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

Heul, H.U. van der

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.

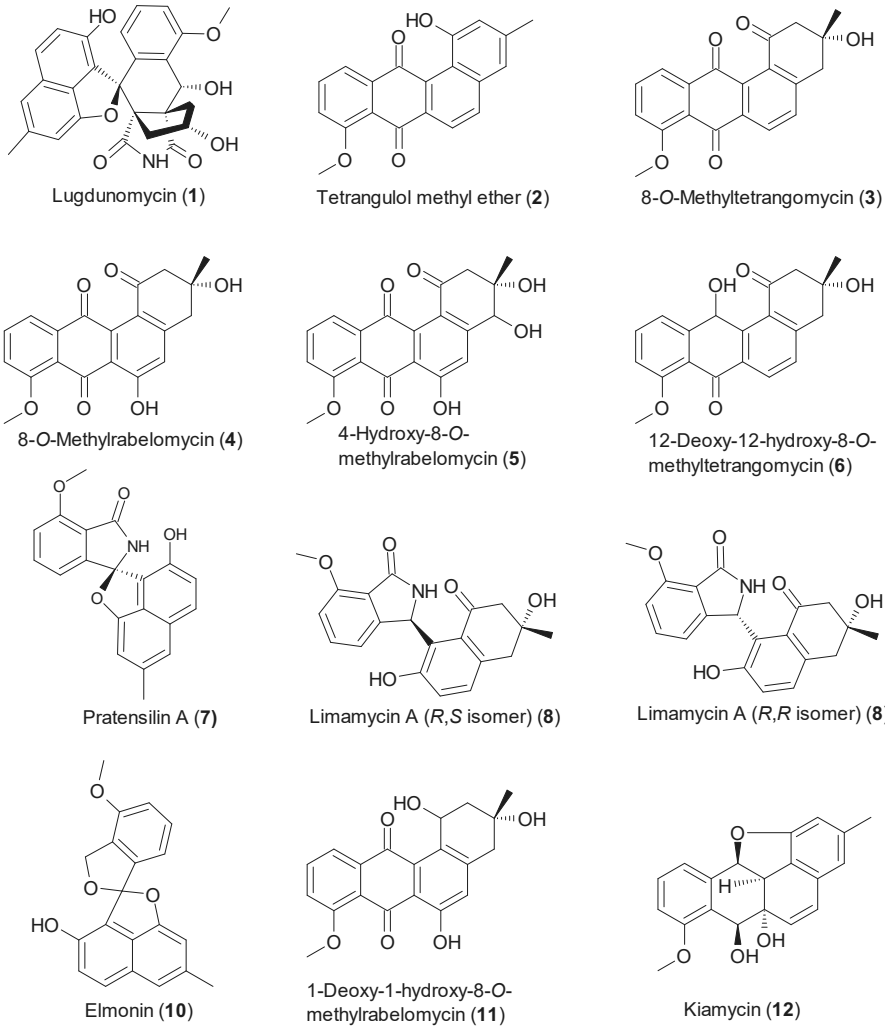


Figure 1 Structures of molecules discussed in this study.
Lugdunomycin (**1**), the non-rearranged angucyclines (**2**, **3**, **4**, **5** and **6**), pratensilin A (**7**) and the limamycins (**8**, **9**) were previously isolated from *Streptomyces* sp. QL37 (Wu *et al.*, 2019). Elmonin (**10**) was 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.

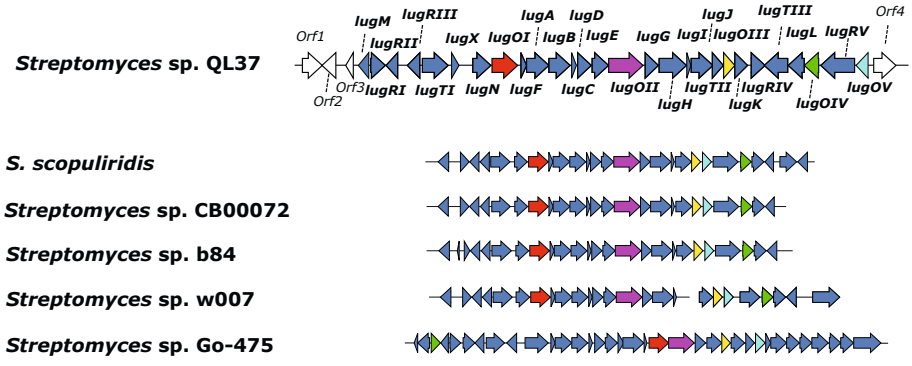


Figure 2A Phylogenetic analysis of oxygenases encoded by angucycline BGCs. Comparison of the *lug* gene cluster with related clusters from other *Streptomyces* strains.

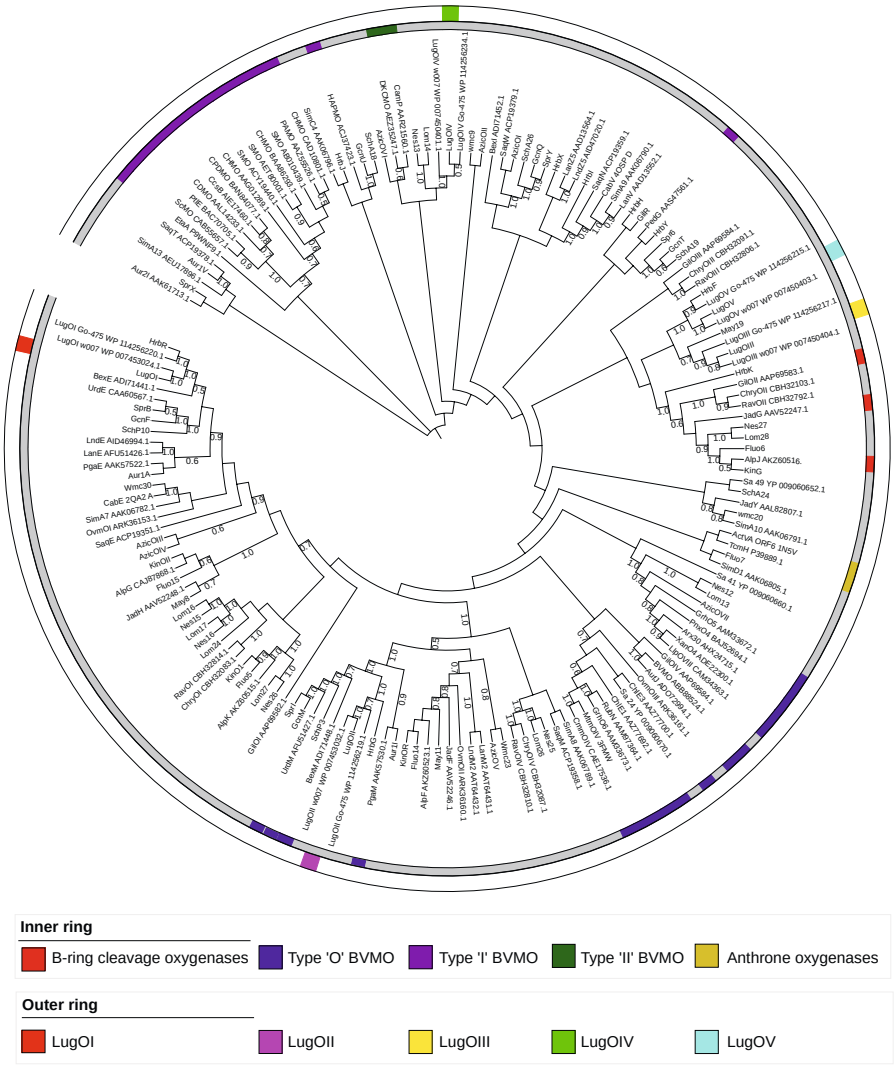


Figure 2B Phylogenetic analysis of oxygenases encoded by angucycline BGCs. Maximum-likelihood tree of oxygenases encoded by all known angucycline BGCs and BVMOs known to be involved in ring opening of aromatic compounds. The analysis involved 169 amino acid sequences and in total 1480 positions. The 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.

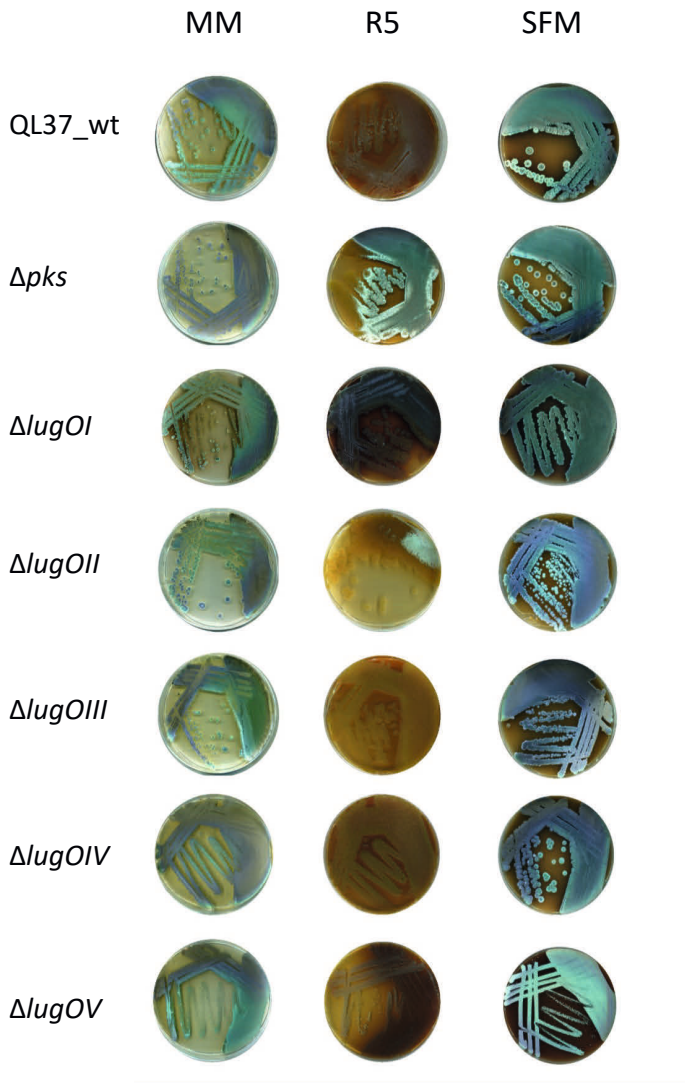


Figure 3 *Streptomyces* sp. QL37, and its *lug-pks* and *lugOI–OV* mutants on different media show differential pigmentation, suggesting differential angucycline production.
The strains were grown on MM (Minimal medium with 0.5% mannitol and 1% glycerol), on R5 (R5 supplemented with 0.8% peptone and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

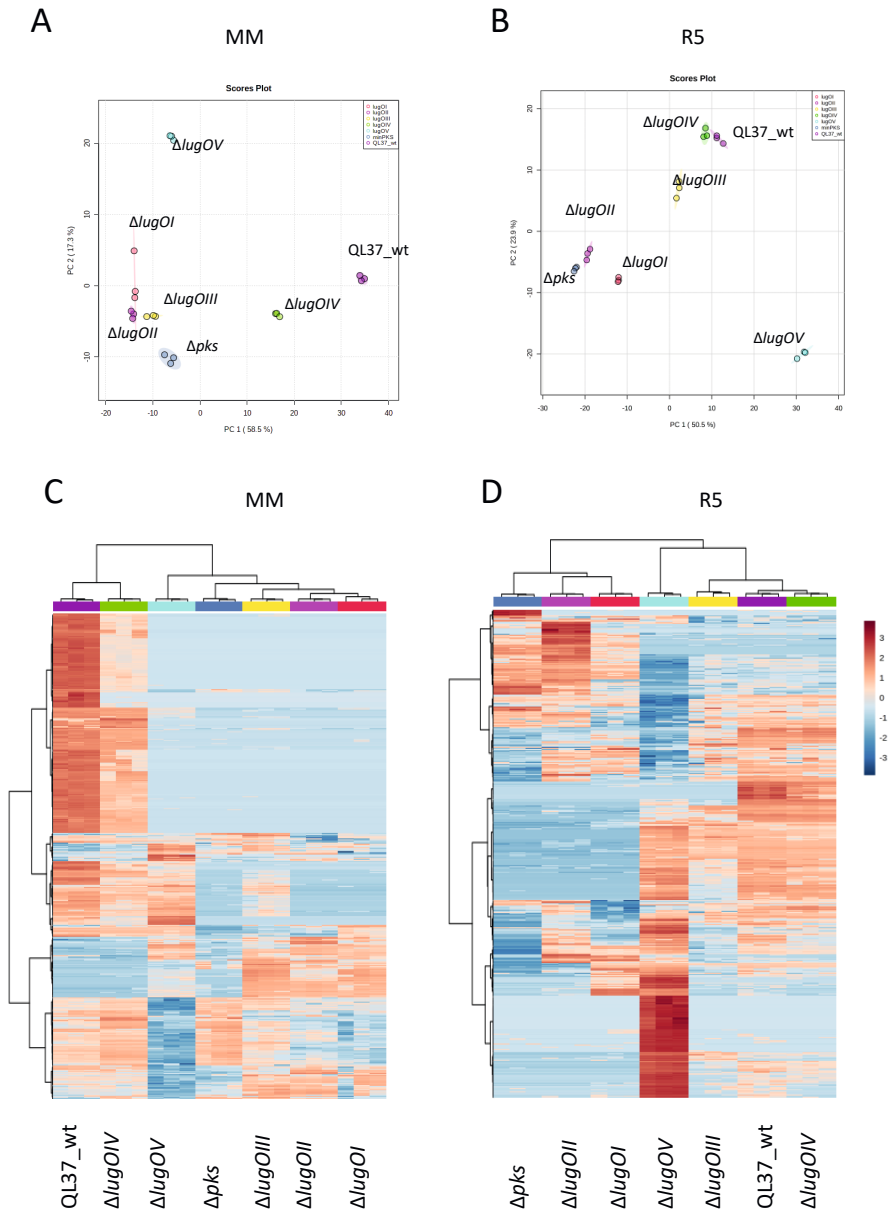


Figure 4 Exploratory statistical analysis of the metabolomic profiles of *Streptomyces* sp. QL37, and its *lugOI–OV* and *pks* mutants grown on MM and R5 agar.
Principle component analysis (PCA) score plots of the extracts from the various strains grown on MM (A) and R5 (B). Heatmaps with additional 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.

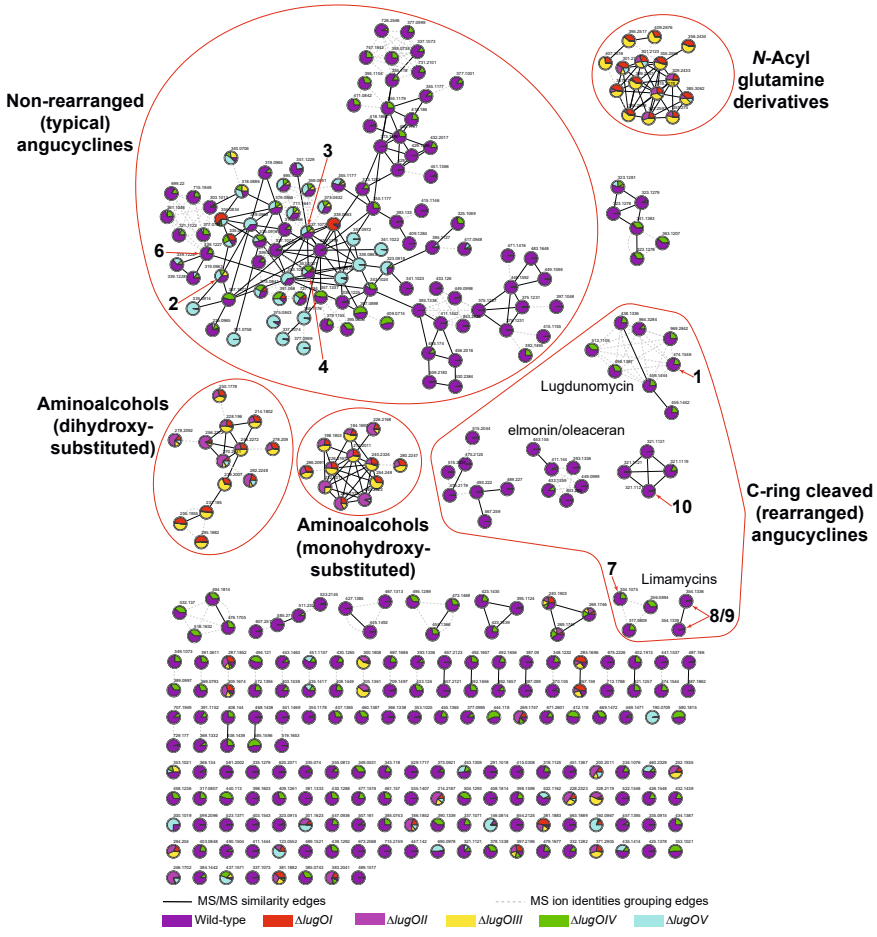


Figure 5 Molecular network of the ions detected in the extracts of *Streptomyces*. sp. QL37 and its *lugOI–lugOV* mutants grown on MM agar.
The nodes are labelled by the precursor mass of their ions and pie charts are mapped to the nodes to indicate the relative intensities of the ions in the different 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.

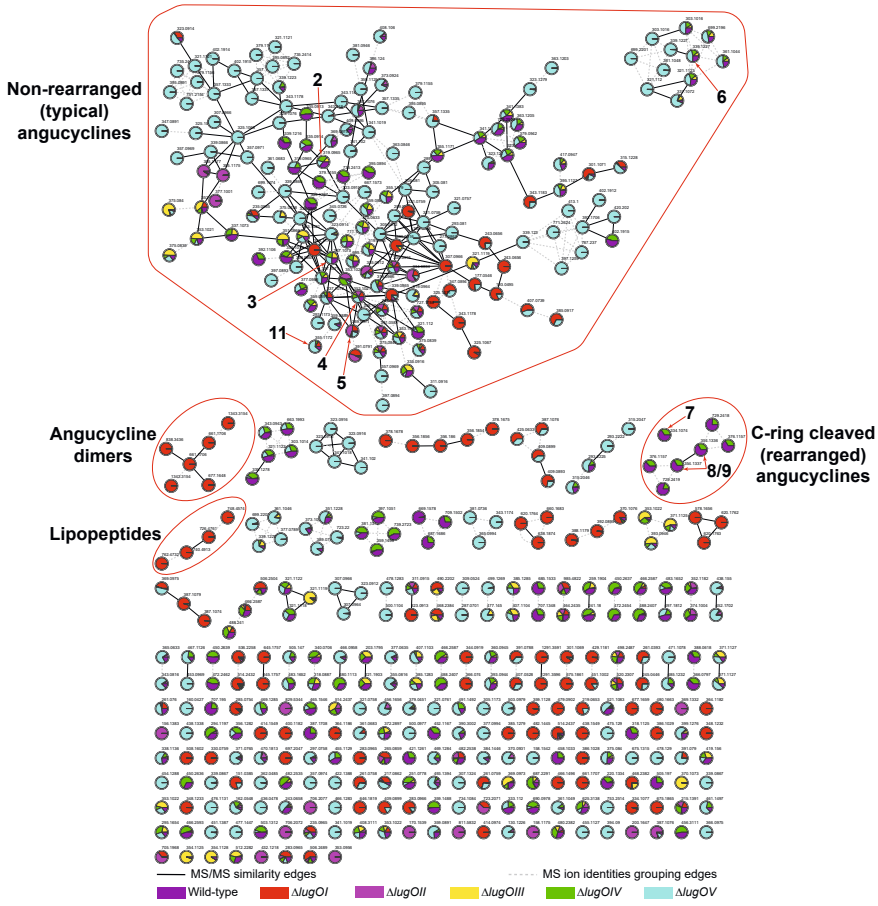


Figure 6 Molecular network of the ions detected in the extracts of *Streptomyces*. sp. QL37 and its *lugOI–lugOV* mutants grown on R5 agar.
The nodes are labelled by the precursor mass of their ions and pie charts are mapped to the nodes to indicate the relative intensities of the ions in the different 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.

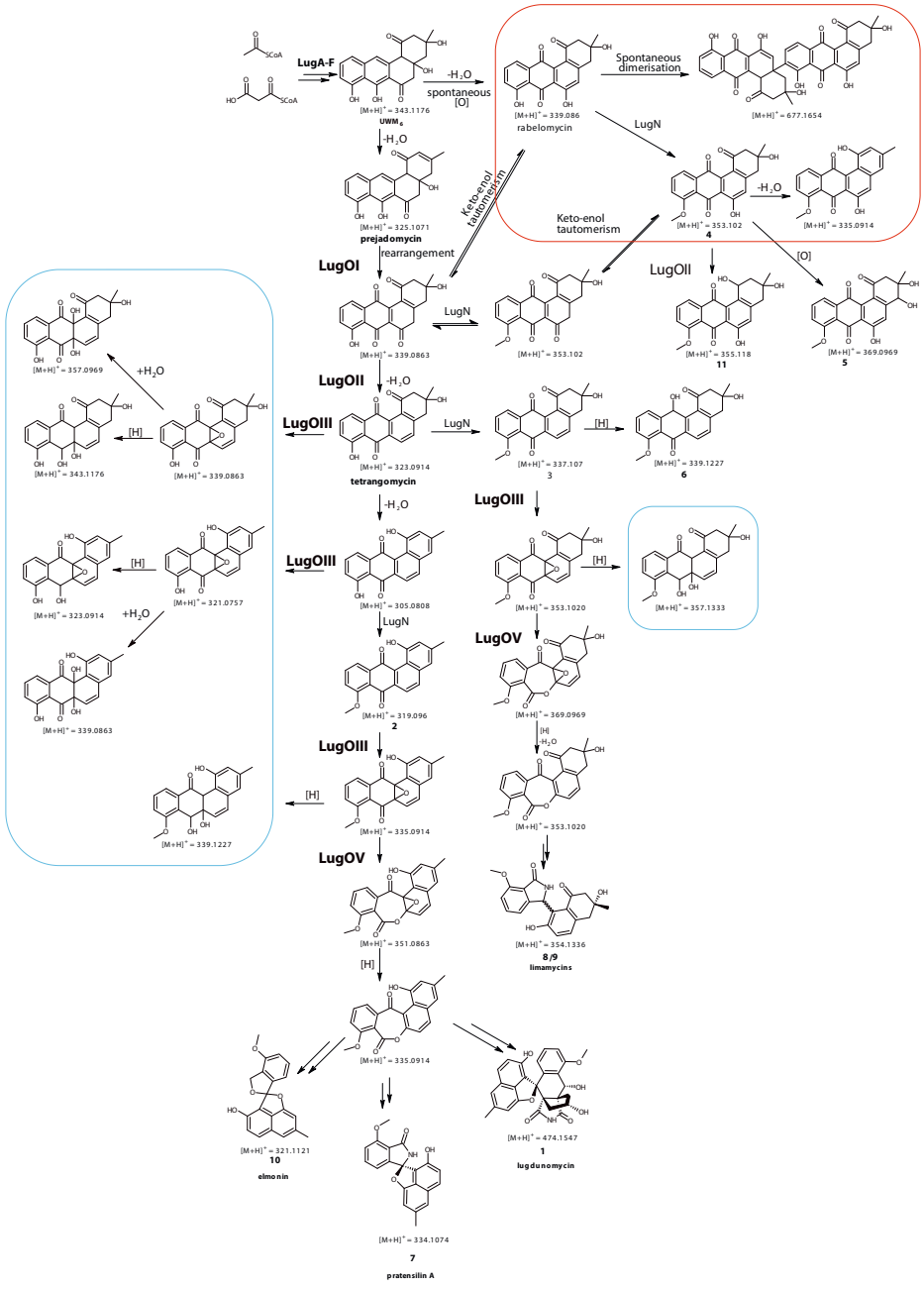


Figure 7 Proposed role of the oxygenases LugOI, LugOII, LugOIII and LugOV in lugdunomycin biosynthesis based on the metabolomics data (this study), and *in vitro* assays.
The calculated masses of the $[M+H]^+$ ions, which correspond to the precursor masses observed in the molecular networks, are given 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

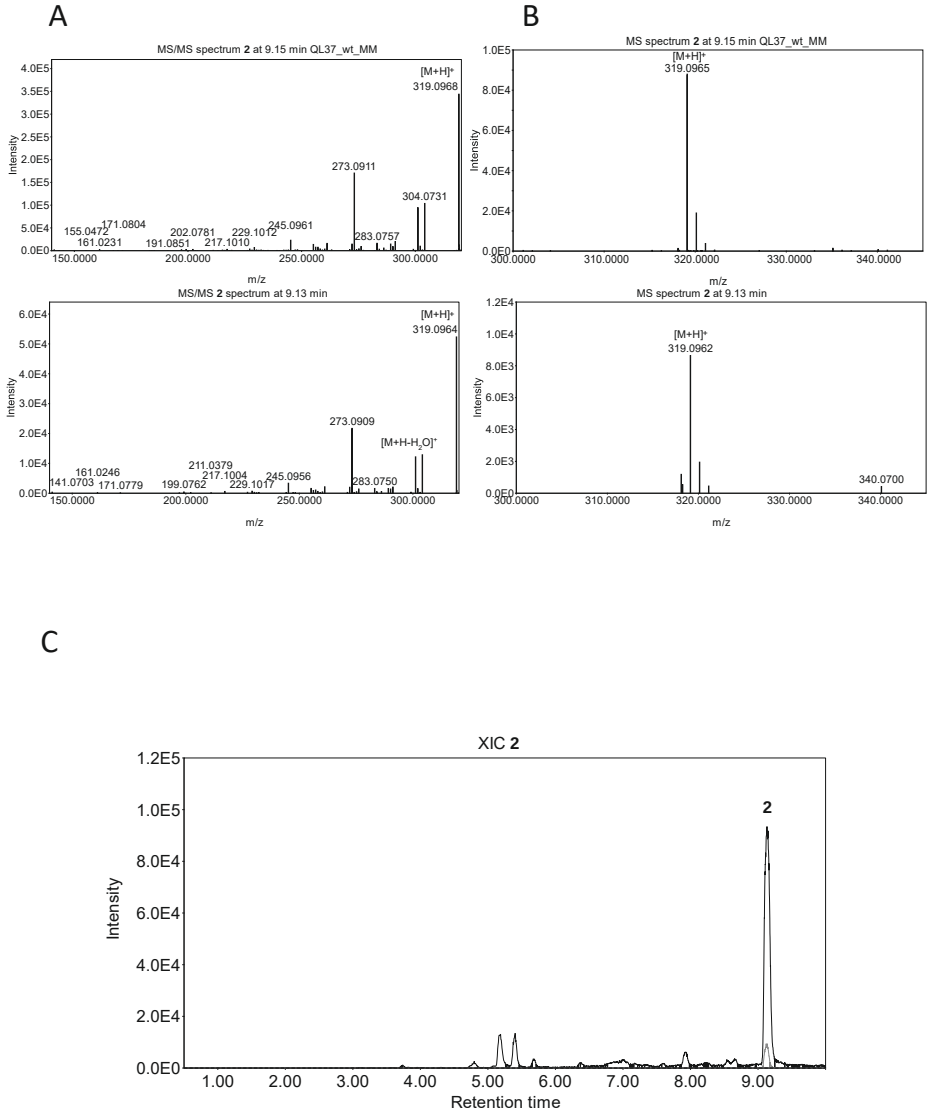


Figure S1 Identification of compound 2 in the extract of *Streptomyces* sp. QL37
Comparison of the MS/MS (A) and MS spectra (B) of the semi-pure compound **2** and its corresponding peak in the crude extract of wild-type *Streptomyces* sp. QL37. The extracted ion chromatograms of the two peaks are shown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

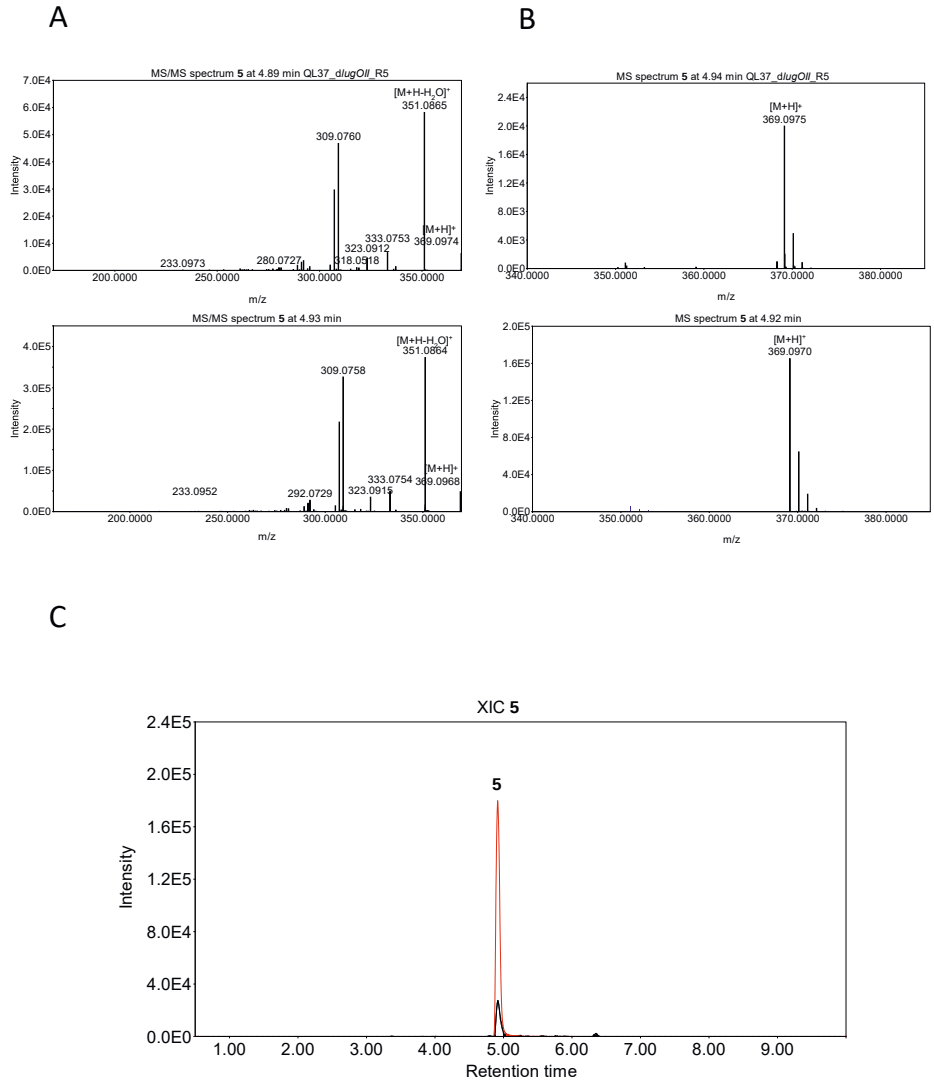


Figure S2 Identification of compound 5 in the extract of *Streptomyces* sp. QL37
Comparison of the MS/MS (A) and MS spectra (B) of the semi-pure compound **5** and its corresponding peak in the crude extract of *lugOII* mutant of *Streptomyces* sp. QL37. The extracted ion chromatograms of the two peaks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

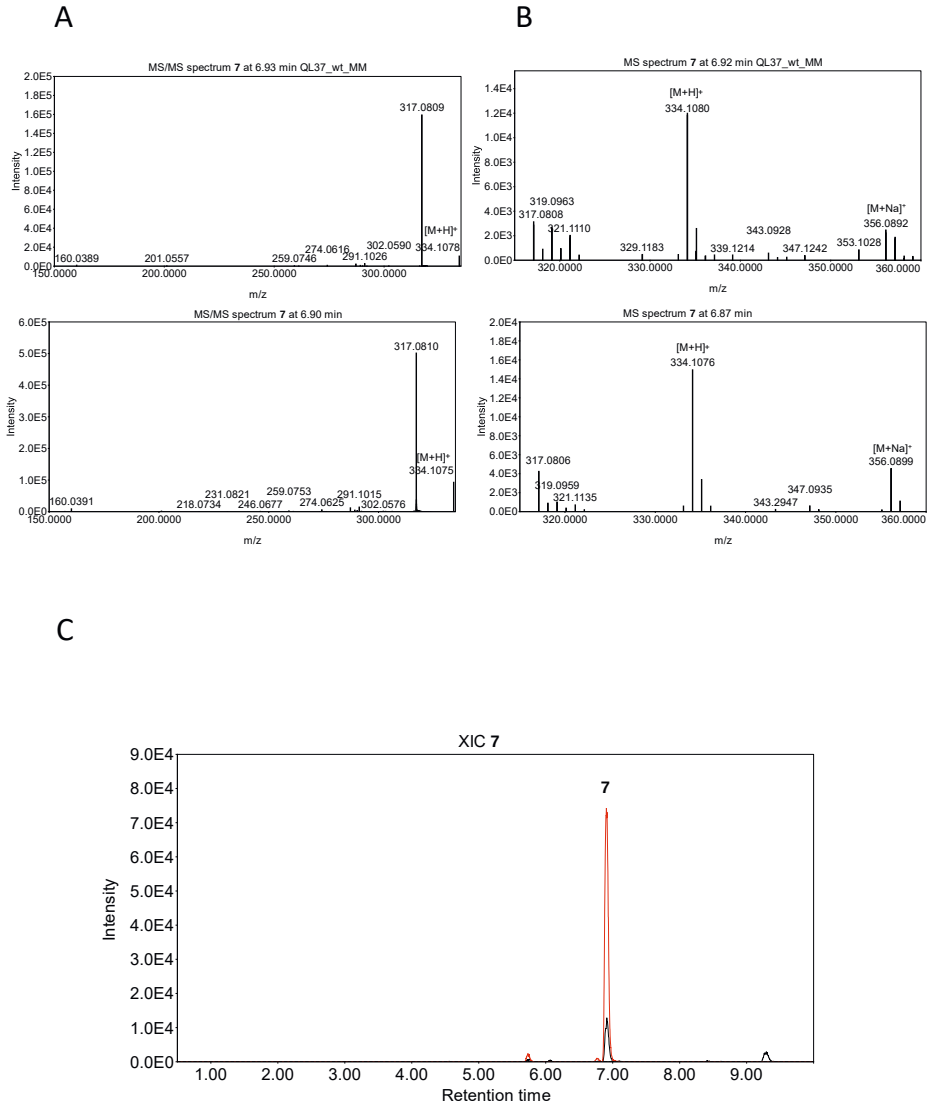


Figure S3 Identification of compound 7 in the extract of *Streptomyces* sp. QL37
Comparison of the MS/MS (A) and MS spectra (B) of the semi-pure compound **7** and its corresponding peak in the crude extract of *Streptomyces* sp. QL37. The extracted ion chromatograms of the two peaks are shown in (C)

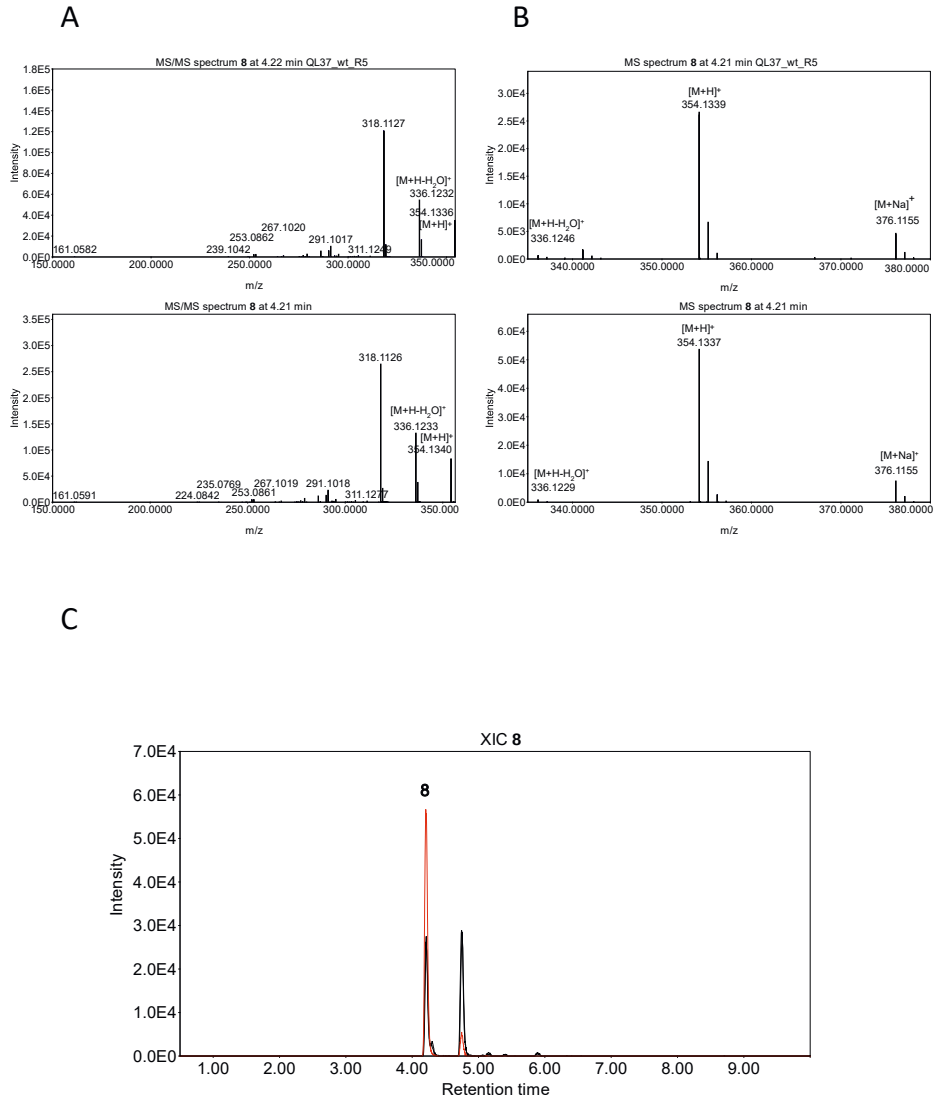


Figure S4 Identification of compound 8 in the extract of *Streptomyces* sp. QL37
Comparison of the MS/MS (A) and MS spectra (B) of the semi-pure compound **8** and its corresponding peak in the crude extract of *Streptomyces* sp. QL37. The extracted ion chromatograms of the two peaks are shown in (C).

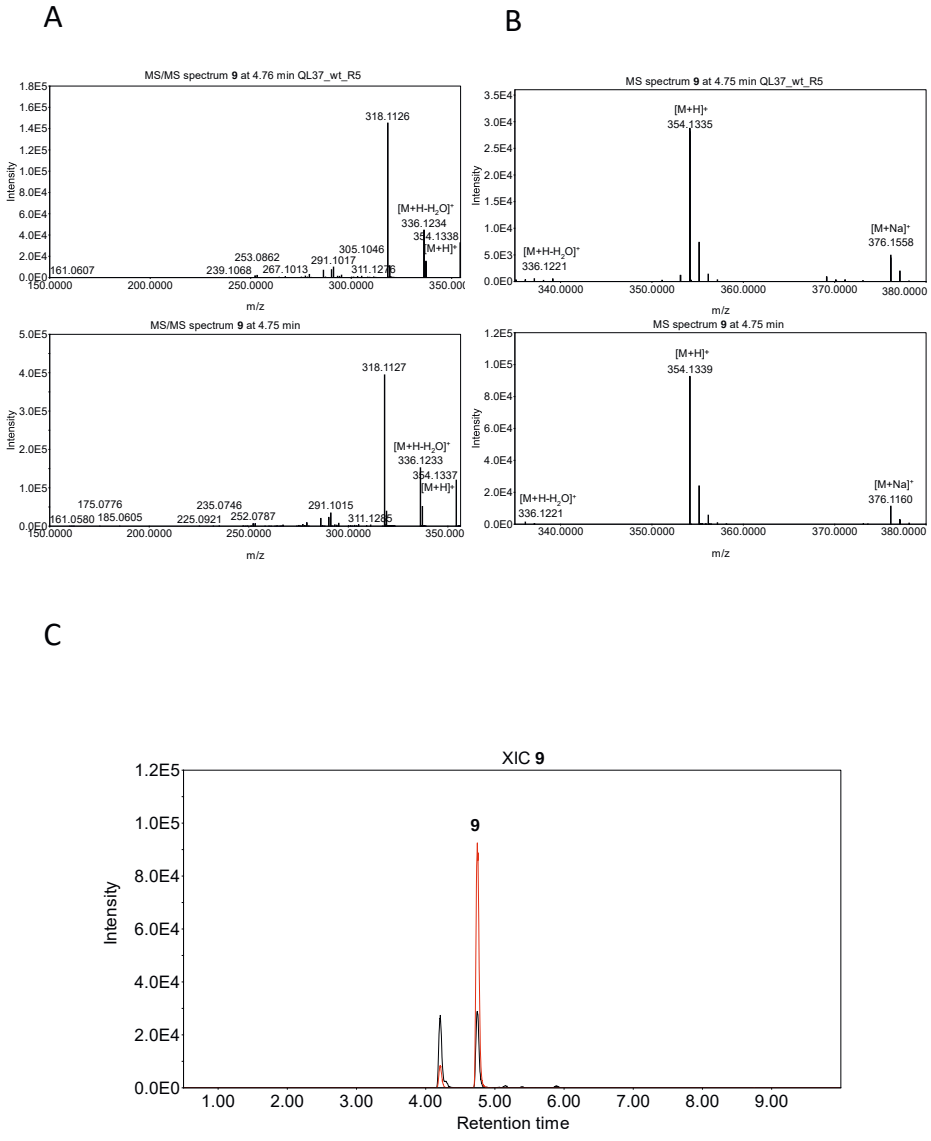


Figure S5 Identification of compound 9 in the extract of *Streptomyces* sp. QL37
Comparison of the MS/MS (A) and MS spectra (B) of the semi-pure compound **9** and its corresponding peak in the crude extract of *Streptomyces* sp. QL37. The extracted ion chromatograms of the two peaks are shown in (C).

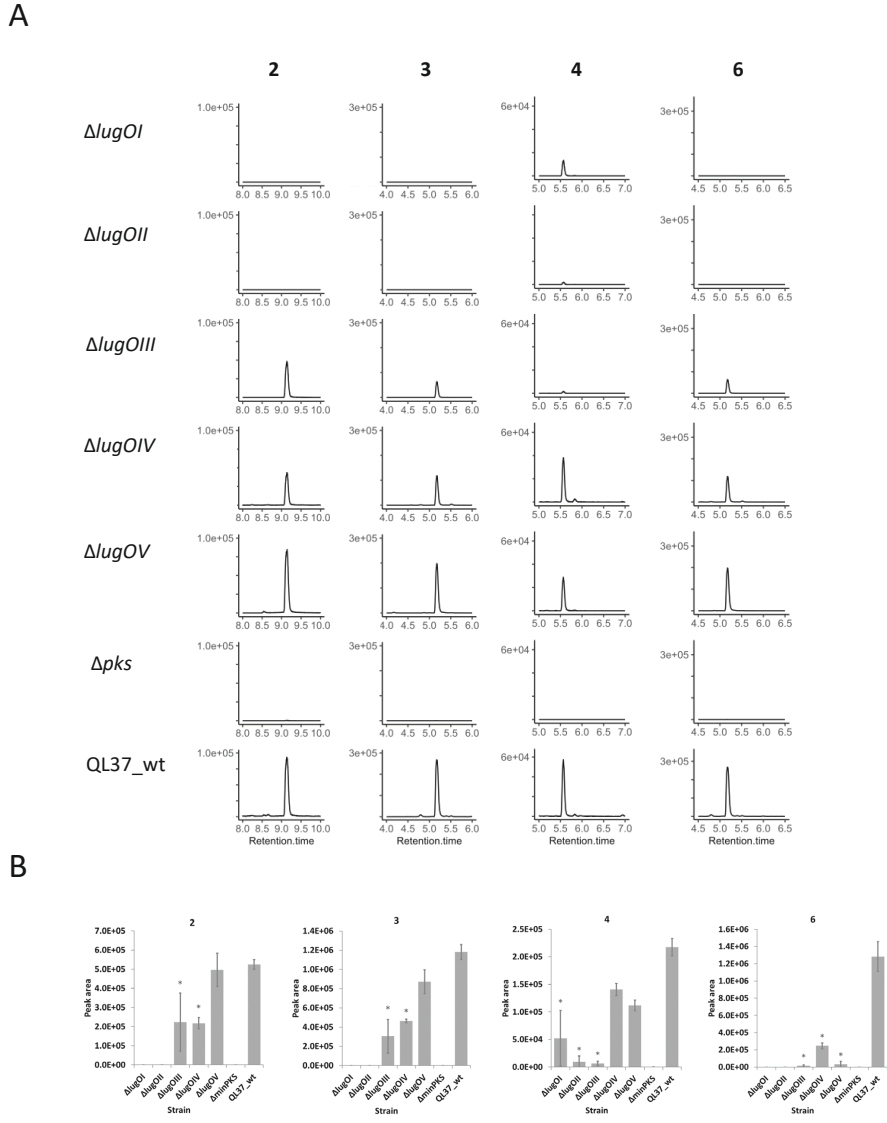


Figure S6 Production of non-rearranged angucyclines by *Streptomyces* sp. QL37, its lug-pks mutant and its *lugOI–lugOV* mutant, when grown on MM.
Extracted ion chromatograms of the non-rearranged angucyclines (**2**, **3, 4**, and **6**) in the extracts of *Streptomyces* sp. QL37 and its mutants of *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.

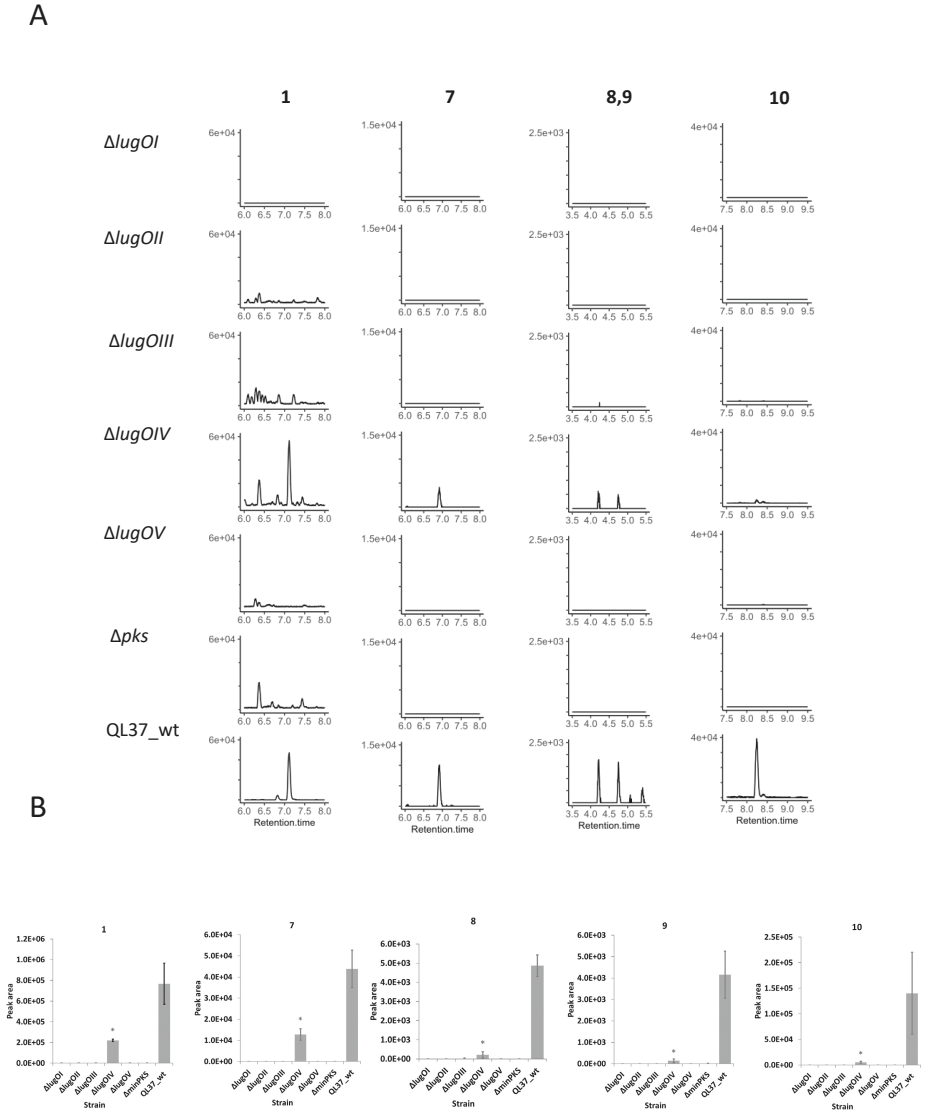


Figure S7 Production of C-ring-rearranged angucyclines by *Streptomyces* sp. QL37, its *lug-pks* mutant and its *lugOI–lugOV* mutant, when grown on MM.
Extracted ion chromatograms of the rearranged angucyclines (**1** $[M+H]^+$, **7** , **8** , **9** and **10**) in the extracts of *Streptomyces* sp. 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.

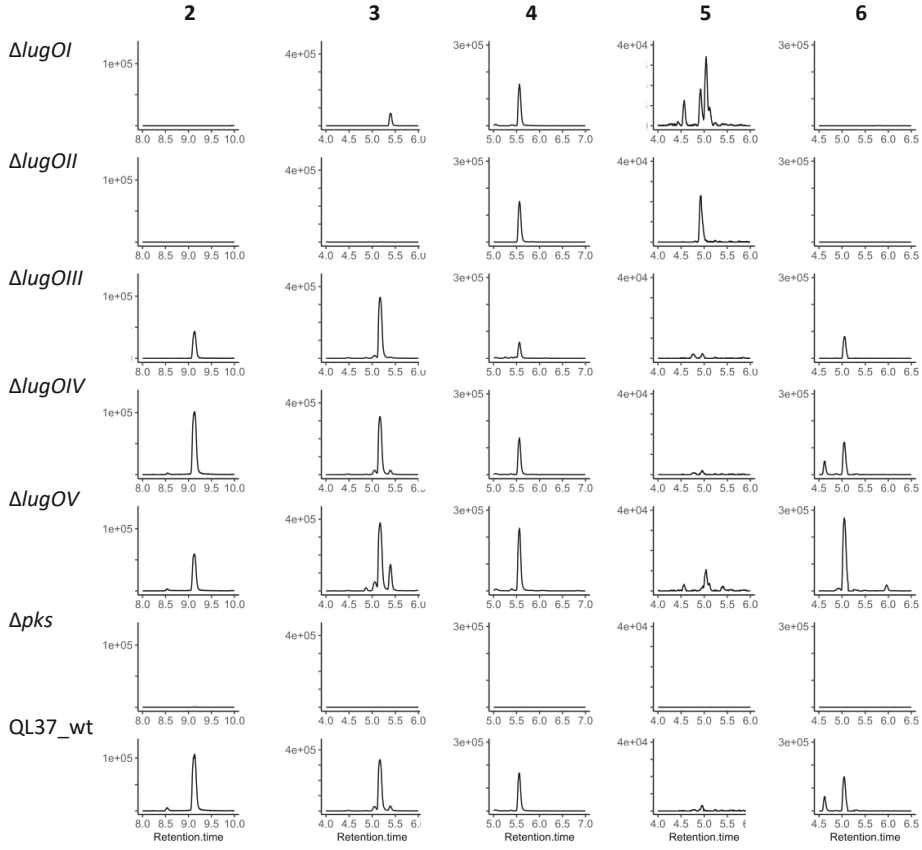


Figure S8A Production of non-rearranged angucyclines by *Streptomyces* sp. QL37, its lug-pks mutant and its *lugOI–lugOV* mutant, when grown on R5.
Extracted ion chromatograms of the typical angucyclines (**2**, **3**, **4**, **5**, and **6**) in the extracts of *Streptomyces* sp. QL37 and its mutants 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.

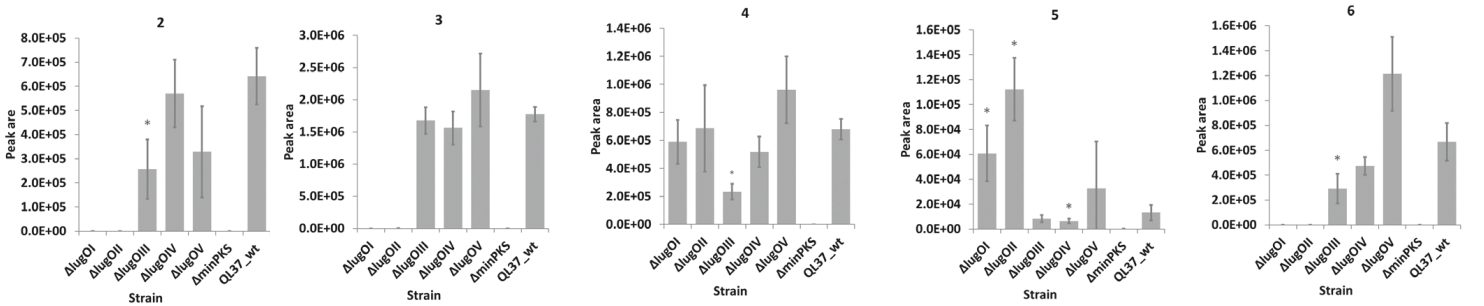


Figure S8B Production of non-rearranged angucyclines by *Streptomyces* sp. QL37, its lug-pks mutant and its *lugOI–lugOV* mutant, when grown on R5.

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.

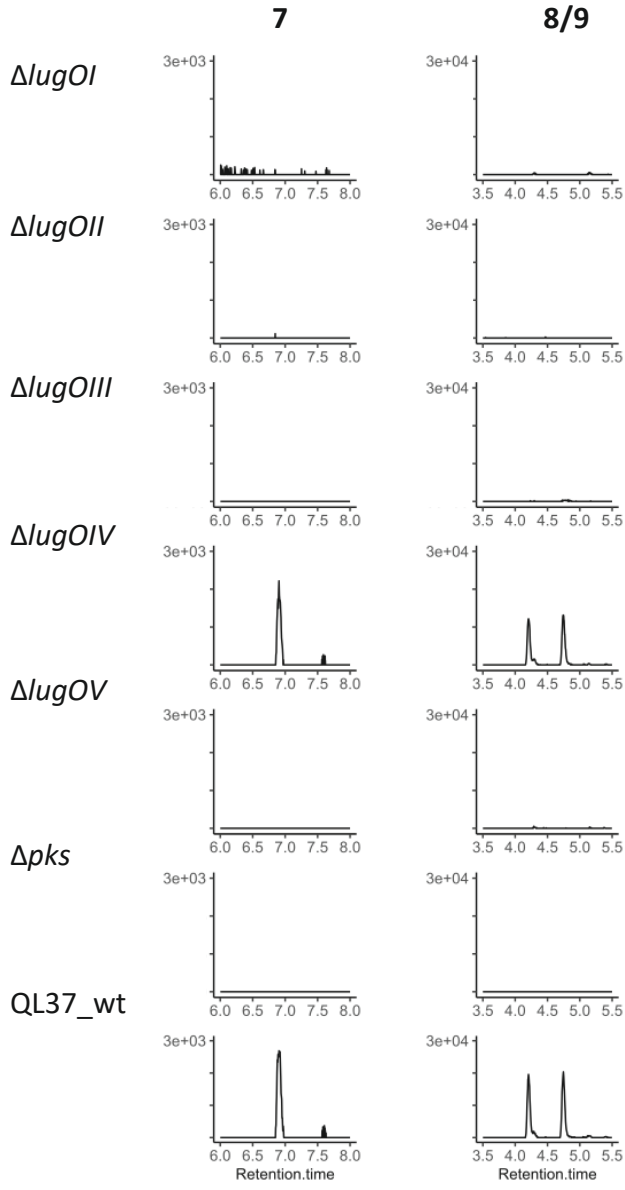


Figure S9A Production of C-ring-rearranged angucyclines by *Streptomyces* sp. QL37, its lug-pks mutant and its *lugOI–lugOV* mutant, when grown on R5.
Extracted ion chromatograms of the typical angucyclines (**7**, **8** and **9**) in the extracts of *Streptomyces* sp. QL37 and its mutants of *lugOI-lugOV*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

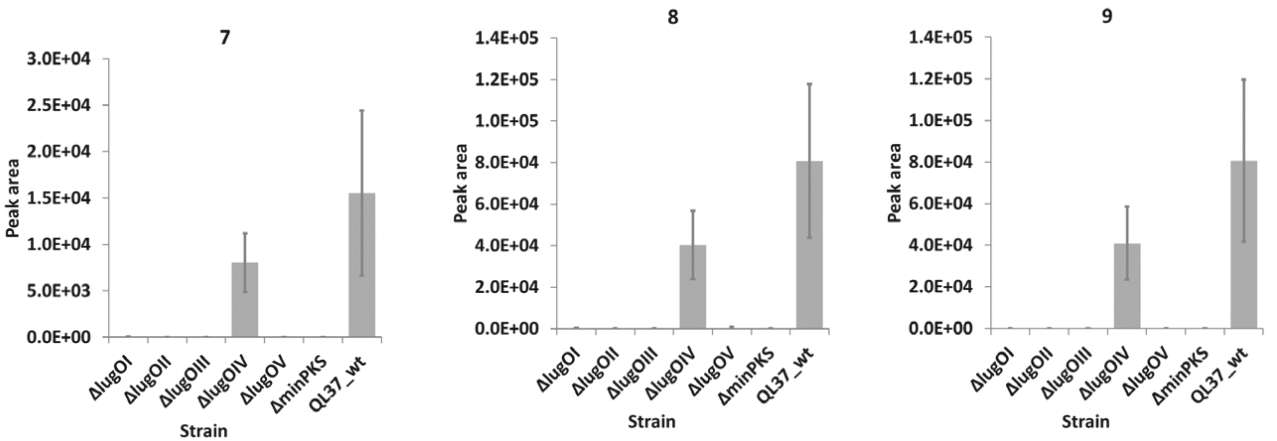


Figure S9B Production of C-ring-rearranged angucyclines by *Streptomyces* sp. QL37, its lug-pks mutant and its *lugOI–lugOV* mutant, when grown on R5.

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.

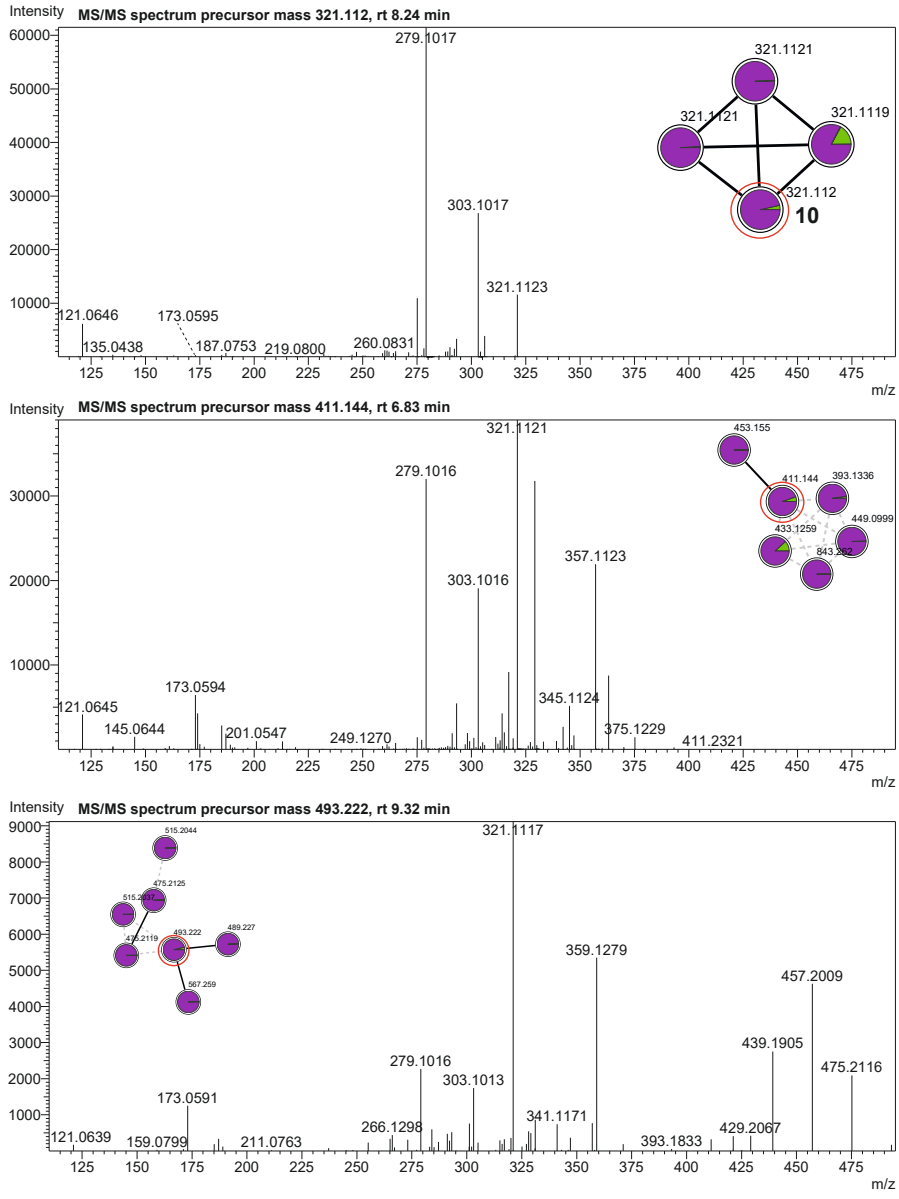


Figure S10 MS/MS spectrum of precursor masses related to elmonin and their molecular families.

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.

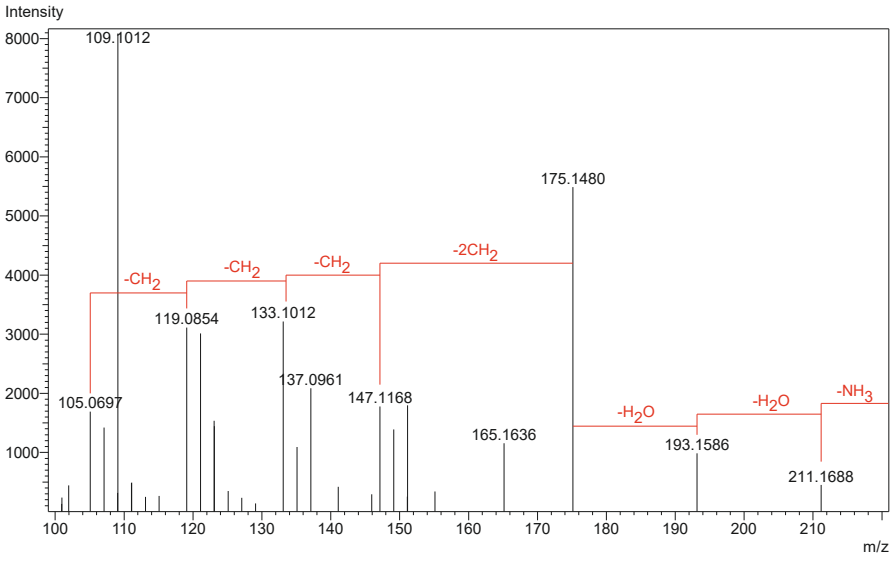


Figure S11 MS/MS spectrum dihydroxy amino alcohol.
MS/MS spectrum of a node in one of the molecular families of the MM extracts, which showed a fragmentation pattern consistent with an unsaturated aliphatic compound containing dihydroxy and a monoamino substitution. This is evident through the observation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

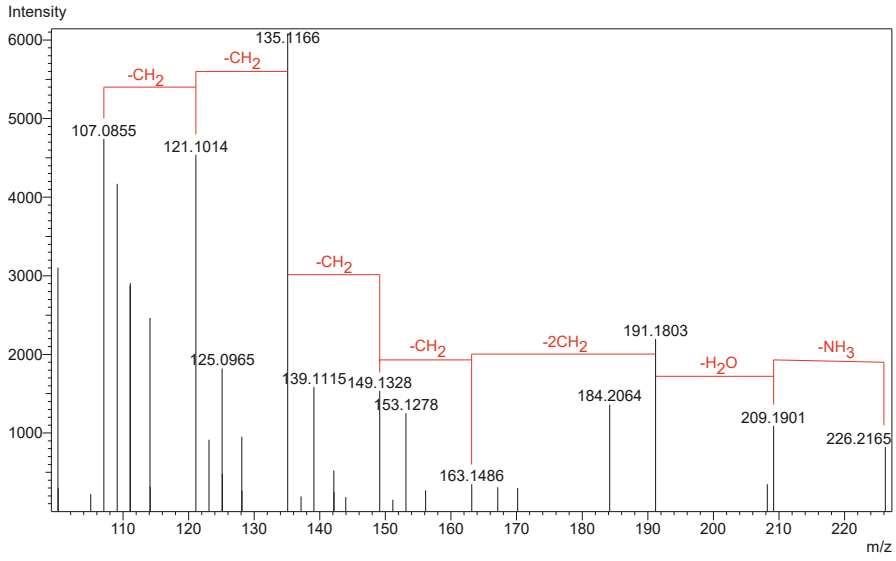


Figure S12 MS/MS spectrum monohydroxy amino alcohol.
MS/MS spectrum of a node in one of the molecular families of the MM extracts, which showed a fragmentation pattern consistent with an unsaturated aliphatic compound containing a monohydroxy and a monoamino substitution. This is evident through the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

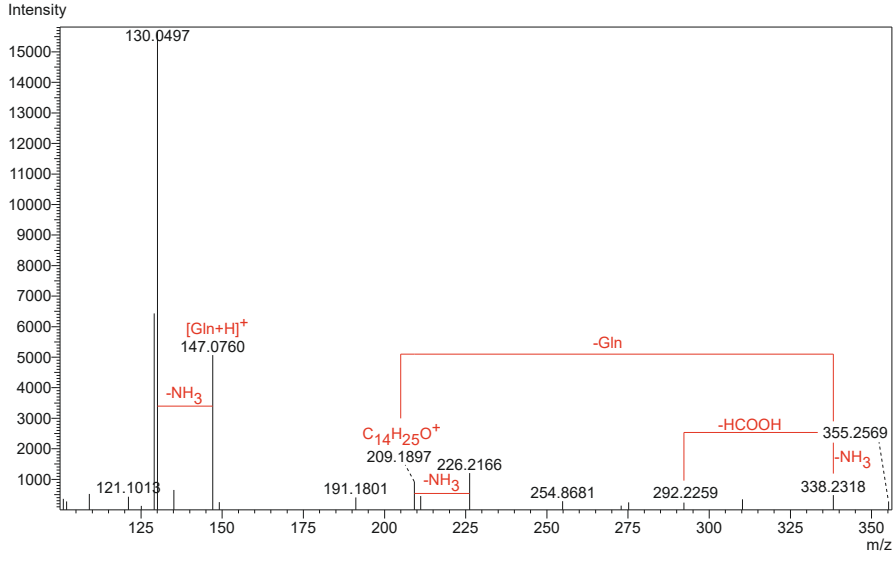


Figure S13 MS/MS spectrum acyl glutamine.
MS/MS spectrum of a node in one of the molecular families of the MM extracts, which showed a fragmentation pattern consistent with an unsaturated fatty acyl chain that is conjugated to glutamine amino acid.

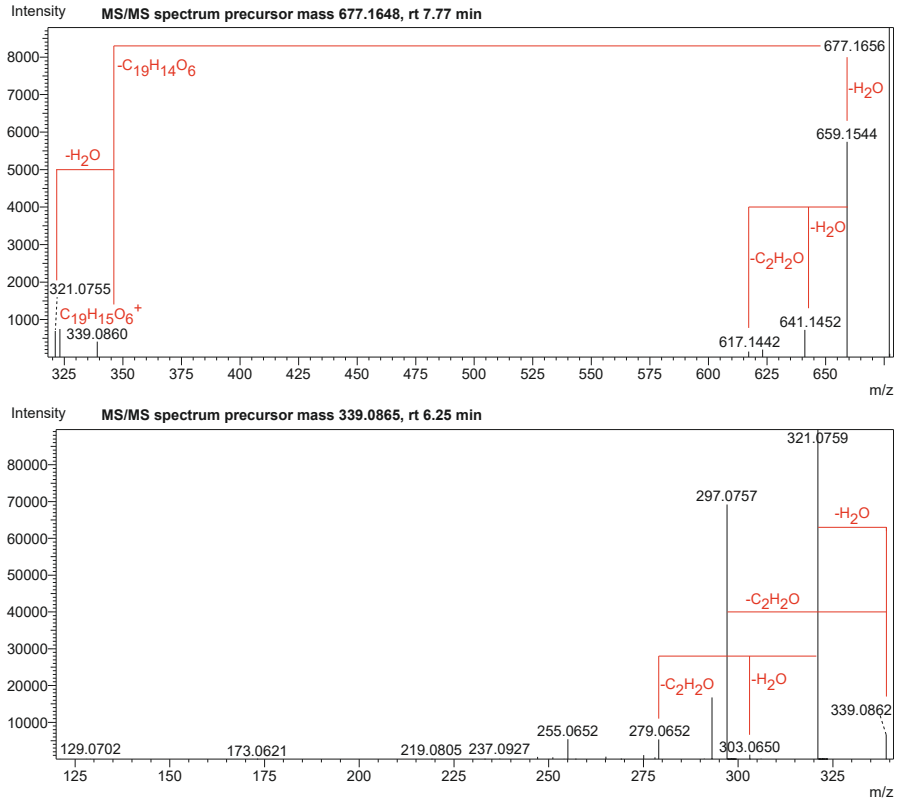
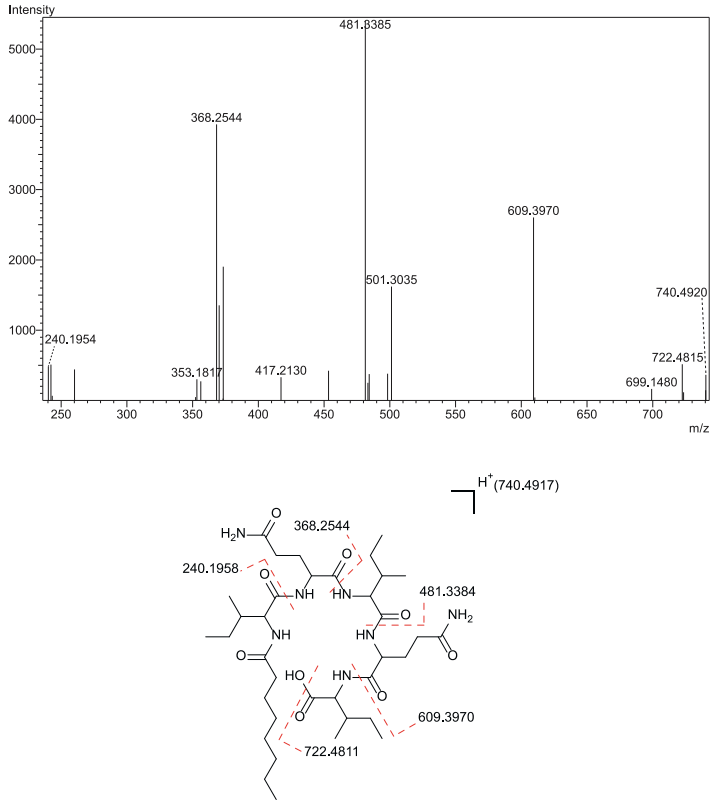


Figure S14 MS/MS spectrum of rabelomycin dimer.
MS/MS spectrum of a node in one of the molecular families of the R5 extracts representing a likely dimer of two rabelomycin molecules (top), as compared to the MS/MS spectrum of rabelomycin (bottom). The dimer shows in the MS/MS spectrum a loss of a rabelomycin 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

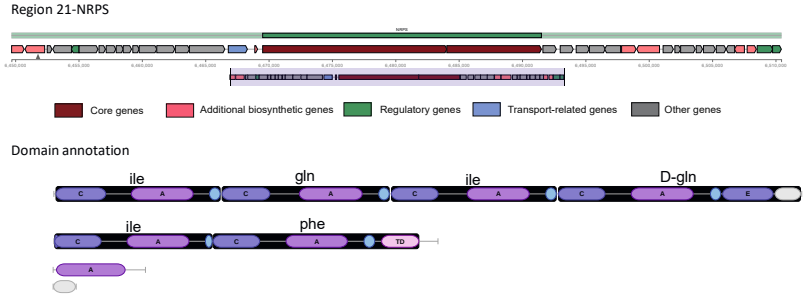


Figure S15 MS/MS spectrum of proposed lipopeptide, its structure and its putative BGC. MS/MS spectrum of a node in one of the molecular families of the R5 extracts, which showed a fragmentation pattern consistent with a lipopeptide metabolite. A proposed structure on which the major fragments are shown 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).

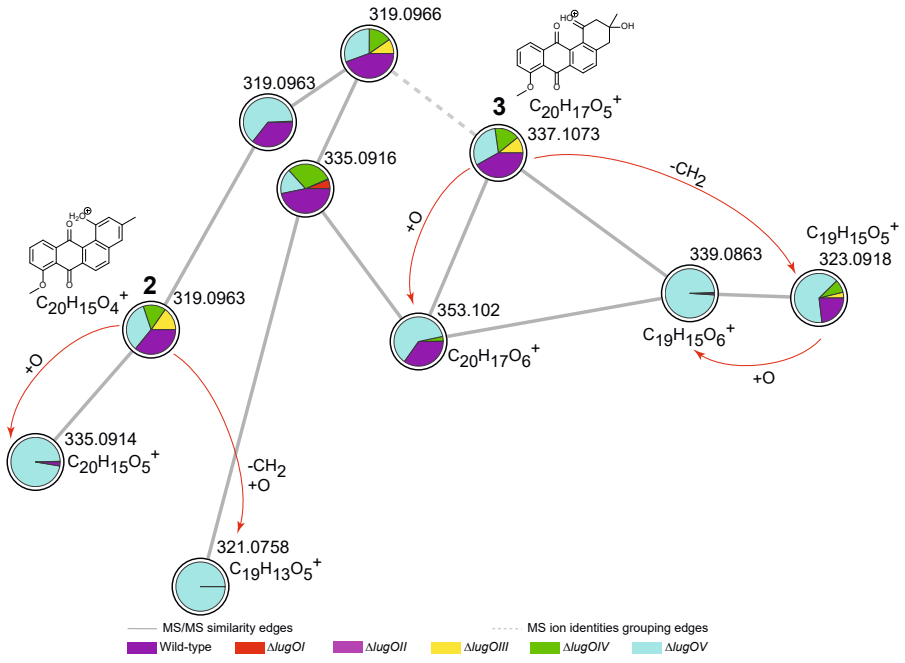


Figure S16 Relationship between nodes upregulated in the *lugOV* mutant grown on MM. Part of the non-rearranged angucycline molecular family in the MM extracts showing the likely structural relationship between the previously identified metabolites **2** and **3**, and the metabolites which were upregulated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

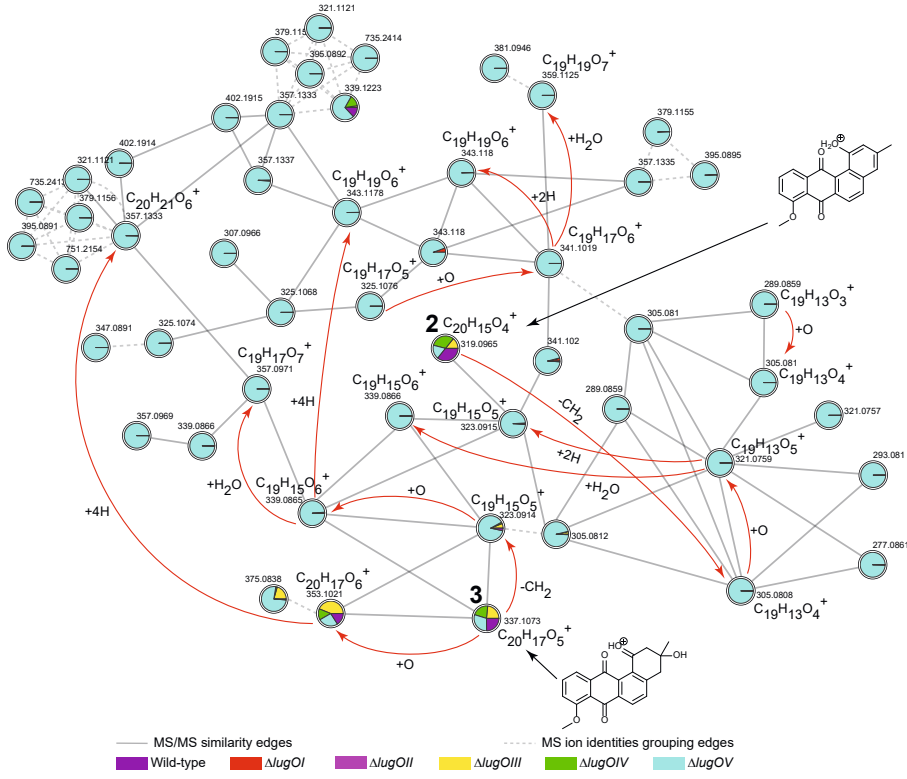


Figure S17 Relationship between nodes upregulated in the *lugOV* mutant grown on R5. Part of the non-rearranged angucycline molecular family in the R5 extracts showing the likely structural relationship between the previously identified metabolites **2** and **3**, and the metabolites which were upregulated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

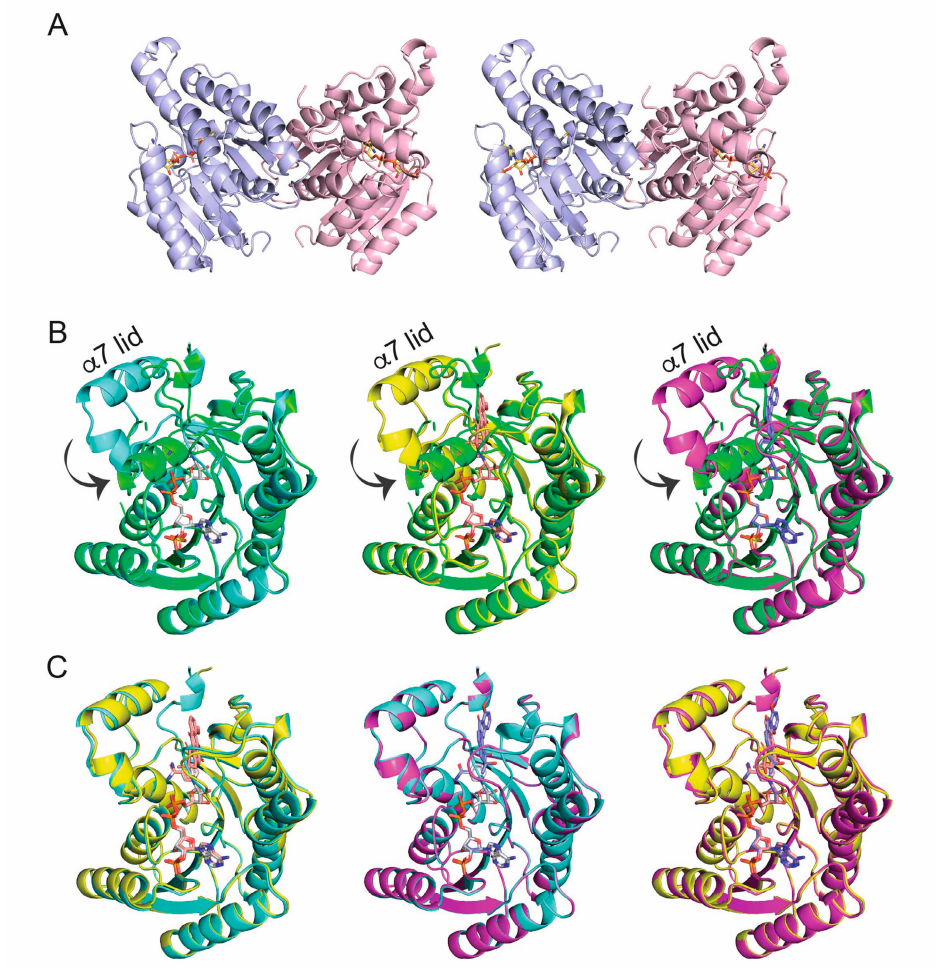


Figure S19 Arrangement of the dimeric structures of LugOIIred and the observed conformational changes between the LugOIIred structures.
A) Stereoscopic view of NADPH bond LugOIIred in dimeric form. The NADPH is displayed as sticks. B) Unliganded LugOIIred (green), LugOIIred/NADPH (cyan), LugOIIred/NADPH/**4** (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)).

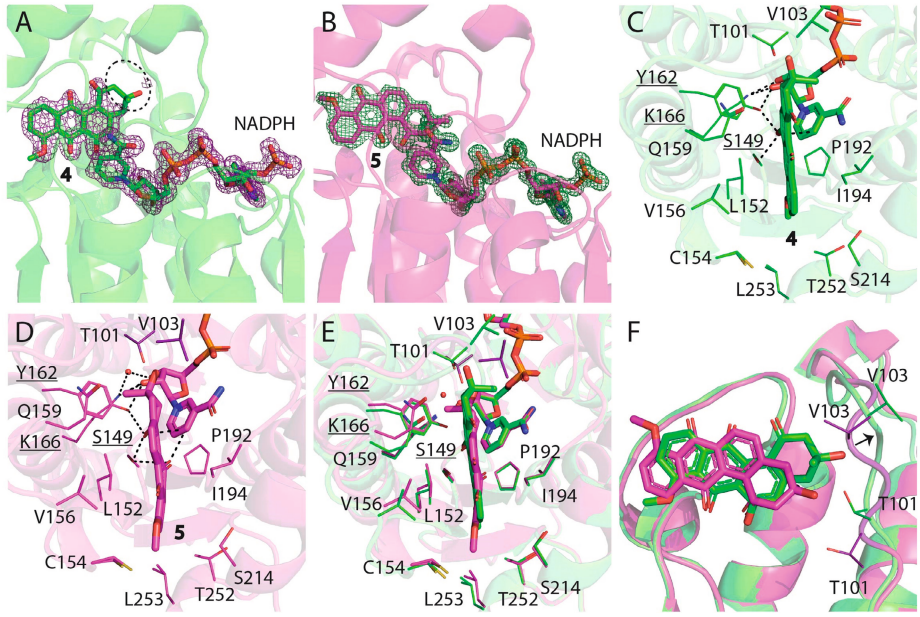


Figure S20 The LugOIIred active site.
The LugOIIred active site. A&B: 2Fo-Fc omit maps contoured at the 1s level corresponding to ligand **3**, **4** and cofactor NADPH. The missing density for **4** is highlighted with black circle. C&D: key residues that surround the binding site of **3** and **4**. 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)).

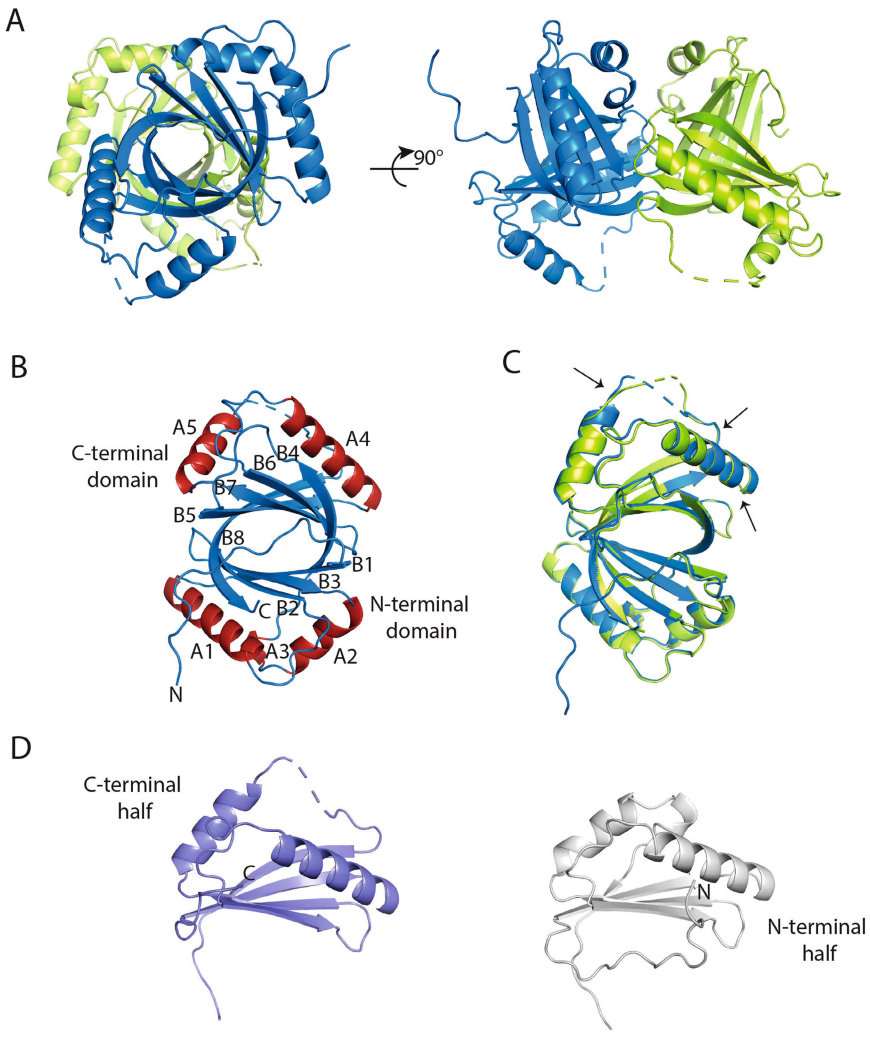


Figure S21 Dimeric arrangement of LugOIII structure.
A) Top and side views of LugOIII in dimeric form. B) LugOIII is consisted by eight b-strands and five a-helices. C) Structural alignment of the two chains of LugOIII structure. Three b-turns are highlighted with arrows. D) The alignment of the N-terminal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