Universiteit Leiden
The Netherlands

Exploring the chemical space of natural products from Streptomyces using multi-omics approaches

Nuñez Santiago, I.

Citation

Nuñez Santiago, I. (2025, April 16). *Exploring the chemical space of natural products from Streptomyces using multi-omics approaches*. Retrieved from https://hdl.handle.net/1887/4212242

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/4212242

Note: To cite this publication please use the final published version (if applicable).

03

CHAPTER 3

Biomimetic total synthesis and paired omics identify an intermolecular Diels-Alder reaction as the key step in lugdunomycin biosynthesis

Michiel T. Uiterweerd*, Isabel Nuñez Santiago*, Remco W.A. Havenith, Chao Du, Le Zhang, Helga U. van der Heul, Somayah Elsayed, Adriaan J. Minnaard and Gilles P. van Wezel

*These authors contributed equally to the work

The work described in this chapter is published in the publication:

Michiel T Uiterweerd*, Isabel Nunez Santiago**,** Remco Havenith, Chao Du, Le Zhang, Helga van der Heul, Somayah Elsayed, Adri Minnaard, Gilles van Wezel. Biomimetic total synthesis and paired omics identify an intermolecular Diels-Alder reaction as the key step in lugdunomycin biosynthesis*. ChemRxiv*. (2024). 10.26434/chemrxiv-2024-d4qs8

Abstract

Microbial natural products are the basis of the majority of the clinical drugs. To tackle antimicrobial resistance (AMR), the discovery of truly novel structural scaffolds is a prerequisite. Lugdunomycin is a highly rearranged angucycline polyketide produced by *Streptomyces* sp. QL37, with an enigmatic biosynthetic pathway. Here, using combined biomimetic chemical synthesis, computational methods, genomics and molecular biology, we show that lugdunomycin is formed by a rare intermolecular Diels-Alder reaction, with elmonin as a masked diene and *iso*-maleimycin as dienophile. Genetics, extracellular complementation and heterologous expression revealed that biosynthesis of the substrates is encoded by distinct biosynthetic gene clusters (BGCs), whereby elmonin is specified by an angucycline BGC, while the biosynthesis of *iso*-maleimycin is encoded by a BGC for a β-lactone-like compound. Biomimetic total synthesis of lugdunomycin showed that the Diels-Alder reaction leads to the production of a diastereomer of lugdunomycin as the main product *in vitro*. The diastereomeric ratio of the *in vitro* Diels-Alder reaction shifted towards lugdunomycin in the presence of proteinaceous material, suggesting that the *in vivo* Diels-Alder reaction is templated. The requirement of distinct biosynthetic pathways and complex chemical reactions indicate the challenges we face in discovering new chemical space.

Introduction

The increasing threat of antibiotic-resistant bacteria, coupled with a declining number of new antibiotic discoveries, poses a serious challenge to public health. Natural products from microorganisms, particularly Actinobacteria, have been the traditional source of antibiotics (Bérdy, 2005, Barka *et al.*, 2016), but we have likely explored only a fraction of their chemical diversity. Compared to the vast number of known molecules in NPAtlas, (van Santen *et al.*, 2019) an estimated 3% of biosynthetic diversity has been accessed (Genilloud, 2014, Loureiro *et al.*, 2022). Thus, there is huge potential to overcome one of the major challenges in human health: discover truly novel chemical scaffolds that could serve as the basis for future antibiotics (Genilloud, 2014).

Lugdunomycin (**1**) is a highly rearranged angucycline derivative produced by soil-derived *Streptomyces* sp. QL37. (Wu *et al.*, 2019b) Angucyclines comprise a very large family of glycosylated natural products (Kharel *et al.*, 2012, Rohr & Thiericke, 1992). Their aglycones, referred to as angucyclinones, originate from a benz[a]anthracene framework synthesized through the polyketide pathway (Xiao *et al.*, 2020, Kharel *et al.*, 2012). They collectively constitute the most extensive category of natural products derived from a type II polyketide synthase (PKS), boasting over 400 identified members, approximately 45% of which exhibit biological activity (Elsayed *et al.*, 2023). Many of the compounds acts have antimicrobial and/or anticancer activity; however, many other bioactivities have been described, such as vasodilator, glutamate receptor agonist, platelet aggregation inhibitor or antidiabetic (Kharel *et al.*, 2012) A major reason for the strong interest in the angucyclines is the occurrence of highly diverse chemical scaffolds, whereby over a hundred different compounds may be produced by the same *Streptomyces* strain (van Bergeijk *et al.*, 2022, Wu *et al.*, 2019b). This still offers great opportunities to discover new synthetic and biosynthetic routes and thus enrich the chemical space for drug discovery.

Lugdunomycin is very different from typical angucyclinones, featuring a rare benzaza[4,3,3]propellane-6-spiro-2'-2H-naphtho[1,8-bc]furan core structure. (Wu *et al.*, 2019b) A putative biosynthetic pathway was proposed for the biosynthesis of lugdunomycin. While the biosynthesis of the core angucycline backbone of lugdunomycin is encoded by a type II PKS biosynthetic gene cluster (BGC), it is yet unclear how the final product is synthesized. Indeed, the lugdunomycin BGC is highly similar to that other angucycline BGCs, while these

strains are not able to produce lugdunomycin. We previously hypothesized that the final steps in lugdunomycin biosynthesis may require the combination of an angucyclinone and maleimycin[7] or its constitutional isomer *iso*-maleimycin (**2**). Both *iso*-maleimycin **2** and maleimycin were produced by chemical synthesis, and comparative metabolomics analysis demonstrated that *iso*-maleimycin, and not maleimycin, is produced by *Streptomyces* sp. QL37 (Uiterweerd *et al.*, 2020). Understanding the biosynthesis of lugdunomycin is of major importance for better insights into how angucyclines and angucyclinones are synthesized. Its highly unusual structure likely requires complex chemical reaction.

In this paper we resolved the final steps in lugdunomycin biosynthesis via a combination of biomimetic total synthesis and paired omics. We harnessed the power of total bioorganic synthesis to synthesize the highly complex lugdunomycin *in vitro* and, in doing so, provide compelling evidence that lugdunomycin is produced from *iso*-maleimycin and elmonin in an intermolecular Diels Alder reaction. This type of reaction is extremely rare in natural product biosynthesis. Our work also shows that *iso*-maleimycin and elmonin are produced from different biosynthetic gene clusters, which explains why lugdunomycin is rarely found in nature, despite the fact that many streptomycetes produce angucyclines. Lugdunomycin is an important example of the major challenges scientists face when searching for novel chemistry and elucidating the biosynthetic pathway.

Results and discussion

Bioinformatics combined with quantitative proteomics points at multiple gene clusters that encode lugdunomycin biosynthesis

We previously showed that lugdunomycin **1** was detected exclusively in solid-grown cultures of *Streptomyces* sp. QL37, and at very low levels; 7.5 L of minimal media (MM) agar afforded some 0.6 mg of the compound, which was sufficient to elucidate its structure (Wu *et al.*, 2019b). Still, expression of metabolites in liquid-grown cultures is the preferred method to enable elicitation approaches and paired omics. These approaches facilitate correlating gene expression profiles to metabolite production, so as to allow the identification of the biosynthetic genes for the compound (s) of interest (Gubbens *et al.*, 2014, Wu *et al.*, 2015a). In an attempt to improve lugdunomycin biosynthesis in liquid-grown cultures, *Streptomyces* sp. QL37 was grown in liquid MM with glucose,

xylose, rhamnose, fructose, arabinose or galactose as the sole carbon source, and extracts were prepared from culture supernatants. Analysis via liquid chromatography combined with mass spectrometry (LC-MS/MS) identified classical angucyclinones in all samples, as well as rearranged-angucyclinones such as elmonin (Elsayed *et al.*, 2023) (Fig. S1a). Interestingly, lugdunomycin was only detected in supernatants of galactose-grown cultures (Fig. S1b).

To establish which genes/proteins relate to lugdunomycin biosynthesis, we applied the proteomining approach (Gubbens *et al.*, 2014), which is based on the fact that protein expression levels correlate very well to metabolite levels. The strong correlation between the two allows connection of biosynthetic proteins to their cognate metabolite (s) (Gubbens *et al.*, 2014). For this, quantitative proteomics was performed to determine the most differentially expressed proteins in mycelia grown in MM with galactose (inducing conditions) and MM with glucose (non-inducing). After three days of growth, biomass was harvested and snap-frozen in liquid nitrogen. Subsequent quantitative proteomics analysis was performed on four replicate samples per growth condition, yielding 2444 quantifiable proteins, of which 42 were differentially expressed with a $\log^2$ fold change higher than 2, between galactose- and glucose-grown cultures (Table S1). Expectedly, most proteins that exhibited enhanced differential expression in galactose-grown cultures were primarily associated with galactose metabolism. However, a subset of genes stood out that belonged to a BGC that was annotated as BGC23 in the *Streptomyces* sp. QL37 genome.

BGC23 (Fig. 1a) was annotated by antiSMASH (Blin *et al.*, 2017) as consisting of two subclusters, namely BGC23a (for a β-lactone-like compound) and BGC23b (for a γ-butyrolactone). Subcluster 23a includes a gene encoding a 2-isopropylmalate synthase (QL37_30710), and a gene encoding an AMP ligase (QL37_30720), similar as seen in β-lactone BGCs responsible for belactosin and cystargolide production (Wolf *et al.*, 2017). Furthermore, BGC23a also includes a gene for an orthologue of LysW (QL37_30715), a carrier protein involved in lysine synthesis (Matsuda *et al.*, 2017). Recent studies have demonstrated that these proteins play a key role in the biosynthesis of non-proteogenic building blocks in natural products, such as vazabitide A (Hasebe *et al.*, 2016). BGC23a shows significant homology in the terms of sequence identity and gene arrangement with the putative maleimycin BGC of *Streptomyces showdoensis* ATCC 15227 (Fig. 1b).

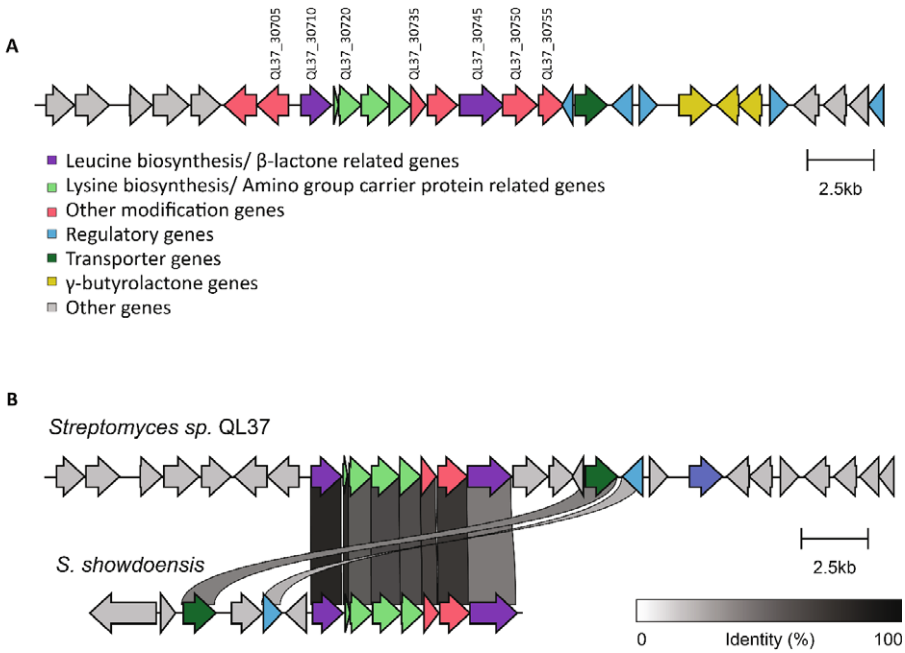


Figure 1. Genomic structure of BGC23 in *Streptomyces* sp. QL37. a. Prediction of the BGC by antiSMASH (Blin *et al.*, 2021). The BGC consists of two subclusters, BGC23a and BGC23b. Colours indicate predicted gene function as indicated. Genes that were identified as upregulated in galactose-grown cultures via proteomics are indicated with the gene name. b. Comparison of BGC23a of *Streptomyces* sp. QL37 with the putative maleimycin BGC of *S. showdoensis* ATCC 15227. Genes sharing a minimum of 40% nucleotide sequence identity are shown by connections between the BGCs. Note the high homology in the gene organization and sequence.

BGC23a is required for the biosynthesis of *iso*-maleimycin

To gain further insights into the precise compound (s) produced by BGC23a, and the possible role in lugdunomycin biosynthesis, a knock-out mutant of *Streptomyces* sp. QL37 was created in which BGC23 was inactivated. To achieve this, genes *lysW* (QL30715) and *rimK* (QL30720), which were predicted to encode an amino carrier protein and its co-factor, respectively (Hasebe *et al.*, 2016), were selected for gene deletion. The two genes were together replaced by the apramycin resistance cassette (*aacC4*) using homologous recombination. For this, a knock-out strategy was applied that is based on the unstable multi-copy plasmid pWHM3 (Swiatek *et al.*, 2012a). The correct mutant was verified by PCR.

To establish if the knock-out strain was still capable of producing *iso*-maleimycin, we compared the metabolomics profiles with chemically synthesized *iso*-maleimycin (Uiterweerd *et al.*, 2020), dissolved in methanol, which serves as a standard due to the challenges in detecting *iso*-maleimycin directly using LC-MS. In this way, *iso*-maleimycin was readily detected as its methanol adduct $[M+CH_3OH+H]^+$ (Fig. S2). Subsequently, extracts from both the parental strain *Streptomyces* sp. QL37 and its Δ*lysWrimK* mutant were analysed for the ability to produce *iso*-maleimycin. The strains were cultivated for 7 days on MM agar supplemented with 0.5% mannitol and 1% glycerol, and extracts were prepared from the culture supernatant. Although *iso*-maleimycin was readily detected in extracts from the wild-type strain, its BGC23a knock-out strain failed to produce *iso*-maleimycin (Fig. S3), thus implying its involvement in *iso*-maleimycin biosynthesis. Next, the effect of the inactivation of BGC23a on the biosynthesis of angucyclines and lugdunomycin was studied. For this purpose, mutant Δ*lysWrimK* and its parent *Streptomyces* sp. QL37, along with the minimal PKS mutant of BGC12, which fails to produce lugdunomycin or angucyclinones (Wu *et al.*, 2019b), were cultivated in the same way as mentioned above, and extracts prepared. LC-MS analysis (Fig. 2a) revealed the presence of canonical angucyclines, elmonin and lugdunomycin in extracts from the parental strain. Expectedly, no angucyclines or elmonin were detected in the minimal PKS mutant. Importantly, while the Δ*lysWrimK* mutant did produce angucyclines, lugdunomycin was not detected (Fig. 2a). Thus, the proteomics, metabolomics and gene deletion experiment collectively suggested a pivotal role for BGC23a in the later stage of lugdunomycin biosynthesis. Further analysis of the metabolome of the PKS null mutant revealed the biosynthesis of *iso*-maleimycin (Fig. S3), while the compound was not produced in D*lysWrimK* mutant. This strongly suggests that the biosynthesis of *iso*-maleimycin depends on BGC23a, while it is independent of the angucycline biosynthetic pathway. From here onwards BGC23a is referred to as the *iso*-maleimycin BGC. The involvement of a reaction partner derived from another BGC in lugdunomycin production is further supported by two lines of evidence. Firstly, heterologous expression of BGC12 in *Streptomyces coelicolor* M1152 resulted in the biosynthesis of angucyclines, while lugdunomycin itself was never detected (van der Heul, 2022). Secondly, over-expression of the angucycline gene cluster (BGC12) resulted in strong overproduction of angucyclinones, but did not improve lugdunomycin production (van der Heul, 2022).

Chemical complementation and heterologous expression support the requirement of two BGCs for lugdunomycin biosynthesis

Our data show that inactivation of either BGC12 (for the biosynthesis of angucyclinones) or BGC23a (for the biosynthesis of *iso*-maleimycin) prevents lugdunomycin production. This suggests that biosynthesis of lugdunomycin requires reaction partners from both BGCs. If indeed the two BGCs independently produce the substrates for the final step in lugdunomycin biosynthesis, the individual mutants lacking either BGC12 or BGC23a should be able to complement each other chemically, and thus produce lugdunomycin. To test this hypothesis, spores of the two strains were mixed in a 1:1 ratio and cultured for 7 days on MM agar supplemented with 0.5% mannitol and 1% glycerol and subsequently extracted with ethyl acetate for analysis by LC-MS (Fig. 2b). As controls, BGC12 null mutant Δ*minPKS*, BGC23a null mutant Δ*lysWrimK* and their parent were grown separately and treated the same way. Importantly, while no lugdunomycin was produced by either of the mutants grown alone, co-cultivation of the mutants restored lugdunomycin production (Fig. 2b). This further supports the notion that both BGCs are necessary for the biosynthesis of lugdunomycin.'

To test if the two BGCs indeed suffice for lugdunomycin biosynthesis, we expressed BGCs in *S. coelicolor* M1152 (Gomez-Escribano & Bibb, 2011) For this, we used *S. coelicolor* M1152 which already contains a copy of the *lug* gene cluster (van der Heul, 2022). This strain was shown to be able to produce angucyclines but fails to produce lugdunomycin. Importantly, additional introduction of BGC23 into *S. coelicolor* M1152+*lug* indeed afforded production of lugdunomycin **1** (Figure S4a) as well as *iso*-maleimycin **2** (Figure S4b). Taken together, all of these results demonstrate that both BGC12 and BGC23 are required for the production of lugdunomycin.

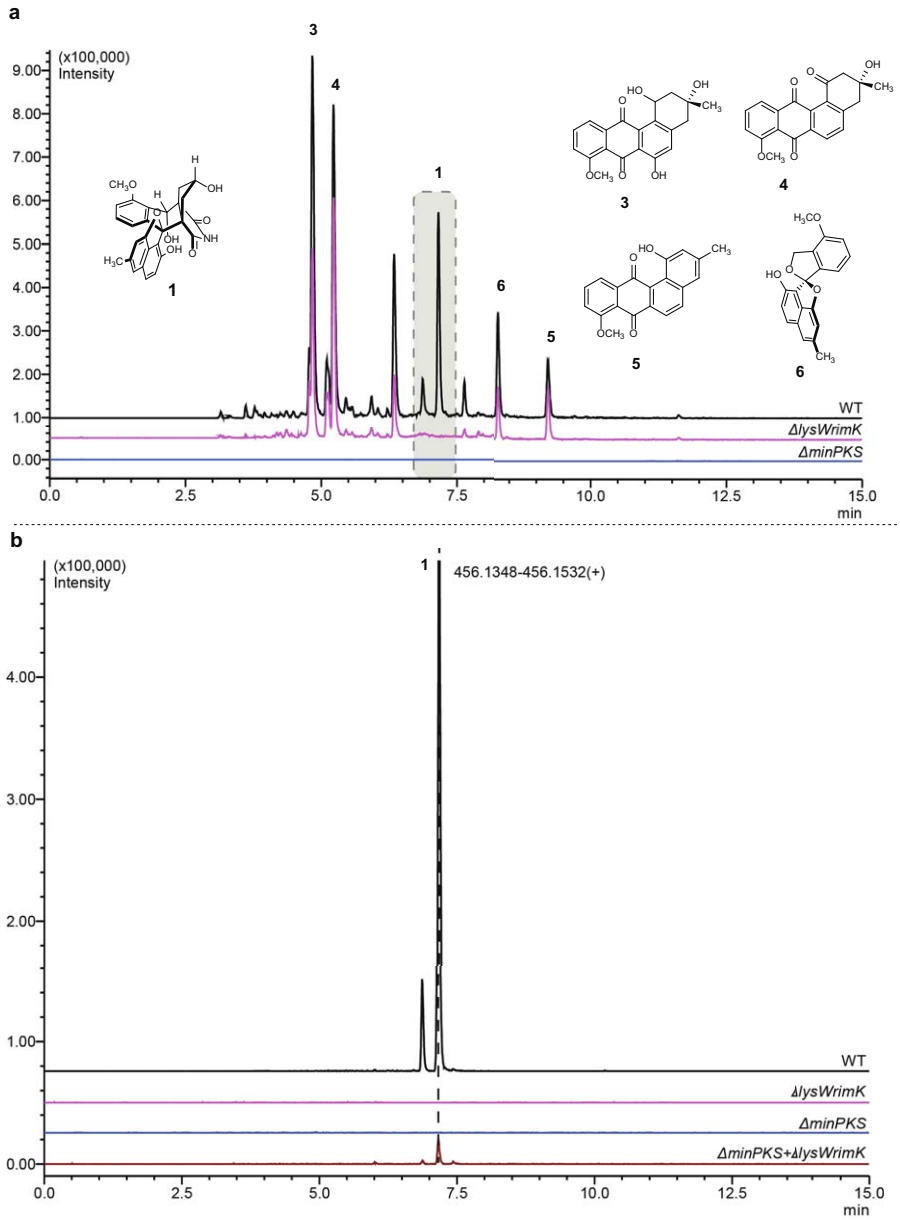


Figure 2. Comparative LC-MS chromatograms of wild-type *Streptomyces* sp. QL37 and mutant strains. a. Comparison of the *lysWrimK* mutant (pink) to the *minPKS* mutant (blue) and the parental strain *Streptomyces* sp. QL37 (black). The chromatograms represent the extracted ion chromatogram (XIC) of some typical angucyclinones (compound 3, 4 and 5), elmonin (6) and lugdunomycin (1) produced by *Streptomyces* sp. QL37. b. Extracted ion chromatogram (XIC, selected [M+H-H2O]+) for lugdunomycin in extracts obtained from solid fermentation of *Streptomyces* sp. QL37 wild type (WT, black), BGC23a mutant (Δ*lysWrimK*, pink) BGC12 mutant (Δ*minPKS*, blue), and combination of BGC12 and BGC23a spores (Δ*minPKS* , Δ*lysWrimK*, brown).

Phylogenomic analysis identifies other streptomycetes as potential lugdunomycin producers

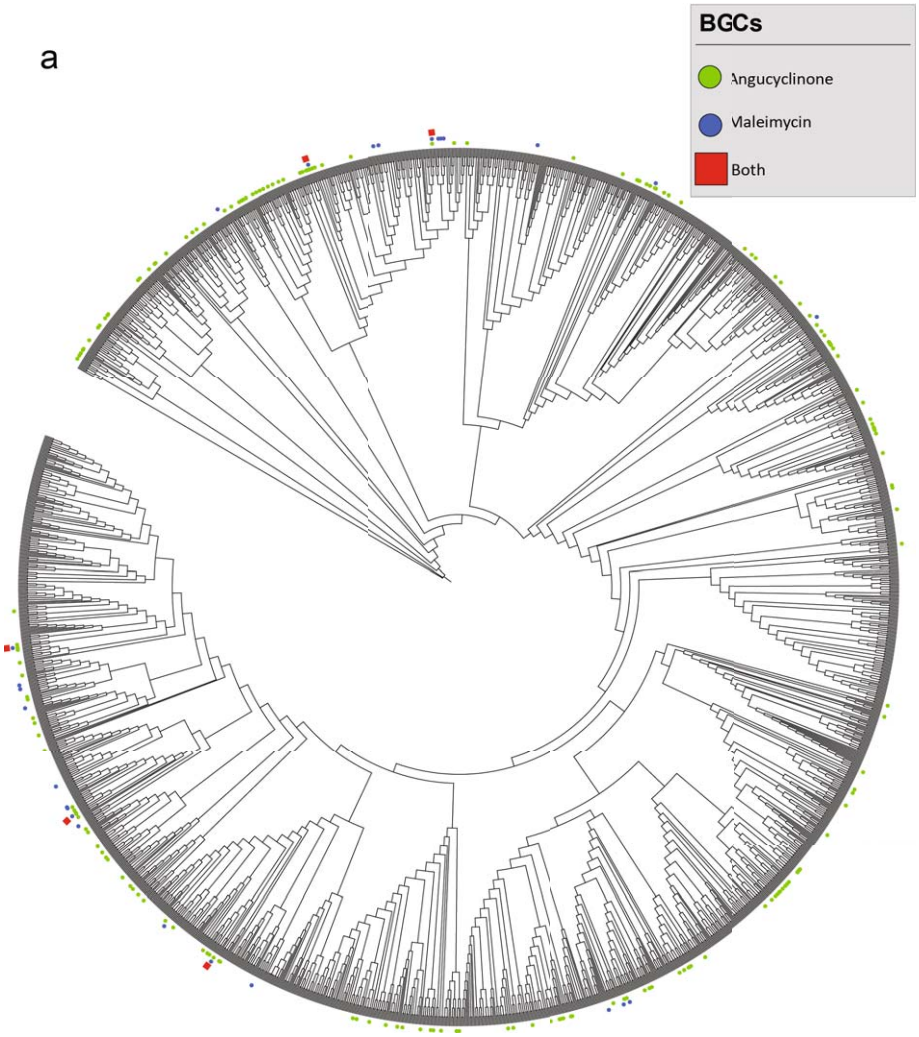
The requirement of two separate BGCs for lugdunomycin biosynthesis serves as a beacon to search for other strains that produce lugdunomycin and/or relate compounds. For this, we queried all available *Streptomyces* and *Kitasatospora* genomes to explore the co-occurrence within a single strain of a BGC homologous to the *iso*-maleimycin BGC together with a angucycline BGC. The genomes were downloaded from RefSeq, and the work was performed using the ALICE compute resources provided by Leiden University. The quality of genomes with more than 400 contigs was considered too low and therefore filtered out. In addition to the genomes obtained from the public databases, 101 genomes from our in-house MBT collection were then added to the collection. A phylogenetic tree was made for this collection of genomes using PhyloPhlAn (version 3.0.60) (Asnicar *et al.*, 2020), based on the protein sequence of 370 core genes. To predict the BGCs for natural products in these genomes, they were analysed using antiSMASH (version 6.1.1) (Blin *et al.*, 2021). This predicted the presence of in total 71624 potential BGCs. All of these BGCs were de-replicated using a modified BiG-MAP script 10.5281/zenodo.10978017. 34582 BGCs showed <40% overall similarity in the translated amino acid sequences, and these BGCs were therefore treated as different clusters. These clusters were submitted to BiG-SCAPE (version 1.1.4) (Navarro-Muñoz *et al.*, 2020) to generate sequence similarity networks. BiG-SCAPE analysed the protein motifs of the core as well as the modifying enzymes of each cluster to compare the BGCs. With a cut-off of 0.2 in distance, clusters of similar BGCs for both BGC12 and BGC23a were identified. Adding the similar BGCs which were filtered out by the BiG-MAP algorithm, 201 strains likely contained an angucycline BGC with significant similarity to BGC12, while 27 contained a BGC similar to BGC23a. Importantly, of these, five strains contained gene clusters with high similarity to both BGCs (Fig. 3). These are *Streptomyces* sp. QL37 itself and in addition *Streptomyces* sp. MBT70 (from our in-house collection), *Streptomyces* sp. SM11, *Streptomyces filamentosus* and *Streptomyces shenzhenensis*. This suggests that the biosynthesis of lugdunomycin-like molecules is not unique to *Streptomyces* sp. QL37.

From *in vivo* to *in vitro*: biomimetic synthesis of lugdunomycin from its predicted substrates

Taken together, all of our *in vivo* genetic and metabolomics data as well as the bioinformatics analyses provided strong evidence that the biosynthesis of lugdunomycin requires two BGCs. We thereby anticipated that BGC12 ensures the biosynthesis of an angucycline-derived diene, while BGC23a ensures the biosynthesis of a dienophile, most likely *iso*-maleimycin. To investigate this further, we embarked on a bioorganic synthesis route, to see if lugdunomycin could be produced in vitro. We previously synthesized *iso*-maleimycin (Uiterweerd *et al.*, 2020). Searching for a diene for the Diels Alder reaction, we realized such a diene could originate from elmonin **6** by an intramolecular elimination reaction. Elmonin is a *C*-ring cleaved, rearranged angucyclinone polyketide that had been independently isolated by two groups from *Streptomyces* spp (Tsukahara *et al.*, 2015, Raju *et al.*, 2013), though not from *Streptomyces* sp. QL37. We achieved the chemical synthesis of elmonin in a thirteen-step synthesis, as described elsewhere (Uiterweerd & Minnaard, 2022). Importantly, LC-MS showed that elmonin is produced by *Streptomyces* sp. QL37 (Fig. S5), and thus the putative dienophile (*iso*-maleimycin) and the precursor of the putative masked diene (elmonin) for the proposed intermolecular Diels-Alder reaction are *de facto* present in *Streptomyces* sp. QL37. The hypothesized pathway is presented in Fig. 4.

Elmonin is a masked diene

Elmonin 6 can be viewed at as a 1-alkoxy 1,3-dihydroisobenzofuran derivative. These compounds are precursors of isobenzofuran derivatives, generated upon base or Brønsted acid promoted 1,4-elimination (Mir-Mohamad-Sadeghy & Rickborn, 1983). Isobenzofurans, in turn, are highly reactive dienes in Diels-alder reactions (Tobia et al., 1993). Next to studies on the reactivity of isobenzofurans in situ, isobenzofuran formation has been applied in the chemical synthesis of galtamycinone (Apsel et al., 2003) and in the synthesis of epithuriferic acid methyl ester. (Koprowski et al., 2016) Considering the structure of elmonin, we anticipated it could act as a masked diene and we would be able to generate the corresponding isobenzofuran by treatment with acid. In an exploratory experiment, a solution of 6 in D3COD was treated with a small amount of p-TsOH (Fig. 5).


a
BGCs
Angucyclinone
Maleimycin
Both

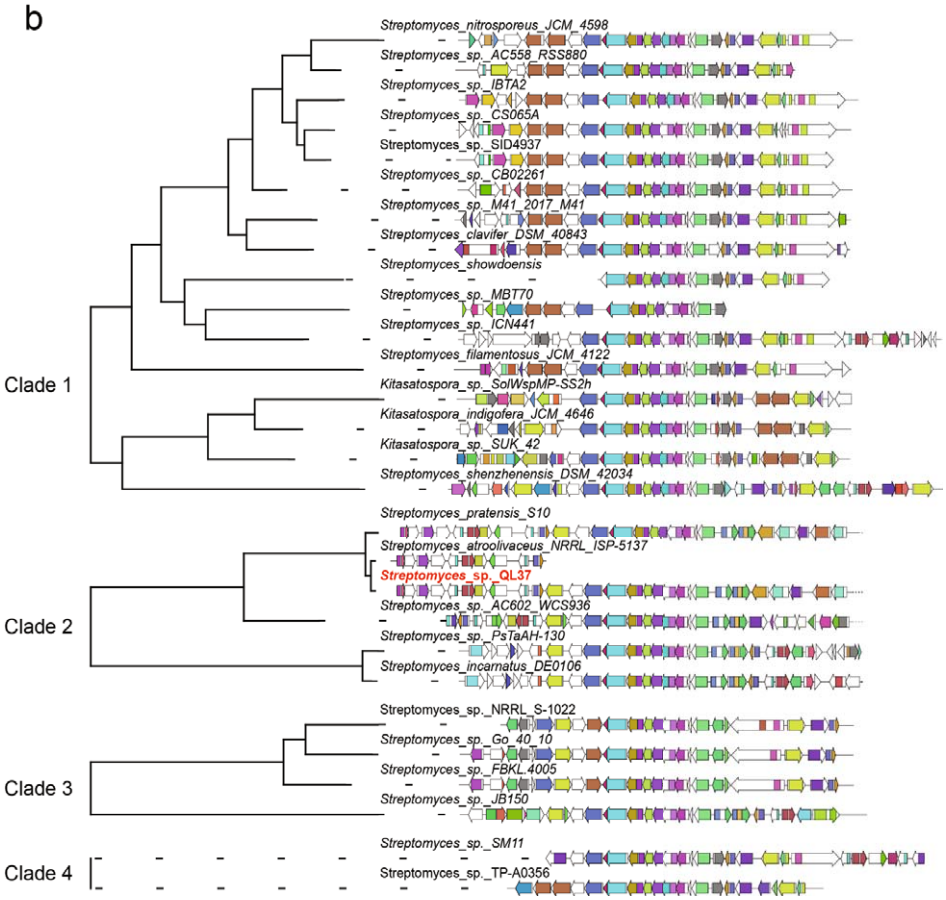


Figure 3. Bioinformatic analysis of *Streptomyces* and *Kitasatospora* genomes identifying BGCs for angucyclines and for *iso*-maleimycin-like molecules.

a. Phylogenetic tree of strains predicted to contain angucycline BGCs (similar to BGC12) and/or BGCs similar to the BGC for maleimycin-like molecules (BGC23a). The phylogenetic tree depicts the relationship among *Streptomyces* and *Kitasatospora* strains analyzed using PhyloPhlAn (version 3.0.60) (Asnicar *et al.*, 2020). Strains containing likely angucycline BGCs are highlighted in green, while those likely to contain BGCs similar to the BGC for maleimycin-like molecules are highlighted in blue. Strains harboring both types of BGCs are denoted in red. Note that co-occurrence of the two different BGCs is infrequent, with five co-occurrences found in our analysis. This co-occurrence is not related to phylogeny. b. Alignment of BGCs similar to the BGC (based on BiG-SCAPE) for maleimycin-like molecules in the analyzed genomes. Four distinct types of BGCs similar to the *iso*-maleimycin BGC23a were grouped by BiG-SCAPE, characterized by the conservation of core genes and their arrangement within the clusters.

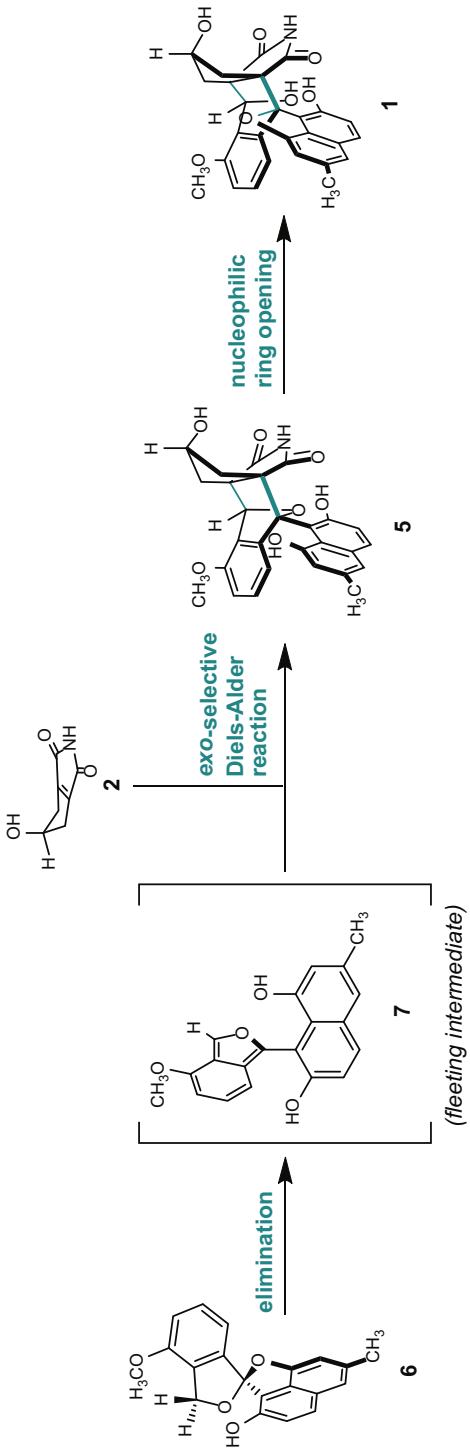


Figure 4. Lugdunomycin and its proposed bio (mimetic)synthesis from elmonin and *iso*-maleimycin.

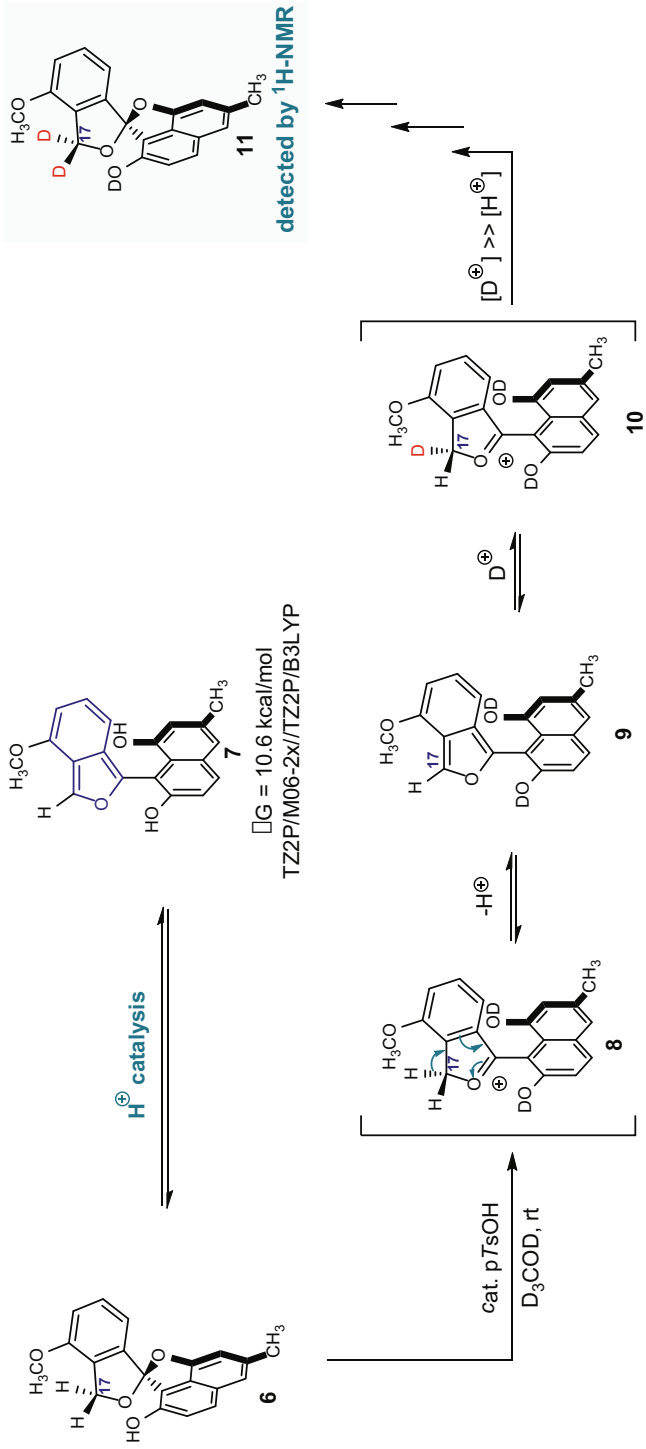


Figure 5. Acid treatment of elmonin leads to deuteration of C17, a result of reversible isobenzofuran formation.

After 24 h, a significant decrease of the signal at 5.30 ppm (corresponding to H17) was observed in ^{1}H-NMR (Fig. S6). This is the result of reversible deprotonation at C17 and supports the hypothesis that **6** is a masked isobenzofuran. Deuteration of the spiro-ketal leads to ring opening and elimination to give isobenzofuran **9**. As this process is reversible, the net result is deuteration **at** C17. A mechanistically similar formation of a quinodimethane under the same conditions was initially considered, but Density Functional Theory (DFT) calculations showed that isobenzofuran **7** is more than 15 kcal/mol lower in energy than the corresponding quinodimethane (Fig. S18). This can be underpinned with qualitative chemical arguments; in **7**, aromaticity is partly retained, whereas in a quinodimethane this is not the case.

A Diels-Alder reaction of *iso*-maleimycin and elmonin yields lugdunomycin and its diastereomers

Upon heating a mixture of elmonin **6** and *iso*-maleimycin **2** in *m*-xylene in a sealed tube at 110 ˚C, one main Diels-Alder product was formed. However, close inspection of the ^{1}H- and ^{13}C-NMR spectra revealed that the product obtained was not lugdunomycin, but instead an isomer (**12**, Fig. 6a; Figs. S9, S10 and S11).

With the help of X-ray crystallography, we established that four out of five stereocenters (16*S**, 17*R**, 19*S**, 21*S**) had identical relative configurations as compared to lugdunomycin (Fig. 6. and Fig. S4). However, the quaternary, fifth stereocenter at C9, being *R** in lugdunomycin,[6] was *S** in the synthetic material. This makes **12** the C9-epimer of lugdunomycin.

Attempts to convert **12** by isomerisation into the desired lugdunomycin were unsuccessful (Fig. 6b). We anticipated that protonation of the ether oxygen at C9 would lead to ring opening, providing the relatively stable carbenium ion **15**. Bond rotation would then provide the other conformer of **15**, which can subsequently undergo a ring closure, leading to **1**. However, when **12** was treated with *p*TsOH, no epimerization product was observed. Instead, it afforded product **14** in high yield, apparently formed due to electronically and entropically favoured dehydration. In addition, DFT calculations showed that the required bond rotation is not feasible because of hindered rotation (Fig. S20). Although the reaction mainly afforded the production of **12**, analysis of the reaction mixture by LC-MS also revealed small quantities of lugdunomycin, accompanied with up to six other products with identical mass and similar retention times, which probably represent diastereomers.

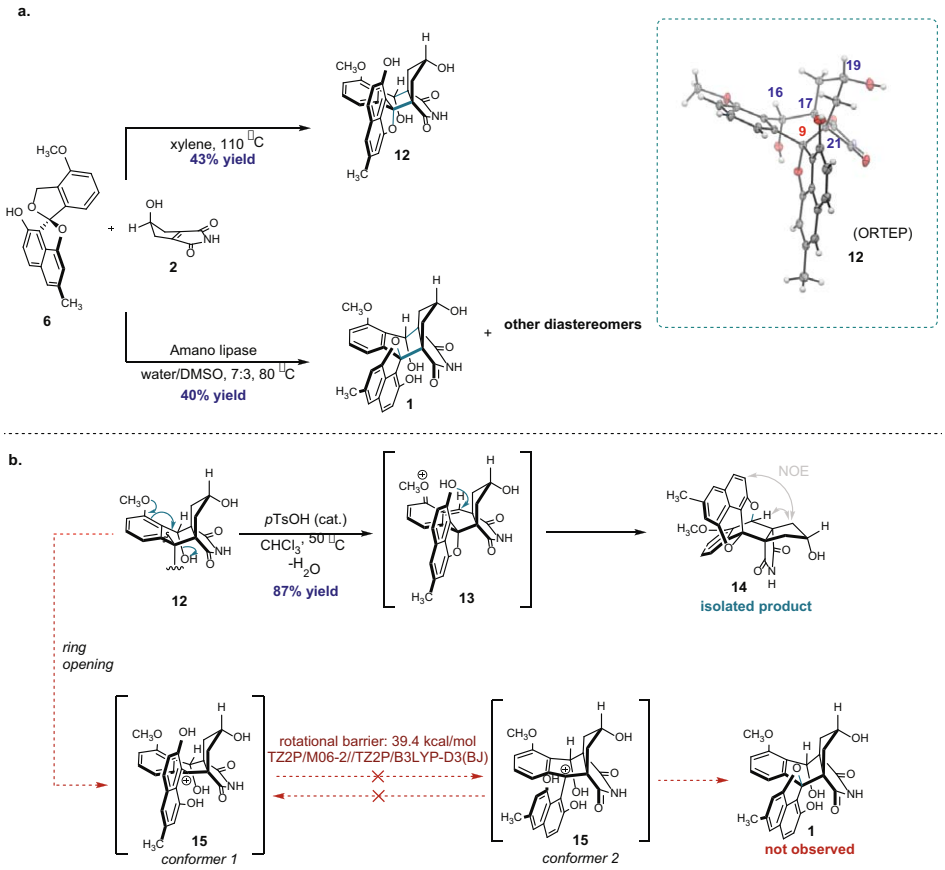
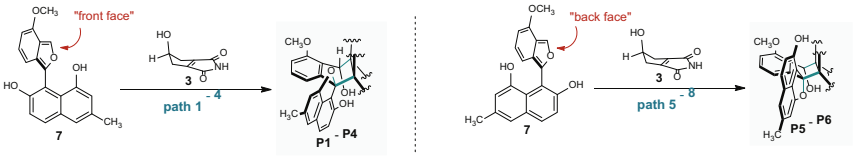


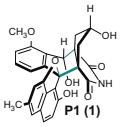
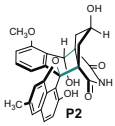
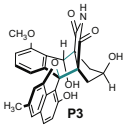
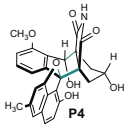
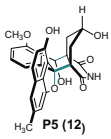
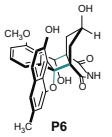
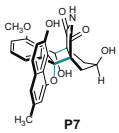
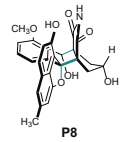
Figure 6. Synthesis experiments. a. Synthesis of epi-lugdunomycin and lugdunomycin via Diels-Alder reaction of elmonin and *iso*-maleimycin. b. Attempted isomerisation of epi-lugdunomycin 12 leads to dehydration instead of lugdunomycin formation.

We were puzzled by the fact that the Diels-Alder reaction in the bacterial host consistently led to lugdunomycin **1**, while during *in vitro* reaction, 9-*epi*-lugdunomycin **12** was obtained as the main product. This conflicts with the idea that lugdunomycin is biosynthesised *in vivo* in an undirected manner, *i.e.* without a steering factor. Involvement of an enzyme, a Diels-Alderase, was considered less likely, also because lugdunomycin crystallized from the isolate as the racemate. However, all attempts to mimic this reaction chemically, including a study of the effects of solvent, temperature, pH, the addition of surfactants or salts, led to just minor changes in the ratio of the diastereomeric products (Table S4).

In order to elucidate the cause of the discrepancy between the *in vivo* and *in vitro* synthesis of lugdunomycin and its diastereomers, all eight possible reaction pathways were calculated for the Diels-Alder reaction between *iso*-maleimycin and the isobenzofuran derived from elmonin (Table 1 and Table S6). From a mechanistic point of view, the dienophile can approach the isobenzofuran from two different faces, arbitrarily called “front face” and “back face”. Four diastereomers can be formed when the dienophile approaches from the “front face”, and another four when the dienophile approaches from the “back-face” (Table 1). Front face path 1 leads to lugdunomycin **P1** (**1**), and “back face” path 5 leads to 9-epi-lugdunomycin **P5** (**12**). The other paths lead to the other diastereomers. At the Density Functional Theory (DFT) level the relative Gibbs energy (ΔG) of the stationary points in paths 1 – 8 was calculated using the Amsterdam Modelling Suite (AMS) software package. (Komissarov *et al.*, 2021) Thermochemical properties were calculated at the TZ2P/M06-2x//TZ2P/B3LYP-D3 (BJ) level of theory, using a similar “hybrid” approach as described by Tang *et al* (Supplementary Information Section 7). (Meredith *et al.*, 2022)

Table 1. Calculated ΔG values for the formation of lugdunomycin and its diastereomers.



path 1 (*exo*)	path 2 (*exo*)	path 3 (*endo*)	path 4 (*endo*)
			
P1 (1)	P2	P3	P4
ΔG = -28 kcal/mol	ΔG = -26 kcal/mol	ΔG = -21 kcal/mol	ΔG = -24 kcal/mol
path 5 (*exo*)	**path 6 (*exo*)**	**path 7 (*endo*)**	**path 8 (*endo*)**
			
P5 (12)	P6	P7	P8
ΔG = -28 kcal/mol	ΔG = -26 kcal/mol	ΔG = -27 kcal/mol	ΔG = -24 kcal/mol

Evaluation of path 1 revealed that the Diels-Alder reaction likely is a true concerted [4+2] cycloaddition, having a moderate asynchronous character. Computation of the charge density and placing an electrostatic potential show that the charges are approximately equally distributed over the maleimide portion of the transition state. An alternative ionic two-step mechanism is therefore unlikely (Fig. S19). With regard to the thermochemical properties, all of the paths seem feasible at ambient temperatures, as the $\Delta G^{\ddagger}$ values for the reactions are ranging from 21 to 26 kcal/mol. Furthermore, the formation of all diastereomeric intermediates is thermodynamically favoured with ΔG values ranging from –0.3 kcal/mol to –14 kcal/mol (not shown, Table S5). Also, the formation of all final products, so after the ring closure, is highly favoured with Δ*G*values ranging from –21 for **P3** to –28 kcal/mol for **P1** and **P5**. Molecular modelling suggested that the large energy gap between the intermediates and the products, is due to strain relief upon ring-opening of the bridgehead ether moiety. These computations are in full agreement with the experimental observations, *e.g.* up to eight different isomeric products are observed by HPLC-MS in the *in vitro* Diels-Alder reaction. The small energy differences between the transition states (Table S6) further underline the agreement between the theoretical modelling and the experimental data. This explains the observation that changes in solvent and temperature slightly changed the diastereoselectivity of the reaction. It also imposed the conclusion that there is an additional "steering factor" leading to lugdunomycin production with high selectivity in *Streptomyces* sp. QL37.

Protein-based templating promotes lugdunomycin synthesis *in vitro*

So far we failed to identify an obvious candidate gene that would encode a Diels-Alderase in the genome of *Streptomyces* sp. QL37. An important property of such an enzyme would consist of a templating effect, by increasing the effective molarity of both *iso*-maleimycin and elmonin, which is needed considering the low concentrations of the reactants in the bacterium. Furthermore, suitable positioning of both reactants would explain the stereoselectivity of the reaction, not necessarily disputing the observation that lugdunomycin is formed with low or small enantioselectivity, considering its isolation as racemic crystals. In order to mimic this templating effect, we added proteins with a hydrophobic pocket or cleft to the *in vitro* Diels-Alder reaction. *Iso*-maleimycin, and in particular elmonin, have a limited solubility in water so offering a hydrophobic "shelter"

would possibly bring both compounds together. Upon adding proteins like BSA to the reaction we immediately saw a significant and positive effect; the diastereomeric ratio shifted in favour of lugdunomycin although not to such an extent that it became the main product. After extensive screening and optimization, it turned out that the use of Amano Lipase from *Pseudomonas fluorescens* in phosphate buffered saline and 30% DMSO at 80 °C, shifted the ratio of **1**/**12**/other from 2:6:1 to 2:2:1. This is a dramatic shift in favour of lugdunomycin and offered support for the template hypothesis, allowing completion of the chemical total synthesis of lugdunomycin. Scale-up of this protein-assisted reaction at 0.4 mmol provided 74 mg of lugdunomycin-diastereomers after column chromatography. Preparative HPLC subsequently provided 8.4 mg of pure lugdunomycin **1** which in all aspects was identical to the lugdunomycin isolated from cultures of *Streptomyces* sp. QL37 (Supplementary sections 3 and 5).

CONCLUSIONS

In this work we show that the angucyclinone elmonin (**6**) (Elsayed *et al.*, 2023), biosynthesized from BGC12, and *iso*-maleimycin (**2**) which is independently biosynthesized from BGC23a, are most likely the substrates required for biosynthesis of the highly rearranged angucyclinone lugdunomycin (**1**). These two biosynthetic intermediates provide the diene and the dienophile of an intermolecular Diels-Alder reaction that leads after subsequent ring closure to **1**. The observation that the diene and the dienophile in this biosynthesis originate from two different BGCs, effectively working in cooperation to generate lugdunomycin, presents an important concept with potentially far-reaching implications. The canonical view involves a single gene cluster that is responsible for the biosynthesis of a molecule or a family of molecules. For one, heterologous expression is becoming more and more feasible for the analysis of BGCs and the identification of their cognate natural products. Expressing BGC12 would result in production of angucyclines, but lugdunomycin would remain undiscovered as it also requires BGC23a. Such collaboration between two or even more BGCs may be more common than currently believed, and our laboratory found several examples.

Intermolecular Diels-Alder reactions are rare in biosynthesis. Among the very few well-documented cases are the biosynthesis of paracaseolide A (Noutsias &

Vassilikogiannakis, 2012) and the flavonoid chalcomoracin (Wang & Hoye, 2015), with the latter catalysed by a *bona fide* Diels-Alderase. Elmonin (**6**), a 1-alkoxy 1,3-dihydroisobenzofuran, acts as a masked isobenzofuran that is formed upon an acid catalysed intramolecular elimination reaction and undergoes a Diels-Alder reaction with the dienophile *iso*-maleimycin. This Diels-Alder reaction is impressive from a reactivity point of view. It involves a tetra-substituted alkene which is a poor dienophile because of its steric encumbrance. The isobenzofuran is a reactive diene but only present as a fleeting intermediate, in particular in a cellular context. The Diels-Alder reaction produces, from two achiral starting materials, after subsequent ring-opening-ring closure lugdunomycin with five stereocenters, formed with high diastereoselectivity.

Biosynthetic intermolecular Diels-Alder reactions may be spontaneous, such as reported for the Diels-Alder reaction in the biosynthesis of paracaseolide A (Noutsias & Vassilikogiannakis, 2012), or catalyzed by an enzyme, such as during the biosynthesis of chalcomoracin (Wang & Hoye, 2015). It is likely that the Diels-Alder reaction that led to lugdunomycin biosynthesis in *Streptomyces* sp. QL37 is templated. This is suggested by the high activation barrier and the low concentration of both reactants, exacerbated by the thermodynamically uphill isobenzofuran formation. The main product of the non-templated reaction (which occurs *in vitro* at higher temperature and at much higher concentrations) is not lugdunomycin but a related diastereomer. Upon addition of proteins with a hydrophobic pocket or cleft, the reaction can be steered in an impressive way to produce lugdunomycin. The addition of micelle-forming surfactants did not induce this effect, and taken together this is provides strong support for the involvement of an enzyme. Natural lugdunomycin, however, is found as a racemate, or a scalemic mixture at best, and a candidate Diels-Alderase (gene) has not been identified. It is tempting to propose that the reaction is catalysed by an “opportunistic Diels-Alderase”. (Bitchagno *et al.*, 2022, Zheng *et al.*, 2016)

The role of cross-talk and regulatory mechanisms in the biosynthesis of lugdunomycin should be further investigated. The discovery of the dependence on two BGCs - and thus two precursors - emphasizes the importance of the study of genomics and structure/reactivity properties of secondary metabolites within one organism. To the best of our knowledge, this is the first case whereby biomimetic chemical synthesis, computational chemistry and genomics have been combined, allowing us to shed important new light on the biosynthesis of a molecule with highly complex chemical scaffold. This emphasizes the need to

consider unconventional biosynthetic relationships and molecular interactions that may drive the synthesis of complex secondary metabolites. In addition, our work shows that unravelling the genetic elements and regulatory mechanisms governing the cooperation between BGCs, combined with chemical mechanistic insights, can help provide novel insights into the biosynthetic machinery responsible for the biosynthesis of natural products of interest. Such novel biosynthetic insights are also key to pathway engineering and synthetic biology approaches, enabling the production of (families of) compounds at higher yields or generation of novel analogues with enhanced properties. These approaches should help scientist to further explore and uncover the complexity of microbial secondary metabolism.

Materials and methods

Microorganisms and culturing conditions

All media and routine *Streptomyces* techniques were done following routine protocols (Kieser, 2000). *Streptomyces* sp. QL37 was obtained from soil collected in the Qinling mountains of the People's Republic of China, as previously described (Zhu *et al.*, 2014, Wu *et al.*, 2019b). The strain was deposited in the Centraal Bureau voor Schimmelcultures (CBS) in Utrecht, The Netherlands, with the accession number 138593.For solid cultivation, *Streptomyces* sp. QL37 was grown confluently on minimal media (MM) agar plates containing 1% glycerol and 0.5 % mannitol (w/v) as the carbon sources. Liquid fermentation was performed in 100 mL Erlenmeyer flasks (Fisher Scientific) were 20 mL of liquid minimal medium (MM without agar) supplemented with R5 trace elements solution (Kieser, 2000) (4mL/L), glycerol (1% w/v) and/or mannitol (0.5% w/v), after autoclaving. Other carbon sources were added at 1% (w/v) prior inoculation.

Preparation of crude extracts and metabolite profiling

In the case of solid fermentations, 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 h (Elsayed *et al.*, 2023). After evaporation extracts were dissolved in methanol (MeOH) to a final concentration of 1 mg/mL. Spent media of liquid-grown cultures were extracted after five days of incubation using HP-20 beads (Supelco) previously soaked in MeOH overnight. Approximately 1 g of HP-20

beads was added to each flask and incubated O/N while shaking. Extracts were prepared as described (Nunez Santiago *et al.*, 2024) and dissolved in methanol (MeOH) to a final concentration of 1 mg/mL.

For metabolite profiling of the crude extracts, liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis was performed using a Shimadzu Nexera X2 ultra high-performance liquid chromatography (UPLC) system, equipped with a photodiode array detector (PDA), coupled to a Shimadzu 9030 QTOF mass spectrometer, equipped with an electrospray ionization (ESI) source unit, which included a calibrant delivery system (CDS). For details see Supplementary Section 2.

To allow the detection of *iso*-maleimycin using LC-MS in the complex bacterial extract, we developed the following method. As standard we used 0.01 mg/mL of synthesized *iso*-maleimycin (Uiterweerd *et al.*, 2020) dissolved in methanol. Then, a multiple reaction monitoring (MRM) method was developed targeting the $[M+CH_3OH+H]^+$ ion of *iso*-maleimycin (186.0759) and the characteristic MS/MS fragment of 109.0283.

Proteomics experiments

Proteomics samples were prepared as previously described (van Bergeijk *et al.*, 2022, Gubbens *et al.*, 2014). See Supplementary Section 2. The desalted peptides solution was separated on an UltiMate 3000 RSLCnano system set in a trap-elute configuration with a nanoEase M/Z Symmetry C18 100 Å, 5 µm, 180 µm × 20 mm (Waters) trap column for peptide loading/retention and nanoEase M/Z HSS C18 T3 100 Å, 1.8 µm, 75 µm × 250 mm (Waters) analytical column for peptide separation. The eluent was introduced by electro-spray ionisation (ESI) via the nanoESI source (Thermo) to QExactive HF (Thermo Scientific) operated in positive mode and mass range 350-1400 m·z-1 at 120,000 resolution. For individual peaks, the data dependent intensity threshold of 2.0 × 104 was applied for triggering an MS/MS event, isotope exclusion on and dynamic exclusion was 20 s. Unassigned, +1 and charges >+8 were excluded with peptide match mode preferred. For MS/MS events, the loop count was set to 10, isolation window at 1.6 m·z-1, resolution at 15,000, fixed first mass of 120 m·z-1, and normalized collision energy (NCE) at 28 eV. The obtained raw data was processed using MaxQuant version 2.1.0.0 (Prianichnikov *et al.*, 2020) with default label free quantification settings. For further details see Supplementary Section 2.

Creation of a mutant of the *lysWrimK* genes in the BGC23

The deletion of the *lysWrimK* genes was done based on a previously reported method (Swiatek *et al.*, 2012a). Briefly, the construct for gene disruption was obtained by amplification of ~1.5 kb regions up- and downstream of the chromosome of *Streptomyces* sp. QL37 using primer pairs 1+2 and 3+4 (Table S2) and cloned as EcoRI-XbaI and XbaI-BamHI fragments into a derivative of the unstable multi-copy plasmid pWHM3 (Vara *et al.*, 1989) that harbours oriT to allow transference by conjugation. 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 (Zhou *et al.*, 2012), and subsequently introduced into *Streptomyces* sp. QL37, following the protocol as described (Kieser, 2000). The correct mutant was selected by resistance to apramycin (50 µg/mL) and sensitivity to thiostrepton (10 µg/mL). Polymerase chain reactions (PCR) were performed on a T100 Thermal Cycler (Bio-Rad, Hercules, CA, USA). The ZR Plasmid Miniprep-Classic kit (Zymo Research, Irvine, CA, USA) was used for plasmid extraction.

Heterologous expression of lugdunomycin in *Streptomyces coelicolor* M1152

For the heterologous expression of BGC23, the BGC was divided into three fragments, each amplified by PCR using the chromosome of *Streptomyces* sp. QL37 as the template. Fragment 1 was amplified with primers BGC23_P1 and BGC23_P2, fragment 2 with primers BGC23_P3 and BGC23_P4, and fragment 3 with primers BGC23_P5 and BGC23_P6 (Table S2). The final construct was assembled by Gibson assembly to combine fragments 1, 2, and 3 with the vector pSET152, which integrates at the øC31 attachment site on the *S. coelicolor* chromosome (van Bergeijk *et al.*, 2022) to create pSET152-BGC23. To generate pMS82-BGC23, the insert was obtained as an HpaI-EcoRV fragment from pSET152-BGC23 and cloned into EcoRV-digested pMS82 (Gregory *et al.*, 2003). The integrity of the constructs was verified using the restriction enzyme combinations KpnI, KpnI and SpeI, ClaI and EcoRV. Non-methylated pMS82-BGC23 was isolated from *E. coli* strain ET12567 and introduced into *S. coelicolor* M1152-*lug* (van der Heul, 2022) via protoplast transformation Correct transformants were checked via PCR using oligonucleotides LysWRimK_check_R2 and WK_check_F primers (Table S2).

Acknowledgements

We are grateful to Changsheng Wu and Hermen Overkleeft for discussions. The work was supported by the NACTAR program of The Netherlands Organization for Scientific Research (NWO), Grant 16439 to G.P.vW and A.J.M. The bioinformatic experiments presented in this work were performed using the compute resources from the Academic Leiden Interdisciplinary Cluster Environment (ALICE) provided by Leiden University.

Supplementary information

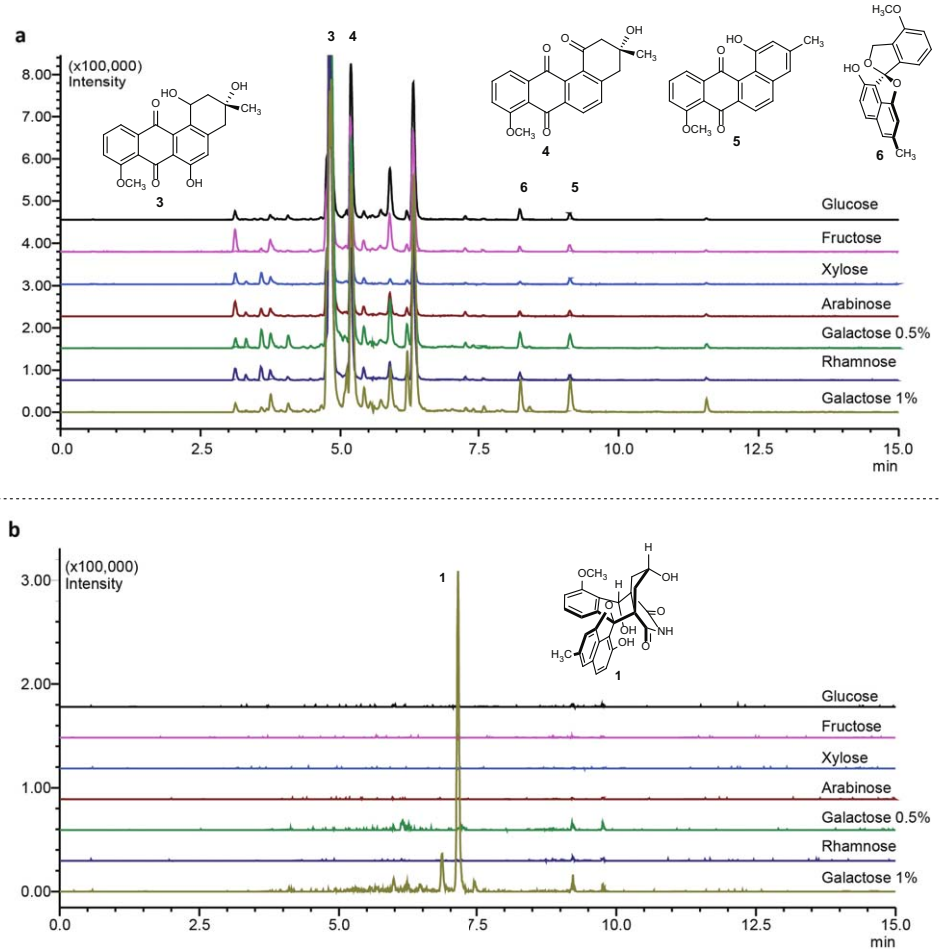


Figure S1. Carbon-source dependent production of angucyclines and lugdunomycin. a. Extracted ion chromatogram (XIC) for some previously identified angucyclinones produced by *Streptomyces* sp. QL37. The XIC correspond to the following m/z: 355.117 (**3**), 339.123 (**4**), 319.096 (**5**), 321.111 (**6**) with a tolerance of 20 ppm. **b**. XIC for 456.144, the m/z of **1** $[M+H-H_2O]^+$ identified in *Streptomyces* sp. QL37 extracts. *Streptomyces* cultures were grown for 7 days in liquid MM supplemented with carbon sources as indicated. Note that lugdunomycin was only observed when cultures contained larger amounts of galactose.

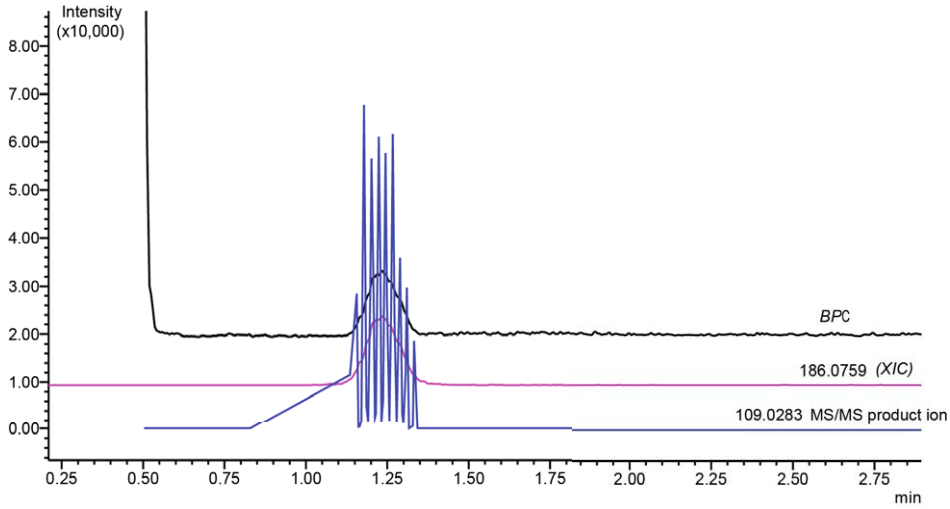


Figure S2. LC-MS chromatogram of chemically synthesized *iso*-maleimycin. LC-MS chromatogram depicting the analysis of a standard solution of chemically synthesized *iso*-maleimycin 2 (0.01 mg/mL, black) dissolved in MeOH. We noted that *iso*-maleimycin reacts with MeOH and the molecular ion $[M+CH3OH+H]^+$ was established at m/z 186.0759 (pink), and a unique product ion at m/z 109.0283 was selected for detection (blue). The retention time (RT) for *iso*-maleimycin was determined to be 1.25 min.

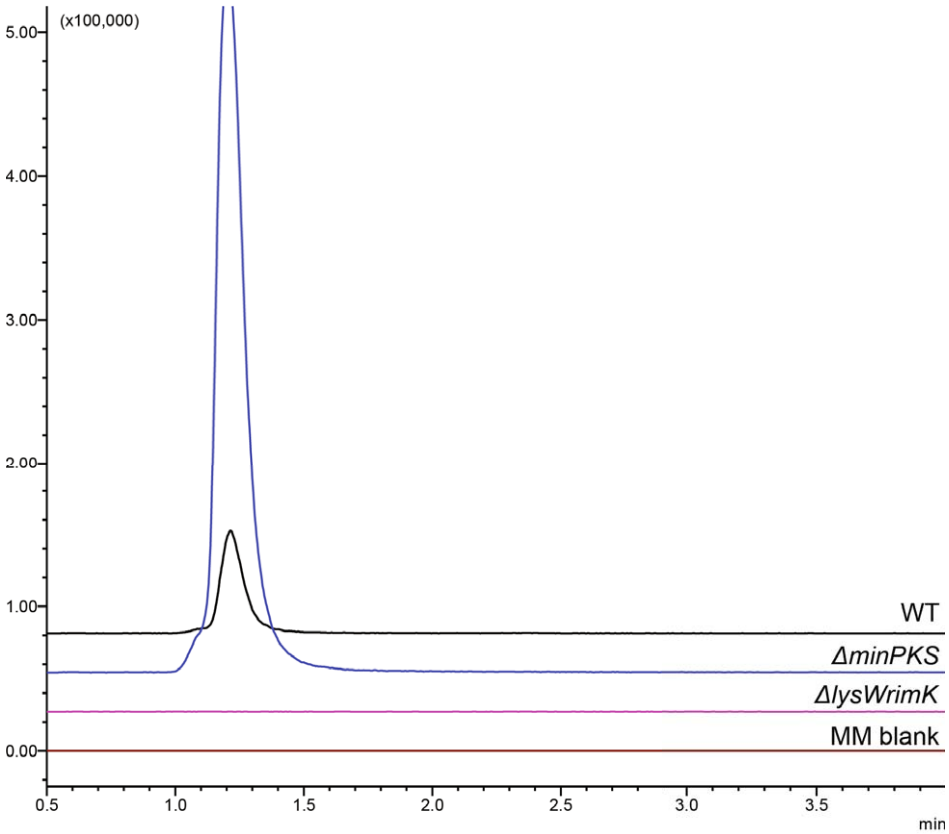


Figure S3. LC-MS/MS chromatograms of *iso*-maleimycin in *Streptomyces* sp. QL37 wild-type, Δ*minPKS* and Δ*lysWrimK* mutant strains. LC-MS/MS chromatograms utilizing multiple reaction monitoring (MRM) searching for the presence of *iso*-maleimycin 2 in various samples: *Streptomyces* sp. QL37 wild-type (WT), Δ*minPKS* and Δ*lysWrimK* mutant. Notably, the *iso*-maleimycin MeOH adduct peak, detected at a retention time of 1.25 min, is observed in the wild-type strain and *ΔminPKS* but is absent in the Δ*lysWrimK* strain. A sample of extract from not inoculated minimal medium was included as a control (MM blank).

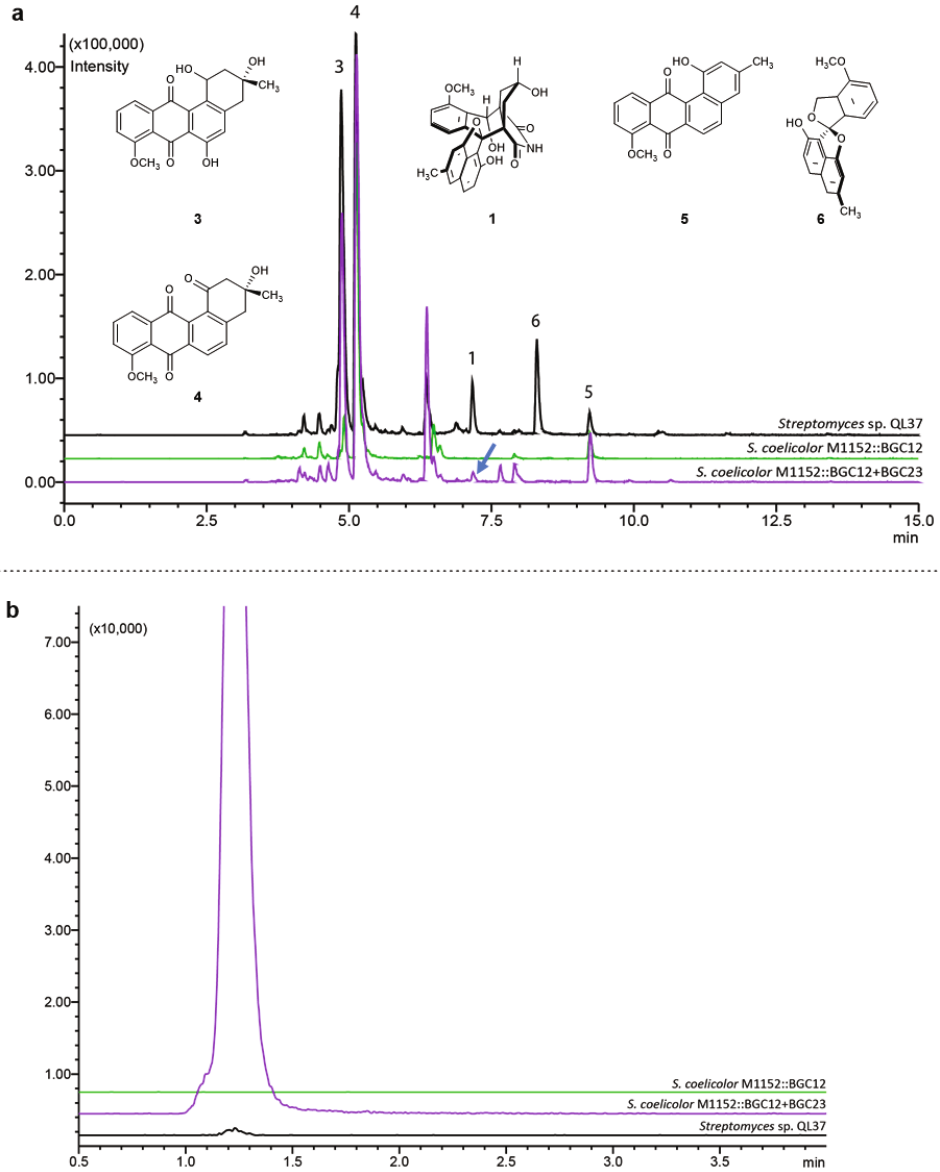


Figure S4. Comparative LC-MS chromatograms of BGC12 and BGC12+BGC23 heterologous expression in *S.coelicolor* M1152 and *Streptomyces* sp. QL37. (a) Extracted ion chromatograms (XIC) for angucyclinones and lugdunomycin produced by control strain *Streptomyces* sp. QL37, corresponding to the following m/z values: 456.144 (1), 355.117 (3), 339.123 (4), 319.096 (5), and 321.111 (6), with a tolerance of 20 ppm. *S. coelicolor* M1152 containing BGC12 (M1152::BGC12) produces angucyclines but not lugdunomycin 1. Production of 1 (see arrow) was seen when both BGCs were expressed heterologously in *S. coelicolor* M1152 (*S. coelicolor* M1152::BGC12+BGC23). (b) LC-MS/MS chromatograms using multiple reaction monitoring (MRM) to detect *iso*-maleimycin. Heterologous strains containing BGC23 show a significant peak of *iso*-maleimycin.

Table S1: Differential protein expression between galactose- and glucose-grown MM cultures of *Streptomyces* sp. QL37.

Proteins annotation #	log2FC ^
QL37_30735 GMP synthase [glutamine-hydrolyzing]	**6.0218079**
QL37_36965 ABC transporter periplasmic-binding protein YtfQ	5.2994111
QL37_08055 hypothetical protein	4.6876532
QL37_15965 Galactokinase	4.4700264
QL37_11300 Multiple sugar-binding periplasmic protein SbpA	4.4042577
QL37_11810 Pyruvate, phosphate dikinase	4.3336984
QL37_30710 2-isopropylmalate synthase	**4.2639892**
QL37_07245 hypothetical protein	4.1272887
QL37_15975 Galactose-1-phosphate uridylyltransferase	4.0765178
QL37_30755 hypothetical protein	**3.9478366**
QL37_30705 Acyl-CoA dehydrogenase	**3.6859316**
QL37_07695 hypothetical protein	3.4194662
QL37_09770 Transcriptional regulator SlyA	3.1588078
QL37_30745 Short-chain-fatty-acid--CoA ligase	**2.6588961**
QL37_08755 Putative prophage major tail sheath protein	2.5847662
QL37_07175 Benzaldehyde dehydrogenase [NAD (+)]	2.5527653
QL37_30750 Aspartate aminotransferase	**2.5430627**
QL37_09775 hypothetical protein	2.5100967
QL37_19275 Acireductone dioxygenase	2.4473463
QL37_30720 Alpha-aminoadipate--LysW ligase LysX	**2.2967689**
QL37_03115 hypothetical protein	2.2783263
QL37_32225 Glycerol kinase	2.2180443
QL37_20690 Daunorubicin/doxorubicin resistance ATP-binding protein DrrA	2.1672179
QL37_36990 Aldose 1-epimerase	2.1152377
QL37_25910 ATP synthase subunit a	2.0875887
QL37_25595 hypothetical protein	2.044777
QL37_25085 Succinyl-diaminopimelate desuccinylase	-2.90E-05
QL37_17640 hypothetical protein	-2.067939
QL37_20340 Putative oxidoreductase YdbC	-2.087165
QL37_29760 hypothetical protein	-2.18598
QL37_24510 hypothetical protein	-2.231756
QL37_33135 putative peptidyl-prolyl cis-trans isomerase B	-2.285745
QL37_26270 hypothetical protein	-2.553681
QL37_12020 30S ribosomal protein S20	-2.566372
QL37_10455 Sugar phosphatase YfbT	-2.666196
QL37_22665 30S ribosomal protein S12	-2.728723
QL37_19985 hypothetical protein	-2.79044
QL37_27255 D-xylose-proton symporter	-3.395676

In bold, the proteins belonging to the predicted BGC23 by antiSMASH.
^ Positive fold change indicate upregulation in galactose and negative fold change indicates upregulation in glucose (n=4).

Table S2. Primer Sequences for knock-out constructs and verifications of mutants ("check") in BGC23 of *Streptomyces* sp. QL37.

Primer name	Sequence
LysWRimK_check_R2	CCGACGATTCCTGCTCGGATC
LysWRimK_LF_FW2	CTAGGAATTCGCCGATTTATCGGACGAATG
LysWRimK_LF_RV2	CTAGTCTAGACGGCTCAGACGTTCTGGAAG
LysWRimK_RF_FW	CTAGTCTAGATCCACGACCGACACAGACATC
LysWRimK_RF_RV	CTAGAAGCTTCGACGAGCTGCTTGTCCTTC
WK_check_F	CGAACTCGTCGGCAACACACG
BGC23_P1	CAGGAAACAGCTATGACATGATTACGAATTCGATATCGTTCGGGCT GTGCGAGCACCA
BGC23_P2	CAGGTTCGCGTGATGGCGGTTCC
BGC23_P3	GGAACCGCCATCACGCGAACCTG
BGC23_P4	GTGCTACGGGCGGTTTCGGAGG
BGC23_P5	CCTCCGAAACCGCCCGTAGCAC
BGC23_P6	GCTTGGGCTGCAGGTCGACTCTAGAGCTGGGACGGAGGCATGG AGAC

Section 2: Supplemental methods

Extended Metabolomics Methods

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) was performed using a Shimadzu Nexera X2 ultra high-performance liquid chromatography (UPLC) system, equipped with a photodiode array detector (PDA), coupled to a Shimadzu 9030 QTOF mass spectrometer, equipped with an electrospray ionization (ESI) source unit, which included a calibrant delivery system (CDS). For the LC separation, 2 µL of the sample was injected into a Waters Acquity HSS C18 column (1.8 µm, 100 Å, 2.1 × 100 mm). The column temperature was maintained at 30 °C, and the separation was carried out at a flow rate of 0.5 mL/min. Solvent A consisted of 0.1% formic acid in H_2O, while solvent B comprised 0.1% formic acid in acetonitrile (ACN). The gradient elution profile started with 5% B for 1 min, followed by a linear increase from 5% to 85% B over 9 min, then a steep increase to 100% B over 1 min, and finally, an isocratic hold at 100% B for 4 min. To re-equilibrate the column, it was set to 5% B for 3 min before the subsequent run (Uiterweerd & Minnaard, 2022, Uiterweerd *et al.*, 2020).

The PDA acquisition was conducted within the wavelength range of 200–600 nm, with a scan rate of 4.2 Hz and a slit width of 1.2 nm. The temperature of the flow cell was maintained at 40 °C throughout the analysis. All samples were

analyzed in positive polarity mode using a data-dependent acquisition method. This involved acquiring full scan MS spectra within the m/z range of 100–1700 at a scan rate of 10 Hz, with intelligent data (ID) enabled. Subsequently, two data-dependent MS/MS spectra within the same m/z range and scan rate were obtained for the two most intense ions detected in each scan. Collision-induced dissociation (CID) was employed for ion fragmentation with a fixed collision energy (CE) of 20 eV. To avoid re-analysis of previously fragmented ions, they were excluded from selection for 1 s before being eligible for fragmentation again. The parameters used for the ESI source were as follows: interface voltage of 4 kV, interface temperature of 300 °C, nebulizing gas flow rate of 3 L/min, and drying gas flow rate of 10 L/min (Uiterweerd *et al.*, 2020).

Extended Proteomics Methods

To obtain samples for proteomics analysis, bacterial cultures were pelleted and lysed in 100 µL lysis buffer (4% SDS, 100 mM Tris-HCl pH 7.6, 50 mM EDTA) and disrupted by sonication. Total protein was precipitated using the chloroform-methanol method (Wu *et al.*, 2019b), and the proteins dissolved in 0.1% RapiGest SF surfactant (Waters) at 95°C. The protein concentration was measured at this step using the BCA method. Protein samples were then reduced by adding 5 mM DTT and incubate at 60°C for 30 min, followed by thiol group protection with 21.6 mM iodoacetamide incubation at room temperature in dark for 30 min. Then 0.1 µg trypsin (recombinant, proteomics grade, Roche) per 10 µg protein was added, and samples were digested at 37°C overnight. After digestion, trifluoroacetic acid was added to 0.5% and samples were incubated at 37°C for 30 min followed by centrifugation to degrade and remove RapiGest SF. Peptide solution containing 6 µg peptide was then cleaned and desalted using STAGE-Tips (Te Velde *et al.*, 2001). Briefly, 6 µg of peptide was loaded on a conditioned StageTip with 2 pieces of 1 mm diameter C18 disk (Empore, product number 2215), washed twice with 0.5% formic acid solution, and eluted with elution solution (80% acetonitrile, 0.5% formic acid). Acetonitrile was then evaporated in a SpeedVac. Final peptide concentration was adjusted to 40 ng·µL-1 using sample solution (3% acetonitrile, 0.5% formic acid) for analysis.

The desalted peptides solution weres separated on an UltiMate 3000 RSLCnano system set in a trap-elute configuration with a nanoEase M/Z Symmetry C18 100 Å, 5 µm, 180 µm × 20 mm (Waters) trap column for peptide loading/retention and nanoEase M/Z HSS C18 T3 100 Å, 1.8 µm, 75 µm ×

250 mm (Waters) analytical column for peptide separation. Mobile phase A composed of 0.1% formic acid (FA) in ULC-MS grade H2O (Biosolve), while mobile phase B composed of 0.1% FA, 10% H2O in ULC-MS grade acetonitrile (ACN, Biosolve). The flow gradients used for analysis was a shallow 113 min gradient of mobile phase A and B controlled by a flow sensor at 0.3 μL·min-1. The gradient was programmed with linear increment from 1% to 5% B from 0 to 2 min, 5% to 13% B from 2 to 63 min, 13% to 22% B from 63 to 85 min, 22% to 40% B from 85 to 104 min, 90% at 105 min and kept at 90% to 113 min.

The eluent was introduced by electro-spray ionisation (ESI) via the nanoESI source (Thermo) to QExactive HF (Thermo Scientific). The QExactive HF was operated in positive mode with data dependent acquisition. The MS survey scan was set with mass range 350-1400 m·z-1 at 120,000 resolution. For individual peaks, the data dependent intensity threshold of 2.0 × 104 was applied for triggering an MS/MS event, isotope exclusion on and dynamic exclusion was 20 s. Unassigned, +1 and charges >+8 were excluded with peptide match mode preferred. For MS/MS events, the loop count was set to 10, isolation window at 1.6 m·z-1, resolution at 15,000, fixed first mass of 120 m·z-1, and normalised collision energy (NCE) at 28 eV.

The obtained raw data was processed using MaxQuant version 2.1.0.0b (Rüger *et al.*) with default label free quantification settings.

Organic chemistry was done at Groningen University by Dr. Michiel Uiterweerd and Prof. Dr. Adri Minnaard. For this thesis version, details on SI of organic chemistry were left out. However, all details can be found in the published version.