGTTCTCGGCCC		

^ restriction sites underlined. GGATCC, BamHI; GAATTC, EcoRI; AAGCTT, HindIII; CATATG, NdeI; TCTAGA, XbaI.
# position relative to the translational start site (+1) of the respective genes.

Table 4 Plasmids and constructs used in this study

Plasmid	Description	Reference
pWHM3-oriT.	*E. coli*/*Streptomyces* shuttle vector, high-copy number in *E. coli.* For conjugative transfer *oriT* was inserted into unique NdeI site. Amp^R, $Thio^R$.	(Wu *et al.*, 2019, Vara *et al.*, 1989)
pSET152	*E. coli*/*Streptomyces* shuttle vector carrying *attP* site and integrase gene of ϕC31 phage for stable integration into chromosomal *attB* site of *Streptomyces*, $Apra^R$	(Bierman *et al.*, 1992)
pHM10a	*E. coli*/*Streptomyces* shuttle vector, harbouring the *ermE** promoter and engineered ribosome binding site (RBS), Amp^R, Hyg^R	(Motamedi *et al.*, 1995)
pUWL-Cre	Cre-recombinase expression plasmid, $Thio^R$	(Khodakaramian *et al.*, 2006)
pOE-*lugRII*	pSET152 with coding region of *lugRII* under the control of the *ermE** promoter	This study
pOE-*lugRIII*	pSET152 with coding region of *lugRIII* downstream *ermE** promoter	This study
pOE-*lugRIV*	pSET152 with coding region of *lugRIV* downstream *ermE** promoter	This study
pOE-*lugRV*	pSET152 with coding region of *lugRV* downstream *ermE**promoter	This study
pΔ*lugRII*	pWHM3 containing flanking regions of *lugRII* interspaced with the *apra*^{R}-loxP cassette	This study

Table 4 Plasmids and constructs used in this study *(continued)*

Plasmid	Description	Reference
pΔ*lugRIII*	pWHM3 containing flanking regions of *lugRIII* interspaced with the *apra*^{R}-loxP cassette	This study
pΔ*lugRIV*	pWHM3 containing flanking regions of *lugRIV* interspaced with the *apra*^{R}-loxP cassette	This study
pΔ*lugRV*	pWHM3 containing flanking regions of *lugRV* interspaced with the *apra*^{R}-loxP cassette	This study

Natural products extraction

Spores of the mutants and the wild type strains were confluently grown on minimal medium (MM) agar plates (25 mL) containing 0.5% mannitol and 1% glycerol and on R5 Difco containing 1% mannitol and 0.8% peptone for seven days at 30 °C (Wu *et al.*, 2019). Agar added to MM was Iberian agar derived from TM Duche & Sons Ltd (batch from 2017). The same agars as described in Chapter 3 were used for the preparation of the media. Spores of the strains for enhanced expression of the *lugR* genes and the control strain (QL37::pSET152) were grown for two days on agar plates (MM and R5) covered with cellophane. Natural products were extracted from the agar and the cellophane. Before extraction, the mycelium was removed from the cellophane and used for proteomics. In addition, the spores for enhanced expression of the *lugR* genes were grown for seven days on MM agar plates. Natural products were extracted from both the spores and the agar. Natural products extractions were further executed as described in Chapter 5, with only one exception; evaporation of organic solvents was executed at room temperature.

LC-MS methods

Mutants of the regulatory genes, the wild wild-type strain and strains for enhanced expression (OlugRIV and OlugRV and control strain) grown on MM for seven days.

LC-MS/MS acquisition was performed using Shimadzu Nexera X2 UHPLC system, with attached PDA, coupled to Shimadzu 9030 QTOF mass spectrometer, equipped with a standard ESI source unit, in which a calibrant delivery system (CDS) is installed. The extracts were injected into a Waters Acquity HSS C_{18} column (1.8 µm, 100 Å, 2.1 × 100 mm). The column was maintained at 30 °C, and run at a flow rate of 0.5 mL/min, using 0.1% formic acid in H_2O as solvent A, and 0.1% formic

acid in acetonitrile as solvent B. A gradient was employed for chromatographic separation starting at 5% B for 1 min, then 5 – 85% B for 9 min, 85 – 100% B for 1 min, and finally held at 100% B for 4 min. The column was re-equilibrated to 5% B for 3 min before the next run was started.

All the samples were analysed in positive polarity, using data dependent acquisition mode. In this regard, full scan MS spectra (m/z 100 – 2000, scan rate 20 Hz) were followed by three data dependent MS/MS spectra (m/z 100 – 2000, scan rate 20 Hz) for the three most intense ions per scan. The ions were selected when they reach an intensity threshold of 1000, isolated at the tuning file Q1 resolution, fragmented using collision induced dissociation (CID) with collision energy ramp (CE 20–50 eV), and excluded for 0.05 s (one MS scan) before being re-selected for fragmentation. The parameters used for the ESI source were: interface voltage 4 kV, interface temperature 300 °C, nebulizing gas flow 3 L/min, and drying gas flow 10 L/min.

LC-MS analysis for the strains with enhanced expression of the regulatory genes (metabolomics combined with proteomics)

The natural product extracts for combined metabolomics and proteomics experiment were analysed with a similar method as described above. The LC-method was identical to the method described for the analysis of the extracts of the mutants, only a few changes were made in the MS acquisition method.

All the samples were analysed in positive polarity, using data dependent acquisition mode. In this regard, full scan MS spectra (*m/z* 100–1700, scan rate 10 Hz, ID enabled) were followed by two data dependent MS/MS spectra (*m/z* 100–1700, scan rate 10 Hz, ID disabled) for the two most intense ions per scan. The ions were selected when they reach an intensity threshold of 1500, isolated at the tuning file Q1 resolution, fragmented using collision induced dissociation (CID) with fixed collision energy (CE 20 eV), and excluded for 1 s before being re-selected for fragmentation. The parameters used for the ESI source were: interface voltage 4 kV, interface temperature 300 °C, nebulizing gas flow 3 L/min, and drying gas flow 10 L/min.

Comparative Metabolomics

Before statistical analysis the data obtained from the LC-MS-runs were processed using MZmine version 2.53 (Pluskal *et al.*, 2010). The data derived from MM-grown cultures and R5- grown cultures were processed separately. Data processing in MZmine for statistical analysis in MetaboAnalyst was executed as in Chapter 4 (Chong *et al.*, 2019)(See Pre-processing LC-MS data in MZmine 2.53) with a few exceptions. With the detection of the mass ion peaks only mass ion peaks at an MS level 1 were detected. Thus, in the chromatogram deconvolution step no additional MS2 scan pairing was performed. Furthermore, within the modules isotopic peak grouper, join aligner and gap-filling an RT tolerance of 0.1 min was applied, instead of 0.05 min for cultures grown on MM. The final step was removing mass features detected before 1 min and after 10 min with peak list row filter (RT 1.0–10.0). Peaks with and without MS/MS data were kept. The data were exported to MetaboAnalyst file and were filtered as described in Chapter 4.

Statistical analysis was performed using MetaboAnalyst V4.0 (Chong *et al.*, 2019) as described in Chapter 5.

GNPS molecular networking

Pre-processing LC-MS data in MZmine 2.53 and MZmine 2.37.1 corr17.7 and data filtering

For molecular networking using GNPS, the Feature Based Molecular Networking (FBMN) platform was used in combination with ion identity networking (IIMN), where the data were first processed using MZmine version 2.53 and subsequently further processed in MZmine version 2.37.1.corr17.7 (Phelan, 2020, Xie *et al.*, 2020). Data processing in MZmine for statistical analysis in MetaboAnalyst was executed as in Chapter 4, with a few exceptions: within the modules isotopic peak grouper, join aligner and gap-filling an RT tolerance of 0.1 min was applied, instead of 0.05 min for cultures grown on MM. Data filtering executed as in Chapter 4

Feature-based ion identity molecular networks

The data were submitted to the GNPS web tool and a network was generated using FBMN. The same filtration method was used as described in Chapter 4, with one exception: For R5-grown cultures the edges were filtered to have a cosine score

above 0.8. The molecular network was visualised using the software Cytoscape (Kohl *et al.*, 2011).

Proteomics

Spores were confluently grown on MM and R5 agar plates covered with cellophane. After two days of growth at 30 °C the mycelium was scraped off the plate and subsequently transferred to eppendorfs containing metal beads. These were subsequently frozen in liquid nitrogen. The cells were disrupted using a pre-cooled TissueLyser (Qiagen, the Netherlands) for two cycles of 30 s at 30 Hz. The disrupted cells were subsequently treated with a lysis buffer (4% SDS, 0.06 M DTT, 100 mM Tris-HCl pH 7.6, 50 mM EDTA) in order to dissolve the proteins. The proteins were precipitated using chloroform-methanol (Wessel & Flugge, 1984). The precipitated proteins were dissolved in 0.1% RapiGest SF surfactant (Waters, USA) at 95 °C. Proteins were digested with trypsin according to Rooden *et.al.* (van Rooden *et al.*, 2018). Protein digestion was followed by addition of trifluoroacetic acid for the complete removal and degradation of RapiGest SF. Using the STAGE-Tipping technique peptide solutions containing 8 µg of peptide were cleaned and desalted (Rappsilber *et al.*, 2007, Gubbens *et al.*, 2014). Protein concentrations were determined with the BCA (bicinchoninic acid assay) assay (Walker, 1994). Final peptide concentration was adjusted to 100 ng/µL using sample solution (3% acetonitrile, 0.5% formic acid) before analysis.

UPLC and mass spectrometry measurement and data processing for proteomics

200 ng (2 µL) of the digested peptide was injected and analysed by reversed-phase liquid chromatography on a nanoAcquity UPLC system (Waters, Massachusetts, U.S.) equipped with HSS-T3 C_{18} 1.8 µm, 75 µm X 250 mm column (Waters, Massachusetts, U.S.). A gradient from 1% to 40% acetonitrile in 110 min (ending with a brief regeneration step to 90% for 3 min) was applied. [Glu^1]-fibrinopeptide B was used as lock mass compound and sampled every 30 s. Online MS/MS analysis was done using Synapt G2-Si HDMS mass spectrometer (Waters, Massachusetts, U.S.) with an $UDMS^E$ method set up as described in (Distler *et al.*, 2014).

Raw data from all samples were first analysed using the vendor software ProteinLynx Global SERVER (PLGS) version 3.0.3. Generally, mass spectrum data were generated using an MS^E processing parameter with charge 2 lock

mass 785.8426, and default energy thresholds. For protein identification, default workflow parameters except an additional acetyl in N-terminal variable modification were used. A reference protein database was made from predicted 7476 coding sequences from the NCBI database (NZ_PTJS00000000). The resulted dataset was imported to ISOQuant (version 1.8, Distler *et al.*, 2014) for label-free quantification. ISOQuant processing includes spectrum alignment, data normalization at spectrum level, peptide quantification by top3 method and subsequent protein quantification. Default high identification parameters were used in the quantification process.

Of the 7476 proteins from the database, a total of 2104 proteins were identified across all the samples in MM-derived protein extracts. For each sample, on average 1069 proteins were identified. In R5-derived protein extracts 1306 proteins were identified across all samples. For each sample, on average 662 proteins were identified.

Statistical analysis of the proteomics data

Heatmaps were generated using Matplotlib package (version 3.5.0) (Hunter, 2007). Violin plots were generated using Seaborn package (version 0.11.2) (Waskom, 2021). Mann-Witney tests were performed using "mannwhitneyu" method in Scipy package (version 1.8.0) (Virtanen *et al.*, 2020).

SUPPLEMENTARY TABLES

Table S1 Shared amino acid identity of LugR proteins (in %) with their orthologues encoded by *lug*-type gene clusters

Strain/Protein	LugRII	LugRIII	LugRIV	LugRV
S. sp. CB00072	53.85	64.86	64.73	58.41
S. scopuliridis	52.75	66.22	66.82	59.72
S. sp. b84	53.85	65.32	61.13	58.57
S. sp. w007	53.85	65.77	61.13	55.53

SUPPLEMENTARY FIGURES

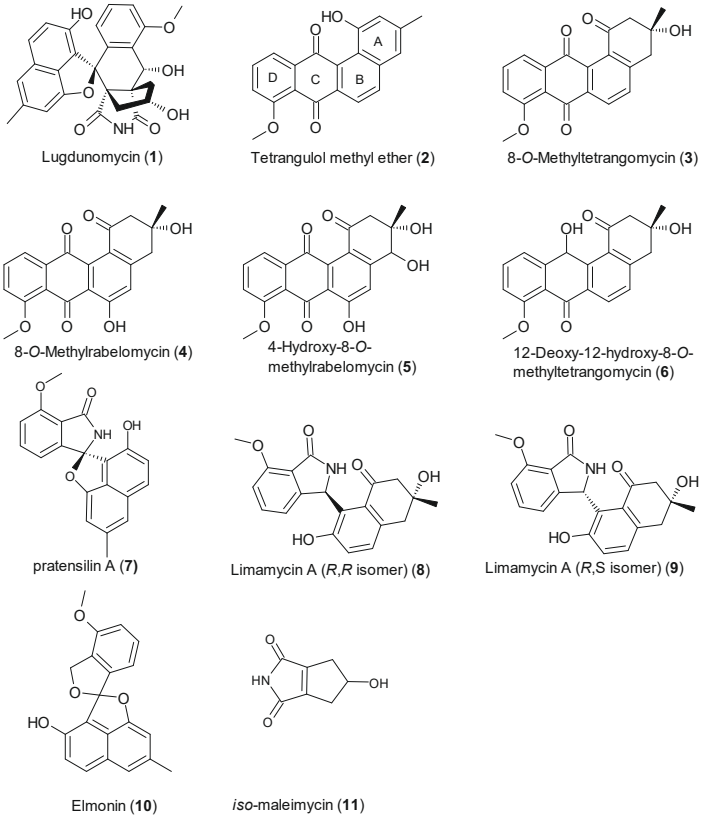


Figure S1 Structures of the compounds discussed in this study
Structures of molecules produced by *Streptomyces* sp. QL37 (**1**–**10)**. Lugdunomycin (**1**), the non-rearranged angucyclines (**2**–**6**) and the rearranged angucyclines (**7**–**9**) were previously isolated from *Streptomyces* sp. QL37 (Wu *et al.*, 2019). Elmonin (**10**) was detected in the extracts of *Streptomyces* sp. QL37 by comparison with a standard (Chapter 4)

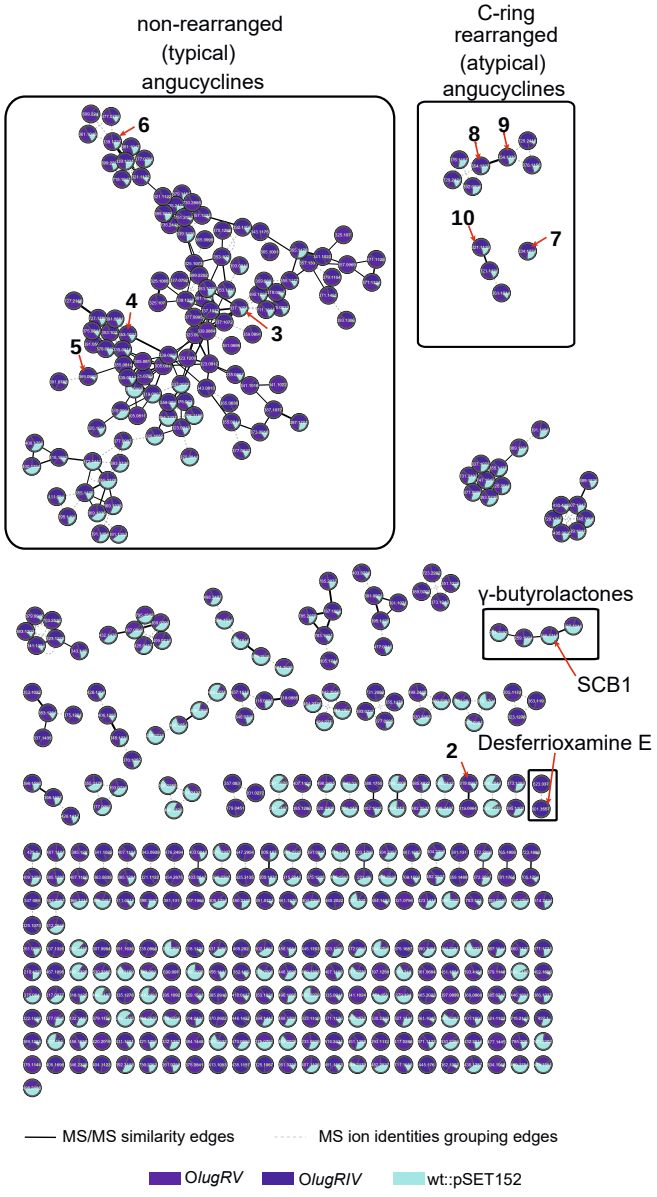


Figure S2 Molecular network of the ions detected in the extracts of *Streptomyces* sp. QL37 O*lugRIV*, O*lugRV* and the control strain grown on R5.
Molecular network of the ions detected in extracts of *Streptomyces*. sp. QL37 overexpressing *lugRIV* and *lugRV* and the control strain grown on R5. The nodes are labelled by the precursor mass of their ions and pie charts are mapped to the nodes to indicate the relative intensities of the ions in the different samples. Arrows highlight metabolites that were previously identified in *Streptomyces* sp. QL37. Molecular families that could be dereplicated are indicated.

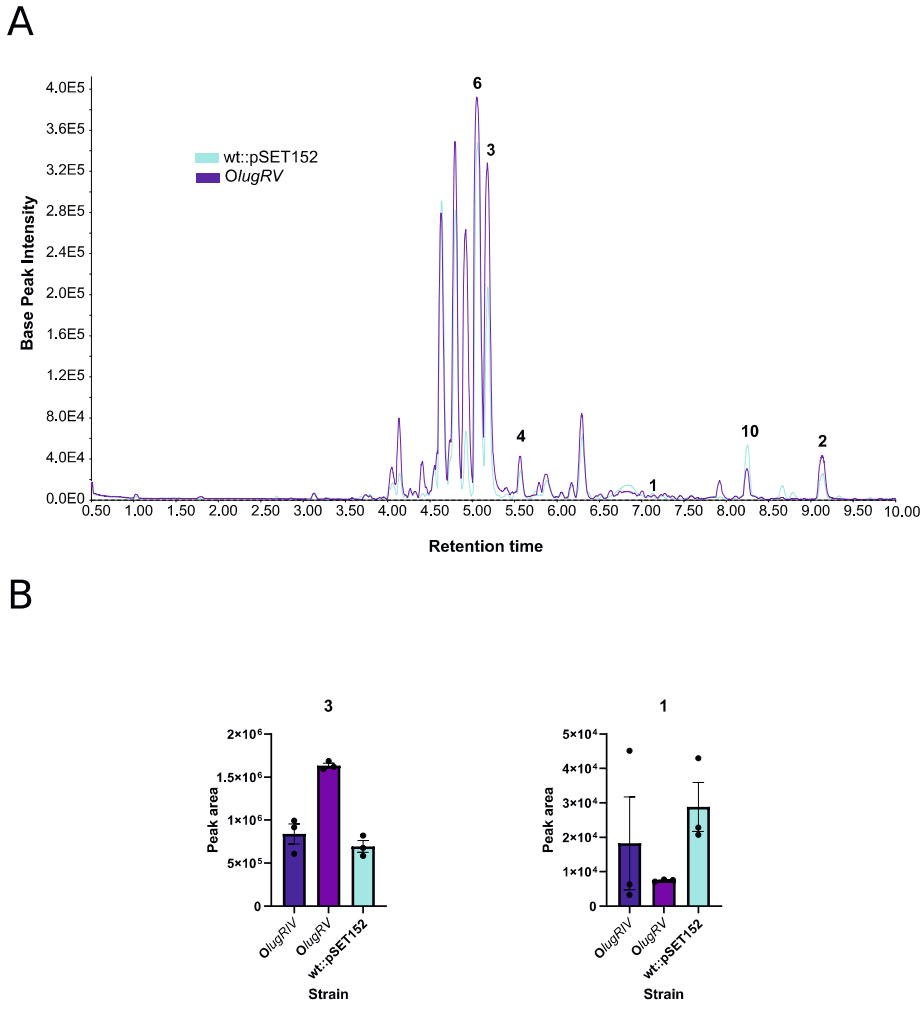


Figure S3 Production of angucyclines and lugdunomycin by of *Streptomyces* sp. QL37 overexpression *lugRIV* and *lugRV* after seven days of growth at 30 °C.
A) Overlay of LC-MS chromatograms of the control strain and O*lugRV*. B) Bar plots representing the peak areas of 8-*O*-methytetrangomycin (3) and lugdunomycin (1) with the error bar indicating the standard error of the mean (SEM).

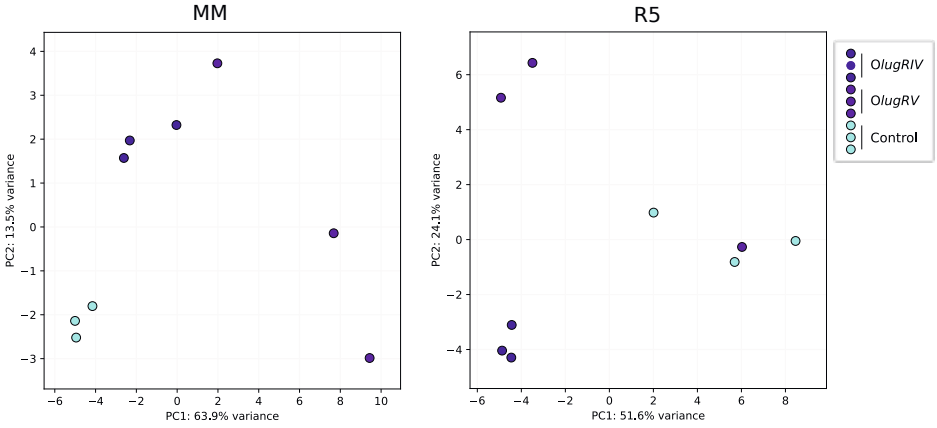


Figure S4 Exploratory statistical analysis of the proteome of *Streptomyces* sp. QL37 with enhanced expression of *lugRIV*, *lugRV* and the control strain.
PCA score plots for the extracts of the strains grown on MM (A) and R5 (B) for two days at 30 °C.

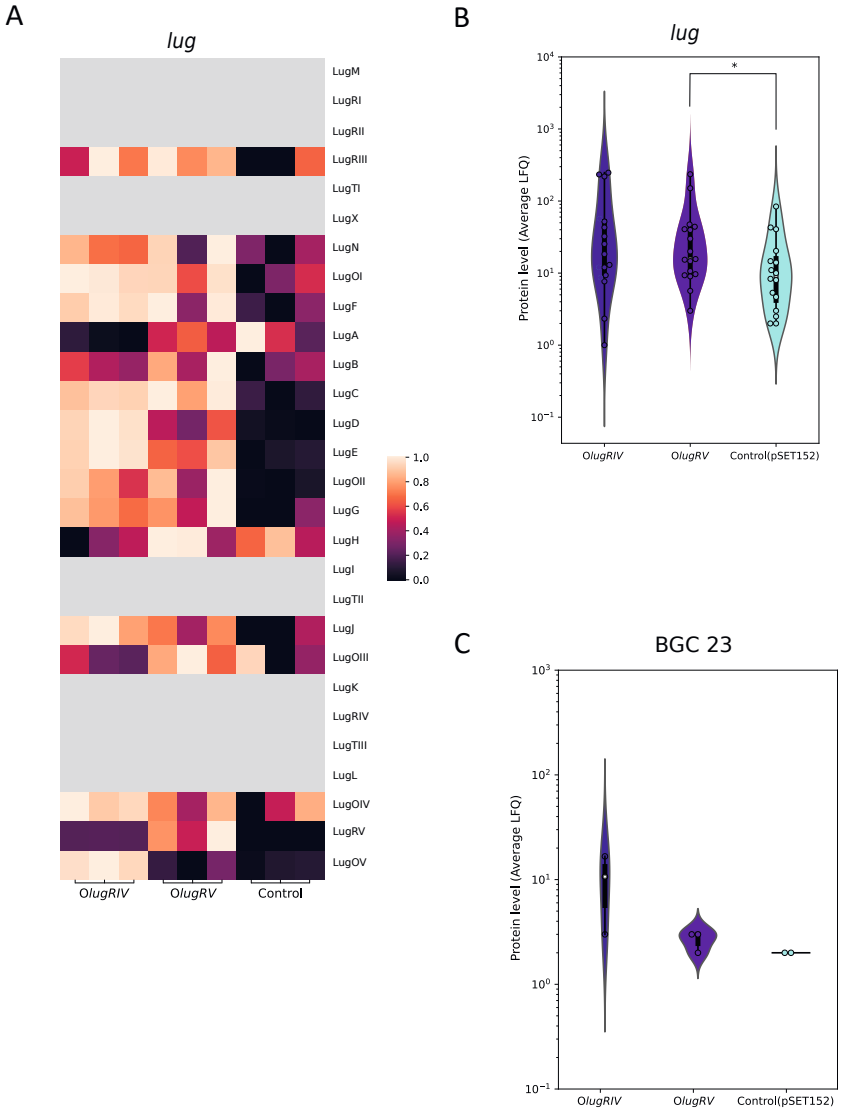


Figure S5 Statistical analysis of the proteins encoded by the gene clusters involved in lugdunomycin biosynthesis in O*lugRIV*, O*lugRV* and the control strain (QL37::pSET152). The strains were grown on R5 for two days at 30 °C. A) Heatmap shows the protein level in each replicate of O*lugRIV*, O*lugRV* and the control strain, normalised by row to the length of 1. Grey tiles indicate that the protein was not detected. B) Distribution of the overall expression level of the Lug proteins represented in a violin plot. Dots indicate the average expression level of each Lug protein; asterisk (*) indicates that the *p*-value was below 0.05 (O*lugRIV* vs control: *p*-value = 0.08372 and O*lugRV* vs control = 0.04749) C) Distribution of the overall expression level of the proteins from BGC 23. The expression of BGC 23 was not significantly different among the strains, and only three proteins were detected (O*lugRIV* vs control: *p*-value = 0.13864; O*lugRV* vs control: *p*-value = 0.31731). The *p*-values were calculated using a Mann-Witney test.

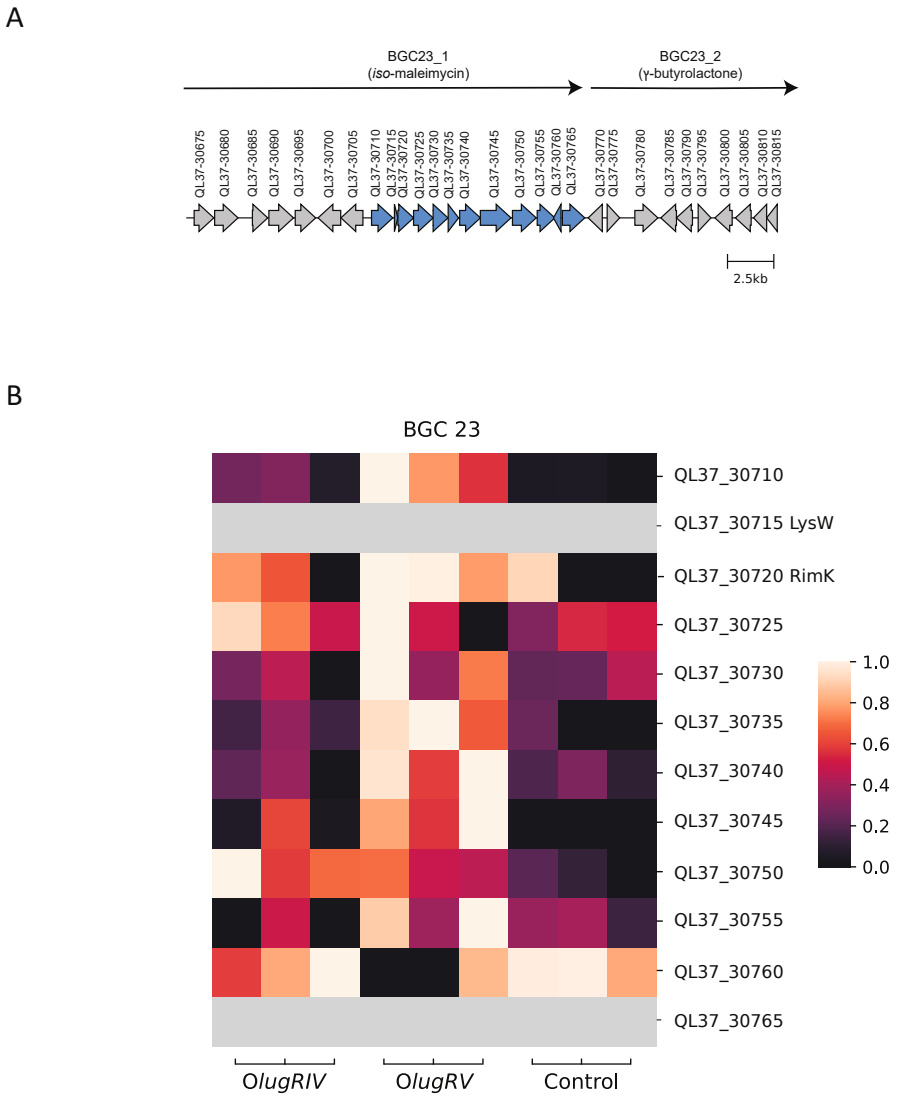
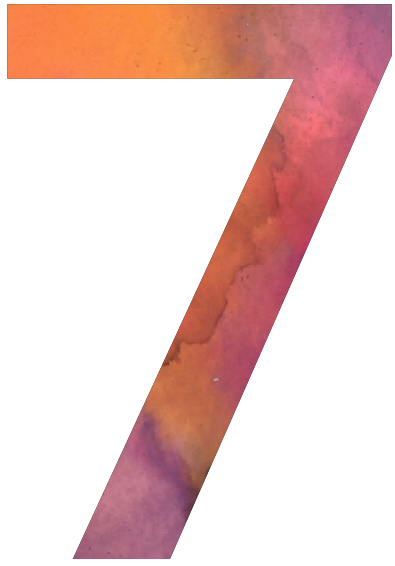


Figure S6 Organization of BGC 23 and the level of proteins encoded by BGC 23 represented in a heatmap.

A) Organization of the predicted *iso*-maleimycin BGC. Blue colours indicate the core biosynthetic genes (QL37_30710-QL37_30755) that are putatively required for the synthesis of *iso*-maleimycin (**11**), shared with the maleimycin BGC and one predicted regulator (QL37_30760) and one predicted transporter gene QL37_30765. For more details see Chapter 4. B) Heatmap shows the protein level of each protein of BGC 23 indicated in blue (A) in each replicate of *OlugRIV*, *OlugRV* and the control strain. Strains were grown on MM for two days at 30 °C. Levels were normalised by row to the length of 1. Grey tiles indicate that the protein was not detected.



General Discussion

GENERAL DISCUSSION

The increase in resistance phenomena, such as antibiotic resistant pathogenic bacteria and drug resistant tumour cells, demands the search for novel drug leads (Wright, 2017, Huang *et al.*, 2021). *Streptomyces* species are a rich source of bioactive natural products with great chemical diversity (Baltz, 2008). These filamentous soil bacteria are responsible for the production of almost half of all antibiotics used in the clinic as well as natural products applied for other medical, biotechnological and agricultural purposes (Berdy, 2005, Hopwood, 2007). Advancements in genome mining strategies revealed that genomes of streptomycetes contain many more biosynthetic gene clusters (BGCs) for natural products than initially thought, many of which are not actively transcribed under routine screening conditions (Bentley *et al.*, 2002, Medema *et al.*, 2011) This offers the perspective of a vast chemical space that may still await discovery. Approaches to activate these BGCs are discussed in Chapter 2 (van der Heul *et al.*, 2018).

One group of compounds that is well known for their antibacterial and anticancer activities are the angucyclines. The biosynthesis pathways of the typical angucyclines have been studied extensively, such as those from gaudimycin, urdamycin and landomycin (Kharel *et al.*, 2012). The initial angucycline framework, characterised by its benz[a]anthracene structure is constructed by the minimal polyketide synthase (PKS) followed by modifications by post-PKS enzymes, leading to a wide array of interesting chemical scaffolds (Kharel *et al.*, 2012). Particularly the opening of one of the aromatic rings offers a drastic structural diversification. This has been demonstrated by the discovery of the jadomycins, kinamycins and gilvocarcins, which are derived from a typical angucycline intermediate of which the B-ring has been opened by a monooxygenase (Fan & Zhang, 2018).

The discovery of the angucycline-derived molecule lugdunomycin shows that even in a well-known compound family a novel structure can be found. This molecule, produced by *Streptomyces* sp. QL37, is characterised by its unique benzaza[4,3,3] propellane-6-spiro-2'-2H-naphtho[1,8-bc]furan backbone and represents a novel subclass of polyketides (Wu *et al.*, 2019). The hypothesised biosynthesis pathway includes a Baeyer-Villiger oxidative cleavage, expansion of the C-ring followed by a Diels-Alder reaction between a C-ring rearranged angucycline as a diene and iso-

maleimycin as a dienophile. Although many different C-ring cleaved angucyclines have been discovered since the identification of lugdunomycin, the reaction mechanism for the fragmentation of the angucycline quinone moiety (the C-ring) is still unknown Thus, this molecule offers great possibilities for the discovery of novel organic chemistry and enzymology that may be applied in combinatorial (bio)synthesis of new natural products or the discovery of new BGCs. (Mikhaylov *et al.*, 2021). Lugdunomycin has antibiotic activity against Gram-positive bacteria (Wu *et al.*, 2019). The molecule is produced in very low quantities by the natural host; 0.5 mg of compound was obtained from 7.5 L of medium, which limits the opportunity to do pharmacological and pharmacokinetic tests (Wu *et al.*, 2019).

In this thesis a fundamental approach was applied to obtain knowledge on the biosynthetic reactions and the control of lugdunomycin biosynthesis, using a combination of bioinformatic analysis, gene deletion experiments, heterologous expression, metabolomics, proteomics, and RNA-seq. These approaches led to increased understanding of the enzymes encoded by the *lug* gene cluster, its regulation and the biosynthesis of lugdunomycin. One of the foreseen applications is to improve the yield of lugdunomycin and expand the angucyclines chemical space.

Characterisation of the *lug* gene cluster

Angucycline production is directed by a type II polyketide synthase (PKS). Analysis of the *Streptomyces* sp. QL37 genome using antiSMASH (Medema *et al.*, 2011) identified a polyketide type II gene cluster (designated *lug*) with high similarity to the BGC of kiamycin, an angucycline derivative produced by *Streptomyces* sp. W007 (Zhang *et al.*, 2012, Cao *et al.*, 2021, Kibret *et al.*, 2018). The genes *lugA–OII*, covering the genes *lugA–C* which encode the so-called minimal PKS enzymes required for the production of the angucycline backbone (KSα, KSβ and ACP), were deleted and the metabolomics profile was compared with that of the wild-type strain (Wu *et al.*, 2019). This *lug-pks* mutant failed to produce any rearranged, non-rearranged angucyclines and lugdunomycin, which confirmed that this gene cluster is indeed required for the production of these molecules (**Chapter 3**).

RNA-seq for transcriptome profiling was conducted on RNA isolated from *Streptomyces* sp. QL37 grown for different incubation times (24, 36, 48 and 60 h) on lugdunomycin production medium. This revealed that the genes *lugRI–lugOV*

showed a similar expression pattern, suggesting that these genes are controlled by the same regulatory mechanism and are thus part of one gene cluster. Thus, according to these data, the *lug* gene cluster comprises 27 genes, including amongst other genes encoding for the minimal PKS, five oxygenase genes and five regulatory ones.

To further verify the extent of the *lug* gene cluster and to find other possible lugdunomycin producers, a collection of over 1000 genomes of *Streptomyces* and *Kitasatospora* strains was searched for the presence of angucycline BGCs. Interestingly, some 25% of these strains contained the minimal PKS, that is characteristic for the angucycline BGCs, suggesting that angucycline production is widespread. We identified 36 strains (incl. *Streptomyces* sp. QL37) that had angucycline BGCs, and these shared at least 18 genes with the *lug* gene cluster. Comparison of the homologous BGCs with the *lug* gene cluster, revealed that it comprises 28 genes covering *lugM–lugOV,* which is close to what was predicted using RNA-seq data. Thus, based on the bioinformatics data in combination with the RNA-seq data we conclude that the *lug* gene cluster comprises 28 genes. Possibly the expression of *lugM* is controlled differently compared to *lugRI–lugOV* resulting in different expression levels.

Among the 35 angucycline BGCs, only one contained an orthologue of the *lugRI* regulator which was in the genome of *Streptomyces* sp. 94 (Hulcr *et al.*, 2011). On the other hand, all of them lacked homologues of *lugX* and *lugK*. Not a single strain had a homologue of *lugX* even outside its identified angucycline BGC, and this gene was not found anywhere in the NCBI database; therefore, it is unclear whether this encodes a functional protein and if so, what its function is. Transcripts of *lugX* were detected, strongly suggesting that it is not a pseudogene. Homologues of *lugK* were found elsewhere in the genomes of 28 out of the 35 strains. *lugK* encodes a phosphopantetheinyl transferase (PPTase), which can be used in multiple pathways (Bunet *et al.*, 2014)

A phylogenetic tree was generated based on housekeeping genes from 1020 *Streptomyces* and *Kitasatospora* strains. This revealed that strains containing a *lug*-type gene cluster were spread all over the phylogenetic tree and thus the *lug* gene cluster is not correlated to phylogeny. Strains that were closely related to

Streptomyces sp. QL37, such as *Streptomyces* sp. LamerLS-316 did not contain the *lug* gene cluster.

Comparison of the *lug* gene cluster with 27 well-characterised angucycline BGCs that direct the production of either non-rearranged, A-ring or B-ring cleaved angucyclines revealed that in total eight genes were specific to the *lug*(-type) gene clusters or those that direct the production of C-ring cleaved molecules. These included the regulatory genes *lugRI–lugRIII* and *lugRV*, the oxygenase genes *lugOIII* and *lugOIV* and the transporter gene *lugTI* (**Chapter 3**). The genes *lugOIII* and *lugOV* are most likely required for C-ring cleavage (see below). Although genes could be assigned that differentiate *lug*-type gene clusters from non-rearranged and B-ring cleaved angucycline BGCs, the question remained why the *lug* gene cluster is so special that it can direct the production of such as complex molecule. Heterologous expression of the *lug* gene cluster and flanking regions in the engineered host *Streptomyces coelicolor* M1152 resulted in the production of rearranged and non-rearranged angucyclines, but not lugdunomycin (**Chapter 4**). This further indicated that the *lug* gene cluster is truly an angucycline gene cluster directing the production of non-rearranged and C-ring cleaved angucyclines. However, which other genes are then required for lugdunomycin biosynthesis?

An additional BGC is required for the final steps of lugdunomycin biosynthesis

Lugdunomycin production requires two substrates: a C-ring cleaved rearranged angucycline and *iso*-maleimycin. We previously hypothesised that *iso*-maleimycin was derived from the limamycins (Wu *et al.*, 2019). However, we found that *iso*-maleimycin was not only observed in the wild-type, but also in the extracts of the *lug-pks* mutant (Uiterweerd, 2020). This suggested that *iso*-maleimycin is not derived from the limamycins, as previously hypothesised (Wu *et al.*, 2019). The temporal RNA-seq data revealed that the *lug* gene cluster was co-expressed with amongst others BGC 23. This BGC contains both β-lactone related genes and amino group carrier protein related genes (BGC 23a), together with γ-butyrolactone related genes (BGC23b) (Wolf *et al.*, 2017a, Robinson *et al.*, 2019, Matsuda *et al.*, 2017). Suggestively, an amino group carrier protein was proposed to be involved in maleimycin biosynthesis, a molecule produced by *S. showdoensis* ATCC 15227 (Makato, 2012-2017, Prima *et al.*, 2017, Matsuda *et al.*, 2017, Elstner *et al.*, 1973). Indeed, BGC23 and the maleimycin BGC are highly similar, and we therefore

propose that this cluster is involved in the production of *iso*-maleimycin, one of the substrates for the final reaction that generate lugdunomycin. The lack of a BGC similar to BGC23 in *S. coelicolor* explains why the heterologous host *S. coelicolor* M1152 carrying the *lug* gene cluster did not produce lugdunomycin (**Chapter 4**).

The role of the oxygenase genes of the *lug* gene cluster in lugdunomycin biosynthesis

One of the key reactions in the production of lugdunomycin is the cleavage of the C-ring in the angucycline, which is required to generate rearranged angucyclines as well as lugdunomycin, and is likely catalysed by a Baeyer–Villiger oxidation (Wu *et al.*, 2019). The *lug* gene cluster encodes five putative oxygenases, namely *lugOI–lugOV*, and their functional role in lugdunomycin production was studied by deletion experiments and metabolomic studies (**Chapter 5**).

The presence of a range of oxygenase genes with different functionalities in angucyclines BGCs is important for the diversification of the angucyclines that are biosynthesised (Fan & Zhang, 2018). Orthologues of LugOI and LugOII, such as PgaE and PgaM are important for the diversification of non-rearranged angucyclines (Patrikainen *et al.*, 2012, Fan & Zhang, 2018). Deletion of *lugOI* or *lugOII* resulted in mutants that had lost the ability to produce most of the unrearranged angucyclines and were incapable of producing the rearranged angucyclines and lugdunomycin. Thus, these genes are likely involved in early post-PKS modifications. These results are well in line with the known role of the homologues of LugOI (PgaE, UrdE, LanE) and LugOII (PgaM, UrdM, LanV), which are involved in the biosynthesis of gaudimycin, urdamycin and landomycin, respectively (Kharel & Rohr, 2012, Mayer *et al.*, 2005). Functional analysis showed that LugOII catalyses C6 ketoreduction, similar to other LugOII-like enzymes. Interestingly, previous *in vitro* experiments also revealed that LugOII can perform C1 ketoreduction, a reaction that has not been reported before for LugOII homologues (Xiao *et al.*, 2020). Deletion of *lugOIV* did not significantly affect angucycline biosynthesis, and its role in lugdunomycin biosynthesis, if any, so far remains elusive.

LugOIII and LugOV showed low similarity to antibiotic monooxygenases. The *lugOIII* and the *lugOV* mutants produced canonical angucyclines, but failed to produce the rearranged ones. These results suggest that these enzymes are candidates for the cleavage of the C-ring. Interestingly, analysis of the molecular

network of the *lugO* mutants revealed a likely hierarchical role between LugOIII and LugOV, whereby LugOIII catalyses the production of an angucycline epoxide that is used by LugOV for the subsequent Baeyer–Villiger oxidation reaction, resulting in the C-ring cleavage of angucyclines. These results are in line with the observation that the production of C-ring cleaved angucyclines governed by the *tac* gene cluster of *Streptomyces* sp. CB00072 requires TacS and TacT, which are orthologues of LugOIII and LugOV, respectively (Cao *et al.*, 2021). Altogether this study lays groundwork for the generation of novel C-ring cleaved rearranged angucyclines using combinatorial biosynthesis. *In vitro* assays using the purified oxygenases should further give insight into the functional role of the oxygenases encoded by the *lug* gene cluster.

Characterisation of the regulatory genes of the *lug* gene cluster

Streptomycetes harbour up to 1000 different transcription factors, forming a hugely complex regulatory circuitry that allows them to respond to the cacophony of signals they receive from the environment, so as to formulate appropriate responses, including decisions when to initiate development and specialised metabolism. Insights into the regulation of the *lug* gene cluster could provide insight into the role of lugdunomycin in the environment of *Streptomyces* sp. QL37 and why it is hardly produced in the laboratory (van Bergeijk *et al.*, 2020). Regulation of BGCs is controlled by both pleiotropic and cluster-situated regulators (Liu *et al.*, 2013). In this study the cluster-situated regulators (CSRs) encoded by the *lug* gene cluster were studied.

A mutational study of *lugRII–lugRV* showed that LugRII (LuxR-type), LugRIV (atypical response regulator) and LugRV (SARP-regulator) are required for angucyclines production in *Streptomyces* sp. QL37 and hence also for lugdunomycin production, while LugRIII (TetR-regulator) only plays a minor role in the regulation of the *lug* gene cluster. Overexpression of *lugRIV* and *lugRV* led to enhanced angucycline production compared to the control strain in early growth stages and - importantly - overexpression of *lugRV* led to improved lugdunomycin production. Based on these results, the role of LugRIV and LugRV was investigated further. The role of LugRIV and LugRV as transcriptional activators was further validated by quantitative proteomics, which showed that the proteins of the lugdunomycin BGC were significantly upregulated. Notably at later stages of growth, overexpression of *lugRV* increased the production of the angucyclines,

but not lugdunomycin (**Chapter 6**). As mentioned above, the likely final reaction required for lugdunomycin includes two substrates; a rearranged angucycline and *iso*-maleimycin (Wu *et al.*, 2019, Uiterweerd, 2020). Therefore, BGC23 also needs to be upregulated to achieve lugdunomycin biosynthesis, assuming that no other enzymes encoded by genes outside the two BGCs are required for the biosynthesis of lugdunomycin

The CSRs of a regulatory network act in a hierarchical process and the DNA activity of the regulators belonging to the class of LuxR, TetR and atypical response regulators can be modulated by the intermediates and the final product of the biosynthesis pathway (van der Heul *et al.*, 2018, Wang *et al.*, 2009). These two aspects lead to (multiple) feedback and/or feedforward loops, which does not yet take pleiotropic regulators into account. Examples of these complex regulatory networks are those of the auricin- and jadomycin BGC of *S. aureofaciens* CCM 3239 and *S. venezuelae* ISP5230, respectively (Kormanec *et al.*, 2014, Zou *et al.*, 2014). Cross-complementation of the generated mutants suggested that the response regulator LugRIV activates the expression of the SARP regulator LugRV that in turn controls the expression of the structural *lug* genes (**Chapter 6**). Further molecular biological studies, such as DNA-binding experiments and extensive transcriptional analysis, are required to obtain more insights into precisely how the LugR proteins control the expression of the *lug* gene cluster.

Chemical diversity of the secondary metabolome of *Streptomyces* sp. QL37.
With 35 BGCs predicted in its genome, *Streptomyces* sp. QL37 has the capacity to produce a plethora of secondary metabolites other than angucyclines and *iso*-maleimycin. Indeed, under the culture conditions tested in this study, the wild-type strain was found to also produce γ-butyrolactones (produced from BGC 16, 23 or 33), tetramate macrolactams (produced from BGC 6) and sceliphrolactam (produced from BGC 9) (Yang *et al.*, 2005, Moree *et al.*, 2014, Low *et al.*, 2018) (**Chapter 4**). Furthermore, some molecular families that were unrelated to angucyclines were specifically observed in some of the *lugO* null mutants. On MM agar the accumulation of molecules related to *N*-acyl glutamine and aminoalcohols were observed, in particular in *lugOI*, *lugOII* and *lugOIII* null mutants (Battista *et al.*, 2019, Won *et al.*, 2014, Harrison *et al.*, 2018). On R5 agar the accumulation of a novel lipopeptide was observed in extracts derived from the *lugOI* null mutant (possibly produced from BGC 21) (**Chapter 5**). A more extensive

OSMAC approach, combined with genetic manipulation of the strain, should shed more light on the extent of the specialised metabolome of *Streptomyces* sp. QL37, and discover the chemical space of the natural products it can produce. Different studies already reported that affecting the production levels of natural products that are abundantly produced by the host, can lead to significant alterations in the production of other unrelated secondary metabolites (Iorio *et al.*, 2021, Culp *et al.*, 2019). This was demonstrated by the fact that inactivation of the BGCs for the highly produced antibiotics streptomycin and streptothricin led to the production of previously hidden rare unknown natural products (Culp *et al.*, 2019). In addition, blocking several steps in the pseudouridimycin (PUM) pathway in *Streptomyces* sp. ID38640 led to the production of previously unseen and unrelated metabolites. Thus, interfering with one BGC can unveil a new potential of the producer strain and possibly be used to discover new secondary metabolites and antibiotics (Iorio *et al.*, 2021).

Morphological changes associated with angucycline production.

Production of angucyclines disturbs growth and blocks development, most likely due to cytotoxicity as a result of their DNA-degrading properties (Kharel *et al.*, 2012). In line with that, we observed major differences in morphological differentiation between the wild-type strain on the one hand, and the mutants that failed to produce angucyclines (Δ*lug-pks*, Δ*lugRII*, Δ*lugRIV* and Δ*lugRV*) on the other, particularly on R5 agar (**Chapter 6**). Angucycline non-producers developed well, whereas the strains producing angucyclines, including the wild-type strain, were inhibited in their development, resulting in a nonsporulating phenotype. Previously it was observed that the cytotoxic DNA-damaging prodiginines produced by *S. coelicolor* postpone sporulation of the strain on rich media. The *redD* mutant of *S. coelicolor*, which is unable to produce these molecules, showed precocious morphological differentiation on R2YE agar plates as compared to the parent strain (Tenconi *et al.*, 2020). The authors proposed that prodiginines mediate the onset of programmed cell death of the vegetative mycelium, postponing the developmental program until sufficient nutrients are produced to support extensive aerial growth. The accelerated sporulation and smaller colonies of *redD* mutants are then explained by the lack of cell death and hence accelerated life cycle. Angucyclines may play a similar role in the control of cell death and development due to their DNA-damaging properties (Tenconi & Rigali, 2018).

Concluding remarks and future perspectives

This study provides new insights into the biosynthesis and regulation of angucyclines and the highly rearranged angucycline derivative lugdunomycin. These new insights may be applied to increase lugdunomycin production by *Streptomyces* sp. QL37 with the aim to isolate significant quantities of lugdunomycin and determine its mode-of-action and that of related molecules.

The *lug* gene cluster consists of 28 genes, including the minimal PKS, five oxygenase genes (*lugOI–lugOV*) and five regulatory genes (*lugRI–lugRV*). LugOIII and LugOV likely mediate C-ring cleavage during angucycline biosynthesis, which is one of the key reactions in lugdunomycin biosynthesis. In addition, we hypothesise that an angucycline epoxide is possibly the required substrate for C-ring opening. Homologous BGCs were found in other *Streptomyces* strains. BGCs that contain *lugOIII* and/or *lugOV* homologues likely specify a diversity of C-ring cleaved angucyclines and perhaps also lugdunomycin. However, the final reaction in the lugdunomycin pathway likely requires two BGCs, one producing the diene (angucycline) and the other producing the dienophile (*iso*-maleimycin) of a Diels-Alder reaction, which is proposed to be spontaneous, based on the fact that a racemic mixture is produced (Wu *et al.*, 2019).

An interesting question that arises is whether lugdunomycin biosynthesis offers an evolutionary advantage to the producer. In a general sense, the concept that molecules may be produced by multiple BGCs may be much more common than we currently anticipate. We already observed this phenomenon for the production of endophenasides by *Kitasatospora* sp. MBT 66 and actinomycin L from *Streptomyces* sp. MBT 27 (Wu *et al.*, 2016b, Machushynets *et al.*, 2022). Indeed, most of the drug discovery research emanates from the idea that one molecule is derived from one BGC, and this is also the basis for synthetic biology approaches. The example of lugdunomycin shows that in this way interesting chemical space may be missed and we should still rely on the chemistry isolated from the producer strain.

Nederlandse Samenvatting

NEDERLANDSE SAMENVATTING

De toename van resistentie, zoals antibioticaresistente pathogene bacteriën en chemotherapie-resistente tumorcellen, vraagt om nieuwe aanknopingspunten voor de ontdekking van nieuwe medicijnen (Wright, 2017, Huang *et al*., 2021). Streptomyceten zijn een rijke bron van secundaire metabolieten met een ruime chemische diversiteit en veelal met een geneeskrachtige werking (Baltz, 2008). Deze filamenteuze bodembacteriën zijn verantwoordelijk voor de productie van bijna de helft van alle in de kliniek gebruikte antibiotica, als ook van secundaire metabolieten die voor andere medische, biotechnologische en landbouwdoeleinden worden gebruikt (Berdy, 2005, Hopwood, 2007). Vooruitgang op het gebied van "genome mining" strategieën heeft aangetoond dat de genomen van streptomyceten veel meer biosynthetische genclusters (BGC's) voor secundaire metabolieten bevatten dan eerder gedacht, waarvan vele niet tot expressie komen onder standaard laboratoriumomstandigheden (Bentley *et al*., 2002, Medema *et al*., 2011). Specifieke signalen of "elicitors" zijn nodig om de expressie van deze BGCs te activeren (van Bergeijk *et al*., 2020). Zoals besproken in **hoofdstuk 2** worden signalen vanuit de omgeving opgevangen en doorgegeven door een complex regulatie systeem in de streptomyceet. Dit systeem reguleert de groei, het primaire metabolisme en de productie van secundaire metabolieten (van der Heul *et al*., 2018).

Angucyclines zijn chemische verbindingen die bekend staan om hun antibacteriële en anti-tumor activiteit. Deze groep moleculen behoort tot een van de grootste familie van secundaire metabolieten: de polyketiden. Veel van deze moleculen zijn klinisch relevant als antibioticum of als chemotherapie, zoals oxytetracycline en doxorubicin (Tibrewal & Tang, 2014). Streptomyceten hebben verschillende polyketide synthasen (PKS) die behoren tot de klassen PKSI, PKSII, of PKSIIII. Deze enzymen creëren verschillende structuren van kleine koolstofverbindingen. De gevormde structuur wordt dan verder gemodificeerd door post-PKS modificatie enzymen zoals oxygenases, methyltransferases en glycosidases (Tibrewal & Tang, 2014). Het angucycline-skelet, gekenmerkt door zijn typische benzo[a]antraceen-structuur, wordt opgebouwd van 1 acetyl-CoA en 9 malonyl-CoA moleculen door een enzymgroep genaamd de minimale "type II polyketide synthase" (PKSII). Het angucycline-skelet wordt vervolgens gemodificeerd door post-PKS-enzymen hetgeen leidt tot een breed scala aan chemische verbindingen (Kharel *et al*., 2012).

De biosynthese van verschillende angucyclines is uitgebreid bestudeerd, zoals die van gaudimycine, urdamycine en landomycine (Kharel *et al.*, 2012). Vooral het openen van één van de aromatische ringen in het angucycline-skelet zorgt voor een drastische structurele diversificatie. Mooie voorbeelden zijn de jadomycines, kinamycines en gilvocarcines, welke zijn afgeleid van een typisch angucycline tussenproduct waarvan de B-ring is gesplitst door een monooxygenase (Fan & Zhang, 2018).

De ontdekking van lugdunomycine binnen de familie van angucyclines laat zien dat zelfs in een bekende moleculaire familie een totaal nieuwe structuur kan worden gevonden. Lugdunomycine, geproduceerd door *Streptomyces* sp. QL37, wordt gekenmerkt door een uniek benzaza[4,3,3]propellaan-6-spiro-2'-2H-nafto[1,8-bc] furan-skelet en vertegenwoordigt een nieuwe subklasse van polyketiden (Wu *et al.*, 2019). De voorspelde biosyntheseroute bevat een Baeyer-Villiger oxidatieve splitsing, expansie van de C-ring, gevolgd door een Diels-Alder reactie tussen een C-ring-gesplitste angucycline als dieen en *iso*-maleimycine als dienofiel. Hoewel sinds de identificatie van lugdunomycine veel verschillende C-ring-gesplitste angucyclines zijn ontdekt, is het reactiemechanisme voor de opening van het angucycline-quinone gedeelte (de C-ring) onbekend. Lugdunomycine biedt dus veel mogelijkheden voor de ontdekking van nieuwe organische chemie en van enzymen die kunnen worden toegepast in de combinatoire (bio)synthese van nieuwe secundaire metabolieten of de ontdekking van nieuwe BGC's (Mikhaylov *et al.*, 2021). Lugdunomycine heeft antibacteriële activiteit tegen Gram-positieve bacteriën (Wu *et al.*, 2019). Het molecuul wordt in zeer kleine hoeveelheden geproduceerd door *Streptomyces* sp. QL37; slechts 0,5 mg van het molecuul kon worden verkregen uit 7,5 Liter agar, wat de mogelijkheid om farmacologische en farmacokinetische testen te doen beperkt (Wu *et al.*, 2019).

In dit proefschrift is een fundamentele benadering toegepast om kennis te verkrijgen over de biochemische reacties en de regulatie van de lugdunomycinebiosynthese, door middel van een combinatie van bioinformatische analyse, gen-deletie-experimenten, heterologe expressie, metabolomics, proteomics en transcriptomics. Dit heeft geleid tot een beter begrip van de enzymen die door het *lug* gencluster worden gecodeerd, de regulatie ervan en de biosynthese van lugdunomycine. Een mogelijk toepassing is het verbeteren van de opbrengst van lugdunomycine en het verder uitbreiden van de chemische diversiteit van angucyclines in het algemeen.

Karakterisering van het *lug* gencluster

Angucyclines worden gesynthetiseerd door een type II polyketide synthase (PKS). Analyse van het *Streptomyces* sp. QL37 genoom met behulp van anti-SMASH (Medema *et al.*, 2011) leidde tot de identificatie van een polyketide type II-gencluster (genaamd *lug*) wat sterk lijkt op de BGC voor kiamycine, een angucycline geproduceerd door *Streptomyces* sp. W007 (Zhang *et al.,* 2012, Cao *et al.,* 2021, Kibret *et al.,* 2018). De genen *lugA-OII*, waaronder *lugA-C* die coderen voor de zogenaamde minimale PKS-enzymen die nodig zijn voor de productie van het angucycline-skelet (KSα, KSβ en ACP), werden uitgeschakeld en het metabolische profiel werd vergeleken met dat van de wild-type stam (Wu *et al.,* 2019). Deze *lug-pks* mutant produceerde geen angucyclines en lugdunomycine, wat bevestigt dat dit gencluster inderdaad nodig is voor de productie van deze moleculen (**Hoofdstuk 3**).

Voor een volledige transcriptie-analyse werd RNA-sequencing toegepast op RNA geïsoleerd uit *Streptomyces* sp. QL37, welke 24, 36, 48 of 60 uur was gegroeid op lugdunomycine-productiemedium. Hieruit bleek dat de genen *lugRI–lugOV* op een zeer vergelijkbare manier zijn gereguleerd, wat suggereert dat deze genen zeer waarschijnlijk deel uitmaken van één gencluster. Volgens deze gegevens omvat het *lug* gencluster daarom 27 genen, waaronder genen die coderen voor de minimale PKS, vijf oxygenases en vijf regulator eiwitten.

Om de omvang van het *lug* gencluster verder te verifiëren en om andere mogelijke lugdunomycine producenten te vinden, werd een verzameling van meer dan 1000 genomen van *Streptomyces*- en *Kitasatospora*-stammen onderzocht op de aanwezigheid van angucycline BGC's. Ongeveer 25% van deze stammen bevatten de minimale PKS die kenmerkend is voor de angucycline BGC's, hetgeen suggereert dat angucyclineproductie wijdverspreid is. We identificeerden 36 stammen (waaronder *Streptomyces* sp. QL37 zelf) die angucycline BGC's hadden, waarbij 18 of meer genen gedeeld waren met het *lug* gencluster. Vergelijking van BGC's die zeer sterk lijken op het *lug* gencluster wees uit dat 28 genen voorkomen in meerdere BGC's, hetgeen dicht in de buurt komt van wat werd voorspeld met behulp van RNA-sequencing data. Op basis van de bioinformatica data in combinatie met de RNA-sequencing data concluderen wij dat het *lug* gencluster 28 genen omvat. Mogelijk wordt de expressie van *lugM* anders gereguleerd dan

die van *lugRI–lugOV,* wat resulteert in verschillende expressielevels voor dit ene gen ten opzichte van de andere 27 genen.

Van de 35 angucycline BGC's bevatte er slechts één een ortholoog van de *lugRI* regulator, namelijk het BGC in het genoom van *Streptomyces* sp. 94 (Hulcr *et al.*, 2011). Anderzijds ontbraken in alle stammen homologen van *lugX* en *lugK*. Geen enkele stam had een homoloog van *lugX*, zelfs niet buiten de geïdentificeerde angucycline BGCs en dit gen werd zelfs überhaupt niet in de NCBI database gevonden. Het is daarom onduidelijk of dit gen een functioneel eiwit codeert en zo ja, wat de functie ervan is. Echter, aangezien *lugX* transcripten werden gedetecteerd, nemen we aan dat het geen pseudogen is. Homologen van *lugK*, dat codeert voor een fosfopantetheinyltransferase (PPTase) dat in meerdere biosyntheseroutes kan worden gebruikt (Bunet *et al.,* 2014), werden elders in het genoom van 28 van de 35 stammen aangetroffen.

Er werd een fylogenetische boom gegenereerd op basis van huishoud genen van 1020 *Streptomyces* en *Kitasatospora* stammen. Hieruit bleek dat stammen die een *lug* gencluster bevatten over de hele fylogenetische boom zijn verspreid en dat het *lug* gencluster dus niet aan de fylogenie is gerelateerd. Stammen die nauw verwant zijn aan *Streptomyces* sp. QL37, zoals *Streptomyces* sp. LamerLS-316 en *S. atroolivaceus* ISP5137 bevatten het *lug* gencluster niet. Vergelijking van het *lug* gencluster met 27 goed gekarakteriseerde angucycline BGC's die nodig zijn voor de productie van typische, A-ring- of B-ring-gesplitste angucyclines, toonde aan dat er in totaal acht genen specifiek zijn voor de *lug*(-type) genclusters ofwel genclusters die verantwoordelijk zijn voor de productie van C-ring-gesplitste moleculen. Deze omvatten regulatie-genen *lugRI–lugRIII* en *lugRV*, de oxygenase-genen *lugOIII* en *lugOIV* en het transporter gen *lugTI* (**Hoofdstuk 3**). De genen *lugOIII* en *lugOV* zijn waarschijnlijk nodig voor de C-ring splitsing (zie hieronder). Hoewel genen konden worden aangewezen die *lug*-genclusters onderscheiden van typische en B-ring-gesplitste angucycline BGC's, bleef de vraag wat er zo bijzonder is aan het *lug* gencluster dat het verantwoordelijk is voor de productie van zo'n complex molecuul. Heterologe expressie van het *lug* gencluster en flankerende gebieden in de heterolge productie stam *Streptomyces coelicolor* M1152 resulteerde in de productie van C-ring-gesplitste en typische angucyclines, maar niet van lugdunomycine (**Hoofdstuk 4**). Dit wees erop dat het *lug* gencluster een angucycline gencluster is dat verantwoordelijk is voor de productie van typische

en C-ring-gesplitste angucyclines. De vraag blijft daarom welke andere genen er nodig zijn voor de biosynthese van lugdunomycine.

Een tweede gencluster is nodig voor de laatste stappen van de biosynthese van lugdunomycine

Lugdunomycine productie vereist twee substraten: een C-ring-gesplitste angucycline en *iso*-maleimycine. Aanvankelijk dachten we dat *iso*-maleimycine zou kunnen worden afgeleid van de limamycines, omdat er geen genen aanwezig zijn die het anders zouden kunnen verklaren (Wu *et al.,* 2019). Echter, dit wordt tegengesproken door experimenten die laten zien dat *iso*-maleimycine ook wordt geproduceerd door mutanten die geen angucycilines kunnen maken (Uiterweerd, 2020). Dit suggereerde dat *iso*-maleimycine niet kan zijn afgeleid van de limamycines. Uit de RNA-sequencing data bleek dat het *lug* gencluster samen met onder andere BGC23 tot expressie kwam. Dit BGC bevat genen die betrokken zijn bij de biosynthese van β-lactone en codere voor een "amino-group-carrier" eiwit (BGC23a) en biosynthese van γ-butyrolactones (BGC23b) (Wolf *et al.,* 2017a, Robinson *et al.,* 2019, Matsuda *et al.,* 2017). Een "amino-group-carrier"-eiwit is betrokken bij de biosynthese van maleimycine, een molecuul dat wordt geproduceerd door *S. showdoensis* ATCC 15227 (Makato, 2012-2017, Prima *et al.,* 2017, Matsuda *et al.,* 2017, Elstner *et al.,* 1973). BGC23 en het BGC voor maleimycine lijken sterk op elkaar, wat sterk suggereert dat dit cluster verantwoordelijk is voor de productie van *iso*-maleimycine. Zo'n soort BGC ontbreekt in *S. coelicolor*, hetgeen zeer waarschijnlijk verklaart waarom de heterologe gastheer *S. coelicolor* M1152 met het *lug* gencluster wel angucyclines en limamycines maar geen lugdunomycine produceert (**Hoofdstuk 4**).

De rol van de oxygenase-genen van het *lug* gencluster bij de biosynthese van lugdunomycine

Een van de belangrijkste reacties bij de productie van lugdunomycine is de splitsing van de C-ring in de angucycline, die nodig is om zowel C-ring-gesplitste-angucyclines als lugdunomycine te genereren en die waarschijnlijk wordt gekatalyseerd door een Baeyer-Villiger oxidatie (Wu *et al.,* 2019). Het *lug* gencluster codeert voor vijf mogelijke oxygenases, namelijk *lugOI–lugOV* en hun functie in de productie van lugdunomycine werd bestudeerd door mutagenese en analyse van het metaboloom (**Hoofdstuk 5**).

De aanwezigheid van een reeks oxygenasegenen met verschillende functies in angucycline BGC's is belangrijk voor de diversificatie van de angucyclines die worden geproduceerd (Fan & Zhang, 2018). Orthologen van LugOI en LugOII, zoals PgaE en PgaM, zijn belangrijk voor de diversificatie van typische angucyclines (Patrikainen *et al.,* 2012, Fan & Zhang, 2018). Deletie van *lugOI* of *lugOII* resulteerde in mutanten die de meeste typische angucyclines niet produceerden. Bovendien produceren deze mutanten noch de C-ring-gesplitste angucyclines noch lugdunomycine zelf. Deze genen zijn dus waarschijnlijk betrokken bij de vroege post-PKS modificaties. Deze resultaten komen overeen met de bekende functie van de homologen van LugOI (PgaE, UrdE, LanE) en LugOII (PgaM, UrdM, LanV), die betrokken zijn bij de biosynthese van respectievelijk gaudimycine, urdamycine en landomycine (Kharel & Rohr, 2012, Mayer *et al.*, 2005). Functionele analyse toonde aan dat LugOII een C6 ketoreductie katalyseert, vergelijkbaar met andere LugOII-achtige enzymen. Interessant is dat eerdere *in vitro* experimenten lieten zien dat LugOII ook een C1 ketoreductie kan katalyseren, wat nog niet eerder voor LugOII-homologen is gerapporteerd (Xiao *et al.,* 2020). Het uitschakelen van *lugOIV* had geen significante invloed op de biosynthese van angucyclines. Of dit gen een rol speelt in de biosynthese van lugdunomycine – en zo ja, welke - is vooralsnog onbekend.

LugOIII en LugOV vertonen weinig gelijkenis met andere mono-oxygenases die bekend zijn van de biosynthese van secundaire metabolieten. De *lugOIII*- en de *lugOV*-mutanten produceren standaard angucyclines, maar niet de C-ring-gesplitsten. Deze resultaten suggereren dat deze enzymen mogelijk betrokken kunnen zijn bij de splitsing van de C-ring. Interessant is dat analyse van het moleculaire netwerk van de *lugO* mutanten een waarschijnlijke hiërarchie tussen LugOIII en LugOV aan het licht bracht, waarbij LugOIII de productie van een angucycline-epoxide katalyseert dat door LugOV wordt gebruikt voor de daaropvolgende Baeyer-Villiger oxidatiereactie, resulterend in de C-ring splitsing van angucyclines. Inderdaad zijn TacS en TacT, orthologen van respectievelijk LugOIII en LugOV, nodig voor de productie van C-ring-gesplitste angucyclines door *Streptomyces* sp. CB00072 (Cao *et al.,* 2021). Al met al legt deze studie de basis voor het genereren van nieuwe C-ring-gesplitste angucyclines met behulp van combinatoire biosynthese. *In vitro* testen met de gezuiverde oxygenases kunnen verder inzicht geven in de functionele rol van de oxygenases gecodeerd door het *lug* gencluster.

Nader inzicht in de regulatie van het *lug* gencluster

Streptomyceten bevatten tot wel 1000 verschillende transcriptiefactoren, die een enorm complex systeem vormen dat hen in staat stelt te reageren op de kakofonie van signalen die ze uit de omgeving ontvangen. Op deze manier komen reacties in gang die aangepast zijn aan de natuurlijk omgeving, waaronder groei en de productie van secundaire metabolieten. Studie van de regulatie van het *lug* gencluster zou inzicht kunnen geven in de rol van lugdunomycine in de natuurlijke omgeving van *Streptomyces* sp. QL37 en waarom het nauwelijks in het laboratorium wordt geproduceerd (van Bergeijk *et al.,* 2020). De expressie van BGC's wordt gecontroleerd door zowel pleiotrope als clustergerichte regulatoren (Liu *et al.*, 2013). In deze studie werden de cluster-gesitueerde regulatoren (CSRs) gecodeerd door het *lug* gencluster bestudeerd.

Uit een mutatiestudie bleek dat LugRII (LuxR-type), LugRIV (atypische response regulator) en LugRV (SARP-regulator) nodig zijn voor de angucyclineproductie in *Streptomyces* sp. QL37 en dus ook voor de productie van lugdunomycine, terwijl LugRIII (TetR-regulator) slechts een ondergeschikte rol speelt in de regulatie van het *lug* gencluster. Overexpressie van LugRIV of LugRV leidde tot een verhoogde angucyclineproductie in vergelijking met het wildtype in vroege groeistadia en – belangrijk – overexpressie van *lugRV* leidde tot een verbeterde lugdunomycine productie. Op basis van deze resultaten werd de rol van LugRIV en LugRV verder onderzocht. De rol van LugRIV en LugRV als transcriptie-activatoren werd verder gevalideerd door kwantitatieve proteomics, waaruit bleek dat de expressie van de Lug-eiwitten afhankelijk was van LugRIV of LugRV. Opmerkelijk was dat in latere groeistadia de expressie van LugRV de productie van de angucyclines verhoogde, maar niet die van lugdunomycine zelf (**Hoofdstuk 6**). Dit wordt zeer waarschijnlijk verklaard door het feit dat lugdunomycinebiosynthese naast een afgeleide van angucyclines zelf ook *iso*-maleimycine nodig heeft (Wu *et al.*, 2019, Uiterweerd, 2020). Daarom moet ook de expressie van BGC23 worden verhoogd.

Regulatoren die specifiek zijn voor de controle van een bepaald gencluster werken vaak hiërarchisch, waarbij de ene de andere activeert; tegelijk kan de activiteit van de regulatoren worden beïnvloed door moleculen die tijdens de biosynthese worden geproduceerd (van der Heul *et al.,* 2018, Wang *et al.,* 2009). Deze twee aspecten leiden tot (meerdere) feedback en/of feedforward loops, waarbij nog geen rekening is gehouden met pleiotrope regulatoren die een meer algemen rol

spelen. Voorbeelden van zulke complexe regulatoire netwerken zijn die van de auricine- en jadomycine BGC's van respectievelijk *S. aureofaciens* CCM 3239 en *S. venezuelae* ISP5230 (Kormanec *et al.,* 2014, Zou *et al.,* 2014). Kruis-complementatie van de mutanten suggereerde dat de responsregulator LugRIV de expressie van de SARP-regulator LugRV activeert, die op zijn beurt de expressie van de structurele *lug* genen controleert (**Hoofdstuk 6**). Verder moleculair biologisch onderzoek, zoals DNA-bindingsexperimenten en uitgebreide transcriptie-analyse, moet meer inzicht verschaffen in het precieze regulatienetwerk van het *lug* gencluster.

Chemische diversiteit van het secundaire metaboloom van *Streptomyces* sp. QL37

Met 35 voorspelde BGC's in zijn genoom heeft *Streptomyces* sp. QL37 het vermogen om een overvloed aan andere secundaire metabolieten dan angucyclines en *iso*-maleimycine te produceren. Onder de in deze studie geteste condities bleek het wildtype ook γ-butyrolactonen (geproduceerd door BGC16, BGC23 en BGC33), tetramaat macrolactams (geproduceerd door BGC6) en sceliphrolactam (geproduceerd door BGC9) te produceren (Yang *et al.,* 2005, Moree *et al.,* 2014, Low *et al.,* 2018) (**Hoofdstuk 4**). Bovendien werden enkele moleculaire families die geen verband hielden met angucyclines gedetecteerd in sommige van de *lugO* deletie mutanten. Op MM agar werden moleculen gerelateerd aan *N-acyl* glutamine en aminoalcoholen gedetecteerd, met name in *lugOI*, *lugOII* en *lugOIII* deletiemutanten (Battista *et al.,* 2019, Won *et al.,* 2014, Harrison *et al.,* 2018). Op R5 agar werd een nieuwe lipopeptide gedetecteerd in extracten afkomstig van de *lugOI* mutant (mogelijk geproduceerd door BGC21) (**Hoofdstuk 5**). Uit verschillende studies is al gebleken dat beïnvloeding van de productieniveaus van secundaire metabolieten die overvloedig door de gastheer worden geproduceerd, kan leiden tot aanzienlijke wijzigingen in de productie van ongerelateerde metabolieten (Iorio *et al.,* 2021, Culp *et al.,* 2019). Uitschakelen van de BGC's voor de antibiotica streptomycine en streptothricine leidde daarbij tot de productie van niet eerder waargenomen secundaire metabolieten (Culp *et al.,* 2019). Bovendien leidde het blokkeren van verschillende stappen in de pseudouridimycine (PUM)-route in *Streptomyes* sp. ID38640 tot de productie van nieuwe metabolieten van andere metabole routes. Het verstoren van de expressie van het ene gencluster kan dus een ander gencluster aanzetten, wat een mogelijke strategie is om nieuwe antibiotica te ontdekken (Iorio *et al.,* 2021).

Morfologische veranderingen geassocieerd met angucyclineproductie

Productie van angucyclines verstoort de groei en blokkeert de ontwikkeling, waarschijnlijk door cytotoxiciteit als gevolg van hun DNA-afbrekende eigenschappen (Kharel *et al.*, 2012). In overeenstemming daarmee werden grote verschillen waargenomen in morfologische differentiatie tussen de wild-type stam enerzijds en angucycline mutanten (Δ*lug-pks*, Δ*lugRII*, Δ*lugRIV* en Δ*lugRV*) anderzijds, met name op R5 agar (**Hoofdstuk 6**). De stammen die geen angucyclines produceerden ontwikkelden zich goed, terwijl de angucycline-producerende stammen, waaronder het wild-type, geremd werden in hun ontwikkeling, wat resulteerde in een blokkade van de sporulatie. Eerder bleek al dat de DNA-beschadigende prodiginines, geproduceerd door *S. coelicolor,* de sporulatie van de stam op rijke media remmen. De *redD* mutant van *S. coelicolor,* die deze moleculen niet kan produceren, vertoonde een vroegtijdige morfologische differentiatie op R2YE-agarplaten in vergelijking met het wild-type (Tenconi *et al.,* 2020). Prodiginines bevorderen de celdood in het vegetatieve mycelium, waardoor de ontwikkeling van de bacterie wordt uitgesteld totdat er zoveel biomassa is afgebroken dat er voldoende voedingsstoffen zijn om de volgende fase van de levenscyclus te doorlopen (Tenconi & Rigali, 2018). Angucyclines kunnen door hun DNA-beschadigende eigenschappen een soortgelijke rol spelen bij de controle van celdood en ontwikkeling.

Slotopmerkingen en toekomstperspectieven

Deze studie verschaft nieuwe inzichten in de biosynthese en regulatie van angucyclines en de sterk gemodificeerde angucycline, lugdunomycine. Deze nieuwe inzichten kunnen worden toegepast om de productie van lugdunomycine door *Streptomyces* sp. QL37 te verhogen met het doel om grotere hoeveelheden lugdunomycine te isoleren en het werkingsmechanisme ervan en dat van verwante moleculen te bepalen. Het *lug* gencluster bestaat waarschijnlijk uit 28 genen, waaronder de genen voor de minimale PKS, voor vijf oxygenases en vijf regulatoren. Oxygenases LugOIII en LugOV katalyseren waarschijnlijk de cruciale C-ring splitsing, één van de sleutelreacties tijdens de biosynthese van lugdunomycine. Een angucycline-epoxide is mogelijk het substraat voor de C-ring splitsing. Streptomyceten met angucycline BGC's die homologen van *lugOIII* en/of *lugOV* bevatten, produceren waarschijnlijk een diversiteit aan C-ring-gesplitste angucyclines en misschien ook lugdunomycine. De laatste reactie in de lugdunomycine-route vereist waarschijnlijk twee BGC's, één die het dieen

(angucycline) produceert en de andere die het dienofiel (*iso*-maleimycine) van een Diels-Alder reactie produceert. Deze Diels-Alder reactie is waarschijnlijk spontaan, een conclusie die we baseren op het feit dat *Streptomyces* sp. QL37 een racemisch mengsel van lugdunomycine produceert (Wu *et al.,* 2019).

Een belangrijke vraag is of lugdunomycinebiosynthese een evolutionair voordeel biedt voor de producent. Het concept dat moleculen door samenwerking van enzymen uit meerdere BGC's kunnen worden geproduceerd lijkt vaker voor te komen dan tot nu toe werd gedacht. Voorbeelden zijn onder meer de biosynthese van endofenasiden door *Kitasatospora* sp. MBT 66, waarbij glycosylering gebeurt door enzymen van een ander gencluster en bij de biosynthese van actinomycine L door *Streptomyces* sp. MBT 27, waarbij anthranilamide wordt ingebouwd om zo een nieuwe variant van het aloude actinomycine te vormen (Wu *et al.,* 2016b, Machushynets *et al.,* 2022). Het idee dat één molecuul wordt afgeleid van één BGC vormt de basis voor het ontwikkelen van nieuwe antibiotica middels de synthetische biologie. Het voorbeeld van lugdunomycine laat opnieuw zien dat op deze manier zeer interessante chemische diversiteit kan worden gemist en dat het belangrijk is om de chemische diversiteit van de oorspronkelijke productiestam nooit uit het oog te verliezen.

References

REFERENCES

Abrudan, M.I., Smakman, F., Grimbergen, A.J., Westhoff, S., Miller, E.L., *et al.* (2015) Socially mediated induction and suppression of antibiosis during bacterial coexistence. *Proc Natl Acad Sci U S A* **112**: 11054-11059.

Ahn, S.K., Cuthbertson, L., and Nodwell, J.R. (2012) Genome context as a predictive tool for identifying regulatory targets of the TetR family transcriptional regulators. *PLoS One* **7**: e50562.

Aigle, B., Pang, X., Decaris, B., and Leblond, P. (2005) Involvement of AlpV, a new member of the *Streptomyces* antibiotic regulatory protein family, in regulation of the duplicated type II polyketide synthase *alp* gene cluster in *Streptomyces ambofaciens*. *J Bacteriol* **187**: 2491-2500.

Akanuma, G., Hara, H., Ohnishi, Y., and Horinouchi, S. (2009) Dynamic changes in the extracellular proteome caused by absence of a pleiotropic regulator AdpA in *Streptomyces griseus*. *Mol Microbiol* **73**: 898-912.

Alduina, R., Lo Piccolo, L., D'Alia, D., Ferraro, C., Gunnarsson, N., *et al.* (2007) Phosphate-controlled regulator for the biosynthesis of the dalbavancin precursor A40926. *J Bacteriol* **189**: 8120-8129.

Allenby, N.E., Laing, E., Bucca, G., Kierzek, A.M., and Smith, C.P. (2012) Diverse control of metabolism and other cellular processes in *Streptomyces coelicolor* by the PhoP transcription factor: Genome-wide identification of *in vivo* targets. *Nucleic Acids Res* **40**: 9543-9556.

Amos, G.C.A., Awakawa, T., Tuttle, R.N., Letzel, A.C., Kim, M.C., *et al.* (2017) Comparative transcriptomics as a guide to natural product discovery and biosynthetic gene cluster functionality. *Proc Natl Acad Sci U S A* **114**: E11121-E11130.

Anborgh, P.H., and Parmeggiani, A. (1991) New antibiotic that acts specifically on the GTP-bound form of elongation factor Tu. *EMBO J* **10**: 779-784.

Andersson, D.I., and Hughes, D. (2014) Microbiological effects of sublethal levels of antibiotics. *Nat Rev Microbiol* **12**: 465-478.

Angell, S., Lewis, C.G., Buttner, M.J., and Bibb, M.J. (1994) Glucose repression in *Streptomyces coelicolor* A3(2): A likely regulatory role for glucose kinase. *Mol Gen Genet* **244**: 135-143.

Angell, S., Schwarz, E., and Bibb, M.J. (1992) The glucose kinase gene of *Streptomyces coelicolor* A3(2): Its nucleotide sequence, transcriptional analysis and role in glucose repression. *Mol Microbiol* **6**: 2833-2844.

Antoraz, S., Rico, S., Rodriguez, H., Sevillano, L., Alzate, J.F., *et al.* (2017) The orphan response regulator Aor1 is a new relevant piece in the complex puzzle of *Streptomyces coelicolor* antibiotic regulatory network. *Front Microbiol* **8**: 2444.

Arias, A.A., Lambert, S., Martinet, L., Adam, D., Tenconi, E., *et al.* (2015) Growth of desferrioxamine-deficient *Streptomyces* mutants through xenosiderophore piracy of airborne fungal contaminations. *FEMS Microbiol Ecol* **91**: fiv080-fiv080.

Arnison, P.G., Bibb, M.J., Bierbaum, G., Bowers, A.A., Bugni, T.S., *et al.* (2013) Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Nat Prod Rep* **30**: 108-160.

Aroonsri, A., Kitani, S., Choi, S.U., and Nihira, T. (2008) Isolation and characterization of *bamA* genes, homologues of the gamma-butyrolactone autoregulator-receptor gene in *Amycolatopsis mediterranei*, a rifamycin producer. *Biotechnol Lett* **30**: 2019-2024.

Aroonsri, A., Kitani, S., Hashimoto, J., Kosone, I., Izumikawa, M., *et al.* (2012) Pleiotropic control of secondary metabolism and morphological development by KsbC, a butyrolactone autoregulator receptor homologue in *Kitasatospora setae*. *Appl Environ Microbiol* **78**: 8015-8024.

Arthur, M., Reynolds, P.E., Depardieu, F., Evers, S., Dutka-Malen, S., *et al.* (1996) Mechanisms of glycopeptide resistance in enterococci. *J Infect* **32**: 11-16.

Asnicar, F., Thomas, A.M., Beghini, F., Mengoni, C., Manara, S., *et al.* (2020) Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. *Nat Commun* **11**: 2500.

Autret, S., Nair, R., and Errington, J. (2001) Genetic analysis of the chromosome segregation protein Spo0J of *Bacillus subtilis*: Evidence for separate domains involved in DNA binding and interactions with Soj protein. *Mol Microbiol* **41**: 743-755.

Bai, C., Zhang, Y., Zhao, X., Hu, Y., Xiang, S., *et al.* (2015) Exploiting a precise design of universal synthetic modular regulatory elements to unlock the microbial natural products in *Streptomyces*. *Proc Natl Acad Sci U S A* **112**: 12181-12186.

Baltz, R.H. (2008) Renaissance in antibacterial discovery from actinomycetes. *Curr Opin Pharmacol* **8**: 557-563.

Baltz, R.H. (2018) Correction to: Synthetic biology, genome mining, and combinatorial biosynthesis of nrpsderived antibiotics: A perspective. *J Ind Microbiol Biotechnol* **45**: 651-655.

Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., *et al.* (2016) Taxonomy, physiology, and natural products of *Actinobacteria*. *Microbiol Mol Biol Rev* **80**: 1-43.

Barna, J.C., and Williams, D.H. (1984) The structure and mode of action of glycopeptide antibiotics of the vancomycin group. *Annu Rev Microbiol* **38**: 339-357.

Bartholomae, M., Buivydas, A., Viel, J.H., Montalban-Lopez, M., and Kuipers, O.P. (2017) Major gene-regulatory mechanisms operating in ribosomally synthesized and post-translationally modified peptide (RiPP) biosynthesis. *Mol Microbiol* **106**: 186-206.

Battista, N., Bari, M., and Bisogno, T. (2019) N-acyl amino acids: Metabolism, molecular targets, and role in biological processes. *Biomolecules* **9**: 822.

Bayles, K.W. (2014) Bacterial programmed cell death: Making sense of a paradox. *Nat Rev Microbiol* **12**: 63-69.

Beam, M.P., Bosserman, M.A., Noinaj, N., Wehenkel, M., and Rohr, J. (2009) Crystal structure of baeyer-villiger monooxygenase MtmOIV, the key enzyme of the mithramycin biosynthetic pathway. *Biochemistry* **48**: 4476-4487.

Beniddir, M.A., Kang, K.B., Genta-Jouve, G., Huber, F., Rogers, S., *et al.* (2021) Advances in decomposing complex metabolite mixtures using substructure- and network-based computational metabolomics approaches. *Nat Prod Rep* **38**: 1967-1993.

Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., *et al.* (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**: 141-147.

Berdy, J. (2005) Bioactive microbial metabolites. *J Antibiot (Tokyo)* **58**: 1-26.

Bermudez, O., Padilla, P., Huitron, C., and Flores, M.E. (1998) Influence of carbon and nitrogen source on synthesis of NADP(+)-isocitrate dehydrogenase, methylmalonyl-coenzyme A mutase, and methylmalonyl-coenzyme A decarboxylase in *Saccharopolyspora erythraea* CA340. *FEMS Microbiol Lett* **164**: 77-82.

Bhatnagar, R.K., Doull, J.L., and Vining, L.C. (1988) Role of the carbon source in regulating chloramphenicol production by *Streptomyces venezuelae*: Studies in batch and continuous cultures. *Can J Microbiol* **34**: 1217-1223.

Bibb, M.J. (2005) Regulation of secondary metabolism in streptomycetes. *Curr Opin Microbiol* **8**: 208-215.

Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Rao, R.N., *et al.* (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* **116**: 43-49.

Bignell, D.R., Seipke, R.F., Huguet-Tapia, J.C., Chambers, A.H., Parry, R.J., *et al.* (2010) *Streptomyces scabies* 87-22 contains a coronafacic acid-like biosynthetic cluster that contributes to plant-microbe interactions. *Mol Plant-Microbe Interact* **23**: 161-175.

Bingol, K., Bruschweiler-Li, L., Li, D., Zhang, B., Xie, M., *et al.* (2016) Emerging new strategies for successful metabolite identification in metabolomics. *Bioanalysis* **8**: 557-573.

Blin, K., Shaw, S., Kloosterman, A.M., Charlop-Powers, Z., van Wezel, G.P., *et al.* (2021) antiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res* **49**: W29-W35.

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., *et al.* (2019) antiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* **47**: W81-W87.

Bo, S.T., Xu, Z.F., Yang, L., Cheng, P., Tan, R.X., *et al.* (2018) Structure and biosynthesis of mayamycin b, a new polyketide with antibacterial activity from *Streptomyces* sp. 120454. *J Antibiot (Tokyo)* **71**: 601-605.

Book, A.J., Lewin, G.R., McDonald, B.R., Takasuka, T.E., Wendt-Pienkowski, E., *et al.* (2016) Evolution of high cellulolytic activity in symbiotic *Streptomyces* through selection of expanded gene content and coordinated gene expression. *PLoS Biol* **14**: e1002475.

Brian, P., Riggle, P.J., Santos, R.A., and Champness, W.C. (1996) Global negative regulation of *Streptomyces coelicolor* antibiotic synthesis mediated by an *absA*-encoded putative signal transduction system. *J Bacteriol* **178**: 3221-3231.

Brown, E.D., and Wright, G.D. (2016) Antibacterial drug discovery in the resistance era. *Nature* **529**: 336-343.

Brückner, R., and Titgemeyer, F. (2002) Carbon catabolite repression in bacteria: Choice of the carbon source and autoregulatory limitation of sugar utilization. *FEMS Microbiol Lett* **209**: 141-148.

Bugg, T.D., Wright, G.D., Dutka-Malen, S., Arthur, M., Courvalin, P., *et al.* (1991) Molecular basis for vancomycin resistance in *Enterococcus faecium* BM4147: Biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins vanh and vana. *Biochemistry* **30**: 10408-10415.

Bunet, R., Riclea, R., Laureti, L., Hotel, L., Paris, C., *et al.* (2014) A single Sfp-type phosphopantetheinyl transferase plays a major role in the biosynthesis of PKS and NRPS derived metabolites in *Streptomyces ambofaciens* ATCC 23877. *PLoS One* **9**: e87607.

Buttner, M.J., Schafer, M., Lawson, D.M., and Maxwell, A. (2018) Structural insights into simocyclinone as an antibiotic, effector ligand and substrate. *FEMS Microbiol Rev* **42**: fux055.

Caffrey, P., Aparicio, J.F., Malpartida, F., and Zotchev, S.B. (2008) Biosynthetic engineering of polyene macrolides towards generation of improved antifungal and antiparasitic agents. *Curr Topics Med Chem* **8**: 639-653.

Cao, G., Zhong, C., Zong, G., Fu, J., Liu, Z., *et al.* (2016) Complete genome sequence of *Streptomyces clavuligerus* F613-1, an industrial producer of clavulanic acid. *Genome Announc* **4**: 1020-1021.

Cao, M., Zheng, C., Yang, D., Kalkreuter, E., Adhikari, A., *et al.* (2021) Cryptic sulfur incorporation in thioangucycline biosynthesis. *Angew Chem Int Ed Engl* **60**: 7140-7147.

Carmody, M., Byrne, B., Murphy, B., Breen, C., Lynch, S., *et al.* (2004) Analysis and manipulation of amphotericin biosynthetic genes by means of modified phage KC515 transduction techniques. *Gene* **343**: 107-115.

Castiglione, F., Cavaletti, L., Losi, D., Lazzarini, A., Carrano, L.*, et al.* (2007) A novel lantibiotic acting on bacterial cell wall synthesis produced by the uncommon actinomycete *Planomonospora* sp. *Biochemistry* **46**: 5884-5895.

Castiglione, F., Lazzarini, A., Carrano, L., Corti, E., Ciciliato, I.*, et al.* (2008) Determining the structure and mode of action of microbisporicin, a potent lantibiotic active against multiresistant pathogens. *Chem Biol* **15**: 22-31.

Cen, X.F., Wang, J.Z., Zhao, G.P., Wang, Y., and Wang, J. (2016) Molecular evidence for the coordination of nitrogen and carbon metabolisms, revealed by a study on the transcriptional regulation of the *agl3EFG* operon that encodes a putative carbohydrate transporter in *Streptomyces coelicolor*. *Biochem Biophys Res Commun* **471**: 510-514.

Ceniceros, A., Dijkhuizen, L., and Petrusma, M. (2017) Molecular characterization of a *Rhodococcus jostii* RHA1 gamma-butyrolactone(-like) signalling molecule and its main biosynthesis gene *gblA*. *Sci Rep* **7**: 17743.

Chakraburtty, R., and Bibb, M. (1997) The ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3(2) plays a conditional role in antibiotic production and morphological differentiation. *J Bacteriol* **179**: 5854-5861.

Challis, G.L., and Hopwood, D.A. (2003) Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc Natl Acad Sci U S A* **100**: 14555-14561.

Champness, W., Riggle, P., Adamidis, T., and Vandervere, P. (1992) Identification of *Streptomyces coelicolor* genes involved in regulation of antibiotic synthesis. *Gene* **115**: 55-60.

Chandra, G., and Chater, K.F. (2008) Evolutionary flux of potentially *bldA*-dependent *Streptomyces* genes containing the rare leucine codon TTA. *Antonie Van Leeuwenhoek* **94**: 111-126.

Chater, K.F. (1972) A morphological and genetic mapping study of white colony mutants of *Streptomyces coelicolor*. *J Gen Microbiol* **72**: 9-28.

Chater, K.F., (2011) Differentiation in *Streptomyces*: The properties and programming of diverse cell-types. In: *Streptomyces*: Molecular biology and biotechnology. P. Dyson (ed). Norfolk, UK: Caister Academic Press, pp. 43-86.

Chater, K.F., Biro, S., Lee, K.J., Palmer, T., and Schrempf, H. (2010) The complex extracellular biology of *Streptomyces*. *FEMS Microbiol Rev* **34**: 171-198.

Chater, K.F., and Chandra, G. (2006) The evolution of development in *Streptomyces* analysed by genome comparisons. *FEMS Microbiol Rev* **30**: 651-672.

Chater, K.F., and Losick, R., (1997) Mycelial life style of *Streptomyces coelicolor* A3(2) and its relatives. In: Bacteria as multicellular organisms. J.A. Shapiro & M. Dworkin (eds). New York: Oxford University Press, pp. 149-182.

Chavali, A.K., and Rhee, S.Y. (2017) Bioinformatics tools for the identification of gene clusters that biosynthesize specialized metabolites. *Briefings in bioinformatics*: 1022–1034.

Chavez, A., Forero, A., Sanchez, M., Rodriguez-Sanoja, R., Mendoza-Hernandez, G.*, et al.* (2011) Interaction of SCO2127 with BldKB and its possible connection to carbon catabolite regulation of morphological differentiation in *Streptomyces coelicolor*. *Appl Microbiol Biotechnol* **89**: 799-806.

Chen, Q., Mulzer, M., Shi, P., Beuning, P.J., Coates, G.W.*, et al.* (2011) *De novo* asymmetric synthesis of fridamycin E. *Org Lett* **13**: 6592-6595.

Chen, Y.H., Wang, C.C., Greenwell, L., Rix, U., Hoffmeister, D.*, et al.* (2005) Functional analyses of oxygenases in jadomycin biosynthesis and identification of JadH as a bifunctional oxygenase/dehydrase. *J Biol Chem* **280**: 22508-22514.

Chevrette, M.G., Aicheler, F., Kohlbacher, O., Currie, C.R., and Medema, M.H. (2017) SANDPUMA: Ensemble predictions of nonribosomal peptide chemistry reveal biosynthetic diversity across Actinobacteria. *Bioinformatics* **33**: 3202-3210.

Choi, S.U., Lee, C.K., Hwang, Y.I., Kinoshita, H., and Nihira, T. (2004) Cloning and functional analysis by gene disruption of a gene encoding a gamma-butyrolactone autoregulator receptor from *Kitasatospora setae*. *J Bacteriol* **186**: 3423-3430.

Choi, S.U., Lee, C.K., Hwang, Y.I., Kinosita, H., and Nihira, T. (2003) Gamma-butyrolactone autoregulators and receptor proteins in non- *Streptomyces* actinomycetes producing commercially important secondary metabolites. *Arch Microbiol* **180**: 303-307.

Chong, J., Wishart, D.S., and Xia, J. (2019) Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis. *Curr Protoc Bioinformatics* **68**: e86.

Chouayekh, H., Nothaft, H., Delaunay, S., Linder, M., Payrastre, B.*, et al.* (2007) Phosphoinositides are involved in control of the glucose-dependent growth resumption that follows the transition phase in *Streptomyces lividans*. *J Bacteriol* **189**: 741-749.

Chouayekh, H., and Virolle, M.J. (2002) The polyphosphate kinase plays a negative role in the control of antibiotic production in *Streptomyces lividans*. *Mol Microbiol* **43**: 919-930.

Claessen, D., Rozen, D.E., Kuipers, O.P., Sogaard-Andersen, L., and van Wezel, G.P. (2014) Bacterial solutions to multicellularity: A tale of biofilms, filaments and fruiting bodies. *Nat Rev Microbiol* **12**: 115-124.

Colson, S., Stephan, J., Hertrich, T., Saito, A., van Wezel, G.P.*, et al.* (2007) Conserved *cis*-acting elements upstream of genes composing the chitinolytic system of streptomycetes are DasR-responsive elements. *J Mol Microbiol Biotechnol* **12**: 60-66.

Colson, S., van Wezel, G.P., Craig, M., Noens, E.E., Nothaft, H.*, et al.* (2008) The chitobiose-binding protein, dasa, acts as a link between chitin utilization and morphogenesis in *Streptomyces coelicolor*. *Microbiology* **154**: 373-382.

Craig, M., Lambert, S., Jourdan, S., Tenconi, E., Colson, S.*, et al.* (2012) Unsuspected control of siderophore production by N-acetylglucosamine in streptomycetes. *Environ Microbiol Rep* **4**: 512-521.

Craney, A., Ozimok, C., Pimentel-Elardo, S.M., Capretta, A., and Nodwell, J.R. (2012) Chemical perturbation of secondary metabolism demonstrates important links to primary metabolism. *Chem Biol* **19**: 1020-1027.

Cruz-Morales, P., Vijgenboom, E., Iruegas-Bocardo, F., Girard, G., Yanez-Guerra, L.A.*, et al.* (2013) The genome sequence of *Streptomyces lividans* 66 reveals a novel tRNA-dependent peptide biosynthetic system within a metal-related genomic island. *Genome Biol Evol* **5**: 1165-1175.

Culp, E.J., Yim, G., Waglechner, N., Wang, W., Pawlowski, A.C.*, et al.* (2019) Hidden antibiotics in actinomycetes can be identified by inactivation of gene clusters for common antibiotics. *Nat Biotechnol* **37**: 1149-1154.

Cuthbertson, L., and Nodwell, J.R. (2013) The TetR family of regulators. *Microbiol Mol Biol Rev* **77**: 440-475.

D'Alia, D., Eggle, D., Nieselt, K., Hu, W.S., Breitling, R.*, et al.* (2011) Deletion of the signalling molecule synthase scba has pleiotropic effects on secondary metabolite biosynthesis, morphological differentiation and primary metabolism in *Streptomyces coelicolor* A3(2). *Microbial Biotechnol* **4**: 239-251.

Daigle, F., Lerat, S., Bucca, G., Sanssouci, E., Smith, C.P.*, et al.* (2015) A *terD* domain-encoding gene (SCO2368) is involved in calcium homeostasis and participates in calcium regulation of a DosR-like regulon in *Streptomyces coelicolor*. *J Bacteriol* **197**: 913-923.

Dairi, T., Hamano, Y., Furumai, T., and Oki, T. (1999) Development of a self-cloning system for *Actinomadura verrucosospora* and identification of polyketide synthase genes essential for production of the angucyclic antibiotic pradimicin. *Appl Environ Microbiol* **65**: 2703-2709.

Davies, J., Spiegelman, G.B., and Yim, G. (2006) The world of subinhibitory antibiotic concentrations. *Curr Opin Microbiol* **9**: 445-453.

de Jong, A., van Heel, A.J., Kok, J., and Kuipers, O.P. (2010) BAGEL2: Mining for bacteriocins in genomic data. *Nucleic Acids Res* **38**: W647-651.

de Jong, A., van Hijum, S.A., Bijlsma, J.J., Kok, J., and Kuipers, O.P. (2006) BAGEL: A web-based bacteriocin genome mining tool. *Nucleic Acids Res* **34**: W273-279.

Demain, A.L. (2014) Importance of microbial natural products and the need to revitalize their discovery. *J Ind Microbiol Biotechnol* **41**: 185-201.

Demain, A.L., and Inamine, E. (1970) Biochemistry and regulation of streptomycin and mannosidostreptomycinase (alpha-D-mannosidase) formation. *Bacteriol Rev* **34**: 1-19.

den Hengst, C.D., Tran, N.T., Bibb, M.J., Chandra, G., Leskiw, B.K., *et al.* (2010) Genes essential for morphological development and antibiotic production in *Streptomyces coelicolor* are targets of bldd during vegetative growth. *Mol Microbiol* **78**: 361-379.

Derouaux, A., Halici, S., Nothaft, H., Neutelings, T., Moutzourelis, G., *et al.* (2004) Deletion of a cyclic AMP receptor protein homologue diminishes germination and affects morphological development of *Streptomyces coelicolor*. *J Bacteriol* **186**: 1893-1897.

Deutscher, J., Francke, C., and Postma, P.W. (2006) How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev* **70**: 939-1031.

Diaz, M., Esteban, A., Fernandez-Abalos, J.M., and Santamaria, R.I. (2005) The high-affinity phosphate-binding protein PstS is accumulated under high fructose concentrations and mutation of the corresponding gene affects differentiation in *Streptomyces lividans*. *Microbiology* **151**: 2583-2592.

Distler, U., Kuharev, J., Navarro, P., Levin, Y., Schild, H., *et al.* (2014) Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nature Methods* **11**: 167-170.

Donadio, S., Sosio, M., Stegmann, E., Weber, T., and Wohlleben, W. (2005) Comparative analysis and insights into the evolution of gene clusters for glycopeptide antibiotic biosynthesis. *Mol Genet Genomics* **274**: 40-50.

Doull, J.L., and Vining, L.C. (1990) Physiology of antibiotic production in actinomycetes and some underlying control mechanisms. *Biotechnol Adv* **8**: 141-158.

Du, C., and van Wezel, G.P. (2018) Mining for microbial gems: Integrating proteomics in the postgenomic natural product discovery pipeline. *Proteomics* **18**: e1700332.

Dun, J., Zhao, Y., Zheng, G., Zhu, H., Ruan, L., *et al.* (2015) PapR6, a putative atypical response regulator, functions as a pathway-specific activator of pristinamycin II biosynthesis in *Streptomyces pristinaespiralis*. *J Bacteriol* **197**: 441-450.

Eccleston, M., Ali, R.A., Seyler, R., Westpheling, J., and Nodwell, J. (2002) Structural and genetic analysis of the BldB protein of *Streptomyces coelicolor*. *J Bacteriol* **184**: 4270-4276.

Elliot, M., Damji, F., Passantino, R., Chater, K., and Leskiw, B. (1998) The *bldD* gene of *Streptomyces coelicolor* A3(2): A regulatory gene involved in morphogenesis and antibiotic production. *J Bacteriol* **180**: 1549-1555.

Elstner, E.F., Carnes, D.M., Suhadolnik, R.J., Kreishman, G.P., Schweizer, M.P., *et al.* (1973) Isolation, structural elucidation, biological properties, and biosynthesis of maleimycin, a new bicyclic maleimide antibiotic isolated from the culture filtrates of *Streptomyces showdoensis*. *Biochemistry* **12**: 4992-4997.

Escalante, L., Lopez, H., Mateos, R.D., Lara, F., and Sanchez, S. (1982) Transient repression of erythromycin formation in *Streptomyces erythraeus*. *J Gen Microbiol* **128**: 2011-2015.
Fabret, C., Feher, V.A., and Hoch, J.A. (1999) Two-component signal transduction in *Bacillus subtilis*: How one organism sees its world. *J Bacteriol* **181**: 1975-1983.
Fakhruzzaman, M., Inukai, Y., Yanagida, Y., Kino, H., Igarashi, M*., et al.* (2015) Study on *in vivo* effects of bacterial histidine kinase inhibitor, waldiomycin, in *Bacillus subtilis* and *Staphylococcus aureus*. *J Gen Appl Microbiol* **61**: 177-184.
Fan, K., Pan, G., Peng, X., Zheng, J., Gao, W*., et al.* (2012a) Identification of JadG as the B ring opening oxygenase catalyzing the oxidative C-C bond cleavage reaction in jadomycin biosynthesis. *Chem Biol* **19**: 1381-1390.
Fan, K., and Zhang, Q. (2018) The functional differentiation of the post-PKS tailoring oxygenases contributed to the chemical diversities of atypical angucyclines. *Synth Syst Biotechnol* **3**: 275-282.
Fan, K., Zhang, X., Liu, H., Han, H., Luo, Y*., et al.* (2012b) Evaluation of the cytotoxic activity of new jadomycin derivatives reveals the potential to improve its selectivity against tumor cells. *J Antibiot (Tokyo)* **65**: 449-452.
Fedoryshyn, M., Welle, E., Bechthold, A., and Luzhetskyy, A. (2008) Functional expression of the Cre recombinase in actinomycetes. *Appl Microbiol Biotechnol* **78**: 1065-1070.
Feitelson, J.S., Malpartida, F., and Hopwood, D.A. (1985) Genetic and biochemical characterization of the *red* gene cluster of *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **131**: 2431-2441.
Feng, W.H., Mao, X.M., Liu, Z.H., and Li, Y.Q. (2011) The ECF sigma factor SigT regulates actinorhodin production in response to nitrogen stress in *Streptomyces coelicolor*. *Appl Microbiol Biotechnol* **92**: 1009-1021.
Fenical, W., and Jensen, P.R. (2006) Developing a new resource for drug discovery: Marine actinomycete bacteria. *Nat Chem Biol* **2**: 666-673.
Fernandez-Martinez, L.T., Gomez-Escribano, J.P., and Bibb, M.J. (2015) A *relA*-dependent regulatory cascade for auto-induction of microbisporicin production in *Microbispora corallina*. *Mol Microbiol* **97**: 502-514.
Fernandez-Martinez, L.T., Santos-Beneit, F., and Martin, J.F. (2012) Is PhoR-PhoP partner fidelity strict? PhoR is required for the activation of the *pho* regulon in *Streptomyces coelicolor*. *Mol Gen Genet* **287**: 565-573.
Fillenberg, S.B., Friess, M.D., Korner, S., Bockmann, R.A., and Muller, Y.A. (2016) Crystal structures of the global regulator DasR from *Streptomyces coelicolor*: Implications for the allosteric regulation of GntR/HutC repressors. *PLoS One* **11**: e0157691.
Fillenberg, S.B., Grau, F.C., Seidel, G., and Muller, Y.A. (2015) Structural insight into operator *dre*-sites recognition and effector binding in the GntR/HutC transcription regulator NagR. *Nucleic Acids Res* **43**: 1283-1296.
Finney, L.A., and O'Halloran, T.V. (2003) Transition metal speciation in the cell: Insights from the chemistry of metal ion receptors. *Science* **300**: 931-936.
Flärdh, K., and Buttner, M.J. (2009) *Streptomyces* morphogenetics: Dissecting differentiation in a filamentous bacterium. *Nat Rev Microbiol* **7**: 36-49.
Flett, F., Mersinias, V., and Smith, C.P. (1997) High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol Lett* **155**: 223-229.
Flinspach, K., Kapitzke, C., Tocchetti, A., Sosio, M., and Apel, A.K. (2014) Heterologous expression of the thiopeptide antibiotic GE2270 from *Planobispora rosea* ATCC 53733 in *Streptomyces coelicolor* requires deletion of ribosomal genes from the expression construct. *PLoS One* **9**: e90499.

Flores, F.J., Barreiro, C., Coque, J.J.R., and Martín, J.F. (2005) Functional analysis of two divalent metal-dependent regulatory genes *dmdR1* and *dmdR2* in *Streptomyces coelicolor* and proteome changes in deletion mutants. *FEBS J* **272**: 725-735.

Flores, F.J., and Martin, J.F. (2004) Iron-regulatory proteins DmdR1 and DmdR2 of *Streptomyces coelicolor* form two different DNA–protein complexes with iron boxes. *Biochem J* **380**: 197-503.

Fotso, S., Mahmud, T., Zabriskie, T.M., Santosa, D.A., and Proteau, P.J. (2008) Rearranged and unrearranged angucyclinones from Indonesian *Streptomyces* spp. *J Antibiot (Tokyo)* **61**: 449-456.

Foulston, L., and Bibb, M. (2011) Feed-forward regulation of microbisporicin biosynthesis in *Microbispora corallina*. *J Bacteriol* **193**: 3064-3071.

Foulston, L.C., and Bibb, M.J. (2010) Microbisporicin gene cluster reveals unusual features of lantibiotic biosynthesis in actinomycetes. *Proc Natl Acad Sci U S A* **107**: 13461-13466.

Fowler-Goldsworthy, K., Gust, B., Mouz, S., Chandra, G., Findlay, K.C*., et al.* (2011) The *Actinobacteria*-specific gene *wblA* controls major developmental transitions in *Streptomyces coelicolor* A3(2). *Microbiology* **157**: 1312-1328.

Freel, K.C., Edlund, A., and Jensen, P.R. (2012) Microdiversity and evidence for high dispersal rates in the marine actinomycete '*Salinispora pacifica*'. *Environ Microbiol* **14**: 480-493.

Fujii, T., Gramajo, H.C., Takano, E., and Bibb, M.J. (1996) *redD* and *act*II-ORF4, pathway-specific regulatory genes for antibiotic production in *Streptomyces coelicolor* A3(2), are transcribed *in vitro* by an RNA polymerase holoenzyme containing sigma *hrdD*. *J Bacteriol* **178**: 3402-3405.

Fürst, M.J.L.J., et al. (2019) Baeyer–villiger monooxygenases: Tunable oxidative biocatalysts. *ACS Catalysis* **9** 11207–11241.

Furuya, K., and Hutchinson, C.R. (1996) The DnrN protein of *Streptomyces peucetius*, a pseudo-response regulator, is a DNA-binding protein involved in the regulation of daunorubicin biosynthesis. *J Bacteriol* **178**: 6310-6318.

Gagnat, J., Chouayekh, H., Gerbaud, C., Francou, F., and Virolle, M.J. (1999) Disruption of *sblA* in *Streptomyces lividans* permits expression of a heterologous alpha-amylase gene in the presence of glucose. *Microbiology* **145** 2303-2312.

Galet, J., Deveau, A., Hotel, L., Frey-Klett, P., Leblond, P*., et al.* (2015) *Pseudomonas fluorescens* pirates both ferrioxamine and ferricoelichelin siderophores from *Streptomyces ambofaciens*. *Appl Environ Microbiol* **81**: 3132-3141.

Gallo, G., Renzone, G., Palazzotto, E., Monciardini, P., Arena, S*., et al.* (2016) Elucidating the molecular physiology of lantibiotic NAI-107 production in *Microbispora* ATCC-pta-5024. *BMC Genomics* **17**: 42.

Gao, C., Hindra, Mulder, D., Yin, C., and Elliot, M.A. (2012) Crp is a global regulator of antibiotic production in *Streptomyces*. *MBio* **3**: 00407-00412.

Garg, R.P., and Parry, R.J. (2010) Regulation of valanimycin biosynthesis in *Streptomyces viridifaciens*: Characterization of VlmI as a *Streptomyces* antibiotic regulatory protein (SARP). *Microbiology (Reading)* **156**: 472-483.

Gaudêncio, S.P., and Pereira, F. (2015) Dereplication: Racing to speed up the natural products discovery process. *Nat Prod Rep* **32**: 779-810.

Gaur, N.K., Oppenheim, J., and Smith, I. (1991) The *Bacillus subtilis sin* gene, a regulator of alternate developmental processes, codes for a DNA-binding protein. *J Bacteriol* **173**: 678-686.

Genilloud, O., Gonzalez, I., Salazar, O., Martin, J., Tormo, J.R*., et al.* (2011) Current approaches to exploit actinomycetes as a source of novel natural products. *J Ind Microbiol Biotechnol* **38**: 375-389.

Ghorbel, S., Smirnov, A., Chouayekh, H., Sperandio, B., Esnault, C., *et al.* (2006) Regulation of *ppk* expression and *in vivo* function of Ppk in *Streptomyces lividans* TK24. *J Bacteriol* **188**: 6269-6276.
Gilchrist, C.L.M., and Chooi, Y.H. (2021) Clinker & clustermap.Js: Automatic generation of gene cluster comparison figures. *Bioinformatics* **37**: 2473–2475.
Girard, G., Traag, B.A., Sangal, V., Mascini, N., Hoskisson, P.A., *et al.* (2013) A novel taxonomic marker that discriminates between morphologically complex actinomycetes. *Open Biol* **3**: 130073.
Goerke, B., and Stulke, J. (2008) Carbon catabolite repression in bacteria: Many ways to make the most out of nutrients. *Nat Rev Microbiol* **6**: 613-624.
Gomez-Escribano, J.P., and Bibb, M.J. (2011) Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* **4**: 207-215.
Gomez-Escribano, J.P., Song, L., Fox, D.J., Yeo, V., Bibb, M.J., *et al.* (2012) Structure and biosynthesis of the unusual polyketide alkaloid coelimycin P1, a metabolic product of the *cpk* gene cluster of *Streptomyces coelicolor* M145. *Chem Sci* **3**: 2716-2720.
Gorke, B., and Stülke, J. (2008) Carbon catabolite repression in bacteria: Many ways to make the most out of nutrients. *Nat Rev Microbiol* **6**: 613-624.
Gottelt, M., Kol, S., Gomez-Escribano, J.P., Bibb, M., and Takano, E. (2010) Deletion of a regulatory gene within the *cpk* gene cluster reveals novel antibacterial activity in *Streptomyces coelicolor* A3(2). *Microbiology* **156**: 2343-2353.
Grace, S.C.a.H., D.A. (2016) Processing and visualization of metabolomics data using r *Itechopen*.
Gramajo, H.C., Takano, E., and Bibb, M.J. (1993) Stationary-phase production of the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2) is transcriptionally regulated. *Mol Microbiol* **7**: 837-845.
Gross, H. (2009) Genomic mining-a concept for the discovery of new bioactive natural products. *Curr Opin Drug Disc Dev* **12**: 207-219.
Gubbens, J., Janus, M., Florea, B.I., Overkleeft, H.S., and van Wezel, G.P. (2012) Identification of glucose kinase dependent and independent pathways for carbon control of primary metabolism, development and antibiotic production in *Streptomyces coelicolor* by quantitative proteomics. *Mol Microbiol* **86**: 1490-1507.
Gubbens, J., Zhu, H., Girard, G., Song, L., Florea, B.I., *et al.* (2014) Natural product proteomining, a quantitative proteomics platform, allows rapid discovery of biosynthetic gene clusters for different classes of natural products. *Chem Biol* **21**: 707-718.
Guerinot, M.L. (1994) Microbial iron transport. *Annu Rev Microbiol* **48**: 743-772.
Gullberg, E., Albrecht, L.M., Karlsson, C., Sandegren, L., and Andersson, D.I. (2014) Selection of a multidrug resistance plasmid by sublethal levels of antibiotics and heavy metals. *MBio* **5**: e01918-01914.
Gunnewijk, M.G., van den Bogaard, P.T., Veenhoff, L.M., Heuberger, E.H., de Vos, W.M., *et al.* (2001) Hierarchical control versus autoregulation of carbohydrate utilization in bacteria. *J Mol Microbiol Biotechnol* **3**: 401-413.
Guo, F., Xiang, S., Li, L., Wang, B., Rajasarkka, J., *et al.* (2015) Targeted activation of silent natural product biosynthesis pathways by reporter-guided mutant selection. *Metab Eng* **28**: 134-142.
Guo, J., Zhao, J., Li, L., Chen, Z., Wen, Y., *et al.* (2010) The pathway-specific regulator AveR from *Streptomyces avermitilis* positively regulates avermectin production while it negatively affects oligomycin biosynthesis. *Mol Genet Genomics* **283**: 123-133.
Guo, L., Zhang, L., Yang, Q., Xu, B., Fu, X., *et al.* (2020) Antibacterial and cytotoxic bridged and ring cleavage angucyclinones from a marine *Streptomyces* sp. *Front Chem* **8**: 586.

Guthrie, E.P., Flaxman, C.S., White, J., Hodgson, D.A., Bibb, M.J., *et al.* (1998) A response-regulator-like activator of antibiotic synthesis from *Streptomyces coelicolor* A3(2) with an amino-terminal domain that lacks a phosphorylation pocket. *Microbiology* **144**: 727-738.

Guzman, S., Carmona, A., Escalante, L., Imriskova, I., Lopez, R., *et al.* (2005) Pleiotropic effect of the SCO2127 gene on the glucose uptake, glucose kinase activity and carbon catabolite repression in *Streptomyces peucetius* var. *Caesius*. *Microbiology* **151**: 1717-1723.

Hackl, S., and Bechthold, A. (2015) The gene *bldA*, a regulator of morphological differentiation and antibiotic production in *Streptomyces*. *Arch Pharm (Weinheim)* **348**: 455-462.

Harrison, P.J., Dunn, T.M., and Campopiano, D.J. (2018) Sphingolipid biosynthesis in man and microbes. *Nat Prod Rep* **35**: 921-954.

Harvey, A.L., Edrada-Ebel, R., and Quinn, R.J. (2015) The re-emergence of natural products for drug discovery in the genomics era. *Nat Rev Drug Discov* **14**: 111-129.

Hasebe, F., Matsuda, K., Shiraishi, T., Futamura, Y., Nakano, T., *et al.* (2016) Amino-group carrier-protein-mediated secondary metabolite biosynthesis in *Streptomyces*. *Nat Chem Biol* **12**: 967-972.

He, J.M., Zhu, H., Zheng, G.S., Liu, P.P., Wang, J., *et al.* (2016) Direct involvement of the master nitrogen metabolism regulator GlnR in antibiotic biosynthesis in *Streptomyces*. *J Biol Chem* **291**: 26443-26454.

Helfrich, E.J.N., Ueoka, R., Dolev, A., Rust, M., Meoded, R.A., *et al.* (2019) Automated structure prediction of trans-acyltransferase polyketide synthase products. *Nat Chem Biol* **15**: 813-821.

Helmann, J.D. (2002) The extracytoplasmic function (ECF) sigma factors. *Adv Microb Physiol* **46**: 47-110.

Hertweck, C., Luzhetskyy, A., Rebets, Y., and Bechthold, A. (2007) Type II polyketide synthases: Gaining a deeper insight into enzymatic teamwork. *Nat Prod Rep* **24**: 162-190.

Hesketh, A., Kock, H., Mootien, S., and Bibb, M. (2009) The role of *absC*, a novel regulatory gene for secondary metabolism, in zinc-dependent antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **74**: 1427-1444.

Hiard, S., Maree, R., Colson, S., Hoskisson, P.A., Titgemeyer, F., *et al.* (2007) Predetector: A new tool to identify regulatory elements in bacterial genomes. *Biochem Biophys Res Commun* **357**: 861-864.

Hindle, Z., and Smith, C.P. (1994) Substrate induction and catabolite repression of the *Streptomyces coelicolor* glycerol operon are mediated through the GylR protein. *Mol Microbiol* **12**: 737-745.

Hirano, S., Tanaka, K., Ohnishi, Y., and Horinouchi, S. (2008) Conditionally positive effect of the TetR-family transcriptional regulator AtrA on streptomycin production by *Streptomyces griseus*. *Microbiology* **154**: 905-914.

Hodgson, D.A. (1982) Glucose repression of carbon source uptake and metabolism in *Streptomyces coelicolor* A3(2) and its perturbation in mutants resistant to 2-deoxyglucose. *J Gen Microbiol* **128**: 2417-2430.

Hong, B., Phornphisutthimas, S., Tilley, E., Baumberg, S., and McDowall, K.J. (2007) Streptomycin production by *Streptomyces griseus* can be modulated by a mechanism not associated with change in the *adpa* component of the A-factor cascade. *Biotechnol Lett* **29**: 57-64.

Hong, H.J., Hutchings, M.I., Neu, J.M., Wright, G.D., Paget, M.S., *et al.* (2004) Characterization of an inducible vancomycin resistance system in *Streptomyces coelicolor* reveals a novel gene (*vanK*) required for drug resistance. *Mol Microbiol* **52**: 1107-1121.
Hong, S.K., Kito, M., Beppu, T., and Horinouchi, S. (1991) Phosphorylation of the AfsR product, a global regulatory protein for secondary-metabolite formation in *Streptomyces coelicolor* A3(2). *J Bacteriol* **173**: 2311-2318.
Hopwood, D.A. (1999) Forty years of genetics with *Streptomyces*: From *in vivo* through *in vitro* to *in silico*. *Microbiology* **145**: 2183-2202.
Hopwood, D.A. (2006) Soil to genomics: The *Streptomyces* chromosome. *Annu Rev Genet* **40**: 1-23.
Hopwood, D.A., (2007) *Streptomyces in nature and medicine: The antibiotic makers*. Oxford University Press, New York.
Hopwood, D.A., and Wright, H.M. (1983) CDA is a new chromosomally-determined antibiotic from *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **129**: 3575-3579.
Horbal, L., Kobylyanskyy, A., Truman, A.W., Zaburranyi, N., Ostash, B., *et al.* (2014) The pathway-specific regulatory genes, *tei15** and *tei16**, are the master switches of teicoplanin production in *Actinoplanes teichomyceticus*. *Appl Microbiol Biotechnol* **98**: 9295-9309.
Horbal, L., Kobylyanskyy, A., Yushchuk, O., Zaburannyi, N., Luzhetskyy, A., *et al.* (2013) Evaluation of heterologous promoters for genetic analysis of *Actinoplanes teichomyceticus*-producer of teicoplanin, drug of last defense. *J Biotechnol* **168**: 367-372.
Horbal, L., Zaburannyy, N., Ostash, B., Shulga, S., and Fedorenko, V. (2012) Manipulating the regulatory genes for teicoplanin production in *Actinoplanes teichomyceticus*. *World J Microbiol Biotechnol* **28**: 2095-2100.
Horinouchi, S. (2003) AfsR as an integrator of signals that are sensed by multiple serine/threonine kinases in *Streptomyces coelicolor* A3(2). *J Ind Microbiol Biotechnol* **30**: 462-467.
Horinouchi, S. (2007) Mining and polishing of the treasure trove in the bacterial genus *Streptomyces*. *Biosci Biotechnol Biochem* **71**: 283-299.
Horinouchi, S., Kito, M., Nishiyama, M., Furuya, K., Hong, S.K., *et al.* (1990) Primary structure of *afsR*, a global regulatory protein for secondary metabolite formation in *Streptomyces coelicolor* A3(2) *Gene* **95**: 49-56.
Hoskisson, P.A., and Seipke, R.F. (2020) Cryptic or silent? The known unknowns, unknown knowns, and unknown unknowns of secondary metabolism. *mBio* **11**: 2642-2646.
Hostalek, Z. (1980) Catabolite regulation of antibiotic biosynthesis. *Folia Microbiol (Praha)* **25**: 445-450.
Hou, B., Lin, Y., Wu, H., Guo, M., Petkovic, H., *et al.* (2018) The novel transcriptional regulator LmbU promotes lincomycin biosynthesis through regulating expression of its target genes in *Streptomyces lincolnensis*. *J Bacteriol* **200**: 447-463.
Hu, H., Zhang, Q., and Ochi, K. (2002) Activation of antibiotic biosynthesis by specified mutations in the *rpoB* gene (encoding the RNA polymerase beta subunit) of *Streptomyces lividans*. *J Bacteriol* **184**: 3984-3991.
Huang, C., Yang, C., Zhang, W., Zhang, L., De, B.C., *et al.* (2018) Molecular basis of dimer formation during the biosynthesis of benzofluorene-containing atypical angucyclines. *Nat Commun* **9**: 2088.
Huang, H., Hou, L., Li, H., Qiu, Y., Ju, J., *et al.* (2016) Activation of a plasmid-situated type III PKS gene cluster by deletion of a *wbl* gene in deepsea-derived *Streptomyces somaliensis* scsio zh66. *Microb Cell Fact* **15**: 116.

Huang, M., Lu, J.J., and Ding, J. (2021) Natural products in cancer therapy: Past, present and future. *Nat Prod Bioprospect* **11**: 5-13.
Huang, X., Ma, T., Tian, J., Shen, L., Zuo, H., *et al.* (2017) *wblA*, a pleiotropic regulatory gene modulating morphogenesis and daptomycin production in *Streptomyces roseosporus*. *J Appl Microbiol* **123**: 669-677.
Hulcr, J., Adams, A.S., Raffa, K., Hofstetter, R.W., Klepzig, K.D., *et al.* (2011) Presence and diversity of *Streptomyces* in *Dendroctonus* and sympatric bark beetle galleries across north america. *Microb Ecol* **61**: 759-768.
Hulst, M.B., Grocholski, T., Neefjes, J.J.C., van Wezel, G.P., and Metsa-Ketela, M. (2021) Anthracyclines: Biosynthesis, engineering and clinical applications. *Nat Prod Rep*.
Hunter, J.D. (2007) Matplotlib: A 2D graphics environment. *Computing in Science and Engineering* **9**: 90-95.
Hutchings, M.I. (2007) Unusual two-component signal transduction pathways in the *Actinobacteria*. *Adv Appl Microbiol* **61**: 1-26.
Hutchings, M.I., Hong, H.J., and Buttner, M.J. (2006) The vancomycin resistance VanRS two-component signal transduction system of *Streptomyces coelicolor*. *Mol Microbiol* **59**: 923-935.
Hutchings, M.I., Hoskisson, P.A., Chandra, G., and Buttner, M.J. (2004) Sensing and responding to diverse extracellular signals? Analysis of the sensor kinases and response regulators of *Streptomyces coelicolor* A3(2). *Microbiology* **150**: 2795-2806.
Hutchings, M.I., Truman, A.W., and Wilkinson, B. (2019) Antibiotics: Past, present and future. *Curr Opin Microbiol* **51**: 72-80.
Hutchinson, C.R., and Colombo, A.L. (1999) Genetic engineering of doxorubicin production in *Streptomyces peucetius*: A review. *J Ind Microbiol Biotechnol* **23**: 647-652.
Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., *et al.* (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* **21**: 526-531.
Inaoka, T., and Ochi, K. (2011) Scandium stimulates the production of amylase and bacilysin in *Bacillus subtilis*. *Appl Environ Microbiol* **77**: 8181-8183.
Iorio, M., Davatgarbenam, S., Serina, S., Criscenzo, P., Zdouc, M.M., *et al.* (2021) Blocks in the pseudouridimycin pathway unlock hidden metabolites in the *Streptomyces* producer strain. *Sci Rep* **11**: 5827.
Iqbal, M., Mast, Y., Amin, R., Hodgson, D.A., Consortium, S., *et al.* (2012) Extracting regulator activity profiles by integration of *de novo* motifs and expression data: Characterizing key regulators of nutrient depletion responses in *Streptomyces coelicolor*. *Nucleic Acids Res* **40**: 5227-5239.
Izawa, M., Kimata, S., Maeda, A., Kawasaki, T., and Hayakawa, Y. (2014) Functional analysis of hatomarubigin biosynthesis genes and production of a new hatomarubigin using a heterologous expression system. *J Antibiot (Tokyo)* **67**: 159-162.
Jensen, P.R., Moore, B.S., and Fenical, W. (2015) The marine actinomycete genus *Salinispora*: A model organism for secondary metabolite discovery. *Nat Prod Rep* **32**: 738-751.
Jones, A.C., Gust, B., Kulik, A., Heide, L., Buttner, M.J., *et al.* (2013) Phage p1-derived artificial chromosomes facilitate heterologous expression of the FK506 gene cluster. *PLoS One* **8**: e69319.
Joynt, R., and Seipke, R.F. (2018) A phylogenetic and evolutionary analysis of antimycin biosynthesis. *Microbiology* **164**: 28-39.
Ju, K.S., Zhang, X., and Elliot, M.A. (2018) New kid on the block: LmbU expands the repertoire of specialized metabolic regulators in *Streptomyces*. *J Bacteriol* **200**: 559-562.

Kallio, P., Liu, Z., Mantsala, P., Niemi, J., and Metsa-Ketela, M. (2008a) A nested gene in *Streptomyces* bacteria encodes a protein involved in quaternary complex formation. *J Mol Biol* **375**: 1212-1221.

Kallio, P., Liu, Z., Mantsala, P., Niemi, J., and Metsa-Ketela, M. (2008b) Sequential action of two flavoenzymes, PgaE and pgam, in angucycline biosynthesis: Chemoenzymatic synthesis of gaudimycin C. *Chem Biol* **15**: 157-166.

Kallio, P., Patrikainen, P., Belogurov, G.A., Mantsala, P., Yang, K.*, et al.* (2013) Tracing the evolution of angucyclinone monooxygenases: Structural determinants for C-12b hydroxylation and substrate inhibition in PgaE. *Biochemistry* **52**: 4507-4516.

Kamjam, M., Sivalingam, P., Deng, Z., and Hong, K. (2017) Deep sea actinomycetes and their secondary metabolites. *Front Microbiol* **8**: 760.

Kang, S.G., Jin, W., Bibb, M., and Lee, K.J. (1998) Actinorhodin and undecylprodigiosin production in wild-type and *relA* mutant strains of *Streptomyces coelicolor* A3(2) grown in continuous culture. *FEMS Microbiol Lett* **168**: 221-226.

Kang, S.H., Huang, J., Lee, H.N., Hur, Y.A., Cohen, S.N.*, et al.* (2007) Interspecies DNA microarray analysis identifies WblA as a pleiotropic down-regulator of antibiotic biosynthesis in *Streptomyces*. *J Bacteriol* **189**: 4315-4319.

Kato, J.Y., Funa, N., Watanabe, H., Ohnishi, Y., and Horinouchi, S. (2007) Biosynthesis of g-butyrolactone autoregulators that switch on secondary metabolism and morphological development in *Streptomyces*. *Proc Natl Acad Sci U S A* **104**: 2378-2383.

Katz, L., and Baltz, R.H. (2016) Natural product discovery: Past, present, and future. *J Ind Microbiol Biotechnol* **43**: 155-176.

Kautsar, S.A., Blin, K., Shaw, S., Navarro-Munoz, J.C., Terlouw, B.R.*, et al.* (2020) MIBiG 2.0: A repository for biosynthetic gene clusters of known function. *Nucleic Acids Res* **48**: D454-D458.

Kawai, K., Wang, G., Okamoto, S., and Ochi, K. (2007) The rare earth, scandium, causes antibiotic overproduction in *Streptomyces* spp. *FEMS Microbiol Lett* **274**: 311-315.

Kearns, D.B., Chu, F., Branda, S.S., Kolter, R., and Losick, R. (2005) A master regulator for biofilm formation by *Bacillus subtilis*. *Mol Microbiol* **55**: 739-749.

Kelemen, G.H., and Buttner, M.J. (1998) Initiation of aerial mycelium formation in *Streptomyces*. *Curr Opin Microbiol* **1**: 656-662.

Kharel, M.K., Pahari, P., Shepherd, M.D., Tibrewal, N., Nybo, S.E.*, et al.* (2012) Angucyclines: Biosynthesis, mode-of-action, new natural products, and synthesis. *Nat Prod Rep* **29**: 264-325.

Kharel, M.K., and Rohr, J. (2012) Delineation of gilvocarcin, jadomycin, and landomycin pathways through combinatorial biosynthetic enzymology. *Curr Opin Chem Biol* **16**: 150-161.

Khodakaramian, G., Lissenden, S., Gust, B., Moir, L., Hoskisson, P.A.*, et al.* (2006) Expression of Cre recombinase during transient phage infection permits efficient marker removal in *Streptomyces*. *Nucleic Acids Res* **34**: 20.

Kibret, M., Guerrero-Garzon, J.F., Urban, E., Zehl, M., Wronski, V.K.*, et al.* (2018) *Streptomyces* spp. From Ethiopia producing antimicrobial compounds: Characterization via bioassays, genome analyses, and mass spectrometry. *Front Microbiol* **9**: 1270.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., and Hopwood, D.A. (2000) Practical *Streptomyces* genetics. *The John Innes Foundation, Norwich, United Kingdom*.

Kilian, R., Frasch, H.J., Kulik, A., Wohlleben, W., and Stegmann, E. (2016) The VanRS homologous two-component system $VnlRS_{Ab}$ of the glycopeptide producer *Amycolatopsis balhimycina* activates transcription of the $vanHAX_{Sc}$ genes in *Streptomyces coelicolor*, but not in *A. Balhimycina*. *Microb Drug Resist* **22**: 499-509.

Kim, E., Moore, B.S., and Yoon, Y.J. (2015a) Reinvigorating natural product combinatorial biosynthesis with synthetic biology. *Nat Chem Biol* **11**: 649-659.
Kim, E.S., Hong, H.J., Choi, C.Y., and Cohen, S.N. (2001) Modulation of actinorhodin biosynthesis in *Streptomyces lividans* by glucose repression of *afsR2* gene transcription *J Bacteriol* **183**: 2969-2969.
Kim, H.J., Kim, M.K., Jin, Y.Y., and Kim, E.S. (2014) Effect of antibiotic down-regulatory gene *wblA* ortholog on antifungal polyene production in rare actinomycetes *Pseudonocardia autotrophica*. *J Microbiol Biotechnol* **24**: 1226-1231.
Kim, S.H., Traag, B.A., Hasan, A.H., McDowall, K.J., Kim, B.G.*, et al.* (2015b) Transcriptional analysis of the cell division-related *ssg* genes i*n Streptomyces coelicolor* reveals direct control of *ssgR* by AtrA. *Antonie Van Leeuwenhoek* **108**: 201-213.
Kitani, S., Doi, M., Shimizu, T., Maeda, A., and Nihira, T. (2010) Control of secondary metabolism by *farX*, which is involved in the gamma-butyrolactone biosynthesis of *Streptomyces lavendulae* FRI-5. *Arch Microbiol* **192**: 211-220.
Kitani, S., Iida, A., Izumi, T.A., Maeda, A., Yamada, Y.*, et al.* (2008) Identification of genes involved in the butyrolactone autoregulator cascade that modulates secondary metabolism in *Streptomyces lavendulae* FRI-5. *Gene* **425**: 9-16.
Kitani, S., Miyamoto, K.T., Takamatsu, S., Herawati, E., Iguchi, H.*, et al.* (2011) Avenolide, a *Streptomyces* hormone controlling antibiotic production in *Streptomyces avermitilis*. *Proc Natl Acad Sci U S A* **108**: 16410-16415.
Kitani, S., Yamada, Y., and Nihira, T. (2001) Gene replacement analysis of the butyrolactone autoregulator receptor (FarA) reveals that FarA acts as a novel regulator in secondary metabolism of *Streptomyces lavendulae* FRI-5. *J Bacteriol* **183**: 4357-4363.
Kloosterman, A.M., Cimermancic, P., Elsayed, S.S., Du, C., Hadjithomas, M.*, et al.* (2020a) Expansion of RiPP biosynthetic space through integration of pan-genomics and machine learning uncovers a novel class of lanthipeptides. *PLoS Biol* **18**: e3001026.
Kloosterman, A.M., Shelton, K.E., van Wezel, G.P., Medema, M.H., and Mitchell, D.A. (2020b) Rre-finder: A genome-mining tool for class-independent RiPP discovery. *mSystems* **5**.
Kohl, M., Wiese, S., and Warscheid, B. (2011) Cytoscape: Software for visualization and analysis of biological networks. *Methods Mol Biol* **696**: 291-303.
Kolter, R., and van Wezel, G.P. (2016) Goodbye to brute force in antibiotic discovery? *Nat Microbiol* **1**: 15020.
Komatsu, M., Komatsu, K., Koiwai, H., Yamada, Y., Kozone, I.*, et al.* (2013) Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS Synth Biol* **2**: 384-396.
Kong, L., Zhang, W., Chooi, Y.H., Wang, L., Cao, B.*, et al.* (2016) A multifunctional monooxygenase XanO4 catalyzes xanthone formation in xantholipin biosynthesis via a cryptic demethoxylation. *Cell Chem Biol* **23**: 508-516.
Kormanec, J., Novakova, R., Mingyar, E., and Feckova, L. (2014) Intriguing properties of the angucycline antibiotic auricin and complex regulation of its biosynthesis. *Appl Microbiol Biotechnol* **98**: 45-60.
Korner, H., Sofia, H.J., and Zumft, W.G. (2003) Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: Exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol Rev* **27**: 559-592.
Korynevska, A., Heffeter, P., Matselyukh, B., Elbling, L., Micksche, M.*, et al.* (2007) Mechanisms underlying the anticancer activities of the angucycline landomycin E. *Biochem Pharmacol* **74**: 1713-1726.
Kulowski, K., Wendt-Pienkowski, E., Han, L., Yang, K., Vining, L.C.*, et al.* (1999) Functional characterization of the *jadI* gene as a cyclase forming angucyclinones. *Journal of the American Chemical Society* **121**: 1786-1794.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* **35**: 1547-1549.

Kurniawan, Y.N., Kitani, S., Iida, A., Maeda, A., Lycklama a Nijeholt, J.*, et al.* (2016) Regulation of production of the blue pigment indigoidine by the pseudo gamma-butyrolactone receptor FarR2 in *Streptomyces lavendulae* FRI-5. *J Biosci Bioeng* **121**: 372-379.

Kurniawan, Y.N., Kitani, S., Maeda, A., and Nihira, T. (2014) Differential contributions of two SARP family regulatory genes to indigoidine biosynthesis in *Streptomyces lavendulae* FRI-5. *Appl Microbiol Biotechnol* **98**: 9713-9721.

Kwakman, J.H.J.M., and Postma, P.W. (1994) Glucose kinase has a regulatory role in carbon catabolite repression in *Streptomyces coelicolor*. *J Bacteriol* **176**: 2694-2698.

Labeda, D.P., Goodfellow, M., Brown, R., Ward, A.C., Lanoot, B.*, et al.* (2012) Phylogenetic study of the species within the family *Streptomycetaceae*. *Antonie Van Leeuwenhoek* **101**: 73-104.

Lai, Z., Yu, J., Ling, H., Song, Y., Yuan, J.*, et al.* (2018) Grincamycins i - k, cytotoxic angucycline glycosides derived from marine-derived actinomycete *Streptomyces lusitanus* scsio lr32. *Planta Med* **84**: 201-207.

Lambert, S., Traxler, M.F., Craig, M., Maciejewska, M., Ongena, M.*, et al.* (2014) Altered desferrioxamine-mediated iron utilization is a common trait of *bald* mutants of *Streptomyces coelicolor*. *Metallomics* **6**: 1390-1399.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.

Lautru, S., Deeth, R.J., Bailey, L.M., and Challis, G.L. (2005) Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nat Chem Biol* **1**: 265-269.

Lawlor, E.J., Baylis, H.A., and Chater, K.F. (1987) Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). *Genes Dev* **1**: 1305-1310.

Le Marechal, P., Decottignies, P., Marchand, C.H., Degrouard, J., Jaillard, D.*, et al.* (2013) Comparative proteomic analysis of *Streptomyces lividans* wild-type and *ppk* mutant strains reveals the importance of storage lipids for antibiotic biosynthesis. *Appl Environ Microbiol* **79**: 5907-5917.

Lechner, A., Eustaquio, A.S., Gulder, T.A., Hafner, M., and Moore, B.S. (2011) Selective overproduction of the proteasome inhibitor salinosporamide A via precursor pathway regulation. *Chem Biol* **18**: 1527-1536.

Lee, H.N., Im, J.H., Lee, M.J., Lee, S.Y., and Kim, E.S. (2009) A putative secreted solute binding protein, SCO6569 is a possible AfsR2-dependent down-regulator of actinorhodin biosynthesis in *Streptomyces coelicolor*. *Process Biochem* **44**: 373-377.

Lee, N., Hwang, S., Kim, J., Cho, S., Palsson, B.*, et al.* (2020) Mini review: Genome mining approaches for the identification of secondary metabolite biosynthetic gene clusters in *Streptomyces*. *Comput Struct Biotechnol J* **18**: 1548-1556.

Lehka, L.V., Panchuk, R.R., Berger, W., Rohr, J., and Stoika, R.S. (2015) The role of reactive oxygen species in tumor cells apoptosis induced by landomycin A. *Ukr Biochem J* **87**: 72-82.

Lei, C., Wang, J., Liu, Y., Liu, X., Zhao, G.*, et al.* (2018) A feedback regulatory model for RifQ-mediated repression of rifamycin export in *Amycolatopsis mediterranei*. *Microb Cell Fact* **17**: 14.

Leskiw, B.K., Lawlor, E.J., Fernandez-Abalos, J.M., and Chater, K.F. (1991) TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. *Proc Natl Acad Sci U S A* **88**: 2461-2465.

Letunic, I., and Bork, P. (2021) Interactive tree of life (itol) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* **49**: W293-W296.
Lewis, K. (2013) Platforms for antibiotic discovery. *Nat Rev Drug Discov* **12**: 371-387.
Li, C., Liu, X., Lei, C., Yan, H., Shao, Z.*, et al.* (2017a) RifZ (AMED_0655) is a pathway-specific regulator for rifamycin biosynthesis in *Amycolatopsis mediterranei*. *Appl Environ Microbiol* **83**: 3201-3210.
Li, L., Liu, X., Jiang, W., and Lu, Y. (2019) Recent advances in synthetic biology approaches to optimize production of bioactive natural products in Actinobacteria. *Front Microbiol* **10**: 2467.
Li, M.H., Ung, P.M., Zajkowski, J., Garneau-Tsodikova, S., and Sherman, D.H. (2009) Automated genome mining for natural products. *BMC Bioinformatics* **10**: 185.
Li, T.L., Huang, F., Haydock, S.F., Mironenko, T., Leadlay, P.F.*, et al.* (2004) Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: Characterization of two glycosyltransferases and the key acyltransferase. *Chem Biol* **11**: 107-119.
Li, W., Ying, X., Guo, Y., Yu, Z., Zhou, X.*, et al.* (2006) Identification of a gene negatively affecting antibiotic production and morphological differentiation in *Streptomyces coelicolor* A3(2). *J Bacteriol* **188**: 8368-8375.
Li, X., Wang, J., Li, S., Ji, J., Wang, W.*, et al.* (2016) Corrigendum: ScbR- and ScbR2-mediated signal transduction networks coordinate complex physiological responses in *Streptomyces coelicolor*. *Sci Rep* **6**: 21574.
Li, X., Wang, J., Shi, M., Wang, W., Corre, C.*, et al.* (2017b) Evidence for the formation of ScbR/ScbR2 heterodimers and identification of one of the regulatory targets in *Streptomyces coelicolor*. *Appl Microbiol Biotechnol* **101**: 5333-5340.
Li, X., Yu, T., He, Q., McDowall, K.J., Jiang, B.*, et al.* (2015) Binding of a biosynthetic intermediate to AtrA modulates the production of lidamycin by *Streptomyces globisporus*. *Mol Microbiol* **96**: 1257-1271.
Liao, C.-H., Yao, L.-l., and Ye, B.-C. (2014a) Three genes encoding citrate synthases in *Saccharopolyspora erythraea* are regulated by the global nutrient-sensing regulators GlnR, DasR, and crp. *Mol Microbiol* **94**: 1065-1084.
Liao, C., Rigali, S., Cassani, C.L., Marcellin, E., Nielsen, L.K.*, et al.* (2014b) Control of chitin and N-acetylglucosamine utilization in *Saccharopolyspora erythraea*. *Microbiology* **160**: 1914-1928.
Liao, C.H., Xu, Y., Rigali, S., and Ye, B.C. (2015a) DasR is a pleiotropic regulator required for antibiotic production, pigment biosynthesis, and morphological development in *Saccharopolyspora erythraea*. *Appl Microbiol Biotechnol* **99**: 10215-10224.
Liao, C.H., Yao, L., Xu, Y., Liu, W.B., Zhou, Y.*, et al.* (2015b) Nitrogen regulator GlnR controls uptake and utilization of non-phosphotransferase-system carbon sources in actinomycetes. *Proc Natl Acad Sci U S A* **112**: 15630-15635.
Linares, J.F., Gustafsson, I., Baquero, F., and Martinez, J.L. (2006) Antibiotics as intermicrobial signaling agents instead of weapons. *Proc Natl Acad Sci U S A* **103**: 19484-19489.
Liu, G., Chater, K.F., Chandra, G., Niu, G., and Tan, H. (2013) Molecular regulation of antibiotic biosynthesis in *Streptomyces*. *Microbiol Mol Biol Rev* **77**: 112-143.
Liu, J., Li, J., Dong, H., Chen, Y., Wang, Y.*, et al.* (2017) Characterization of an Lrp/AsnC family regulator SCO3361, controlling actinorhodin production and morphological development in *Streptomyces coelicolor*. *Appl Microbiol Biotechnol* **101**: 5773-5783.
Liu, X., Liu, D., Xu, M., Tao, M., Bai, L.*, et al.* (2018) Reconstitution of kinamycin biosynthesis within the heterologous host *Streptomyces albus* J1074. *J Nat Prod* **81**: 72-77.

Lo Grasso, L., Maffioli, S., Sosio, M., Bibb, M., Puglia, A.M., *et al.* (2015) Two master switch regulators trigger A40926 biosynthesis in *Nonomuraea* sp. Strain ATCC 39727. *J Bacteriol* **197**: 2536-2544.

Loria, R., Bignell, D.R., Moll, S., Huguet-Tapia, J.C., Joshi, M.V., *et al.* (2008) Thaxtomin biosynthesis: The path to plant pathogenicity in the genus *Streptomyces*. *Antonie Van Leeuwenhoek* **94**: 3-10.

Love, M.I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Low, Z.J., Pang, L.M., Ding, Y., Cheang, Q.W., Le Mai Hoang, K., *et al.* (2018) Identification of a biosynthetic gene cluster for the polyene macrolactam sceliphrolactam in a *Streptomyces* strain isolated from mangrove sediment. *Sci Rep* **8**: 1594.

Lu, Y., He, J., Zhu, H., Yu, Z., Wang, R., *et al.* (2011) An orphan histidine kinase, OhkA, regulates both secondary metabolism and morphological differentiation in *Streptomyces coelicolor*. *J Bacteriol* **193**: 3020-3032.

Ludwig, W., Euzeby, J., Schumann, P., Busse, H.-J., Trujillo, M.E., *et al.*, (2012) Road map of the phylum *Actinobacteria*. In: Bergey's manual of systematic bacteriology. M. Goodfellow, P. Kämpfer, H.-J. Busse, M.E. Trujillo, K.-I. Suzuki, W. Ludwig & W.B. Whitman (eds). New York: Springer, pp. 1-28.

Luo, S., Sun, D., Zhu, J., Chen, Z., Wen, Y., *et al.* (2014) An extracytoplasmic function sigma factor, sigma(25), differentially regulates avermectin and oligomycin biosynthesis in *Streptomyces avermitilis*. *Appl Microbiol Biotechnol* **98**: 7097-7112.

Ma, M., Rateb, M.E., Teng, Q., Yang, D., Rudolf, J.D., *et al.* (2015) Angucyclines and angucyclinones from *Streptomyces* sp. CB01913 featuring C-ring cleavage and expansion. *J Nat Prod* **78**: 2471-2480.

Machushynets, N.V., Elsayed, S.S., Du, C., Siegler, M.A., de la Cruz, M., *et al.* (2022) Discovery of actinomycin L, a new member of the actinomycin family of antibiotics. *Sci Rep* **12**: 2813.

Machushynets, N.V., Wu, C., Elsayed, S.S., Hankemeier, T., and van Wezel, G.P. (2019) Discovery of novel glycerolated quinazolinones from *Streptomyces* sp. MBT27. *J Ind Microbiol Biotechnol* **46**: 483-492.

MacNeil, D.J. (1988) Characterization of a unique methyl-specific restriction system in *Streptomyces avermitilis*. *J Bacteriol* **170**: 5607-5612.

MacNeil, D.J., Gewain, K.M., Ruby, C.L., Dezeny, G., Gibbons, P.H., *et al.* (1992) Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**: 61-68.

Makato, N. (2012-2017) Elucidation of mechanisms of biomaterial conversion modified by amino group-modifying carrier protein and application.

Manteca, A., Fernandez, M., and Sanchez, J. (2005) A death round affecting a young compartmentalized mycelium precedes aerial mycelium dismantling in confluent surface cultures of *Streptomyces antibioticus*. *Microbiology* **151**: 3689-3697.

Mao, X.M., Sun, Z.H., Liang, B.R., Wang, Z.B., Feng, W.H., *et al.* (2013) Positive feedback regulation of stgr expression for secondary metabolism in *Streptomyces coelicolor*. *J Bacteriol* **195**: 2072-2078.

Marcone, G.L., Carrano, L., Marinelli, F., and Beltrametti, F. (2010) Protoplast preparation and reversion to the normal filamentous growth in antibiotic-producing uncommon actinomycetes. *J Antibiot (Tokyo)* **63**: 83-88.

Martin-Martin, S., Rodriguez-Garcia, A., Santos-Beneit, F., Franco-Dominguez, E., Sola-Landa, A., *et al.* (2017) Self-control of the PHO regulon: The PhoP-dependent protein PhoU controls negatively expression of genes of PHO regulon in *Streptomyces coelicolor*. *J Antibiot (Tokyo)* **71**: 113–122.

Martin, J.F., and Liras, P. (2012) Cascades and networks of regulatory genes that control antibiotic biosynthesis. *Subcell Biochem* **64**: 115-138.
Martinet, L., Naome, A., Baiwir, D., De Pauw, E., Mazzucchelli, G.*, et al.* (2020) On the risks of phylogeny-based strain prioritization for drug discovery: *Streptomyces lunaelactis* as a case study. *Biomolecules* **10**: 1027.
Martinez-Castro, M., Salehi-Najafabadi, Z., Romero, F., Perez-Sanchiz, R., Fernandez-Chimeno, R.I.*, et al.* (2013) Taxonomy and chemically semi-defined media for the analysis of the tacrolimus producer '*Streptomyces tsukubaensis*'. *Appl Microbiol Biotechnol* **97**: 2139-2152.
Maskey, R.P., Helmke, E., and Laatsch, H. (2003) Himalomycin A and B: Isolation and structure elucidation of new fridamycin type antibiotics from a marine *Streptomyces* isolate. *J Antibiot (Tokyo)* **56**: 942-949.
Mast, Y., Guezguez, J., Handel, F., and Schinko, E. (2015) A complex signaling cascade governs pristinamycin biosynthesis in *Streptomyces pristinaespiralis*. *Appl Environ Microbiol* **81**: 6621-6636.
Mast, Y., Weber, T., Golz, M., Ort-Winklbauer, R., Gondran, A.*, et al.* (2011) Characterization of the 'pristinamycin supercluster' of *Streptomyces pristinaespiralis*. *Microbial Biotechnol* **4**: 192-206.
Mast, Y., and Wohlleben, W. (2014) Streptogramins - two are better than one! *Int J Med Microbiol* **304**: 44-50.
Matsuda, K., Hasebe, F., Shiwa, Y., Kanesaki, Y., Tomita, T.*, et al.* (2017) Genome mining of amino group carrier protein-mediated machinery: Discovery and biosynthetic characterization of a natural product with unique hydrazone unit. *ACS Chem Biol* **12**: 124-131.
Matsumoto, A., Hong, S.K., Ishizuka, H., Horinouchi, S., and Beppu, T. (1994) Phosphorylation of the AfsR protein involved in secondary metabolism in *Streptomyces* species by a eukaryotic-type protein kinase. *Gene* **146**: 47-56.
Matulova, M., Feckova, L., Novakova, R., Mingyar, E., Csolleiova, D.*, et al.* (2019) A structural analysis of the angucycline-like antibiotic auricin from *Streptomyces lavendulae* subsp. *Lavendulae* CCM 3239 revealed its high similarity to griseusins. *Antibiotics (Basel)* **8**: 102.
Mayer, A., Taguchi, T., Linnenbrink, A., Hofmann, C., Luzhetskyy, A.*, et al.* (2005) LanV, a bifunctional enzyme: Aromatase and ketoreductase during landomycin A biosynthesis. *Chembiochem* **6**: 2312-2315.
McDowall, K.J., Thamchaipenet, A., and Hunter, I.S. (1999) Phosphate control of oxytetracycline production by *Streptomyces rimosus* is at the level of transcription from promoters overlapped by tandem repeats similar to those of the DNA-binding sites of the OmpR family. *J Bacteriol* **181**: 3025-3032.
McKenzie, N.L., and Nodwell, J.R. (2007) Phosphorylated AbsA2 negatively regulates antibiotic production in *Streptomyces coelicolor* through interactions with pathway-specific regulatory gene promoters. *J Bacteriol* **189**: 5284-5292.
McLean, T.C., Hoskisson, P.A., and Seipke, R.F. (2016) Coordinate regulation of antimycin and candicidin biosynthesis. *mSphere* **1**: 305-318.
Medema, M.H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P.*, et al.* (2011) antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* **39**: W339-346.
Medema, M.H., de Rond, T., and Moore, B.S. (2021) Mining genomes to illuminate the specialized chemistry of life. *Nat Rev Genet* **22**: 553-571.
Medema, M.H., Kottmann, R., Yilmaz, P., Cummings, M., Biggins, J.B.*, et al.* (2015) Minimum information about a biosynthetic gene cluster. *Nat Chem Biol* **11**: 625-631.

Medema, M.H., Paalvast, Y., Nguyen, D.D., Melnik, A., Dorrestein, P.C., *et al.* (2014) Pep2Path: Automated mass spectrometry-guided genome mining of peptidic natural products. *PLoS Comput Biol* **10**: e1003822.
Merrick, M.J. (1976) A morphological and genetic mapping study of *bald* colony mutants of *Streptomyces coelicolor*. *J Gen Microbiol* **96**: 299-315.
Merrick, M.J., and Edwards, R.A. (1995) Nitrogen control in bacteria. *Microbiol Rev* **59**: 604-622.
Metsa-Ketela, M., Palmu, K., Kunnari, T., Ylihonko, K., and Mantsala, P. (2003) Engineering anthracycline biosynthesis toward angucyclines. *Antimicrob Agents Chemother* **47**: 1291-1296.
Miguelez, E.M., Hardisson, C., and Manzanal, M.B. (2000) Streptomycetes: A new model to study cell death. *Int Microbiol* **3**: 153-158.
Mikhaylov, A.A., Ikonnikova, V.A., and Solyev, P.N. (2021) Disclosing biosynthetic connections and functions of atypical angucyclinones with a fragmented C-ring. *Nat Prod Rep* **38**: 1506-1517.
Mingyar, E., Feckova, L., Novakova, R., Bekeova, C., and Kormanec, J. (2015) A gamma-butyrolactone autoregulator-receptor system involved in the regulation of auricin production in *Streptomyces aureofaciens* CCM 3239. *Appl Microbiol Biotechnol* **99**: 309-325.
Miyamoto, K.T., Kitani, S., Komatsu, M., Ikeda, H., and Nihira, T. (2011) The autoregulator receptor homologue AvaR3 plays a regulatory role in antibiotic production, mycelial aggregation and colony development of *Streptomyces avermitilis*. *Microbiology* **157**: 2266-2275.
Monciardini, P., Iorio, M., Maffioli, S., Sosio, M., and Donadio, S. (2014) Discovering new bioactive molecules from microbial sources. *Microb Biotechnol* **7**: 209-220.
Moree, W.J., McConnell, O.J., Nguyen, D.D., Sanchez, L.M., Yang, Y.L., *et al.* (2014) Microbiota of healthy corals are active against fungi in a light-dependent manner. *ACS Chem Biol* **9**: 2300-2308.
Motamedi, H., Shafiee, A., and Cai, S.J. (1995) Integrative vectors for heterologous gene expression in *Streptomyces* spp. *Gene* **160**: 25-31.
Munch, D., Muller, A., Schneider, T., Kohl, B., Wenzel, M., *et al.* (2014) The lantibiotic NAI-107 binds to bactoprenol-bound cell wall precursors and impairs membrane functions. *J Biol Chem* **289**: 12063-12076.
Nah, H.J., Pyeon, H.R., Kang, S.H., Choi, S.S., and Kim, E.S. (2017) Cloning and heterologous expression of a large-sized natural product biosynthetic gene cluster in *Streptomyces* species. *Front Microbiol* **8**: 394.
Nah, J.H., Park, S.H., Yoon, H.M., Choi, S.S., Lee, C.H., *et al.* (2012) Identification and characterization of *wblA*-dependent *tmcT* regulation during tautomycetin biosynthesis in *Streptomyces* sp. Ck4412. *Biotechnol Adv* **30**: 202-209.
Nazari, B., Kobayashi, M., Saito, A., Hassaninasab, A., Miyashita, K., *et al.* (2012) Chitin-induced gene expression involved in secondary metabolic pathways in *Streptomyces coelicolor* A3(2) grown in soil. *Appl Environ Microbiol* **79**: 707-713.
Nett, M., Ikeda, H., and Moore, B.S. (2009) Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Natural Prod Rep* **26**: 1362-1384.
Nguyen, D.D., Wu, C.H., Moree, W.J., Lamsa, A., Medema, M.H., *et al.* (2013) MS/MS networking guided analysis of molecule and gene cluster families. *Proc Natl Acad Sci U S A* **110**: E2611-2620.
Niu, G., Chater, K.F., Tian, Y., Zhang, J., and Tan, H. (2016) Specialised metabolites regulating antibiotic biosynthesis in *Streptomyces* spp. *FEMS Microbiol Rev* **40**: 554-573.

Noens, E.E., Mersinias, V., Willemse, J., Traag, B.A., Laing, E.*, et al.* (2007) Loss of the controlled localization of growth stage-specific cell-wall synthesis pleiotropically affects developmental gene expression in an *ssgA* mutant of *Streptomyces coelicolor*. *Mol Microbiol* **64**: 1244-1259.

Noh, J.H., Kim, S.H., Lee, H.N., Lee, S.Y., and Kim, E.S. (2010) Isolation and genetic manipulation of the antibiotic down-regulatory gene, *wblA* ortholog for doxorubicin-producing *Streptomyces* strain improvement. *Appl Microbiol Biotechnol* **86**: 1145-1153.

Nothaft, H., Dresel, D., Willimek, A., Mahr, K., Niederweis, M.*, et al.* (2003) The phosphotransferase system of *Streptomyces coelicolor* is biased for N-acetylglucosamine metabolism. *J Bacteriol* **185**: 7019-7023.

Nothaft, H., Rigali, S., Boomsma, B., Swiatek, M., McDowall, K.J.*, et al.* (2010) The permease gene nage2 is the key to N-acetylglucosamine sensing and utilization in *Streptomyces coelicolor* and is subject to multi-level control. *Mol Microbiol* **75**: 1133-1144.

Nothias, L.F., Petras, D., Schmid, R., Duhrkop, K., Rainer, J.*, et al.* (2020) Feature-based molecular networking in the GNPS analysis environment. *Nat Methods* **17**: 905-908.

Novakova, R., Homerova, D., Feckova, L., and Kormanec, J. (2005) Characterization of a regulatory gene essential for the production of the angucycline-like polyketide antibiotic auricin in *Streptomyces aureofaciens* CCM 3239. *Microbiology* **151**: 2693-2706.

Novotna, G.B., Kwun, M.J., and Hong, H.J. (2015) *In vivo* characterization of the activation and interaction of the VanR-VanS two-component regulatory system controlling glycopeptide antibiotic resistance in two related *Streptomyces* species. *Antimicrob Agents Chemother* **60**: 1627-1637.

O'Rourke, S., Wietzorrek, A., Fowler, K., Corre, C., Challis, G.L.*, et al.* (2009) Extracellular signalling, translational control, two repressors and an activator all contribute to the regulation of methylenomycin production in *Streptomyces coelicolor*. *Mol Microbiol* **71**: 763-778.

Ochi, K., Tanaka, Y., and Tojo, S. (2014) Activating the expression of bacterial cryptic genes by *rpoB* mutations in RNA polymerase or by rare earth elements. *Journal of industrial microbiology & biotechnology* **41**: 403-414.

Ohnishi, Y., Ishikawa, J., Hara, H., Suzuki, H., Ikenoya, M.*, et al.* (2008) Genome sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350. *J Bacteriol* **190**: 4050-4060.

Ohnishi, Y., Kameyama, S., Onaka, H., and Horinouchi, S. (1999) The A-factor regulatory cascade leading to streptomycin biosynthesis in *Streptomyces griseus*: Identification of a target gene of the A-factor receptor. *Mol Microbiol* **34**: 102-111.

Ohnishi, Y., Yamazaki, H., Kato, J.Y., Tomono, A., and Horinouchi, S. (2005) AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. *Biosci Biotechnol Biochem* **69**: 431-439.

Okada, B.K., and Seyedsayamdost, M.R. (2017) Antibiotic dialogues: Induction of silent biosynthetic gene clusters by exogenous small molecules. *FEMS Microbiol Rev* **41**: 19-33.

Olano, C., Garcia, I., Gonzalez, A., Rodriguez, M., Rozas, D.*, et al.* (2014) Activation and identification of five clusters for secondary metabolites in *Streptomyces albus* J1074. *Microbial Biotechnol* **7**: 242-256.

Olano, C., Mendez, C., and Salas, J.A. (2009) Antitumor compounds from actinomycetes: From gene clusters to new derivatives by combinatorial biosynthesis. *Nat Prod Rep* **26**: 628-660.

Oliynyk, M., Samborskyy, M., Lester, J.B., Mironenko, T., Scott, N.*, et al.* (2007) Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL 23338. *Nature Biotechnol* **25**: 447-453.

Onaka, H., Ando, N., Nihira, T., Yamada, Y., Beppu, T.*, et al.* (1995) Cloning and characterisation of the A-factor receptor protein from *Streptomyces griseus*. *J Bacteriol* **177**: 6083-6092.

Onaka, H., and Horinouchi, S. (1997) DNA-binding activity of the A-factor receptor protein and its recognition DNA sequences. *Mol Microbiol* **24**: 991-1000.

Ostash, B., Yushchuk, O., Tistechok, S., Mutenko, H., Horbal, L.*, et al.* (2015) The *adpA*-like regulatory gene from *Actinoplanes teichomyceticus*: *In silico* analysis and heterologous expression. *World J Microbiol Biotechnol* **31**: 1297-1301.

Paget, M.S., Chamberlin, L., Atrih, A., Foster, S.J., and Buttner, M.J. (1999) Evidence that the extracytoplasmic function sigma factor sigmae is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). *J Bacteriol* **181**: 204-211.

Paget, M.S.B., Hong, H.J., Bibb, M.J., and Buttner, M.J., (2002) The ECF sigma factors of *Streptomyces coelicolor* A3(2). In: Sgm symposium 61. D.A. Hodgson & C.M. Thomas (eds). Cambridge: Cambridge University Press, pp. 105-125.

Pait, I.G.U., Kitani, S., Kurniawan, Y.N., Asa, M., Iwai, T.*, et al.* (2017) Identification and characterization of *IbpA*, an indigoidine biosynthetic gene in the gamma-butyrolactone signaling system of *Streptomyces lavendulae* FRI-5. *J Biosci Bioeng* **124**: 369-375.

Pan, G., Gao, X., Fan, K., Liu, J., Meng, B.*, et al.* (2017) Structure and function of a C-C bond cleaving oxygenase in atypical angucycline biosynthesis. *ACS Chem Biol* **12**: 142-152.

Panina, E.M., Mironov, A.A., and Gelfand, M.S. (2003) Comparative genomics of bacterial zinc regulons: Enhanced ion transport, pathogenesis, and rearrangement of ribosomal proteins. *Proc Natl Acad Sci U S A* **100**: 9912-9917.

Park, S.S., Yang, Y.H., Song, E., Kim, E.J., Kim, W.S.*, et al.* (2009) Mass spectrometric screening of transcriptional regulators involved in antibiotic biosynthesis in *Streptomyces coelicolor* A3(2). *J Ind Microbiol Biotechnol* **36**: 1073-1083.

Patrikainen, P. (2015) Structural and functional studies of angucycline tailoring enzymes. *PhD thesis University of Turku* ISBN:978-951-929-6156-6156.

Patrikainen, P., Kallio, P., Fan, K., Klika, K.D., Shaaban, K.A.*, et al.* (2012) Tailoring enzymes involved in the biosynthesis of angucyclines contain latent context-dependent catalytic activities. *Chem Biol* **19**: 647-655.

Pawlik, K., Kotowska, M., Chater, K.F., Kuczek, K., and Takano, E. (2007) A cryptic type I polyketide synthase (*cpk*) gene cluster in *Streptomyces coelicolor* A3(2). *Arch Microbiol* **187**: 87-99.

Payne, D.J., Gwynn, M.N., Holmes, D.J., and Pompliano, D.L. (2007) Drugs for bad bugs: Confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* **6**: 29-40.

Pelzer, S., Sussmuth, R., Heckmann, D., Recktenwald, J., Huber, P.*, et al.* (1999) Identification and analysis of the balhimycin biosynthetic gene cluster and its use for manipulating glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908. *Antimicrob Agents Chemother* **43**: 1565-1573.

Perez-Redondo, R., Santamarta, I., Bovenberg, R., Martin, J.F., and Liras, P. (2010) The enigmatic lack of glucose utilization in *Streptomyces clavuligerus* is due to inefficient expression of the glucose permease gene. *Microbiology* **156**: 1527-1537.

Phelan, V.V. (2020) Feature-based molecular networking for metabolite annotation. *Methods Mol Biol* **2104**: 227-243.

Piette, A., Derouaux, A., Gerkens, P., Noens, E.E., Mazzucchelli, G.*, et al.* (2005) From dormant to germinating spores of *Streptomyces coelicolor* A3(2): New perspectives from the *crp* null mutant. *J Proteome Res* **4**: 1699-1708.

Pluskal, T., Castillo, S., Villar-Briones, A., and Oresic, M. (2010) MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**: 395.

Polkade, A.V., Mantri, S.S., Patwekar, U.J., and Jangid, K. (2016) Quorum sensing: An under-explored phenomenon in the phylum *Actinobacteria*. *Front Microbiol* **7**: 131.

Pootoolal, J., Thomas, M.G., Marshall, C.G., Neu, J.M., Hubbard, B.K., *et al.* (2002) Assembling the glycopeptide antibiotic scaffold: The biosynthesis of A47934 from *Streptomyces toyocaensis* nrrl15009. *Proc Natl Acad Sci U S A* **99**: 8962-8967.

Pope, M.K., Green, B., and Westpheling, J. (1998) The *bldB* gene encodes a small protein required for morphogenesis, antibiotic production, and catabolite control in *Streptomyces coelicolor*. *J Bacteriol* **180**: 1556-1562.

Pope, M.K., Green, B.D., and Westpheling, J. (1996) The *bld* mutants of *Streptomyces coelicolor* are defective in the regulation of carbon utilization, morphogenesis and cell-cell signalling. *Mol Microbiol* **19**: 747-756.

Postma, P.W., Lengeler, J.W., and Jacobson, G.R. (1993) Phosphoenolpyruvate:Carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev* **57**: 543-594.

Prija, F., Srinivasan, P., Das, S., Kattusamy, K., and Prasad, R. (2017) DnrI of *Streptomyces peucetius* binds to the resistance genes, *drrAB* and *drrC* but is activated by daunorubicin. *J Basic Microbiol* **57**: 862-872.

Prima, P.M., Kazuhiro, N., Kenichi, M., Takeo, T., Kazuo, S., *et al.*, (2017) *Streptomyces* sp. SpE090715-01 produces a maleimycin-like compound using amino-group carrier protein (AmCP). In., pp.

Raaijmakers, J.M., and Mazzola, M. (2012) Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. *Annu Rev Phytopathol* **50**: 403-424.

Rabyk, M., Ostash, B., Rebets, Y., Walker, S., and Fedorenko, V. (2011) *Streptomyces ghanaensis* pleiotropic regulatory gene $wblA_{(gh)}$ influences morphogenesis and moenomycin production. *Biotechnol Lett* **33**: 2481-2486.

Raju, R., Gromyko, O., Fedorenko, V., Luzhetskyy, A., and Muller, R. (2013) Oleaceran: A novel spiro[isobenzofuran-1,2'-naptho[1,8-bc]furan] isolated from a terrestrial *Streptomyces* sp. *Org Lett* **15**: 3487-3489.

Ramos, I., Guzman, S., Escalante, L., Imriskova, I., Rodriguez-Sanoja, R., *et al.* (2004) Glucose kinase alone cannot be responsible for carbon source regulation in *Streptomyces peucetius* var. Caesius. *Res Microbiol* **155**: 267-274.

Ramos, J.L., Martinez-Bueno, M., Molina-Henares, A.J., Teran, W., Watanabe, K., *et al.* (2005) The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev* **69**: 326-356.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using stagetips. *Nat Protoc* **2**: 1896-1906.

Ratcliff, W.C., and Denison, R.F. (2011) Microbiology. Alternative actions for antibiotics. *Science* **332**: 547-548.

Reitzer, L., and Schneider, B.L. (2001) Metabolic context and possible physiological themes of sigma(54)-dependent genes in *Escherichia coli*. *Microbiol Mol Biol Rev* **65**: 422-444, table of contents.

Reuther, J., and Wohlleben, W. (2007) Nitrogen metabolism in *Streptomyces coelicolor*::Transcriptional and post-translational regulation. *J Mol Microbiol Biotechnol* **12**: 139-146.

Reyes-Caballero, H., Campanello, G.C., and Giedroc, D.P. (2011) Metalloregulatory proteins: Metal selectivity and allosteric switching. *Biophys Chem* **156**: 103-114.

Rice, K.C., and Bayles, K.W. (2003) Death's toolbox: Examining the molecular components of bacterial programmed cell death. *Mol Microbiol* **50**: 729-738.
Rice, L.B. (2008) Federal funding for the study of antimicrobial resistance in nosocomial pathogens: No ESKAPE. *J Infect Dis* **197**: 1079-1081.
Rico, S., Santamaria, R.I., Yepes, A., Rodriguez, H., Laing, E.*, et al.* (2014a) Deciphering the regulon of *Streptomyces coelicolor* AbrC3, a positive response regulator of antibiotic production. *Appl Environ Microbiol* **80**: 2417-2428.
Rico, S., Yepes, A., Rodriguez, H., Santamaria, J., Antoraz, S.*, et al.* (2014b) Regulation of the AbrA1/A2 two-component system in *Streptomyces coelicolor* and the potential of its deletion strain as a heterologous host for antibiotic production. *PLoS One* **9**: e109844.
Rigali, S., Nothaft, H., Noens, E.E., Schlicht, M., Colson, S.*, et al.* (2006) The sugar phosphotransferase system of *Streptomyces coelicolor* is regulated by the GntR-family regulator DasR and links N-acetylglucosamine metabolism to the control of development. *Mol Microbiol* **61**: 1237-1251.
Rigali, S., Titgemeyer, F., Barends, S., Mulder, S., Thomae, A.W.*, et al.* (2008) Feast or famine: The global regulator DasR links nutrient stress to antibiotic production by *Streptomyces*. *EMBO Rep* **9**: 670-675.
Risdian, C., Mozef, T., and Wink, J. (2019) Biosynthesis of polyketides in *Streptomyces*. *Microorganisms* **7**: 124.
Rix, U., Remsing, L.L., Hoffmeister, D., Bechthold, A., and Rohr, J. (2003) Urdamycin L: A novel metabolic shunt product that provides evidence for the role of the *urdM* gene in the urdamycin A biosynthetic pathway of *Streptomyces fradiae* Tu 2717. *Chembiochem* **4**: 109-111.
Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y.*, et al.* (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* **4**: 651-657.
Robinson, S.L., Christenson, J.K., and Wackett, L.P. (2019) Biosynthesis and chemical diversity of beta-lactone natural products. *Nat Prod Rep* **36**: 458-475.
Rodriguez, H., Rico, S., Diaz, M., and Santamaria, R.I. (2013) Two-component systems in *Streptomyces*: Key regulators of antibiotic complex pathways. *Microb Cell Fact* **12**: 127.
Rogers, L.A., and Whittier, E.O. (1928) Limiting factors in the lactic fermentation. *J Bacteriol* **16**: 211-229.
Rogers, S., Ong, C.W., Wandy, J., Ernst, M., Ridder, L.*, et al.* (2019) Deciphering complex metabolite mixtures by unsupervised and supervised substructure discovery and semi-automated annotation from MS/MS spectra. *Faraday Discuss* **218**: 284-302.
Romano, S., Jackson, S.A., Patry, S., and Dobson, A.D.W. (2018) Extending the "one strain many compounds" (OSMAC) principle to marine microorganisms. *Mar Drugs* **16**: 244.
Romero-Rodríguez, A., Rocha, D., Ruiz-Villafan, B., Tierrafría, V., Rodríguez-Sanoja, R.*, et al.* (2016) Transcriptomic analysis of a classical model of carbon catabolite regulation in *Streptomyces coelicolor*. *BMC microbiology* **16**: 1.
Romero, A., Ruiz, B., Sohng, J.K., Koirala, N., Rodriguez-Sanoja, R.*, et al.* (2015) Functional analysis of the GlcP promoter in *Streptomyces peucetius* var. *Caesius*. *Appl Biochem Biotechnol* **175**: 3207-3217.
Romero, D., Traxler, M.F., Lopez, D., and Kolter, R. (2011) Antibiotics as signal molecules. *Chem Rev* **111**: 5492-5505.
Romero, D.A., Hasan, A.H., Lin, Y.F., Kime, L., Ruiz-Larrabeiti, O.*, et al.* (2014) A comparison of key aspects of gene regulation in *Streptomyces coelicolor* and *Escherichia coli* using nucleotide-resolution transcription maps produced in parallel by global and differential RNA sequencing. *Mol Microbiol*.

Rosenberg, S.M. (2009) Life, death, differentiation, and the multicellularity of bacteria. *PLoS genetics* **5**: e1000418.
Rottig, M., Medema, M.H., Blin, K., Weber, T., Rausch, C.*, et al.* (2011) NRPSpredictor2-a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res* **39**: W362-367.
Rudd, B.A., and Hopwood, D.A. (1979) Genetics of actinorhodin biosynthesis by *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **114**: 35-43.
Rutledge, P.J., and Challis, G.L. (2015) Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol* **13**: 509-523.
Saier, M.H., Jr., and Reizer, J. (1992) Proposed uniform nomenclature for the proteins and protein domains of the bacterial phosphoenolpyruvate: Sugar phosphotransferase system. *J Bacteriol* **174**: 1433-1438.
Sambrook J., F.E.F., Maniatis T. (1989) Molecular cloning: A laboratory manual.
Sanchez, S., Chavez, A., Forero, A., Garcia-Huante, Y., Romero, A.*, et al.* (2010) Carbon source regulation of antibiotic production. *J Antibiot (Tokyo)* **63**: 442-459.
Sanchez, S., and Demain, A.L. (2002) Metabolic regulation of fermentation processes. *Enzyme Microb Technol* **31**: 895-906.
Santos-Aberturas, J., Payero, T.D., Vicente, C.M., Guerra, S.M., Canibano, C.*, et al.* (2011) Functional conservation of PAS-LuxR transcriptional regulators in polyene macrolide biosynthesis. *Metab Eng* **13**: 756-767.
Santos-Beneit, F. (2015) The Pho regulon: A huge regulatory network in bacteria. *Front Microbiol* **6**: 402.
Santos-Beneit, F., Rodriguez-Garcia, A., Franco-Dominguez, E., and Martin, J.F. (2008) Phosphate-dependent regulation of the low- and high-affinity transport systems in the model actinomycete *Streptomyces coelicolor*. *Microbiology* **154**: 2356-2370.
Santos-Beneit, F., Rodriguez-Garcia, A., and Martin, J.F. (2011) Complex transcriptional control of the antibiotic regulator *afsS* in *Streptomyces*: PhoP and AfsR are overlapping, competitive activators. *J Bacteriol* **193**: 2242-2251.
Santos-Beneit, F., Rodríguez-García, A., and Martín, J.F. (2012) Overlapping binding of PhoP and AfsR to the promoter region of *glnR* in *Streptomyces coelicolor*. *Microbiol Res* **167**: 532-535.
Santos-Beneit, F., Rodríguez-García, A., Sola-Landa, A., and Martín, J.F. (2009) Cross-talk between two global regulators in *Streptomyces*: PhoP and AfsR interact in the control of *afsS, pstS* and *phoRP* transcription. *Mol Microbiol* **72**: 53-68.
Sasaki, E., Ogasawara, Y., and Liu, H.W. (2010) A biosynthetic pathway for BE-7585A, a 2-thiosugar-containing angucycline-type natural product. *J Am Chem Soc* **132**: 7405-7417.
Schafer, M., Le, T.B., Hearnshaw, S.J., Maxwell, A., Challis, G.L.*, et al.* (2015) SimC7 is a novel NAD(P)H-dependent ketoreductase essential for the antibiotic activity of the DNA gyrase inhibitor simocyclinone. *J Mol Biol* **427**: 2192-2204.
Schmid, R., et. al. (2020) Ion identity networking in the GNPS environment *bioRxiv*.
Schorn, M.A., Verhoeven, S., Ridder, L., Huber, F., Acharya, D.D.*, et al.* (2021) A community resource for paired genomic and metabolomic data mining. *Nat Chem Biol* **17**: 363-368.
Sciara, G., Kendrew, S.G., Miele, A.E., Marsh, N.G., Federici, L.*, et al.* (2003) The structure of ActVA-Orf6, a novel type of monooxygenase involved in actinorhodin biosynthesis. *EMBO J* **22**: 205-215.
Seemann, T. (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**: 2068-2069.
Seipke, R.F. (2015) Strain-level diversity of secondary metabolism in *Streptomyces albus*. *PLoS One* **10**: e0116457.

Seipke, R.F., and Hutchings, M.I. (2013) The regulation and biosynthesis of antimycins. *Beilstein J Org Chem* **9**: 2556-2563.
Seipke, R.F., Kaltenpoth, M., and Hutchings, M.I. (2012) *Streptomyces* as symbionts: An emerging and widespread theme? *FEMS Microbiol Rev* **36**: 862-876.
Seipke, R.F., Patrick, E., and Hutchings, M.I. (2014) Regulation of antimycin biosynthesis by the orphan ECF RNA polymerase sigma factor sigma (anta.). *PeerJ* **2**: e253.
Seipke, R.F., Song, L., Bicz, J., Laskaris, P., Yaxley, A.M., *et al.* (2011) The plant pathogen *Streptomyces* scabies 87-22 has a functional pyochelin biosynthetic pathway that is regulated by TetR- and AfsR-family proteins. *Microbiology* **157**: 2681-2693.
Sekurova, O.N., Brautaset, T., Sletta, H., Borgos, S.E.F., Jakobsen, O.M., *et al.* (2004) *In vivo* analysis of the regulatory genes in the nystatin biosynthetic gene cluster of *Streptomyces noursei* ATCC 11455 reveals their differential control over antibiotic biosynthesis. *J Bacteriol* **186**: 1345-1354.
Seno, E.T., and Chater, K.F. (1983) Glycerol catabolic enzymes and their regulation in wild-type and mutant strains of *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **129**: 1403-1413.
Seo, J.W., Ohnishi, Y., Hirata, A., and Horinouchi, S. (2002) ATP-binding cassette transport system involved in regulation of morphological differentiation in response to glucose in *Streptomyces griseus*. *J Bacteriol* **184**: 91-103.
Seyedsayamdost, M.R. (2014) High-throughput platform for the discovery of elicitors of silent bacterial gene clusters. *Proc Natl Acad Sci U S A* **111**: 7266-7271.
Shaikh, A.A., Nothias, L.F., Srivastava, S.K., Dorrestein, P.C., and Tahlan, K. (2021) Specialized metabolites from ribosome engineered strains of *Streptomyces clavuligerus*. *Metabolites* **11**: 239.
Sharif, E.U., and O'Doherty, G.A. (2012) Biosynthesis and total synthesis studies on the jadomycin family of natural products. *European J Org Chem* **2012**: 2095-2108.
Shawky, R.M., Puk, O., Wietzorrek, A., Pelzer, S., Takano, E., *et al.* (2007) The border sequence of the balhimycin biosynthesis gene cluster from *Amycolatopsis balhimycina* contains *bbr,* encoding a StrR-like pathway-specific regulator. *J Mol Microbiol Biotechnol* **13**: 76-88.
Sherwood, E.J., and Bibb, M.J. (2013) The antibiotic planosporicin coordinates its own production in the actinomycete *Planomonospora alba*. *Proc Natl Acad Sci U S A* **110**: E2500-2509.
Shin, J.H., Oh, S.Y., Kim, S.J., and Roe, J.H. (2007) The zinc-responsive regulator Zur controls a zinc uptake system and some ribosomal proteins in *Streptomyces coelicolor* A3(2). *J Bacteriol* **189**: 4070-4077.
Shu, D., Chen, L., Wang, W., Yu, Z., Ren, C., *et al.* (2009) *afsQ1-Q2-sigQ* is a pleiotropic but conditionally required signal transduction system for both secondary metabolism and morphological development in *Streptomyces coelicolor*. *Appl Microbiol Biotechnol* **81**: 1149-1160.
Sidda, J.D., Poon, V., Song, L., Wang, W., Yang, K., *et al.* (2016) Overproduction and identification of butyrolactones SCB1-8 in the antibiotic production superhost *Streptomyces* M1152. *Org Biomol Chem* **14**: 6390-6393.
Silver, L.L. (2011) Challenges of antibacterial discovery. *Clin Microbiol Rev* **24**: 71-109.
Skinnider, M.A., Dejong, C.A., Rees, P.N., Johnston, C.W., Li, H., *et al.* (2015) Genomes to natural products prediction informatics for secondary metabolomes (PRISM). *Nucleic Acids Res* **43**: 9645-9662.
Skinnider, M.A., Merwin, N.J., Johnston, C.W., and Magarvey, N.A. (2017) PRISM 3: Expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res* **45**: W49-W54.

Sola-Landa, A., Moura, R.S., and Martin, J.F. (2003) The two-component PhoR-PhoP system controls both primary metabolism and secondary metabolite biosynthesis in *Streptomyces lividans*. *Proc Natl Acad Sci U S A* **100**: 6133-6138.

Sola-Landa, A., Rodríguez-García, A., Amin, R., Wohlleben, W., and Martín, J.F. (2013) Competition between the GlnR and PhoP regulators for the *glnA* and *amtB* promoters in *Streptomyces coelicolor*. *Nucleic Acids Res* **41**: 1767-1782.

Sola-Landa, A., Rodriguez-Garcia, A., Franco-Dominguez, E., and Martin, J.F. (2005) Binding of PhoP to promoters of phosphate-regulated genes in *Streptomyces coelicolor*: Identification of PHO boxes. *Mol Microbiol* **56**: 1373-1385.

Som, N.F., Heine, D., Holmes, N., Knowles, F., Chandra, G.*, et al.* (2017) The MtrAB two-component system controls antibiotic production in *Streptomyces coelicolor* A3(2). *Microbiology* **163**: 1415-1419.

Sosio, M., and Donadio, S. (2006) Understanding and manipulating glycopeptide pathways: The example of the dalbavancin precursor A40926. *J Ind Microbiol Biotechnol* **33**: 569-576.

Sosio, M., Giusino, F., Cappellano, C., Bossi, E., Puglia, A.M.*, et al.* (2000) Artificial chromosomes for antibiotic-producing actinomycetes. *Nature Biotechnology* **18**: 343-345.

Sosio, M., Kloosterman, H., Bianchi, A., de Vreugd, P., Dijkhuizen, L.*, et al.* (2004) Organization of the teicoplanin gene cluster in *Actinoplanes teichomyceticus*. *Microbiology* **150**: 95-102.

Sosio, M., Stinchi, S., Beltrametti, F., Lazzarini, A., and Donadio, S. (2003) The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by *Nonomuraea* species. *Chem Biol* **10**: 541-549.

Spohn, M., Kirchner, N., Kulik, A., Jochim, A., Wolf, F.*, et al.* (2014) Overproduction of ristomycin A by activation of a silent gene cluster in *Amycolatopsis japonicum* MG417-CF17. *Antimicrob Agents Chemother* **58**: 6185-6196.

Spohn, M., Wohlleben, W., and Stegmann, E. (2016) Elucidation of the zinc-dependent regulation in *Amycolatopsis japonicum* enabled the identification of the ethylenediamine-disuccinate ([s,s]-edds) genes. *Environ Microbiol* **18**: 1249-1263.

Starcevic, A., Zucko, J., Simunkovic, J., Long, P.F., Cullum, J.*, et al.* (2008) ClustScan: An integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and *in silico* prediction of novel chemical structures. *Nucleic Acids Res* **36**: 6882-6892.

Stock, A.M., Robinson, V.L., and Goudreau, P.N. (2000) Two-component signal transduction. *Annu Rev Biochem* **69**: 183-215.

Sultan, S.P., Kitani, S., Miyamoto, K.T., Iguchi, H., Atago, T.*, et al.* (2016) Characterization of AvaR1, a butenolide-autoregulator receptor for biosynthesis of a *Streptomyces* hormone in *Streptomyces avermitilis*. *Appl Microbiol Biotechnol* **100**: 9581-9591.

Sun, J.H., Hesketh, A., and Bibb, M. (2001) Functional analysis of *relA* and *rsha*, two *relA/spot* homologues of *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **183**: 3488-3498.

Sun, Y.Q., Busche, T., Ruckert, C., Paulus, C., Rebets, Y.*, et al.* (2017) Development of a biosensor concept to detect the production of cluster-specific secondary metabolites. *ACS Synth Biol* **6**: 1026-1033.

Suroto, D.A., Kitani, S., Miyamoto, K.T., Sakihama, Y., Arai, M.*, et al.* (2017) Activation of cryptic phthoxazolin A production in *Streptomyces avermitilis* by the disruption of autoregulator-receptor homologue AvaR3. *J Biosci Bioeng*: 611-617.

Suzuki, H., Takahashi, S., Osada, H., and Yoshida, K. (2011) Improvement of transformation efficiency by strategic circumvention of restriction barriers in *Streptomyces griseus*. *J Microbiol Biotechnol* **21**: 675-678.

Swiatek-Polatynska, M.A., Bucca, G., Laing, E., Gubbens, J., Titgemeyer, F*., et al.* (2015) Genome-wide analysis of *in vivo* binding of the master regulator DasR in *Streptomyces coelicolor* identifies novel non-anonical targets. *PLoS One* **10**: e0122479.

Swiatek, M.A., Gubbens, J., Bucca, G., Song, E., Yang, Y.H*., et al.* (2013) The ROK family regulator Rok7B7 pleiotropically affects xylose utilization, carbon catabolite repression, and antibiotic production in *Streptomyces coelicolor*. *J Bacteriol* **195**: 1236-1248.

Swiatek, M.A., Tenconi, E., Rigali, S., and van Wezel, G.P. (2012) Functional analysis of the N-acetylglucosamine metabolic genes of *Streptomyces coelicolor* and role in control of development and antibiotic production. *J Bacteriol* **194**: 1136-1144.

Tahlan, K., Ahn, S.K., Sing, A., Bodnaruk, T.D., Willems, A.R*., et al.* (2007) Initiation of actinorhodin export in *Streptomyces coelicolor*. *Mol Microbiol* **63**: 951-961.

Takano, E. (2006) Gamma-butyrolactones: *Streptomyces* signalling molecules regulating antibiotic production and differentiation. *Curr Opin Microbiol* **9**: 287-294.

Takano, E., Chakraburtty, R., Nihira, T., Yamada, Y., and Bibb, M.J. (2001) A complex role for the gamma-butyrolactone SCB1 in regulating antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **41**: 1015-1028.

Takano, E., Kinoshita, H., Mersinias, V., Bucca, G., Hotchkiss, G*., et al.* (2005) A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide biosynthetic gene cluster of *Streptomyces coelicolor*. *Mol Microbiol* **56**: 465-479.

Tanaka, A., Takano, Y., Ohnishi, Y., and Horinouchi, S. (2007) AfsR recruits RNA polymerase to the afsS promoter: A model for transcriptional activation by sarps. *J Mol Biol* **369**: 322-333.

Tanaka, Y., Hosaka, T., and Ochi, K. (2010) Rare earth elements activate the secondary metabolite-biosynthetic gene clusters in *Streptomyces coelicolor* A3(2). *J Antibiot (Tokyo)* **63**: 477-481.

Tenconi, E., Jourdan, S., Motte, P., Virolle, M.J., and Rigali, S. (2012) Extracellular sugar phosphates are assimilated by *Streptomyces* in a PhoP-dependent manner. *Antonie Van Leeuwenhoek* **102**: 425-433.

Tenconi, E., and Rigali, S. (2018) Self-resistance mechanisms to DNA-damaging antitumor antibiotics in *Actinobacteria*. *Curr Opin Microbiol* **45**: 100-108.

Tenconi, E., Traxler, M., Tellatin, D., van Wezel, G.P., and Rigali, S. (2020) Prodiginines postpone the onset of sporulation in *Streptomyces coelicolor*. *Antibiotics (Basel)* **9**: 847.

Tenconi, E., Urem, M., Swiatek-Polatynska, M.A., Titgemeyer, F., Muller, Y.A*., et al.* (2015) Multiple allosteric effectors control the affinity of DasR for its target sites. *Biochem Biophys Res Commun* **464**: 324-329.

Tibrewal, N., Pahari, P., Wang, G., Kharel, M.K., Morris, C*., et al.* (2012) Baeyer-villiger C-C bond cleavage reaction in gilvocarcin and jadomycin biosynthesis. *J Am Chem Soc* **134**: 18181-18184.

Tibrewal, N., and Tang, Y. (2014) Biocatalysts for natural product biosynthesis. *Annu Rev Chem Biomol Eng* **5**: 347-366.

Tiffert, Y., Franz-Wachtel, M., Fladerer, C., Nordheim, A., Reuther, J*., et al.* (2011) Proteomic analysis of the GlnR-mediated response to nitrogen limitation in *Streptomyces coelicolor* M145. *Appl Microbiol Biotechnol* **89**: 1149-1159.

Tiffert, Y., Supra, P., Wurm, R., Wohlleben, W., Wagner, R*., et al.* (2008) The *Streptomyces coelicolor* GlnR regulon: Identification of new GlnR targets and evidence for a central role of GlnR in nitrogen metabolism in actinomycetes. *Mol Microbiol* **67**: 861-880.

Titgemeyer, F., and Hillen, W. (2002) Global control of sugar metabolism: A Gram-positive solution. *Antonie Van Leeuwenhoek* **82**: 59-71.

Tocchetti, A., Donadio, S., and Sosio, M. (2018) Large inserts for big data: Artificial chromosomes in the genomic era. *FEMS Microbiol Lett* **365**: fny064.

Tolmie, C., Smit, M.S., and Opperman, D.J. (2019) Native roles of baeyer-villiger monooxygenases in the microbial metabolism of natural compounds. *Nat Prod Rep* **36**: 326-353.

Tomono, A., Tsai, Y., Yamazaki, H., Ohnishi, Y., and Horinouchi, S. (2005) Transcriptional control by A-factor of *strR*, the pathway-specific transcriptional activator for streptomycin biosynthesis in *Streptomyces griseus*. *J Bacteriol* **187**: 5595-5604.

Traag, B.A., Kelemen, G.H., and Van Wezel, G.P. (2004) Transcription of the sporulation gene *ssgA* is activated by the IclR-type regulator SsgR in a *whi*-independent manner in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **53**: 985-1000.

Traag, B.A., and van Wezel, G.P. (2008) The SsgA-like proteins in actinomycetes: Small proteins up to a big task. *Antonie Van Leeuwenhoek* **94**: 85-97.

Traxler, M.F., Seyedsayamdost, M.R., Clardy, J., and Kolter, R. (2012) Interspecies modulation of bacterial development through iron competition and siderophore piracy. *Mol Microbiol* **86**: 628-644.

Tschowri, N., Schumacher, M.A., Schlimpert, S., Chinnam, N.B., Findlay, K.C., *et al.* (2014) Tetrameric c-di-GMP mediates effective transcription factor dimerization to control *Streptomyces* development. *Cell* **158**: 1136-1147.

Tunca, S., Barreiro, C., Coque, J.J.R., and Martín, J.F. (2009) Two overlapping antiparallel genes encoding the iron regulator DmdR1 and the Adm proteins control sidephore and antibiotic biosynthesis in *Streptomyces coelicolor* A3(2). *FEBS J* **276**: 4814-4827.

Tunca, S., Barreiro, C., Sola-Landa, A., Coque, J.J.R., and Martín, J.F. (2007) Transcriptional regulation of the desferrioxamine gene cluster of *Streptomyces coelicolor* is mediated by binding of DmdR1 to an iron box in the promoter of the *desA* gene. *FEBS J* **274**: 1110-1122.

Udwary, D.W., Zeigler, L., Asolkar, R.N., Singan, V., Lapidus, A., *et al.* (2007) Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc Natl Acad Sci U S A* **104**: 10376-10381.

Uguru, G.C., Stephens, K.E., Stead, J.A., Towle, J.E., Baumberg, S., *et al.* (2005) Transcriptional activation of the pathway-specific regulator of the actinorhodin biosynthetic genes in *Streptomyces coelicolor*. *Mol Microbiol* **58**: 131-150.

Uiterweerd, M.T., *et al.* (2020) *Iso*-maleimycin, a constitutional isomer of maleimycin, from *Streptomyces* sp. QL37. *Eur. J. Org. Chem.* **2020**: 5145-5152.

Umeyama, T., Lee, P.-C., and Horinouchi, S. (2002) Protein serine/threonine kinases in signal transduction for secondary metabolism and morphogenesis in *Streptomyces*. *Appl Microbiol Biotechnol* **59**: 419-425.

Urem, M., Swiatek-Polatynska, M.A., Rigali, S., and van Wezel, G.P. (2016a) Intertwining nutrient-sensory networks and the control of antibiotic production in *Streptomyces*. *Mol Microbiol* **102**: 183-195.

Urem, M., van Rossum, T., Bucca, G., Moolenaar, G.F., Laing, E., *et al.* (2016b) OsdR of *Streptomyces coelicolor* and the dormancy regulator DevR of *Mycobacterium tuberculosis* control overlapping regulons. *mSystems* **1**: e00014-00016.

van Bergeijk, D.A., Elsayed, S.S., Du, C., Nuñez Santiago, I., Roseboom, A.M., *et al.* (2022) The ubiquitous catechol moiety elicits siderophore and angucycline production in *Streptomyces*. *Communications Chemistry* **5**: 14.

van Bergeijk, D.A., Terlouw, B.R., Medema, M.H., and van Wezel, G.P. (2020) Ecology and genomics of *Actinobacteria*: New concepts for natural product discovery. *Nat Rev Microbiol* **18**: 546-558.

van der Aart, L.T., Lemmens, N., van Wamel, W.J., and van Wezel, G.P. (2016) Substrate inhibition of vana by d-alanine reduces vancomycin resistance in a VanX-dependent manner. *Antimicrob Agents Chemother* **60**: 4930-4939.
van der Heul, H.U., Bilyk, B.L., McDowall, K.J., Seipke, R.F., and van Wezel, G.P. (2018) Regulation of antibiotic production in Actinobacteria: New perspectives from the post-genomic era. *Nat Prod Rep* **35**: 575-604.
van der Meij, A., Worsley, S.F., Hutchings, M.I., and van Wezel, G.P. (2017) Chemical ecology of antibiotic production by actinomycetes. *FEMS Microbiol Rev* **41**: 392-416.
van Dissel, D., Claessen, D., and Van Wezel, G.P. (2014) Morphogenesis of *Streptomyces* in submerged cultures. *Adv Appl Microbiol* **89**: 1-45.
van Keulen, G., and Dyson, P.J. (2014) Production of specialized metabolites by *Streptomyces coelicolor* A3(2). *Adv Appl Microbiol* **89**: 217-266.
van Opijnen, T., Bodi, K.L., and Camilli, A. (2009) Tn-seq: High-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods* **6**: 767-772.
van Rooden, E.J., Florea, B.I., Deng, H., Baggelaar, M.P., van Esbroeck, A.C.M.*, et al.* (2018) Mapping *in vivo* target interaction profiles of covalent inhibitors using chemical proteomics with label-free quantification. *Nat Protoc* **13**: 752-767.
van Wezel, G.P., Konig, M., Mahr, K., Nothaft, H., Thomae, A.W.*, et al.* (2007) A new piece of an old jigsaw: Glucose kinase is activated posttranslationally in a glucose transport-dependent manner in *Streptomyces coelicolor* A3(2). *J Mol Microbiol Biotechnol* **12**: 67-74.
van Wezel, G.P., Krabben, P., Traag, B.A., Keijser, B.J., Kerste, R.*, et al.* (2006) Unlocking *Streptomyces* spp. For use as sustainable industrial production platforms by morphological engineering. *Appl Environ Microbiol* **72**: 5283-5288.
van Wezel, G.P., Mahr, K., Konig, M., Traag, B.A., Pimentel-Schmitt, E.F.*, et al.* (2005) GlcP constitutes the major glucose uptake system of *Streptomyces coelicolor* A3(2). *Mol Microbiol* **55**: 624-636.
van Wezel, G.P., and McDowall, K.J. (2011) The regulation of the secondary metabolism of *Streptomyces*: New links and experimental advances. *Nat Prod Rep* **28**: 1311-1333.
van Wezel, G.P., McKenzie, N.L., and Nodwell, J.R. (2009) Chapter 5. Applying the genetics of secondary metabolism in model actinomycetes to the discovery of new antibiotics. *Methods Enzymol* **458**: 117-141.
van Wezel, G.P., van der Meulen, J., Kawamoto, S., Luiten, R.G., Koerten, H.K.*, et al.* (2000a) *ssgA* is essential for sporulation of *Streptomyces coelicolor* A3(2) and affects hyphal development by stimulating septum formation. *J Bacteriol* **182**: 5653-5662.
van Wezel, G.P., White, J., Hoogvliet, G., and Bibb, M.J. (2000b) Application of *redD*, the transcriptional activator gene of the undecylprodigiosin biosynthetic pathway, as a reporter for transcriptional activity in *Streptomyces coelicolor* A3(2) and *Streptomyces lividans*. *J Mol Microbiol Biotechnol* **2**: 551-556.
van Wezel, G.P., White, J., Young, P., Postma, P.W., and Bibb, M.J. (1997) Substrate induction and glucose repression of maltose utilization by *Streptomyces coelicolor* A3(2) is controlled by *malR*, a member of the *lacI-galR* family of regulatory genes. *Mol Microbiol* **23**: 537-549.
Vara, J., Lewandowska-Skarbek, M., Wang, Y.G., Donadio, S., and Hutchinson, C.R. (1989) Cloning of genes governing the deoxysugar portion of the erythromycin biosynthesis pathway in *Saccharopolyspora erythraea* (*Streptomyces erythreus*). *J Bacteriol* **171**: 5872-5881.
Ventola, C.L. (2015) The antibiotic resistance crisis: Part 1: Causes and threats. *P T* **40**: 277-283.

Vicente, C.M., Payero, T.D., Santos-Aberturas, J., Barreales, E.G., de Pedro, A., *et al.* (2015) Pathway-specific regulation revisited: Cross-regulation of multiple disparate gene clusters by PAS-LuxR transcriptional regulators. *Appl Microbiol Biotechnol* **99**: 5123-5135.

Vicente, C.M., Santos-Aberturas, J., Payero, T.D., Barreales, E.G., de Pedro, A., *et al.* (2014) PAS-LuxR transcriptional control of filipin biosynthesis in *S. Avermitilis*. *Appl Microbiol Biotechnol* **98**: 9311-9324.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., *et al.* (2020) SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nat Methods* **17**: 261-272.

Vujaklija, D., Horinouchi, S., and Beppu, T. (1993) Detection of an A-factor-responsive protein that binds to the upstream activation sequence of *strR*, a regulatory gene for streptomycin biosynthesis in *Streptomyces griseus*. *J Bacteriol* **175**: 2652-2661.

Walker, J.M. (1994) The bicinchoninic acid (BCA) assay for protein quantitation. *Methods Mol Biol* **32**: 5-8.

Walker, M., Pohl, E., Herbst-Irmer, R., Gerlitz, M., Rohr, J., *et al.* (1999) Absolute configurations of emycin D, E and F; mimicry of centrosymmetric space groups by mixtures of chiral stereoisomers. *Acta Crystallogr B* **55**: 607-616.

Wang, B., Ren, J., Li, L., Guo, F., Pan, G., *et al.* (2015) Kinamycin biosynthesis employs a conserved pair of oxidases for B-ring contraction. *Chem Commun (Camb)* **51**: 8845-8848.

Wang, J., Wang, W., Wang, L., Zhang, G., Fan, K., *et al.* (2011) A novel role of 'pseudo'gamma-butyrolactone receptors in controlling gamma-butyrolactone biosynthesis in *Streptomyces*. *Mol Microbiol* **82**: 236-250.

Wang, J., and Zhao, G.-P. (2009) GlnR positively regulates *nasA* transcription in *Streptomyces coelicolor*. *Biochem Biophys Res Commun* **386**: 77-81.

Wang, J.B., Zhang, F., Pu, J.Y., Zhao, J., Zhao, Q.F., *et al.* (2014a) Characterization of AvaR1, an autoregulator receptor that negatively controls avermectins production in a high avermectin-producing strain. *Biotechnol Lett* **36**: 813-819.

Wang, L., Tian, X., Wang, J., Yang, H., Fan, K., *et al.* (2009) Autoregulation of antibiotic biosynthesis by binding of the end product to an atypical response regulator. *Proc Natl Acad Sci U S A* **106**: 8617-8622.

Wang, L., and Vining, L.C. (2003) Control of growth, secondary metabolism and sporulation in *Streptomyces venezuelae* ISP5230 by *jadW(1)*, a member of the *afsA* family of gamma-butyrolactone regulatory genes. *Microbiology* **149**: 1991-2004.

Wang, L., Wang, L., Zhou, Z., Wang, Y.J., Huang, J.P., *et al.* (2019) Cangumycins A-F, six new angucyclinone analogues with immunosuppressive activity from *Streptomyces*. *Chin J Nat Med* **17**: 982-987.

Wang, M., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., *et al.* (2016) Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat Biotechnol* **34**: 828-837.

Wang, W., Ji, J., Li, X., Wang, J., Li, S., *et al.* (2014b) Angucyclines as signals modulate the behaviors of *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A* **111**: 5688-5693.

Wang, X.K., and Jin, J.L. (2014) Crucial factor for increasing the conjugation frequency in *Streptomyces netropsis* SD-07 and other strains. *FEMS Microbiol Lett* **357**: 99-103.

Warner, J.B., and Lolkema, J.S. (2003) Ccpa-dependent carbon catabolite repression in bacteria. *Microbiol Mol Biol Rev* **67**: 475-490.

Waskom, M.L. (2021) Seaborn: Statistical data visualization. *Journal of Open Source Software* **6**: 3021.

Weber, T., Rausch, C., Lopez, P., Hoof, I., Gaykova, V.*, et al.* (2009) CLUSEAN: A computer-based framework for the automated analysis of bacterial secondary metabolite biosynthetic gene clusters. *J Bacteriol* **140**: 13-17.
Wenzel, S.C., Meiser, P., Binz, T.M., Mahmud, T., and Muller, R. (2006) Nonribosomal peptide biosynthesis: Point mutations and module skipping lead to chemical diversity. *Angew Chem Int Ed Engl* **45**: 2296-2301.
Wessel, D., and Flugge, U.I. (1984) A method for the quantitative recovery of protein in dilute-solution in the presence of detergents and lipids. *Anal Biochem* **138**: 141-143.
Westhoff, S., van Leeuwe, T.M., Qachach, O., Zhang, Z., van Wezel, G.P.*, et al.* (2017) The evolution of no-cost resistance at sub-MIC concentrations of streptomycin in *Streptomyces coelicolor*. *ISME J* **11**: 1168-1178.
White, J., and Bibb, M. (1997) *bldA* dependence of undecylprodigiosin production in *Streptomyces coelicolor* A3(2) involves a pathway-specific regulatory cascade. *J Bacteriol* **179**: 627-633.
Whitworth, D.E., (2012) Classification and organization of two-component systems. In: Two-component systems in bacteria. R. Gross & D. Beier (eds). Poole, UK: Caister Academic Press, pp. 1-20.
Wietzorrek, A., and Bibb, M. (1997) A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. *Mol Microbiol* **25**: 1181-1184.
Willems, A.R., Tahlan, K., Taguchi, T., Zhang, K., Lee, Z.Z.*, et al.* (2008) Crystal structures of the *Streptomyces coelicolor* TetR-like protein actr alone and in complex with actinorhodin or the actinorhodin biosynthetic precursor (s)-dnpa. *J Mol Biol* **376**: 1377-1387.
Willemse, J., Borst, J.W., de Waal, E., Bisseling, T., and van Wezel, G.P. (2011) Positive control of cell division: FtsZ is recruited by SsgB during sporulation of *Streptomyces*. *Genes Dev* **25**: 89-99.
Willey, J.M., and Gaskell, A.A. (2011) Morphogenetic signaling molecules of the streptomycetes. *Chem Rev* **111**: 174-187.
Willey, J.M., and van der Donk, W.A. (2007) Lantibiotics: Peptides of diverse structure and function. *Annu Rev Microbiol* **61**: 477-501.
Wolf, F., Bauer, J.S., Bendel, T.M., Kulik, A., Kalinowski, J.*, et al.* (2017a) Biosynthesis of the beta-lactone proteasome inhibitors belactosin and cystargolide. *Angew Chem Int Ed Engl* **56**: 6665-6668.
Wolf, T., Droste, J., Gren, T., Ortseifen, V., Schneiker-Bekel, S.*, et al.* (2017b) The MalR type regulator AcrC is a transcriptional repressor of acarbose biosynthetic genes in Actinoplanes sp. Se50/110. *BMC Genomics* **18**: 562.
Wolfender, J.L., Nuzillard, J.M., van der Hooft, J.J.J., Renault, J.H., and Bertrand, S. (2019) Accelerating metabolite identification in natural product research: Toward an ideal combination of liquid chromatography-high-resolution tandem mass spectrometry and NMR profiling, in silico databases, and chemometrics. *Anal Chem* **91**: 704-742.
Won, T.H., You, M., Lee, S.H., Rho, B.J., Oh, D.C.*, et al.* (2014) Amino alcohols from the ascidian pseudodistoma sp. *Mar Drugs* **12**: 3754-3769.
Wood, M.J. (1996) The comparative efficacy and safety of teicoplanin and vancomycin. *J Antimicrob Chemother* **37**: 209-222.
Wray, L.V., Atkinson, M.R., and Fisher, S.H. (1991) Identification and cloning of the *glnR* locus, which is required for transcription of the *ginA* gene in *Streptomyces coelicolor* A3(2). *J Bacteriol* **173**: 7351-7360.
Wright, G.D. (2014) Something old, something new: Revisiting natural products in antibiotic drug discovery. *Can J Microbiol* **60**: 147-154.

Wright, G.D. (2017) Opportunities for natural products in 21(st) century antibiotic discovery. *Nat Prod Rep* **34**: 694-701.
Wright, L.F., and Hopwood, D.A. (1976) Identification of the antibiotic determined by the SCP1 plasmid of *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **95**: 96-106.
Wu, C. (2016) Discovery of novel antibiotics from actinomycetes by integrated metabolomics and genomic approaches. *PhD thesis Leiden University*.
Wu, C., Choi, Y.H., and van Wezel, G.P. (2016a) Metabolic profiling as a tool for prioritizing antimicrobial compounds. *J Ind Microbiol Biotechnol* **43**: 299-312.
Wu, C., Du, C., Gubbens, J., Choi, Y.H., and van Wezel, G.P. (2015a) Metabolomics-driven discovery of a prenylated isatin antibiotic produced by *Streptomyces* species MBT28. *J Nat Prod* **78**: 2355-2363.
Wu, C., Kim, H.K., van Wezel, G.P., and Choi, Y.H. (2015b) Metabolomics in the natural products field-a gateway to novel antibiotics. *Drug Discov Today Technol* **13**: 11-17.
Wu, C., Medema, M.H., Lakamp, R.M., Zhang, L., Dorrestein, P.C*., et al.* (2016b) Leucanicidin and endophenasides result from methyl-rhamnosylation by the same tailoring enzymes in *Kitasatospora* sp. MBT66. *ACS Chem Biol* **11**: 478-490.
Wu, C., van der Heul, H.U., Melnik, A.V., Lubben, J., Dorrestein, P.C*., et al*. (2019) Lugdunomycin, an angucycline-derived molecule with unprecedented chemical architecture. *Angew Chem Int Ed Engl* **58**: 2809-2814.
Xiao, X. (2020) Structural and functional analysis of proteins involved in natural product biosynthesis and morphological differentiation in *Streptomyces PhD thesis Leiden University*.
Xiao, X., Elsayed, S.S., Wu, C., van der Heul, H.U., Metsa-Ketela, M*., et al.* (2020) Functional and structural insights into a novel promiscuous ketoreductase of the lugdunomycin biosynthetic pathway. *ACS Chem Biol*: 2529–2538.
Xie, H.F., Kong, Y.S., Li, R.Z., Nothias, L.F., Melnik, A.V*., et al.* (2020) Feature-based molecular networking analysis of the metabolites produced by *in vitro* solid-state fermentation reveals pathways for the bioconversion of epigallocatechin gallate. *J Agric Food Chem* **68**: 7995-8007.
Xie, P., Zeng, A., and Qin, Z. (2009) *cmdABCDEF*, a cluster of genes encoding membrane proteins for differentiation and antibiotic production in *Streptomyces coelicolor* A3(2). *BMC Microbiol* **9**: 157.
Xu, F., Nazari, B., Moon, K., Bushin, L.B., and Seyedsayamdost, M.R. (2017) Discovery of a cryptic antifungal compound from *Streptomyces albus* J1074 using high-throughput elicitor screens. *J Am Chem Soc* **139**: 9203-9212.
Xu, G., Wang, J., Wang, L., Tian, X., Yang, H*., et al.* (2010) "Pseudo" gamma-butyrolactone receptors respond to antibiotic signals to coordinate antibiotic biosynthesis. *J Biol Chem* **285**: 27440-27448.
Xu, Q., van Wezel, G.P., Chiu, H.J., Jaroszewski, L., Klock, H.E*., et al.* (2012) Structure of an MmyB-like regulator from *C. aurantiacus*, member of a new transcription factor family linked to antibiotic metabolism in actinomycetes. *PLoS One* **7**: e41359.
Yadav, G., Gokhale, R.S., and Mohanty, D. (2003) Searchpks: A program for detection and analysis of polyketide synthase domains. *Nucleic Acids Res* **31**: 3654-3658.
Yamanaka, K., Oikawa, H., Ogawa, H.O., Hosono, K., Shinmachi, F*., et al.* (2005) Desferrioxamine E produced by *Streptomyces griseus* stimulates growth and development of *Streptomyces tanashiensis*. *Microbiology* **151**: 2899-2905.
Yamanaka, K., Reynolds, K.A., Kersten, R.D., Ryan, K.S., Gonzalez, D.J*., et al.* (2014) Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc Natl Acad Sci U S A* **111**: 1957-1962.

Yan, X., Ge, H., Huang, T., Hindra, Yang, D*., et al.* (2016) Strain prioritization and genome mining for enediyne natural products. *mBio* **7**: e02104-02116.
Yang, R., Liu, X., Wen, Y., Song, Y., Chen, Z*., et al.* (2015) The PhoP transcription factor negatively regulates avermectin biosynthesis in *Streptomyces avermitilis*. *Appl Microbiol Biotechnol* **99**: 10547-10557.
Yang, Y.H., Joo, H.S., Lee, K., Liou, K.K., Lee, H.C*., et al.* (2005) Novel method for detection of butanolides in *Streptomyces coelicolor* culture broth, using a His-tagged receptor (ScbR) and mass spectrometry. *Appl Environ Microbiol* **71**: 5050-5055.
Yang, Y.H., Song, E., Kim, E.J., Lee, K., Kim, W.S*., et al.* (2009) Ndgr, an IclR-like regulator involved in amino-acid-dependent growth, quorum sensing, and antibiotic production in *Streptomyces coelicolor*. *Appl Microbiol Biotechnol* **82**: 501-511.
Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.
Yepes, A., Rico, S., Rodriguez-Garcia, A., Santamaria, R.I., and Diaz, M. (2011) Novel two-component systems implied in antibiotic production in *Streptomyces coelicolor*. *PLoS One* **6**: e19980.
Yixizhuoma, Ishikawa, N., Abdelfattah, M.S., and Ishibashi, M. (2017) Elmenols C-h, new angucycline derivatives isolated from a culture of *Streptomyces* sp. IFM 11490. *J Antibiot (Tokyo)* **70**: 601-606.
Yoon, S.Y., Lee, S.R., Hwang, J.Y., Benndorf, R., Beemelmanns, C*., et al.* (2019) Fridamycin A, a microbial natural product, stimulates glucose uptake without inducing adipogenesis. *Nutrients* **11**: 765.
Yoon, V., and Nodwell, J.R. (2014) Activating secondary metabolism with stress and chemicals. *Journal of industrial microbiology & biotechnology* **41**: 415-424.
Yu, H., Yao, Y., Liu, Y., Jiao, R., Jiang, W*., et al.* (2007) A complex role of *Amycolatopsis mediterranei* GlnR in nitrogen metabolism and related antibiotics production. *Arch Microbiol* **188**: 89-96.
Yu, L., Gao, W., Li, S., Pan, Y., and Liu, G. (2016a) GntR family regulator SCO6256 is involved in antibiotic production and conditionally regulates the transcription of myo-inositol catabolic genes in *Streptomyces coelicolor* A3(2). *Microbiology* **162**: 537-551.
Yu, L., Pan, Y., and Liu, G. (2016b) A regulatory gene SCO2140 is involved in antibiotic production and morphological differentiation of *Streptomyces coelicolor* A3(2). *Curr Microbiol* **73**: 196-201.
Yu, Z., Zhu, H., Dang, F., Zhang, W., Qin, Z*., et al.* (2012) Differential regulation of antibiotic biosynthesis by DraR-K, a novel two-component system in *Streptomyces coelicolor*. *Mol Microbiol* **85**: 535-556.
Yushchuk, O., Kharel, M., Ostash, I., and Ostash, B. (2019) Landomycin biosynthesis and its regulation in *Streptomyces*. *Appl Microbiol Biotechnol* **103**: 1659-1665.
Zarins-Tutt, J.S., Barberi, T.T., Gao, H., Mearns-Spragg, A., Zhang, L*., et al.* (2016) Prospecting for new bacterial metabolites: A glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. *Nat Prod Rep* **33**: 54-72.
Zdouc, M.M., Iorio, M., Vind, K., Simone, M., Serina, S*., et al.* (2021) Effective approaches to discover new microbial metabolites in a large strain library. *J Ind Microbiol Biotechnol* **48**: kuab017.
Zerikly, M., and Challis, G.L. (2009) Strategies for the discovery of new natural products by genome mining. *Chembiochem* **10**: 625-633.

Zhang, H., Wang, H., Wang, Y., Cui, H., Xie, Z., *et al.* (2012) Genomic sequence-based discovery of novel angucyclinone antibiotics from marine *Streptomyces* sp. W007. *FEMS Microbiol Lett* **332**: 105-112.

Zhang, H., Xie, Z., Lou, T., and Jiang, P. (2015a) Isolation, identification, and cytotoxicity of a new isobenzofuran derivative from marine *Streptomyces* sp. W007. *Chinese Journal of Oceanology and Limnology* **34**: 386–390

Zhang, J., Sun, Y., Wang, Y., Chen, X., Xue, L., *et al.* (2021) Genome mining of novel rubiginones from *Streptomyces* sp. CB02414 and characterization of the post-PKS modification steps in rubiginone biosynthesis. *Microb Cell Fact* **20**: 192.

Zhang, M.M., Wong, F.T., Wang, Y., Luo, S., Lim, Y.H., *et al.* (2017a) Crispr-cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nat Chem Biol* **13**: 607–609.

Zhang, P., Zhao, Z., Li, H., Chen, X.L., Deng, Z., *et al.* (2015b) Production of the antibiotic FR-008/candicidin in *Streptomyces* sp. FR-008 is co-regulated by two regulators, FscRI and FscRIV, from different transcription factor families. *Microbiology* **161**: 539-552.

Zhang, Q., Chen, Q., Zhuang, S., Chen, Z., Wen, Y., *et al.* (2015c) A MarR family transcriptional regulator, *dptr3*, activates daptomycin biosynthesis and morphological differentiation in *Streptomyces roseosporus*. *Appl Environ Microbiol* **81**: 3753-3765.

Zhang, S., Yang, Q., Guo, L., Zhang, Y., Feng, L., *et al.* (2017b) Isolation, structure elucidation and racemization of (+)- and (-)-pratensilins A-C: Unprecedented spiro indolinone-naphthofuran alkaloids from a marine *Streptomyces* sp. *Chem Commun (Camb)* **53**: 10066-10069.

Zhang, X., Chen, Z., Li, M., Wen, Y., Song, Y., *et al.* (2006) Construction of ivermectin producer by domain swaps of avermectin polyketide synthase in *Streptomyces avermitilis*. *Appl Microbiol Biotechnol* **72**: 986-994.

Zheng, J.T., Wang, S.L., and Yang, K.Q. (2007) Engineering a regulatory region of jadomycin gene cluster to improve jadomycin B production in *Streptomyces venezuelae*. *Appl Microbiol Biotechnol* **76**: 883-888.

Zhu, H., Sandiford, S.K., and van Wezel, G.P. (2014a) Triggers and cues that activate antibiotic production by actinomycetes. *Journal of industrial microbiology & biotechnology* **41**: 371-386.

Zhu, H., Swierstra, J., Wu, C., Girard, G., Choi, Y.H., *et al.* (2014b) Eliciting antibiotics active against the ESKAPE pathogens in a collection of actinomycetes isolated from mountain soils. *Microbiology* **160**: 1714-1725.

Zhu, J., Sun, D., Liu, W., Chen, Z., Li, J., *et al.* (2016) AvaR2, a pseudo gamma-butyrolactone receptor homologue from *Streptomyces avermitilis*, is a pleiotropic repressor of avermectin and avenolide biosynthesis and cell growth. *Mol Microbiol* **102**: 562-578.

Ziemert, N., Alanjary, M., and Weber, T. (2016) The evolution of genome mining in microbes - a review. *Nat Prod Rep* **33**: 988-1005.

Zou, Z., Du, D., Zhang, Y., Zhang, J., Niu, G., *et al.* (2014) A gamma-butyrolactone-sensing activator/repressor, JadR3, controls a regulatory mini-network for jadomycin biosynthesis. *Mol Microbiol* **94**: 490-505.

Curriculum Vitae

CURRICULUM VITAE

Hendrikje Ursusla (Helga) van der Heul was born in Schiedam, the Netherlands on the 6th of November 1990. In 2009 she attained her pre-university degree with the "Nature and Health" profile from the Groen van Prinsterer Lyceum in Vlaardingen. In that same year she started the study Biology and Medical Laboratory Research at the Hogeschool Rotterdam, and graduated with a BSc degree, cum laude, in 2012. During this period, she completed her first internship at the department of Microbiology and Infectious Diseases at the Erasmus MC on the detection of beta-lactamases in *Klebsiella pneumoniae*, under the joint supervision of Marianne ten Kate and dr. Wil Goessens. Her second bachelor internship, also at the Erasmus MC, was led by dr. Kees Vink and focused on antigenic variation in *Mycoplasma pneumoniae*.

In 2012 she started the master's study Molecular and Cellular Biology at Leiden University where she studied the ferritins of *Clostridium difficile* as part of an internship supervised by dr. Wiep Klaas Smits at the Leiden University Medical Center (LUMC). This was followed by an internship at the Institute of Biology (IBL) at Leiden University on the cell division protein SsgB from *Streptomyces coelicolor*, supervised by prof. Gilles van Wezel and dr. Joost Willemse. After receiving her MSc degree in Biology in 2014, she continued to work in the same group as research assistant. In this period, she spent two months at Naicons (Milan, Italy), a biotech company specialised in the discovery of novel bioactive compounds. Her work at the company was supervised by dr. Margherita Sosio.

In 2016 she started her PhD under the supervision of prof. Gilles van Wezel and dr. Somayah Elsayed as co-promotor. She worked on the regulation and biosynthesis of lugdunomycin production in *Streptomyces* sp. QL37, which is presented in this thesis. During her PhD study she participated in several Dutch and international conferences and presented at CHAINS 2018 in Veldhoven and at the Conference of Genetics of Industrial Microorganisms (GIM 2019) in Pisa, Italy. In 2021 she started working as lab manager and researcher in the group of prof. Gilles van Wezel.

Publications

PUBLICATIONS

van der Heul, H.U., Bilyk, B.L., McDowall, K.J., Seipke, R.F., and van Wezel, G.P. (2018) Regulation of antibiotic production in Actinobacteria: New perspectives from the post-genomic era. *Nat Prod Rep* **35**: 575–604.

Wu, C., van der Heul, H.U., Melnik, A.V., Lubben, J., Dorrestein, P.C., Minnaard, A.J., Choi, Y. H., Wezel, G.P. (2019) Lugdunomycin, an angucycline-derived molecule with unprecedented chemical architecture. *Angew Chem Int Ed Engl* **58**: 2809–2814.

Uiterweerd, M.T., Nuñez-Santiago, I., van der Heul, H.U., Wezel, G.P., Minnaard, A.J. (2020) *Iso*-maleimycin, a constitutional isomer of maleimycin, from *Streptomyces* sp. QL37. *Eur. J. Org. Chem.* **2020**: 5145-5152.

Xiao, X., Elsayed, S.S., Wu, C., van der Heul, H.U., Metsä-Ketelä, Du, C., Prota, A.E., Chen, C.-C., Liu, W., Guo, R.-T., Abrahams, J.P., van Wezel, G.P. (2020) Functional and structural insights into a novel promiscuous ketoreductase of the lugdunomycin biosynthetic pathway. *ACS Chem Biol.* **15**: 2529-2538