



Universiteit
Leiden
The Netherlands

Discovery of novel antibiotics from actinomycetes by integrated metabolomics & genomics approaches

Wu, Changsheng

Citation

Wu, C. (2016, October 26). *Discovery of novel antibiotics from actinomycetes by integrated metabolomics & genomics approaches*. Retrieved from <https://hdl.handle.net/1887/43768>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/43768>

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

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/43768> holds various files of this Leiden University dissertation

Author: Wu, Shangsheng

Title: Discovery of novel antibiotics from actinomycetes by integrated metabolomics & genomics approaches

Issue Date: 2016-10-26

Discovery of Novel Antibiotics from Actinomycetes by Integrated Metabolomics & Genomics Approaches

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 26 oktober 2016
klokke 16:15 uur

door

Changsheng Wu

吴长生

geboren te Fuzhou, China
13 March 1985

PROMOTIECOMMISSIE

Promotor: Prof. dr. G.P. van Wezel
Co-promotor: Dr. Y.H. Choi
Overige leden: Prof. dr. H.P. Spaink
Prof. dr. P.J.J. Hooykaas
Prof. dr. H.S. Overkleeft
Prof. dr. J. Piel
Prof. dr. O.P. Kuipers



This research was supported by a fellowship from the China Scholarship Council. The cover was designed by Xinyu Ma and Changsheng Wu.

CONTENTS

Chapter 1	General Introduction	1
Chapter 2	Metabolomics in the natural products field – a gateway to novel antibiotics	9
Chapter 3	Metabolic profiling as a tool for prioritizing antimicrobial compounds	21
Chapter 4	Metabolomics-driven discovery of a prenylated isatin antibiotic produced by <i>Streptomyces</i> species MBT28	41
Chapter 5	Expanding the chemical space for natural products by <i>Aspergillus-Streptomyces</i> co-cultivation and biotransformation	57
Chapter 6	Metabolomics-guided analysis of isocoumarin production by <i>Streptomyces</i> species MBT76 and biotransformation of flavonoids and phenylpropanoids	71
Chapter 7	The cryptic <i>qin</i> gene cluster of <i>Streptomyces</i> sp. MBT76 specifies C-glycosylpyranonaphthoquinones with new chemistry	87
Chapter 8	Lugdunomycin, a novel angucycline derivative with unprecedented chemical architecture	103
Chapter 9	Identification of novel endophenaside antibiotics produced by <i>Kitasatospora</i> sp. MBT66	123
Chapter 10	Leucanicidin and endophenasides result from methyl-rhamnosylation by the same tailoring enzymes in <i>Kitasatospora</i> sp. MBT66	137
Chapter 11	Utilization of cellulose as a solid sorbent for antibiotic extraction	159
Chapter 12	Summary and Discussion	169
Samenvatting		181
Appendix	Appendix I: Supplemental data belonging to Chapter 4	193
	Appendix II: Supplemental data belonging to Chapter 5	205
	Appendix III: Supplemental data belonging to Chapter 6	217
	Appendix IV: Supplemental data belonging to Chapter 7	221
	Appendix V: Supplemental data belonging to Chapter 8	229
	Appendix VI: Supplemental data belonging to Chapter 9	239
	Appendix VII: Supplemental data belonging to Chapter 10	269
	Appendix VIII: Supplemental data belonging to Chapter 11	295
	Publication List	297
	Curriculum Vitae	299

Chapter 1

General Introduction

Fleming's fortuitous discovery of penicillin in 1928 opened up a journey to find antibiotics that fight against bacterial infections, which has been regarded as one of the greatest triumphs of modern medicine. However, the overuse of antibacterial drugs over several decades, in combination with extensive misuse in humans and in particular in animals led to the selection and spread of drug-resistant bacteria. Consequently, antibacterial drugs have become less effective or even ineffective, resulting in an accelerating global health security emergency that is rapidly outpacing available treatment options.¹ The explosive increase in infections by pathogens such as multidrug-resistant *Mycobacterium tuberculosis* (MDR-TB), methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant *Enterococcus faecium* (VRE) and fluoroquinolone-resistant *Pseudomonas aeruginosa* is a major problem in the clinic today, and the most recent occurrence of pan-antibiotic-resistant infections pose the grave threat of completely untreatable infections.² This thesis was written within this context, with the concept of "finding novel antibiotics from actinomycetes" in mind.

A rich source of natural products, actinomycetes are high-G+C, Gram-positive, free-living, filamentous bacteria that are ubiquitous in nature, and produce some 70% of all known antibiotics.³ Of the actinomycetes, the members of the genus *Streptomyces* are particularly prolific antibiotic producers and have a complex life cycle. The life cycle starts with the germination of a spore under the appropriate conditions, to produce a vegetative mycelium (also known as substrate mycelium) consisting of an intricate network of branching hyphae. In response to nutrient exhaustion or other stressed signals, an aerial mycelium is formed, whereby the vegetative mycelium is autolytically degraded to provide the necessary nutrients.⁴ Eventually, the aerial hyphae differentiate into long chains of pre-spore compartments by a complex cell division event, followed by separation of the mature spores which spread and then enter a new life cycle.⁵ The transition from vegetative to aerial growth roughly coincides with, or slightly precedes, that of chemical differentiation, *i.e.* the production of secondary metabolites (Figure 1).^{6,7}

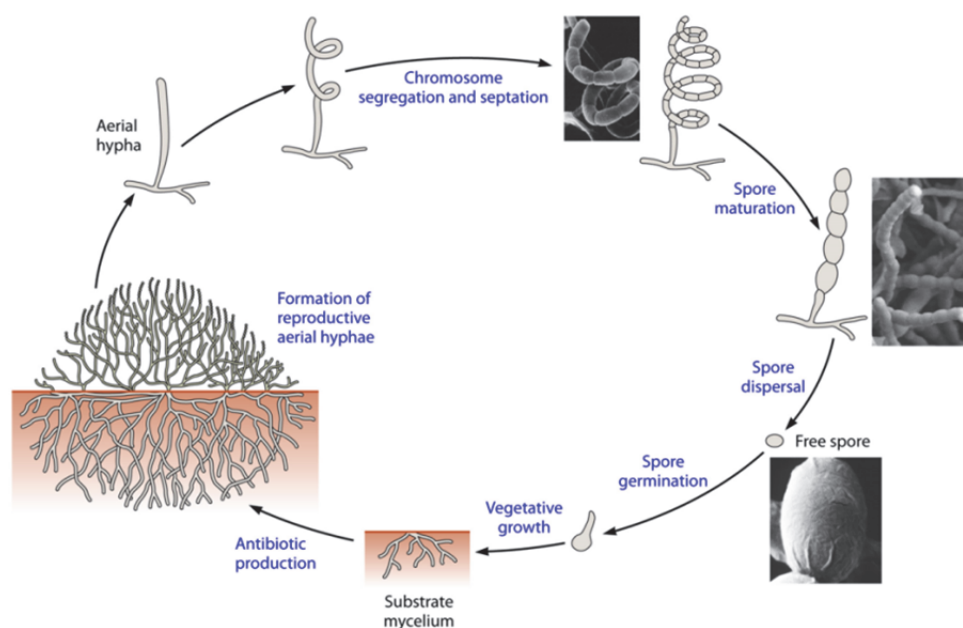


Figure 1. Life cycle of *Streptomyces*. *Streptomyces* has a very complex life cycle. During the initiation of aerial growth, it starts to produce a plethora of antibiotics. The picture is reproduced with permission from Barka et al.⁸

At the basis of this thesis lies a unique collection of antibiotic-producing actinomycetes present at Leiden University in van Wezel lab,⁹ derived from unusual and underexplored environments, such as unspoilt areas in the Qinling and Himalaya mountains. The aim of this thesis was to look for compounds with totally new scaffolds and therefore hopefully also a new mode of action. The research involves both chemistry and biology, *e.g.* 1) activating silent gene clusters for antibiotics in actinomycetes; 2) isolation of compounds from a complex mixture, in particular application of NMR-based metabolomics; 3) identifying the structures of compounds by using spectroscopic techniques like nuclear magnetic resonance (NMR) or mass spectrometry (MS) and where needed complemented by X-ray crystallography; 4) testing antimicrobial activity of isolated compounds; and 5) elucidation of the biosynthetic pathway for the new compounds.

Genome-mining approach as a major trend for future antibiotics discovery

In the past, antibiotics were typically discovered through high-throughput antibacterial screening programs, usually by cultivating bacteria from soil samples and isolating their secondary metabolites. However, in more recent years this traditional approach has yielded disappointing returns owing to the high frequency of rediscovery of known compounds.¹⁰ Because of this high reoccurrence, it was thought that the biosynthetic potential of the traditional producers had been fully explored, so that pharmaceutical companies started turning away from natural products for drug discovery programs in the 1990s.¹¹ The massive acceleration in the development of sequencing technologies over the past decade has resulted in easy and affordable access to bacterial genomes, and this has brought these “over-explored” secondary metabolite producers back into the spotlight because genetic data have shown that the metabolic capabilities of these natural resources are severely underestimated.^{12–14} Additionally, through genomics, it was shown that underexplored or neglected organisms are capable producers of secondary metabolites. New technologies based on genome sequence data, like metagenomics, allow scientists access to the ‘dark matter’, namely the vast biosynthetic potential of microbes harbored in environmental niches, especially those produced by laboratory-uncultivable microbes.¹⁵ One can easily envision that new natural products possessing unprecedented structural features as well as novel modes of action can be isolated from such underexplored organisms.¹⁶

The approach of genomics-driven discovery of natural products (Figure 2) is rendered feasible by the fact that all of the genes encoding the large number of enzymes required for the synthesis of a typical secondary metabolite are clustered in a tight locus that is termed as biosynthetic gene cluster (BGC). A variety of computational tools has been developed for their automatic genomic identification, such as the “Antibiotics and Secondary Metabolite Analysis SHell” (antiSMASH) which has served as the most comprehensive resource for identifying and analyzing novel secondary metabolite biosynthetic pathways in microorganisms.^{17,18} Gene clusters such as those for polyketide synthase (PKS) or nonribosomal peptide synthase (NRPS) that follow a linear bioassembly line are readily identified using bioinformatics, and to some extent the domain structures of the biosynthetic proteins allow prediction of the molecule that they specify.¹⁶ In this sense, the utility of genomics is complementary to spectroscopy, to enable rapid determination of complex structures.^{19,20} Bioinformatics analyses of biosynthetic gene cluster can also predict important

structural information, e.g. physiochemical properties like chromophore, which can guide researchers in the identification and isolation of compounds specified by particular gene clusters.¹⁹

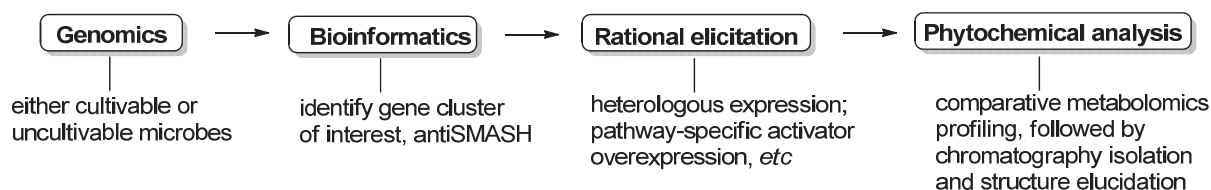


Figure 2. The genome mining approach for natural product discovery in microorganisms. Following the identification and growth of a microorganism of interest, the genome sequence is obtained. Bioinformatics tools are then used to analyze the genome and identify cryptic biosynthetic gene clusters (BGCs). Silent BGCs are activated using one or a combination of pleiotropic or pathway-specific approaches. Novel metabolites are identified in culture extracts or supernatants, followed by their purification and characterization, typically using a combination of high-resolution mass spectrometry, and 1D and 2D NMR spectroscopy.¹⁷

Tapping untapped microbial resources

Well-known bacterial sources of antibiotics include Actinomycetes, predominantly from the genus *Streptomyces*,²¹ as well as myxobacteria,²² cyanobacteria,²³ bacilli²⁴ and pseudomonads²⁵. Previously unexplored (or under investigated) bacterial sources are being tapped for their potential to produce novel compounds with new activities. These “neglected producers” include bacteria from unusual habitats and distant branches of the eubacterial phylogenetic tree, such as genera *Lysobacter*, plant- and insect-associated bacteria and anaerobic bacteria.²⁶ Many “neglected bacteria” have been shown not only able to generate secondary metabolites, but that many of these compounds have novel chemical structures and, in some cases, unique modes of action and activities. From the anaerobic bacterium *Clostridium*, Hertweck and co-workers recently discovered clostrubins A and B with unusual pentacyclic polyphenolic topology, which show pronounced activity against various pathogenic bacteria.^{27,28} In addition, the vast majority of soil microorganisms cannot be cultivated under laboratory conditions.^{29,30} To circumvent this problem, methods have been developed for the growth of such microorganisms in conditions that simulate their natural environment³⁰ and for growing such ‘uncultivated microorganisms’ in their natural environment.²⁹ The iChip (isolation chip) technique allows microbial cultivation in field by using a perforated plate in which each well captures a single microbial cell,^{29,31,32} resulting in the discovery of lassomycin³² and teixobactin.³¹

Objectives and outline of this thesis

The aim of this study was to discover novel secondary metabolites with potential antimicrobial properties from actinomycetes, and the thesis is generally classified into two sections according to the methodology used in finding new antibiotics. In the first section including Chapters 2—7, different strategies were utilized to enforce fluctuations in the production of bioactive compounds, and NMR-based metabolic profiling was used to facilitate uncovering those elicited molecules, which are less commonly applied in microbial drug discovery; The second section consisting of Chapters 8—11 is more focused on systematic chemical extraction and isolation approach that is classic in the field of natural products chemistry. The brief introduction for each chapter is displayed as following.

Chapters 2 and 3 are reviews of metabolomics in natural products field, with emphasis on the analysis of NMR-based metabolomics. Metabolomics is a useful tool to dereplicate the known compounds and thus enable efficient discovery of novel molecules.

Chapters 4—7 offer proofs of the concept that NMR-based metabolic profiling tool facilitates new microbial NPs discovery. We used strategies of inducing streptomycin resistance (Chapter 4), microbial co-cultivation between *Streptomyces* and *Aspergillus* (Chapter 5), different harvesting times (Chapter 6), and overexpression of a pathway-specific regulator (Chapter 7), to induce and diversify the expression of natural products in actinomycetes. ^1H NMR was subsequently used to profile the metabolomes elicited through these ecological, chemical or genetic perturbations, in comparison with the respective control experiments. NMR spectral data were furthermore interrogated statistically using multivariate data analyses to differentiate chemical profiles and correlate specific NMR signals to bioactivity. The discriminating signals for elicited molecules were used as probe(s) for the NMR-guided isolation of bioactive compounds, which allowed efficient identification of new molecules and avoid chemical redundancy. Besides, as demonstrated in Chapters 5 and 6, NMR-based metabolic profiling streamlines microbial biotransformation that is a very important tool to add new chemical diversity to known natural products.

In Chapters 8 and 9, different culture media were first tested to increase secondary metabolites production and antimicrobial activity in actinomycetes, including *Kitasatospora* sp. MBT66 and *Streptomyces* sp. QL37. The up-scale fermentation in the resulting optimized growth medium was conducted to accumulate enough material. After extraction with ethyl acetate, efforts were made to chromatographically isolate as many compounds as possible from the referred microbial strains, *i.e.* a systematical isolation approach. All the purified compounds were subjected to spectroscopy analysis for structure elucidation, and the identified new compounds were tested for antimicrobial activities. Moreover, the biosynthetic pathway of new compounds was proposed and their corresponding biosynthetic gene clusters were identified through bioinformatics analysis and genetic manipulation. Especially, in *Streptomyces* sp. QL37 (Chapter 8), I characterized a novel antibiotic, called lugdunomycin as well as a number of new angucycline-type antibiotics. Lugdunomycin is particularly exciting as it has a completely novel skeleton. The polyketide lugdunomycin possesses a benzaza[4,3,3]propellane skeleton adorned with a spirocyclic 2*H*-naphtho[1,8-*bc*]furan moiety and two all-carbon quaternary centers embedded within five contiguous stereogenic carbons. The striking backbone of benzaza[4,3,3]propellane-6-spiro-2'-2*H*-naphtho[1,8-*bc*]furan has a so far unprecedented chemistry. The biosynthesis of lugdunomycin involves an unusual carbon-carbon cleavage in the quinone ring of angucycline, which could be explored to modify a wide range of polyketide-based antibiotics, such as tetracyclines.

The work in Chapter 10 answers the question that remained unresolved in Chapter 9, namely which genes are responsible for the methyl-rhamnosylation of the new endophenazines that had been discovered in *Kitasatospora* sp. MBT66. The work revealed biosynthetic “crosstalk” among chromosomally and phylogenetically distant gene clusters enabled *de novo* synthesis of different classes of rhamnosylated natural products, and confirmed the promiscuity of a glycosyltransferase by bioinformatics analysis and gene knockout. In Chapter 11, a natural polymer, cellulose was investigated for its potential application in the field of selective purification and isolation of antibiotics. Cellulose was

found to selectively bind water-soluble antibiotics like aminoglycosides and glycosylated peptides.

Finally, a general discussion of the thesis is presented in Chapter 12, which also includes a summary of the most important data and observations, and provides perspectives for future antibiotic discovery from microbial sources.

REFERENCES

- (1) WHO Geneva, Switzerland, 2014.
- (2) Arias, C. A.; Murray, B. E. *N. Engl. J. Med.* **2009**, *360*, 439–443.
- (3) Genilloud, O.; González, I.; Salazar, O.; Martín, J.; Tormo, J. R.; Vicente, F. *J. Ind. Microbiol. Biotechnol.* **2011**, *38*, 375–389.
- (4) Rigali, S.; Titgemeyer, F.; Barends, S.; Mulder, S.; Thomae, A. W.; Hopwood, D. A.; van Wezel, G. P. *EMBO Rep.* **2008**, *9*, 670–675.
- (5) Flärdh, K.; Buttner, M. J. *Nat. Rev. Microbiol.* **2009**, *7*, 36–49.
- (6) Bibb, M. J. *Curr. Opin. Microbiol.* **2005**, *8*, 208–215.
- (7) van Wezel, G. P.; McDowall, K. J. *Nat. Prod. Rep.* **2011**, *28*, 1311–1333.
- (8) Barka, E. A.; Vatsa, P.; Sanchez, L.; Gaveau-vaillant, N.; Jacquard, C.; Klenk, H.; Clément, C.; Ouhdouch, Y.; van Wezel, Gilles, P. *Microbiol. Mol. Biol. Rev.* **2016**, *80*, 1–43.
- (9) Zhu, H.; Swierstra, J.; Wu, C.; Girard, G.; Choi, Y. H.; Van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. *Microbiol. (United Kingdom)* **2014**, *160*, 1714–1726.
- (10) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat. Rev. Drug Discov.* **2007**, *6*, 29–40.
- (11) Koehn, F. E.; Carter, G. T. *Nat. Rev. Drug Discov.* **2005**, *4*, 206–220.
- (12) Corre, C.; Challis, G. L. *Nat. Prod. Rep.* **2009**, *26*, 977–986.
- (13) Bentley, S. D.; Cerdeño-Tárraga, K. F. C. A.-M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Bentley, S. D.; Harper, D.; Bateman, A.; Brown, S.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Larke, L.; Murphy, L.; Oliver, K.; Rabinowitsch, E.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Woodward, J.; Barrell, B. G.; Parkhill, J. *Nature* **2002**, *3*, 141–147.
- (14) Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S. *Nat. Biotechnol.* **2003**, *21*, 526–531.
- (15) Brady, S. F.; Simmons, L.; Kim, J. H.; Schmidt, E. W. *Nat. Prod. Rep.* **2009**, *26*, 1488–503.
- (16) Winter, J. M.; Behnken, S.; Hertweck, C. *Curr. Opin. Chem. Biol.* **2011**, *15*, 22–31.
- (17) Rutledge, P. J.; Challis, G. L. *Nat. Rev. Microbiol.* **2015**, *13*, 509–523.
- (18) Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Brucoleri, R.; Lee, S. Y.; Fischbach, M. A.; Müller, R.; Wohlleben, W.; Breitling, R.; Takano, E.; Medema, M. H. *Nucleic Acids Res.* **2015**, *43*, W237–W243.

-
- (19) McAlpine, J. B.; Bachmann, B. O.; Pirae, M.; Tremblay, S.; Alarco, A. M.; Zazopoulos, E.; Farnet, C. M. *J. Nat. Prod.* **2005**, *68*, 493–496.
- (20) Tietz, J. I.; Mitchell, D. *Curr. Top. Med. Chem.* **2015**, *16*, 1–1.
- (21) Watve, M. G.; Tickoo, R.; Jog, M. M.; Bhole, B. D. *Arch. Microbiol.* **2001**, *176*, 386–390.
- (22) Wenzel, S. C.; Müller, R. *Mol. Biosyst.* **2009**, *5*, 567–574.
- (23) Welker, M.; Dittmann, E.; von Döhren, H. *Methods Enzymol.* **2012**, *517*, 23–46.
- (24) Fickers, P. *Am. J. Biochem. Biotechnol.* **2012**, *8*, 38–43.
- (25) Gross, H.; Loper, J. E. *Nat. Prod. Rep.* **2009**, *26*, 1408–1446.
- (26) Pidot, S. J.; Coyne, S.; Kloss, F.; Hertweck, C. *Int. J. Med. Microbiol.* **2014**, *304*, 14–22.
- (27) Pidot, S.; Ishida, K.; Cyrulies, M.; Hertweck, C. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 7856–7859.
- (28) Shabuer, G.; Ishida, K.; Pidot, S. J.; Roth, M.; Dahse, H.-M.; Hertweck, C. *Science (80-.)*. **2015**, *350*, 670–675.
- (29) Nichols, D.; Cahoon, N.; Trakhtenberg, E. M.; Pham, L.; Mehta, A.; Belanger, A.; Kanigan, T.; Lewis, K.; Epstein, S. S. *Appl. Environ. Microbiol.* **2010**, *76*, 2445–2450.
- (30) Kaeberlein, T.; Lewis, K.; Epstein, S. S. *Science* **2002**, *296*, 1127–1129.
- (31) Ling, L. L.; Schneider, T.; Peoples, A. J.; Spoering, A. L.; Engels, I.; Conlon, B. P.; Mueller, A.; Hughes, D. E.; Epstein, S.; Jones, M.; Lazarides, L.; Steadman, V. a; Cohen, D. R.; Felix, C. R.; Fetterman, K. A.; Millett, W. P.; Nitti, A. G.; Zullo, A. M.; Chen, C.; Lewis, K. *Nature* **2015**, *517*, 455–459.
- (32) Gavrish, E.; Sit, C. S.; Cao, S.; Kandrör, O.; Spoering, A.; Peoples, A.; Ling, L.; Fetterman, A.; Hughes, D.; Bissell, A.; Torrey, H.; Akopian, T.; Mueller, A.; Epstein, S.; Goldberg, A.; Clardy, J.; Lewis, K. *Chem. Biol.* **2014**, *21*, 509–518.

Chapter 2

Metabolomics in the natural products field – a gateway to novel antibiotics

Changsheng Wu^{1,2}, Hye Kyong Kim¹, Gilles P. van Wezel², and Young Hae Choi¹

¹*Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE Leiden, The Netherlands*

²*Molecular Biotechnology, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE Leiden, The Netherlands*

Drug Discov. Today Technol. **2015**, 13, 11–17.

ABSTRACT

Metabolomics is a high throughput analytical technique used to globally measure low molecular weight metabolites, allowing simultaneous metabolic comparison of different biological samples and thus highlighting differential compounds as potential biomarkers. Though microbes are renowned as prolific sources of antibiotics, the traditional approach for new anti-infectives discovery is time-consuming and labor-intensive. In this review, the use of NMR- or MS-based metabolomics is proposed as an efficient approach to find antimicrobials in microbial single- or co-cultures.

1. INTRODUCTION

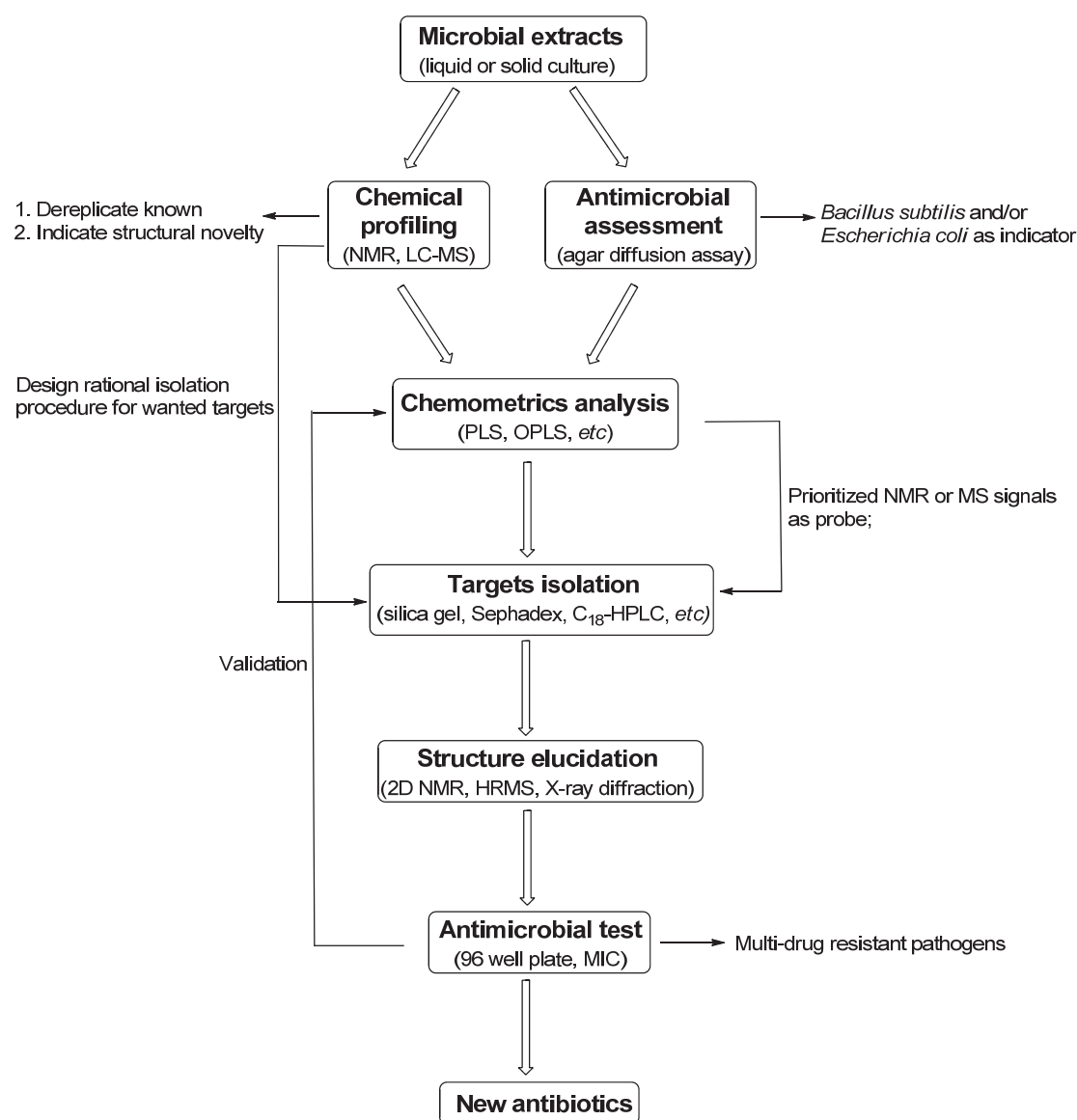
The number of multi-drug resistant bacteria is rising alarmingly and the treatment of infections caused by these microbes is extremely challenging.¹ Microorganisms are a renowned source of effective antibiotics. The traditional design of methods used for this research is considerably time and labor-intensive, and even worse is high rediscovery rate as much as 99.9%.¹ Thus, in view of the urgent need of new antibiotics, this search could be revitalized with a more efficient approach directed at speeding up the discovery process.

Metabolomics is a high throughput analytical technique used to globally measure low molecular weight metabolites in biological samples. Using a combination of chemical profiles built by LC-HRMS or NMR with multivariate data analysis (MDA), it is possible to compare and detect differential metabolites in active/inactive biological samples, narrowing the search to potential biomarkers and avoiding chemical redundancy in an early stage. In this paper, we have reviewed examples of the use of metabolomics as an effective tool to increase the efficiency of antibiotic discovery from microbial sources.

2. Metabolomics-guided discovery of new antibiotics

Metabolomics methodology, either NMR- or MS- based, is an ideal tool for the chemical screening and subsequent detailed comparison of the secondary metabolomes of a set of bacterial fermentations, rapidly revealing the differences/biomarkers among experimental groups. The next question is how to handle the compounds responsible for the separation in the score plot of supervised and/or unsupervised MDA. Since NMR and MS techniques can give structural information of compounds in an early stage, there are three possibilities for the identification of valuable biomarkers. Firstly, it is possible to compare the structural information provided by NMR (¹H NMR chemical shifts) and/or high-resolution mass spectrometry (molecular formula and isotopic pattern) with literature and/or databases. If known structures, discriminators can be identified without the need for time-consuming isolation process and thus metabolomics serves as an approach for dereplication; Another possible situation is the new compounds that possess known structural core but unknown functional groups—"known unknowns". In this case, their structural characterization can be done within the context of the crude biological matrix with the aid of various 2D NMR techniques as reviewed.^{2,3} Besides this, the fragmentation patterns observed in the MS/MS spectra in conjunction with molecular networking evaluation can enable the identification of previously unknown close relatives within the same chemical families through spectral correlations;^{4,5} Last but not least, if the compounds have unprecedented nuclei, chromatographic isolation is inevitable for full *de novo* NMR structural characterization, but at the same time are the most promising in terms of truly new antimicrobials discovery. For this step, the discriminating signals generated in the comparative analysis (MDA) could be used as probes to track the target of interest ("star antibiotic") throughout the chromatographic isolation, which is referred to as NMR-^{6,7} and/or MS-guided^{8,9} separation method. This can prevent losing track of the target that is unfortunately quite common in conventional TLC-guided isolation. Furthermore, some degree of structure identification, albeit incomplete, can be achieved beforehand with the 2D NMR measurements since they allow a partial prediction of the novel compounds. This can also provide information on chemical characteristics and help optimize isolation procedure, which makes senses for minimizing the

loss of compounds (e.g. ruling out the use of bare silica gel for polyphenolic polyketides due to their practically irreversible adsorption as in the case of clostrubin¹⁰). Taken together, the strategy of metabolomics-guided discovery of new antibiotics is illustrated in Scheme 1.



Scheme 1. Metabolomics-guided discovery of new antibiotics.

The use of metabolomics as a methodology for de-replication of known microbial secondary metabolites and prioritization of promising candidates is exemplified in the following cases. The LC-MS metabolome profiling of nine *Myxococcus* species allowed the rapid identification of known myxobacterial metabolites and the subsequent compound-based principal component analysis (PCA) was successfully highlighted the putative novel compound determined by its exact mass and isotopic pattern.¹¹ Tim S. Bugni¹² and his colleagues used UHPLC-HRMS-based metabolomics to dereplicate known structures and predict novel compounds in microbial crude extract. They then applied MS-guided separation approach to target the molecules highlighted by metabolomics analysis, and successfully discovered a series of antibiotics with novel skeletons, such as microtermolides,¹³

bottromycin D,¹⁴ and more recently forazoline A.¹⁵ Forazoline A (Figure 1) is a potent polyketide with a highly unusual skeleton, exhibiting *in vivo* antifungal efficacy against *Candida albicans* (MIC 16 µg/mL)—comparable to that of amphotericin B in a mouse model of *C. albicans* and no toxicity.¹⁵ Furthermore, they utilized a LC-MS-SPE-NMR hyphenation system that integrates MS-guided separation and *de novo* NMR structure elucidation¹⁶ to facilitate rapid structure elucidation of rare phenyl-acetyl-desferrioxamines indicated by metabolomics analysis.¹² In case of lowly abundant compounds, targeted LC-MS microisolation strategy reported by Jean-Luc Wolfender *et. al*^{17,18} for the purification of novel metabolites is necessary for full structure determination by means of microflow NMR analysis that just requires micrograms of compounds.^{19,20}

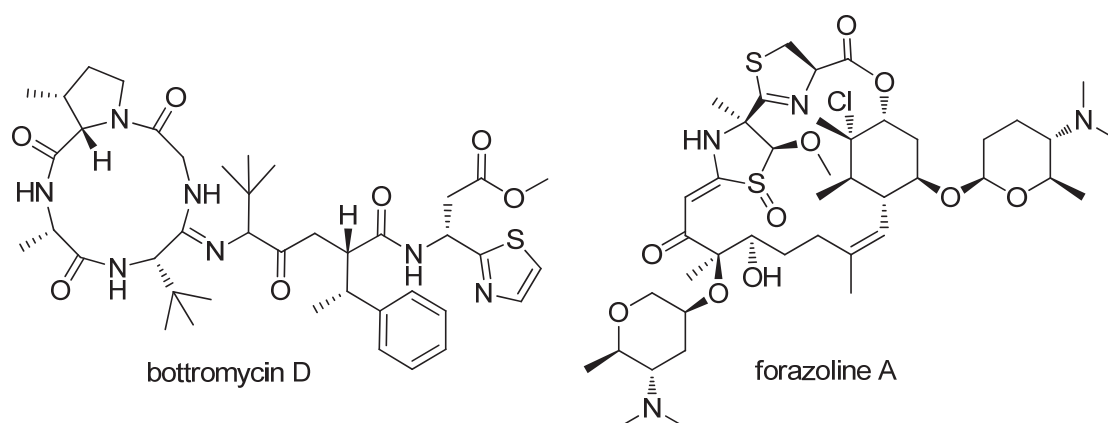


Figure 1. Structurally novel antibiotics discovered through the framework of metabolomics methodology.

3. Metabolomics applied to microbial cocultivation

Conventionally, scientists isolate pure strains of microbes collected from various biological samples, such as soil, marine sediments, symbionts of plants or sponges, *etc.* They then scale-up the single-strain fermentation under specific laboratory conditions to accumulate sufficient material for the next-stage systematical or bioassay-guided isolation. However, the isolation of novel antibiotics from microbes by means of this well-established routine is difficult. One of the reasons is many putative gene clusters remain silent under standard laboratory growth conditions when microbes are cultivated singly. Thus, it might be fruitful to consider not only single microorganism but more complex biological systems consisting of different organisms. This novel approach in which microbes are grown together (co-culture or confrontation experiments) has received increasing interest for its potential to produce new leads, and to decipher specific biosynthetic pathways mainly related to defense.²¹ The focus should be set on antagonistic interactions among the different microbes (fungi and/or bacteria) coexisting in the same defined environment, because they use chemical warfare consisting generally of small organic compounds for defense. In some cases, clear long-distance growth inhibition can be observed during pairwise co-culture on solid media. This phenomena is also expected to be linked to the production of antimicrobial compounds as a defense mechanism by one of the two co-inhabiting microorganisms.^{18,22} Coculturing the fungal endophyte *Fusarium tricinctum* with the bacterium *Bacillus subtilis* 168 trpC2 on solid rice medium resulted in up to 78 times the production of the fungal antibiotics enniatins A1 and B1.²³ Microbial cocultivation is also effective for the induction of new molecules with antimicrobial effect, as verified by the discovery of the new antibiotics pestalone²⁴, emericellamides A and

B,²⁵ *O*-methymllelin.²⁶ It is quite likely that one microbe produce molecules that could be precursors for biotransformation into other final products by its partner, which can increase the chemodiversity of induced metabolites and thus the chance of finding novel antibiotic.²⁷

Considering the great potential of cocultivation as a fertile source for new antibiotics, it is necessary to develop effective methods for their detection and identification of induced metabolites. Metabolomics is particularly appropriate for this purpose as indicated in ²². Untargeted generic fingerprinting using ¹H-NMR, but more by hyphenated MS methods (considered by some to be more powerful because of their sensitivity) is aimed to capture metabolites comprehensively. The comparison of the metabolome of the coculture with that of the individual microorganisms can highlight the induced metabolites present in coculture but absent in pure strains, since these *de novo*-produced molecules have high chances of being new compounds (even new skeletons) with antimicrobial activity. However, it is important to bear in mind that the conventional chemometric methods, such as PCA and (O)PLS-DA, are not absolutely ideal for coculture metabolomics studies, because they might not effectively filter the up-regulated metabolites in monocultures and thus not always highlight the coculture-induced metabolites specifically. A novel data mining approach referred to as projected orthogonalized chemical encounter monitoring (POChEMon) was recently developed and has proved to be more appropriate for this purpose.²¹

The detection of antibiotics using the microbial cocultivation strategy may be further revolutionized by the application of advanced analytical technologies introduced into metabolic profiling, one of which is imaging mass spectrometry (IMS).^{28–30} This technique can provide distinct chemical/structural data on the spatial distributions of metabolites directly on the agar Petri dishes avoiding complex sample pre-treatment.^{31–34} For this, the antagonistic interaction between tested microorganisms is favored. After the antagonistic phenotypes are observed, the inhibition regions are subjected to IMS screening to visualize the metabolites. The microbial metabolites with a spatial distribution that is superimposed with the zone of inhibition are considered to be potential anti-infective candidates. The exact mass and isotope pattern provided by high resolution instruments such as MALDI-TOF-IMS, in combination with MS/MS network analysis can help estimate the structural novelty of the induced metabolites.^{4,35} This technique was applied in individual co-cultures of the pathogens *Staphylococcus aureus* and *Staphylococcus epidermidis* with *Streptomyces roseosporus* NRRL 15998. Using matrix-assisted laser desorption ionization (MALDI) imaging mass spectrometry, it was possible to observe the occurrence of three molecules (*m/z* 863, 877, and 891) that were responsible for the inhibition of both *Staphylococcus* pathogens. These stress-induced ions were subsequently identified as arylomycins, a class of broad-spectrum antibiotics that target type I signal peptidases.³⁶

4. Metabolomics reveals naturally occurring microbial interactions

Throughout thousands of years of evolution and survival adaption, microorganisms have established all types of close relationships with other organisms, including symbiosis, parasitism, competition and antagonism. Revealing the underneath molecular mechanisms involved in these “invisible” interactions would presumably contribute to the discovery of new chemical entities with antibiotic activity, because virulence factors in their natural contexts might turn out to have beneficial applications in human anti-infective therapy.³⁷

Particularly, deciphering the microbial pathogenesis of some ecologically relevant systems, such as associations with plants, mushrooms, insects, and sponges,³⁸ opens new perspectives for the exploration for new antimicrobials, as exemplified by the discovery of the novel antibiotics, rhizoxin³⁹ and desoxyhavannahine (Figure 2).⁴⁰ A very promising but yet quite unexplored reservoir for new antibiotics discovery is the insect-associated microbiome. As reviewed by Helge B. Bode,⁴¹ insects could be true pioneers as providers of anti-infectives for human beings. Paenilamicins (Figure 2), a class of novel hybrid nonribosomal peptide/polyketide antibiotics, were discovered in the bee pathogen, *Paenibacillus larvae*. Bee larval co-infection assays revealed that the paenilamicins are employed by *P. larvae* to fight ecological niche competitors and are not directly involved in mortality of the bee larvae.⁴²

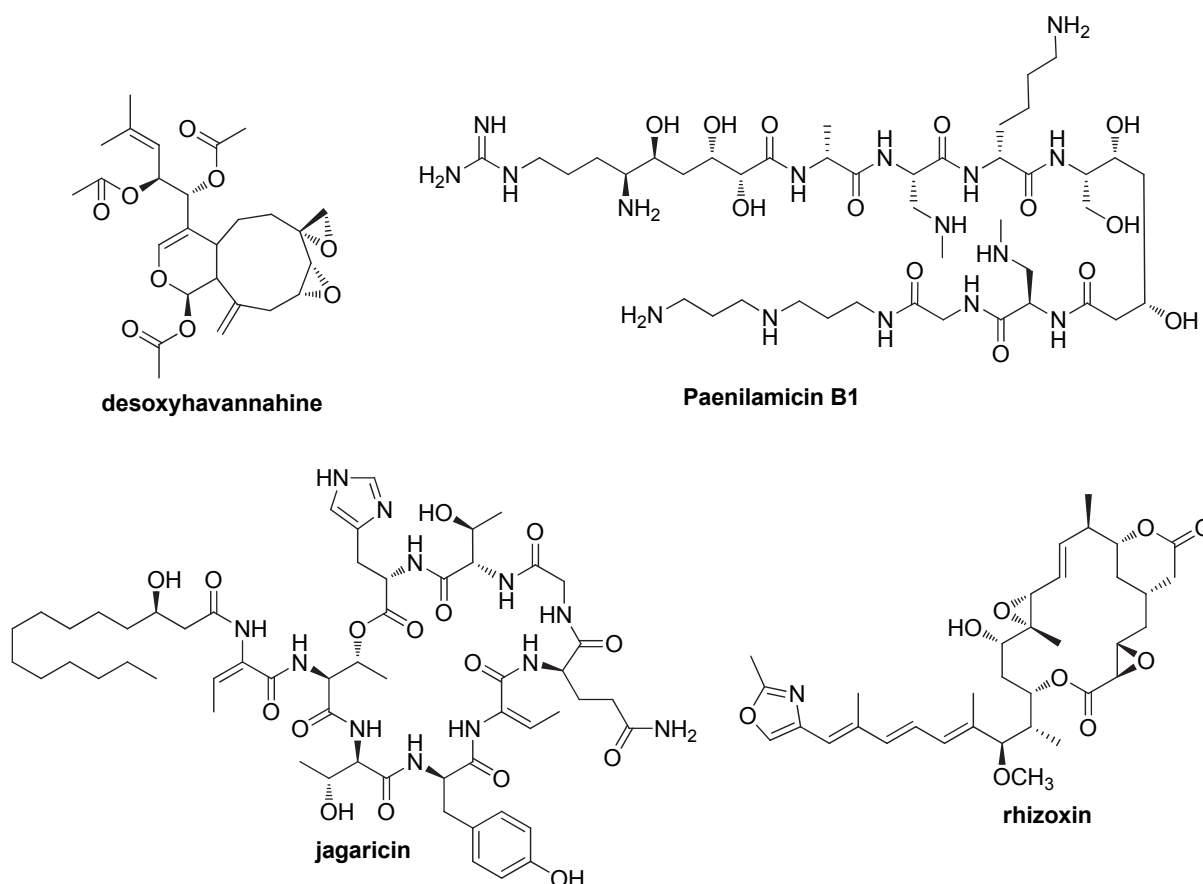


Figure 2. Structures of several novel antibiotics discovered through the investigation of microbial interactions.

For cultivable symbionts, it could be expected that secondary metabolites produced in a natural context are probably different from those in a standard laboratory setting, since the growth condition considerably influences microbial secondary metabolism. Many cryptic biosynthetic gene clusters encoding molecules, remain silent in the laboratory setting due to the lack of appropriate environmental stimuli, and it is extremely challenging to mimic their natural environment in a laboratory. More so, many symbiotic microbes are not even amenable to artificial cultivation. In other words, antibiotics that would occur due to *in situ* microbial interactions are impossible to be discovered in a traditional culture-dependent

approach. Therefore, it is crucial to find a way to shed light on “hidden antibiotics” in a natural setting. The advances in metabolomics provide an opportunity to make what appeared to be impossible possible, using for example the above described IMS metabolome profiling techniques. The motile Gram-negative bacterium *Janthinobacterium agaricidamnosum* is known to be the causative agent of soft rot disease in the cultured button mushroom *Agaricus bisporus*. Driven by the hypothesis that the pathobiology of this mushroom pathogen was potentially correlated to antifungal search, MALDI-IMS was used for *in situ* profiling of the secondary metabolites harbored in the mushroom-associated ecological system of *Agaricus bisporus* vs *Janthinobacterium agaricidamnosum*. The visualization of a unique ion of m/z 1181 $[M+H]^+$ in the diseased region of the mushroom prompted the efforts to determine the structure of this compound and it was identified as an unprecedented cyclic lipopeptide product, jagaricin (Figure 2). Jagaricin turned out to be a fungicide that is highly active against human pathogenic fungi including *Candida albicans*, *Aspergillus fumigatus*, and *Aspergillus terreus* at submicromolar concentrations. Notably, jagaricin production by the bacterium *J. agaricidamnosum* tended to occur exclusively in the mushroom (*A. bisporus*) landscape and was otherwise impaired under a variety of regular laboratory cultivation conditions.⁴³ This example serves as a strong support for our arguments that understanding naturally-occurring microbial pathogenesis is beneficial for antibiotic search; and that IMS-based metabolomics can be applied to reveal the cryptic antibiotics that are only produced in a particular natural environment.

Revealing cryptic new antibiotics hidden in environmental samples might benefit from the surge in the development of single cell metabolomics, a tool used in metabolic analysis at cellular and subcellular level. As relevant perspectives of single cell metabolomics have been thoroughly reviewed just recently,^{44–46} our focus is to discuss its potential use in unveiling the metabolic communication between symbioses and hosts. Based on the morphological phenotypes or other properties such as fluorescence, the selective metabolic analysis of certain microbial cells in environmental samples could be used to view the unique secondary metabolites generated by involved microorganisms during the naturally occurring interspecies interactions. For instance, halogen *in situ* hybridization-secondary ion mass spectroscopy (HISH-SIMS) has allowed the simultaneous identification and quantitation of metabolic activities of environmental microbial assemblage at single-cell level without the need for traditional microbial separation and laboratory cultivation.⁴⁷ Single-cell mass spectrometry (SCMS) is a rapidly growing field in analytical chemistry, and there are a number of examples of its applicability to different cells.⁴⁴ In this case, the possibility of accurate mass measurement and acquiring structural information as provided by tandem MS, facilitate the identification of numerous metabolites from a single cell.

5. CONCLUSION

Metabolomics methods provide an innovative strategy for the discovery of new antibiotics from microbial sources. Allowing the comprehensive chemometric comparison of secondary metabolomes, it makes the dream of “finding a needle in a haystack” more realistic. The ability to measure the entire metabolomes within microorganisms that interact with hosts expressing cryptic gene clusters can be crucial for the exploitation of these untapped

reservoirs of potentially new anti-infectives.⁴⁸ Thus, metabolomics can be complementary to metagenomics methodology in accessing previously inaccessible natural products,⁴⁹ even has the advantage of delving into the biosynthetic potential of uncultivable microorganisms when applied to *in situ* analysis of live, undamaged biological tissues.

REFERENCES

- (1) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat. Rev. Drug Discov.* **2007**, *6*, 29–40.
- (2) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Trends Biotechnol.* **2011**, *29*, 267–275.
- (3) Verpoorte, R.; Choi, Y. H.; Kim, H. K. *Phytochem. Rev.* **2007**, *6*, 3–14.
- (4) Watrous, J.; Roach, P.; Alexandrov, T.; Heath, B. S.; Yang, J. Y.; Kersten, R. D.; van der Voort, M.; Pogliano, K.; Gross, H.; Raaijmakers, J. M.; Moore, B. S.; Laskin, J.; Bandeira, N.; Dorrestein, P. C. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1743–1752.
- (5) Sidebottom, A. M.; Johnson, A. R.; Karty, J. a; Trader, D. J.; Carlson, E. E. *ACS Chem. Biol.* **2013**, *8*, 2009–2016.
- (6) Grkovic, T.; Pouwer, R. H.; Vial, M.-L.; Gambini, L.; Noël, A.; Hooper, J. N. a; Wood, S. a; Mellick, G. D.; Quinn, R. J. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 6070–6074.
- (7) Williams, D. E.; Dalisay, D. S.; Patrick, B. O.; Matainaho, T.; Andrusiak, K.; Deshpande, R.; Myers, C. L.; Piotrowski, J. S.; Boone, C.; Yoshida, M.; Andersen, R. J. *Org. Lett.* **2011**, *13*, 3936–3939.
- (8) Geng, C.-A.; Chen, X.-L.; Zhou, N.-J.; Chen, H.; Ma, Y.-B.; Huang, X.-Y.; Zhang, X.-M.; Chen, J.-J. *Org. Lett.* **2014**, *16*, 370–373.
- (9) Kersten, R. D.; Yang, Y.-L.; Xu, Y.; Cimermancic, P.; Nam, S.-J.; Fenical, W.; Fischbach, M. a; Moore, B. S.; Dorrestein, P. C. *Nat. Chem. Biol.* **2011**, *7*, 794–802.
- (10) Pidot, S.; Ishida, K.; Cyrulies, M.; Hertweck, C. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 7856–7859.
- (11) Krug, D.; Zurek, G.; Schneider, B.; Garcia, R.; Müller, R. *Anal. Chim. Acta* **2008**, *624*, 97–106.
- (12) Hou, Y.; Braun, D. R.; Michel, C. R.; Klassen, J. L.; Adnani, N.; Wyche, T. P.; Bugni, T. S. *Anal. Chem.* **2012**, *84*, 4277–4283.
- (13) Carr, G.; Poulsen, M.; Klassen, J. L.; Hou, Y.; Wyche, T. P.; Bugni, T. S.; Currie, C. R.; Clardy, J. *Org. Lett.* **2012**, *14*, 2822–2825.
- (14) Hou, Y.; Tianero, M. D. B.; Kwan, J. C.; Wyche, T. P.; Michel, C. R.; Ellis, G. a; Vazquez-Rivera, E.; Braun, D. R.; Rose, W. E.; Schmidt, E. W.; Bugni, T. S. *Org. Lett.* **2012**, *14*, 5050–5053.
- (15) Wyche, T. P.; Piotrowski, J. S.; Hou, Y.; Braun, D.; Deshpande, R.; McIlwain, S.; Ong, I. M.; Myers, C. L.; Guzei, I. a.; Westler, W. M.; Andes, D. R.; Bugni, T. S. *Angew. Chemie Int. Ed.* **2014**, *53*, 11583–11586.
- (16) Seger, C.; Godejohann, M.; Tseng, L.-H.; Spraul, M.; Girtler, A.; Sturm, S.; Stuppner, H. *Anal. Chem.* **2005**, *77*, 878–885.

-
- (17) Glauser, G.; Guillarme, D.; Grata, E.; Boccard, J.; Thiocone, A.; Carrupt, P.-A.; Veuthey, J.-L.; Rudaz, S.; Wolfender, J.-L. *J. Chromatogr. A* **2008**, *1180*, 90–98.
- (18) Bertrand, S.; Schumpp, O.; Bohni, N.; Monod, M.; Gindro, K.; Wolfender, J.-L. *J. Nat. Prod.* **2013**, *76*, 1157–1165.
- (19) Wolfender, J.-L.; Queiroz, E. F.; Hostettmann, K. *Magn. Reson. Chem.* **2005**, *43*, 697–709.
- (20) Schroeder, F. C.; Gronquist, M. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 7122–7131.
- (21) Bertrand, S.; Bohni, N.; Schnee, S.; Schumpp, O.; Gindro, K.; Wolfender, J.-L. *Biotechnol. Adv.* **2014**, *32*, 1180–1204.
- (22) Bertrand, S.; Schumpp, O.; Bohni, N.; Bujard, A.; Azzollini, A.; Monod, M.; Gindro, K.; Wolfender, J.-L. *J. Chromatogr. A* **2013**, *1292*, 219–228.
- (23) Ola, A.; Thomy, D. *J. Nat. Prod.* **2013**, *76*, 2094–2099.
- (24) Cueto, M.; Jensen, P.; Kauffman, C.; Fenical, W.; Lobkovsky, E.; Clardy, J. *J. Nat. Prod.* **2001**, *64*, 1444–1446.
- (25) Oh, D.-C.; Kauffman, C. a; Jensen, P. R.; Fenical, W. *J. Nat. Prod.* **2007**, *70*, 515–520.
- (26) Glauser, G.; Gindro, K.; Fringeli, J.; De Joffrey, J.-P.; Rudaz, S.; Wolfender, J.-L. *J. Agric. Food Chem.* **2009**, *57*, 1127–1134.
- (27) Scherlach, K.; Busch, B.; Lackner, G.; Paszkowski, U.; Hertweck, C. *Angew. Chemie* **2012**, *124*, 9753–9756.
- (28) Watrous, J. D.; Dorrestein, P. C. *Nat. Rev. Microbiol.* **2011**, *9*, 683–694.
- (29) Krug, D.; Müller, R. *Nat. Prod. Rep.* **2014**, *31*, 768–783.
- (30) Shih, C.-J.; Chen, P.-Y.; Liaw, C.-C.; Lai, Y.-M.; Yang, Y.-L. *Nat. Prod. Rep.* **2014**, *31*, 739–755.
- (31) Hsu, C.-C.; ElNaggar, M. S.; Peng, Y.; Fang, J.; Sanchez, L. M.; Mascuch, S. J.; Møller, K. a; Alazzeh, E. K.; Pikula, J.; Quinn, R. a; Zeng, Y.; Wolfe, B. E.; Dutton, R. J.; Gerwick, L.; Zhang, L.; Liu, X.; Månsson, M.; Dorrestein, P. C. *Anal. Chem.* **2013**, *85*, 7014–7018.
- (32) Watrous, J.; Hendricks, N.; Meehan, M.; Dorrestein, P. C. *Anal. Chem.* **2010**, *82*, 1598–1600.
- (33) Liu, W.-T.; Yang, Y.-L.; Xu, Y.; Lamsa, A.; Haste, N. M.; Yang, J. Y.; Ng, J.; Gonzalez, D.; Ellermeyer, C. D.; Straight, P. D.; Pevzner, P. a; Pogliano, J.; Nizet, V.; Pogliano, K.; Dorrestein, P. C. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16286–16290.
- (34) Hoe, B. C.; Gorzelnik, K. V.; Yang, J. Y.; Hendricks, N.; Dorrestein, P. C. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 13082–13087.
- (35) Traxler, M. F.; Watrous, J. D.; Alexandrov, T.; Dorrestein, P. C.; Kolter, R. *MBio* **2013**, *4*, 1–12.
- (36) Liu, W.; Kersten, R.; Yang, Y. *J. Am. Chem. Soc.* **2011**, *133*, 18010–18013.
- (37) Schmidt, E. W. *Nat. Chem. Biol.* **2008**, *4*, 466–473.
- (38) Pidot, S. J.; Coyne, S.; Kloss, F.; Hertweck, C. *Int. J. Med. Microbiol.* **2014**, *304*, 14–22.
- (39) Partida-Martinez, L. P.; Hertweck, C. *Nature* **2005**, *437*, 884–888.
- (40) Kelman, D.; Kashman, Y.; Hill, R. T.; Rosenberg, E.; Loya, Y. *Pure Appl. Chem.* **2009**, *81*, 1113–1121.
- (41) Bode, H. B. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 6394–6396.

-
- (42) Müller, S.; Garcia-Gonzalez, E.; Mainz, A.; Hertlein, G.; Heid, N. C.; Mösker, E.; van den Elst, H.; Overkleeft, H. S.; Genersch, E.; Süßmuth, R. D. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 1–6.
- (43) Graupner, K.; Scherlach, K.; Bretschneider, T.; Lackner, G.; Roth, M.; Gross, H.; Hertweck, C. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 13173–13177.
- (44) Rubakhin, S. S.; Lanni, E. J.; Sweedler, J. V. *Curr. Opin. Biotechnol.* **2013**, *24*, 95–104.
- (45) Trouillon, R.; Passarelli, M. K.; Wang, J.; Kurczy, M. E.; Ewing, A. G. *Anal. Chem.* **2013**, *85*, 522–542.
- (46) Zenobi, R. *Science (80-.)*. **2013**, *342*, 1201–1211.
- (47) Musat, N.; Halm, H.; Winterholler, B.; Hoppe, P.; Peduzzi, S.; Hillion, F.; Horreard, F.; Amann, R.; Jørgensen, B. B.; Kuypers, M. M. M. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17861–17866.
- (48) Perić-Concha, N.; Long, P. F. *Drug Discov. Today* **2003**, *8*, 1078–1084.
- (49) Brady, S. F.; Simmons, L.; Kim, J. H.; Schmidt, E. W. *Nat. Prod. Rep.* **2009**, *26*, 1488–503.

Chapter 3

Metabolic profiling as a tool for prioritizing antimicrobial compounds

Changsheng Wu ^{1,2}, Young Hae Choi ², and Gilles P. van Wezel ^{1,3}

¹ *Molecular Biotechnology, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands;*

² *Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands;*

³ *Department of Microbial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, The Netherlands;*

J. Ind. Microbiol. Biotechnol. **2016**, 43, 299–312

ABSTRACT

Metabolomics is an analytical technique that allows scientists to globally profile low molecular weight metabolites between samples in a medium- or high-throughput environment. Different biological samples are statistically analysed and correlated to a bioactivity of interest, highlighting differentially produced compounds as potential biomarkers. Here, we review NMR- and MS-based metabolomics as technologies to facilitate the identification of novel antimicrobial natural products from microbial sources. Approaches to elicit the production of poorly expressed (cryptic) molecules are thereby key to allow statistical analysis of samples to identify bioactive markers, while connection of compounds to their biosynthetic gene cluster is a determining step in elucidating the biosynthetic pathway and allows downstream process optimization and upscaling. The review focuses on approaches built around NMR-based metabolomics, which enables efficient dereplication and guided fractionation of (antimicrobial) compounds.

1. INTRODUCTION

The discovery and development of antibiotics to treat bacterial infections is one of the greatest triumphs of modern medicine. However, the exponential increase of antimicrobial resistance means that bacterial infections now once more pose a major threat to human health.¹ The high frequency of re-discovery of known molecules thereby frustrates screening efforts, and this necessitates new approaches to bolster the antibiotic pipelines.²⁻⁵ Filamentous fungi and bacteria of the order of *Actinomycetales* are the major producers of biomedical and agricultural natural products, and these microorganisms are responsible for producing the vast majority of the known antibiotics.^{6, 7} Some two thirds of all antibiotics are produced by actinomycetes, the majority of which are sourced by members of the genus *Streptomyces*. However, this is likely to represent only a tiny portion of the repertoire of total chemical space of bioactive compounds that these microorganisms may produce.^{6, 7} Firstly, while the number of filamentous microbes in soil and marine environments is enormous, the bulk of them resist cultivation in the laboratory and thus escape screening and exploitation.⁸ Recent work showed that enabling growth of these ‘uncultivable’ microorganisms might open up a new area of the chemical space of natural products.^{9, 10} Secondly, many natural product biosynthetic gene clusters (BGCs) of cultivable microbes remain silent during standard laboratory cultivation.^{11, 12} As an example, genome sequencing of actinomycetes revealed that the producing capacity of even the best-studied model organisms has been grossly underestimated.¹³⁻¹⁶ And thirdly, those gene clusters that are expressed, in particular if expression is relatively low as compared to other bioactive compounds produced by the same organism, often escape discovery due to the very demanding chemical analysis or chromatographic purification that is required.¹⁷ New approaches are therefore required to exploit the huge unexplored reservoir of bioactive natural products and rejuvenate the drug-discovery pipelines.

The preferred method of identifying compounds in complex mixtures is metabolomics, which is a high-throughput analytical technique that offers a global analysis of the products of all cellular metabolic reactions, regardless of the reactions that lead to their production.¹⁸ Spectroscopic techniques, and in particular nuclear magnetic resonance (NMR) and mass spectrometry (MS), are applied to chemically profile the metabolites that are produced by microorganisms. In connection with a wide range of chemometric methods, the metabolic differences among experimental groups are subsequently unraveled by comparison.^{19, 20} In this review, we discuss ways to enforce fluctuations in the production of bioactive compounds, and the application of metabolomics to facilitate uncovering those compounds, whereby we zoom in on NMR-based metabolic profiling methods, which are less commonly applied in microbial drug discovery.

2. APPROACHES TO ELICIT ANTIBIOTIC PRODUCTION

Genome mining of actinomycetes on average reveals the presence of 30-40 secondary metabolite biosynthetic gene clusters (BGCs) per species.²¹ Only a small number of these gene clusters has been matched to a natural product. A major problem is that many BGCs remain silent or poorly expressed under routine screening conditions. If we are to even start attempting to visualize these compounds by metabolomics approaches, their activation is the first essential step. Generic approaches to activate silent antibiotics to uncover the hidden chemical diversity include: (i) perturbed growth media with additives,^{22, 23} (ii) inducing

antibiotic resistance, *e.g.* to streptomycin²⁴ and rifampicin,²⁵⁻²⁷ (iii) the addition of chemical elicitors, such as *N*-acetylglucosamine,²⁸ γ -butyrolactones,^{29, 30} or histone deacetylase inhibitors;³¹ and (iv) microbial cocultivation.³²⁻³⁴ Instead of inducing the production of an antibiotic, an interesting alternative may be offered by screening for inducible resistance, in particular associated with classes of antibiotics that are less frequently identified in classical screens. Growth on media containing vancomycin was used to enrich a library of actinomycetes for producers of glycopeptide antibiotics, and subsequent phylogeny-based screening according to specific fingerprints of biosynthetic genes, lead to the identification of the novel glycopeptide antibiotic pekiskomycin.³⁵

In terms of the activation of antibiotic production, a concept study was recently done using a collection of some 800 strains that were grown under 40 different growth conditions and then assayed for changes in the production of antimicrobial activity against the so-called ESKAPE pathogens.²³ This showed that many strains that fail to produce antibiotics under routine growth conditions, produce bioactivities that inhibits the growth of MDR pathogens under specific nutritional conditions or after addition of elicitors.²³

2.1. Harnessing the regulatory mechanisms that control antibiotic biosynthesis

To develop new directed approaches, it is imperative that we better understand the underlying regulatory networks. The regulation of antibiotic production involves multiple regulatory cascades and networks. Here, we will guide the reader with some general principles gleaned from the model organism *Streptomyces coelicolor*; for more extensive reviews on the control of antibiotic production we refer the reader elsewhere.³⁶⁻³⁹

Antibiotic BGCs specify biosynthesis and modification, resistance, transport and pathway-specific control. The best-studied example of cluster-encoded, pathway-specific activators is that of the *Streptomyces* antibiotic regulatory proteins (SARPs). SARPs typically bind to direct repeats in promoter regions of biosynthetic genes, thereby activating transcription of some or all of the antibiotic biosynthetic genes in the cluster.⁴⁰ In *S. coelicolor*, the pathway-specific regulatory genes *actII-ORF4*, *cdaR* and *redD* encode the SARPs for actinorhodin (ACT), calcium-dependent antibiotic (CDA) and undecylprodigiosin (RED) production, respectively. Of these, *actII-ORF4* and *redD* have been extensively studied. Deletion of these genes abolishes the production of actinorhodin or undecylprodigiosin, respectively,⁴¹ and their transcription is activated in a growth phase-dependent manner.^{42, 43} Interestingly, positioning *redD* behind a developmental or nitrogen-responsive promoter is sufficient to produce undecylprodigiosin in aerial hyphae or bring it under nitrogen control, respectively, which suggests a lack of downstream control.⁴⁴ Hence, one effective approach to activate gene clusters is overexpression of their activator genes, but this involves genetic manipulation which is not amenable to high-throughput approaches.

Recent years have underlined the importance of global regulators as a higher order regulatory network, and their possible application for the activation of antibiotic production. Many pleiotropic regulators characterized so far are required only under specific environmental conditions.^{45, 46} The nutrient sensory GntR-family regulator DasR controls among others aminosugar metabolism and transport, the chitinolytic system and antibiotic production.^{28, 47-51} DasR is a highly pleiotropic regulator, as demonstrated by recent systems biology analysis of chitin- or *N*-acetylglucosamine-induced cultures of *S. coelicolor*.^{52, 53}

DasR directly controls the transcription of *actII-ORF4* and *redD* as well as *redZ*, a response regulator required for the activation of *redD* and thus of undecylprodigiosin production.²⁸ Systems-wide DNA binding experiments using ChIP-chip analysis revealed that in *S. coelicolor* DasR likely controls all pathway-specific activator genes,⁵³ but this appears to be more an exception than a rule. The activity of DasR is modulated by metabolic derivatives of *N*-acetylglucosamine, and addition of this amino sugar to the culture media activates antibiotic production in several actinomycetes, including cryptic antibiotics.²⁸

Interesting cross-talk is seen for DasR with other higher order antibiotic regulators. One is the TetR-family regulatory protein AtrA that in *S. coelicolor* is required for the transcription of *actII-ORF4*⁵⁴ and in *S. griseus* of *strR*, the pathway-specific activator gene of streptomycin production.⁵⁵ The antibiotic activating role of AtrA may be widespread, as suggested by recent data that AtrA also controls biosynthesis of the important lipoglycopeptide antibiotic daptomycin in *Streptomyces roseosporus*.⁵⁶ AtrA is controlled by the level of phosphate and repressed by the PhoRP system.⁵⁷ AtrA activates development⁵⁸ as well as *N*-acetylglucosamine import,⁵⁹ and thus antagonizes the repressing activities of DasR, whereby the metabolic balance likely plays a deciding role on the net outcome. Further complexity is offered by Rok7B7, a member of the ROK family of proteins, which are predominantly sugar regulatory proteins and sugar kinases, including glucose kinase.^{60, 61} Rok7B7 also pleiotropically affects primary and secondary metabolism, and is like AtrA required for actinorhodin production.^{62, 63} Recent evidence suggests that Rok7B7 may be activated by a derivative of the C5 sugar xylose.⁶³

Other pleiotropic antibiotic regulators involved in the control of *actII-ORF4* are the large SARP-family regulator AfsR and the phosphate regulator PhoP. AfsR derives its name from a putative relationship to synthesis of the hormone-like signaling molecule A-factor, belonging to the γ -butyrolactones. The A-factor-responsive AdpA controls the onset of development and antibiotic production in *S. griseus*,^{64, 65} but the precise relationship with AfsR (if any) so far remains unresolved. In *S. coelicolor*, AfsR is conditionally required for Act and Red production⁴¹ and its control is somehow transmitted through control of the small downstream gene *afsS*, thereby activating its transcription.⁶⁶ While the precise function of *afsS* is unclear, similarly to AtrA, its pleiotropic and positive effect on antibiotic production makes it an attractive target for the activation of antibiotic production. PhoP represses actinorhodin production in response to phosphate,⁶⁷ and it may also relay its control via *afsS*.⁵⁷ These are examples to highlight the complexity of the control of antibiotic production, whereby an astonishing number of around 15 regulatory proteins have been shown to control *actII-ORF4* alone.⁴⁶ Improved understanding of these multiple and intertwined regulatory networks will allow scientists to design new approaches to activate poorly expressed BGCs.

2.2. Connecting eliciting approaches to metabolomics-driven lead discovery

In the framework of this review, the activation of gene expression is important to activate silent gene clusters and particularly to achieve significant fluctuation of the production of (cryptic) antibiotics and thus differential bioactivity between cultures; this serves as enabling technology for statistical correlation between a bioactivity of interest and the responsible (sought-after) compounds. Having established such differential expression, the next challenge is to efficiently identify the induced antimicrobial compounds against a background of an

inevitably highly complex metabolome matrix.

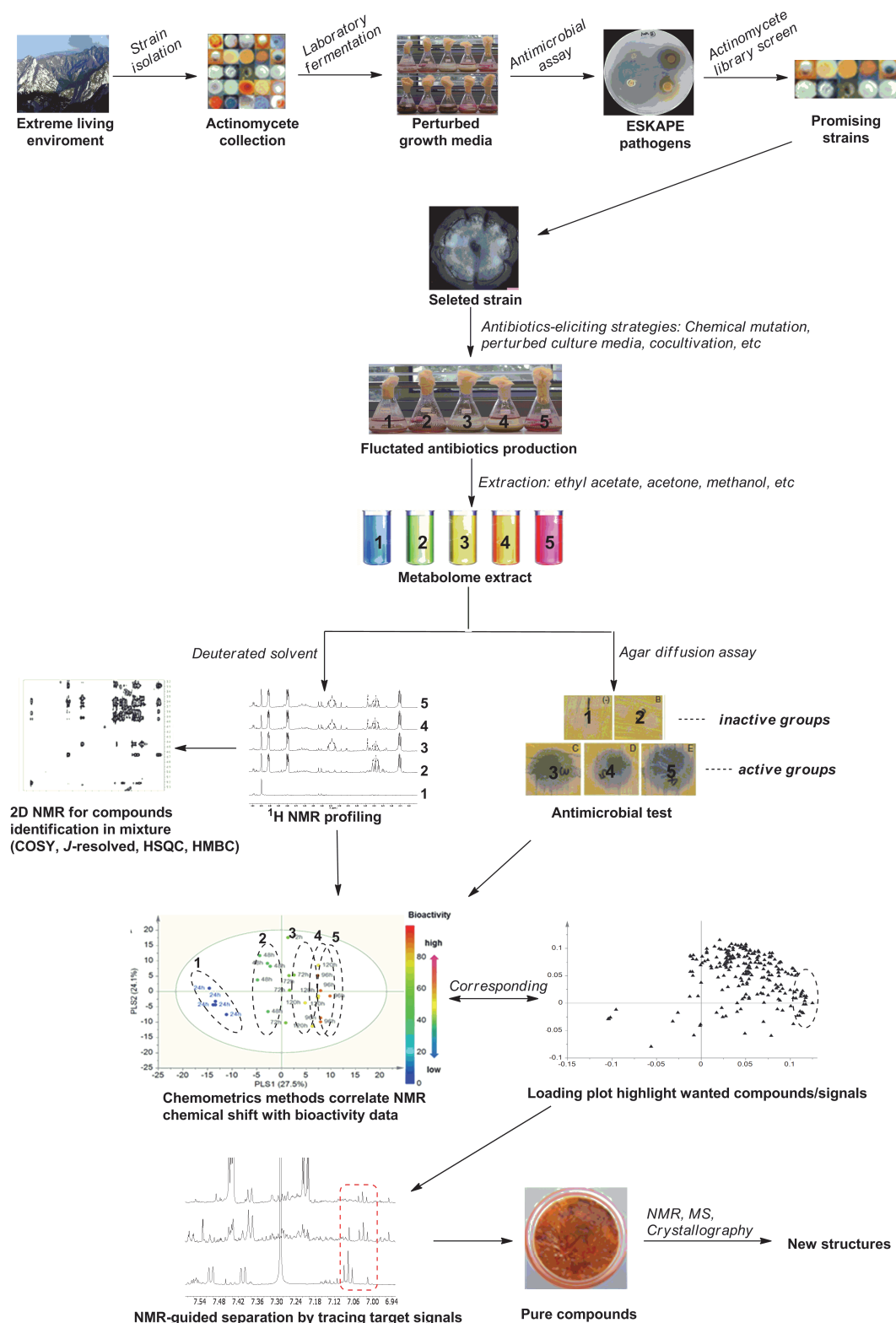


Figure 1. A pipeline for the discovery of antibiotics produced by actinomycetes based on NMR-based metabolomics. Strains are subjected to eliciting approaches to activate the production of poorly expressed antimicrobials, followed by NMR-based metabolomics to identify the bioactive molecules. ^1H NMR profiling of active and inactive groups, aided by 2D NMR, allows dereplication of known molecules and chemometric methods then correlate proton signals in the active fractions to the bioactive compound(s).

Metabolomics is an effective tool for facilitating the discovery of new antibiotics, because it allows the multivariate comparison of active and inactive metabolic samples, highlighting particularly the differentially produced compounds which serve as potential biomarkers, thereby avoiding chemical redundancy in the very early stage.⁶⁸⁻⁷⁰ Integrating metabolomics approaches with new eliciting strategies should allow scientists to streamline their drug-discovery pipeline, so as to prioritize novel molecules over repeated (and frustrating) isolation of known molecules. A schematic representation of a feasible antibiotics-discovery pipeline based on NMR-based metabolomics is presented in Figure 1. Compounds can be readily identified in complex biological matrices without time-consuming chromatographic separation, further aided by 2D NMR experiments. In the next section, we detail some of the NMR techniques that are most appropriate for metabolomics approaches.

3. METABOLOMICS OF NATURAL PRODUCTS BASED ON NMR SPECTROSCOPY

NMR techniques that are most frequently used in NMR metabolomics are 1J -resolved, COSY, TOCSY, HSQC, and HMBC.⁷⁰⁻⁷³ ^1H NMR signals corresponding to molecules that exist in both the active and the inactive groups are discarded based on lack of statistical relevance.⁷⁴ Chemometric analysis is then used for further data mining of the ^1H NMR data, extracting information that is intractable to visual detection. In principle, the methodology of NMR measurements, spectral data preprocessing, and multivariate data analysis basically follow a protocol that was originally designed for analysis of plant metabolites.⁷¹ The ^1H NMR signals correlated with the better bioactivity and separation of the active from the inactive groups are used as probing tools for NMR-guided isolation of the sought-after bioactivity. Advantages of NMR-based over MS-based (LC-MS or GC-MS) metabolomics are: (i) the chemical shifts and splitting patterns of proton resonance data provide partial structural information that allows early assessment of the type of compound causing the bioactivity, thus aiding in avoiding redundancy; (ii) the analysis of crude mixtures can be done in very short time (in the order of minutes), with high signal robustness (*i.e.* minimum technical variation), so as to guide scientists in a medium- to high-throughput environment. These advantages should be weighed against the lack of true high-throughput application which is a key selling point of MS-based approaches. As soon as the major interfering signals are removed from the NMR spectra, the material (even if not yet pure) is likely of sufficient quality for full structure elucidation. In some cases, structural elucidation of unknown compounds is challenging, especially for those entities with unprecedented molecular scaffold. Two-dimensional NMR (2D NMR) techniques have formed the foundation of most contemporary approaches for structure elucidation,⁷⁵ allowing assembly of the molecular architecture by defining atom-to-atom connectivity. In general, NMR-based metabolomics uses a combination of 2D NMR techniques for deconvoluting metabolites in congested mixtures, namely J -resolved, ^1H -based 2D (COSY and TOCSY), ^{13}C -based 2D NMR (HSQC and HMBC).⁷⁵ J -resolved solves the signal purity of ^1H NMR by showing splitting pattern with coupling constants in F2 axis which requires for chemical structures of corresponding signals. COSY and TOCSY give information of ^1H connectivity within the same molecule. Along with the connectivity, the ^1H -based COSY and TOCSY also help signal identification in overlapped regions. Many signals are overlapped which obscures some of the peaks in the mixture, but these can be deconvoluted by the correlations in the 2D NMR spectra. ^{13}C -related spectra such as HSQC

and HMBC give more detailed structural information, particularly for the carbon skeleton. Although important for identifying the carbon skeleton of organic compounds, ^{13}C -NMR is rarely applied in NMR-based metabolomics because of the low abundance of ^{13}C in nature and its long relaxation time, and also because ^{13}C broad-band decoupling cannot be used quantitatively. HSQC and HMBC are therefore primarily applied to support the ^1H NMR data.

Deciphering the structure of organic molecules by using conventional routine suites, based on for example J -resolved-COSY (or TOCSY), HSQC and HMBC⁷⁶ often gives rise to misassignment. This partly arises from the inherent challenges of differentiating two-bond heteronuclear ($\text{H}\cdots\text{C}$, $^2J_{\text{CH}}$) correlations from three-bond correlations ($^3J_{\text{CH}}$) in the HMBC spectrum. In addition, as HMBC normally just provides $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ but rarely $^4J_{\text{CH}}$ correlations, structure elucidation becomes progressively challenging the more molecules become proton-deficient.⁷⁷ To address these problems, many other approaches have been tried such as combined HSQC-TOCSY.⁷⁸ Researchers from Merck proposed that adequate sensitivity double-quantum (1, n -ADEQUATE) NMR spectroscopy should be reconsidered as a complement for HMBC in structure elucidation.⁷⁹ 1,1-ADEQUATE exclusively presents $^1J_{\text{CC}}$ correlations and can help distinguish $^2J_{\text{CH}}$ from $^3J_{\text{CH}}$ correlations in the HMBC spectrum, which would be a powerful method for determining the proton-rich small molecules by using a combination of 1,1-ADEQUATE + HMBC.^{80, 81} 1, n -ADEQUATE experiments predominantly yield homonuclear $^3J_{\text{CC}}$ correlations (equal to $^4J_{\text{CH}}$) that are difficult to obtain with HMBC NMR. This feature makes 1, n -ADEQUATE a complement for HMBC in defining proton-deficient structures, such as a large group of aromatic polyketides produced by actinomycetes.^{80, 82} Furthermore, newly developed inverted $^1J_{\text{CC}}$ 1, n -ADEQUATE spectroscopy allows the integration of 1,1-ADEQUATE and 1, n -ADEQUATE information in a single spectrum, while they remain distinguished by plotting the output in different colors (Figure 2).⁸³

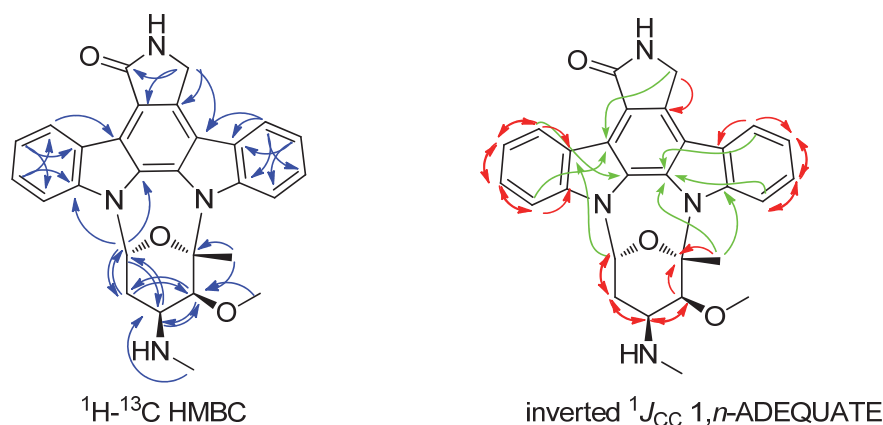


Figure 2. Diagram of HMBC ($\rightarrow$), and inverted $^1J_{\text{CC}}$ 1, n -ADEQUATE correlations by using the proton-deficient alkaloid staurosporine as an example. The skeleton of this model compound can be unequivocally assigned by HMBC + inverted $^1J_{\text{CC}}$ 1, n -ADEQUATE strategy. In inverted $^1J_{\text{CC}}$ 1, n -ADEQUATE, $^1J_{\text{CC}}$ correlations are shown as $\rightarrow$, while important $^3J_{\text{CC}}$ correlations are shown as $\rightarrow$.

Though NMR spectroscopy is undoubtedly powerful in the structure elucidation, it works better on pure organic compounds, as otherwise signal-overlapping issues come into play. To address this, chromatography-hyphenated NMR (HPLC-NMR) has been introduced into metabolic analysis.⁸⁴ The major advantages of on-flow HPLC-NMR lie in the real-time

access to structure information of chromatographic peaks by using NMR spectrometers as detector.⁸⁵ To balance the inherent lower sensitivity of NMR spectra and the limited injection volume onto the HPLC unit, solid-phase-extraction (SPE) is coupled to trap the analytes onto an identical SPE cartridge when repeated HPLC analysis, followed by final elution with a deuterated solvent into miniaturization tube for NMR analysis.⁸⁶ By using this LC-SPE-NMR approach,^{87, 88} comprehensive structural information about chemical composition of crude extracts can be obtained online, thus enabling early and prospective assessment of the valuable components of an extract or fraction.⁸⁹ Furthermore, when integrating MS into the system, like LC-DAD-SPE-MS-NMR instrumentation,^{90, 91} the architecture of novel compounds can be elucidated *de novo* and online without laborious purification of the analyte(s) from its highly complex matrix.

4. CONNECTING GENOTYPE TO CHEMOTYPE: INTEGRATING GENOMICS WITH METABOLOMICS

Before looking at applications of metabolomics to identify novel compounds, in this section we will look into how metabolomics may aid in connecting a natural product of interest to its BGC. This is a crucial step as it facilitates molecular approaches such as directed mutagenesis, overexpression of pathway-specific regulators and heterologous expression. Furthermore, identification of the gene cluster is often required for determining the precise biosynthetic pathway.

Next-generation sequencing (NGS) technologies have uncovered the genetic architecture of thousands of BGCs, and this has revolutionized the drug-discovery approaches.⁹² Exploiting this rapidly increasing source of information, a recent network analysis of the phylogeny and distribution of BGCs in microbial genomes revealed a vast number of yet underexplored genetic resources, belonging to many hundreds of gene cluster families.⁹³

Gene clusters such as those for polyketides (PKS) or nonribosomal peptides (NRPS) are readily identified using bioinformatics, and to some extent the domain structures of the biosynthetic proteins allows prediction of the molecule that they specify.^{94, 95} In recent years, excellent bioinformatics packages have been developed for the identification of BGCs. These include antiSMASH (ANTIbiotics & Secondary Metabolite Analysis SHell)⁹⁶ and SMURF (Secondary Metabolite Unknown Regions Finder), as well as packages that enable the identification of biosynthesis genes for specific subclasses of NPs (reviewed in⁹⁷). Examples are SBSPKS (Structure Based Sequence Analysis of Polyketide Synthases) for polyketide BGCs,⁹⁸ NRPSPredictor for nonribosomal peptide gene clusters⁹⁹ and BAGEL for the identification of biosynthetic clusters for bacteriocins and lantibiotics.¹⁰⁰

4.1. From genotype to chemotype: bioinformatics-guided natural product discovery

How can we determine whether the compound that is derived from a given BGC is novel (*genotype to chemotype*) or, conversely, link a natural product of interest to its gene cluster (*chemotype to genotype*)? The *genotype-to-chemotype* roadmap (Figure 3A)^{101, 102} provides a glimpse of ‘hidden treasures’, *e.g.* aided by the exploitation of pathway-specific regulators^{103, 104} or via heterologous gene expression.^{105, 106} Metagenomic libraries from environmental samples thereby allow scientists access to the ‘dark matter’, namely the vast biosynthetic potential of uncultivable microbes.^{107, 108} However, an intrinsic disadvantage of genome-based

bioprospecting is that current technologies target a single BGC at the time, which is time-consuming, while the discovery efficiency is low as it is very difficult to prioritize ‘promising’ gene clusters. After all, if a gene cluster appears novel, it does not mean that the compound has not already been identified in HT screening regimes in the pre-genomic era. We predict, however, that synthetic biology and heterologous expression of gene clusters will soon become part of a more automated environment, which would make these approaches much more feasible. In addition, approaches directed at the (heterologous) expression of gene clusters depend on the functional analysis of biosynthetic enzymes and/or detection of intact clusters, while biosynthetic reactions may be catalyzed by enzymes that are not encoded by the gene cluster itself and/or by enzymes with low specificity,¹⁰⁹ or are catalyzed non-enzymatically.^{110, 111} Indeed, metabolites are the end products of many cellular processes, so that the bioactive end product is not necessarily specified by individual genes or clusters.¹⁸

In the case of intertwined metabolic and biosynthetic pathways giving rise to a single compound, targeted genomic approaches lose their effectiveness.¹¹¹⁻¹¹³ In such cases, metabolomic approaches are required to bridge the gap between bioinformatics-driven gene cluster analysis and NP discovery.^{114, 115} The fundamental principle is based on the fact that structurally related NPs are typically characterized by similar MS/MS fragmentation patterns. The MS/MS structural relatedness among molecules can be detected in an automated manner, and can subsequently generate a molecular network wherein analogues cluster together. In this way, known structures and “known unknowns” can be identified in extracts, offering an effective dereplication strategy and a way of prioritizing novel compounds.¹¹⁶

4.2. From chemotype to genotype: metabolomics to identify biosynthetic gene clusters

MS/MS analysis allows the detection of “structural unit tags” or building blocks whose biosynthesis needs specific enzymes bearing substrate specificity. Candidate genes can be used as query against the genome databases to locate the target gene cluster for the compound(s) of interest. MS-based metabolomics in conjunction with molecular networking builds up a *chemotype-to-genotype* roadmap (Figure 3B)^{101, 102, 115, 117} for NP discovery, directly linking secondary metabolites to their corresponding gene clusters. Examples for this concept include peptidogenomics,^{101, 118} glycogenomics¹⁰² and natural product proteomining.¹¹⁹ Peptidogenomics is an MS-based genome-mining method that links peptide natural products to their BGCs through iteratively matching tandem MS (MSⁿ) sequence tagging with genomics-derived peptide structures,¹⁰¹ supported by software automation packages such as the recently developed Pep2Path.¹²⁰ For an excellent review of peptidogenomics techniques and their application we refer the reader elsewhere.¹¹⁵ Glycogenomics is a *chemotype-to-genotype* approach for fast characterization of glycosylated natural products (GNPs) and their biosynthetic pathways. Tandem mass spectrometry (MSⁿ) allows rapid identification of the structural information of sugar substituents of GNPs in microbial metabolic extracts.^{121, 122} The biosynthesis of certain sugars need specific enzymes; as an example, genes for forosamine production encode 2,3- and 3,4-dehydratases, 3-ketoreductase, aminotransferase and *N,N*-dimethyltransferase. Applying the bacterial sugar biosynthetic knowledge, MSⁿ-characterized glycosyl groups can be matched to corresponding glycosylation genes. In microbial genomes, the glycosylation-associated biosynthetic genes are usually co-clustered with those for a wide range of aglycones. Programs such as

antiSMASH are specifically designed with the identification of such signature BGCs in mind. The relevant aglycone biosynthetic genes are then identified, which in turn guides further isolation and structural identification of glycoconjugates by applying knowledge of the aglycone-specific BGC. As aglycones could be any type of structural scaffolds, for sugar-modified NPs, glycogenomics has an advantage over peptidogenomics. This technique therefore expands the applicability of MS-guided genome mining for different classes of NPs, such as nonribosomal and ribosomal peptides, polyketides, terpenes, and alkaloids.¹²² Furthermore, since glycosylated products are produced via the decoration of an aglycone, which is often bioactive by itself, we envisage that glycogenomics can be extended to discover even non-glycosylated NPs that serve as intermediates for final GNPs.

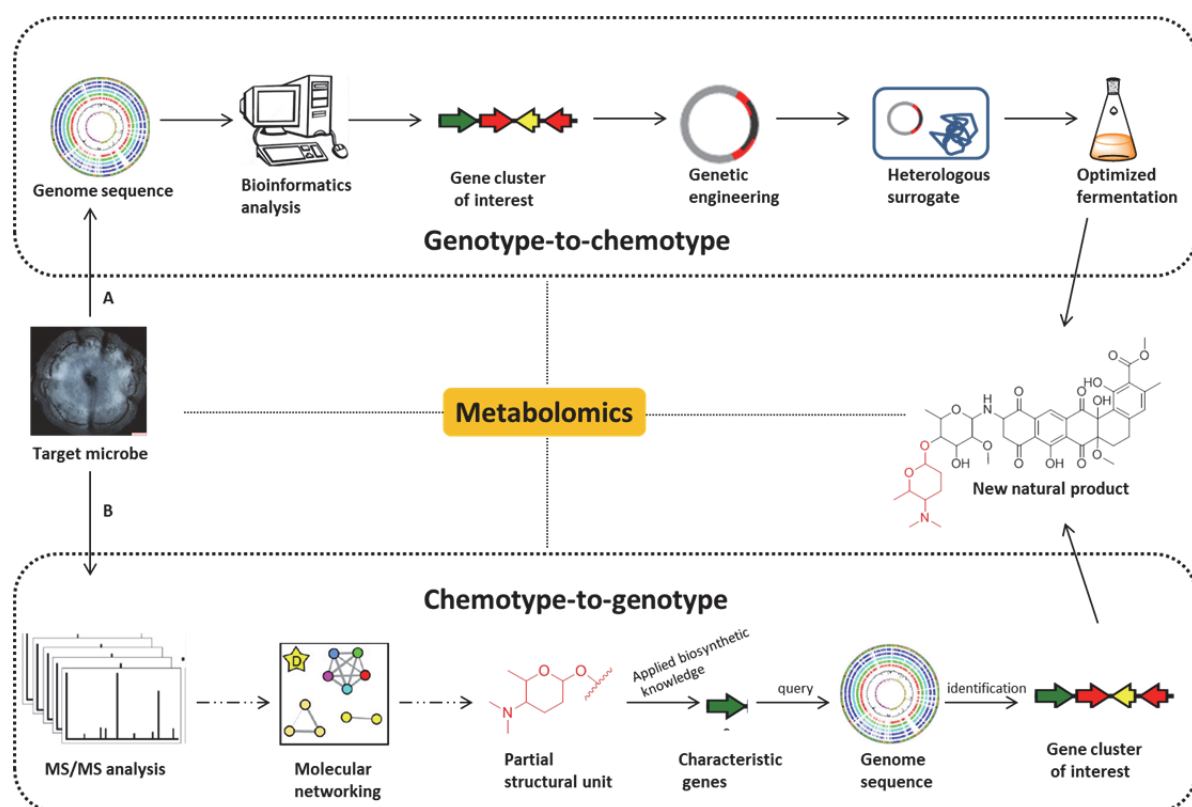


Figure 3. General workflows for genotype-to-chemotype (A) and chemotype-to-genotype (B) genome mining for natural product discovery.

Natural product proteomining is based on the concept that gene cluster expression profiles correlate directly to the level of the corresponding natural product.¹¹⁹ When growth conditions are chosen such that a bioactivity of interest is differentially produced, *e.g.* high under some conditions and intermediate or not at all in others, statistical correlation between bioactivity (metabolomics) and gene expression profiles (proteomics, RNAseq) will allow prioritizing those gene clusters with the optimal match. Since the metabolomics will also provide clues on the type of molecule, conclusive evidence of the BGC responsible for the product of interest can thus be obtained, depending on the number of candidate BGCs present in the genome and the number of samples analyzed.¹¹⁹ The concept of metabolomics aiding genomics in finding new NPs is also exemplified by the case of myxoprincomide from the *Myxococcus xanthus*.¹²³ Initially, 13 BGCs in *M. xanthus* DK1622 were found for which the

metabolites remained elusive. After targeted inactivation of these cryptic gene clusters, MS-based metabolomics was applied to pairwise compare extracts of mutants and the wild-type strain, highlighting the subtle metabolic differences by statistical filtering. Thus, a metabolite with molecular formula $C_{45}H_{76}N_{10}O_{16}$ was assigned to a NRPS/PKS hybrid gene cluster and identified as a molecule with the unprecedented structural skeleton of myxoprincomide. Gene knockout/untargeted metabolomics strategy could also facilitate discovery of intermediates and/or new side products of biosynthesis pathway, as exemplified by characterization of 2-alkyl-4,5-dihydrothiazole-4-carboxylates from *Pseudomonas aeruginosa*.^{124,125} Comparison of the extracellular metabolome of *P. aeruginosa pch* mutants with that of the wild-type strain identified 198 secondary metabolites regulated by these pathway-specific *pch* genes. Therefore, perturbation of the expression of global or pathway-specific regulators followed by metabolomics is a promising strategy to rapidly connect molecules to genes or gene clusters.

5. EXPERIMENTAL APPROACHES BASED ON NMR METABOLOMICS

As explained above, metabolomics approaches allow elucidating the nature of poorly bioactive molecules in complex mixtures, and for this achieving differential production of the bioactivity of interest is a key step. Such fluctuation may be achieved by among others different culturing conditions, inducing drug-resistance or by co-culturing.^{19, 20} NMR-based metabolomics is very common for identifying plant NPs, but has yet seen relatively little application in microbial drug discovery. A general scheme is presented in Figure 1. We recently applied NMR-based metabolomics for the mining of novel soil isolates with promising antibiotic-producing potential, of which we provide a few examples here for illustration purposes. To identify a bioactivity produced by the soil isolate *Streptomyces* sp. MBT70, the strain was grown in media supplemented with different additives followed by 1H NMR spectroscopy and statistical comparison of the extracts by partial least square modeling-discriminant analysis (PLS-DA). A main discriminator between bioactive and inactive cultures was a signature H-5 residue belonging to a naphthoquinone. Further HPLC fractionation targeting naphthoquinones then led to the isolation and identification of juglomycin C amide (Figure 4).¹¹⁹ The generation of streptomycin resistant (Str^R) mutants in soil isolate *Streptomyces* sp. MBT28 resulted in the specific activation of a bioactive compound, which was subsequently identified by comparing the active (Str^R -induced) and non-active fractions by 1H NMR and statistical analysis by projection to latent structures (PLS); this allowed correlation of the enhanced bioactivity to a set of distinctive aromatic signals. NM-guided separation by tracking characteristic 1H NMR signals as probes, resulted in the characterization of the isatin-type antibiotic 7-prenylisatin (our unpublished data). As a third example, co-cultivation of the filamentous model microbes *S. coelicolor* and *Aspergillus niger* substantially influenced their NPs profiles. NMR spectroscopy of extracts obtained from the separate strains and the coculture combined with multivariate data analysis revealed diketopiperazine-type compounds like cyclo(Phe-Phe) (Figure 4), which were exclusively produced by *A. niger* when grown in co-culture with *S. coelicolor*.¹²⁶

1H NMR based profiling also finds its application in lead discovery. Quinn and coworkers integrated solid phase extraction (SPE) with NMR fingerprinting. An extract library was prepared featuring lead-like enhanced (LLE) fractions selected on the basis of favorable physicochemical properties¹²⁷, and this LLE library was subsequently profiled by

^1H NMR spectroscopy to present the structural information of the small molecules contained in the library. In this way, the complicated and overlapping ^1H NMR spectra typical of crude extracts were avoided by prefractionation, facilitating the detection of minor compounds. Using the proton signals as tracking signals for NMR-guided isolation, an unprecedented scaffold of iotrochotazine A was identified in LLE-fractionated extracts from sponges¹²⁸ and low-abundant venulosides C and D (Figure 4) in extracts from plants.¹²⁹

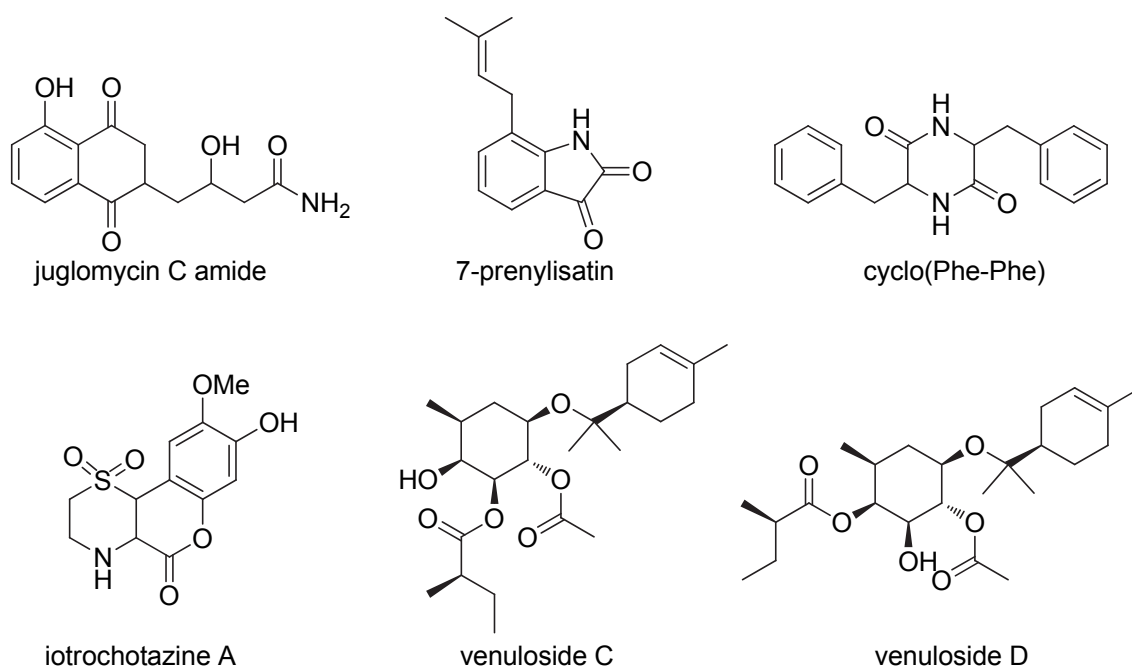


Figure 4. Chemical structures of the compounds mentioned in the final section of the review (“Experimental approaches based on NMR metabolomics”).

An important class of NPs in the fight against antibiotic resistance is that of resistance inhibitors. Penicillins are still the most administered antibiotics,^{130, 131} but resistance caused by β -lactamases was discovered almost as soon as the antibiotic itself¹³² and has become a major problem for their clinical application.¹ The combined use of β -lactam antibiotics together with β -lactamase inhibitors is exemplified by the combination of amoxicillin and clavulanic acid, marketed as Augmentin®.^{133, 134} A target-based in-cell NMR approach was recently proposed for use in high-throughput screening approaches to identify resistance inhibitors with activity against a specific β -lactamase.¹³⁵ The New Delhi metallo- β -lactamase (NDM-1), which opens the cyclic amide ring in β -lactams antibiotics,¹³⁶ was monitored in real time in *E. coli* cells by ^1H NMR spectroscopy.¹³⁷ The approach is summarized in Figure 5. The chemical shifts of four characteristic methyl groups are different between the substrate meropenem and its NDM-1-degraded product, which was used as a reporter of the lytic reaction. Inhibition of NMD-1 enzymatic activity was readily detected since the characteristic methyl groups of meropenem remain unshifted in the ^1H NMR spectrum. NMR-guided fractionation was then applied to trace the responsible antibiotic adjuvants in the active fractions.

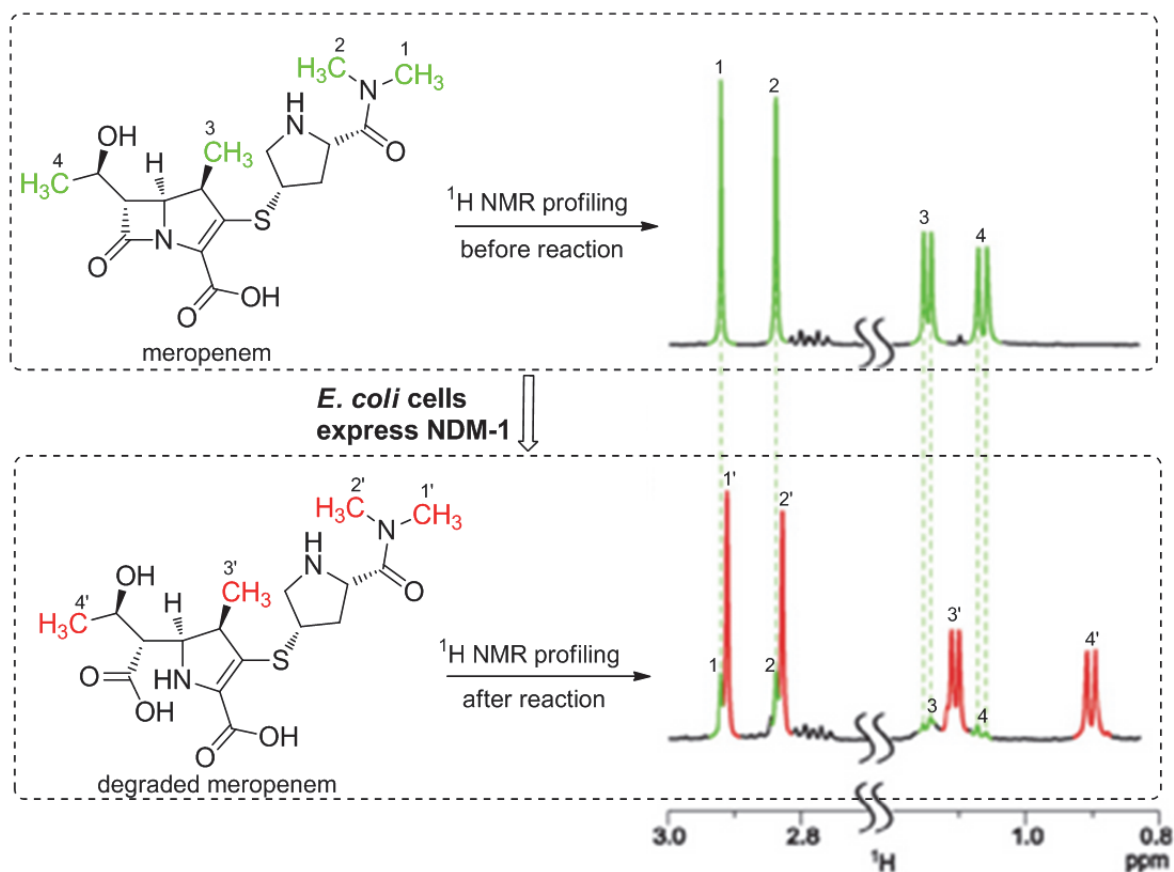


Figure 5. *In vivo* ^1H NMR spectroscopy monitors NDM-1 activity in living *Escherichia coli* cells. Cyclic amide ring in β -lactam antibiotic meropenem is hydrolyzed under catalysis of enzyme NDM-1. The ^1H NMR chemical shifts of four characteristic methyls (green) in the substrate accordingly move upfield in the hydrolytic product (red). This reaction can be monitored in real-time without disturbing the experimental system.

6. CONCLUSION

In a time where dereplication is one of the major challenges in screening efforts, metabolomics offers an effective strategy to prioritize novel molecules produced by microorganisms. Combination of eliciting strategies with comprehensive chemometric comparison of secondary metabolomes and genomics is an excellent alternative over traditional HT approaches. Metabolomics thereby complements (meta-)genomics and synthetic biology approaches in the mining of producing organisms, in particular for poorly expressed (cryptic) molecules. In addition, connecting chemotype to genotype and *vice versa* is important to facilitate molecular approaches and to accelerate unraveling of the biosynthetic pathway, and thus enable metabolic engineering and upscaling. Also taking the huge promise of the unculturable microorganisms into account, it is clear that natural product discovery is far from an echo from the past. Instead, as noted by many scientists worldwide, we may well be witnessing a new era of drug discovery, where many different disciplines such as genomics, metabolomics, bioinformatics, synthetic biology, ecology, chemical biology and industrial screening come together to meet the drug-discovery challenges and deliver new solutions to treat infectious diseases associated with drug resistance.

REFERENCES

- (1) WHO, *Antimicrobial Resistance: Global Report on Surveillance*. Geneva, Switzerland, 2014.
- (2) Baltz, R. H. *Curr Opin Pharmacol* **2008**, *8*, 557-563.
- (3) Cooper, M. A.; Shlaes, D. *Nature* **2011**, *472*, 32.
- (4) Genilloud, O.; Gonzalez, I.; Salazar, O.; Martin, J.; Tormo, J. R.; Vicente, F. *J Ind Microbiol Biotechnol* **2011**, *38*, 375-389.
- (5) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat Rev Drug Discov* **2007**, *6*, 29-40.
- (6) Bérdy, J. *J Antibiot (Tokyo)* **2005**, *58*, 1-26.
- (7) Newman, D. J.; Cragg, G. M. *J Nat Prod* **2007**, *70*, 461-477.
- (8) Kaeberlein, T.; Lewis, K.; Epstein, S. S. *Science* **2002**, *296*, 1127-1129.
- (9) Ling, L. L.; Schneider, T.; Peoples, A. J.; Spoering, A. L.; Engels, I.; Conlon, B. P.; Mueller, A.; Schaberle, T. F.; Hughes, D. E.; Epstein, S.; Jones, M.; Lazarides, L.; Steadman, V. A.; Cohen, D. R.; Felix, C. R.; Fetterman, K. A.; Millett, W. P.; Nitti, A. G.; Zullo, A. M.; Chen, C.; Lewis, K. *Nature* **2015**, *517*, 455-459.
- (10) Wilson, M. C.; Mori, T.; Ruckert, C.; Uriá, A. R.; Helf, M. J.; Takada, K.; Gernert, C.; Steffens, U. A.; Heycke, N.; Schmitt, S.; Rinke, C.; Helfrich, E. J.; Brachmann, A. O.; Gurgui, C.; Wakimoto, T.; Kracht, M.; Crusemann, M.; Hentschel, U.; Abe, I.; Matsunaga, S.; Kalinowski, J.; Takeyama, H.; Piel, J. *Nature* **2014**, *506*, 58-62.
- (11) Bergmann, S.; Schumann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. *Nat Chem Biol* **2007**, *3*, 213-217.
- (12) Zhu, H.; Sandiford, S. K.; van Wezel, G. P. *J Ind Microbiol Biotechnol* **2014**, *41*, 371-386.
- (13) Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A. *Nature* **2002**, *417*, 141-147.
- (14) Cruz-Morales, P.; Vijgenboom, E.; Iruegas-Bocardo, F.; Girard, G.; Yanez-Guerra, L. A.; Ramos-Aboites, H. E.; Pernodet, J. L.; Anne, J.; van Wezel, G. P.; Barona-Gomez, F. *Genome Biol Evol* **2013**, *5*, 1165-1175.
- (15) Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S. *Nat Biotechnol* **2003**, *21*, 526-531.
- (16) van Keulen, G.; Dyson, P. J. *Adv Appl Microbiol* **2014**, *89*, 217-266.
- (17) Pauli, G. F.; Chen, S. N.; Friesen, J. B.; McAlpine, J. B.; Jaki, B. U. *J Nat Prod* **2012**, *75*, 1243-1255.
- (18) Perić-Concha, N.; Long, P. F. *Drug Disc Today* **2003**, *8*, 1078-1084.
- (19) Gaudêncio, S. P.; Pereira, F. *Nat Prod Rep* **2015**, *32*, 779-810.
- (20) Wu, C.; Kim, H. K.; van Wezel, G. P.; Choi, Y. H. *Drug Disc Today: Technol* **2015**, *13*,

- 11–17.
- (21) Weber, T.; Charusanti, P.; Musiol-Kroll, E. M.; Jiang, X.; Tong, Y.; Kim, H. U.; Lee, S. Y. *Trends Biotechnol* **2015**, *33*, 15-26.
 - (22) Bode, H. B. r.; Bethe, B.; H[^]fs, R.; Zeeck, A. *ChemBioChem* **2002**, *3*, 619-627.
 - (23) Zhu, H.; Swierstra, J.; Wu, C.; Girard, G.; Choi, Y. H.; van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. *Microbiology* **2014**, *160*, 1714-1725.
 - (24) Ohnishi, Y.; Ishikawa, J.; Hara, H.; Suzuki, H.; Ikenoya, M.; Ikeda, H.; Yamashita, A.; Hattori, M.; Horinouchi, S. *J Bacteriol* **2008**, *190*, 4050-4060.
 - (25) Fu, P.; Jamison, M.; La, S.; Macmillan, J. B. *Org Lett* **2014**, *16*, 5656-5659.
 - (26) Hosaka, T.; Ohnishi-Kameyama, M.; Muramatsu, H.; Murakami, K.; Tsurumi, Y.; Kodani, S.; Yoshida, M.; Fujie, A.; Ochi, K. *Nat Biotechnol* **2009**, *27*, 462-464.
 - (27) Ochi, K.; Tanaka, Y.; Tojo, S. *J Ind Microbiol Biotechnol* **2014**, *41*, 403-414.
 - (28) Rigali, S.; Titgemeyer, F.; Barends, S.; Mulder, S.; Thomae, A. W.; Hopwood, D. A.; van Wezel, G. P. *EMBO Rep* **2008**, *9*, 670-675.
 - (29) Willey, J. M.; Gaskell, A. A. *Chem Rev* **2011**, *111*, 174-187.
 - (30) Hsiao, N. H.; Gottelt, M.; Takano, E. *Methods Enzymol* **2009**, *458*, 143-157.
 - (31) Albright, J. C.; Henke, M. T.; Soukup, A. A.; McClure, R. A.; Thomson, R. J.; Keller, N. P.; Kelleher, N. L. *ACS Chem Biol* **2015**, In press.
 - (32) Bertrand, S.; Bohni, N.; Schnee, S.; Schumpp, O.; Gindro, K.; Wolfender, J.-L. *Biotechnol Adv* **2014**, *32*, 1180-1204.
 - (33) Hoshino, S.; Wakimoto, T.; Onaka, H.; Abe, I. *Org Lett* **2015**, *17*, 1501-1504.
 - (34) Sugiyama, R.; Nishimura, S.; Ozaki, T.; Asamizu, S.; Onaka, H.; Kakeya, H. *Org Lett* **2015**, *17*, 1918-1921.
 - (35) Thaker, M. N.; Wang, W.; Spanogiannopoulos, P.; Waglechner, N.; King, A. M.; Medina, R.; Wright, G. D. *Nat Biotechnol* **2013**, *31*, 922-927.
 - (36) Bibb, M. J. *Curr Opin Microbiol* **2005**, *8*, 208-215.
 - (37) Liu, G.; Chater, K. F.; Chandra, G.; Niu, G.; Tan, H. *Microbiol Mol Biol Rev* **2013**, *77*, 112-143.
 - (38) Sanchez, S.; Chavez, A.; Forero, A.; Garcia-Huante, Y.; Romero, A.; Sanchez, M.; Rocha, D.; Sanchez, B.; Avalos, M.; Guzman-Trampe, S.; Rodriguez-Sanoja, R.; Langley, E.; Ruiz, B. *J Antibiot (Tokyo)* **2010**, *63*, 442-459.
 - (39) van Wezel, G. P.; McDowall, K. J. *Nat Prod Rep* **2011**, *28*, 1311-1333.
 - (40) Wietzorrek, A.; Bibb, M.; Chakraborty, R. *Mol Microbiol* **1997**, *25*, 1181-1184.
 - (41) Floriano, B.; Bibb, M. *Mol Microbiol* **1996**, *21*, 385-396.
 - (42) Arias, P.; Fernandez-Moreno, M. A.; Malpartida, F. *J Bacteriol* **1999**, *181*, 6958-6968.
 - (43) Takano, E.; Gramajo, H. C.; Strauch, E.; Andres, N.; White, J.; Bibb, M. J. *Mol Microbiol* **1992**, *6*, 2797-2804.
 - (44) van Wezel, G. P.; White, J.; Hoogvliet, G.; Bibb, M. J. *J Mol Microbiol Biotechnol* **2000**, *2*, 551-556.
 - (45) Bibb, M.; Hesketh, A. *Methods Enzymol* **2009**, *458*, 93-116.
 - (46) van Wezel, G. P.; McKenzie, N. L.; Nodwell, J. R. *Methods Enzymol* **2009**, *458*, 117-141.
 - (47) Colson, S.; Stephan, J.; Hertrich, T.; Saito, A.; van Wezel, G. P.; Titgemeyer, F.; Rigali, S. *J Mol Microbiol Biotechnol* **2007**, *12*, 60-66.

-
- (48) Colson, S.; van Wezel, G. P.; Craig, M.; Noens, E. E.; Nothaft, H.; Mommaas, A. M.; Titgemeyer, F.; Joris, B.; Rigali, S. *Microbiology* **2008**, *154*, 373-382.
- (49) Craig, M.; Lambert, S.; Jourdan, S.; Tenconi, E.; Colson, S.; Maciejewska, M.; Ongena, M.; Martin, J. F.; van Wezel, G.; Rigali, S. *Environ Microbiol Rep* **2012**, *4*, 512-521.
- (50) Rigali, S.; Nothaft, H.; Noens, E. E.; Schlicht, M.; Colson, S.; Muller, M.; Joris, B.; Koerten, H. K.; Hopwood, D. A.; Titgemeyer, F.; van Wezel, G. P. *Mol Microbiol* **2006**, *61*, 1237-1251.
- (51) Swiatek, M. A.; Tenconi, E.; Rigali, S.; van Wezel, G. P. *J Bacteriol* **2012**, *194*, 1136-1144.
- (52) Nazari, B.; Kobayashi, M.; Saito, A.; Hassaninasab, A.; Miyashita, K.; Fujii, T. *Appl Environ Microbiol* **2012**, *79*, 707-713.
- (53) Swiatek-Polatynska, M. A.; Bucca, G.; Laing, E.; Gubbens, J.; Titgemeyer, F.; Smith, C. P.; Rigali, S.; van Wezel, G. P. *PLoS One* **2015**, *10*, e0122479.
- (54) Uguru, G. C.; Stephens, K. E.; Stead, J. A.; Towle, J. E.; Baumberg, S.; McDowall, K. J. *Mol Microbiol* **2005**, *58*, 131-150.
- (55) Hong, B.; Phornphisutthimas, S.; Tilley, E.; Baumberg, S.; McDowall, K. J. *Biotechnol Lett* **2007**, *29*, 57-64.
- (56) Mao, X. M.; Luo, S.; Zhou, R. C.; Wang, F.; Yu, P.; Sun, N.; Chen, X. X.; Tang, Y.; Li, Y. Q. *J Biol Chem* **2015**, *290*, 7992-8001.
- (57) Santos-Beneit, F.; Rodriguez-Garcia, A.; Sola-Landa, A.; Martin, J. F. *Mol Microbiol* **2009**, *72*, 53-68.
- (58) Kim, S. H.; Traag, B. A.; Hasan, A. H.; McDowall, K. J.; Kim, B. G.; van Wezel, G. P. *Antonie Van Leeuwenhoek* **2015**, *108*, 201-213.
- (59) Nothaft, H.; Rigali, S.; Boomsma, B.; Swiatek, M.; McDowall, K. J.; van Wezel, G. P.; Titgemeyer, F. *Mol Microbiol* **2010**, *75*, 1133-1144.
- (60) Mahr, K.; van Wezel, G. P.; Svensson, C.; Krenkel, U.; Bibb, M. J.; Titgemeyer, F. *Antonie Van Leeuwenhoek* **2000**, *78*, 253-261.
- (61) Titgemeyer, F.; Reizer, J.; Reizer, A.; Saier, M. H., Jr. *Microbiology* **1994**, *140*, 2349-2354.
- (62) Martinez, A.; Kolvek, S. J.; Hopke, J.; Yip, C. L.; Osburne, M. S. *Appl Environ Microbiol* **2005**, *71*, 1638-1641.
- (63) Swiatek, M. A.; Gubbens, J.; Bucca, G.; Song, E.; Yang, Y. H.; Laing, E.; Kim, B. G.; Smith, C. P.; van Wezel, G. P. *J Bacteriol* **2013**, *195*, 1236-1248.
- (64) Tomono, A.; Tsai, Y.; Yamazaki, H.; Ohnishi, Y.; Horinouchi, S. *J Bacteriol* **2005**, *187*, 5595-5604.
- (65) Higo, A.; Hara, H.; Horinouchi, S.; Ohnishi, Y. *DNA Res* **2012**, *19*, 259-273.
- (66) Lee, P. C.; Umeyama, T.; Horinouchi, S. *Mol Microbiol* **2002**, *43*, 1413-1430.
- (67) Sola-Landa, A.; Moura, R. S.; Martin, J. F. *Proc Natl Acad Sci U S A* **2003**, *100*, 6133-6138.
- (68) Patti, G. J.; Yanes, O.; Siuzdak, G. *Nat Rev Mol Cell Biol* **2012**, *13*, 263-269.
- (69) Kuehnbaum, N. L.; Britz-Mckibbin, P. *Chem Rev* **2013**, *113*, 2437-2468.
- (70) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Trends Biotechnol* **2011**, *29*, 267-275.
- (71) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Nat Protoc* **2010**, *5*, 536-549.
- (72) Mahrous, E. A.; Farag, M. A. *J Adv Res* **2015**, *6*, 3-15.

- (73) Novoa-Carballal, R.; Fernandez-Megia, E.; Jimenez, C.; Riguera, R. *Nat Prod Rep* **2011**, *28*, 78-98.
- (74) Sidebottom, A. M.; Johnson, A. R.; Karty, J. A.; Trader, D. J.; Carlson, E. E. *ACS Chem Biol* **2013**, *8*, 2009-2016.
- (75) Breton, R. C.; Reynolds, W. F. *Nat Prod Rep* **2013**, *30*, 501-524.
- (76) Nicolaou, K. C.; Snyder, S. A. *Angew Chem Intl Ed* **2005**, *44*, 1012-1044.
- (77) Senior, M. M.; Williamson, R. T.; Martin, G. E. *J Nat Prod* **2013**, *76*, 2088-2093.
- (78) Kövér, K. E.; Hruby, V. J.; Uhrin, D. *J Magn Res* **1997**, *129*, 125-129.
- (79) Martin, G. E., Using 1,1- and 1,*n*-ADEQUATE 2D NMR data in structure elucidation protocols. In *Annual Rep NMR Spectr*, Elsevier: London, 2011; pp 215-291.
- (80) Meyer, S. W.; Köck, M. *J Nat Prod* **2008**, *71*, 1524-1529.
- (81) Cheatham, S. F.; Kline, M.; Sasaki, R. R.; Blinov, K. A.; Elyashberg, M. E.; Molodtsov, S. G. *Magn Res Chem* **2010**, *48*, 571-574.
- (82) Buevich, A. V.; Williamson, R. T.; Martin, G. E. *J Nat Prod* **2014**, *77*, 1942-1947.
- (83) Martin, G. E.; Williamson, R. T.; Dormer, P. G.; Bermel, W. *Magn Res Chem* **2012**, *50*, 563-568.
- (84) Wolfender, J.-L.; Ndjoko, K.; Hostettmann, K. *Phytochem Anal* **2001**, *12*, 2-22.
- (85) Månsson, M.; Phipps, R. K.; Gram, L.; Munro, M. H. G.; Larsen, T. O.; Nielsen, K. F. *J Nat Prod* **2010**, *73*, 1126-1132.
- (86) Seger, C.; Sturm, S.; Stuppner, H. *Nat Prod Rep* **2013**, *30*, 970-987.
- (87) Sprogøe, K.; Stærk, D.; Jäger, A. K.; Adsersen, A.; Hansen, S. H.; Witt, M.; Landbo, A. K.; Meyer, A. S.; Jaroszewski, J. W. *J Nat Prod* **2007**, *70*, 1472-1477.
- (88) Lambert, M.; Stærk, D.; Hansen, S. H.; Sairafianpour, M.; Jaroszewski, J. W. *J Nat Prod* **2005**, *68*, 1500-1509.
- (89) Johansen, K. T.; Wubshet, S. G.; Nyberg, N. T.; Jaroszewski, J. W. *J Nat Prod* **2011**, *74*, 2454-2461.
- (90) Liu, H.; Zheng, A.; Liu, H.; Yu, H.; Wu, X.; Xiao, C.; Dai, H.; Hao, F.; Zhang, L.; Wang, Y.; Tang, H. *J Agr Food Chem* **2012**, *60*, 129-135.
- (91) Exarchou, V.; Godejohann, M.; Beek, T. A. V.; Gerothanassis, I. P.; Vervoort, J. *Anal Chem* **2003**, *75*, 6288-6294.
- (92) Corre, C.; Challis, G. L. *Nat Prod Rep* **2009**, *26*, 977-986.
- (93) Cimermancic, P.; Medema, M. H.; Claesen, J.; Kurita, K.; Wieland Brown, L. C.; Mavrommatis, K.; Pati, A.; Godfrey, P. A.; Koehrsen, M.; Clardy, J.; Birren, B. W.; Takano, E.; Sali, A.; Linington, R. G.; Fischbach, M. A. *Cell* **2014**, *158*, 412-421.
- (94) Bachmann, B. O.; Ravel, J. *Methods Enzymol* **2009**, *458*, 181-217.
- (95) Lautru, S.; Challis, G. L. *Microbiology* **2004**, *150*, 1629-1636.
- (96) Medema, M.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. *Nucleic Acids Res* **2011**, *39*, W339-W346.
- (97) Fedorova, N. D.; Moktali, V.; Medema, M. H. *Methods Mol Biol* **2012**, *944*, 23-45.
- (98) Anand, S.; Prasad, M. V.; Yadav, G.; Kumar, N.; Shehara, J.; Ansari, M. Z.; Mohanty, D. *Nucleic Acids Res* **2010**, *38*, W487-496.
- (99) Rottig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. *Nucleic Acids Res* **2011**, *39*, W362-367.
- (100) de Jong, A.; van Heel, A. J.; Kok, J.; Kuipers, O. P. *Nucleic Acids Res* **2010**, *38*,

- W647-651.
- (101) Kersten, R. D.; Yang, Y. L.; Xu, Y.; Cimermancic, P.; Nam, S. J.; Fenical, W.; Fischbach, M. A.; Moore, B. S.; Dorrestein, P. C. *Nat Chem Biol* **2011**, *7*, 794-802.
 - (102) Kersten, R. D.; Ziemert, N.; Gonzalez, D. J.; Duggan, B. M.; Nizet, V.; Dorrestein, P. C.; Moore, B. S. *Proc Natl Acad Sci U S A* **2013**, *110*, 4407-4416.
 - (103) Laureti, L.; Song, L.; Huang, S.; Corre, C.; Leblond, P.; Challis, G. L.; Aigle, B. *Proc Natl Acad Sci U S A* **2011**, *108*, 6258-6263.
 - (104) Gottelt, M.; Kol, S.; Gomez-Escribano, J. P.; Bibb, M.; Takano, E. *Microbiology* **2010**, *156*, 2343-2353.
 - (105) Baltz, R. H. *J Ind Microbiol Biotechnol* **2010**, *37*, 759-772.
 - (106) Ongley, S. E.; Bian, X.; Neilan, B. A.; Müller, R. *Nat Prod Rep* **2013**, *30*, 1121-1138.
 - (107) Brady, S. F.; Simmons, L.; Kim, J. H.; Schmidt, E. W. *Nat Prod Rep* **2009**, *26*, 1488-1503.
 - (108) Kang, H. S.; Brady, S. F. *ACS Chem Biol* **2014**, *9*, 1267-1272.
 - (109) Schwab, W. *Phytochemistry* **2003**, *62*, 837-849.
 - (110) Mayer, M.; Thiericke, R. *J Antibiot (Tokyo)* **1993**, *46*, 1372-1380.
 - (111) Wu, C.; van Wezel, G. P.; Choi, Y. H. *J Antibiot (Tokyo)* **2015**, In press.
 - (112) Vingadassalon, A.; Lorieux, F.; Juguet, M.; Gof, G. r. L.; Gerbaud, C.; Pernodet, J.-L.; Lautru, S. *ACS Chem Biol* **2015**, *10*, 601-610.
 - (113) Tsukioka, Y.; Yamashita, Y.; Nakano, Y.; Oho, T.; Koga, T. *J Bacteriol* **1997**, *179*, 4411-4414.
 - (114) Doroghazi, J. R.; Albright, J. C.; Goering, A. W.; Ju, K.-S.; Haines, R. R.; Tchalukov, K. A.; Labeda, D. P.; Kelleher, N. L.; Metcalf, W. W. *Nat Chem Biol* **2014**, *10*, 963-970.
 - (115) Krug, D.; Müller, R. *Nat Prod Rep* **2014**, *31*, 768-783.
 - (116) Yang, J. Y.; Sanchez, L. M.; Rath, C. M.; Liu, X.; Boudreau, P. D.; Bruns, N.; Glukhov, E.; Wodtke, A.; De Felicio, R.; Fenner, A.; Wong, W. R.; Linington, R. G.; Zhang, L.; Debonisi, H. M.; Gerwick, W. H.; Dorrestein, P. C. *J Nat Prod* **2013**, *76*, 1686-1699.
 - (117) Helfrich, E. J. N.; Reiter, S.; Piel, J. *Curr Opin Biotechnol* **2014**, *29*, 107-115.
 - (118) Nguyen, D. D.; Wu, C. H.; Moree, W. J.; Lamsa, A.; Medema, M. H.; Zhao, X.; Gavilan, R. G.; Aparicio, M.; Atencio, L.; Jackson, C.; Ballesteros, J.; Sanchez, J.; Watrous, J. D.; Phelan, V. V.; van de Wiel, C.; Kersten, R. D.; Mehnaz, S.; De Mot, R.; Shank, E. A.; Charusanti, P.; Nagarajan, H.; Duggan, B. M.; Moore, B. S.; Bandeira, N.; Palsson, B. O.; Pogliano, K.; Gutierrez, M.; Dorrestein, P. C. *Proc Natl Acad Sci U S A* **2013**, *110*, E2611-2620.
 - (119) Gubbens, J.; Zhu, H.; Girard, G.; Song, L.; Florea, B. I.; Aston, P.; Ichinose, K.; Filippov, D. V.; Choi, Y. H.; Overkleeft, H. S.; Challis, G. L.; van Wezel, G. P. *Chem Biol* **2014**, *21*, 707-718.
 - (120) Medema, M. H.; Paalvast, Y.; Nguyen, D. D.; Melnik, A.; Dorrestein, P. C.; Takano, E.; Breitling, R. *PLoS Comput Biol* **2014**, *10*, e1003822.
 - (121) Herget, S.; Toukach, P. V.; Ranzinger, R.; Hull, W. E.; Knirel, Y. A.; von der Lieth, C.-W. *BMC Struct Biol* **2008**, *8*, 35.
 - (122) Elshahawi, S. I.; Shaaban, K. A.; Kharel, M. K.; Thorson, J. S. *Chem Soc Rev* **2015**, In press.

-
- (123) Cortina, N. S.; Krug, D.; Plaza, A.; Revermann, O.; Müller, R. *Angew Chem Intl Ed* **2012**, *51*, 811-816.
- (124) Kersten, R. D.; Dorrestein, P. C. *ACS Chem Biol* **2009**, *4*, 599-601.
- (125) Vinayavekhin, N.; Saghatelian, A. *ACS Chem Biol* **2009**, *4*, 617-623.
- (126) Wu, C.; Zacchetti, B.; Ram, A. F. J.; van Wezel, G. P.; Claessen, D.; Choi, Y. H. *Sci Rep* **2015**, *5*, 10868.
- (127) Camp, D.; Davis, R. A.; Campitelli, M.; Ebdon, J.; Quinn, R. J. *J Nat Prod* **2012**, *75*, 72-81.
- (128) Grkovic, T.; Pouwer, R. H.; Vial, M. L.; Gambini, L.; Noël, A.; Hooper, J. N. A.; Wood, S. A.; Mellick, G. D.; Quinn, R. J. *Angew Chem Intl Ed* **2014**, *53*, 6070-6074.
- (129) Grkovic, T.; Pouwer, R. H.; Wang, Q.; Guymer, G. P.; Holst, J.; Quinn, R. J. *J Nat Prod* **2015**, In press.
- (130) Demain, A. L. *Proc Natl Sci Counc Repub China B* **1991**, *15*, 251-265.
- (131) Southgate, R.; Elson, S. *Fortschr Chem Org Naturst* **1985**, *47*, 1-106.
- (132) Abraham, E. P.; Chain, E. *Nature* **1940**, *46*, 837.
- (133) Drawz, S. M.; Bonomo, R. A. *Clin Microbiol Rev* **2010**, *23*, 160-201.
- (134) Fisher, J. F.; Meroueh, S. O.; Mobashery, S. *Chem Rev* **2005**, *105*, 395-424.
- (135) Ma, J.; Cao, Q.; Mcleod, S. M.; Ferguson, K.; Gao, N.; Breeze, A. L. *Angew Chem Intl Ed* **2015**, *54*, 4764-4767.
- (136) Yong, D.; Toleman, M. A.; Giske, C. G.; Cho, H. S.; Sundman, K.; Lee, K.; Walsh, T. R. *Antimicrob Agents Chemother* **2009**, *53*, 5046-5054.
- (137) Ma, J.; McLeod, S.; MacCormack, K.; Sriram, S.; Gao, N.; Breeze, A. L.; Hu, J. *Angew Chem Intl Ed* **2014**, *53*, 2130-2133.

Chapter 4

Metabolomics-driven discovery of a prenylated isatin antibiotic produced by *Streptomyces* species MBT28

Changsheng Wu ^{1,2}, Chao Du ¹, Jacob Gubbens ³, Young Hae Choi ², and Gilles P. van Wezel ¹

¹ *Molecular Biotechnology, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands*

² *Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands*

³ *Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands*

J. Nat. Prod. **2015**, 78, 2355–2363

ABSTRACT

Actinomycetes are a major source of antimicrobials, anti-cancer compounds and other medically natural products, and their genomes harbor extensive biosynthetic potential. Major challenges in the screening of these microorganisms are to activate the expression of cryptic biosynthetic gene clusters and the development of technologies for efficient dereplication of known molecules. Here we report the identification of a previously unidentified isatin-type antibiotic produced by *Streptomyces* sp. MBT28, following a strategy based on NMR-based metabolomics combined with the introduction of streptomycin resistance in the producer strain. NMR-guided isolation by tracking the target proton signal resulted in the characterization of 7-prenylisatin (**1**) with antimicrobial activity against *Bacillus subtilis*. The metabolite-guided genome mining of *Streptomyces* sp. MBT28 combined with proteomics identified a gene cluster with an indole prenyltransferase that catalyzes the conversion of tryptophan into 7-prenylisatin. This study underlines the applicability of NMR-based metabolomics in facilitating the discovery of novel antibiotics.

1. INTRODUCTION

The discovery and development of antibiotics to fight bacterial diseases is one of the greatest triumphs of modern medicine. However, the increase of antimicrobial resistance means that bacterial infections which once seemed conquered, now form a major threat to human health.¹ The high frequency of re-discovery of known compounds by high-throughput (HT) screening regimes necessitates new approaches to discover the antimicrobial lead compounds.²⁻⁵ As producers of approximately two thirds of all known antibiotics and many other medically relevant natural products, actinomycetes are a major source of clinical drugs.^{6, 7} It is probable that the antimicrobial compounds so far been identified represent only a fraction of the total products.⁸ In support of this, sequencing the genomes of actinomycetes revealed the potential of even the best-studied model organisms as producers of natural products had been grossly underestimated.⁹⁻¹¹ The biosynthetic gene clusters are under extensive genetic control.^{12, 13} Many of these gene clusters are poorly expressed under routine laboratory conditions and are therefore generally referred as cryptic or silent gene clusters (reviewed in ¹⁴⁻¹⁷).

A drug-discovery pipeline that is rapidly gaining momentum combines genome sequence information with different culturing conditions to allow fluctuation of the production of (cryptic) antibiotics, followed by the metabolic profiling-based identification of the bioactivity of interest. A major challenge thereby lies in finding the appropriate chemical triggers or ecological cues to elicit the production of cryptic antibiotics (recently reviewed in ^{18, 19}). Strategies that have been employed include heterologous expression,²⁰ the use of chemical elicitors,^{21, 22} and inducing antibiotic resistance.^{23, 24} Selection of streptomycin- or rifampicin-resistant mutants, typically caused by mutations in *rpsL* (for ribosomal protein S12) or *rpoB* (for the β -subunit of RNA polymerase), respectively, enhances the production of antibiotics^{23, 25, 26} and may also lead to the production of novel antibiotics.²⁷

Once differential expression of a bioactivity is achieved in a producing organism, *e.g.*, by varying culturing conditions or by making mutants, the next challenge is to rapidly establish whether the natural product is sufficiently novel to warrant full structure elucidation. Robust chemical analysis is required to monitor the differences in metabolic complexity and antibiotic production when using different strains or culturing conditions. Metabolomics approaches globally identify the low molecular weight metabolites in biological samples. Chemical profiling techniques using LC-MS or NMR followed by multivariate data analysis allows scientists to compare and detect molecules that are differentially synthesized between diverse biological samples, thereby effectively narrowing down the search for the sought-after biomarker and avoiding chemical redundancy in an early stage.^{28, 29}

In this study, we describe the activation of antimicrobial activity in *Streptomyces* sp. MBT28, by selecting streptomycin-resistant mutants. Multivariate data analysis correlated the bioactivity to the corresponding NMR signals, facilitating the characterization of a previously undescribed antibiotic, 7-prenylisatin (**1**). The likely genes encoding the key biosynthetic enzymes for the synthesis of 7-prenylisatin were uncovered using natural product proteomining,³⁰ a proteomics-based genome mining approach.

2. RESULTS AND DISCUSSION

2.1. Induced antibiotic production by selection for streptomycin resistance

Previous screening of over 800 actinomycetes from our in-house actinomycete collection

obtained from remote mountain soils, showed that a subset of 96 isolates was of particular relevance for drug discovery as they produced antibiotics against many ESKAPE pathogens under specific growth conditions.²² One of these isolates is *Streptomyces* sp. MBT28, which has a broad spectrum activity against multidrug-resistant (MDR) bacterial pathogens, including *Enterococcus faecium*, *Klebsiella pneumonia* and *Acinetobacter baumannii*. The strain produced the known antibiotic borrelidin as shown by bioactivity-guided separation,²² but other bioactive molecules remained uncharacterized due to their low abundance.

To enhance the expression of antimicrobials in *Streptomyces* sp. MBT28, we made use of the well-established fact that the introduction of streptomycin resistance induces antibiotic production by actinomycetes.³¹ *Streptomyces* sp. MBT28 spores were plated onto SFM agar containing 10 µg/ml streptomycin, which resulted in the formation of streptomycin resistant (Str^R) mutants after around nine days of growth. A surprisingly high number of these spontaneous Str^R mutants showed altered colony size and/or pigmentation, with many white colonies appearing. White colonies also arose in the absence of streptomycin, but at a much lower frequency (Figure S1). As a control, we also selected Str^R colonies of the phylogenetically related *Streptomyces lividans* 1326. The resulting *S. lividans* Str^R mutants did not show the same high frequency of white mutants (Figure S1), suggesting that MBT28 is genetically instable.

Around 200 randomly chosen Str^R mutants were transferred to MM agar plates, allowed to grow for six days at 30 °C and then overlaid with *Bacillus subtilis* to assess their antimicrobial activity. 14% of the Str^R mutants had enhanced antimicrobial activity as compared to the parental strain, while more than half of the mutants showed reduced bioactivity as compared to the parental strain. Furthermore, no clear correlation between the phenotype (*i.e.* pigmentation) and antibiotic production was observed.

One white mutant derivative (MBT28-91) with particularly strong bioactivity was chosen for further investigation, and a second mutant (MBT28-30) with less enhanced bioactivity was selected for comparison. On SFM agar plates, *Streptomyces* sp. MBT28 and Str^R mutant MBT28-30 produced grey-pigmented spores after two days of incubation, while MBT28-91 remained white even after nine days of incubation at 30 °C (Figure 1A). After prolonged incubation (two weeks), MBT28-91 eventually turned light grey. To compare the antimicrobial activity of the strains, antibiotic production was examined at 24 h intervals during growth on MM agar plates (Figure 1B). MBT28-91 produced a much larger zone of inhibition against *B. subtilis* than the parent, indicative of enhanced antibiotic production, while MBT28-30 exhibited slightly higher antibiotic activity than the parent.

The *rpsL* gene, which encodes ribosomal protein S12, is the major hotspot for streptomycin-resistant mutations and therefore also referred to as the *str* gene.³² An alternative site for Str^R mutations, particularly associated with lower streptomycin resistance, is the *rsmG* gene for an rRNA methyltransferase.³³ In an attempt to elucidate the nature of the Str^R mutations, the *rpsL* and *rsmG* genes of the two mutants were amplified by PCR and the DNA products were sequenced. Neither of the mutants had sustained a mutation in *rpsL*. However, the *rsmG* gene of MBT28-30 carried a G to A mutation at nt position 92, which resulted in a Gly→Asp mutation at aa position 31 of RsmG. Surprisingly, Mutant MBT28-91 carried wild-type copies of both *rpsL* and *rsmG*, and the origin of the enhanced Str^R is therefore as yet unclear.

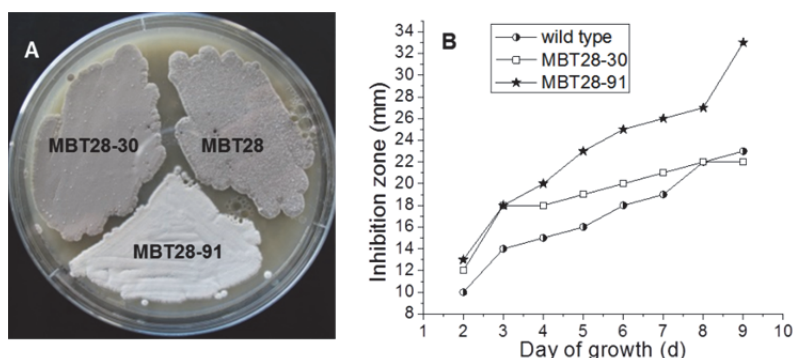


Figure 1. Antibiotic activity of *Streptomyces* sp. MBT28 and its streptomycin-resistant mutants. Morphologies of colonies of *Streptomyces* sp. MBT28 and its Str^R mutants MBT28-30 and MBT28-91 grown on SFM agar for six days at 30 °C (A), and time course of inhibition activity against *Bacillus subtilis* (B). For the activity test, the three strains were grown on MM agar supplemented with 0.5% mannitol and 1% glycerol, and samples taken at 24 h intervals (*X* axis) were overlaid with *B. subtilis* to give inhibition zone (*Y* axis). Note the strong increase in bioactivity of MBT-28-91 as compared to the parental strain MBT28 or to mutant MBT28-30.

2.2. Metabolic profiling and multivariate data analysis

¹H nuclear magnetic resonance (¹H NMR) based metabolic profiling was conducted to analyze and compare the metabolomes of wild-type *Streptomyces* sp. MBT28 and its Str^R-mutant derivatives, so as to identify the compound(s) responsible for the elevated bioactivity. NMR spectroscopy, and in particular ¹H NMR which is commonly used for natural product profiling, is well suited for such comparative metabolomics as it is fast, reproducible, and benefits from relatively straightforward sample preparation.³⁴ Multivariate data analysis (MVDA) was used to reduce the dimensionality of the multivariate dataset and to discriminate the samples. For this, the strains were grown on MM agar plates supplemented with 0.5% (w/v) mannitol and 1% (w/v) glycerol as carbon sources. For metabolomics experiments, ten biological replicates were analyzed for each sample.

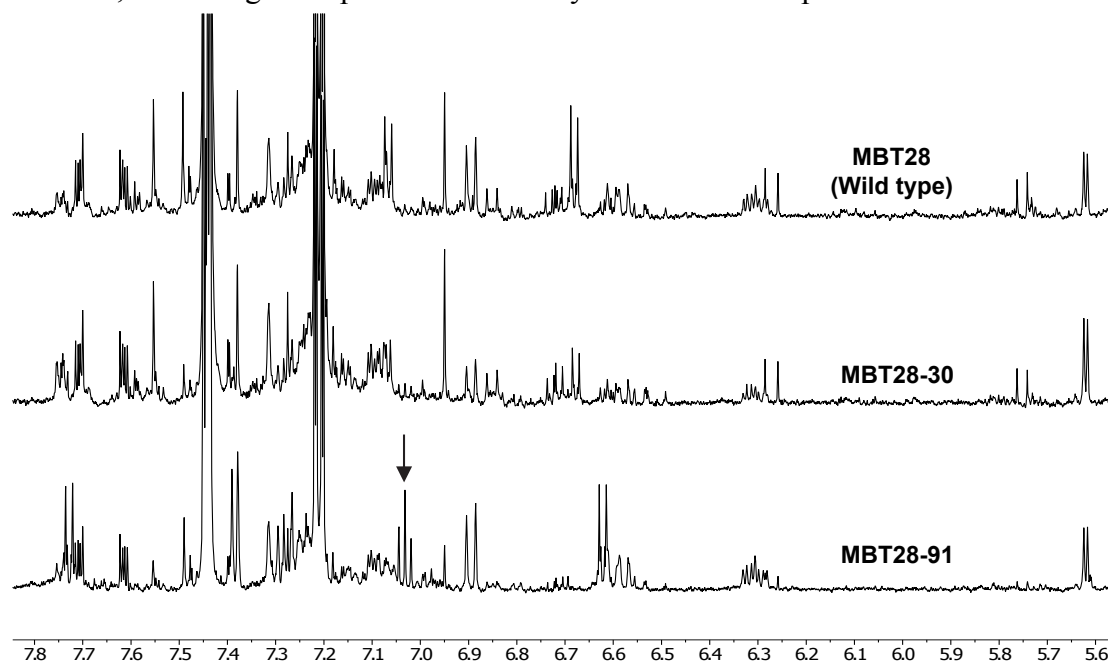


Figure 2. Comparison of ¹H NMR spectra. ¹H NMR (600 MHz) was done on samples from *Streptomyces* sp. MBT28 and its Str^R mutants MBT28-30 and MBT28-91. The solvent was CH₃OH-*d*₄, and peaks are shown for the region of δ 5.6—7.8 ppm. The signal δ 7.03 (t, *J* = 7.2 Hz) that was exclusively identified in mutant MBT28-91 is highlighted by an arrow.

While only small differences were found in the ^1H NMR spectra between the parental strain and its mutant MBT28-30, those of the antibiotic-overproducing mutant MBT28-91 displayed distinctly different NMR signals, particularly in the aromatic region (δ 5.7–7.8, Figure 2). The compounds were elucidated by 2D NMR including COSY, HSQC, and HMBC experiments (Table S1). The signals of the main aromatic compounds in the wild-type strain were attributed to caffeic acid, *p*-cresol, and syringic acid, which were not detected in MBT28-91. Conversely, ^1H NMR spectra of MBT28-91 specifically exhibited resonances for 4-hydroxybenzoic acid and an unknown compound with typical resonances at δ 7.03 (t, J = 7.2 Hz), which were absent in the other two strains. The known antimicrobial compound borrelidin (δ 6.60, 6.90, and 6.31), which was previously shown to be produced by MBT28,²² was detected in all samples, but its production level was 10% lower in MBT28-91 as compared to the parent or mutant MBT28-30. This strongly suggests that borrelidin was not the causative agent of the increased bioactivity.

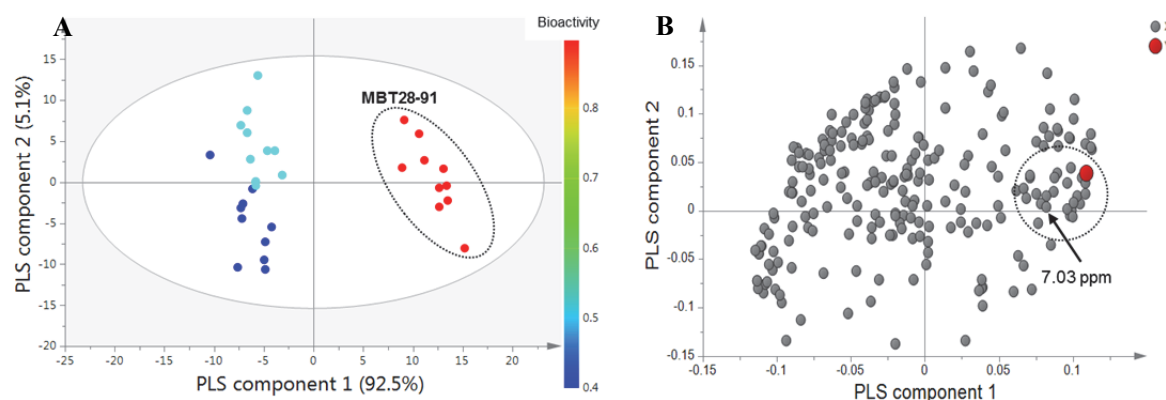


Figure 3. Multivariate data analysis. The PLS-DA score plot (A) reveals good separation between wild-type *Streptomyces* sp. MBT28 (blue) and its mutant derivatives MBT28-30 (light blue) and MBT28-91 (red) with low, medium, and high antibiotic activity against *B. subtilis*, respectively. The corresponding loading plot (B) presents the NMR signals (dashed circles) that make major contributions to the separation of the samples (and the bioactivity) of MBT28-91 from those of the other two strains. The arrow refers to the characteristic proton signal δ 7.03 (t, J = 7.2 Hz). X, primary variable of the chemical shift; Y, bioactivity.

Multivariate data analysis was conducted to correlate NMR signals with the bioactivity, which does not require extensive pre-fractionation and chromatographic separation of the crude extract.³⁵ An unsupervised principal component analysis (PCA) revealed that MBT28-91 fingerprints did not overlap with those of the two other strains, but it was rather difficult to correlate the significant signals to the improved activity in MBT28-91. Projection to latent structures (PLS) is a robust multivariate linear regression technique to predict a set of dependent variables from a set of independent variables or predictors.³⁶ To investigate the detailed correlation between bioactivity and metabolites, PLS was employed as a supervised data reduction algorithm in which bioactivity values - determined as the diameter of the inhibition zone in an agar diffusion assay - were used instead of dummy Y-variables. The score plot of the PLS analysis (Figure 3A) showed that not only samples from the highly bioactive MBT28-91 were well separated, but that separation was also found between all independent samples of MBT28-30 and those of the wild-type strain. Examining the corresponding loading plot again identified characteristic signals at δ 7.03 (t, J = 7.2 Hz) that correlated very well to the enhanced bioactivity in MBT28-91 (Figure 3B), but full structure

elucidation could not be done on crude extracts. The supervised model was validated by a permutation test that is a calculation of quality of the fit and the predictive ability. For antimicrobial activity, the R² and Q² values for PLS analysis were 0.98 and 0.95, respectively, indicating an excellent reproducibility and predictability.

2.3. Identification of 7-prenylisatin as the elicited bioactivity

To elucidate the structure of the bioactive component produced by mutant MBT28-91, NMR-guided fractionation³⁷ tracing the characteristic signal δ 7.03 (t, J = 7.2 Hz) was conducted. Extracts of MBT28-91 grown on MM agar plates were sequentially defatted using *n*-hexane, followed by silica gel, and Sephadex LH-20 chromatography. After each step, the collected fractions were analysed by ¹H NMR for target signal tracking, and relevant fractions were subjected to preparative TLC chromatography. A bright orange compound (0.35 mg, R_f = 0.45 on TLC, developed in CHCl₃) was readily obtained. Its ¹H NMR spectrum (600 MHz, CDCl₃) showed an aromatic ring ABC spin system, δ_H 7.39 (brd, J = 7.8 Hz, H-6), 7.06 (t, J = 7.8 Hz, H-5), 7.49 (brd, J = 7.8 Hz, H-4). A set of typical signals attributable to a prenyl group, including δ_H 3.31 (brd, J = 7.2 Hz, H-8), 5.24 (m, H-9), 1.80 (d, J = 1.2 Hz, H₃-11), and 1.78 (brs, H₃-12) were also observed (Table 1). The prenyl chain was located at the C-7 position of a benzene ring, which was established from the key HMBC correlation from δ_H 3.31 (d, J = 7.2 Hz, H-8) to a heteroatom-bearing carbon δ_C 147.9 (C-7a). The HMBC correlation from H-4 to the carboxyl carbon δ_C 183.1 (C-3) proved its conjugation to the benzene ring. The 1*H*-indole-2,3-dione backbone (isatin) was ultimately confirmed by the chemical formula C₁₃H₁₃NO₂ established from high-resolution mass [M + H]⁺ peak at m/z 216.1018 (calculated for C₁₃H₁₄NO₂ 216.1019). Thus, compound **1** was identified as 7-prenylisatin, not previously identified in nature (Figure 4).

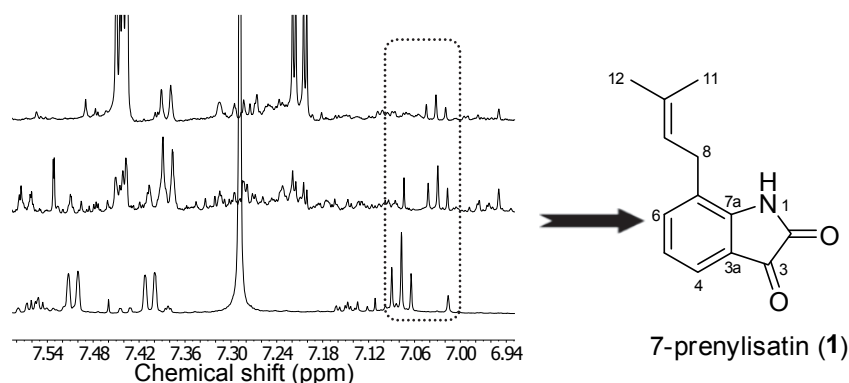


Figure 4. NMR-guided purification of 7-prenylisatin from *Streptomyces* sp. MBT28-91. Characteristic proton signals δ 7.03 (t, J = 7.2 Hz) that correlated to the high bioactivity are boxed (dashed box), which were used as tracking signals for chromatographic separation. Chemical shift divergence was due to the different deuterated solvents used (methanol-*d*₄ for top and middle rows, and CDCl₃ for bottom row).

Isatins exhibit antibacterial, antifungal, and cytotoxic properties.³⁸ However, only a few isatin-type compounds of microbial origin have been identified.³⁹⁻⁴¹ 5-prenylisatin from the fungus *Chaetomium globosum* was reported to have antifungal activity,³⁹ while 6-prenylisatin was the first example of an isatin derivative discovered in a streptomycete, namely *Streptomyces albus*, which showed antimicrobial activity against Gram-positive bacteria.⁴⁰ Purified 7-prenylisatin was tested for its efficacy against *B. subtilis* using a microtitre plate-based test for the minimal inhibitory concentration (MIC) (see Materials and Methods

section for details). This showed that 7-prenylisatin acts as an antibiotic with an MIC of around 25 $\mu\text{g/ml}$ against *B. subtilis*, which is similar to that of 6-prenylisatin. No antimicrobial activity was observed against the Gram-positive *Staphylococcus aureus* or the Gram-negative *Escherichia coli* K12.

Table 1. ^1H and ^{13}C NMR data assignments for 7-prenylisatin (1) ^a

NO.	δ_{C} , type	δ_{H} (J in Hz)	HMBC correlations	COSY correlations
2	159.2, C			
3	183.1, C			
3a	118.4, C			
4	123.7, CH	7.49, brd (7.8)	C-3, C-6, C-7a	H-5
5	124.0, CH	7.06, t (7.8)	C-7, C-3a	H-4, H-6
6	139.1, CH	7.39, brd (7.8)	C-8, C-7a, C-4	H-5
7	125.1, C			
7a	147.9, C			
8	29.7, CH ₂	3.31, brd (7.2)	C-7a, C-6, C-7, C-9, C-10	
9	120.0, CH	5.24, m	C-8, C-11, C-12	H-8, H-11, H-12
10	136.1, C			
11	25.9, CH ₃	1.80, d (1.2)	C-9, C-10, C-12	H-9, H-8
12	18.3, CH ₃	1.78, brs	C-9, C-10, C-11	H-9, H-8

^a **1** recorded in CDCl_3 . Proton coupling constants (J) in Hz are given in parentheses. ^1H NMR and ^{13}C APT NMR spectra were recorded at 600 MHz. All chemical shift assignments were done on the basis of 1D and 2D NMR techniques.

2.4. Identification of the biosynthetic gene cluster for 7-prenylisatin

A biosynthetic route for 7-prenylisatin was proposed based on the existing literature for related isatin compounds (Figure 5). The initial step of isatin biosynthesis is prenylation at the C-7 carbon of tryptophan, followed by cleavage of the amino acid chain to generate the intermediate 7-prenylindole.^{42, 43} This intermediate is then further oxidized at C-2 and/or C-3 into the isatin skeleton through indoxyl or oxindole.⁴⁴⁻⁴⁶

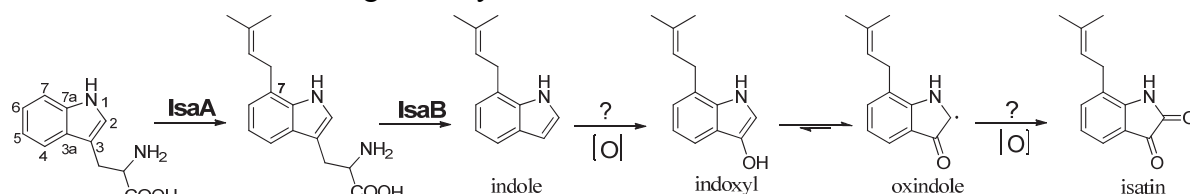


Figure 5. Proposed biosynthetic pathway of 7-prenylisatin. IsaA is the likely enzyme for prenylation at the C-7 carbon of tryptophan and IsaB for the subsequent cleavage of the amino acid chain. The intermediate 7-prenylindole is then further oxidized into 7-prenylisatin by a yet uncharacterized oxygenase.

To characterize the 7-prenylisatin biosynthetic pathway, the genome of *Streptomyces* sp. MBT28 was sequenced by Illumina paired end sequencing, and the resulting draft genome sequence was annotated as described.⁴⁷ The genes responsible for synthesis of the 1*H*-indole-2,3-dione scaffold (isatin) have not previously been identified in bacteria. However, the prenyltransferase (PTase) catalyzing the initial prenylation of the indole nucleus has been well characterized. For instance, enzymes that prenylate at all non-bridgehead atoms of the indole ring (N-1, C-2, C-3, C-4, C-5, C-6, and C-7) have been identified (Figure S2 and Table 2).^{48, 49} Therefore, the genome was scanned for genes encoding possible PTases, which might give clues as to which gene cluster is responsible for 7-prenylisatin biosynthesis. Genes for indole PTases have been found in actinomycetes in two different genomic contexts,⁴² Type A in combination with a tryptophanase, and Type B in combination with a flavin-dependent

monooxygenase (FMO). For instance, in *Streptomyces* sp. SN-593, IptA (Type A) installs a dimethylallyl group on the C-6 position of the tryptophan indole core, while in *Streptomyces coelicolor* A3(2), a Type B PTase prenylates tryptophan at C-5.⁴³ Interestingly, the *Streptomyces* sp. MBT28 genome contained both types of gene clusters (Tables S2 and S3). However, considering that a tryptophanase is required for cleavage of the side chain (Figure 5), the Type A cluster was the most likely candidate for the bioassembly of 7-prenylisatin.

Table 2. Examples of prenyltransferases (PTs) for prenylation of indoles. All possible positions in the indole core can be prenylated, depending on the enzyme. See also Figure S2.

Enzyme	Function	PT site	Enzyme origin	Homology	Accession number
IsaA	regular	C-7	<i>Streptomyces</i> sp. MBT28	100%	PRJNA279348
LtxC	reverse	C-7	<i>Lyngbya majuscula</i>	48%	NP_631515
EAL92290	regular	C-7	<i>Aspergillus fumigatus</i>	23%	ABS89001.1
IptA	regular	C-6	<i>Streptomyces</i> sp. SN-593	46%	BAJ07990.1
EAW08391	regular	C-5	<i>Aspergillus clavatus</i> NRRL 1	29%	EAW08391
SCO7467	regular	C-5	<i>Streptomyces coelicolor</i> A3(2)	48%	NP_631515
FgaPT2	regular	C-4	<i>Aspergillus fumigatus</i>	33%	AAX08549.1
MaPT	regular	C-4	<i>Malbranchea aurantiaca</i>	36%	ABZ80611.1
AnaPT	reverse	C-3	<i>Neosartorya fischeri</i>	23%	EAW16181.1
CdpNPT	reverse	C-3	<i>Aspergillus fumigatus</i>	45%	ABR14712.1
FgaPT1	reverse	C-2	<i>Aspergillus fumigatus</i>	22%	EAL94098
FtmPT1	regular	C-2	<i>Aspergillus fumigatus</i>	35%	AAX56314.1
FtmPT2	regular	N-1	<i>Aspergillus fumigatus</i>	35%	AAX56314.1
CymD	reverse	N-1	<i>Salinispora arenicola</i>	31%	ABW00334

To validate that 7-prenylisatin is synthesized by a Type A indole PTase, we made use of the fact that the expression of biosynthetic genes correlates closely with the bioactivity of interest, so that fluctuations in the proteome (as assessed by quantitative proteomics) correlate directly to the level of the bioactivity.³⁰ Protein extracts of MBT28 and its mutants MBT28-30 and MBT28-91 had been obtained from the same cultures as used for metabolomics, and stable isotope dimethyl labeling (see Materials and Methods section for details) and subsequent LC-MS/MS analysis were used for relative quantification of protein levels between the three strains. 1,382 proteins were positively identified and the changes in expression levels of 1,035 proteins could be quantified. Based on the intensity-dependent significance-*B* value,⁵⁰ 177 of the quantified proteins were significantly enhanced or reduced ($p < 0.05$) in at least one comparison (Table S4 and Figure S3). Among the proteins that were significantly enhanced in strain MBT28-91 in comparison to the wild-type strain, we found the two key enzymes of the Type A indole PTase gene cluster, namely the aromatic prenyltransferase (IsaA) and the tryptophanase (IsaB) (Table 3 and Table S2). The two enzymes were also upregulated in strain MBT28-30 compared to the wild-type strain, but much less than for strain MBT28-91. Thus, the expression level of Type A indole PTase gene cluster showed excellent correlation to the amount of 7-prenylisatin produced. In contrast, none of the proteins encoded by the Type B cluster were detected, which may therefore not be expressed under these conditions (Table S3).

Tryptophan is a known key precursor for the biosynthesis of the prenylated indole alkaloids and acts as the biogenetic origin of indole rings.^{51, 52} Indole alkaloids that contain prenylated tryptophan moieties at their core represent an important class of natural products for which considerable synthetic and biosynthetic interest exists. Gene clusters containing genes for indole prenyltransferases are widely distributed in actinomycetes. Although some

prenylated indole derivatives have been isolated from *Streptomyces* species,^{42, 43, 53, 54} only two studies on the biosynthesis of prenylated indoles have been reported.^{41, 42} IsaA is the third actinomycete prenyltransferase known to catalyze the addition of a dimethylallyl group to the C-7 of tryptophan. It is yet unclear which gene(s) are responsible for the oxidation of indole into isatin backbone, as no such genes were found within or in close proximity of the *isa* gene cluster. Three proteins with an oxidation-related PQQ domain were all found significantly up-regulated ($p < 0.05$) in the overproducing strain as compared to the parental strain (Table S4). Further analysis is required to establish whether one of these enzymes may execute the oxidation reaction to generate the final isatin moiety.

Table 3. Differential expression (in fold change mutant/wt) of the proteins of the Type A indole PTase gene cluster as determined by quantitative proteomics. See also Table S4. The sequences have been deposited as Genbank Bioproject Accession PRJNA279348.

Protein	MBT28-91/WT (² log)	Significance -B	MBT28-30/WT (² log)	Significance -B	Description*
Isa Regulator	-	-	-	-	Transcriptional regulator
IsaA	3.0	0.023	0.9	0.116	Aromatic prenyltransferase
IsaB	2.8	0.031	1.6	0.040	Tryptophanase

* based on blast homology searches

3. CONCLUSION

The increasingly high frequency of re-discovery of known antibiotics using routine screening methods poses the challenge of prioritizing the bioactive molecules of interest.^{2, 5, 55, 56} Starting from a ‘novel’ biosynthetic cluster identified by genome sequencing is tricky, because the compound may already have been seen. Our work followed a strategy based on (i) fluctuations in the production level of the compound of interest; (ii) metabolomics combined with multivariate data analysis to correlate bioactive molecules to the observed bioactivity, followed by NMR-guided purification; and (iii) correlating the changes in gene expression profiles to the bioactivity to identify candidate gene clusters. This efficiently identified 7-prenylisatin, an antibiotic with activity against the Gram-positive bacterium *B. subtilis*, and also partially elucidated the biosynthetic pathway and the responsible biosynthetic genes. Further experiments are required to elucidate the mode of action and the activity spectrum of 7-prenylisatin and related compounds of the (prenyl)isatin family. The integrated approach of metabolomics as an efficient way to elucidate the nature of any bioactivity of interest, even in complex crude extracts, combined with eliciting strategies for the activation of novel bioactive molecules, forms a promising pipeline in the search for novel antibiotics and other medically relevant natural products.

4. EXPERIMENTAL SECTION

4.1. General Experimental Procedures

FT-IR spectra were recorded on a Shimadzu FT-IR 83000 spectrometer. UV measurements were performed using a Shimadzu UV mini-1240. NMR spectra were recorded in MeOD on a

Bruker 600 MHz calibrated to a residual $\text{CH}_3\text{OH-}d_4$ (3.30 ppm). The UHPLC-TOF-MS analyses were performed on an Ultimate 3000 UHPLC system (ThermoScientific, Pittsburgh, PA, USA) coupled to a micro-ToF-2Q mass spectrometer from Bruker Daltonics (Bremen, Germany) with an electrospray (ESI) interface. Silica gel (pore size 60 Å, 70–230 mesh) for open column chromatography was purchased from Sigma-Aldrich (St. Louis, MO, USA). Sephadex LH-20 (GE Healthcare Life Sciences, Eindhoven, The Netherlands) was used for size-exclusion chromatography. Silica gel 60 F254, 1 mm (Merck, Darmstadt, Germany) was used for preparative TLC separation. All organic solvents and chemicals were of analytical or HPLC grade, depending on the experiment.

4.2. Strains and culturing conditions

Streptomyces sp. MBT28 (Biosample Accession 3444561) was obtained from the Leiden University strain collection.²² Spores were stored in 20% glycerol and maintained in -20°C . Soy flower agar medium (SFM⁵⁷) was used to generate streptomycin-resistant (Str^R) mutants and for preparing spore suspensions. Minimal medium (MM) agar plates⁵⁷ supplemented with 0.5% (w/v) mannitol and 1% (w/v) glycerol as carbon sources were used for activity tests and to prepare samples for metabolomics and proteomics experiments. *Bacillus subtilis* 168 was grown in Luria-Bertani broth (LB) at 37°C . Imaging was done using a Zeiss standard 25 phase-contrast microscope and a Zeiss Lumar V12 stereo microscope as described.⁵⁸

4.3. Selection of streptomycin-resistant mutants and antibiotic activity assays

Str^R mutants were selected by plating 10 μl of spores (1×10^7 cfu/ml) onto SFM agar plates containing streptomycin (10 $\mu\text{g/ml}$) at 30°C . Colonies typically developed after nine days. For bioactivity screening, colonies were grown for six days on MM agar plates and overlaid with LB agar containing 100 μl of exponentially *B. subtilis* cells (OD_{600} of 0.4). The activity was initially assessed as the zone of growth inhibition after overnight incubation at 37°C . MIC values against *B. subtilis* were carried out as described.²² All MIC determinations were performed in triplicate. Ampicillin was used as positive control, and the solvent chloroform as the negative control.

4.4. PCR amplification

PCRs were performed in a minicycler (MJ Research, Watertown, MA, USA), using Phusion polymerase (Stratagene, La Jolla, CA) as described.⁵⁹ *rpsL* was amplified from genomic DNA using oligonucleotides *rpsL*_For (5'-CGGCACACAGAAACCGGAGAAG) and *rpsL*_Rev (5'-GTCGATGATGACCGGGCGCTTCG), and *rsmG* was amplified using oligonucleotides *rsmG*_For (5'-TGACGAATTCGTCAGCCTGATAGTTCTGGTGG) and *rsmG*_Rev (5'-TCAGTAAGCTTGCCGTGCAGCAGTGAGCGACG).

4.5. Metabolomics sampling and NMR-guided separation

Streptomyces were grown on MM agar plates supplemented with 0.5% mannitol and 1% glycerol. The agar with mycelia was cut into small pieces and extracted with ethyl acetate by soaking in the solvent overnight at room temperature. Samples were dried under reduced pressure at 40°C and redissolved in methanol. The solvent was then evaporated at room temperature under nitrogen gas flow, and subsequently dipped into liquid nitrogen and

lyophilized using a freeze dryer (Edwards Ltd., Crawley, UK). Ten replicates were performed for each sample.

For enrichment of isatin compounds, crude extracts were partitioned between methanol and *n*-hexane to remove the lipids. The resolved methanol fraction was subsequently separated by silica gel (pore size 60 Å, 70–230 mesh) column chromatography employing step-wise gradient elution by *n*-hexane, *n*-hexane-CHCl₃ (1:1), CHCl₃, CHCl₃-acetone (1:1), and acetone. All fractions were subsequently dissolved in CH₃OH-*d*₄ and subjected to NMR measurement. The *n*-hexane-CHCl₃ (1:1) elution fraction containing the target proton signals was firstly separated by Sephadex LH-20 (GE Healthcare Life Sciences, Eindhoven, The Netherlands) eluting with methanol, and further purified by preparative TLC (Merck, Darmstadt, Germany), migrated with solvent system CHCl₃-acetone (20:1) and detected under UV light at 254 nm wavelength.

7-prenylisatin (1): orange, amorphous powder, UV (MeOH) $\lambda_{\max}$ (log ϵ) 244 (4.26), 314 (3.32) nm; IR $\nu_{\max}$ 2922, 1732, 1716, 1558, 1456 cm⁻¹; ¹H and ¹³C NMR data, see Table 1; HRMS (positive mode) *m/z* 216.1018 [M + H]⁺ (calculated for C₁₃H₁₄NO₂ 216.1019), 238.0816 [M + Na]⁺ (calculated for C₁₃H₁₃NO₂Na 238.0844).

4.6. NMR measurement and multivariate data analysis

NMR sample preparation and NMR measurement was essentially done as described previously.^{60, 61} 500 µl of CH₃OH-*d*₄ was added to the freeze-dried sample. Subsequently, the mixture was vortexed for 10 s and sonicated for 20 min at the frequency of 42 kHz using an Ultrasonicator 5510E-MT (Branson, Danbury, CT, USA), followed by centrifugation at 13,000 g at room temperature for 5 min. The supernatant (300 µl) was transferred to a 3 mm micro NMR tube and analyzed. ¹H NMR spectra were recorded at 25 °C on a 600 MHz Bruker DMX-600 spectrometer (Bruker, Karlsruhe, Germany) operating at a proton NMR frequency of 600.13 MHz. Deuterated methanol was used as the internal lock. Each ¹H NMR spectrum consisted of 128 scans using the following parameters: 0.16 Hz/point, pulse width (PW) = 30 (11.3 s) and relaxation delay (RD) = 1.5 s. Free induction decays (FIDs) were Fourier transformed with a line broadening (LB) = 0.3 Hz. The resulting spectra were manually phased and baseline corrected, and calibrated to methanol at 3.30 ppm, using XWIN NMR (version 3.5, Bruker). The ¹H NMR data files were processed as described.⁶² The AMIX software (Bruker Biospin GmbH) was used to convert the ¹H NMR spectra to an ASCII file, with total intensity scaling and calibration to methanol signal (δ 3.30 ppm). Bucketing or binning was performed and the spectral data were reduced to included regions of equal width (0.04 ppm) equivalent to the region of δ 0.30–10.02. The regions of δ 4.85 – 4.95 and δ 3.25 – 3.35 were removed in the analysis because of the remaining signal of residual solvents of HDO and CD₃OD, respectively. Multivariate data analysis was performed with the SIMCA-P+ software (Version 13.0, Umetrics, Umeå, Sweden).

4.7. Genome sequencing, assembly, and annotation

Genome sequencing and annotation was done essentially as described.⁴⁷ Illumina/Solexa sequencing on Genome Analyzer IIx and sequencing on PacBio RS were outsourced at BaseClear BV (Leiden, Netherlands). 100-nt paired-end reads were obtained and the quality of the short reads verified using FastQC

(<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/>). Depending on quality, reads were trimmed to various lengths at both ends. Processed raw reads were subsequently used as input for the Velvet assembly algorithm. The resulting contigs were analyzed using the GeneMark.hmm algorithm with the *Streptomyces coelicolor* genome as the reference for identification and annotation of coding sequences (CDSs).⁶³ The genome has been deposited under Bioproject Accession PRJNA279348.

4.8. Quantitative proteomics analysis

To prepare samples for proteomics, strains were grown on MM agar plates overlaid with 0.1 µm polycarbonate filter papers prior to inoculation. After six days of incubation at 30 °C, mycelia were scraped from the filters. Protein extraction, in-solution digestion, C₁₈ column dimethyl labeling, strong cation-exchange chromatography (SCX) fractionation, and LC-MS/MS analysis were performed as previously described.^{30, 64} Generally, mycelia were washed and then sonicated for disruption of the cell wall. After removal of cell debris, protein concentration was measured using a Bradford assay and 0.167 mg of protein was precipitated for each sample. Proteins were dissolved using RapiGest SF Surfactant (Waters) and treated with iodoacetamide before trypsin digestion. The primary amines of the peptides were dimethyl labeled using three combinations of isotopomers of formaldehyde and cyanoborohydride on Ultra-Clean SPE C₁₈ columns (Altech, 200 mg, Thermo Fisher Scientific, USA) as described, via CH₂O + NaBH₃CN, CD₂O + NaBH₃CN, and ¹³CD₂O + NaBD₃CN.⁶⁵ Light, medium, and heavy-labelled samples with 4 Da mass differences were mixed to obtain 0.5 mg of peptides for fractionation by SCX. 24 fractions were collected and subjected to LC-MS/MS analysis on an LTQ-orbitrap MS (ThermoFisher Scientific, Pittsburgh, Pennsylvania, USA).⁶⁴ Data analysis was performed using MaxQuant 1.4.1.2.⁵⁰ Dimethyl labeling was selected as the quantification method, carbamidomethylation of cysteine was selected as fixed modification, and oxidation of methionine was set as variable modification.⁶⁶ MS/MS spectra were searched against a database of translated coding sequences. Peptide and protein identification FDR were set to 0.01, and a minimum of two quantification events was specified to obtain a protein quantification. To determine the significance of the found expression ratios, intensity-dependent B-significance values⁵⁰ were calculated.

REFERENCES

- (1) WHO, *Antimicrobial Resistance: Global Report on Surveillance*. Geneva, Switzerland, 2014.
- (2) Baltz, R. H. *Cur.r Opin. Pharmacol.* **2008**, 8, 557-563.
- (3) Cooper, M. A.; Shlaes, D. *Nature* **2011**, 472, 32.
- (4) Genilloud, O.; Gonzalez, I.; Salazar, O.; Martin, J.; Tormo, J. R.; Vicente, F. *J. Ind. Microbiol. Biotechnol.* **2011**, 38, 375-389.
- (5) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat. Rev. Drug. Discov.* **2007**, 6, 29-40.

- (6) Bérdy, J. J. *Antibiot. (Tokyo)* **2005**, 58, 1-26.
- (7) Hopwood, D. A., *Streptomyces in nature and medicine: the antibiotic makers*. Oxford University Press: New York, 2007.
- (8) Pimm, S.; Russell, G. J.; Gittleman, J. L.; Brooks, T. M. *Science* **1995**, 269, 347-350.
- (9) Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A. *Nature* **2002**, 417, 141-147.
- (10) Cruz-Morales, P.; Vijgenboom, E.; Iruegas-Bocardo, F.; Girard, G.; Yanez-Guerra, L. A.; Ramos-Aboites, H. E.; Pernodet, J. L.; Anne, J.; van Wezel, G. P.; Barona-Gomez, F. *Genome Biol. Evol.* **2013**, 5, 1165-1175.
- (11) Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S. *Nat. Biotechnol.* **2003**, 21, 526-531.
- (12) Bibb, M. J. *Curr. Opin. Microbiol.* **2005**, 8, 208-215.
- (13) van Wezel, G. P.; McDowall, K. J. *Nat. Prod. Rep.* **2011**, 28, 1311-1333.
- (14) Gross, H. *Curr. Opin. Drug. Discov. Devel.* **2009**, 12, 207-219.
- (15) Medema, M. H.; Breitling, R.; Bovenberg, R.; Takano, E. *Nat Rev Microbiol* **2011**, 9, 131-137.
- (16) Nett, M.; Ikeda, H.; Moore, B. S. *Nat. Prod. Rep.* **2009**, 26, 1362-1384.
- (17) van Wezel, G. P.; McKenzie, N. L.; Nodwell, J. R. *Methods Enzymol* **2009**, 458, 117-141.
- (18) Yoon, V.; Nodwell, J. R. *J. Ind. Microbiol. Biotechnol.* **2014**, 41, 415-424.
- (19) Zhu, H.; Sandiford, S. K.; van Wezel, G. P. *J. Ind. Microbiol. Biotechnol.* **2014**, 41, 371-386.
- (20) Schumann, J.; Hertweck, C. *J. Biotechnol.* **2006**, 124, 690-703.
- (21) Craney, A.; Ozimok, C.; Pimentel-Elardo, S. M.; Capretta, A.; Nodwell, J. R. *Chem. Biol.* **2012**, 19, 1020-1027.
- (22) Zhu, H.; Swierstra, J.; Wu, C.; Girard, G.; Choi, Y. H.; van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. *Microbiology* **2014**, 160, 1714-1725.
- (23) Hosaka, T.; Ohnishi-Kameyama, M.; Muramatsu, H.; Murakami, K.; Tsurumi, Y.; Kodani, S.; Yoshida, M.; Fujie, A.; Ochi, K. *Nat. Biotechnol.* **2009**, 27, 462-464.
- (24) Ochi, K.; Tanaka, Y.; Tojo, S. *J. Ind. Microbiol. Biotechnol.* **2014**, 41, 403-414.
- (25) Tamehiro, N.; Hosaka, T.; Xu, J.; Hu, H.; Otake, N.; Ochi, K. *Appl. Environ. Microbiol.* **2003**, 69, 6412-6417.
- (26) Wang, G.; Hosaka, T.; Ochi, K. *Appl. Environ. Microbiol.* **2008**, 74, 2834-2840.
- (27) Fu, P.; Jamison, M.; La, S.; MacMillan, J. B. *Org. Lett.* **2014**, 16, 5656-5659.
- (28) Krug, D.; Muller, R. *Nat. Prod. Rep.* **2014**, 31, 768-783.
- (29) Wu, C.; Kim, H. K.; van Wezel, G. P.; Choi, Y. H. *Drug. Discov. Today. Technol.* **2015**, 13, 11-17.

-
- (30) Gubbens, J.; Zhu, H.; Girard, G.; Song, L.; Florea, B. I.; Aston, P.; Ichinose, K.; Filippov, D. V.; Choi, Y. H.; Overkleeft, H. S.; Challis, G. L.; van Wezel, G. P. *Chem. Biol.* **2014**, *21*, 707-718.
- (31) Tanaka, Y.; Komatsu, M.; Okamoto, S.; Tokuyama, S.; Kaji, A.; Ikeda, H.; Ochi, K. *Appl. Environ. Microbiol.* **2009**, *75*, 4919-4922.
- (32) Jaskunas, S. R.; Lindahl, L.; Nomura, M. *Nature* **1975**, *257*, 458-462.
- (33) Nishimura, K.; Hosaka, T.; Tokuyama, S.; Okamoto, S.; Ochi, K. *J Bacteriol* **2007**, *189*, 3876-3883.
- (34) Wu, C.; Choi, Y. H.; van Wezel, G. P. *J. Ind. Microbiol. Biotechnol* **2016**, *43*, 299-312.
- (35) Choi, H.-K.; Choi, Y. H.; Verberne, M.; Lefeber, A. W. M.; Erkelens, C.; Verpoorte, R. *Phytochemistry* **2004**, *65*, 857-864.
- (36) Abdi, H. *Wiley Interdisciplinary Reviews: Computational Statistics* **2010**, *2*, 97-106.
- (37) Grkovic, T.; Pouwer, R. H.; Vial, M.-L.; Gambini, L.; Noël, A.; Hooper, J. N. a.; Wood, S. a.; Mellick, G. D.; Quinn, R. J. *Angew. Chem. (Intl ed)* **2014**, *53*, 6070-6074.
- (38) Vine, K. L.; Matesic, L.; Locke, J. M.; Ranson, M.; Skropeta, D. *Anticancer Agents Med. Chem.* **2009**, *9*, 397-414.
- (39) Breinholt, J.; Demuth, H.; Heide, M.; Jensen, G. W.; Moller, I. L.; Nielsen, R. I.; Olsen, C. E.; Rosendahl, C. N. *Acta Chem. Scand.* **1996**, *50*, 443-445.
- (40) Graefe, U.; Radics, L. *J. Antibiot. (Tokyo)* **1986**, *39*, 162-163.
- (41) Sobolevskaya, M. P.; Denisenko, V. A.; Fotso, S.; Laach, H.; Menzorova, N. I.; Sibirtsev, Y. T.; Kuznetsova, T. A. *Russ. Chem. Bull.* **2009**, *57*, 665-668.
- (42) Ozaki, T.; Nishiyama, M.; Kuzuyama, T. *J. Biol. Chem.* **2013**, *288*, 9946-9956.
- (43) Takahashi, S.; Takagi, H.; Toyoda, A.; Uramoto, M.; Nogawa, T.; Ueki, M.; Sakaki, Y.; Osada, H. *J. Bacteriol.* **2010**, *192*, 2839-2851.
- (44) Gillam, E. M.; Notley, L. M.; Cai, H.; De Voss, J. J.; Guengerich, F. P. *Biochemistry* **2000**, *39*, 13817-13824.
- (45) Madsen, E.; Bollag, J. *Arch. Microbiol.* **1988**, *151*, 71-76.
- (46) Russell, G. A.; Kaupp, G. *J. Am. Chem. Soc.* **1969**, *228*, 3851-3859.
- (47) Girard, G.; Willemse, J.; Zhu, H.; Claessen, D.; Bukarasam, K.; Goodfellow, M.; van Wezel, G. P. *Antonie Van Leeuwenhoek* **2014**, *106*, 365-380.
- (48) Tanner, M. E. *Nat. Prod. Rep.* **2015**, *32*, 88-101.
- (49) Walsh, C. *ACS Chem. Biol.* **2014**, *9*, 2718-2728.
- (50) Cox, J.; Mann, M. *Nat. Biotechnol.* **2008**, *26*, 1367-1372.
- (51) Ishikura, M.; Yamada, K.; Abe, T. *Nat. Prod. Rep.* **2010**, *27*, 1630-1680.
- (52) Steffan, N.; Grundmann, a.; Yin, W.-B.; Kremer, A.; Li, S.-M. *Curr. Med. Chem.* **2009**, *16*, 218-231.
- (53) Motohashi, K.; Irie, K.; Toda, T.; Matsuo, Y. *J Antibiot. (Tokyo)* **2008**, *61*, 75-80.
- (54) Sasaki, T.; Igarashi, Y.; Ogawa, M.; Furumai, T. *J Antibiot. (Tokyo)* **2002**, *55*, 11-14.
- (55) Donadio, S.; Monciardini, P.; Alduina, R.; Mazza, P.; Chiocchini, C.; Cavaletti, L.; Sosio, M.; Puglia, A. M. *J. Biotechnol.* **2002**, *99*, 187-198.
- (56) Genilloud, O. *Recent Pat. Antiinfect. Drug. Discov.* **2012**, *7*, 189-204.
- (57) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A., *Practical Streptomyces genetics*. John Innes Foundation: Norwich, U.K., 2000.

- (58) Colson, S.; van Wezel, G. P.; Craig, M.; Noens, E. E. E.; Nothaft, H.; Mommaas, A. M.; Titgemeyer, F.; Joris, B.; Rigali, S. *Microbiology* **2008**, *154*, 373-382.
- (59) Colson, S.; Stephan, J.; Hertrich, T.; Saito, A.; van Wezel, G. P.; Titgemeyer, F.; Rigali, S. *J. Mol. Microbiol. Biotechnol.* **2007**, *12*, 60-66.
- (60) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Nat. Protoc.* **2010**, *5*, 536-549.
- (61) Wu, C.; van Wezel, G. P.; Hae Choi, Y. *J Antibiot. (Tokyo)* **2015**, *68*, 445-452.
- (62) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Nat. Protoc.* **2010**, *5*, 536-549.
- (63) Lukashin, A. V.; Borodovsky, M. *Nucleic. Acids Res.* **1998**, *26*, 1107-1115.
- (64) Li, N.; Kuo, C. L.; Paniagua, G.; van den Elst, H.; Verdoes, M.; Willems, L. I.; van der Linden, W. A.; Ruben, M.; van Genderen, E.; Gubbens, J.; van Wezel, G. P.; Overkleeft, H. S.; Florea, B. I. *Nat. Protoc.* **2013**, *8*, 1155-1168.
- (65) Boersema, P. J.; Raijmakers, R.; Lemeer, S.; Mohammed, S.; Heck, A. J. R. *Nat. Protoc.* **2009**, *4*, 484-494.
- (66) Gubbens, J.; Janus, M.; Florea, B. I.; Overkleeft, H. S.; van Wezel, G. P. *Mol. Microbiol.* **2012**, *86*, 1490-1507.

Chapter 5

Expanding the chemical space for natural products by *Aspergillus-Streptomyces* co-cultivation and biotransformation

Changsheng Wu ^{1,2}, Boris Zacchetti ¹, Arthur F.J. Ram ¹, Gilles P. van Wezel ¹, Dennis Claessen ¹, and Young Hae Choi ²

¹ *Molecular Biotechnology, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands;*

² *Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands*

Sci. Rep. **2015**, 5, 10868.

ABSTRACT

Actinomycetes and filamentous fungi produce a wide range of bioactive compounds, with applications as antimicrobials, anticancer agents or agrochemicals. Their genomes contain a far larger number of gene clusters for natural products than originally anticipated, and novel approaches are required to exploit this potential reservoir of new drugs. Here, we show that co-cultivation of the filamentous model microbes *Streptomyces coelicolor* and *Aspergillus niger* has a major impact on their secondary metabolism. NMR-based metabolomics combined with multivariate data analysis revealed several compounds that correlated specifically to co-cultures, including the cyclic dipeptide cyclo(Phe-Phe) and 2-hydroxyphenylacetic acid, both of which were produced by *A. niger* in response to *S. coelicolor*. Furthermore, biotransformation studies with *o*-coumaric acid and caffeic acid resulted in the production of the novel compounds (*E*)-2-(3-hydroxyprop-1-en-1-yl)-phenol and (2*E*,4*E*)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid, respectively. This highlights the utility of microbial co-cultivation combined with NMR-based metabolomics as an efficient pipeline for the discovery novel natural products.

1. INTRODUCTION

Filamentous microorganisms have an enormous potential for the production of a wide range of bioactive compounds, with application as among others antibiotic, anticancer or antifungal drugs.¹⁻³ In particular, the emergence of infectious diseases involving multi-drug resistant (MDR) bacterial pathogens since the 1980s has reinforced the screening efforts using filamentous microorganisms. It is likely that in terms of the biosynthetic potential of microbes we have only seen the tip of the iceberg.⁴ However, the return of investment on high-throughput screening campaigns has dropped alarmingly, resulting in an almost complete withdrawal of BigPharma from the antibiotic scene.⁵ However, genome sequencing has now revealed that prolific antibiotic producers, and in particular fungi⁶⁻⁸ and actinomycetes,⁹⁻¹¹ which had been intensively studied for decades, harbor far more biosynthetic gene clusters for natural products than originally anticipated. The challenge scientists now face is to elicit the expression of these potential treasures, via chemical, genetic or biological triggers.^{12, 13} Strategies that have been employed include heterologous expression,¹⁴ screening or selecting for antibiotic resistant strains,^{15, 16} and the use of chemical elicitors.^{17, 18} Besides the activation of cryptic gene clusters in known species, a second major reservoir for natural products is formed by microorganisms that have not yet been cultivated, which recently led to the discovery of novel antibiotics in unexpected microbial sources.^{19, 20}

Growth in microbial communities or interactions between different microorganisms is the next logical step in the search for new molecules, as many natural products can be expected to be activated specifically by competitors or symbionts in the natural habitat. Microbial co-cultures have been shown to elicit the biosynthesis of novel metabolites, thus increasing the pre-existent chemical diversity.^{21, 22} The model microorganisms that are studied intensively in our laboratory are *Streptomyces coelicolor*, a Gram-positive bacterium with a complex multicellular lifestyle,²³ and the fungus *Aspergillus niger*. It was shown previously that interaction with bacteria may have profound effects on the producing ability of *Aspergillus* species.^{24, 25} Here, we report the application of co-cultivation of *S. coelicolor* and *A. niger* for the discovery of novel natural products. In addition, the structurally related *o*-coumaric acid and caffeic acid were specifically converted into the novel molecules (*E*)-2-(3-hydroxyprop-1-en-1-yl)-phenol and (*2E,4E*)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid, respectively. We thereby apply NMR-based metabolomics, a technology that is particularly effective in identifying molecules in complex biological mixtures, which therefore facilitates the direct biochemical analysis of the community metabolism.²⁶ While routinely applied in plant metabolomics, its application in microbial research is rare.²⁷ The combination of microbial co-cultivation with NMR-based metabolomics is presented as an efficient way to discover new molecules.

2. RESULTS AND DISCUSSION

The interaction between the two filamentous model microbes *Streptomyces coelicolor* A3(2) M145 and *Aspergillus niger* N402 was studied in co-cultivations in submerged NMMP cultures. Preliminary experiments indicated that the pH of the medium had contrasting effects on the growth of both microbes: whereas *Streptomyces* species grow well at neutral pH, those of *Aspergillus* grow optimally at lower pH values. A phosphate buffer of pH 5 was optimal for

obtaining sufficient biomass of both microorganisms for a mutual interaction to occur. As the growth rate of *A. niger* was higher under these conditions, the fungal spores were added to the culture flasks 72 h after inoculation of *S. coelicolor*. To allow the ready visualization of the interaction between the two microorganisms, we used a test setup with *A. niger* AR19#1 which expresses eGFP.²⁸ *A. niger* AR19#1 formed pellets with a diameter of 100-500 μm that were rapidly colonized by *S. coelicolor* within 24 h after addition of *A. niger* to the *Streptomyces* culture. The colonization initiated with the adhesion of *S. coelicolor* to the *A. niger* biomass (Figure 1A & 1D), eventually leading to pellets that were largely covered with *Streptomyces* biomass after 72 h of co-cultivation (Figures 1B & 1E). Notably, loose fungal hyphae were evident in the culture broth in the following 24 h, inferring that the fungal mycelium was almost completely degraded by the bacterium (Figure 1C & 1F).

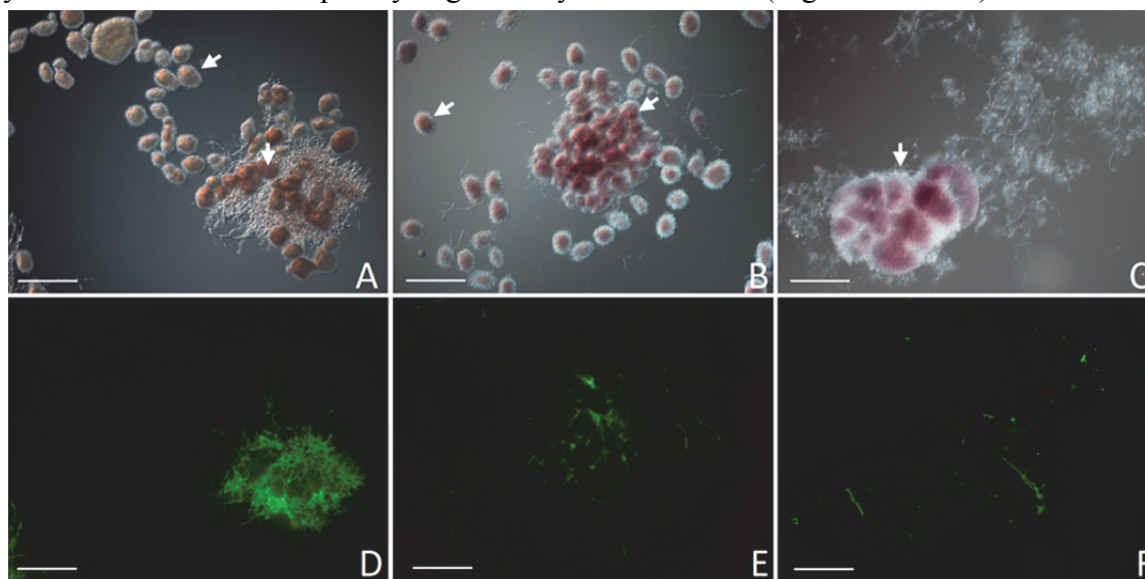


Figure 1. Mycelial interactions between *Aspergillus niger* AR19#1 (large open mycelial structures) and *Streptomyces coelicolor* A3(2) M145 (small red-pigmented pellets indicated by arrow heads) during co-culture. Strain AR19#1 is a derivative of *A. niger* N402 which expresses eGFP. Samples were observed following inoculation of *A. niger* at 24 h (A, D), 72 h (B, E), and 96 h (C, F). The reduction in fluorescence intensity in the eGFP-positive hyphae of *A. niger* (D-F) highlights the strong decline in viable fungal biomass in later stages of the co-culture. Bar, 250 μm (A, B) or 125 μm (C).

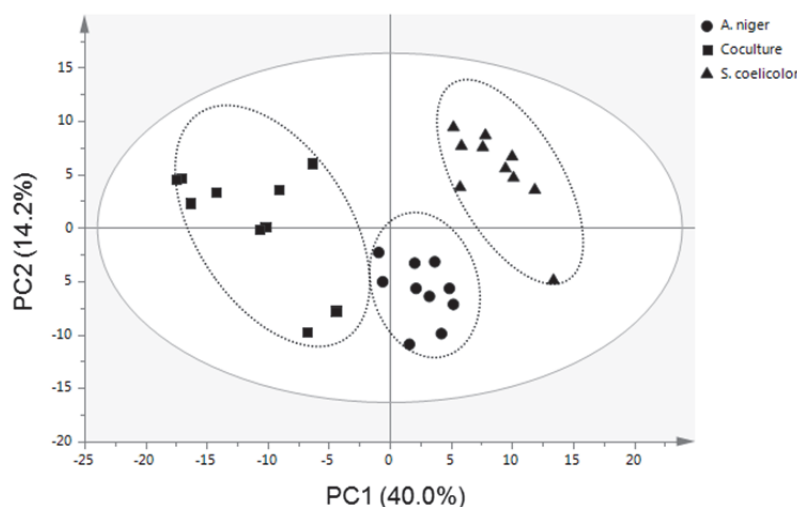


Figure 2. Unsupervised (PCA) multivariate data analysis of the ^1H NMR fingerprint data included in Figure S1. Triangles represent the *S. coelicolor* monoculture, circles the *A. niger* monoculture, while the co-culture of *S. coelicolor* and *A. niger* is represented by squares.

Previous studies have shown that evident metabolic changes occur when different organisms are grown together.²⁹ An NMR-based metabolomics approach was employed to study such effects in the *Aspergillus-Streptomyces* co-cultivations, using *S. coelicolor* A3(2) M145 and *A. niger* N402 (the plasmid-free parent of strain AR19#1). The method of NMR-based metabolomics was performed according to our previously published protocol, including sample preparation, NMR analysis, metabolite identification and multivariate data analysis.²⁶ In the current study, a total of 30 liquid cultures (10 for each monoculture and 10 for the coculture) were separately extracted followed by global NMR-profiling of the secondary metabolites (Figure S1). The ¹H NMR spectra were then subjected to statistical analysis. An unsupervised principal component analysis (PCA) revealed clustering of the samples in three groups, indicating that the co-culture fingerprints did not overlap with the two corresponding monoculture clusters (Figure 2). This showed that the dataset contained information that allowed the discrimination of the chemical composition of the co-culture from that of the monocultures. Features exclusively related to the co-culture were found within the dataset of induced compounds (see below), implying that microbial interactions modulated the biosynthetic pathways for the production of secondary metabolites. The PCA loading plot (not shown) allowed detecting biomarkers responsible for the discrimination among the three groups. The compounds were identified with different NMR techniques (1D and/or 2D experiments) and UHPLC-TOF-MS analysis (positive and/or negative modes) (Table 1) and the results were compared with spectroscopic data from literature. The major discriminators for the *S. coelicolor* culture were the well-studied pigmented antibiotics undecylprodigiosin and actinorhodin,^{30, 31} while the *A. niger* monoculture was abundant in the γ -pyrone derivative carbonarone A,³² as well as several minor components, namely naphtho- γ -pyrone aurasperone B, and fumonisins B₂ and B₄.^{33, 34} Notably, some compounds accounting for the PCA separation of the metabolites in the co-culture were identified as the previously described compounds cyclo-(Phe-Phe),³⁵ cyclo-(Phe-Tyr),³⁶ phenylacetic acid,³⁷ 2-hydroxyphenylacetic acid,³⁷ and furan-2-carboxylic acid (Figure 3).³⁸

Recently, the implementation of the co-cultivation strategy attracted considerable interest and application, since it proved to be an effective way to harvest unique structures with pronounced biological activities.²⁹ It may trigger biosynthetic pathways for ‘cryptic’ natural products that would otherwise remain silent under standard laboratory culture conditions. However, often it has not been established which of the co-cultivated microorganisms is actually responsible for manufacturing these metabolites. Here, we focused on the production of cyclo(Phe-Phe) that was produced abundantly in the co-culture of *A. niger* with *S. coelicolor*. Since *Aspergilli* produce 2,5-diketopiperazine-type compounds,^{29, 39} *A. niger* was the most obvious candidate. To test this hypothesis, *A. niger* was inoculated into the culture filtrate (supernatant without bacterial biomass) of a 5-day-old *S. coelicolor* culture and incubated for another two days. ¹H NMR metabolic profiling (Figure S2) unambiguously demonstrated that the culture filtrate of *S. coelicolor* effectively elicited the production of cyclo(Phe-Phe) and phenylacetic acid by *A. niger*. We hypothesized that *A. niger* metabolizes phenylalanine into phenylacetic acid, which is then further oxidized to 2-hydroxyphenylacetic acid,⁴⁰ while cyclo(Phe-Phe) can be synthesized by condensation and cyclization of two molecules of phenylalanine.^{41, 42}

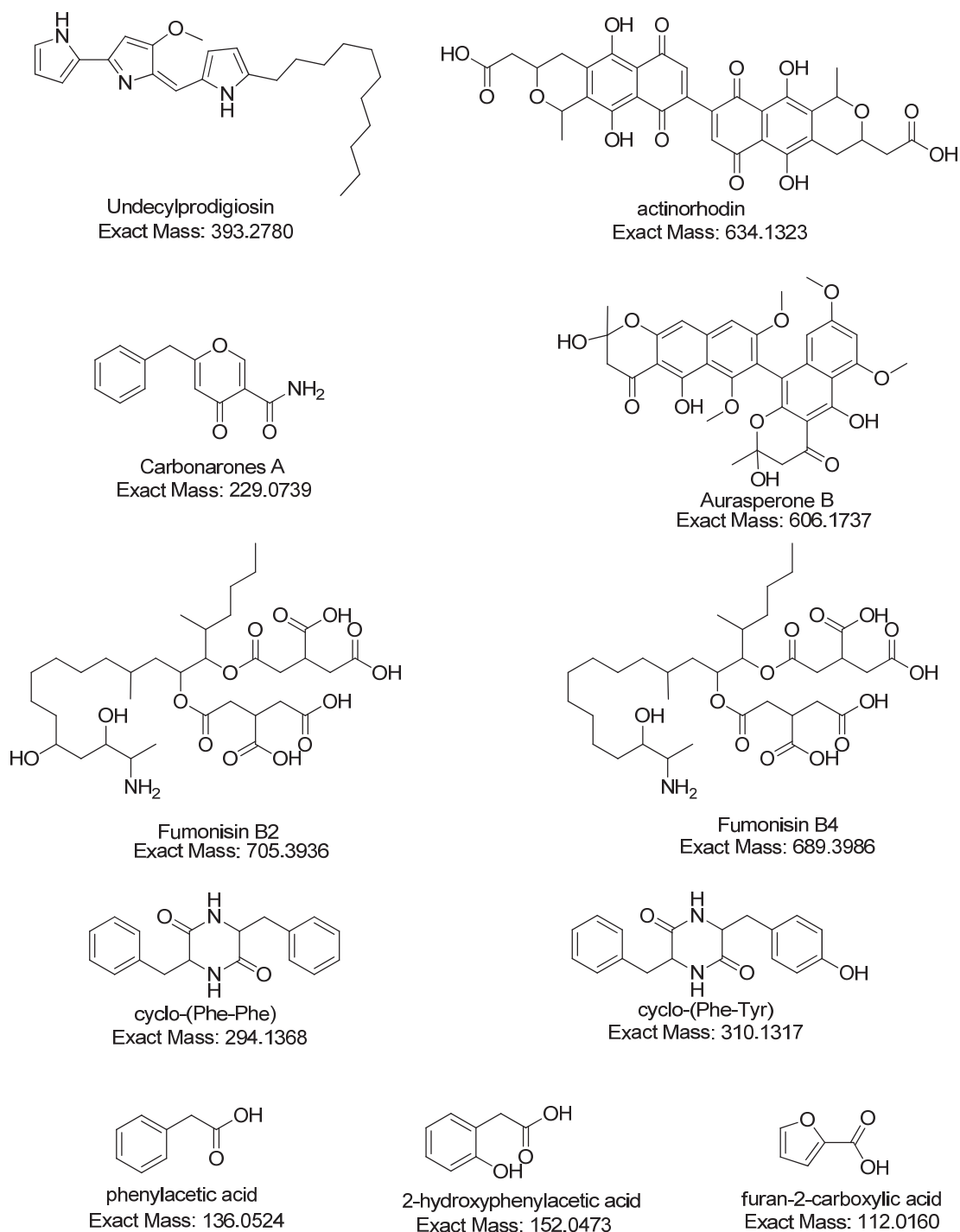


Figure 3. Major discriminating compounds responsible for the PCA separation (Figure 2) of the *S. coelicolor* and *A. niger* monocultures from their co-culture. *S. coelicolor* monoculture: undecylprodigiosin and actinorhodin; *A. niger* monoculture: carbonarones A, aurasperone B, fumonisin B2 and fumonisin B4; Co-culture: cyclo(Phe-Phe), cyclo(Phe-Tyr), phenylacetic acid, 2-hydroxyphenylacetic acid, and furan-2-carboxylic acid. The shown theoretical exact mass of each compound was calculated with *ChemDraw Ultra 12.0* software. The structure elucidation was done on the basis of NMR and/or experimental UHPLC-TOF-MS high resolution mass, and the spectral data assignments were summarized in Table 1.

Table 1. Spectral data assignments for the compounds displayed in Figure 3. All compounds were summarized according to their corresponding producers, namely *S. coelicolor* monoculture, *A. niger* monoculture, and referred co-culture. Compound identification was based on ¹H NMR and/or high resolution mass spectrometry, and compared with literature. Proton coupling constants (J in Hz) are given in parentheses.

Producer	Compounds	Molecular Formula	Characteristic ¹ H NMR Chemical shifts	Exact Mass	Experimental Mass	Ref
<i>S. coelicolor</i>	Undecylprodigiosin	C ₂₅ H ₃₅ N ₃ O	7.32 (dd, <i>J</i> = 2.4, 1.2); 6.37 (dd, <i>J</i> = 4.2, 1.2); 6.18 (d, <i>J</i> = 4.2); 2.78 (m)	393.2780	394.2854 [M + H] ⁺	31
	Actinorhodin	C ₃₂ H ₂₆ O ₁₄	7.38 (s)	634.1323	657.3169 [M + Na] ⁺ ; 633.1256 [M - H] ⁻	30
<i>A. niger</i>	Carbonarones A	C ₁₃ H ₁₁ NO ₃	8.88 (s); 7.24 (d, <i>J</i> = 8.4); 7.37 (m); 7.31 (m); 6.38 (s); 3.88 (s)	229.0739	230.0809 [M + H] ⁺ ; 252.0630 [M + Na] ⁺ ; 269.1353 [M + K] ⁺	32
	Fumonisin B2	C ₃₄ H ₅₉ NO ₁₄		705.3936	706.4009 [M + H] ⁺	34
	Fumonisin B4	C ₃₄ H ₅₉ NO ₁₃		689.3986	690.4063 [M + H] ⁺ ; 688.3917 [M - H] ⁻	33
	Aurasperone B	C ₃₂ H ₃₀ O ₁₂	6.64 (d, <i>J</i> = 2.4); 3.93 (s); 3.72 (s); 3.60 (s); 3.13 (d, <i>J</i> = 17.4); 3.12 (d, <i>J</i> = 17.4)	606.1737	607.1808 [M + H] ⁺ ; 629.1605 [M + Na] ⁺ ; 605.1665 [M - H] ⁻	33
	Cyclo-(Phe-Phe)	C ₁₈ H ₁₈ N ₂ O ₂	7.33 (t, <i>J</i> = 7.8); 7.25 (m); 7.10 (brd, <i>J</i> = 7.8); 4.08 (dd, <i>J</i> = 7.2, 4.2); 2.78 (dd, <i>J</i> = 13.2, 4.2); 2.14 (dd, <i>J</i> = 13.2, 7.2)	294.1368	295.1429 [M + H] ⁺ ; 317.1238 [M + Na] ⁺	35
Co-culture	Cyclo-(Phe-Tyr)	C ₁₈ H ₁₈ N ₂ O ₃	7.33 (t, <i>J</i> = 7.8); 7.25 (m); 7.10 (brd, <i>J</i> = 7.8); 6.88 (d, <i>J</i> = 8.4); 6.68 (d, <i>J</i> = 8.4)	310.1317	311.1379 [M + H] ⁺ ; 333.1185 [M + Na] ⁺	35
	Phenylacetic acid	C ₈ H ₈ O ₂	3.58 (s); 6.77 (m); 7.09 (m)	136.0524	135.0442 [M - H] ⁻	37
	2-hydroxyphenylacetic acid	C ₈ H ₈ O ₃	3.47 (s)	152.0473	153.0551 [M + H] ⁺	40
	Furan-2-carboxylic acid	C ₅ H ₄ O ₃	7.61 (dd, <i>J</i> = 1.8, 1.2); 7.06 (dd, <i>J</i> = 3.6, 0.6); 6.51 (dd, <i>J</i> = 3.6, 1.8)	112.0160	135.0413 [M + Na] ⁺ ; 111.0090 [M - H] ⁻	38

Thus, compounds in the culture fluid derived from *S. coelicolor* induced the production of secondary metabolites by *A. niger*, without the need for physical interaction between them. In contrast, direct contact was required for the biosynthesis of the aromatic polyketides orsellinic acid, lecanoric acid, F-9775A and F-9775B in a co-culture of *Aspergillus nidulans* and *Streptomyces hygroscopicus* ⁴³. We then investigated whether exogenously added compounds could also act as starter material for bioconversion to new compounds. For this, co-cultivation was done in the presence of one of a variety of hydroxycinnamic acids. Seven structurally related molecules, namely cinnamic acid, *o*-, *m*- and *p*- coumaric acid, caffeic acid, ferulic acid, and sinapinic acid, were fed as substrates to the co-culture and to each of the monocultures. After incubation, the cultures were extracted with ethyl acetate and subsequently subjected to ¹H NMR and UHPLC-TOF-MS analysis (positive and/or negative modes) without further chromatographic fractionation. All of the compounds obtained by bioconversion by the three types of cultures were unambiguously identified on the basis of spectral data (Table S1-S7) and summarized (Figure S3-S9). All structures could be readily elucidated based on spectra of known compounds, except for the novel compound **1** (see below). Under these culturing conditions, the majority of substrates were almost completely consumed by *A. niger*, whereas *S. coelicolor* exhibited a limited capacity to convert these substrates, yielding at best trace amounts of biotransformation products. However, although the biotransformation process in co-cultures was dominated by *A. niger*, both qualitative and quantitative differences were found between the mono- and co-cultures in terms of the biotransformation compounds. The addition of *o*-coumaric acid to the co-culture yielded

(*E*)-2-(3-hydroxyprop-1-en-1-yl)-phenol, which was not produced in either of the two monocultures (Figure 4).

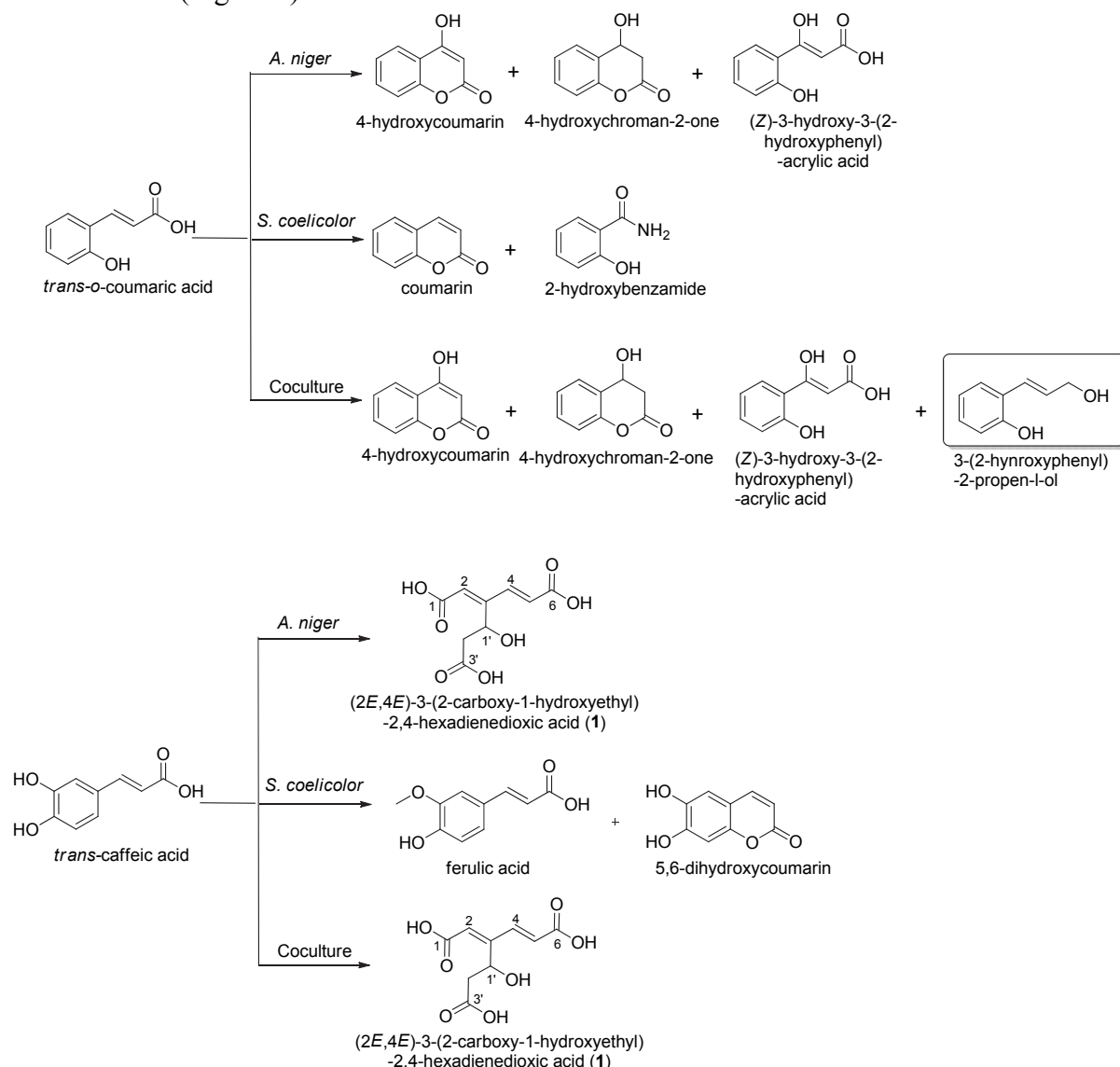


Figure 4. Biotransformation products of *o*-coumaric acid and caffeic acid by *S. coelicolor* and *A. niger* monocultures and their co-culture. The boxed compound (*E*)-2-(3-hydroxyprop-1-en-1-yl)-phenol was exclusively detected in co-culture. The structurally novel molecule (2*E*,4*E*)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid was derived from caffeic acid biotransformation by fungus, while conversion rate was around 3.5 times in co-culture than that of fungal monoculture.

Unexpectedly, caffeic acid was converted by *A. niger* in both mono- and co-culture into a previously undescribed compound (**1**). Its molecular formula, C₉H₁₀O₇, was determined using UHPLC-TOF-2Q high-resolution mass analysis with an ion at *m/z* 211.0254 [M – H₂O – H][–] (calculated mass for C₉H₈O₆ is 211.0254). The ¹H NMR spectrum exhibited a characteristic signal corresponding to a *trans*-substituted olefinic bond, δ_H 6.42 (d, *J* = 16.2 Hz), 7.47 (dt, *J* = 16.2, 0.6 Hz), which indicated the acrylic acid side chain of caffeic acid resisted the biodegradation. The disappearance of the typical ABX system proton signals for 1,2,4-trisubstituted benzene moiety implied that this aromatic ring had been opened. With the aid of 2D NMR experiments (¹H-¹H COSY, HSQC, HMBC, Figure S10-S14), the planar structure of compound **1** was unambiguously identified as (2*E*,4*E*)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid (Table 2) shown in Figure 1.

The formation of **1** was likely due to oxygenation of caffeic acid resulting in cleavage of the aromatic ring between two adjacent hydroxyl groups, thus producing two carboxyl groups at the site of ring opening⁴⁴. We proposed that the resulting product 4-carboxymethylene-2,5-heptadienoic acid⁴⁴ reacted with one molecule of water to generate the final adduct (2*E*,4*E*)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid (Figure S7). In view of the oxidative cleavage of the benzene ring, the geometry at the $\Delta^{2,3}$ double bond was established as *E*. Because the nucleophilic addition of olefin with H₂O tends to generate racemates, the stereochemistry at C-1' was assigned to be both *R* and *S*. The (2*E*,4*E*)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid bioconversion rate was accelerated when *A. niger* was co-cultivated with *S. coelicolor*, to around 3.5 times the rate of that occurring in the fungal monoculture.

Table 2. ¹H and ¹³C NMR data assignment for new compound (2*E*,4*E*)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid (**1**) in CD₃OD. Proton coupling constants (*J* in Hz) are given in parentheses. ¹H NMR and ¹³C NMR spectra were recorded at 600 MHz. All chemical shift assignments were done on the basis of 1D- and 2D -NMR techniques.

NO.	δ_c	δ_H (<i>J</i> in Hz)	HMBC (¹ H→ ¹³ C)	COSY (¹ H→ ¹ H)
1	172.4			
2	120.6	6.42 (d, <i>J</i> = 0.6)	C-1, C-4, C-1'	H-1'
3	162.5			
4	132.0	7.47 (dt, <i>J</i> = 16.2, 0.6)	C-1', C-2, C-6, C-5, C-3	H-5, H-1'
5	128.5	6.42 (d, <i>J</i> = 16.2)	C-6, C-3, C-4	H-4
6	167.1			
1'	79.0	5.68 (dddd, <i>J</i> = 7.2, 3.6, 1.8, 0.6)		H-2, H-4, H-2'
2'	37.6	3.07 (dd, <i>J</i> = 16.8, 3.6); 2.58 (dd, <i>J</i> = 16.8, 3.6)	C-3', C-3	H-1'
3'	171.1			

3. CONCLUSION

Co-cultivation of *S. coelicolor* and *A. niger* allowed the production of the compounds cyclo(Phe-Phe), 2-hydroxyacetic acid and phenylacetic acid by *A. niger* that were not produced by its monoculture. The fact that the bacterial cell-free filtrate was sufficient for cyclo(Phe-Phe) and phenylacetic acid production by *A. niger* proved that the change in fungal phenylalanine metabolism was induced by secreted compounds that act as starter molecules or as elicitors of silent biosynthetic pathways, rather than by cell-to-cell contact with or nutrient depletion by *S. coelicolor*. NMR-based metabolomics coupled to eliciting compounds by microbial co-cultivation is thereby a very effective strategy to dereplicate known molecules and effectively highlighted the compounds induced specifically by co-cultivation. In addition, chemical modification of hydroxycinnamic acids was achieved by *Aspergillus* and/or *Streptomyces* biotransformation, whereby only the co-culture specifically converted *o*-coumaric acid into (*E*)-2-(3-hydroxyprop-1-en-1-yl)-phenol. Importantly, besides novel metabolic activity, the biotransformation also allowed the discovery of the novel compound (2*E*,4*E*)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid, which was produced following addition of caffeic acid. Thus, the implementation of co-cultivation and biotransformation has the capacity to increase chemical diversity, which offers new leads for drug-discovery efforts.

4. EXPERIMENTAL SECTION

4.1. Bacterial strains and culturing conditions

Streptomyces coelicolor A3(2) M145⁴⁵ was obtained from the John Innes Centre strain collection, *Aspergillus niger* N402 (*cspA1*) is a derivative of *A. niger* ATCC9029. *A. niger* AR19#1²⁸ expresses eGFP under control of the *A. nidulans* *gpdA* promoter. NMMP⁴⁵ was used as liquid minimal media for all co-cultivation experiments. Growth curves of *S. coelicolor* M145 were set up in 50 ml NMMP with glucose (0.5% w/v) as described⁴⁶, with an inoculum of 10⁶ spores in 250 ml flasks equipped with a spring, and grown at 30 °C with constant shaking at 160 rpm. After 72 h, 10⁵ spores of *A. niger* N402 were added to the culture, and cultivation continued for another 72 h. In parallel, monocultures of *S. coelicolor* and *A. niger* were grown for 144 h and 72 h, respectively, corresponding to the cultivation time of each strain in the co-culture. Ten replicates of each experiment, namely *Aspergillus niger* monoculture, *Streptomyces coelicolor* monoculture, and the coculture (so 30 samples in total), were prepared for the metabolomics studies. Fluorescence and corresponding light micrographs were obtained with a Zeiss Axioscope A1 upright fluorescence microscope (with an AxioCam Mrc5 camera at a resolution of 37.5 nm/pixel) as described⁴⁷. The green fluorescent images were created using 470/40 nm band pass (bp) excitation and 525/50 bp detection.

4.2. Biotransformation of hydroxycinnamic acids

Seven hydroxycinnamic acid congeners (cinnamic acid, *o*-coumaric acid, *m*-coumaric acid, *p*-coumaric acid, caffeic acid, ferulic acid, and sinapinic acid) were selected as substrates for the biotransformation study. For this, 10 mg of each substrate dissolved in 200 µl DMSO were added to the co-culture, 24 h after addition of *A. niger* to the *S. coelicolor* culture. In parallel, the substrates were added to monocultures of *S. coelicolor* (after 96 h) and *A. niger* (after 24 h), corresponding to the cultivation time of each strain in the co-culture. The biotransformation products were analyzed 48 h after adding the substrates. All biotransformation experiments were conducted in triplicate.

4.3. Extraction of metabolites

Culture replicates were harvested by centrifugation at 4000 rpm for 10 min. The culture broth (50 ml) was extracted twice with 20 ml of ethyl acetate. The organic phase was washed with 30 ml of water and subsequently dried with 5 g of anhydrous Na₂SO₄. Finally, the EtOAc was removed under vacuum at 38 °C and the residue was dissolved in 2.0 ml of EtOAc in a microtube (Eppendorf type-5415C, Hamburg, Germany). The solvent was then evaporated at room temperature under nitrogen gas, and subsequently dipped into liquid nitrogen and lyophilized using a freeze dryer (Edwards Ltd., Crawley, England).

4.4. NMR measurements

NMR sample preparation and measurements were performed according to a protocol that was published previously²⁶. Briefly, 500 µl of methanol-*d*₄ were added to freeze-dried samples, and the resultant mixtures were vortexed for 10 sec and sonicated for 20 min at 42 kHz using an Ultrasonicator 5510E-MT (Branson, Danbury, CT, USA), followed by centrifugation at 16,000x g at room temperature for 5 min. The supernatant (300 µl) was transferred to a 3 mm

micro NMR tube and analyzed. The ^1H NMR spectra were recorded at 25 °C on a 600 MHz Bruker DMX-600 spectrometer (Bruker, Karlsruhe, Germany) operating at a proton NMR frequency of 600.13 MHz. Deuterated methanol was used as the internal lock. Each ^1H NMR spectrum consisted of 128 scans using the following parameters: 0.16 Hz/point, pulse width (PW) = 30 (11.3 ls) and relaxation delay (RD) = 1.5 s. Free induction decays (FIDs) were Fourier transformed with a line broadening (LB) = 0.3 Hz. The resulting spectra were manually phased and baseline corrected, and calibrated to methanol at 3.30 ppm, using XWIN NMR (version 3.5, Bruker).

4.5. UHPLC-TOF mass spectrometry

The UHPLC-TOF-MS analyses were performed on an Ultimate 3000 UHPLC system (ThermoScientific, USA) coupled to a micro-ToF-2Q mass spectrometer from Bruker Daltonics (Bremen, Germany) with an electrospray (ESI) interface⁴⁸. In separate runs, detection was achieved in both positive and negative ion modes. The m/z range was set to be 100–1000 and the ESI conditions were as follows: capillary voltage of 3500 V, source temperature of 250 °C, desolvation temperature of 250 °C, dry gas flow of 10.0 l/min and the nebulizer at 2.0 Bar. For positive ionization, all conditions were the same and the capillary voltage was 4000V. Internal calibration was performed using a 10 mM sodium formate solution from Sigma–Aldrich (Steinheim, Germany). Samples of 3 μl were injected onto a Kinetex C18, 150 x 2.0 mm column, packed with 1.7 μm particles (Phenomenex, USA), and eluted in gradient mode at a flow rate of 0.3 ml/min with the following solvent system: (A) 0.1 % formic acid (v/v) in water; (B) 0.1 % (v/v) formic acid in acetonitrile. Analysis began with a gradient of 5% to 90 % B to in 19.5 min, followed by an isocratic step of 90% B for 1 min and a re-equilibration of 1 minute with 5% B. The total run time was 21.5 min. The temperature was maintained at 30 °C.

4.6. Multivariate Data analysis

The ^1H NMR data files were processed as described^{26, 49}. The ^1H NMR spectra were converted to an ASCII file using AMIX software (Bruker Biospin GmbH), with total intensity scaling. Bucketing or binning was performed and the spectral data were reduced to include regions of equal width (0.04 ppm) equivalent to the region of δ 0.30–10.02. The regions of δ 4.85 – 4.95 and δ 3.25 – 3.35 were removed in the analysis because of the remnant signals of the solvents, HDO and CD₃OD, respectively. Principal component analysis (PCA) was performed with the SIMCA-P+ software (Version 13.0, Umetrics, Umeå, Sweden). Pareto-scaling method was used for PCA.

REFERENCES

- (1) Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J. *Nat Rev Drug Discov* **2015**, *14*, 111-129.
- (2) Hopwood, D. A., *Streptomyces in nature and medicine: the antibiotic makers*. Oxford University Press: New York, 2007.

- (3) Newman, D. J.; Cragg, G. M. *J Nat Prod* **2012**, *75*, 311-335.
- (4) Pimm, S. L.; Russell, G. J.; Gittleman, J. L.; Brooks, T. M. *Science* **1995**, *269*, 347-350.
- (5) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat Rev Drug Discov* **2007**, *6*, 29-40.
- (6) Bergmann, S.; Schumann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. *Nat Chem Biol* **2007**, *3*, 213-217.
- (7) Pel, H. J.; de Winde, J. H.; Archer, D. B.; Dyer, P. S.; Hofmann, G.; Schaap, P. J.; Turner, G.; de Vries, R. P.; Albang, R.; Albermann, K.; Andersen, M. R.; Bendtsen, J. D.; Benen, J. A.; van den Berg, M.; Breestraat, S.; Caddick, M. X.; Contreras, R.; Cornell, M.; Coutinho, P. M.; Danchin, E. G.; Debets, A. J.; Dekker, P.; van Dijck, P. W.; van Dijk, A.; Dijkhuizen, L.; Driessen, A. J.; d'Enfert, C.; Geysens, S.; Goosen, C.; Groot, G. S.; de Groot, P. W.; Guillemette, T.; Henrissat, B.; Herweijer, M.; van den Hombergh, J. P.; van den Hondel, C. A.; van der Heijden, R. T.; van der Kaaij, R. M.; Klis, F. M.; Kools, H. J.; Kubicek, C. P.; van Kuyk, P. A.; Lauber, J.; Lu, X.; van der Maarel, M. J.; Meulenberg, R.; Menke, H.; Mortimer, M. A.; Nielsen, J.; Oliver, S. G.; Olsthoorn, M.; Pal, K.; van Peij, N. N.; Ram, A. F.; Rinas, U.; Roubos, J. A.; Sagt, C. M.; Schmoll, M.; Sun, J.; Ussery, D.; Varga, J.; Vervecken, W.; van de Vondervoort, P. J.; Wedler, H.; Wosten, H. A.; Zeng, A. P.; van Ooyen, A. J.; Visser, J.; Stam, H. *Nat Biotechnol* **2007**, *25*, 221-231.
- (8) van den Berg, M. A.; Albang, R.; Albermann, K.; Badger, J. H.; Daran, J. M.; Driessen, A. J.; Garcia-Estrada, C.; Fedorova, N. D.; Harris, D. M.; Heijne, W. H.; Joardar, V.; Kiel, J. A.; Kovalchuk, A.; Martin, J. F.; Nierman, W. C.; Nijland, J. G.; Pronk, J. T.; Roubos, J. A.; van der Klei, I. J.; van Peij, N. N.; Veenhuis, M.; von Dohren, H.; Wagner, C.; Wortman, J.; Bovenberg, R. A. *Nat Biotechnol* **2008**, *26*, 1161-1168.
- (9) Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A. *Nature* **2002**, *417*, 141-147.
- (10) Cruz-Morales, P.; Vijgenboom, E.; Iruegas-Bocardo, F.; Girard, G.; Yanez-Guerra, L. A.; Ramos-Aboites, H. E.; Pernodet, J. L.; Anne, J.; van Wezel, G. P.; Barona-Gomez, F. *Genome Biol Evol* **2013**, *5*, 1165-1175.
- (11) Ohnishi, Y.; Ishikawa, J.; Hara, H.; Suzuki, H.; Ikenoya, M.; Ikeda, H.; Yamashita, A.; Hattori, M.; Horinouchi, S. *J Bacteriol* **2008**, *190*, 4050-4060.
- (12) van Wezel, G. P.; McKenzie, N. L.; Nodwell, J. R. *Methods Enzymol* **2009**, *458*, 117-141.
- (13) Zhu, H.; Sandiford, S. K.; van Wezel, G. P. *J Ind Microbiol Biotechnol* **2014**, *41*, 371-386.
- (14) Schumann, J.; Hertweck, C. *J Biotechnol* **2006**, *124*, 690-703.
- (15) Ochi, K.; Tanaka, Y.; Tojo, S. *J Ind Microbiol Biotechnol* **2014**, *41*, 403-414.
- (16) Thaker, M. N.; Wang, W.; Spanogiannopoulos, P.; Waglechner, N.; King, A. M.; Medina, R.; Wright, G. D. *Nat Biotechnol* **2013**, *31*, 922-927.

-
- (17) Craney, A.; Ozimok, C.; Pimentel-Elardo, S. M.; Capretta, A.; Nodwell, J. R. *Chem Biol* **2012**, *19*, 1020-1027.
- (18) Zhu, H.; Swierstra, J.; Wu, C.; Girard, G.; Choi, Y. H.; van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. *Microbiology* **2014**, *160*, 1714-1725.
- (19) Ling, L. L.; Schneider, T.; Peoples, A. J.; Spoering, A. L.; Engels, I.; Conlon, B. P.; Mueller, A.; Schaberle, T. F.; Hughes, D. E.; Epstein, S.; Jones, M.; Lazarides, L.; Steadman, V. A.; Cohen, D. R.; Felix, C. R.; Fetterman, K. A.; Millett, W. P.; Nitti, A. G.; Zullo, A. M.; Chen, C.; Lewis, K. *Nature* **2015**, *517*, 455-459.
- (20) Wilson, M. C.; Mori, T.; Ruckert, C.; Uria, A. R.; Helf, M. J.; Takada, K.; Gernert, C.; Steffens, U. A.; Heycke, N.; Schmitt, S.; Rinke, C.; Helfrich, E. J.; Brachmann, A. O.; Gurgui, C.; Wakimoto, T.; Kracht, M.; Crusemann, M.; Hentschel, U.; Abe, I.; Matsunaga, S.; Kalinowski, J.; Takeyama, H.; Piel, J. *Nature* **2014**, *506*, 58-62.
- (21) Angell, S.; Bench, B. J.; Williams, H.; Watanabe, C. M. *Chem Biol* **2006**, *13*, 1349-1359.
- (22) Oh, D. C.; Kauffman, C. A.; Jensen, P. R.; Fenical, W. *J Nat Prod* **2007**, *70*, 515-520.
- (23) Claessen, D.; Rozen, D. E.; Kuipers, O. P.; Sogaard-Andersen, L.; van Wezel, G. P. *Nat Rev Microbiol* **2014**, *12*, 115-124.
- (24) Benoit, I.; van den Esker, M. H.; Patyshakuliyeva, A.; Mattern, D. J.; Blei, F.; Zhou, M.; Dijksterhuis, J.; Brakhage, A. A.; Kuipers, O. P.; de Vries, R. P.; Kovacs, A. T. *Environ Microbiol* **2014**.
- (25) Konig, C. C.; Scherlach, K.; Schroeckh, V.; Horn, F.; Nietzsche, S.; Brakhage, A. A.; Hertweck, C. *Chembiochem* **2013**, *14*, 938-942.
- (26) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Nat Protoc* **2010**, *5*, 536-549.
- (27) Wu, C. S.; Kim, H. K.; van Wezel, G. P.; Choi, Y. K. *Drug Disc Today Technol.* **2015**, *In press*.
- (28) Vinck, A.; Terlouw, M.; Pestman, W. R.; Martens, E. P.; Ram, A. F.; van den Hondel, C. A.; Wosten, H. A. *Mol Microbiol* **2005**, *58*, 693-699.
- (29) Bertrand, S.; Bohni, N.; Schnee, S.; Schumpp, O.; Gindro, K.; Wolfender, J. L. *Biotechnol Adv* **2014**, *32*, 1180-1204.
- (30) Wright, L. F.; Hopwood, D. A. *J Gen Microbiol* **1976**, *96*, 289-297.
- (31) Rudd, B. A.; Hopwood, D. A. *J Gen Microbiol* **1980**, *119*, 333-340.
- (32) Zhang, Y.; Zhu, T.; Fang, Y.; Liu, H.; Gu, Q.; Zhu, W. *J Antibiot (Tokyo)* **2007**, *60*, 153-157.
- (33) Nielsen, K. F.; Mogensen, J. M.; Johansen, M.; Larsen, T. O.; Frisvad, J. C. *Anal Bioanal Chem* **2009**, *395*, 1225-1242.
- (34) Frisvad, J. C.; Smedsgaard, J.; Samson, R. A.; Larsen, T. O.; Thrane, U. *J Agric Food Chem* **2007**, *55*, 9727-9732.
- (35) Wang, F. W. *Chin J Nat Med* **2012**, *10*, 72-76.
- (36) Nishanth Kumar, S.; Dileep, C.; Mohandas, C.; B., N.; Jayaprakas, C. A. *J Peptide Sci* **2014**, *20*, 173-185.
- (37) Kishore, G.; Sugumaran, M.; Vaidyanathan, C. S. *J Bacteriol* **1976**, *128*, 182-191.
- (38) Cai, T.; Qi, W.; Yang, L.; Tu, G.; Yang, R.; Xie, K.; Fu, H. *J Chin Pharm Sci* **2012**, *21*, 428-435.
- (39) Wang, F.; Fang, Y.; Zhu, T.; Zhang, M.; Lin, A.; Gu, Q.; Zu, W. *Tetrahedron* **2008**, *64*,

- 7986–7991.
- (40) Faulkner, J. K.; Woodcock, D. *Phytochem* **1968**, *7*, 1741-1742.
 - (41) Guo, C. J.; Yeh, H. H.; Chiang, Y. M.; Sanchez, J. F.; Chang, S. L.; Bruno, K. S.; Wang, C. C. *J Am Chem Soc* **2013**, *135*, 7205-7213.
 - (42) Sauguet, L.; Moutiez, M.; Li, Y.; Belin, P.; Seguin, J.; Le Du, M. H.; Thai, R.; Masson, C.; Fonvielle, M.; Pernodet, J. L.; Charbonnier, J. B.; Gondry, M. *Nucleic Acids Res* **2011**, *39*, 4475-4489.
 - (43) Schroeckh, V.; Scherlach, K.; Nutzmann, H. W.; Shelest, E.; Schmidt-Heck, W.; Schuemann, J.; Martin, K.; Hertweck, C.; Brakhage, A. A. *Proc Natl Acad Sci U S A* **2009**, *106*, 14558-14563.
 - (44) Finkle, B. J.; Nozaki, M.; Fujisawa, H. *Phytochemistry* **1971**, *10*, 235-242.
 - (45) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A., *Practical Streptomyces genetics*. John Innes Foundation: Norwich, U.K., 2000.
 - (46) van Wezel, G. P.; Konig, M.; Mahr, K.; Nothhaft, H.; Thomae, A. W.; Bibb, M.; Titgemeyer, F. *J Mol Microbiol Biotechnol* **2007**, *12*, 67-74.
 - (47) Colson, S.; van Wezel, G. P.; Craig, M.; Noens, E. E.; Nothhaft, H.; Mommaas, A. M.; Titgemeyer, F.; Joris, B.; Rigali, S. *Microbiology* **2008**, *154*, 373-382.
 - (48) Li, N.; Kuo, C. L.; Paniagua, G.; van den Elst, H.; Verdoes, M.; Willems, L. I.; van der Linden, W. A.; Ruben, M.; van Genderen, E.; Gubbens, J.; van Wezel, G. P.; Overkleeft, H. S.; Florea, B. I. *Nat Protoc* **2013**, *8*, 1155-1168.
 - (49) Wu, C. S.; van Wezel, G. P.; Choi, Y. *J Antibiotics (Tokyo)* **2015**, *68*, 445–452.

Chapter 6

Metabolomics-guided analysis of isocoumarin production by *Streptomyces* species MBT76 and biotransformation of flavonoids and phenylpropanoids

Changsheng Wu ^{1,2}, Hua Zhu ¹, Gilles P. van Wezel ¹, and Young Hae Choi ²

¹ *Molecular Biotechnology, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands*

² *Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands*

Metabolomics **2016**, 12:90.

ABSTRACT

Actinomycetes produce the majority of the antibiotics currently in clinical use. The efficiency and timing of antibiotic production is affected by factors such as nutrients, pH, temperature and growth phase. Finding the optimal harvesting time is crucial for successful isolation of the desired bioactive metabolites, but for this conventional chemical analysis has limitations due to the metabolic complexity. In this study, the growth phase-dependence of secondary metabolites production by *Streptomyces* sp. MBT76 was analyzed using NMR-based metabolomics. Isocoumarins (**1—9**), undecylprodiginine (**10**), streptorubin B (**11**), 1*H*-pyrrole-2-carboxamide (**12**), acetyltryptamine (**13**), and fervenulin (**14**) were identified and their optimal production time was determined in crude extracts, avoiding time-consuming fractionation. Of these compounds, 5,6,7,8-tetramethoxyl-3-methyl-isocoumarin (**9**) is as a novel compound, which was most likely synthesized by a type I iterative polyketide synthase (PKS) encoded by the *icm* gene cluster. Multivariate data analysis of the ¹H NMR spectra showed that acetyltryptamine (**13**) and *tri*-methoxylated isocoumarins (**7** and **8**) were the major determinants of antibiotic activity during later time points. The methoxylation was exploited to allow bioconversion of exogenously added genistein into a suite of methoxylated isoflavones (**15—18**). Methoxylation increased the antimicrobial efficacy of isocoumarins, but decreased that of the isoflavones. Our results show the applicability of NMR-based metabolic profiling to streamline microbial biotransformation and to determine the optimal harvesting time of actinomycetes for antibiotic production.

1. INTRODUCTION

The alarming increase of pathogen resistance underlines the urgent need for new antibiotics.¹ As producers of two thirds of all known antibiotics and of many other bioactive compounds, actinomycetes are a rich source of clinical drugs, with the majority produced by members of the genus *Streptomyces*.^{2,3} A problem is the high frequency of re-discovery of known compounds that has frustrated high-throughput (HT) screening regimes in the past decennia.^{4,5} Sequencing of the genomes of actinomycetes revealed that the potential of even the best-studied model organisms as producers of natural products have been grossly underestimated.^{6–8} However, many antibiotic biosynthetic gene clusters are poorly expressed under routine laboratory conditions. A new drug-discovery pipeline is rapidly gaining momentum, which is combining fluctuation of the (cryptic) antibiotics production with metabolic profiling-based identification of the bioactivity of interest.^{9–11}

In actinomycetes, the onset of antibiotic production typically coincides with the transition from vegetative to aerial growth in solid-grown cultures, which corresponds roughly to the onset of stationary phase in liquid-grown cultures.^{12,13} It is therefore very important to closely monitor the metabolic changes during the entire fermentation time, so as to optimize the fermentation for production of the bioactive molecules of interest. Conventional analytical methods, *e.g.* HPLC-UV, are often used to monitor the production of molecules but fail to provide structural information and the detection of analytes tends to be demanding on chromophores. To capture all the molecular features in a microbial mixture without prior information, techniques with a broader analytical capacity should be employed. Moreover, to take a snapshot of multiple cellular processes at for example any given time-point of a growth curve of a producing microorganism, a robust analytical platform is needed, which should be rapid, reproducible, stable in time, and simple in sample preparation.¹⁴ Proton nuclear magnetic resonance (¹H NMR) spectroscopy provides both qualitative (chemical shift and splitting pattern) and quantifiable (integrative area) peaks for structure elucidation (sometimes aided by 2D NMR), allowing tracing of metabolic changes over time by visual comparison of the metabolomes harvested at several time-points.¹⁵ Subtle metabolic differences could be further highlighted by unbiased statistical filtering of the ¹H NMR spectra, such as chemometric methods, principle component analysis (PCA), partial least squares discriminant analysis (PLS-DA), *etcetera*.

Actinomycetes do not only producing a plethora of natural products, but they can also efficiently perform metabolic conversions on molecules with a range of physicochemical properties.¹⁶ Deciphering the metabolic tailoring reactions of natural products produced by a certain producing strain over time, such as glycosylation, hydroxylation, or *O*-methylation or prenylation,¹⁷ will provide insights into the optimal time of providing starter molecules in biotransformation experiments, since many enzymes show significant promiscuity in terms of substrate specificity.¹⁸

In this study, NMR-based metabolomics was used to determine the variation in metabolites produced by *Streptomyces* sp. MBT76 at five different time-points and correlated these to the observed bioactivity using multivariate data analysis (MVDA). Aided by UHPLC-ToF-MS analysis, various bioactive molecules (**1–14**) including a new *tetra*-methoxylated isocoumarin (**9**), were thus identified in the culture broth. Biotransformation of the isoflavone genistein by MBT76 led to a group of methoxylated

derivatives **15**—**18**. Methoxylation increased the efficacy of isocoumarins against both Gram-positive and Gram-negative bacteria, while, conversely, methoxylation decreased the antimicrobial efficacy of isoflavones.

2. RESULTS AND DISCUSSION

2.1. Correlated growth phase with chemotype

Streptomyces sp. MBT76 was previously identified as a prolific producer of antibiotics, including two highly methoxylated isocoumarins.¹⁹ To assess the correlation between growth and antibiotic production, MBT76 was grown in the liquid minimal medium (NMMP) supplemented with 0.8% (w/v) peptone, which has a positive effect on the strain's antibiotic producing capacity.¹⁹ Samples for the determination of biomass (dry weight) and antibiotic production (measured as zones of inhibition against indicator strains) were collected at 24 h time intervals until 120 h. A long lag phase of growth was observed that covered the first 24 h, after which the strain grew exponentially from 24—50 h. After around 80 h, the culture entered a phase where biomass decreased rapidly, most likely due to nutrient exhaustion (Figure S1, A). Antibiotic activity against *Bacillus subtilis* 168 was not seen before 48 h (Figure S1, B), and the onset of antibiotic production was seen after around 60 h of growth. The largest inhibition zone (25 mm) was found for supernatants obtained after 96 h, when cultures had already entered the death phase.

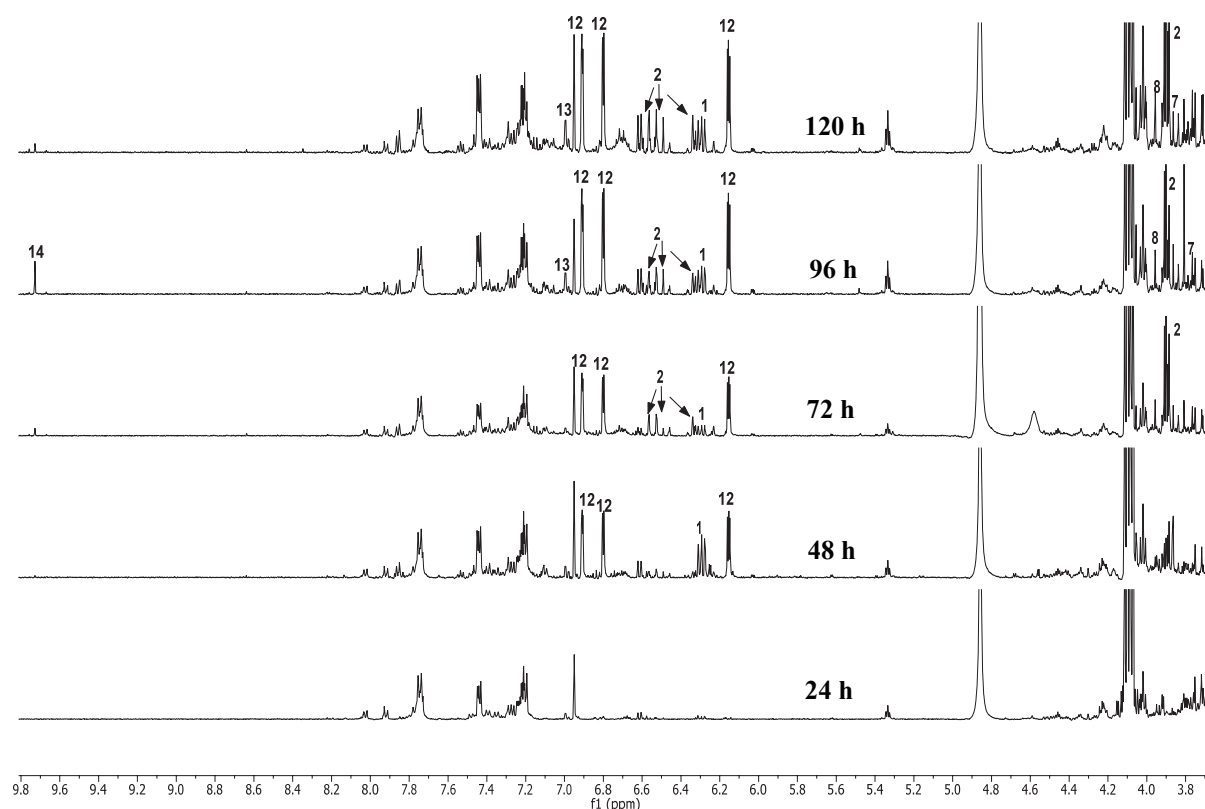


Figure 1. NMR profiling of *Streptomyces* sp. MBT76 harvested at 24-hour intervals. The ¹H NMR spectra varied significantly in aromatic region δ 6.0—8.0, and methoxyls region δ 3.45—4.25, with more complex proton signals as time progressed. Assigned signals are numbered as summarized in Table 1.

NMR-based metabolomics was employed to correlate bacterial growth and antimicrobial activity with the corresponding metabolic changes over time. At five time-points, namely after 24 h, 48 h, 72 h, 96 h, and 120 h of inoculation, the biomass of MBT76 was removed by centrifugation and the compounds present in the spent media were extracted with ethyl acetate. The obtained metabolomes were subjected to ^1H NMR profiling and antimicrobial assay against *B. subtilis* 168. The ^1H NMR spectra varied significantly in aromatic region δ 6.0–8.0, and methoxyls region δ 3.45–4.25, with more complex proton signals as time progressed (Figure 1). Low levels of secondary metabolites were produced before 48 h, while 72 h appeared to be a turning point in terms of secondary metabolites production. 96 h samples exhibited the most complex metabolome as compared to the other groups, coinciding with the highest antimicrobial activity against *B. subtilis*. The compounds in the crude extracts were elucidated with ^1H NMR spectra and UPLC-ToF-MS analysis (retention time, UV spectrum, and high resolution mass). The metabolites included methoxylated isocoumarins (**1–9**);¹⁹ red-pigmented undecylprodiginine (**10**) and streptorubin B (**11**);²⁰ 1*H*-pyrrole-2-carboxamide (**12**);²¹ acetyltryptamine (**13**);²² and fervenulin (**14**) (Figure 2 and Table 1).²³ Among these, the highly methoxylated compound 5,6,7,8-tetramethoxy-3-methyl-isocoumarin (**9**) has not been previously described. This novel molecule was determined based on the similarity of its UV spectrum with those of congeners **1–8** and a high resolution mass at m/z $[\text{M} + \text{H}]^+$ of 281.1039, calculated molecular formula $\text{C}_{14}\text{H}_{17}\text{O}_6$ (Figure S2).

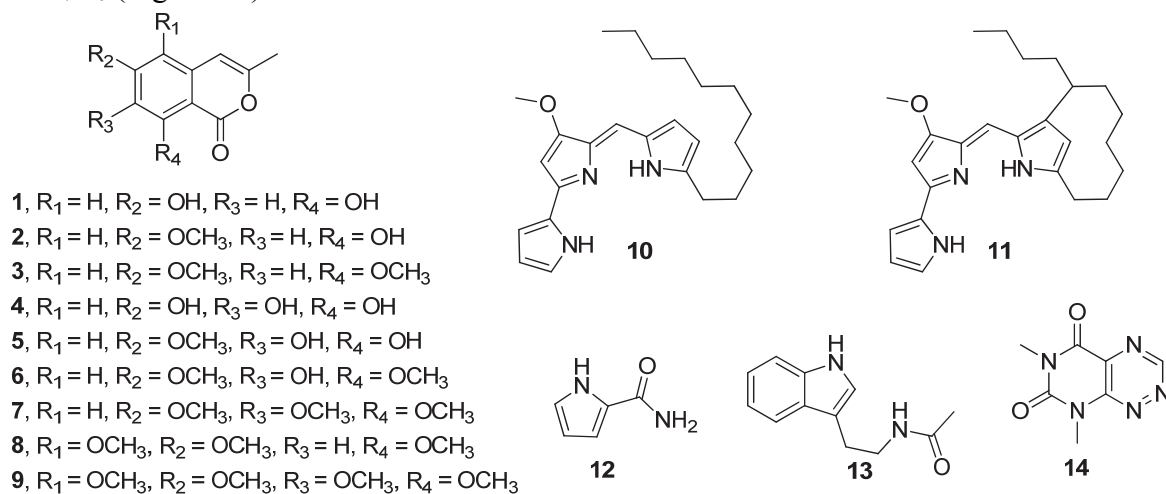


Figure 2. Secondary metabolites produced by *Streptomyces* sp. MBT76.

A plausible biosynthetic pathway (Figure 3) was proposed for isocoumarins **1–9**. Although isocoumarin polyketides are frequently reported in actinomycetes,^{24,25} the underlying genetic basis is poorly understood. In view of the biosynthesis of the isocoumarin backbone in fungi,^{26–29} along with dihydroisocoumarin assembly in the actinomycete *Saccharopolyspora erythraea*³⁰ and in sponge-associated bacteria,³¹ we tentatively propose that isocoumarin biosynthesis in *Streptomyces* sp. MBT76 is executed by a Type I iterative modular polyketide synthase (PKS). Indeed, antiSMASH³² analysis of the genome of *Streptomyces* sp. MBT76 identified a gene cluster (*icm*, Table S1) that encodes the candidate PKS components. This contains the *icmM* gene that encodes multiple catalytic domains including ketosynthase (KS), acyl transferase (AT), dehydrase (DH), acyl carrier protein

(ACP), which is likely responsible for the assembly of the precursor 6,8-dihydroxy-3-methyl-isocoumarin (**1**). This low-order isocoumarin could be further tailored by the oxygenase encoded by *icmF*, and an *O*-methyltransferase encoded by *icmI*.

The coexistence of compounds **7**, **8**, and **9** confirmed all the non-bridged carbons (positions 5, 6, 7, and 8) in benzene ring of isocoumarin could be methoxylated. This suggests that MBT76 could produce more new isocoumarin-type intermediates than those that have been elucidated here, like **4a**, **4b**, **6a**, and **6b**. Because both C-5 and C-7 of **1** could be hydroxylated, the substituted pattern (5,6,8-trihydroxylation or 6,7,8-trihydroxylation) in compounds **5** and **6** remained unclear. Some ambiguity remained regarding the *mono*- and/or *di*-methylation patterns in compounds **2**, **5**, and **6**, as all four —OH groups on ring A could be methylated. Further up-scale fermentation followed by extensive phytochemical isolation is needed to clarify ambiguities in structures of **2**, **5**, and **6**, and to unveil other new derivatives of isocoumarin family.

Table 1. Spectral data assignments for the compounds displayed in Figure 2. Compound identification was based on ^1H NMR, high resolution mass spectrometry, and UV absorption spectrum. Proton coupling constants (J in Hz) are given in parentheses.

NO	Compounds	Characteristic ^1H NMR chemical shifts (ppm)	Experimental mass (Da)	UV λ_{max} (nm)
1	6,8-dihydroxy-3-methyl-isocoumarin	6.29 (d, $J = 1.2$, 1H); 6.24 (d, $J = 2.4$, 1H); 6.27 (d, $J = 2.4$, 1H); 2.21 (d, $J = 1.2$, 3H)	193.0511 [M + H] $^+$; 215.0312 [M + Na] $^+$	244, 328
2	6-methoxyl-8-hydroxy-3-methyl-isocoumarin		207.0662 [M + H] $^+$; 229.0478 [M + Na] $^+$; 435.1053 [2M + Na] $^+$	244, 326
3	6,8-dimethoxyl-3-methyl-isocoumarin	6.29 (d, $J = 1.2$, 1H); 6.47 (d, $J = 2.4$, 1H); 6.52 (d, $J = 2.4$, 1H); 2.23 (d, $J = 1.2$, 3H); 3.89 (s, 3H); 3.90 (s, 3H)	221.0830 [M + H] $^+$; 243.0646 [M + Na] $^+$	244, 328
4	6,7,8-trihydroxy-3-methyl-isocoumarin		209.0441 [M + H] $^+$	244, 328
5	6-methoxyl-7,8-dihydroxy-3-methyl-isocoumarin	6.57 (d, $J = 1.2$, 1H); 6.38 (s, 1H); 3.75 (s, 3H); 2.26 (s, 3H)	223.0612 [M + H] $^+$; 245.0424 [M + Na] $^+$	244, 340
6	6,8-dimethoxy-7-hydroxy-3-methyl-isocoumarin		237.0779 [M + H] $^+$; 259.0593 [M + Na] $^+$; 495.1256 [2M + Na] $^+$	240, 340
7	6,7,8-trimethoxy-3-methyl-isocoumarin	6.35 (d, $J = 1.2$, 1H); 6.81 (s, 1H); 2.22 (d, $J = 1.2$); 3.95 (s, 3H); 3.83 (s, 3H); 3.90 (s, 3H)	251.0935 [M + H] $^+$; 273.0756 [M + Na] $^+$; 523.1596 [2M + Na] $^+$	246, 336
8	5,6,8-trimethoxy-3-methyl-isocoumarin	6.55 (d, $J = 1.2$, 1H); 6.70 (s, 1H); 2.20 (d, $J = 1.2$, 3H); 3.76 (s, 3H); 4.00 (s); 3.95 (s, 3H)	251.0935 [M + H] $^+$; 273.0756 [M + Na] $^+$; 523.1596 [2M + Na] $^+$	246, 336
9	5,6,7,8-tetramethoxy-3-methyl-isocoumarin		281.1039 [M + H] $^+$; 303.0861 [M + Na] $^+$; 583.1804 [2M + Na] $^+$	240, 337
10	undecylprodiginine	7.23 (s, 1H); 6.56 (s, 1H); 6.46 (dd, $J = 4.8$, 3.0, 1H); 6.39 (d, $J = 4.8$, 1H); 4.10 (s, 3H); 2.35 (t, $J = 7.2$, 2H)	394.2857 [M + H] $^+$	
11	streptorubin B	6.73 (s, 1H); 6.53 (s, 1H); 6.44 (dd, $J = 4.8$, 3.0, 1H); 6.41 (d, $J = 4.8$, 1H); 4.11 (s, 3H)	392.2692 [M + H] $^+$	
12	1H-pyrrole-2-carboxamide	6.92 (dd, $J = 2.4$, 1.2, 1H); 6.81 (dd, $J = 3.6$, 1.2, 1H); 6.17 (t, $J = 3.6$, 2.4, 1H)	111.0562 [M + H] $^+$; 133.0372 [M + Na] $^+$	268
13	acetyltryptamine	7.54 (dd, $J = 7.8$, 1.2, 1H); 7.31 (brd, $J = 8.4$, 1H); 7.70 (td, $J = 7.8$, 1.2, 1H); 6.99 (td, $J = 7.8$, 1.2, 1H); 7.05 (s, 1H); 3.45 (t, $J = 7.2$, 2H); 2.56 (t, $J = 7.2$, 2H); 1.90 (s, 3H)	203.1211 [M + H] $^+$; 225.1011 [M + Na] $^+$	224, 282
14	fervenulin	9.72 (s, 1H); 3.80 (s, 3H); 3.47 (s, 3H);	194.0705 [M + H] $^+$; 216.0482 [M + Na] $^+$	240, 344

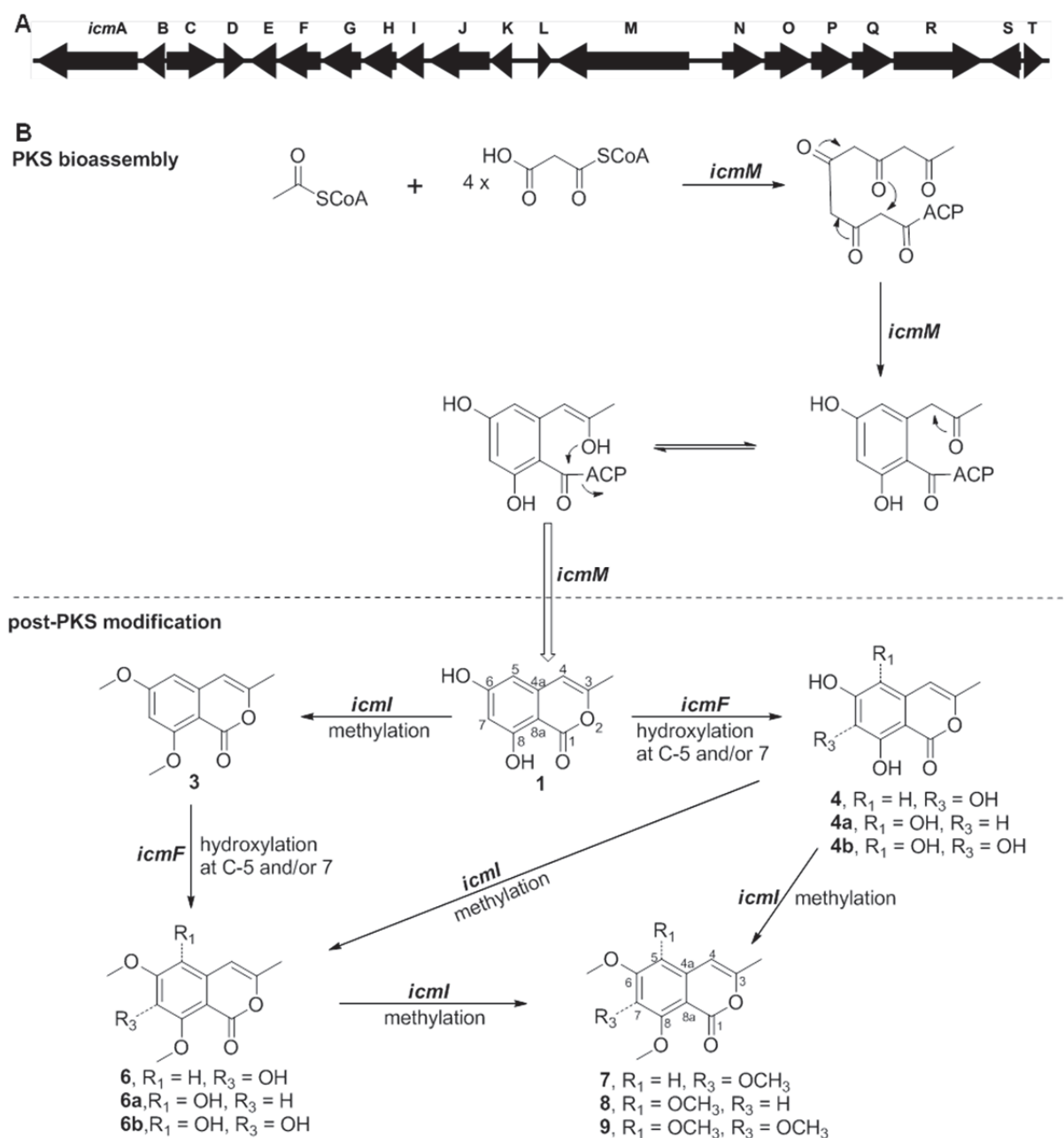


Figure 3. Biosynthetic pathway of isocoumarins produced by *Streptomyces* sp. MBT76. (A) Organization of the type I iterative PKS gene cluster (*icm*) in *Streptomyces* sp. MBT76 (Table S1). (B) Proposed biosynthetic route to highly methoxylated isocoumarins **7–9**. The modular non-reducing PKS gene *icmM* was iteratively used to assemble the basic backbone 6,8-dihydroxy-3-methyl-isocoumarin (**1**), which was further furnished into a range of methoxylated isocoumarins by the oxidoreductase *icmF* and the methyltransferase *icmI* in the post-PKS stage.

Multivariate data analysis (MVDA) was utilized to correlate the antimicrobial activity with the responsible signals in the ¹H NMR spectra. An unsupervised MVDA, principal component analysis (PCA) was first used to discriminate samples, but in this way only the non-active 24 h group was separated from other time-points (data not shown). As a next step to correlate metabolites detected in the ¹H NMR spectra with the bioactivity, orthogonal projection to latent structures (OPLS) was applied to the data using Y-variable of

antimicrobial activity (inhibition percentage), allowing the removal of spectral X-variables unrelated to the sample classes chosen and filtration of the biomarkers from the biological varieties.³³ The regression has been validated using cross validation analysis of variance (CV-ANOVA), with a P -value <0.001 . Separation of all the harvesting times could be observed in the OPLS score plot, with the active group (72 h, 96 h, and 120 h) clustered in the positive side along the OPLS1 axis (Figure 4, A). The S -plot presented the signals contributable to the discrimination among the samples with diversified antibacterial activities. The major signals responsible for the bioactivity were assigned as 6,7,8-trimethoxyl-3-methyl-isocoumarin (**7**), 5,7,8-trimethoxyl-3-methyl-isocoumarin (**8**), and acetyltryptamine (**13**) (Figures 4, B and C). The antibiotic activity may at least in part be attributed to acetyltryptamine (**13**), which was previously reported to have broad-spectrum antibiotic properties.²²

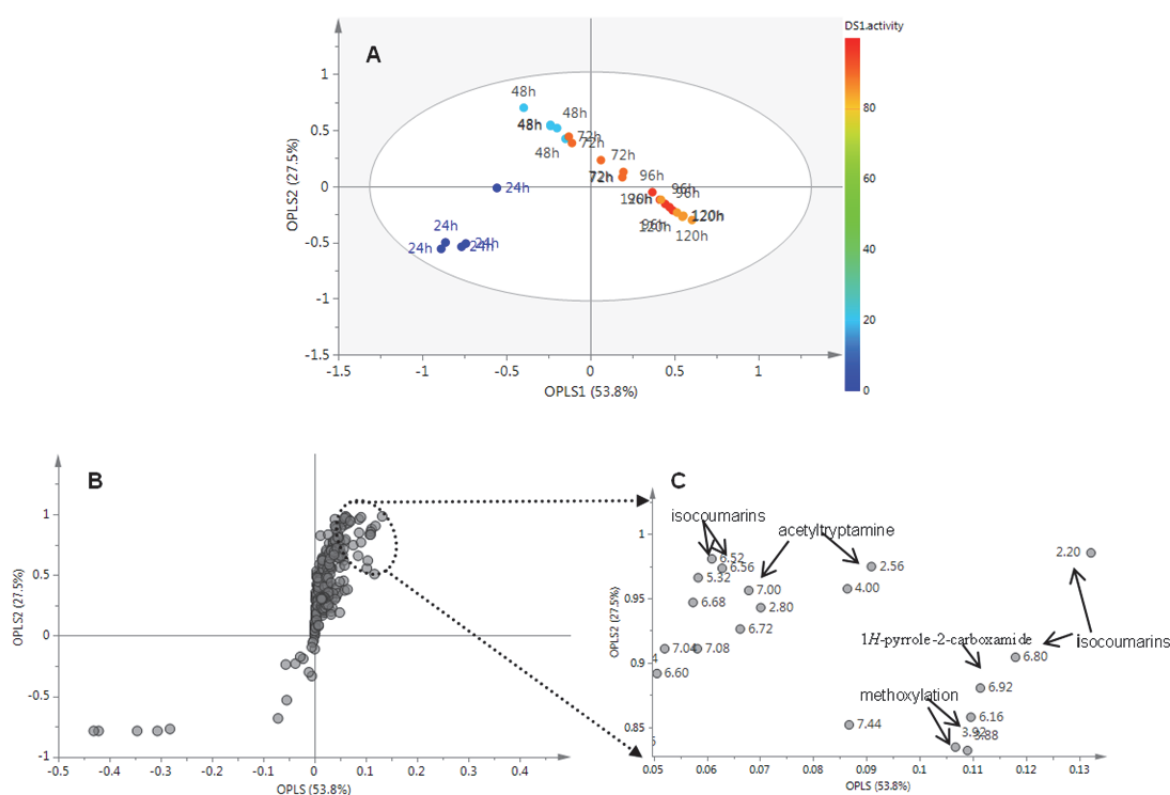


Figure 4. Supervised multivariate data analysis of the NMR spectra. (A) orthogonal projection to latent structures (OPLS) score plot showed separation of different time-points along OPLS1, and (B) the corresponding S -plot highlighted the compounds responsible for better activity in latter time points (96 h, and 120 h). (C) These compounds were identified as acetyltryptamine (**13**), and methoxylated isocoumarins (**7** and **8**) as annotated. Y value in OPLS analysis was antimicrobial bioactivity against *Bacillus subtilis* 168.

To test whether methoxylation helped increase the antimicrobial activity of isocoumarins, we isolated four major isocoumarins (**1**, **3**, **7**, and **8**) with low and/or high level of methoxylation. Comparison of the antimicrobial activity of these compounds was performed by agar diffusion assay, against the Gram-positive *B. subtilis* 168 and the Gram-negative *E. coli* K12 (Figure S3). The *tri*-methoxylated isocoumarins **7** and **8** exhibited pronounced inhibition of growth of *B. subtilis*, while compounds **1** and **3** only gave negligible inhibition zones. The same trend was also observed for inhibition against *E. coli*, though the bioactivity

of the compounds against this strain was much lower than against *B. subtilis*. This result validated the OPLS analysis that attributed *tri*-methoxylated isocoumarins **7** and **8** rather than non-methoxylated precursor **1** as the discriminators for the higher bioactivity at 96 h and 120 h. Based on this observation, we anticipated that the new compound 5,6,7,8-tetramethoxyl-3-methyl-isocoumarin (**9**) should have even better antimicrobial property than **7** and **8**, but its purification failed due to intrinsic low abundance (Figure S2). To correlate harvesting time to specific metabolites, we used time as the Y-variable in the OPLS analysis, resulting in a similar separation pattern of the five time-points along OPLS1 (Figure S4, A). The major discriminators for 96 h and 120 h groups were the same to those for activity as Y-variable OPLS analysis (Figure S4, B). Taken together, the dynamic ¹H NMR profiling and MVDA demonstrated the time-dependent biosynthesis of antibiotics (**1**–**14**) produced by *Streptomyces* sp. MBT76. This strain began to produce precursor 6,8-dihydroxy-3-methyl-isocoumarin (**1**) at 48 h, whose methoxylation into **2**–**9** at all the non-bridged carbons of the benzene ring occurred after 72 h, whereby the highly methoxylated isocoumarins **7** and **8** peaked at 96 h. Prodiginines (**10** and **11**) were observed throughout growth, which was reflected by the abundant by-product 1*H*-pyrrole-2-carboxamide (**12**) initiated at 48 h and the red appearance of the culture broth.²⁰ The antibiotics acetyltryptamine (**13**)

2.2. Metabolomics-driven biotransformation of phenylpropanoids

Inspired by the post-PKS methoxylation of 6,8-dihydroxy-3-methyl-isocoumarin (**1**, Figure 3), we then wanted to see what would happen in a biotransformation experiment whereby the isoflavone genistein was fed to cultures of *Streptomyces* sp. MBT76. Genistein and other isoflavones contain a similar fused benzopyranone moiety as isocoumarins and the same *meta*-hydroxylation on ring A. In accordance with the dynamic metabolism of isocoumarins as shown by the NMR-based metabolomics, genistein was fed to MBT76 culture broth at 48 h, *i.e.* in the ‘late exponential’ growth phase, and incubation was allowed to continue for another 72 h. The ethyl acetate extract was then subjected to NMR and TLC analysis and compared to control samples. To identify genistein-derived biotransformation products, we performed NMR-guided chromatographic separation³⁴ by tracking the downfield aromatic singlet (H-2 around δ 8.0) that is characteristic of isoflavones in the ¹H NMR spectrum. The purified products were subsequently subjected to NMR for structure elucidation. The ¹H NMR (methanol-*d*₄, 600 MHz) spectrum of compound **16** presented two methoxyls, besides two *meta*-coupled aromatic doublets, which allowed identification of one of the genistein-derived biotransformation products as 4'-hydroxy-5,7-dimethoxy-isoflavone (**16**).³⁵ In comparison, the major discrepancy in the ¹H NMR spectrum of compound **15** was a single *O*-CH₃, which indicated that a *mono*-methylation had occurred in the ring A. When DMSO-*d*₆ was used as solvent, the downfield-shifted exchangeable hydroxyl δ 11.5–13.5 that originates from the intramolecular hydrogen bonding with ketone (C-4) had disappeared, which was indicative of methylation at the 5-OH. Thus, compound **15** was identified as 4',7-dihydroxy-5-methoxy-isoflavone. The ¹H NMR spectrum of compound **17** identified three methoxyls at δ 3.94, 4.02, and 3.85, together with one uncoupled singlet at δ 6.68, which indicated *tri*-methoxylation at ring A. The possibility of methylation on ring B (4'-OH) was ruled out by HMBC and HSQC NMR. Furthermore, the two correlations H-6 / 5-OCH₃ and

H-6 / 7-OCH₃ in the ¹H-¹H NOESY spectrum confirmed that compound **17** was 4'-hydroxy-5,7,8-trimethoxy-isoflavone. Compound **18** shared the same substitution mode with **17** in ring A because of one uncoupled singlet at δ 6.48, but it presented two methoxyls in the ¹H NMR spectrum. One methyl was located at 8-OH based on the HMBC correlation from 8-OCH₃ (δ 3.86) to C-8 (δ_C 134.4). The other methyl was located at the 5-OH rather than at the 7-OH because no signal was observed in the range of δ 11.5—13.5 when measured in DMSO-*d*₆. Taken together, genistein underwent exactly the same methylation pattern of isocoumarin **1**, as summarized in Figure 5. In consideration of the structural similarity between isocoumarin **1** and genistein, it is conceivable that the same *O*-methyltransferase (*icmI*) is involved in the methylation of both substrates. Interestingly, this *O*-methyltransferase methylated isoflavone at all the three hydroxyl groups on ring A, but not at the single *p*-hydroxyl group on ring B.

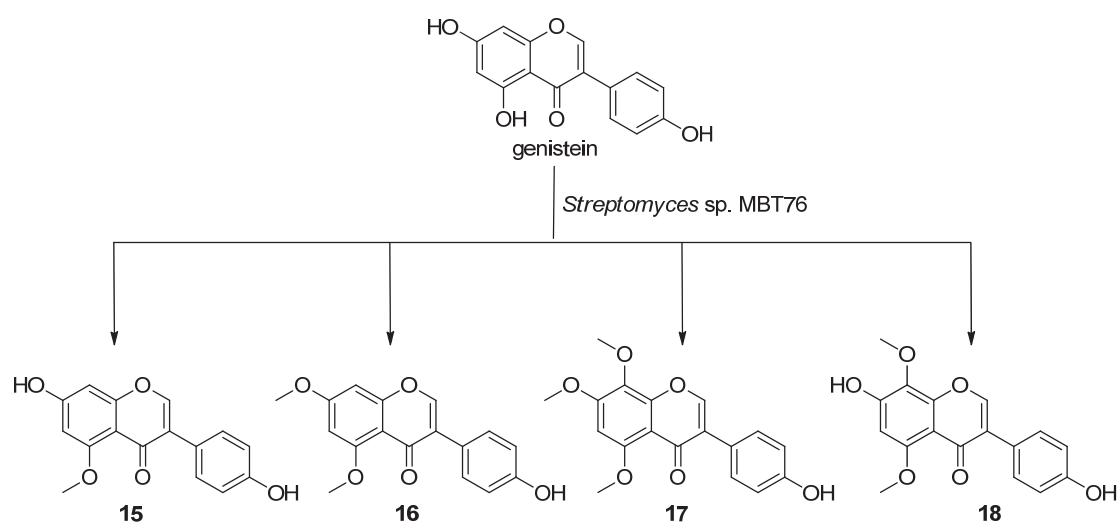


Figure 5. Biotransformation of genistein by *Streptomyces* sp. MBT76. Products were identified on the basis of NMR and/or UPLC-ToF-MS high resolution mass.

Motivated by the biotransformation of genistein, we also investigated whether the structurally related flavone apigenin and the dihydroflavonol naringenin could also be methoxylated. Unexpectedly, both apigenin and naringenin resisted biotransformation by MBT76, as evidenced by TLC and NMR analysis. The linkage pattern of the benzene ring B on chromone (at either C-2 or C-3) might influence the efficiency of methylation. From the perspective of biosynthesis it is interesting to note that in plants, ring B of flavonoids originates from cinnamic acid.³⁶ Therefore, we further tested whether hydroxycinnamic acids could be methylated. Six commercially available congeners, including cinnamic acid, *o*, *m*, *p*-coumaric acid, caffeic acid, and ferulic acid, were fed to the *Streptomyces* sp. MBT76. The corresponding biotransformation products were identified in mixture through ¹H NMR profiling and confirmed by UPLC-ToF-MS analysis. Amidation at the carboxyl group was the exclusive modification for all the (hydroxyl)cinnamic acids, with the exception of ferulic acid that could also be methylated at the 4-OH group (Figure S5). Unlike the situation of isocoumarins, where activity increases on methylation, methylated isoflavones had lower bioactivity than the unmethylated precursors. As shown in Figure S3, genistein showed highest antibiotic activity against both *B. subtilis* and *E. coli*, while *tri*-methoxylated

biotransformation product **19** failed to inhibit growth of the indicator strains. This observation was in accordance with previous results that the free phenolic OH groups of flavonoids are crucial for potent antimicrobial activity against Gram-positive and Gram-negative bacteria.³⁷

The *in vitro* biochemical reaction may in some cases be more advantageous than *in vivo* biotransformation for functional characterization of enzymes, *e.g.* due to limited uptake of the substrate, or because substrates may have an effect on overall gene expression of the various biosynthetic pathways. While the expected *in vivo* methylation of apigenin, naringenin, and most (hydroxyl)cinnamic acids was not observed, the potential promiscuity of methyltransferase IcmI suggested by post-PKS modification of isocoumarin (Figure 3) and biotransformation of genistein (Figure 5) is intriguing and may allow the methylation of a broader range of chemical skeletons. This is currently being investigated in our laboratory.

3. CONCLUSION

The growth phase had a major influence on the production of secondary metabolites in *Streptomyces* sp. MBT76. Four days of incubation was optimal to accumulate antimicrobial compounds identified as 6,7,8-trimethoxyl-3-methyl-isocoumarin (**7**), 5,7,8-trimethoxyl-3-methyl-isocoumarin (**8**), 5,6,7,8-tetramethoxyl-3-methyl-isocoumarin (**9**), and acetyltryptamine (**13**). However, production of 6,8-dihydroxyl-3-methyl-isocoumarin (**1**) was optimal after two days of incubation. NMR spectroscopy in conjugation with multivariate data analysis effectively identified specialized metabolites in the crude extract of MBT76, but also highlighted differences between the metabolomes at the different time-points. The methodology applied in this work should be useful for scientists who seek to identify the optimal fermentation time to harvest desirable bioactive compounds from cultures of in principle any microorganism that produces promising bioactive molecules, such as those with potential industrial or medical application.

4. EXPERIMENTAL SECTION

4.1. Culture media and condition

Streptomyces sp. MBT76 was obtained from the culture collection of Molecular Biotechnology, IBL, Leiden University. Liquid Minimal Medium (NMMP)³⁸ was used as medium for fermentation. Fully supplemented NMMP contained per liter 2 g (NH₄)₂SO₄, 5 g casaminoacids, 0.6 g MgSO₄·7H₂O, 1 ml minor elements solution, 15 mM Na⁺-K⁺ phosphate buffer pH 6.8, 0.5 % mannitol, 1% glycerol and 8 g peptone. Spores of *Streptomyces* sp. MBT76 were inoculated into 50 ml of culture medium, and cultures were incubated at 30 °C with continuous shaking at 220 rpm.

4.2. Isolation of bioactive compounds

Streptomyces sp. MBT76 mycelia were removed by centrifugation at 4,500x g for 10 min. The spent media (culture supernatant) was extracted twice with 20 ml of ethyl acetate. The organic phase was washed with 30 ml of distilled water and dried with 5 g of anhydrous Na₂SO₄. Finally, the ethyl acetate solvent was evaporated under vacuum at 40 °C and the resultant dried material was dissolved in 1.5~2.0 ml of ethyl acetate in a microtube

(Eppendorf type-5415C, Hamburg, Germany). The crude extract was evaporated under room temperature to remove organic solvent, and subsequently dipped into liquid nitrogen and lyophilized by freeze drier (Edwards Ltd., Crawley, UK). The obtained residue was independently dissolved in methanol- d_4 for ^1H NMR analysis (see below). All experiments were conducted in five replicates.

After ^1H NMR measurement, the five 96 h samples were combined and dissolved in methanol (50 ml), which was defatted twice with 20 ml of *n*-hexane. The resulting crude extract was directly separated by semi-preparative HPLC, which was performed with a Agilent 1200 series HPLC apparatus (Agilent technologies Inc, Santa Clara, CA, USA) equipped with a reversed-phase C_{18} column (Phenomenex Luna C18 (2) 100 Å 5 micron 250 $\times$ 10 mm, Torrance, CA, USA). The separation was eluted with a gradient of acetonitrile in water from 20% to 70% in 50 min at 2 ml/min flow rate, to obtain compounds **1** (t_{R} = 31.52 min, 1.14 mg), **3** (t_{R} = 35.75 min, 1.62 mg), **7** (t_{R} = 38.34 min, 0.34 mg), **8** (t_{R} = 39.50 min, 0.28 mg).

4.3. NMR measurement and data analysis

NMR sample preparation and measurements were done according to our previously published protocol.³⁹ Briefly, 0.5 ml of methanol- d_4 was added to the freeze-dried sample. Subsequently, the mixture was vortexed for 10 s and sonicated for 20 min at 42 kHz using an Ultrasonicator 5510E-MT (Branson, Danbury, CT, USA), followed by centrifugation at 13,000x *g* at room temperature for 5 min. The supernatant (0.3 ml) was transferred to a 5 mm micro NMR tube and analyzed. The ^1H NMR spectra were recorded at 25 °C on a 600 MHz Bruker DMX-600 spectrometer (Bruker, Karlsruhe, Germany) operating at a proton NMR frequency of 600.13 MHz. Deuterated methanol was used as the internal lock. Each ^1H NMR spectrum consisted of 128 scans using the following parameters: 0.16 Hz/point, pulse width (PW) = 30 (11.3 s) and relaxation delay (RD) = 1.5 s. Free induction decays (FIDs) were Fourier transformed with a line broadening (LB) = 0.3 Hz. The resulting spectra were manually phased and baseline corrected, and calibrated to residual methanol- d_4 at 3.30 ppm, using XWIN NMR (version 3.5, Bruker). The ^1H NMR data files were processed as described by Kim *et al.*³⁹ The AMIX software (Bruker Biospin GmbH) was used to reduce the ^1H NMR spectra to an ASCII file, with total intensity scaling. Bucketing or binning was performed and the spectral data were reduced to included regions of equal width (0.04 ppm) equivalent to the region of δ 0.30–10.00. The regions of δ 4.85 – 4.95 and δ 3.25 – 3.35 were removed in the analysis because of the remaining signal of solvents of HDO and CD_3OD , respectively. Principal component analysis (PCA) and orthogonal projection to latent structures (OPLS) based on Pareto scaling were performed with the SIMCA-P+ software (version 13.0, Umetrics, Umeå, Sweden).

4.4. UPLC-ToF-MS analysis

UPLC-ToF-MS analyses were performed on an UPLC system (Ultimate 3000, ThermoScientific, Germany) coupled to an ESI-IIQ-ToF spectrometer (microToF-QII, Bruker Daltonics, Germany) in the positive mode. The chromatographic separation was done using a Kinetex C_{18} UPLC 2.6 μm particle size column 150 $\times$ 2.0 mm (Phenomenex, Torrance, CA, USA) at a flow rate of 0.3 ml/min and a column temperature of 30 °C. Samples (3 μl) were

eluted using a gradient of solvent A (water) and B (acetonitrile), both with 0.1% formic acid (v/v). The initial percentage of B was 5%, which was linearly increased to 90% in 19.5 min, followed by a 2 min isocratic period and, then re-equilibrated with original conditions in 2 min. Nitrogen was used as drying and nebulizing gas. The gas flow was set at 10.0 l/min at 250 °C and the nebulizer pressure was 2.0 bar. The MS data were acquired over m/z range of 100–1000. The capillary voltage was 3.5 kV. For internal calibration, a 10 mM solution of sodium formate (Fluka, Steinheim, Germany) was infused. Formic acid, water and acetonitrile were LCMS grade, Optima (Fisher Scientific, Waltham, MA, USA).

4.5. Biotransformation

Substrates (10 mg) dissolved in 200 μ l DMSO were fed to two-day old *Streptomyces sp.* MBT76 culture broth. This biotransformation period lasted for another three days. The control experiment included both negative control of feeding 200 μ l DMSO solvent to MBT76, and positive control of 10 mg substrates dissolved in 200 μ l DMSO added to sterile culture media. Each experiment was done in triplicate. The method used to harvest the biotransformation products was as described for the metabolomics study.

To isolate the biotransformation products of genistein (Sigma), the combined mixture in 50 ml methanol was first defatted twice with 20 ml of *n*-hexane, which was subsequently fractionated by Sephadex LH-20 chromatography (GE Healthcare Life Sciences, Eindhoven, The Netherlands) eluting with methanol, to give seven fractions fr1—fr7. These fractions were subjected to NMR profiling to discard fr1—fr3, and fr6 that gave no isoflavone signals. Fr7 was purified by silica gel (pore size 60 Å, 70–230 mesh, St. Louis, MO, USA) to yield compound **15** (0.7 mg). Fr4 was separated by preparative TLC on a silica gel plate (Si60, Merck, Darmstadt, Germany), migrated with solvent system of CHCl₃-MeOH (20:3) to give pure compound **16** (0.5 mg) and semi-pure compound **18** (0.2 mg). Fr5 was purified by preparative TLC to give pure compound **17** (0.3 mg) using the same TLC conditions.

4.6. Antimicrobial activity of the isocoumarins and isoflavones

Antimicrobial test were performed by agar diffusion assay, against the Gram-positive indicator strain *Bacillus subtilis* 168 and the Gram-negative *Escherichia coli* K12, as we previously described.⁴⁰ The respective compounds were dissolved in methanol (2 mg/ml), and 20 μ l of the solution was applied on a paper disc (6 mm diameter). The discs were then placed onto an agar plate containing a soft agar overlay with either *Bacillus subtilis* 168 or *Escherichia coli* K12. After incubation at 37 °C for 18 h, growth inhibition zones (in mm) were monitored visually as a measure of the antimicrobial activity.

4.7. Genome sequencing, assembly, and annotation

Genome sequencing and annotation was done as described previously.⁴¹ Illumina/Solexa sequencing on Genome Analyzer IIX and sequencing on PacBio RS were outsourced to BaseClear BV (Leiden, The Netherlands). 100-nt paired-end reads were obtained and the quality of the short reads verified using FastQC (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/>). Depending on quality, reads were trimmed to various lengths at both ends. Processed raw reads were subsequently used as input for the Velvet assembly algorithm. Genomes were annotated using the RAST server with

default options. Contigs were also annotated using GeneMark.hmm for ORF prediction, BLASTP for putative function prediction and HMMER for protein-domain prediction, manually inspected for some and visualized using Artemis. *Streptomyces* sp. MBT76 was subjected to Illumina/Solexa whole genomic sequencing, and the genome was assembled in 13 contigs, with a total genome size of 8.64 Mb. In total 7974 coding sequences (CDS) were predicted. Analysis of the contigs by AntiSMASH³² presented a possible 55 putative biosynthetic gene clusters (BGCs) specifying secondary metabolites. The genome has been deposited at GenBank under the accession LNBE00000000.

REFERENCES

- (1) WHO Geneva, Switzerland, 2014.
- (2) Barka, E. A.; Vatsa, P.; Sanchez, L.; Gaveau-vaillant, N.; Jacquard, C.; Klenk, H.; Clément, C.; Ouhdouch, Y.; van Wezel, Gilles, P. *Microbiol. Mol. Biol. Rev.* **2016**, *80*, 1–43.
- (3) Bérdy, J. *J. Antibiot. (Tokyo)*. **2005**, *58*, 1–26.
- (4) Cooper, M. A.; Shlaes, D. *Nature* **2011**, *472*, 32.
- (5) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat. Rev. Drug Discov.* **2007**, *6*, 29–40.
- (6) Bentley, S. D.; Cerdeño-Tárraga, K. F. C. A.-M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Bentley, S. D.; Harper, D.; Bateman, A.; Brown, S.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Larke, L.; Murphy, L.; Oliver, K.; Rabinowitsch, E.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Woodward, J.; Barrell, B. G.; Parkhill, J. *Nature* **2002**, *3*, 141–147.
- (7) Cruz-Morales, P.; Vijgenboom, E.; Iruegas-Bocardo, F.; Girard, G.; Yáñez-Guerra, L. A.; Ramos-Aboites, H. E.; Pernodet, J.-L.; Anné, J.; van Wezel, G. P.; Barona-Gómez, F. *Genome Biol. Evol.* **2013**, *5*, 1165–1175.
- (8) Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S. *Nat. Biotechnol.* **2003**, *21*, 526–531.
- (9) Scherlach, K.; Hertweck, C. *Org. Biomol. Chem.* **2009**, *7*, 1753–1760.
- (10) Wu, C.; Kim, H. K.; van Wezel, G. P.; Choi, Y. H. *Drug Discov. Today Technol.* **2015**, *13*, 11–17.
- (11) Zhu, H.; Sandiford, S. K.; van Wezel, G. P. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 371–386.
- (12) Bibb, M. J. *Curr. Opin. Microbiol.* **2005**, *8*, 208–215.
- (13) van Wezel, G. P.; McDowall, K. J. *Nat. Prod. Rep.* **2011**, *28*, 1311–1333.
- (14) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Trends Biotechnol.* **2011**, *29*, 267–275.
- (15) Novoa-Carballal, R.; Fernandez-Megia, E.; Jimenez, C.; Riguera, R. *Nat. Prod. Rep.* **2011**, *28*, 78–98.
- (16) Wu, C.; Zacchetti, B.; Ram, A. F. J.; van Wezel, G. P.; Claessen, D.; Choi, Y. H. *Sci.*

- Rep.* **2015**, *5*, 10868.
- (17) Medema, M. H.; Kottmann, R.; Yilmaz, P.; Cummings, M.; Biggins, J. B.; Blin, K.; de Bruijn, I.; Chooi, Y. H.; Claesen, J.; Coates, R. C.; Cruz-Morales, P.; Duddela, S.; Düsterhus, S.; Edwards, D. J.; Fewer, D. P.; Garg, N.; Geiger, C.; Gomez-Escribano, J. P.; Greule, A.; Hadjithomas, M.; Haines, A. S.; Helfrich, E. J. N.; Hillwig, M. L.; Ishida, K.; Jones, A. C.; Jones, C. S.; Jungmann, K.; Kegler, C.; Kim, H. U.; Kötter, P.; Krug, D.; Masschelein, J.; Melnik, A. V.; Mantovani, S. M.; Monroe, E. a; Moore, M.; Moss, N.; Nützmänn, H.-W.; Pan, G.; Pati, A.; Petras, D.; Reen, F. J.; Rosconi, F.; Rui, Z.; Tian, Z.; Tobias, N. J.; Tsunematsu, Y.; Wiemann, P.; Wyckoff, E.; Yan, X.; Yim, G.; Yu, F.; Xie, Y.; Aigle, B.; Apel, A. K.; Balibar, C. J.; Balskus, E. P.; Barona-Gómez, F.; Bechthold, A.; Bode, H. B.; Borriss, R.; Brady, S. F.; Brakhage, A. a; Caffrey, P.; Cheng, Y.-Q.; Clardy, J.; Cox, R. J.; De Mot, R.; Donadio, S.; Donia, M. S.; van der Donk, W. a; Dorrestein, P. C.; Doyle, S.; Driessen, A. J. M.; Ehling-Schulz, M.; Entian, K.-D.; Fischbach, M. a; Gerwick, L.; Gerwick, W. H.; Gross, H.; Gust, B.; Hertweck, C.; Höfte, M.; Jensen, S. E.; Ju, J.; Katz, L.; Kaysser, L.; Klassen, J. L.; Keller, N. P.; Kormanec, J.; Kuipers, O. P.; Kuzuyama, T.; Kyrpides, N. C.; Kwon, H.-J.; Lautru, S.; Lavigne, R.; Lee, C. Y.; Linqun, B.; Liu, X.; Liu, W.; Luzhetskyy, A.; Mahmud, T.; Mast, Y.; Méndez, C.; Metsä-Ketelä, M.; Micklefield, J.; Mitchell, D. a; Moore, B. S.; Moreira, L. M.; Müller, R.; Neilan, B. a; Nett, M.; Nielsen, J.; O’Gara, F.; Oikawa, H.; Osbourn, A.; Osburne, M. S.; Ostash, B.; Payne, S. M.; Pernodet, J.-L.; Petricek, M.; Piel, J.; Ploux, O.; Raaijmakers, J. M.; Salas, J. a; Schmitt, E. K.; Scott, B.; Seipke, R. F.; Shen, B.; Sherman, D. H.; Sivonen, K.; Smanski, M. J.; Sosio, M.; Stegmann, E.; Süßmuth, R. D.; Tahlan, K.; Thomas, C. M.; Tang, Y.; Truman, A. W.; Viaud, M.; Walton, J. D.; Walsh, C. T.; Weber, T.; van Wezel, G. P.; Wilkinson, B.; Willey, J. M.; Wohlleben, W.; Wright, G. D.; Ziemert, N.; Zhang, C.; Zotchev, S. B.; Breitling, R.; Takano, E.; Glöckner, F. O. *Nat. Chem. Biol.* **2015**, *11*, 625–631.
 - (18) Schwab, W. *Phytochemistry* **2003**, *62*, 837–849.
 - (19) Zhu, H.; Swierstra, J.; Wu, C.; Girard, G.; Choi, Y. H.; van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. *Microbiology* **2014**, *160*, 1714–1725.
 - (20) Cerdeño, A. M.; Bibb, M. J.; Challis, G. L. *Chem. Biol.* **2001**, *8*, 817–829.
 - (21) Salem, S. M.; Kancharla, P.; Florova, G.; Gupta, S.; Lu, W.; Reynolds, K. A. *J. Am. Chem. Soc.* **2014**, *136*, 4565–4574.
 - (22) Mehdi, R. B. A.; Shaaban, K. A.; Rebai, I. K.; Smaoui, S.; Bejar, S.; Mellouli, L. *Nat. Prod. Res.* **2009**, *23*, 1095–1107.
 - (23) Ruanpanun, P.; Laatsch, H.; Tangchitsomkid, N.; Lumyong, S. *World J. Microbiol. Biotechnol.* **2011**, *27*, 1373–1380.
 - (24) Zinad, D. S.; Shaaban, K. A.; Abdalla, M. A.; Islam, M. T.; Schöffler, A.; Laatsch, H. *Nat. Prod. Commun.* **2011**, *6*, 45–48.
 - (25) Singh, B.; Parshad, R.; Khajuria, R. K.; Guru, S. K.; Pathania, A. S.; Sharma, R.; Chib, R.; Aravinda, S.; Gupta, V. K.; Khan, I. a.; Bhushan, S.; Bharate, S. B.; Vishwakarma, R. A. *Tetrahedron Lett.* **2013**, *54*, 6695–6699.
 - (26) Ishiuchi, K.; Nakazawa, T.; Ookuma, T.; Sugimoto, S.; Sato, M.; Tsunematsu, Y.; Ishikawa, N.; Noguchi, H.; Hotta, K.; Moriya, H.; Watanabe, K. *ChemBioChem* **2012**, *13*, 846–854.

-
- (27) Nakazawa, T.; Ishiuchi, K.; Praseuth, A.; Noguchi, H.; Hotta, K.; Watanabe, K. *Chembiochem* **2012**, *13*, 855–861.
- (28) Zaehle, C.; Gressler, M.; Shelest, E.; Geib, E.; Hertweck, C.; Brock, M. *Chem. Biol.* **2014**, *21*, 719–731.
- (29) Cochrane, R. V. K.; Sanichar, R.; Lambkin, G. R.; Reiz, B.; Xu, W.; Tang, Y.; Vederas, J. C. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 664–668.
- (30) Sun, H. H.; Ho, C. L.; Ding, F. Q.; Soehano, I.; Liu, X. W.; Liang, Z. X. *J. Am. Chem. Soc.* **2012**, *134*, 11924–11927.
- (31) Fisch, K. M.; Gurgui, C.; Heycke, N.; van der Sar, S. A.; Anderson, S. A.; Webb, V. L.; Taudien, S.; Platzer, M.; Rubio, B. K.; Robinson, S. J.; Crews, P.; Piel, J. *Nat. Chem. Biol.* **2009**, *5*, 494–501.
- (32) Blin, K.; Medema, M. H.; Kazempour, D.; Fischbach, M. A.; Breitling, R.; Takano, E.; Weber, T. *Nucleic Acids Res.* **2013**, *41*, 1–9.
- (33) Trygg, J.; Wold, S. *J. Chemom.* **2003**, *17*, 53–64.
- (34) Wu, C.; Du, C.; Gubbens, J.; Choi, Y. H.; van Wezel, G. P. *J. Nat. Prod.* **2015**, *78*, 2355–2363.
- (35) Miyazawa, M.; Takahashi, K.; Araki, H. *J. Chem. Technol. Biotechnol.* **2006**, *81*, 674–678.
- (36) Park, S. R.; Yoon, J. A.; Paik, J. H.; Park, J. W.; Jung, W. S.; Ban, Y.-H.; Kim, E. J.; Yoo, Y. J.; Han, A. R.; Yoon, Y. J. *J. Biotechnol.* **2009**, *141*, 181–188.
- (37) Thongnest, S.; Lhinhatrakool, T.; Wetprasit, N.; Sutthivaiyakit, P.; Sutthivaiyakit, S. *Phytochemistry* **2013**, *96*, 353–359.
- (38) Hodgson, D. *J. Gen. Microbiol.* **1982**, *3*, 2417–2430.
- (39) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Nat. Protoc.* **2010**, *5*, 536–549.
- (40) Wu, C.; van Wezel, G. P.; Choi, Y. H. *J. Antibiot. (Tokyo)*. **2015**, *68*, 445–452.
- (41) Girard, G.; Willemse, J.; Zhu, H.; Claessen, D.; Bukarasam, K.; Goodfellow, M.; van Wezel, G. P. *Antonie Van Leeuwenhoek* **2014**, *106*, 365–380.

Chapter 7

The cryptic *qin* gene cluster of *Streptomyces* sp. MBT76 specifies C-glycosylpyranonaphthoquinones with new chemistry

Changsheng Wu ^{1,2}, Chao Du ¹, Koji Ichinose ³, Young Hae Choi ², and Gilles P. van Wezel ¹

¹ *Molecular Biotechnology, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands*

² *Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands*

³ *Research Institute of Pharmaceutical Sciences, Musashino University, Shinmachi, Nishitokyo-shi, Tokyo 202-8585, Japan*

Submitted for publication

ABSTRACT

Antimicrobial resistance is one of the major threats to human health care. Mining of microbial genomes has revealed that actinomycetes harbor far more biosynthetic potential than anticipated. The activation of (cryptic) biosynthetic gene clusters and subsequent identification of the corresponding metabolites is a major strategy in modern drug discovery approaches. Here, we elucidated the structure of and the biosynthetic pathway for novel *C*-glycosylpyranonaphthoquinones in *Streptomyces* sp. MBT76, via an integrated genomics and metabolomics approach. Following activation of the cryptic *qin* gene cluster for a type II PKS by constitutive expression of its pathway-specific activator, bioinformatics coupled to NMR profiling facilitated the chromatographic isolation and structural elucidation of qinimycins A—C (**1—3**). The intriguing structural features of the qinimycins, including 8-*C*-glycosylation, 5,14-epoxidation and 13-hydroxylation, distinguished these molecules from the model pyranonaphthoquinones actinorhodin, medermycin, and granaticin. Another novelty lies in the unusual fusion of a deoxyaminosugar to the pyranonaphthoquinone backbone during biosynthesis of the antibiotics BE-54238 A and B (**4**, **5**). Qinimycins showed antimicrobial activity against Gram-positive but not against Gram-negative bacteria. Our work shows the utility of combining bioinformatics, targeted activation of cryptic gene clusters and NMR-based metabolic profiling as a pipeline to discover microbial natural products with distinctive skeletons.

1. INTRODUCTION

Actinomycetes are prolific sources of bioactive natural products (NPs). From the roughly 18,000 known bioactive bacterial compounds, more than 10,000 were discovered from the actinomycete genus *Streptomyces*.^{1,2} Even so, their biosynthetic potential is far from exhausted, though the enthusiasm to discover new molecules produced by these bacteria is dampened by chemical redundancy and the consequential poor return of investment of high-throughput screening campaigns.^{3–5} On a more positive note, whole genome sequencing revealed that actinomycetes harbor numerous silent and hence likely untapped biosynthetic gene clusters (BGCs) that may not be associated with known metabolites; indeed, even the very extensively studied model organism *Streptomyces coelicolor* was shown to possess a far greater producing potential than anticipated.⁶ Furthermore, bioinformatics tools developed specifically for mining genome sequences for the identification of BGCs allow the prediction of the chemical output of many BGCs on the basis of accumulated biosynthetic knowledge.^{7,8} These giant developments in genome mining mark the start of a new era of genomics-based drug discovery, with the potential of greatly expanding the chemical space of bioactive natural products.

A bottle-neck is that many of the biosynthetic pathways uncovered by genome sequencing are in a dormant state under routine laboratory conditions, generally referred to as cryptic or silent gene clusters, and specific approaches to activate their expression are required.^{9,10} The transcription of genes encoding the biosynthetic machinery for secondary metabolites in actinomycetes involves multiple regulatory cascades and networks.¹¹ The regulatory signals are transmitted through global regulatory networks and ultimately transmitted to the pathway-specific regulatory genes that control the expression of the biosynthetic genes.¹² Strategies to circumvent the regulatory networks that silence cryptic BGCs and trigger the biosynthesis of the corresponding natural product(s) include manipulating the pleiotropic regulatory networks as well as targeting pathway-specific regulators.^{10,13}

Another challenge is how to rapidly identify the relevant NPs within the context of the complex metabolic matrix. Metabolomic profiling is an efficient tool to achieve this goal. The available bioinformatic information allows prediction of the metabolite produced by the BGC, and this knowledge can be directly related to the ¹H NMR spectrum by examining the expected chemical shift and/or splitting pattern of typical protons in the predicted molecular motifs. In turn, as one compound usually contains multiple ¹H resonances, any unique NMR feature(s) in its structure can be used to check for the expression of the corresponding BGC and subsequently for NMR-guided chromatographic separation. This should allow scientists to bridge the gap between bioinformatics-driven gene cluster analysis and experimental NPs discovery.¹⁴

In this study, we activated the expression of a cryptic BGC identified in the genome of *Streptomyces* sp. MBT76, followed by NMR-based metabolomics to facilitate the characterization of the NPs. Constitutive expression of the likely pathway-specific activator of the type II PKS gene cluster (designated *qin*) activated the biosynthesis of a family of pyranonaphthoquinones with intriguing chemical architecture. Genomics and structural analysis identified a family of 8-*C*-glycosyl-pyranonaphthoquinones (**1**–**5**). The qinimycins A–C (**1**–**3**) encompass structural features of a rare 5,14-epoxidation and unprecedented

13-hydroxylation in pyranonaphthoquinones. In addition, an unusual fusion of a deoxyaminosugar into a pyranonaphthoquinone backbone was seen for the previously described antibiotics BE-54238A (4) and BE-54238B (5). Antimicrobial property of qinimycins was evaluated.

2. RESULTS AND DISCUSSION

2.1. Identification and activation of the cryptic type II PKS gene cluster (*qin*) for glycosylated pyranonaphthoquinones

Streptomyces sp. MBT76, which originates from the Qinling mountains in China, was previously identified as a prolific producer of antibiotics, including those with efficacy against multiple Gram-positive and Gram-negative MDR pathogens.¹⁵ Further detailed metabolic characterization of the strain identified many natural products with often interesting chemistry, including isocoumarins, prodiginines, acetyltryptamine, and fervenulin among others.¹⁶ *Streptomyces* sp. MBT76 was subjected to Illumina/Solexa whole genomic sequencing, and the genome was assembled in 13 contigs, with a total genome size of 8.64 Mb. In total 7974 coding sequences (CDS) were predicted using the GeneMark algorithm.¹⁷ Analysis of the contigs by AntiSMASH¹⁸ presented a possible 55 putative biosynthetic gene clusters (BGCs) specifying secondary metabolites, 22 of which encoding polyketide synthases (PKS).

One 41 kb biosynthetic gene cluster for a type II PKS, designated *qin* (Table 1) was of particular interest considering its potential to specify pyranonaphthoquinones, a well-studied family of aromatic polyketides with highly complex chemical architecture and pronounced bioactivities,^{19,20} including the representative members actinorhodin,²¹ medermycin,²² and granaticin.²³ In the *qin* BGC, besides the central PKS genes that are responsible for the biosynthesis of the pyranonaphthoquinone backbone, the presence of genes for the deoxyaminosugar D-forosamine in the upstream strongly suggested that the end product should be glycosylated. This genetic organization was similar to the clusters for synthesis of the glycosylated pyranonaphthoquinone medermycin in *Streptomyces* sp. AM-7161 (*med*)²² and granaticin in *Streptomyces violaceoruber* Tü22 (*gra*).²³ Coclustering of glycosylation-associated biosynthetic genes with those for the aglycones is typical of microbial genomes, which facilitates matching the BGC to the corresponding NPs.²⁴ In comparison, *qin*-ORF29 for an NADPH-dependent FMN reductase was absent in either the *med* or *gra* BGCs, and decoration with a D-forosamine is unprecedented in the pyranonaphthoquinone family. This promoted investigation into the potentially novel product(s) of the *qin* BGC. However, despite our previous detailed phytochemical investigations of *Streptomyces* sp. MBT76,^{15,16} the corresponding molecules had not been identified, suggesting that the gene cluster may be cryptic under the many different growth conditions that had been tested.

The *qin* BGC contains two putative regulatory genes, namely the SARP-family transcriptional regulatory gene *qin*-ORF11, and the TetR-family regulatory gene *qin*-ORF35 (Table 1). Many pathway-specific activators for secondary metabolite production in streptomycetes belong to the SARP family,¹² while TetR-family regulators often act as repressors.²⁵ As an example, in the type II PKS gene cluster *aur1* of *Streptomyces aureofaciens* CCM 3239, which specifies the angucycline-family auricin²⁶, *aur1P* and

aur1PR3 encode SARP-family activators,²⁷ while the *tetR*-type *aur1R* encodes a negative regulator.²⁸ To activate the *qin* gene cluster, we opted to over-express the likely pathway-specific activator gene *qin*-ORF11. For this, the gene was amplified by PCR and inserted behind the *ermE** promoter²⁹ in the conjugative and integrative vector pSET152. The resulting plasmid was then integrated into the *attB* site of the chromosome of MBT76 to create the recombinant derivative MBT76-1.

Table 1. Predicted functions of the gene products of the qinimycins biosynthetic gene cluster (*qin*) of *Streptomyces* sp. MBT76.

CDS	Length	Putative function	Nearest orthologue found in	aa identity	Accession number
1	322	dTDP-glucose 4,6-dehydratase	<i>Streptomyces</i> sp. NRRL S-623	79%	WP_031121809.1
2	359	dTDP-1-glucose synthase	<i>Streptomyces vietnamensis</i>	69%	ADO32770.1
3	280	N-methyltransferase	<i>Streptomyces clavuligerus</i>	39%	WP_003953258.1
4	365	C-glycosyltransferase	<i>Streptomyces</i> sp. SCC 2136	48%	CAF31363.2
5	385	NDP-hexose aminotransferase	<i>Streptomyces ambofaciens</i>	72%	CAM96587.1
6	421	Hypothetical protein	<i>Streptomyces violaceoruber</i>	45%	CAA09645.1
7	436	NDP-hexose 2,3-dehydratase	<i>Streptomyces</i> sp. NRRL S-623	67%	WP_031121795.1
8	351	NDP-hexose 3-ketoreductase	<i>Streptomyces</i> sp. NRRL S-623	64%	WP_032774990.1
9	432	NDP-hexose-3,4-dehydratase	<i>Streptomyces lydicus</i>	79%	CBA11561.1
10	378	flavin-dependent monooxygenase	<i>Streptomyces</i> sp. AM-7161	59%	BAC79043.1
11	267	SARP family transcriptional regulator	<i>Streptomyces cyaneofuscatus</i>	70%	WP_030562180.1
12	312	polyketide cyclase	<i>Streptomyces aureofaciens</i>	73%	YP_009060633.1
13	287	4-phosphopantetheinyl transferase	<i>Streptomyces albus</i>	45%	WP_031175178.1
14	571	type I polyketide synthase component	<i>Streptomyces scabiei</i>	44%	YP_003491848.1
15	142	hydroxylacyl-CoA dehydrogenase	<i>Streptomyces</i> sp. NRRL F-5135	69%	WP_030744938.1
16	334	bifunctional cyclase/dehydratase	<i>Streptomyces lavenduligriseus</i>	70%	WP_030792829.1
17	87	acyl carrier protein	<i>Streptomyces</i> sp. R1128	69%	AAG30201.1
18	410	Chain-Length Factor (polyketide beta-ketoacyl synthase beta subunit)	<i>Streptomyces lividans</i>	73%	WP_003973890.1
19	418	actinorhodin polyketide beta-ketoacyl synthase alpha subunit	<i>Streptomyces lavenduligriseus</i>	85%	WP_030792818.1
20	420	putative transporter	<i>Streptomyces viridochromogenes</i>	50%	WP_003988487.1
21	166	flavin reductase	<i>Streptomyces fulvoviolaceus</i>	57%	WP_052425229.1
22	261	ketoacyl reductase	<i>Streptomyces chartreusis</i>	72%	WP_01003439
23	309	hydroxylacyl-CoA dehydrogenase	<i>Streptomyces</i> sp. NRRL S-31	67%	WP_030735501.1
24	303	Quinone oxidoreductase	<i>Nocardia</i> sp. CNY236	72%	WP_028476080.1
25	217	DSBA-type oxidoreductase	<i>Streptosporangium roseum</i>	63%	WP_012891327.1
26	191	dehydratase	<i>Streptomyces</i> sp. AM-7161	55%	BAC79041.1
27	576	stereospecific keto reductase	<i>Streptomyces</i> sp. AM-7161	64%	BAC79036.1
28	328	Quinone oxidoreductase	<i>Streptomyces</i> sp. AM-7161	71%	BAC79039.1
29	201	NADPH-dependent FMN reductase	<i>Streptomyces</i> sp. NRRL S-623	64%	WP_031121793.1
30	402	S-adenosylmethionine synthetase	<i>Streptomyces roseoverticillatus</i>	95%	WP_030365339.1
31	339	adenosine kinase	<i>Streptomyces albus</i>	72%	AIA08031.1
32	1167	methionine synthase	<i>Streptomyces roseoverticillatus</i>	93%	WP_030367617.1
33	312	5,10-methylenetetrahydrofolate reductase	<i>Streptomyces</i> sp. NRRL F-5135	81%	WP_030742694.1
34	473	adenosylhomocysteinase	<i>Streptomyces roseoverticillatus</i>	90%	WP_030365247.1
35	227	TetR family transcriptional regulator	<i>Streptomyces lividans</i>	49%	WP_003972639.1

2.2. NMR-based metabolic profiling expedited characterization of novel glycosylated pyranonaphthoquinones

Ex-conjugant MBT76-1 and its parent MBT76 were fermented in parallel in liquid modified

NMMP media.¹⁶ After five days of growth, the cultures were harvested by centrifugation and extracted with EtOAc. The obtained metabolites were subjected to ^1H NMR profiling.³⁰ As shown in Figure 1, MBT76-1 presented a very different metabolic profile in comparison to the parental strain. The production of many of the metabolites usually produced by the wild-type strain, such as 1*H*-pyrrole-2-carboxamide, acetyltryptamine, fervenulin and 2-hydroxy-3-methoxy-benzamide,¹⁶ was aborted in MBT76-1. While the production of 6,8-dihydroxy-3-methyl-isocoumarin (**9**) was not affected by ORF11 overexpression, its post-PKS methoxylation was hampered as shown by the disappearance of methoxyl groups in the region δ 3.45–4.25.¹⁶ Conversely, the ^1H NMR spectra of MBT76-1 now exhibited many previously unseen proton resonances absent in the parental strain, indicating that different compounds were produced. These included molecules with characteristic NMR signals such as δ 4.45 (dd, J = 4.2, 1.8 Hz) for the methine in a substituted pyran ring, and the methyl doublet at δ 1.29 (d, J = 6.6 Hz) which were in agreement with kalafungin (16-CH₃) or 6'-deoxy-CH₃ of a sugar moiety predicted by bioinformatics; In the aromatic region, coupling doublets at δ 8.01 (d, J = 8.4 Hz), and 6.56 (d, J = 8.4 Hz) were indicative of an α,β -unsaturated ketone aryl moiety typical of pyranonaphthoquinones. All these NMR features indicated that indeed the *qin* gene cluster was actively expressed in MBT76-1.

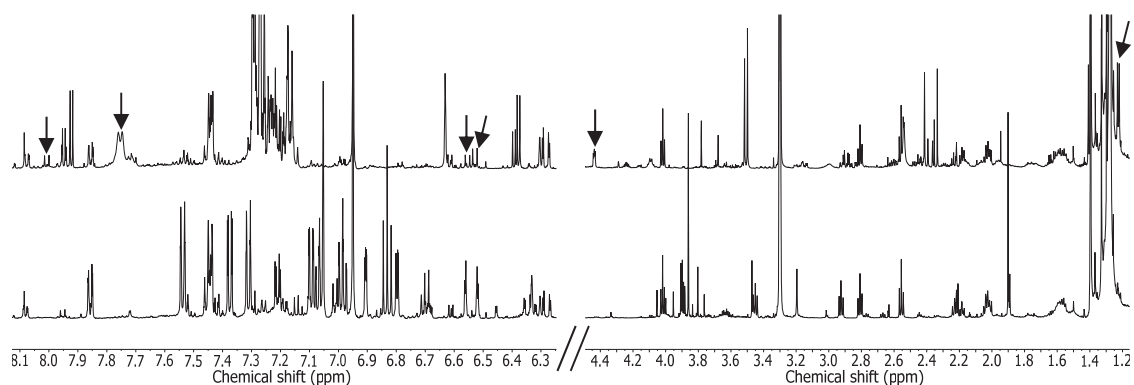


Figure 1. Comparison of ^1H NMR spectra. ^1H NMR (600 MHz, in methanol- d_4) was obtained on the crude extracts from wild-type strain *Streptomyces* sp. MBT76 (bottom) and its ex-conjugant MBT76-1 (top). The presented chemical shifts are for aromatic region δ 6.5–8.1 and aliphatic region δ 1.2–4.5. The signals highlighted by arrows are attributable to glycosylated pyranonaphthoquinone molecules.

To identify the metabolic product(s) of the *qin* gene cluster, the molecules were separated chromatographically by semi-preparative HPLC-UV chromatography. The resulting fractions (Fr1–16) were analysed by ^1H NMR spectroscopy, which showed that Fr1 (2.1 mg) contained the sought-after deoxysugar-pyranonaphthoquinones. Further UHPLC-UV-ToF-HRMS analysis confirmed Fr1 mainly contained three structurally related compounds, with molecular formulas $\text{C}_{23}\text{H}_{27}\text{NO}_8$ (**1**), $\text{C}_{23}\text{H}_{29}\text{NO}_9$ (**2**), and $\text{C}_{24}\text{H}_{31}\text{NO}_9$ (**3**) with a relative abundance of 6:2:1 (Figure S1). Elucidation of the final structure of major compound **1** was done on the basis of extensive NMR analysis, including ^1H NMR, HSQC, HMBC, COSY, and APT. The ^1H NMR spectrum of compound **1** (Table 2) presented characteristic resonances for the bioinformatics-predicted deoxyaminosugar forosamine, including a 6'-deoxy-CH₃ doublet at δ_{H} 1.29 (d, J = 6.6 Hz), a 7'-*N*-CH₃ singlet at δ_{H} 2.89 (s), and highly overlapping signals for 2',3'-deoxy-CH₂ at δ_{H} 2.20–2.40. Aided by COSY and HSQC experiments (Figure S2), the backbone of the aglycone was confirmed as a

kalafungin-type molecule, as established by the spin system H₂-2/H-3/H-4 for a γ -lactone ring, together with two coupling protons at δ_{H} 1.41 (d, J = 6.6 Hz, H-16) and 4.93 (q, J = 6.6 Hz, H-15) for a pyran ring. The additional two aromatic coupling doublets δ_{H} 7.65 (d, J = 8.4 Hz, H-9), and 7.25 (d, J = 8.4 Hz, H-10) indicated that the deoxyaminosugar was linked to the aglycone through the C-8 residue of the benzene ring, which was confirmed by the HMBC correlations (Figure S2) from H₂-2' (δ_{H} 2.30, 2.24) to C-8 (δ_{C} 124.2) and from H-9 (δ_{H} 7.65) to C-1' (δ_{C} 68.7). More importantly, the key HMBC correlations from the H-13 (δ_{H} 5.48) to C-11 (δ_{C} 158.1), C-12 (δ_{C} 129.2), C-7 (δ_{C} 131.1), C-5 (δ_{C} 60.6), C-14 (δ_{C} 68.1), and C-15 (δ_{C} 65.9) demonstrated that the pyranonaphthoquinone skeleton of kalafungin (**6**) acquired two variations, via the epoxidation of the $\Delta^{5,14}$ double bond and hydroxylation of the ketone at C-13. The stereochemistry at C-13 was established as *R*, based on the NOESY correlation between H-13/H₃-16, but no correlation was observed for H-13/H-15. However, the configuration of 5,14-epoxy remained unclear. Consequently, the final structure of major compound **1** (Figure 2) in Fr1 was elucidated as a novel glycosylated pyranonaphthoquinone, and named qinimycin A. Similar NMR spectral analyses revealed that compound **2** was a hydration product of **1** in the γ -lactone ring (Table 2), and accordingly named qinimycin B. Though compound **3** (qinimycin C) was not resolved by NMR spectra due to its low abundance, the additional mass of 14.0153 (Figure S1) corresponding exactly to a CH₂ unit suggested that the imide in **2** was further methylated into a tertiary amine, since the *N*-methyltransferase ORF3 is able to *di*-methylate the nitrogen atom of dTDP-D-forosamine in spinosyn pathway.³¹ The known compounds **4**—**14** (Figure 2) were identified in Fr2—Fr16 by ¹H NMR and/or UHPLC-UV-ToF-HRMS, and the results were compared with spectroscopic data from the literature (Table S1).

Table 2. ¹H and ¹³C NMR data for compounds **1** and **2**^a

NO.	1		2	
	δ_{C} , type	δ_{H} (J in Hz)	δ_{C} , type	δ_{H} (J in Hz)
1	177.4, C		174.4, C	
2	37.3, CH ₂	2.91, dd (15.6, 4.2); 2.40, dd (15.6, 1.8)	36.5, CH ₂	2.58, d (7.2)
3	68.0, CH	4.45, dd (4.2, 1.8)	67.1, CH	4.16, m
4	74.0, CH	4.98, d (1.2)	64.0, CH	4.32, d (12.6)
5	60.6, C		63.0, C	
6	194.6, C		197.1, C	
7	131.1, C		131.5, C	
8	124.2, C		124.1, C	
9	132.2, CH	7.65, d (8.4)	132.1, CH	7.62, d (8.4)
10	121.4, CH	7.25, d (8.4)	121.3, CH	7.23, d (8.4)
11	158.1, C		158.1, C	
12	129.2, C		129.3, C	
13	61.0, CH	5.48, d (0.6)	61.2, CH	5.43, d (0.6)
14	68.1, C		67.7, C	
15	65.9, CH	4.93, q (6.6)	65.7, CH	4.90, q (6.6)
16	14.1, CH ₃	1.41, d (6.6)	14.4, CH ₃	1.40, d (6.6)
1'	68.7, CH	5.36, m	68.7, CH	5.36, m
2'	32.1, CH ₂	2.30, m; 2.24, m	32.1, CH ₂	2.30, m; 2.24, m
3'	22.5, CH ₂	2.38, m; 2.21, m	22.5, CH ₂	2.38, m; 2.21, m
4'	75.0, CH	3.63, td (7.2, 1.8)	75.0, CH	3.63, td (7.2, 1.8)
5'	63.8, CH	4.27, m	63.8, CH	4.27, m
6'	19.9, CH ₃	1.29, d (6.6)	19.9, CH ₃	1.29, d (6.6)
7'	39.0, CH ₃	2.89, s	39.0, CH ₃	2.89, s

^a **1** and **2** were recorded in CD₃OD. Proton coupling constants (J) in Hz are given in parentheses. All chemical shift assignments were done on the basis of 1D and 2D NMR techniques.

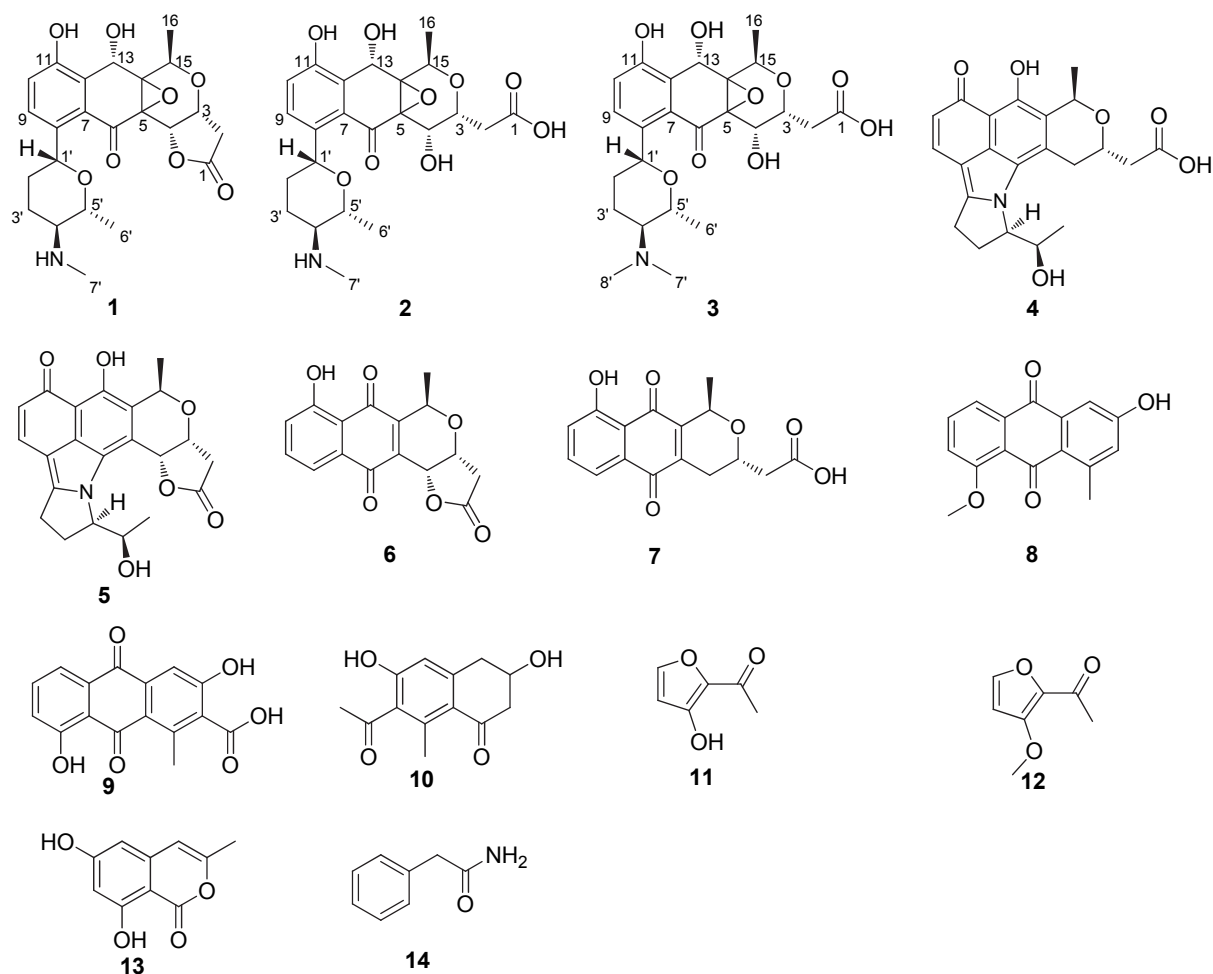


Figure 2. Secondary metabolites produced by *Streptomyces* sp. MBT76-1. Compounds identification was done on the basis of NMR and HRMS techniques, and their spectroscopy data were summarized in Table 2 and Table S1.

The qinimycins have intriguing chemical features, representing a new branch within the extensively studied pyranonaphthoquinone family of natural products. The 5,14-epoxidation was previously described in pyranonaphthoquinone aglycones,^{32,33} but hydroxylation of the pyranonaphthoquinone backbone at C-13 is unprecedented. The 8-C-glycosylation of pyranonaphthoquinone is rare, though previously seen in the antibiotic Sch 38519,^{34,35} and xiakemycin A.³⁶ Within the pyranonaphthoquinones, glycosylation with the deoxyaminosugar forosamine is uniquely seen in the qinimycins. With this chemical architectural complexity, the qinimycins rival the well-known pyranonaphthoquinone glycosides medermycin²² or granaticin²³ in terms of structural complexity. The antimicrobial activity of qinimycins were evaluated by agar diffusion assays. Growth inhibition was seen for the Gram-positive bacteria *Bacillus subtilis* 168 and *Staphylococcus aureus* CECT976, but not for Gram-negative *Escherichia coli* JM109 or *Pseudomonas aeruginosa* PAO1 (Table 3).

Table 3. Antimicrobial activity of qinimycins.

Compound NO.	Inhibition zone (mm)			
	<i>Bacillus subtilis</i>	<i>Escherichia coli</i>	<i>Staphylococcus aureus</i>	<i>Pseudomonas aeruginosa</i>
Fr1	15	0	20	0
AMP	23	20	10	7
APRA	10	7	7	15
NC	0	0	0	0

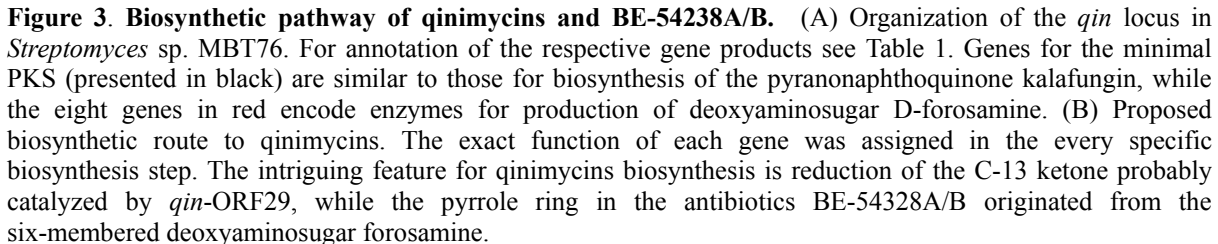
For Fr1, 25 μ L was spotted of a 2 mg/mL solution in methanol; For AMP and APRA, 5 μ L was spotted of a 1 mg/mL solution in miliQ water. AMP, ampicillin; APRA, apramycin; NC, negative control (methanol).

2.3. Proposed biosynthetic pathway for qinimycins based on bioinformatics analysis

A biosynthetic model for qinimycins is proposed based on the functional assignments from sequence analysis and the known metabolites that are produced (Figure 3). The assembly line includes three steps: (i) biosynthesis of an activated form of dTDP-D-forosamine, similar to that found in the spinosyn pathway;³¹ (ii) assembly of the kalafungin aglycone by a minimal PKS similar to that found in the BGCs for medermycin²² and granaticin;²³ and (iii) regioselective installing of a dTDP-D-forosamine moiety at the C-8 position of kalafungin by a C-glycosyltransferase (most likely encoded by *qin*-ORF4).

In a pilot study of actinorhodin biosynthesis, a two-component flavin-dependent monooxygenase (FMO), encoded by *actVA*-ORF5 (oxygenase) and *actVB* (flavin reductase), was demonstrated to perform the two consecutive oxygenations at C-6 and C-8.³⁷ The FMO genes are commonly present in the *med* and *gra* BGCs for medermycin²² and granaticin,²³ respectively. The ActVA homologue Gra-21 was shown to be bifunctional at C-6 and C-8, whereas Med-7 mono-functionally oxygenizes the C-6 position. Interestingly, the gene product of *qin*-ORF10 shows higher similarity to Med-7 (59% aa identity) than to ActVA-ORF5 (53% aa identity), suggesting possible mono-functionality at C-6. The single oxygenation at C-6 leading to a pyranonaphthoquinone skeleton is in good agreement with the subsequent C-glycosylation for qinimycins at C-8 and medermycin at C-10, where the selectivity of the responsible glycosyl transferase *qin*-ORF4 is of great interest. Further studies³³ on ActVA-ORF5/ActVB revealed its additional *in vitro* epoxidation activity of kalafungin. The co-occurrence of three qinimycins (**1**—**3**) is a clear example of *in vivo* activity of a two-component FMO encoded by *qin*-ORF10/*qin*-ORF21. Compared with the *act*,⁶ *med*,²² and *gra*²³ gene clusters, *qin*-ORF29 for an NADPH-dependent FMN reductase is unique in the *qin* BGC. We propose that this enzyme may execute the unprecedented reduction of the C-13 ketone.

It is noteworthy that, based on the biosynthetic pathway (Figure 3), we predict that also qinimycins other than compounds (**1**—**3**) are produced by *Streptomyces* sp. MBT76, which may have been missed due to lower yields. For instance, UHPLC-UV-ToF-HRMS analysis of MBT76-1 crude extract showed molecules with the molecular formulae C₂₃H₂₇NO₉ and C₂₃H₂₅NO₈ for the putative compounds **1a** and **1b**, respectively. The post-PKS glycosylation of 8-C-glycosylpyranonaphthoquinones BE-54238A (**4**) and BE-54238B (**5**)³⁸ involved rearrangement of the deoxyaminosugar moiety. After glycosylation of kalafungin into **5a**, the six-membered amino pyran ring was first rearranged into a five-membered 2H-3,4-dihydropyrrole ring through a *Mannich* reaction, whereby the gene product of *qin*-ORF6 might mediate the ring-opening of the appended forosamine to a linear ketone. The newly formed imide was then further cyclized with the aglycone quinone at C-6 to give an intermediary iminium ion, which subsequently underwent proton tautomerization to produce antibiotic BE-54238A/B. Interestingly, this naturally occurring biosynthetic pathway mimics the total synthesis of BE-54238B from rhamnose.^{39,40} The anthraquinones 11-*O*-methyl-aloesaponarin II (**8**) and 3,8-dihydroxy-1-methyl-anthraquinone-carboxylic acid (**9**) were shunt products of the kalafungin biosynthetic pathway,⁴¹ but it remains unclear whether polyketide molecules **10**—**12** were also derived from the *qin* BGC.



3. CONCLUSION

The replication issues that have frustrated high-throughput screening regimes necessitate new approaches to expand the chemical diversity of natural products and keep filling the discovery pipelines with new lead compounds. The combination of metabolomics and genome mining is an efficient way to streamline the discovery of novel molecules from actinomycetes. Prior knowledge of the types of NPs that can be expected based on bioinformatics thereby simplifies the phytochemical isolation process and final structure determination of target compound(s), while NMR profiling of molecules in highly complex matrices allows rapid linkage of chemotype to genotype. We here provide proof of concept for this principle, whereby activation of the cryptic *qin* type II PKS gene cluster by constitutive expression of its pathway-specific activator gene (*qin*-ORF11) followed by NMR-based metabolic profiling, identified novel glycosylated pyranonaphthoquinones with exciting chemical architecture (**1—5**). The elucidated biosynthetic pathway for the qinimycins in the genetically tractable *Streptomyces* sp. MBT76 offers new insights into the biosynthesis of this exciting family of natural products. The best studied pyranonaphthoquinone is undoubtedly actinorhodin produced by *Streptomyces coelicolor* A3(2), which is the archetype of this family of polyketides. The qinimycins described in this work have intriguing structural features that make them stand out, namely 8-C-glycosylation by the deoxyaminosugar D-forosamine, as well as 5,14-epoxidation and 13-hydroxylation. Another novelty lies in the unusual fusion of the deoxyaminosugar to the pyranonaphthoquinone backbone during biosynthesis of the antibiotics BE-54238 A and B. Thus, the qinimycins form a new branch of the family of pyranonaphthoquinone-type antibiotics and provides important new insights into the biosynthesis of this important class of natural products.

4. EXPERIMENTAL SECTION

4.1. Bacterial strains and culturing conditions

Streptomyces sp. MBT76 was obtained from the culture collection of Molecular Biotechnology, IBL, Leiden University. *Escherichia coli* JM109⁴² was used for routine cloning. *Escherichia coli* ET12567⁴³ containing pUZ8002⁴⁴ was used for introducing non-methylated DNA into *Streptomyces* by conjugation. The basal medium for *Streptomyces* sp. MBT76 growth was modified minimal liquid medium NMMP⁴⁴ without PEG6000 and containing 1% (w/v) glycerol and 0.5% (w/v) mannitol as the carbon sources, which was further supplemented with 0.8% (w/v) Bacto peptone.¹⁶ Tryptone soy broth with 10% sucrose (TSBS) was used to grow MBT76 mycelia as receptor for conjugation experiments. Soy flour mannitol (SFM) agar plates⁴⁴ were used to grow overexpression conjugants of MBT76. For culturing of MBT76 for chemical analysis, 50 mL modified NMMP was inoculated with 10⁶ spores of *Streptomyces* sp. MBT76 in 250 mL flasks equipped with a spring, and grown at 30 °C with constant shaking at 220 rpm. The incubation lasted for 120 h.

4.2. Genome sequencing, assembly, and annotation

DNA was extracted from *Streptomyces* sp. MBT76 as described.⁴⁴ Genome sequencing and annotation was done essentially as described previously.⁴⁵ Illumina/Solexa sequencing on Genome Analyzer Iix and sequencing on PacBio RS were outsourced to BaseClear BV

(Leiden, The Netherlands). 100-nt paired-end reads were obtained and the quality of the short reads verified using FastQC (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/>). Depending on quality, reads were trimmed to various lengths at both ends. Processed raw reads were subsequently used as input for the Velvet assembly algorithm. Genomes were annotated using the RAST server with default options. Contigs were also annotated using GeneMark.hmm¹⁷ for ORF prediction, BLASTP for putative function prediction and HMMER for protein-domain prediction, manually inspected for some and visualized using Artemis. The genome has been deposited at GenBank under the accession LNBE00000000.

4.3. Overexpression of the pathway-specific SARP regulator *qin*-ORF11

qin-ORF11 was amplified by PCR from *Streptomyces* sp. MBT76 genomic DNA as described⁴⁶ using primers SC06_0044_F_EcoRI 5'-CGATGAATTCCGTCCGGCGCTTCGTGTG and SC06_0044_R_EcoRI_NdeI

5'-CGATGAATTCGCACCCGGTACAGGAGTGTGTGCATATGCGATTCAACCTCATC.

The PCR product was subcloned as an EcoRI fragment and subsequently ligated as an NdeI-HindIII fragment behind the constitutive *ermE* promoter in pSET152,²⁹ which integrates at the ϕ C31 attachment site in the *Streptomyces* chromosome. The final recombinant plasmid pCSW01 (pSET152/*ermE*/*qin*-ORF11) was subsequently transformed into *Escherichia coli* ET12567/pUZ8002 and the positive transformant was selected by 100 μ g/mL apramycin. This strain was then used to conjugate two-day old mycelia of *Streptomyces* sp. MBT76. The ex-conjugant was confirmed by PCR using primers SC06_0044_SF 5'-TTTCCCAGTCACGACGTTG and SC06_0044_SR 5'-GGATAACAATTCACACAGG.

4.4. Metabolomics and compound purification

50 mL cultures of *Streptomyces* sp. MBT76 or MBT76-1 were harvested by centrifugation at 4000 rpm for 10 min, and the supernatant extracted twice with 20 mL of ethyl acetate (EtOAc). The organic phase was washed with 30 mL of water and subsequently dried with 5 g of anhydrous Na₂SO₄. EtOAc was removed under vacuum at 38 °C and the residue was dissolved in 2.0 mL of EtOAc in a microtube (Eppendorf type-5415C, Hamburg, Germany). The solvent was then evaporated at room temperature under nitrogen gas, and subsequently dipped into liquid nitrogen and lyophilized using a freeze dryer (Edwards Ltd., Crawley, England). The NMR sample preparation and measurements were done according to our previously published protocol.³⁰ For details see⁴⁷.

For separation of metabolites, 15 replicates of *Streptomyces* sp. MBT76-1 culture (750 mL in total) were pooled. After EtOAc extraction, 0.26 g of crude extract was partitioned between methanol and *n*-hexane to remove the lipids. The resolved methanol fraction was subsequently separated by semipreparative reversed-phase HPLC (Phenomenex Luna 5 μ m C18 (2) 100 Å column, 250 $\times$ 10 mm) on an Agilent 1200 series HPLC (Agilent technologies Inc., Santa Clara, CA, USA), eluting with a gradient of ACN in H₂O adjusted with 0.1% TFA from 20% to 60% at a flow rate of 2 mL/min in 40 min. 16 fractions (Fr1—Fr16) were manually collected by peak detection at 254 nm, which were numbered in ascending order of retention time. After rotary evaporation at 42 °C under vacuum, these 16 fractions were further subjected to ¹H NMR profiling and UHPLC-UV-ToF-MS analysis. Qinimycins were

contained in Fr1 (2.1 mg) at retention time 8.12 min.

4.5. UHPLC-UV-ToF-MS analysis

UHPLC-ToF-MS analyses were performed on an UHPLC system (Ultimate 3000, ThermoScientific, Germany) coupled to an ESI-IIQ-TOF spectrometer (micrOTOF-QII, Bruker Daltonics, Germany) in the positive mode.⁴⁸ The chromatographic separation was done using a Kinetex C₁₈ UHPLC 2.6 μ m particle size column 150 $\times$ 2.0 mm (Phenomenex, USA) at a flow rate of 0.3 mL/min and a column temperature of 30 °C. Samples (3 μ L) were eluted using a gradient of solvent A (water) and B (acetonitrile), both with 0.1% formic acid (v/v). The initial percentage of B was 5%, which was linearly increased to 90% in 19.5 min, followed by a 2 min isocratic period and, then re-equilibrated with original conditions in 2 min. Nitrogen was used as drying and nebulizing gas. The gas flow was set at 10.0 L/min at 250 °C and the nebulizer pressure was 2.0 bar. The MS data were acquired over m/z range of 100–1000. The capillary voltage was 3.5 kV. For internal calibration, a 10 mM solution of sodium formate (Fluka, Steinheim, Germany) was infused. Formic acid, water and acetonitrile were LCMS grade (Optima, Fisher Scientific, NJ, USA).

4.6. Antimicrobial activity assays

Antimicrobial activity of purified new compounds was determined according to a disc diffusion method as described.^{49, 50} 25 μ L of Fr1 (2 mg/mL in methanol) was spotted onto paper discs (6 mm diameter) placed on agar plates containing a soft agar overlay with indicator bacteria. Indicator bacteria were *Bacillus subtilis* 168, *Escherichia coli* ASD19, *Staphylococcus aureus* CECT976, or *Pseudomonas aeruginosa* PAO1. Ampicillin and apramycin were used as positive controls, whereby 5 μ L was spotted of a 1 mg/mL solution in miliQ water. The solvent methanol was used as the negative control. After incubation at 37 °C for 18 h, growth inhibition zones (in mm) were recorded as antimicrobial activity.

REFERENCES

- (1) Bérdy, J. J. *Antibiot. (Tokyo)*. **2012**, 65, 385–395.
- (2) Barka, E. A.; Vatsa, P.; Sanchez, L.; Gaveau-vaillant, N.; Jacquard, C.; Klenk, H.; Clément, C.; Ouhdouch, Y.; van Wezel, Gilles, P. *Microbiol. Mol. Biol. Rev.* **2016**, 80, 1–43.
- (3) Cooper, M. A.; Shlaes, D. *Nature* **2011**, 472, 32.
- (4) Kolter, R.; van Wezel, G. P. *Nat. Microbiol.* **2016**, 1, 15020.
- (5) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat. Rev. Drug Discov.* **2007**, 6, 29–40.
- (6) Bentley, S. D.; Cerdeño-Tárraga, K. F. C. A.-M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Bentley, S. D.; Harper, D.; Bateman, A.; Brown, S.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Larke, L.; Murphy, L.; Oliver, K.; Rabinowitsch, E.; Rutherford, K.; Rutter, S.; Seeger,

- K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Woodward, J.; Barrell, B. G.; Parkhill, J. *Nature* **2002**, *3*, 141–147.
- (7) Medema, M. H.; Kottmann, R.; Yilmaz, P.; Cummings, M.; Biggins, J. B.; Blin, K.; de Bruijn, I.; Chooi, Y. H.; Claesen, J.; Coates, R. C.; Cruz-Morales, P.; Duddela, S.; Düsterhus, S.; Edwards, D. J.; Fewer, D. P.; Garg, N.; Geiger, C.; Gomez-Escribano, J. P.; Greule, A.; Hadjithomas, M.; Haines, A. S.; Helfrich, E. J. N.; Hillwig, M. L.; Ishida, K.; Jones, A. C.; Jones, C. S.; Jungmann, K.; Kegler, C.; Kim, H. U.; Kötter, P.; Krug, D.; Masschelein, J.; Melnik, A. V.; Mantovani, S. M.; Monroe, E. a; Moore, M.; Moss, N.; Nützmann, H.-W.; Pan, G.; Pati, A.; Petras, D.; Reen, F. J.; Rosconi, F.; Rui, Z.; Tian, Z.; Tobias, N. J.; Tsunematsu, Y.; Wiemann, P.; Wyckoff, E.; Yan, X.; Yim, G.; Yu, F.; Xie, Y.; Aigle, B.; Apel, A. K.; Balibar, C. J.; Balskus, E. P.; Barona-Gómez, F.; Bechthold, A.; Bode, H. B.; Borriss, R.; Brady, S. F.; Brakhage, A. a; Caffrey, P.; Cheng, Y.-Q.; Clardy, J.; Cox, R. J.; De Mot, R.; Donadio, S.; Donia, M. S.; van der Donk, W. a; Dorrestein, P. C.; Doyle, S.; Driessen, A. J. M.; Ehling-Schulz, M.; Entian, K.-D.; Fischbach, M. a; Gerwick, L.; Gerwick, W. H.; Gross, H.; Gust, B.; Hertweck, C.; Höfte, M.; Jensen, S. E.; Ju, J.; Katz, L.; Kaysser, L.; Klassen, J. L.; Keller, N. P.; Kormanec, J.; Kuipers, O. P.; Kuzuyama, T.; Kyrpides, N. C.; Kwon, H.-J.; Lautru, S.; Lavigne, R.; Lee, C. Y.; Linquan, B.; Liu, X.; Liu, W.; Luzhetskyy, A.; Mahmud, T.; Mast, Y.; Méndez, C.; Metsä-Ketelä, M.; Micklefield, J.; Mitchell, D. a; Moore, B. S.; Moreira, L. M.; Müller, R.; Neilan, B. a; Nett, M.; Nielsen, J.; O’Gara, F.; Oikawa, H.; Osbourn, A.; Osburne, M. S.; Ostash, B.; Payne, S. M.; Pernodet, J.-L.; Petricek, M.; Piel, J.; Ploux, O.; Raaijmakers, J. M.; Salas, J. a; Schmitt, E. K.; Scott, B.; Seipke, R. F.; Shen, B.; Sherman, D. H.; Sivonen, K.; Smanski, M. J.; Sosio, M.; Stegmann, E.; Süssmuth, R. D.; Tahlan, K.; Thomas, C. M.; Tang, Y.; Truman, A. W.; Viaud, M.; Walton, J. D.; Walsh, C. T.; Weber, T.; van Wezel, G. P.; Wilkinson, B.; Willey, J. M.; Wohlleben, W.; Wright, G. D.; Ziemert, N.; Zhang, C.; Zotchev, S. B.; Breitling, R.; Takano, E.; Glöckner, F. O. *Nat. Chem. Biol.* **2015**, *11*, 625–631.
 - (8) Fedorova, N. D.; Muktali, V.; Medema, M. H. In *Methods in Molecular Biology*; 2012; Vol. 944, pp. 23–45.
 - (9) Zhu, H.; Sandiford, S. K.; van Wezel, G. P. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 371–386.
 - (10) Rutledge, P. J.; Challis, G. L. *Nat. Rev. Microbiol.* **2015**, *13*, 509–523.
 - (11) van Wezel, G. P.; McDowall, K. J. *Nat. Prod. Rep.* **2011**, *28*, 1311–1333.
 - (12) Bibb, M. J. *Curr. Opin. Microbiol.* **2005**, *8*, 208–215.
 - (13) Abdelmohsen, U. R.; Grkovic, T.; Balasubramanian, S.; Kamel, M. S.; Quinn, R. J.; Hentschel, U. *Biotechnol. Adv.* **2015**, *33*, 798–811.
 - (14) Wu, C.; Choi, Y. H.; van Wezel, G. P. *J. Ind. Microbiol. Biotechnol.* **2016**, *43*, 299–312.
 - (15) Zhu, H.; Swierstra, J.; Wu, C.; Girard, G.; Choi, Y. H.; van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. *Microbiology* **2014**, *160*, 1714–1725.
 - (16) Wu, C.; Zhu, H.; van Wezel, G. P.; Choi, Y. H. *Metabolomics* **2016**, *12*, DOI 10.1007/s11306-016-1025-6.
 - (17) Lukashin, A. V.; Borodovsky, M. *Nucleic Acids Res.* **1998**, *26*, 1107–1115.
 - (18) Blin, K.; Medema, M. H.; Kazempour, D.; Fischbach, M. A.; Breitling, R.; Takano, E.;

- Weber, T. *Nucleic Acids Res.* **2013**, *41*, 1–9.
- (19) Metsä-Ketelä, M.; Oja, T.; Taguchi, T.; Okamoto, S.; Ichinose, K. *Curr. Opin. Chem. Biol.* **2013**, *17*, 562–570.
- (20) Oja, T.; Galindo, P. S. M.; Taguchi, T.; Manner, S.; Vuorela, P. M.; Ichinose, K.; Metsä-Ketelä, M.; Fallarero, A. *Antimicrob. Agents Chemother.* **2015**, *59*, 6046–6052.
- (21) Okamoto, S.; Taguchi, T.; Ochi, K.; Ichinose, K. *Chem. Biol.* **2009**, *16*, 226–236.
- (22) Ichinose, K.; Ozawa, M.; Itou, K.; Kunieda, K.; Ebizuka, Y. *Microbiology* **2003**, *149*, 1633–1645.
- (23) Ichinose, K.; Bedford, D. J.; Tornus, D.; Bechthold, A.; Bibb, M. J.; Revill, W. P.; Floss, H. G.; Hopwood, D. A. *Chem. Biol.* **1998**, *5*, 647–659.
- (24) Kersten, R. D.; Ziemert, N.; Gonzalez, D. J.; Duggan, B. M.; Nizet, V.; Dorrestein, P. C.; Moore, B. S. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 4407–4416.
- (25) Ramos, J. L.; Martinez-Bueno, M.; Molina-henares, A. J.; Teran, W.; Watanabe, K.; Zhang, X.; Gallegos, M. T.; Brennan, R.; Tobes, R. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 326–356.
- (26) Novakova, R.; Homerova, D.; Feckova, L.; Kormanec, J. *Microbiology* **2005**, *151*, 2693–2706.
- (27) Novakova, R.; Rehakova, A.; Kutas, P.; Feckova, L.; Kormanec, J. *Microbiology* **2011**, *157*, 1629–1639.
- (28) Novakova, R.; Kutas, P.; Feckova, L.; Kormanec, J. *Microbiology* **2010**, *156*, 2374–2383.
- (29) Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Rao, R. N.; Schoner, B. E. *Gene* **1992**, *116*, 43–49.
- (30) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Nat. Protoc.* **2010**, *5*, 536–549.
- (31) Hong, L.; Zhao, Z.; Melanc, C. E.; Zhang, H. *J. Am. Chem. Soc.* **2008**, *130*, 4954–4967.
- (32) Hara, M.; Soga, S.; Shono, K.; Eishima, J.; Mizukami, T. *J. Antibiot. (Tokyo)*. **2001**, *54*, 182–186.
- (33) Taguchi, T.; Okamoto, S.; Hasegawa, K.; Ichinose, K. *ChemBioChem* **2011**, *12*, 2767–2773.
- (34) Hegde, V. R.; King, A. H.; Patel, M. G.; Puar, M. S. *Tetrahedron Lett.* **1987**, *28*, 4485–4488.
- (35) Patel, M.; Hegde, V.; Horan, A.; Barrett, T.; Bishop, R.; King, A.; Marquez, J.; Hare, R.; Gullo, V. *J. Antibiot. (Tokyo)*. **1989**, *42*, 1063–1069.
- (36) Jiang, Z.; Guo, L.; Chen, C.; Liu, S.; Zhang, L.; Dai, S.; He, Q.; You, X.; Hu, X.; Tuo, L.; Jiang, W.; Sun, C. *J. Antibiot. (Tokyo)*. **2015**, *68*, 771–774.
- (37) Taguchi, T.; Yabe, M.; Odaki, H.; Shinozaki, M.; Metsä-Ketelä, M.; Arai, T.; Okamoto, S.; Ichinose, K. *Chem. Biol.* **2013**, *20*, 510–520.
- (38) Tsukamoto, M.; Nakajima, S.; Murooka, K.; Hirayama, M.; Hirano, K.; Yoshida, S.; Kojiri, K.; Suda, H. *J. Antibiot. (Tokyo)*. **2000**, *53*, 26–32.
- (39) Kuniaki, T.; Tomokazu, H.; Masazumi, K.; Yoshikazu, S.; Takashi, O. *J. Antibiot. (Tokyo)*. **2004**, *57*, 291–297.
- (40) Tatsuta, K.; Hosokawa, S. *Sci. Technol. Adv. Mater.* **2006**, *7*, 397–410.
- (41) Taguchi, T.; Itou, K.; Ebizuka, Y.; Malpartida, F.; Hopwood, D. A.; Surti, C. M.; Booker-Milburn, K. I.; Stephenson, G. R.; Ichinose, K. *J. Antibiot. (Tokyo)*. **2000**, *53*,

- 144–152.
- (42) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular cloning: a laboratory manual*; Cold spring harbor laboratory press: New York, 1989.
 - (43) MacNeil, D. J.; Gewain, K. M.; Ruby, C. L.; Dezeny, G.; Gibbons, P. H.; MacNeil, T. *Gene* **1992**, *111*, 61–68.
 - (44) Kieser T, Bibb M, Buttner M, Chater K, H. D. *Practical Streptomyces Genetics*; John Innes Foundation: Norwich, United Kingdom, 2000.
 - (45) Girard, G.; Willemse, J.; Zhu, H.; Claessen, D.; Bukarasam, K.; Goodfellow, M.; van Wezel, G. P. *Antonie Van Leeuwenhoek* **2014**, *106*, 365–380.
 - (46) Colson, S.; Stephan, J.; Hertrich, T.; Saito, A.; van Wezel, G. P.; Titgemeyer, F.; Rigali, S. *J. Mol. Microbiol. Biotechnol.* **2007**, *12*, 60–66.
 - (47) Wu, C.; Zacchetti, B.; Ram, A. F. J.; van Wezel, G. P.; Claessen, D.; Choi, Y. H. *Sci. Rep.* **2015**, *5*, 10868.
 - (48) Wu, C.; Medema, M. H.; Läkamp, R. M.; Zhang, L.; Dorrestein, P. C.; Choi, Y. H.; van Wezel, G. P. *ACS Chem. Biol.* **2016**, *11*, 478–490.
 - (49) Wu, C.; van Wezel, G. P.; Choi, Y. H. *J. Antibiot. (Tokyo)*. **2015**, *68*, 445–452.

Chapter 8

Lugdunomycin, a novel angucycline derivative with unprecedented chemical architecture

Changsheng Wu ^{1,2}, Helga van der Heul ¹, Young Hae Choi ², and Gilles P. van Wezel ¹

¹ *Molecular Biotechnology, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands*

² *Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands*

*Manuscript in preparation
Patent application PCT/NL2016/050398*

ABSTRACT

The discovery of new lead compounds from natural sources is pivotal to replenish the drug discovery pipelines. Here we describe lugdunomycin (**1**), a novel angucycline polyketide bearing unprecedented chemical skeleton produced by *Streptomyces* species QL37. Lugdunomycin shows unique structural characteristics: a heptacyclic 6/6/5/6/6/5/5 ring system, a spiroatom, two all-carbon stereogenic centers, and a benzaza[4,3,3]propellane motif. Detailed metabolic analysis of the producer strain allowed the characterization of in total 31 angucyclines, whereby **5–7**, **10**, **11**, **16–21**, **24**, **26–31** were new compounds. The derivatization mechanisms of the characterized angucyclines **2–31** provided insights to understand the structural complexity in biosynthesis pathway of lugdunomycin (**1**), which involves both enzymatic catalysis and nonenzymatic chemistry, including Baeyer-Villiger oxidative cleavage, amidation, Diels–Alder reaction, intermolecular Michael addition, and nucleophilic Favorskii rearrangement. Metabolite-guided genome mining of *Streptomyces* sp. QL37 combined with genetic disruption identified a type II PKS gene cluster (*lug*) responsible for biosynthesis of compounds **1–31**, among which, *lugOII* was the gene encoding the key oxygenase responsible for the Baeyer-Villiger oxidative cleavage of quinone in angucycline backbone.

1. INTRODUCTION

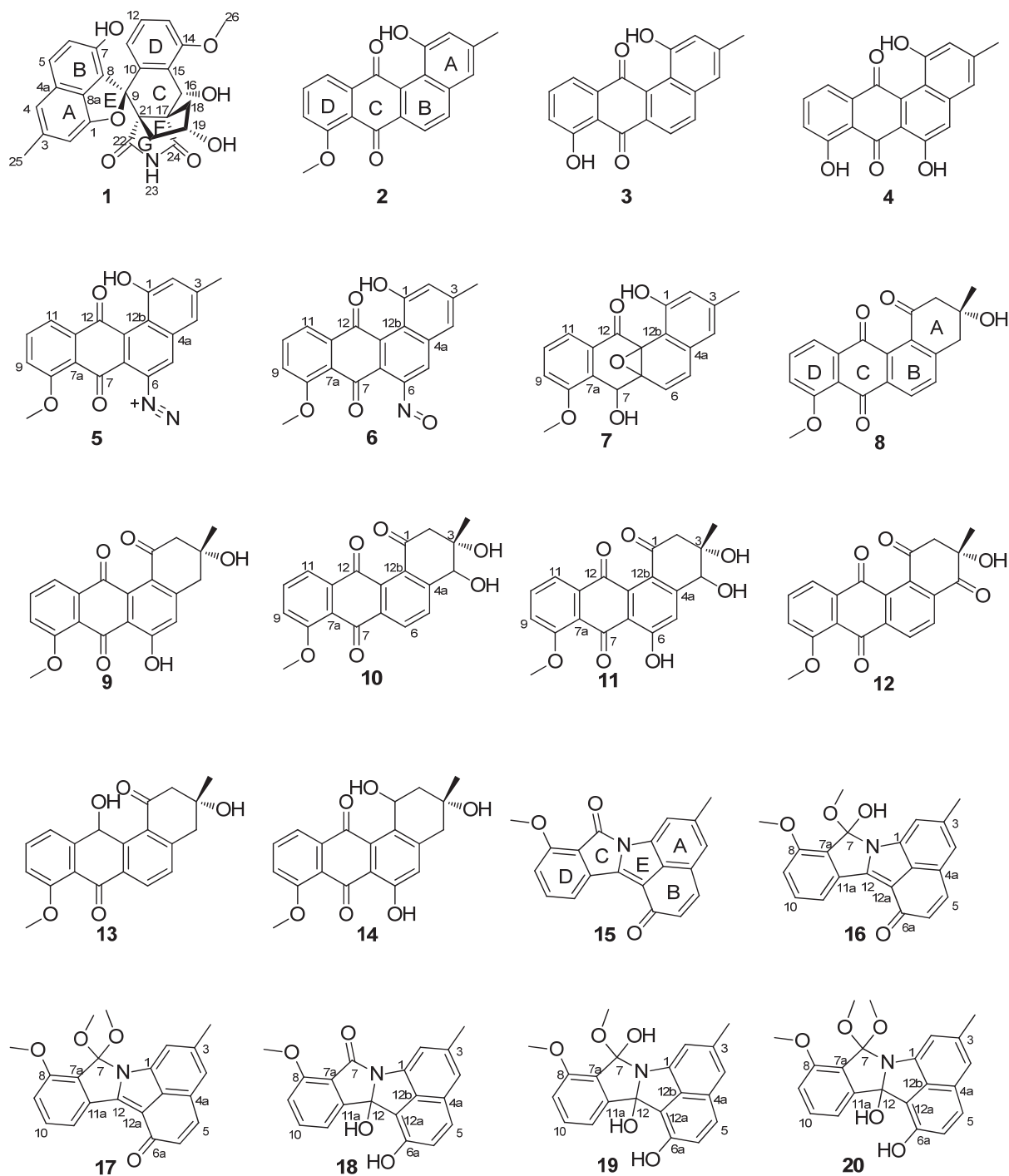
While the twentieth century witnessed the dramatic successes of antimicrobial agents delivered by the pharmaceutical industry, with an explosive growth in the golden era during 1940s to 1960s, the introduction of major new classes of natural product-based antibacterial agents has now come to a near standstill.¹ The approval of the last antibiotic daptomycin in 2003 marked the launch of the first natural product-based antibiotic in decades.^{2,3} The continuous emergence of “untreatable superbugs” threatens to outpace the discovery of new anti-infectives. It is therefore imperative that novel antimicrobials are developed to reverse this negative trend.⁴

Actinomycetes are industrially and medicinally important microorganisms for their ability to produce a plethora of secondary metabolites. More than two-thirds of the known antibiotics are produced by this type of filamentous bacteria, with *Streptomyces* as the major producers.⁵ The advent of next-generation sequencing (NGS) technologies has uncovered that even the best-studied model actinomycete *Streptomyces coelicolor* A(3)2 still possesses many yet underexplored resources for natural products.⁶ As the increasing number of sequenced genomes of actinomycetes reveal a plethora of novel biosynthetic genes, natural product drug discovery is entering an exciting second golden age.⁷ Efforts are therefore required to exploit this huge unexplored reservoir of bioactive natural products.⁸

Angucyclines and angucyclinones bearing an unsymmetrically assembled benz[a]anthracene frame represents the largest family of polycyclic aromatic polyketides from actinomycetes, and exhibit a broad range of biological activities, predominantly anticancer and antibacteria. All angucycline/ones are biosynthesized by type II polyketide synthases (PKSs), via aldol decarboxylative condensations of an acyl-CoA starter and nine malonyl-CoA extenders, to form the initial framework that could be further modified by a wide array of post-PKS tailoring enzymes. The vast structural diversity of angucycline group antibiotics could partially arise from functionality variations, such as hydroxylation, epoxidation, glycosylation at various positions, while the basic angular tetracyclic backbone remains intact. More intriguingly, the referred scaffold could be drastically modified through oxidative ring-cleavages, followed by rearrangement reactions or incorporations of additional carbon units to form new ring(s) system.^{9,10} Therefore, angucycline-based polyketides are a rich family of chemical structures with many different features and bioactivities.

Herein, we report the isolation, structure elucidation, and antimicrobial property of a novel angucycline-derived compound, lugdunomycin (**1**) with striking backbone of benzaza[4,3,3]propellane-6-spiro-2'-2*H*-naphtho[1,8-*bc*]furan, from a soil-derived *Streptomyces* sp. QL37. Application of the OSMAC (one strain many compounds) strategy to this strain allowed the characterization of 30 other rearranged and/or unrearranged angucyclines **2—31**, whereby 18 compounds were new structures featuring unique ring rearrangement, oxidation, and amidation. The structural relationship among compounds **2—31** shed light on the convergent biosynthesis of lugdunomycin (**1**) through a Diels–Alder reaction. Metabolite-guided genome mining of *Streptomyces* sp. QL37 in combination with molecular biology approaches identified a type II PKS gene cluster (*lug*) responsible for the biosynthesis of the characterized angucycline/ones. The new structural features described in this study, especially the C-ring cleavage and expansion, enrich the existent diversity of the angucycline/one type natural products, while the biosynthetic logic for lugdunomycin (**1**)

could be potentially exploited to add chemical diversity and/or pharmaceutical significance to other polyketide skeletons (such as tetracycline antibiotics), either through biochemical or synthetic approaches.



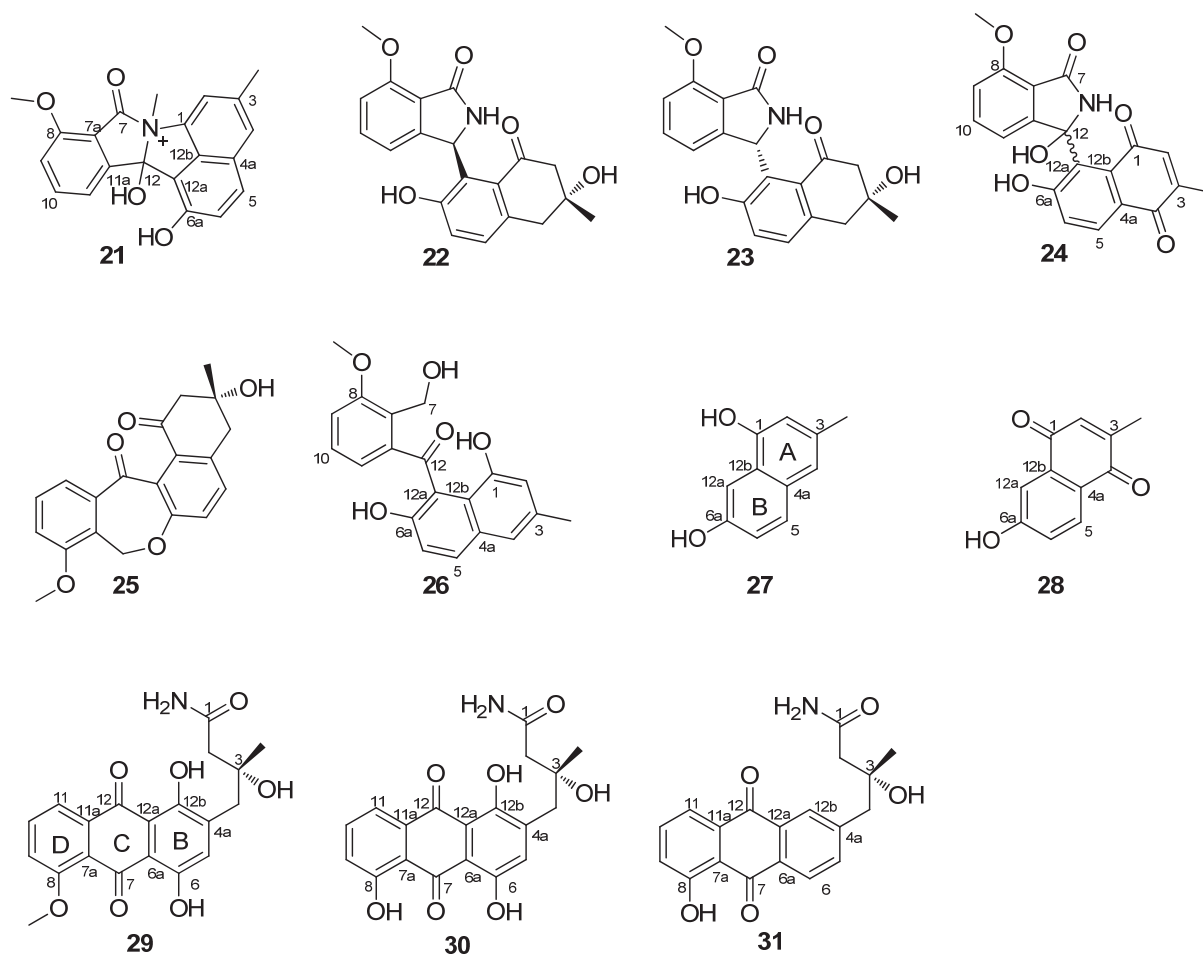


Figure 1. Rearranged and unrearranged angucyclines from *Streptomyces* sp. QL37. Lugdunomycin (**1**) is a novel angucycline derivative with unprecedented skeleton. All compounds are biosynthetically related (see Figure 3), whereby **5–7**, **10**, **11**, **16–21**, **24**, **26–31** with carbon numbering, are novel compounds.

2. RESULTS AND DISCUSSION

2.1. Discovery of lugdunomycin (**1**)

In our search for novel antibiotics, seven strains showing distinctive pigmentation were prioritized out of a unique collection of 816 actinomycetes,¹¹ because the presence of distinctive color is a marker for secondary metabolite biosynthesis.¹² The metabolic profile of a given actinomycete is culture media-dependent, and varying growth conditions or the addition of chemical elicitors can often be applied to activate the biosynthesis of poorly expressed natural products.¹¹ Selected actinomycetes were grown in parallel in six different culture media, and the respective metabolomes were compared by TLC analysis and antimicrobial activity assays against *Bacillus subtilis* 168. Among these, *Streptomyces* sp. QL37 grown in static MM yielded the richest metabolic profile, and was therefore subjected to large-scale (7.5 liter) fermentation. Repeated chromatographic separation of TLC-detectable compounds in the resolved crude extract (2.3 g) resulted in **1** (0.5 mg), **2** (27 mg), **8/9** (27 mg), **10** (3.4 mg), and **13** (1 mg).

Lugdunomycin (**1**) was obtained as a colorless amorphous powder. UHPLC-ToF-MS analysis identified $[M + H]^+$ ion peak at m/z 474.1553 (calculated for $C_{27}H_{24}NO_7$ 474.1547) that established its molecular composition $C_{27}H_{23}NO_7$. This deduced chemical formula was

corroborated by the attached proton test (APT) that exhibited 27 carbons in total. The APT experiment separated 11 carbons of CH and/or CH₃ downwards from 16 carbons of C and/or CH₂ upwards. Protonated carbons were subsequently assigned by HSQC spectroscopy as two methyl groups (one methoxyl), two methylenes, nine methines (seven of which were aromatic units). Based on chemical shifts, 14 quaternary carbons were classified as two carbonyls (δ_C 182.4, and 182.5), nine olefinic carbons (three are oxygenated $\delta_C > 145$, and six are non-oxygenated $\delta_C < 145$), two sp^3 hybridization saturated atoms (δ_C 47.9, 47.3), and one unique *O*-bearing spiroatom at δ_C 94.8 (see below). The elucidation of three aromatic rings (ring A, B, and D) in **1** were readily assigned by the interpretation of the proton splitting pattern and COSY correlations, and further confirmed by HMBC (Table 1 and Figure S1). ¹H NMR spectrum exhibited typical signals attributable to a 1,2,3-trisubstituted phenyl moiety at δ_H 6.92 (brd, $J = 7.8$ Hz, H-11), 7.22 (t, $J = 7.8$ Hz, H-12), 6.96 (brd, $J = 7.8$ Hz, H-13) for ring D. Ring A presented a group of coupling singlets, including two characteristic *meta* benzene singlets at δ_H 6.57 (brs, H-2), 6.95 (brs, H-4) secluded by an olefinic methyl at δ_H 2.46 (brs, H₃-25). Two distinctive aromatic doublets at δ_H 7.39 (d, $J = 8.4$ Hz, H-5), 6.78 (d, $J = 8.4$ Hz, H-6) were assigned to ring B. Benzene rings A and B were fused into a naphthalene system by sharing a double bond between C-4a and C-8a, which was unambiguously confirmed by several HMBC correlations, such as H-4/C-8a, H-5/C-8a, H-4/C-5, H-5/C-4, and H-6/C-4a. However, ring D was located outside the ring A/B system because no HMBC correlations were observed. Adjacent to benzene ring D, ring C was identified by the HMBC correlations from the only proton singlet at δ_H 5.62 (s, H-16) to three aromatic carbons of ring D, via δ_C 139.2 (C-10), 158.4 (C-14), and 123.1 (C-15). On the other hand, the correlations of H-16 with saturated quaternary carbons δ_C 59.8 (C-17), and 62.3 (C-21), demonstrated ring C bridging to other non-aromatic rings. Because ring C was proton-deficient (carrying only a single proton), the continued flow of structure elucidation transferred to a hydrogen-enriched fragment (-CH₂-CH(OH)-CH₂-) that was readily resolved by COSY and HMBC. Key HMBC correlations such as H-19/C-17, H-19/C-21, H-18/C-16, H-20/C-17, and H-20/C-21, *etc* (Table 1), unequivocally demonstrated the presence of cyclopentanol ring F, joined to ring C by sharing the bond between C-17 and C-21. Furthermore, H-18 and H-20 showed ³*J*_{CH} HMBC correlations with two carbonyl groups at δ_C 182.4, and 182.5, respectively, which indicated the presence of another ring fused to two all-carbon quaternary stereocentres¹³ C-17 and C-21 besides rings C and F; these two carbonyls had to be cyclized by a nitrogen atom in line with both the molecular formula and the chemical shifts, which was consistent with succinimide ring G. Consequently, rings D/C/F/G constituted a benzaza[4,3,3]propellane motif which is a molecule with unprecedented chemical topology [2,4-dioxo-7,8-(6-methoxy-benzo)-3-aza[4,3,3]propellan-6,11-diol]. This ring system stereoscopically resembles a propeller: the rings D/C, F, G would be the propeller's blades, and the shared C₁₇-C₂₁ bond would be its axis. The aforementioned ring systems A/B and D/C/F/G were linked at spiroatom C-9 (δ_C 94.8), which was established by the key HMBC correlations H-20/C-9, H-11/C-9, and H-6/C-9, although usual long range couplings (⁴*J*_{CH}) such as H-5/C-1, and H-6/C-9 were considerably misleading (Table 1 and Figure S1). From this we concluded that an additional five-membered furan ring (ring E) was formed to make a 2*H*-naphtho[1,8-*bc*]furan module [7-methy-2*H*-naphtho[1,8-*bc*]furan-3-ol] to join the whole structure. Due to the rigidity of naphthalene, the system of A/B/E was on the same spatial

plane. Taken together, the benzaza[4,3,3]propellane skeleton is adorned with a spirocyclic 2*H*-naphtho[1,8-*bc*]furan moiety and two all-carbon quaternary centers embedded within five contiguous stereogenic carbons. The striking architecture of benzaza[4,3,3]propellane-6-spiro-2'-2*H*-naphtho[1,8-*bc*]furan has a so far unprecedented chemistry.

Table 1. ¹H and ¹³C NMR data for Lugdunomycin (**1**)^a

NO.	δ_{C}	δ_{H} (<i>J</i> in Hz)	HMBC ^b	COSY	NOESY
1	158.8				
2	102.4	6.57 (brs)	C-8a, C-4, C-25	H-4, H-25	H-25
3	136.2				
4	115.7	6.95 (brs)	C-8a, C-2, C-25, C-5	H-2, H-25	H-25
4a	127.4				
5	127.1	7.39 (d, <i>J</i> = 8.4 Hz, 1H)	C-4, C-8a, C-7, C-6, C-1*, C-8*	H-6	
6	123.9	6.78 (d, <i>J</i> = 8.4 Hz, 1H)	C-4a, C-8, C-7, C-9*	H-5	
7	148.2				
8	122.3				
8a	129.1				
9	94.8				
10	139.2				
11	119.7	6.92 (brd, <i>J</i> = 7.8 Hz, 1H)	C-9, C-13, C-15	H-12, H-13	
12	130.9	7.22 (t, <i>J</i> = 7.8 Hz, 1H)	C-10, C-14, C-13, C-11	H-11, H-13	
13	111.1	6.96 (brd, <i>J</i> = 7.8 Hz, 1H)	C-11, C-15, C-14	H-12, H-11	H-26
14	158.4				
15	123.1				
16	62.6	5.62 (s)	C-15, C-14, C-10, C-17, C-18, C-21		H-18b, H-20a
17	59.8				
18	47.9	2.64 (m, H-18a); 1.62 (dd, <i>J</i> = 13.8, 3.6 Hz, H-18b)	C-24, C-19, C-20, C-16 C-24, C-17, C-16, C-19	H-18b H-19, H-18a	H-18b, H-20b H-20a, H-16, H-19, H-18a
19	70.5	4.09 (t, <i>J</i> = 3.6 Hz, 1H)	C-17, C-18, C-20, C-21	H-18b, H-20b	H-18b, H-20b, H-20a
20	47.3	2.63 (m, H-20a) 2.40 (dd, <i>J</i> = 14.4, 3.6 Hz, H-20b)	C-17, C-19, C-18, C-22 C-22, C-9, C-21	H-20b H-19, H-20a	H-19, H-18b H-19, H-18a, H-20a
21	62.3				
22	182.5				
23	-NH				
24	182.4				
25	22.1	2.46 (brs, 3H)	C-2, C-3, C-4	H-2, H-4	H-2, H-4
26	56.1	3.90 (s, 3H)	C-14		H-13

^a **1** recorded in CD₃OD. Proton coupling constants (*J*) in Hz are given in parentheses. ¹H NMR and ¹³C APT NMR spectra were recorded at 600 MHz. All chemical shift assignments were done on the basis of 1D and 2D NMR techniques. ^b All observed HMBC, COSY, and NOESY correlations are summarized, and long range coupling (⁴*J*_{CH}) in HMBC was marked with asterisk (*).

The planar structure was confirmed by single-crystal X-ray diffraction analysis of lugdunomycin (**1**) crystalized from CHCl₃/MeOH 10:1 (Table S1), and the absolute configurations of five chiral centers were accordingly determined as 9*R*, 16*R*, 17*R*, 19*R*, 21*S* (Figure 2). Crystallization of **1** in a chiral space group was initially puzzling, as it exhibited a centrosymmetric space group $P\bar{1}$, suggesting presence of a racemic mixture (Figure S2), but enantiomers were not in a 1:1 ratio. In 3D space, an intramolecular hydrogen-bonding interaction between 16-OH (H acceptor) and 7-OH (H donor), further ensured the rigidification of backbone and fixed geometric conformation. To complicate the story more, we encountered two different crystal forms in the crystallization trial. Unfortunately, one crystal form did not diffract well enough to resolve the structure. Judging by the space group,

these two crystals are diastereomers. The Diels-Alder mechanism (see Figure 3) echoed with the observed stereochemistry of lugdunomycin (**1**) in the crystallization trial. Asymmetrical 3-hydroxypthalimide employing either “*endo*” or “*exo*” transition states can lead to adducts of enantiomeric stereochemistry at four carbons, via C-9, C-17, C-19, and C-21 of lugdunomycin. Electron-withdrawing carbonyls are oriented towards the diene π system in the *endo* orientation, while away from it in the *exo* orientation. It is noteworthy that the “*endo*” is typically favored for rigid dienophile, which explained the unbalanced titer of enantiomers in lugdunomycin crystal. The diastereochemistry implied in crystallization experiment could arise from Michael addition step at C-16, because H₂O attacked the double bond from either rear or front face.

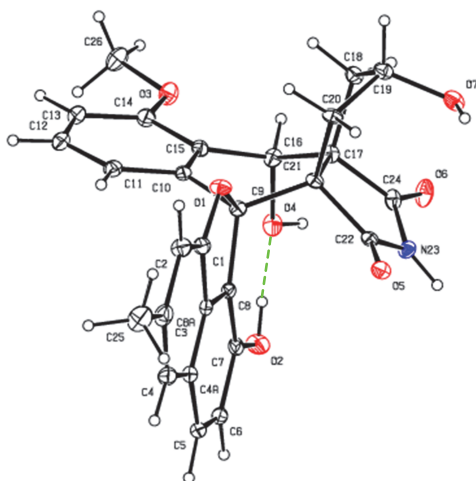


Figure 2. ORTEP drawing of the crystal structure of lugdunomycin (1**).** The dashed green line represents an intramolecular H-bond between two hydroxyl groups (16-OH and 7-OH). The relative configurations of five chiral centers are 9*R*, 16*R*, 17*R*, 19*R*, 21*S*.

2.2. The biosynthetic pathway of lugdunomycin (**1**)

Elucidating the biosynthesis pathway of lugdunomycin (**1**) is challenging, because of the high structural complexity of the molecule. It is not uncommon that a single organism simultaneously produces many structurally related analogues. The structures of the lower order intermediates or side products reflect the way the initial precursor is made, which could provide clues to elucidate the bioassembly mechanism of a higher order final product. In this sense, systematic isolation of all compounds, despite being time-consuming and labor-intensive, is not only efficient to bring to light the chemical diversity of the natural products produced by a given producer organism, but also allows dissecting unfamiliar chemical skeletons in terms of structure elucidation (2D planar connection and even 3D absolute configuration) and the biosynthetic pathway. Herein, lugdunomycin (**1**) was proposed to be an angucycline/one derivative, because i) compound **1** was co-isolated with angucycline/ones **2**, **8**—**10**, and **13**; ii) the structural residue of rings A/B/D in **1** is exactly the same as in **2**. Thereby, **2** was supposed to be the precursor of **1**, but it was difficult to explain the origin of rings C, E—G, especially for the introduction of an unique nitrogen atom into the angucycline/one backbone.

To improve the production of lugdunomycin (**1**) and its structurally related congeners, *Streptomyces* sp. QL37 was fermented in 77 different culture media (Table S2), varied in terms of culturing mode (liquid or solid), source of carbon, phosphate or nitrogen, pH, additional additives, *etcetera*. The EtOAc-resolved metabolome was analyzed by HPLC-UV (detected at 210 nm), which showed that growth on R5 agar plates with 0.8% peptone and 1%

mannitol triggered the production of a considerable number of compounds with a UV spectrum analogous to lugdunomycin (Figure S3). Further UHPLC-ToF-MS analysis confirmed that these compounds indeed contained the sought-after nitrogen atom, though lugdunomycin (**1**) itself was not detected in the mixture. Up-scale refermentation (20 liter) of *Streptomyces* sp. QL37 followed by the extensive systematic isolation, enabled the identification of unrearranged (**2–14**) and rearranged (**15–31**) angucycline/ones (Figure 1), whereby **5–7**, **10**, **11**, **16–21**, **24**, **26–31** were previously undescribed structures. The structure identification was done on the basis of NMR (1D and/or 2D) and UHPLC-ToF-MS analysis in positive and/or negative modes, which were compared with literature spectroscopic data. The 18 new compounds featured among others unique ring rearrangement/cleavage (**15–31**), hydration (**18–21**), methanolysis (**16**, **17**, **19**, **20**), epoxidation (**7**), *N*-quaternary methylation (**21**), and amidation (**15–24**, and **29–31**), which adds further chemical diversity and new members to the family of angucycline antibiotics. Particularly, compounds **5** and **6** possess the rare substituents diazo and nitroso group at C-6, respectively.

The co-identification of compounds **2–31** in *Streptomyces* sp. QL37 provided evidences for the biosynthetic route of lugdunomycin (**1**), as proposed in Figure 3. The type II PKS-catalyzed assembly of the angucycline benz[a]anthracene-7,12-dione in itself has been well documented,^{9,10} whereas the post-PKS modification of angucycline backbone was central to the biosynthesis of lugdunomycin. A pivotal Baeyer–Villiger oxidation at the C-6a/C-7 bond of C ring initiated the structural rearrangements of precursors **2** and/or **8** into **15–31**, whereas the cleavage at C-1/12b bond of A ring generated tricyclic anthraquinone scaffold **29–31**. The oxidation product **2a** was susceptible to hydration to afford **2b** that likely serves as the key intermediate for lugdunomycin (**1**), as well as for all the other rearranged angucyclines **15–31**. On the one hand, the 7-COOH group in **2b** could be furnished by a putative transaminase, and a subsequent intramolecular cyclization between 7-NH₂ and 12-CO in the resultant amidation product **2d** generated a five-membered 2-pyrrolidone ring. However, it is not impossible that the amidation reaction was conducted at the 12-CO instead of 7-COOH. Another oxidative C-C cleavage at the bond C-12/12a of limamycin¹⁴-like intermediate **18a** (or other similar rearranged angucyclines such as **22–24**) resulted in a free 3-hydroxyphthalimide, which could be deduced by the isolation of **27** and **28**. On the other hand, the intermediate **2b** could instead be transformed through intramolecular hemiacetal reaction between 1-OH and 12-CO, followed by an elimination of H₂O and tautomerization to produce an unsaturated aldehyde intermediate **2c**. Convergent coupling of reactive **2c** and 3-hydroxyphthalimide dienophile through Diels–Alder [4+2] cycloaddition constructed a complex system of benzaza[4,4,3]propellane in **1a**. Interestingly, this unconventional mechanism is recently used by Melchiorre et al. to synthesize hydroxy-*o*-quinodimethanes.¹⁵ The α,β -unsaturated carbonyl structural unit in **1a** is susceptible to polar protic solvent nucleophiles like H₂O and MeOH, whereby Michael addition could proceed without heating or use of a catalyst.¹⁶ Consequently, lugdunomycin (**1**) was finally synthesized from **1a** through spontaneous domino reactions, involving a sequence of Michael addition of α,β -conjugated enone with nucleophile reagent H₂O, nucleophilic Favorskii rearrangement, and release of one molecule of CO₂. It is noteworthy that the Favorskii-type rearrangement was also seen in the biosynthesis of the antibiotic enterocin, which was catalyzed by a

Table 2. Gene organization of the angucycline biosynthetic gene cluster (*lug*) in *Streptomyces* sp. QL37.

ORF	Protein	Distribution	AA	Putative function	Nearest homologue	Homology	Accession
1	LugRI	Prokka_02560	379	XRE-family regulator	<i>Streptomyces griseoaurantiacus</i> M045	80%	EGG46652.1
2	LugIU	Prokka_02561	284	ADP-ribose pyrophosphatase	<i>Streptomyces aurantiacus</i> JA 4570	84%	EPH39897.1
3	LugV	Prokka_02562	144	hypothetical protein	<i>Streptomyces aureofaciens</i>	76%	WP_052839114.1
4	LugW	Prokka_02563	199	NADPH-dependent FMN reductase	<i>Streptomyces</i> sp. 303MFCol5.2	85%	WP_020127613.1
5	LugRII	Prokka_02564	280	XRE family transcriptional regulator	<i>Streptomyces sviveus</i>	81%	WP_007383482.1
6	LugRIII	Prokka_02565	223	LuxR family transcriptional regulator	<i>Streptomyces</i> sp. W007	52%	WP_007453015.1
7	LugRIV	Prokka_02566	239	TetR family transcriptional regulator	<i>Streptomyces</i> sp. W007	67%	WP_007453018.1
8	LugTI	Prokka_02567	493	putative export protein	<i>Streptomyces</i> sp. W007	76%	WP_007453020.1
9	LugX	Prokka_02568	144	hypothetical protein	---	---	---
10	LugM	Prokka_02569	346	<i>O</i> -methyltransferase (<i>tcuO</i>)	<i>Streptomyces</i> sp. W007	71%	WP_007453021.1
11	LugOI	Prokka_02570	490	Monooxygenase (<i>urdE</i>)	<i>Streptomyces</i> sp. W007	81%	WP_007453024.1
12	LugF	Prokka_02571	109	polyketide cyclase (<i>urdF</i>)	<i>Streptomyces</i> sp. SCC 2136	80%	CAH10118.1
13	LugA	Prokka_02572	427	polyketide α -ketoacyl synthase II (<i>urdA</i>)	<i>Streptomyces</i> sp. W007	87%	WP_007453026.1
14	LugB	Prokka_02573	407	polyketide β -ketoacyl synthase (<i>urdB</i>)	<i>Streptomyces</i> sp. W007	80%	WP_007453027.1
15	LugC	Prokka_02574	91	Acyl carrier protein (<i>urdC</i>)	<i>Streptomyces</i> sp. W007	71%	WP_007453028.1
16	LugD	Prokka_02575	262	ketoacyl reductase (<i>urdD</i>)	<i>Streptomyces</i> sp. W007	83%	WP_007453030.1
17	LugE	Prokka_02576	316	aromatase	<i>Streptomyces</i> sp. W007	79%	WP_007453031.1
18	LugOII	Prokka_02577	656	monooxygenase (<i>urdM</i>)	<i>Streptomyces</i> sp. W007	76%	WP_007453032.1
19	LugG	Prokka_02578	262	NAD(P)H dependent dehydrogenase	<i>Streptomyces</i> sp. W007	76%	WP_007453033.1
20	LugH	Prokka_02579	527	methylmalonyl-CoA carboxyltransferase	<i>Streptomyces</i> sp. W007	87%	WP_007453034.1
21	LugI	Prokka_02580	79	putative acetyl-CoA carboxylase	<i>Streptomyces rapamycinicus</i>	61%	WP_020868207.1
22	LugJ	Prokka_02581	417	putative MFS-type transporter EfpA	<i>Streptomyces</i> sp. 303MFCol5.2	62%	WP_020130977.1
23	LugK	Prokka_02582	195	NAD(P)H-dependent FMN reductase	<i>Streptomyces fradiae</i>	66%	KDS84998.1
24	LugOIII	Prokka_02583	214	putative monooxygenase	<i>Streptomyces</i> sp. W007	62%	WP_007450404.1
25	LugL	Prokka_02584	245	phosphopantetheinyl transferase	<i>Streptomyces venezuelae</i>	51%	WP_015037371.1
26	LugRV	Prokka_02585	267	transcriptional regulatory protein BaeR	<i>Streptomyces</i> sp. W007	66%	WP_007450396.1
27	LugTII	Prokka_02586	458	MFS transporter	<i>Streptomyces ochraceiscleroticus</i>	76%	WP_051862838.1
28	LugN	Prokka_02587	307	thioesterase	<i>Streptomyces scopuliridis</i>	69%	WP_030349255.1
29	LugOIV	Prokka_02588	275	oxidoreductase	<i>Streptomyces</i> sp. W007	70%	WP_050987713.1
30	LugRVI	Prokka_02589	646	SARP family transcriptional regulator	<i>Streptomyces</i> sp. W007	57%	WP_007450402.1
31	LugOV	Prokka_02590	229	putative dehydrogenase-methyltransferase	<i>Streptomyces</i> sp. W007	57%	WP_007450403.1
32	LugP	Prokka_02591	425	phosphoribosyltransferase	<i>Streptomyces pratensis</i>	90%	WP_014156222.1
33	LugTIII	Prokka_02592	1032	secreted protein	<i>Streptomyces pristinaespiralis</i> ATCC 25486	81%	EDY62729.1
34	LugQ	Prokka_02593	869	hypothetical protein	<i>Streptomyces pristinaespiralis</i>	73%	WP_053557574.1
35	LugS	Prokka_02594	1134	putative Type 1 glutamine amidotransferase	<i>Streptomyces pristinaespiralis</i>	75%	WP_005320663.1
36	LugY	Prokka_02595	1093	hypothetical protein	<i>Streptomyces pristinaespiralis</i>	78%	WP_005320665.1
37	LugZ	Prokka_02596	185	acetyltransferase	<i>Streptomyces</i> sp. W007	72%	WP_032790623.1

Elucidation of the biosynthesis pathway of lugdunomycin (**1**) by the OSMAC strategy shed light on its underlying biosynthetic machinery. Genomic mining of *Streptomyces* sp. QL37 identified a type II polyketide synthase (PKS) gene cluster (*lug*, Table 2 and Figure 3A) displaying a high degree of similarity to PKS gene of *urd*.¹⁸ Genetic inactivation of minimal PKS genes *lugA–C* unambiguously confirmed *lug* was responsible for the production of angucycline/ones in *Streptomyces* sp. QL37 (Figure 4). Within the *lug* biosynthetic gene cluster, five genes encoding putative oxygenases, namely *lugOI–lugOV*, attracted special attention, since multiple post-PKS oxidations were observed in the isolated angucyclines **1–31**, including cleavage of C-C bonds at C-6a/C-7, C-12/12a, and C-1/12b, epoxidation at

C-6a/12a, hydroxylation at C-4 and C-6. The similarity of *lugOI* with *urdE*¹⁹ suggested that this gene is most likely involved in the generation of the *p*-quinone motif to form the basic backbone of angucycline/one. *lugOII* was the best candidate for the C-6a/C-7 Baeyer-Villiger oxygenation, because it shows homologue to *urdM* that is reported to execute this kind of ring oxidative cleavage.²⁰ To test this assumption, we replaced *lugOII* and *lugOIV* genes with apramycin resistance cassette. HPLC-UV analysis (210 nm) confirmed the biosynthesis of rearranged angucyclines **18**, **22**, and **23**, which requires Baeyer-Villiger ring-opening as requisite, was abrogated in null *lugOII* mutant. Null *lugOIV* mutant gave similar HPLC profiling to wild type, excluding its involvement in the C-6a/C-7 bond cleavage. In addition, gene *lugS* harboring a putative amidotransferase domain was proposed to be responsible for the introduction of nitrogen atom into lugdunomycin (**1**) and rearranged angucyclines **15–24**, and **29–31**. Subsequent gene knockout of *lugS* confirmed this assumption, since the production of nitrogen-containing angucyclines **18**, **22**, and **23** was abolished in the mutant, while the biosynthetically upstream unrearranged angucyclines **2**, **8**, **9**, and **13** was not affected.

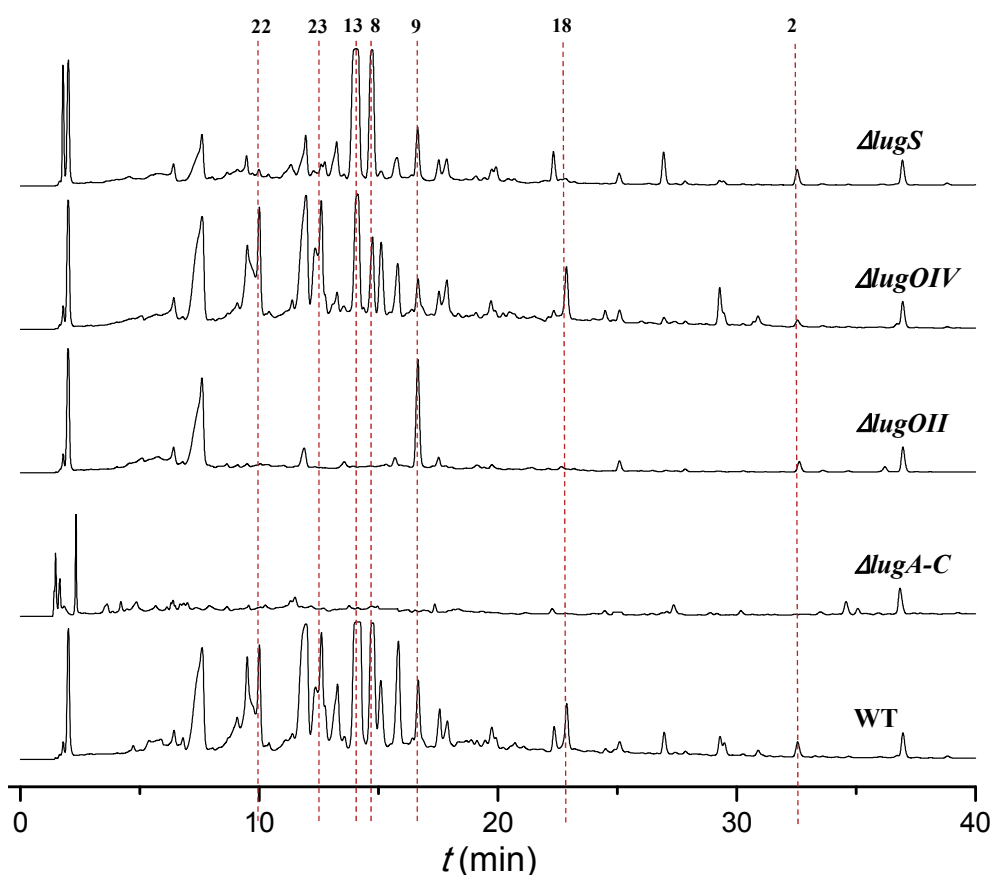


Figure 4. HPLC-UV profiling (210 nm) confirmed functions of selected genes in *lug* cluster. Selected dashed-lined peaks were denoted as compounds with corresponding numbering in Figure 1.

Efforts can be made to improve the supply of two key intermediates for Diels–Alder coupling, via **2b** and 3-hydroxypthalimide. Presumably, *lugOII* could be exploited to increase lugdunomycin production through biochemical transformation (either *in vivo* or *in vitro*) of the abundant precursor **2**. However, it is comparatively more important to overexpress the oxygenase that cleaves the bond C-12/12a, because a considerable amount of limamycins

were already seen in cultures grown on R5 agar plates with peptone and mannitol, implying that the Baeyer–Villiger cleavage of C-6a/12a was not the rate-limiting step for lugdunomycin (**1**) synthesis. The removal of *lugM* encoding an *O*-methyltransferase probably elevate the production of lugdunomycin variant (26-*O*-demethylated lugdunomycin), because the methyl group blocks the proton tautomerization of 3-hydroxyphthalimide into its reactive α,β -unsaturated ketone conformation, and thus hinder the key Diels–Alder reaction for the construction of benzaza[4,4,3]propellane motif in **1a** (Figure 3). Alternatively, the addition of commercial 3-hydroxyphthalimide into the culture medium seems more convenient. Taken together, we propose that 1) the biosynthetic route of lugdunomycin, could be used to guide *in situ* optimization of lugdunomycin production in *Streptomyces* sp. QL37. For instance, supplementation of Lewis acid(s) into culture medium to promote Diels–Alder reaction; 2) Lugdunomycin can be constructed through *in vitro* enzymatic Baeyer–Villiger oxidation to generate reactive intermediate, in tandem with semisynthetic method involving Diels–Alder coupling with 3-hydroxyphthalimide; 3) By varying the dienophile reagent for Diels–Alder reaction, such as maleimide, *N*-hydroxyphthalimide, maleic anhydride, *etc*, angucycline-based lugdunomycin variants could be potentially synthesized to afford a series of variants of lugdunomycin (Figure S4); 4) The post-PKS modification mode in lugdunomycin biosynthesis, via oxidative quinone ring opening followed by Diels–Alder coupling of extra structural unit at the reactive aldehyde site, can be used to modify other PKS-type natural products, allowing the production of industrially and medicinally important chemicals based on angucyclines, but also other PKS like the linear tetracycline;

2.3. Antimicrobial properties of lugdunomycin (**1**)

Selected compounds displayed in Figure 1 were tested for antimicrobial activity in an agar diffusion assay, using the Gram-positive *Bacillus subtilis* 168 and the Gram-negative *Escherichia coli* K12 as indicator strains. All tested compounds showed significant antimicrobial activity. Lugdunomycin (**1**) showed antimicrobial activity against *B. subtilis*, but not against *E. coli* (Figure 5). More compound is required to carry out detailed bioactivity assays as well as cytotoxicity and mode of action studies.

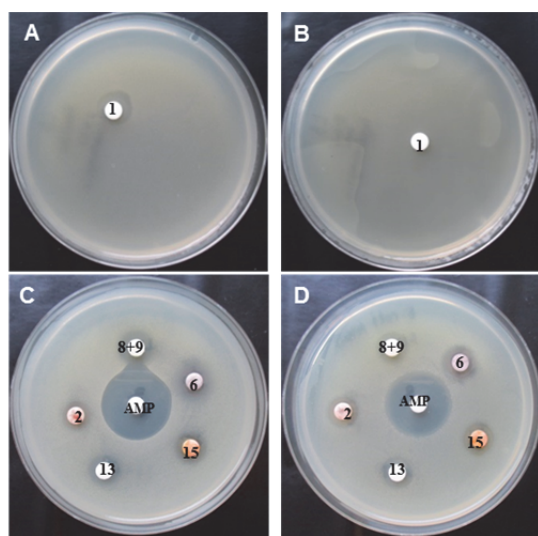


Figure 5. Preliminary antimicrobial activity test of lugdunomycin (1**) and selected angucyclines.** Numbering on the paper disc corresponds to the numbering of the compounds as presented in Figure 1, e.g. **1** stands for lugdunomycin, **6** for the novel angucycline derivative with nitroso group. Lugdunomycin (**1**) showed inhibition efficacy against Gram-positive bacterium *Bacillus subtilis* (A), but not against Gram-negative *Escherichia coli* JM 109 (B); All the selected angucyclines simultaneously inhibited the growth of both *Bacillus subtilis* (C) and *Escherichia coli* JM109 (D). Note: compounds **8** and **9** were obtained as a mixture, which were tested activity together; AMP: positive control, ampicillin.

3. CONCLUSION

Though synthetic biology approaches tend to dominate the industry of microbial natural products (NPs) discovery,²¹ they appear futile in handling intertwined metabolic pathways, such as one biosynthetic gene cluster (BGC) many NPs, many BGCs one NP, and many BGCs many NPs. The OSMAC (one strain many compounds) strategy can complement the genomics-based strain prioritization through direct assessment of all (related) natural products produced under different culturing conditions. Herein, OSMAC applied in *Streptomyces* sp. QL37 allowed the discovery of lugdunomycin (**1**), a novel angucycline-derived molecule harboring a unparalleled backbone of benzaza[4,3,3]propellane-6-spiro-2'-2*H*-naphtho[1,8-*bc*]furan. Tens of co-isolated rearranged and unrearranged angucycline/ones illuminated the “biosynthetic black box” of the unprecedented architectural complexity in lugdunomycin, which also shed light on the characterization of their responsible genetic basis, a type II polyketide synthase gene cluster (*lug*). The genetic and synthetic knowledge revealed in the elucidated biosynthetic pathway for lugdunomycin, especially the Baeyer–Villiger oxidative cleavage and convergent Diels–Alder coupling, is anticipated to not only guide the future up-scale production of lugdunomycin and its potential variants, but also expand the existent structural diversity of known polyketides, which will have profound impact on new drugs development.

4. EXPERIMENTAL SECTION

4.1. General Experimental Procedures

FT-IR was measured on Perkin-Elmer FT-IR Spectrometer Paragon 1000. UV measurements were performed using a Shimadzu UV mini-1240. NMR spectra were recorded on a Bruker DMX 600 MHz, either in methanol-*d*₄ calibrated to 3.30 ppm or CDCl₃ calibrated to 7.27 ppm. Semi-preparative HPLC (pHPLC) was performed with a Shimadzu HPLC system and a 5 ml Rheodyne manual injection loop, equipped with a reversed-phase C₁₈ column (Phenomenex Luna C₁₈ (2) 100 Å 5 micron 250 × 10 mm). All the pHPLC experiments used 2 ml/min flow rate and fraction collection based on detected peak. Silica gel (pore size 60 Å, 230–400 mesh) for open column chromatography was purchased from Sigma-Aldrich (St. Louis, MO, USA). Pre-coated silica gel 60 F₂₅₄ TLC plates (Merck, Darmstadt, Germany) were used for TLC bioautography analysis. PLC silica gel 60 F₂₅₄, 1 mm (Merck, Darmstadt, Germany) was used for preparative TLC separation. Analytical TLC was performed with silica gel 60 (Merck, Darmstadt, Germany) plates and visualized with UV lamp (254 nm and 365 nm) and anisaldehyde/sulfuric acid reagent. All solvents and chemicals were of analytical and HPLC grade.

4.2. Microorganisms and culturing conditions

Streptomyces sp. QL37 was isolated from soil in the Qinling mountains (P. R. China) as described previously¹¹. The strain was deposited to the collection of the Centraal Bureau voor Schimmelcultures (CBS) in Utrecht, The Netherlands. *Streptomyces* sp. QL37 was cultivated on minimal media agar plates (MM)²² with 0.5% glycerol and 1% mannitol (w/v) as the carbon sources, and R5 culture medium supplemented with 0.8% peptone and 1% mannitol (w/v). Square agar plates (12 cm × 12 cm) were inoculated with 1 × 10⁷ spores from a fresh

spore suspension. As indicator strains for the antibacterial assays we used *Bacillus subtilis* 168 and *Escherichia coli* JM109.

4.3. Extraction and Isolation of lugdunomycin (1)

The first round of systematic separation was done on MM culture medium. After seven days of growth, 225 MM agar plates inoculated with *Streptomyces* sp. QL37 were combined and cut into small blocks, which were then homogenated with a pestle. The resultant agar suspension was extracted with ethyl acetate (EtOAc) by soaking overnight at room temperature. The supernatant was filtered and subsequently evaporated under reduced pressure at 38 °C to obtain 2.3 g crude extract. This extract was adsorbed by silica gel and chromatographed on a silica gel (pore size 60 Å, 70–230 mesh, St. Louis, MO, USA) column chromatography employing a gradient elution from *n*-hexane via chloroform to methanol. The combinations of fractions derived from silica gel chromatography separation were done on the basis of thin-layer chromatography (TLC) analysis (Merck, Darmstadt, Germany) using developing solvent system of chloroform and methanol (10:1). The fractions containing lugdunomycin (**1**) were combined by TLC detection that gave a dark spot under UV at 254 nm and a distinctive blue color when further stained with anisaldehyde/sulfuric acid reagent by heating. Purification of lugdunomycin was first defatted by partition between methanol and *n*-hexane, and the resulting methanol phase was subsequently rechromatographed on Sephadex LH-20 column (GE Healthcare Life Sciences, Eindhoven, The Netherlands) eluting by methanol to give five subfractions. The fraction containing lugdunomycin (**1**) was purified by preparative TLC (Merck, Darmstadt, Germany), migrated with solvent system chloroform and methanol (9:1) and detected under UV light at 254 nm. Finally, 0.5 mg pure lugdunomycin (**1**) was obtained. Compounds **2** (3.4 mg), **8/9** (27 mg), **10** (3.4 mg), and **13** (1 mg) were isolated from the same MM agar plates as well.

The second round of separation was done on R5 culture media supplemented with 1% mannitol and 0.8% peptone. A total volume of 20 L R5 agar media was used. The subsequent fermentation and extraction methods were the same with MM culture media. UV and MS-guided separation was used to accumulate the amount of lugdunomycin (**1**) and isolate lugdunomycin analogues. Crude extract (20.5 g) adsorbed by silica gel was first chromatographed on a silica gel column chromatography employing gradient elution from *n*-hexane, chloroform, to methanol, to obtain 17 fractions. These were subsequently subjected to HPLC-UV and UHPLC-ToF-MS analysis, the combination of which was further done by UV spectrum and chemical formula. This identified among others the presence or absence of a nitrogen atom. As a result, 12 final fractions (Fr1–Fr12) were obtained. The fractions (Fr1, Fr4) abundant in already identified compounds **2**, **8–10** in MM medium were discarded. Fr2 was successively chromatographed on silica gel eluting with a gradient of chloroform in *n*-hexane, and Sephadex LH-20 eluting with methanol, to afford pure compound **3** (2.5 mg). Fr3 was separated by silica gel eluting with a gradient of chloroform in *n*-hexane, to give 8 subfractions sfr3.1–sfr3.8. Sfr3.2 was separated by Sephadex LH-20, to afford the pure semi-pure purple compound **6** (0.6 mg); Sfr3.3 was purified by Sephadex LH-20, to afford the pure orange compound **15** (20 mg); Sfr3.5 was firstly separated by Sephadex LH-20 and followed by preparative TLC (PLC Silica gel 60 F₂₅₄, 1 mm, Merck, Darmstadt, Germany), migrated with solvent system of chloroform/methanol (10:1), to afford the semi-pure black

compound **5** (0.6 mg). Fr5 was chromatographed on silica gel eluting with a gradient of methanol in chloroform, to give 5 subfractions sfr5.1–sfr5.5. Sfr5.2 was separated by Sephadex LH-20 to afford pure compound **18** (10 mg); Sfr5.4 was separated by semi-preparative reversed-phase HPLC (Phenomenex Luna C18 (2) 100 Å 5 micron 250 × 10 mm) on an Agilent 1200 series HPLC (Agilent technologies Inc, Santa Clara, CA, USA), eluting with a gradient of ACN in H₂O from 15% to 80% in 40 min. HPLC peaks were manually collected, resulting in the isolation of compound **10** (2.8 mg), **11** (impure, 1.0 mg), **24** (impure, 0.50 mg), **29** (semi-pure, 0.82 mg), **28** (semi-pure, 1.1 mg), **31** (semi-pure, 0.90 mg), and **30** (impure, 0.68 mg). Fr6 was directly separated by semi-preparative HPLC, eluting with a gradient of ACN in H₂O from 15% to 80% in 40 min, which resulted in the separation of compound **14** (semi-pure, 0.80 mg), **7** (semi-pure, 0.40 mg), **25** (semi-pure, 0.30 mg), a mixture of **27** and **28** (semi-pure, 0.60 mg), and **26** (0.50 mg). Fr7 was directly separated by semi-preparative HPLC, eluting with a gradient of ACN in H₂O from 20% to 85% in 30 min, which resulted in the isolation of **16** (semi-pure, 0.50 mg), **19** (semi-pure, 0.40 mg), and **21** (semi-pure, 0.60 mg). Fr9 was separated by silica gel eluting with a gradient of methanol in chloroform, to give 6 subfractions sfr9.1–sfr9.6. Sfr9.3 was purified by preparative TLC, migrated with solvent system of chloroform/methanol (5:1), to afford the pure compound **22** (5 mg) and **23** (5.5 mg).

Lugdunomycin (**1**): colorless, needle crystal, UV (MeOH) $\lambda_{\max}$ (log ϵ) 349 (3.32), 287 (4.06), 250 (4.42) nm; ECD (MeOH) $\lambda_{\max}$ ($\Delta\epsilon$) 204 (+4.49), 227 (−4.04), 251 (+0.92) nm; IR $\nu_{\max}$ 669, 831, 1130, 1203, 1271, 1471, 1558, 1683, 1716, 2324, 2349, 2378, 2850, 2920 cm^{−1}; ¹H and ¹³C NMR data, see Table 1; HRESIMS (positive mode) m/z 456.14481 [M+H−H₂O]⁺ (calcd for C₂₇H₂₂NO₆ 456.14526), 474.15530 [M+H]⁺ (calcd for C₂₇H₂₄NO₇ 474.15473).

4.4. X-ray Crystal Structure Determination of lugdunomycin (**1**).

Colorless crystals of lugdunomycin (**1**) were obtained from a solvent system of CHCl₃/MeOH (10:1). Suitable single crystals for X-ray structural analysis of lugdunomycin were mounted at room temperature in Paratone N inert oil. Single crystal X-ray diffraction data were collected in the home laboratory on a Bruker three circle diffractometer equipped with a Bruker TXS-Mo rotating anode, INCOATEC mirror optics and an APEX II detector. The data were integrated with SAINT and an empirical absorption correction with SADABS²³ was applied. The structure was solved by direct methods.²⁴ The structure model was refined against all data by full-matrix least-squares methods on F^2 with the program shelxl2014.²⁵ All non-hydrogen-atoms were refined with anisotropic displacement parameters. The hydrogen atoms were refined isotropically on calculated positions using a riding model with Uiso values constrained to 1.2/1.5 U_{eq} of their parent atoms.

4.5. Antimicrobial activity assays

Antimicrobial activity was determined using a disk diffusion assay. For this, compounds were dissolved in chloroform to a concentration of (2 µg/µl) and of this solution, 25 µl was applied on a paper disk (GE Healthcare Bio-Sciences, Pittsburgh), except for lugdunomycin, for which 10 µl of a 1µg/µl solution in methanol was used. The disks were then placed onto an LB agar plate overlaid with 3 ml of soft agar (LB with 0.6% (w/v) agar) containing around 5×10⁷ cells of exponentially growing *Bacillus subtilis* 168 or *Escherichia coli* JM109 cells.

Ampicillin was used as positive control. After incubation at 37 °C for 18 h, growth inhibition zones (in mm) were recorded as antimicrobial activity.

4.6. Gene knockouts of *lugA—C*, *lugOII*, *lugOIV*, and *lugS*

The methodology for disruption of target genes was based on homologous recombination as illustrated in Figure S5. In general, ~1.5 kb of left and right flanks of target gene were respectively amplified from the *Streptomyces* sp. QL37 genome using the primers listed in Table S3. After digestion with restriction enzymes, the left and right flanks were linked through ~1.0 kb apramycin resistance cassette *aac(3)IV* engineered by enzyme XbaI. The obtaining ~4.0 kb fragment, via left flank—apramycin resistance cassette—right flank, was cloned into the unstable multi-copy vector pWHM3²⁶ incorporated with *oriT* sequence, to generate the knockout construct. The integrity of the construct was verified by sequencing and restriction enzyme analysis. The correct knockout construct was transformed to the methylase-deficient strain *E. coli* ET12567/pUZ8002, which was further introduced into *Streptomyces* sp. QL37 by conjugation, following the protocol as described.²⁷ The correct mutant going through double crossover was resistant to 50 µg/mL apramycin and sensitive to 10 µg/mL thiostrepton. The apramycin resistance cassette *aac(3)IV* flanked by loxP sites was removed from the chromosome following the introduction of plasmid pUWLCRE that expresses the Cre recombinase.²⁸

4.7. UHPLC-ToF-MS analysis

UHPLC-ToF-MS analyses were performed on an UHPLC system (Ultimate 3000, ThermoScientific, Germany) coupled to an ESI-IIQ-TOF spectrometer (microTOF-QII, Bruker Daltonics, Germany) in the positive mode. The chromatographic separation was done using a Kinetex C₁₈ UHPLC 2.6 µm particle size column 150 × 2.0 mm (Phenomenex, USA) at a flow rate of 0.3 mL/min and a column temperature of 30 °C. Samples (3 µL) were eluted using a gradient of solvent A (water) and B (acetonitrile), both with 0.1% formic acid (v/v). The initial percentage of B was 5%, which was linearly increased to 90% in 19.5 min, followed by a 2 min isocratic period and, then re-equilibrated with original conditions in 2 min. Nitrogen was used as drying and nebulizing gas. The gas flow was set at 10.0 L/min at 250 °C and the nebulizer pressure was 2.0 bar. The MS data were acquired over *m/z* range of 100–1000. The capillary voltage was 3.5 kV. For internal calibration, a 10 mM solution of sodium formate (Fluka, Steinheim, Germany) was infused. Formic acid, water and acetonitrile were LCMS grade (Optima, Fisher Scientific, NJ, USA).

4.8. HPLC-UV analysis

HPLC analysis was performed with an Agilent 1200 series HPLC apparatus (Agilent technologies Inc, Santa Clara, CA, USA), using a 150 × 4.6 mm Luna 5 micron C18 (2) 100 Å column equipped with a guard column containing C18 4 × 3 mm cartridges (Phenomenex Inc, Torrance, CA, USA). The mobile phase consisted of water (A) and acetonitrile (B, HPLC grade) in a linear gradient program from 10% B to 100% B in 50 minutes at a flow rate of 1.0 ml/min. Chromatograms were recorded at 210 nm, 254 nm, and 280 nm. The injection volume was 10 µl.

REFERENCES

- (1) WHO Geneva, Switzerland, 2014.
- (2) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat. Rev. Drug Discov.* **2007**, *6*, 29–40.
- (3) Raja, A.; LaBonte, J.; Lebbos, J.; Kirkpatrick, P. *Nat. Rev. Drug Discov.* **2003**, *2*, 611–612.
- (4) Lewis, K. *Nat. Rev. Drug Discov.* **2013**, *12*, 371–387.
- (5) Baltz, R. H. *Curr. Opin. Pharmacol.* **2008**, *8*, 557–563.
- (6) Bentley, S. D.; Cerdeño-Tárraga, K. F. C. A.-M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Bentley, S. D.; Harper, D.; Bateman, A.; Brown, S.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Larke, L.; Murphy, L.; Oliver, K.; Rabinowitsch, E.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Woodward, J.; Barrell, B. G.; Parkhill, J. *Nature* **2002**, *3*, 141–147.
- (7) Zarins-Tutt, J. S.; Barberi, T. T.; Gao, H.; Mearns-Spragg, A.; Zhang, L.; Newman, D. J.; Goss, R. J. M. *Nat. Prod. Rep.* **2016**, *33*, 54–72.
- (8) Cooper, M. A.; Schlaes, D. *Nature* **2011**, *472*, 32.
- (9) Rohr, J.; Thiericke, R. *Nat. Prod. Rep.* **1992**, *9*, 103–37.
- (10) Kharel, M. K.; Pahari, P.; Shepherd, M. D.; Tibrewal, N.; Nybo, S. E.; Shaaban, K. a; Rohr, J. *Nat. Prod. Rep.* **2012**, *29*, 264–325.
- (11) Zhu, H.; Swierstra, J.; Wu, C.; Girard, G.; Choi, Y. H.; van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. *Microbiology* **2014**, *160*, 1714–1725.
- (12) Brady, S. F.; Chao, C. J.; Handelsman, J.; Clardy, J. *Org. Lett.* **2001**, *3*, 1981–1984.
- (13) Minko, Y.; Pasco, M.; Lercher, L.; Botoshansky, M.; Marek, I. *Nature* **2012**, *490*, 522–6.
- (14) Fotso, S.; Mahmud, T.; Zabriskie, T. M.; Santosa, D. A.; Proteau, P. J. *J. Antibiot. (Tokyo)*. **2008**, *61*, 449–456.
- (15) Dell’amico, L.; Vega-Peñaloza, A.; Cuadros, S.; Melchiorre, P. *Angew. Chemie - Int. Ed.* **2016**, *55*, 3313–3317.
- (16) Cao, C. M.; Zhang, H.; Gallagher, R. J.; Timmermann, B. N. *J. Nat. Prod.* **2013**, *76*, 2040–2046.
- (17) Teufel, R.; Miyanaga, A.; Michaudel, Q.; Stull, F.; Louie, G.; Noel, J. P.; Baran, P. S.; Palfe, B.; Moore, B. S. *Nature* **2013**, *503*, 552–556.
- (18) Faust, B.; Hoffmeister, D.; Weitnauer, G.; Westrich, L.; Haag, S.; Schneider, P.; Decker, H.; Künzel, E.; Rohr, J.; Bechthold, A. *Microbiology* **2000**, *146*, 147–154.
- (19) Decker, H.; Haag, S. *J. Bacteriol.* **1995**, *177*, 6126–6136.
- (20) Rix, U.; Remsing, L. L.; Hoffmeister, D.; Bechthold, A.; Rohr, J. *ChemBioChem* **2003**, *4*, 109–111.
- (21) Smanski, M. J.; Zhou, H.; Claesen, J.; Shen, B.; Fischbach, M. A.; Voigt, C. A. *Nat. Rev. Microbiol.* **2016**, *14*, 135–149.
- (22) Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. John Innes Foundation, Norwich, UK, 2000.
- (23) George M. Sheldrick *SADABS, Program for Empirical Absorption Correction of Area*

- Detector Data*; University of Göttingen: Göttingen, Germany, 1996; Vol. 218.
- (24) Schneider, T. R.; Sheldrick, G. M. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2002**, 58, 1772–1779.
 - (25) Sheldrick, G. M. *Acta Crystallogr. A.* **2008**, 64, 112–122.
 - (26) Vara, J.; Lewandowska-Skarbek, M.; Wang, Y. G.; Donadio, S.; Hutchinson, C. R. *J. Bacteriol.* **1989**, 171, 5872–5881.
 - (27) Kieser T, Bibb M, Buttner M, Chater K, H. D. *Practical Streptomyces Genetics*; John Innes Foundation: Norwich, United Kingdom, 2000.
 - (28) Fedoryshyn, M.; Welle, E.; Bechthold, A.; Luzhetskyy, A. *Appl. Microbiol. Biotechnol.* **2008**, 78, 1065–1070.

Chapter 9

Identification of novel endophenaside antibiotics produced by *Kitasatospora* sp. MBT66

Changsheng Wu ^{1,2}, Gilles P. van Wezel ¹, and Young Hae Choi ²

¹ *Molecular Biotechnology, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands*

² *Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands*

J. Antibiot. (Tokyo). **2015**, 68, 445–452.

ABSTRACT

Actinomycetes are a major source of bioactive secondary metabolites and are a focal point in the search for novel antimicrobial compounds that are needed to combat multidrug-resistant pathogens. Here we report the discovery of several novel phenazine-type antibiotics produced by *Kitasatospora* sp. MBT66. These include the novel glycosylated endophenazines A–E (**1–5**), together with *N*-prenylated endophenazine F1 (**6**). Compounds **1** and **3** contain a 2'-*O*-methylation of the sugar moiety, which is rare in nature and reported for the first time in connection with phenazines. The structures of the new compounds were determined on the basis of their spectral data, including NMR and HR-MS, and the gene cluster responsible for the biosynthesis of phenazines was identified. All phenazine derivatives showed antimicrobial activity against the Gram-positive *Bacillus subtilis*, while compounds **1–3** and **5** also inhibited growth of the Gram-negative *Escherichia coli*.

1. INTRODUCTION

Infections caused by multiple drug-resistant bacteria are a major threat to human health, both in the nosocomial environment and in the general community.^{1–3} It has become increasingly difficult to find novel antibiotics with efficacy against multidrug-resistant pathogens such as MDR-TB (multidrug-resistant *Mycobacterium tuberculosis*), MRSA (methicillin-resistant *Staphylococcus aureus*), and multidrug-resistant Gram-negative pathogens.³ Indeed, despite intensive screening programs in the pharmaceutical industry, hardly any new drugs have been identified in this millenium, which underlines the urgent need for novel approaches to find antimicrobial drugs.⁴ Actinomycetes are Gram-positive mycelial bacteria with a complex multicellular life style.^{5,6} They are industrially highly relevant due to their ability to produce a plethora of natural products, including anticancer, antifungal, anthelmantic and antibacterial drugs.⁷ Of all known naturally occurring antibiotics, some two-third is derived from actinomycetes, the majority of which are produced by streptomycetes.⁸ Sequencing of *Streptomyces* genomes established the presence of silent antibiotic biosynthetic gene clusters, suggesting that the potential of these organisms for novel drug production is much larger than originally anticipated.⁹ As an example, many novel gene clusters were identified even in the extensively studied model streptomycetes *S. coelicolor*,¹⁰ *S. griseus*¹¹ and *S. lividans*.¹² This has led to novel approaches to identifying specific conditions or agents that elicit the expression of antibiotics.^{13–15}

In a recent study, we screened a collection of over 800 actinomycetes using 40 different growth conditions, which identified among others *Kitasatospora* sp. MBT66 that was shown to be a prolific producer of antibiotics with activity against Gram-positive and Gram-negative pathogens.¹⁶ Preliminary analysis predicted that it produced a number of compounds exhibiting similar chemical properties (TLC and HPLC-UV detection) to 1-carbomethoxy phenazine that was described in another actinomycete *Streptomyces* sp. MBT70.¹⁶ The phenazines are a group of heterocyclic, nitrogen-containing compounds isolated primarily from *Pseudomonas* and *Streptomyces* bacteria.¹⁷ They differ in chemical and physical properties according to the nature and/or position of functionalities, and thus give rise to a full spectrum of colours ranging from the deep red to the light blue.¹⁸ Phenazines are redox-active and can reduce molecular oxygen, leading to the generation of toxic reactive oxygen species and explaining their broad-spectrum antibiotic activity toward bacteria, fungi, and plant and animal tissues.^{19,20} Clofazimine is an example of a phenazine compound that is successfully applied in the clinic, where it is used for the treatment of leprosy, while clofazimine also has efficacy against MDR *Mycobacterium tuberculosis*.²¹ In addition, phenazines play multiple roles in the producing bacteria for survival and interaction with the environment, such as protection of plants by rhizobacteria against bacterial and fungal pathogens,²² reducing environmental Fe³⁺ to the more soluble Fe²⁺ to aid iron acquisition,²³ and in primary energy metabolism by mediating the reoxidation of NADH and utilization of pyruvate under oxygen-limiting conditions.²⁴ Although more than 6000 phenazine-containing compounds have been described over the past century, fewer than 100 are of natural origin.²⁵

Very few glycosylated phenazines have so far been found in nature, and these are without exception compounds with bioactivity that are derived from 6-deoxy-L-glycopyranosides.²⁰ In this work, we report the isolation, structure determination and antimicrobial activity of novel

glycosylated endophenazines (**1–5**) and an *N*-prenylated endophenazine (**6**) produced by *Kitasatospora* species MBT66, and present the gene cluster responsible for their biosynthesis.

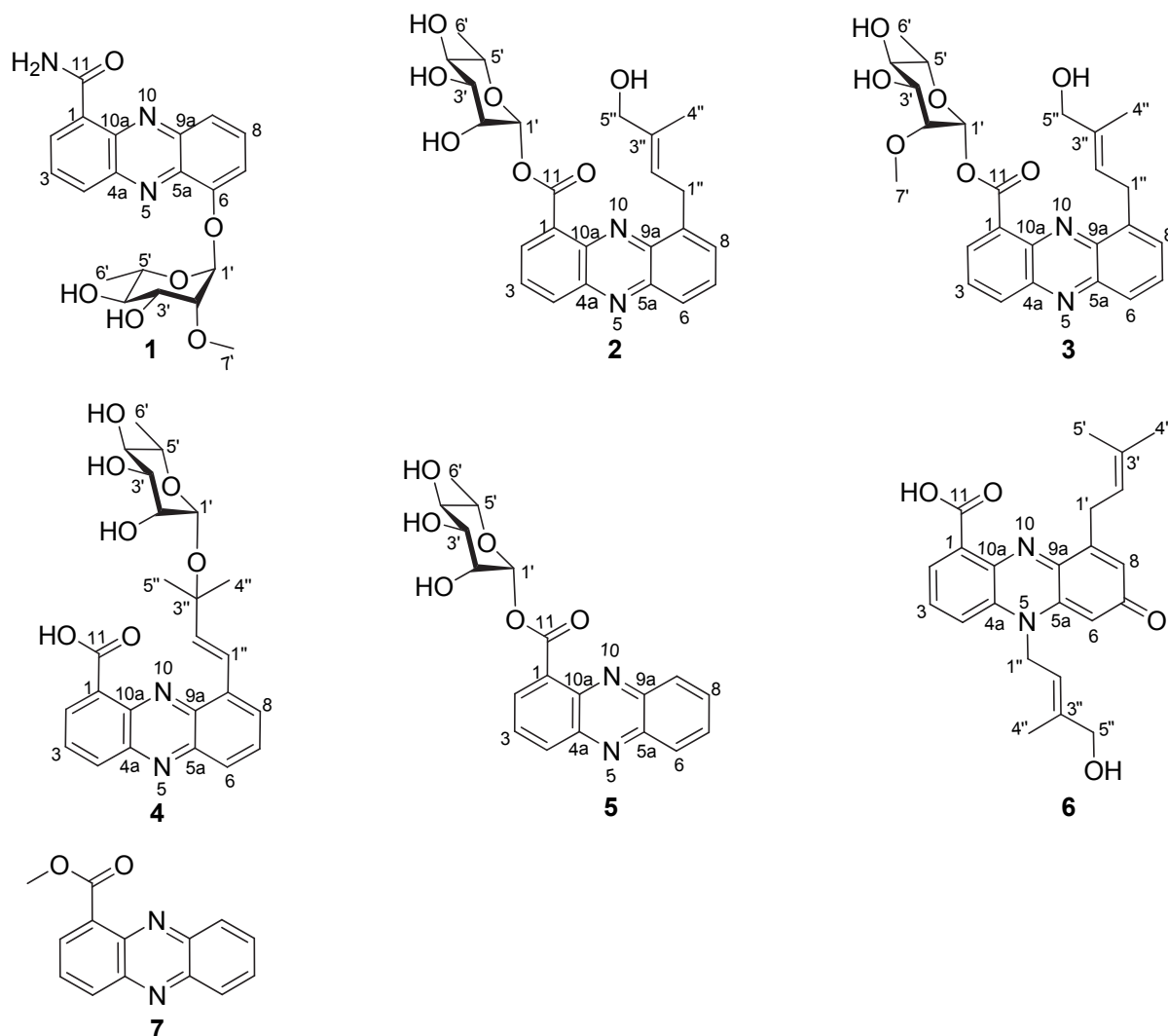


Figure 1. Phenazine derivatives from *Kitasatospora* sp. MBT66. Compounds **1–5** are glycosylated endophenazines, compound **6** is an *N*-prenylated endophenazine. Compounds **1**, and **3** were methylated at the 2'-*O* position of the rhamnose moiety.

2. RESULTS AND DISCUSSION

We previously performed a screen of a collection of some 800 actinomycetes obtained from mountain soils, so as to identify optimal laboratory conditions for antimicrobial secondary metabolites production.¹⁶ *Kitasatospora* sp. MBT66²⁶ was one of our 96 most prolific antibiotic-producing identified actinomycetes,¹⁶ which was selected for further phytochemical investigation. Preliminary TLC detection of the crude extract of compounds produced by *Kitasatospora* sp. MBT66 on MM agar plates showed several yellow UV-active spots (254 nm) and orange fluorescence at 366 nm, which turned reddish-brown by staining with anisaldehyde/sulfuric acid reagent. Repeated chromatographic separation of up-scaled extracts (2.0 g) coupled with subsequent spectroscopy data analysis resulted in the characterization of five previously undescribed rhamnosylated phenazines, endophenazines A–E (**1–5**) and one *N*-prenylated phenazine endophenazine F1 (**6**) with minor yield, together with the previously identified 1-carbomethoxy phenazine (**7**).¹⁶ The structures of the compounds are presented in

Figure 1.

2.1. Resolution of the endophenaside structures

Compound **1** was obtained as a yellowish solid with a molecular formula of $C_{20}H_{21}N_3O_6$ established from high resolution mass $[M + H]^+$ peak at m/z 400.15076. The 1H NMR (600 MHz, CD_3OD) of **1** (Table 1) showed six aromatic protons in two 1,2,3-trisubstituted aromatic spin systems, via δ_H 8.88 (dd, $J = 7.2, 1.2$ Hz, H-2), 8.07 (dd, $J = 9.0, 7.2$ Hz, H-3), and 8.54 (dd, $J = 9.0, 1.2$ Hz, H-4), along with 7.69 (dd, $J = 7.2, 0.6$ Hz, H-7), 7.97 (dd, $J = 9.0, 7.2$ Hz, H-8), and 8.15 (dd, $J = 9.0, 0.6$ Hz, H-9). Five protons in the range of δ_H 6.0–3.50, together with one methyl doublet at δ_H 1.26 (d, $J = 6.0$ Hz, H_3-6') in high magnetic field constituted another spin system by the confirmation of COSY experiment, which suggested the existence of a 6-deoxyhexapyranose sugar moiety (Figure 2). In addition, one distinct methoxyl δ_H 3.64 (s, H_3-7') was identified in the sugar region. The APT spectrum of compound **1** (Table 2) revealed 20 signals, which were assigned to carbonyl δ_C 168.7 (C-11), one oxygenated sp^2 carbon δ_C 152.8 (C-6), six sp^2 aromatic methine carbons δ_C 136.2 (C-2), 130.6 (C-3), 135.0 (C-4), 114.8 (C-7), 133.0 (C-8) and 123.5 (C-9) and five sp^2 quaternary carbons (δ_C 130.1 (C-1), 143.1 (C-4a), 137.8 (C-5a), 143.7 (C-9a), 141.7 (C-10a). Five oxygenated methine carbons δ_C 97.9 (C-1'), 81.5 (C-2'), 71.8 (C-3'), 73.7 (C-4'), and 70.9 (C-5'), together with one methyl group δ_C 17.7 (C-6'), were corresponding to the 6-deoxyhexapyranose by HSQC and APT spectra. The final structure of compound **1** was determined by analyses of the 2D NMR spectra (1H - 1H COSY, HSQC and HMBC). The HMBC spectrum of compound **1** also suggested that it contained a 6-hydroxy-phenazine-1-carboxylic acid aglycone (H-2/C-11, C-10a, C-4; H-3/C-1; H-4/C-4a, C-2; H-7/C-9, C-6, C-5a; H-8/C-9a, C-6 and H-9/C-10a, C-7, C-5a). The aglycone was also confirmed by comparison of the spectral data with the literature.^{27–29} Analysis of the COSY spectrum revealed one contiguous spin system formed by correlations of five oxygen-bearing methines and one doublet methyl signal at δ_H 1.39, confirming the deoxyhexapyranose. The connectivity of the 6-deoxysugar to the aglycone was confirmed from the HMBC long-range coupling of H-8 (δ_H 7.97) and H-1' (δ_H 6.03) to the C-6 (δ_C 152.8) of the phenazine nucleus. The additional characteristic methoxy group was attached onto C-2', based on the two key HMBC correlation H-2'/C-7' and $H_3-7'/C-2'$ (Figure 2). The key HMBC correlation H-2/C-11 (δ_C 168.7) established a carbonyl substituent at C-1. To satisfy the molecular formula $C_{20}H_{21}N_3O_6$ deduced by HRMS, the substituent at C-1 had to be an amide instead of a carboxyl. Consequently, the structure of **1** was assigned as a new hydroxyphenazine-1-carboxamide glycoside and designed as endophenaside A (Figure 1).

Compounds **2–4** and **6** possess the basic phenazine skeleton in comparison with **1**, and the structure elucidation was done by NMR and HRMS. Compound **2** gave a molecular formula of $C_{24}H_{26}N_2O_7$ established from a high-resolution mass measurement of the $[M + H]^+$ peak at m/z 455.18187. 1H NMR (500 MHz, CD_3OD) comparison with compound **1** indicated the presence of a disubstituted phenazine ring, as well as the same existence of deoxyhexopyranose moiety in compound **2**. In contrast to compound **1**, the key HMBC correlation from H-1' (δ_H 6.41) to the carbonyl C-11 (δ_C 166.2) established the linkage of 6-deoxyhexopyranose moiety with phenazine ring through ester bond rather than ether bond. Moreover, an additional hydroxylated prenyl chain in **2** was established by intensive 2D NMR spectral analysis (Figure 2). NOESY experiments assigned the *trans*-position of the

hydroxymethyl group owing to coupling signals for 2" and 5" in combination with lacking signals for 2" and 4". The difference between endophenaside B (**2**) and the known compound phenazoviridin is the terminal hydroxylation in prenyl substitute.³⁰ The structure identification of compound **3** (endophenaside C) was effortless, since both high-resolution mass and ¹H NMR demonstrated an additional methyl group was added to **2**. Methylation pattern of the sugar moiety as seen for **1** was confirmed by HMBC correlation from H₃-7' to C-2' (δ_{C} 80.2). Endophenaside C (**3**) was a new glycosylated phenazine derivative, possessing the methylated sugar structural feature of compound **1**, as well as the hydroxylation attribute in the isoprenyl chain as in compound **2**. Compound **5** shared the same molecular formula C₂₄H₂₆N₂O₇ with **2**, as deduced from [M + H]⁺ peak at *m/z* 455.18243. Though disubstituted phenazine and 6-deoxyhexopyranose feature were found in ¹H NMR spectrum (500 MHz, CD₃OD), difference was obviously observed as well: a *trans* double bond, δ_{H} 7.89 (d, *J* = 16.0 Hz, H-1"), 7.12 (d, *J* = 16.0 Hz, H-2"), and two methyls linked to saturated carbon, δ_{H} 1.54 (s, H₃-4"), 1.53 (s, H₃-5"). The double bond of the isoprenyl group was assumed to be rearranged from C-2"=C-3" to C-2"=C-1" due to the oxidation of C-3", which was ultimately verified by the HSQC and HMBC spectral analysis (Figure 2). Remarkably, the 6-deoxyhexopyranose sugar of compound **4** (endophenaside D) was located at prenyl side chain through 3"-OH instead of 11-COOH, based on the key HMBC correlation from sugar anomeric proton H-1' (δ_{H} 6.43) to the oxygenated quaternary carbon C-3" (δ_{C} 71.7). Compound **6** was isolated as a violet amorphous powder, whose molecular formula was deduced as C₂₃H₂₄N₂O₄ by experimental high resolution mass [M + H]⁺ peak at *m/z* 393.18115. ¹H NMR spectrum (600 MHz, CD₃OD) of **6** indicated its attribution to phenazine type. Both ¹H NMR and APT spectra showed two prenyl side chains, but without the 6-deoxyhexopyranose sugar moiety. These two prenyl substituents were respectively positioned at nitrogen atom 5 and carbon atom 9 by intensive 2D NMR analysis. The distinctive downfield carbonyl signal at δ_{C} 185.8 (C-7) in APT spectrum in combination with the two high field shifted narrow doublets in ¹H NMR spectrum, via δ_{H} 6.19 (d, *J* = 2.4 Hz, H-6) and 6.88 (d, *J* = 2.4 Hz, H-8), revealed oxidation pattern happening at C-7 of phenazine scaffold to form a conjugated ketone like chromophenazine C,³¹ which was further reflected by additional UV/vis absorption at 518 nm and finally confirmed by the HMBC experiment. Also, the geometry of the double bond of the prenyl group was assigned as *trans* based on the NOESY correlation between H-2" (δ_{H} 5.45) and H-5" (δ_{H} 4.00). Notably, Heine and colleagues recently described endophenazines produced by *Kitasatospora* strain HK714, which were similar to but distinct from compound **1**.³² Intensive spectroscopy data analysis for compounds **2–4** showed that their aglycone moieties were identical to those of previously described endophenazines,³² except that compounds **2–4** carried an additional sugar moiety. Compound **6** (endophenazine F1) resembled endophenazine F, but contained a hydroxylated prenyl group at the *N*-5 position, with the hydroxyl group at a terminal methyl group.

Compound **5** (endophenazine E) featured a phenazine nucleus and a 6-deoxyhexopyranose sugar as shown by ¹H NMR. However, seven aromatic protons with overlapped and complicated coupling split patterns suggested a monosubstituted instead of a disubstituted pattern. The key HMBC correlation from sugar anomeric proton H-1' (δ_{H} 6.39) to C-11 (166.7) confirmed the connectivity of a sugar substituent to the phenazine core through a carbonyl group. Furthermore, endophenaside E (**5**) was prone to methanolysis to

generate 1-carbomethoxy phenazine (7) during NMR experiments (Figure S1).

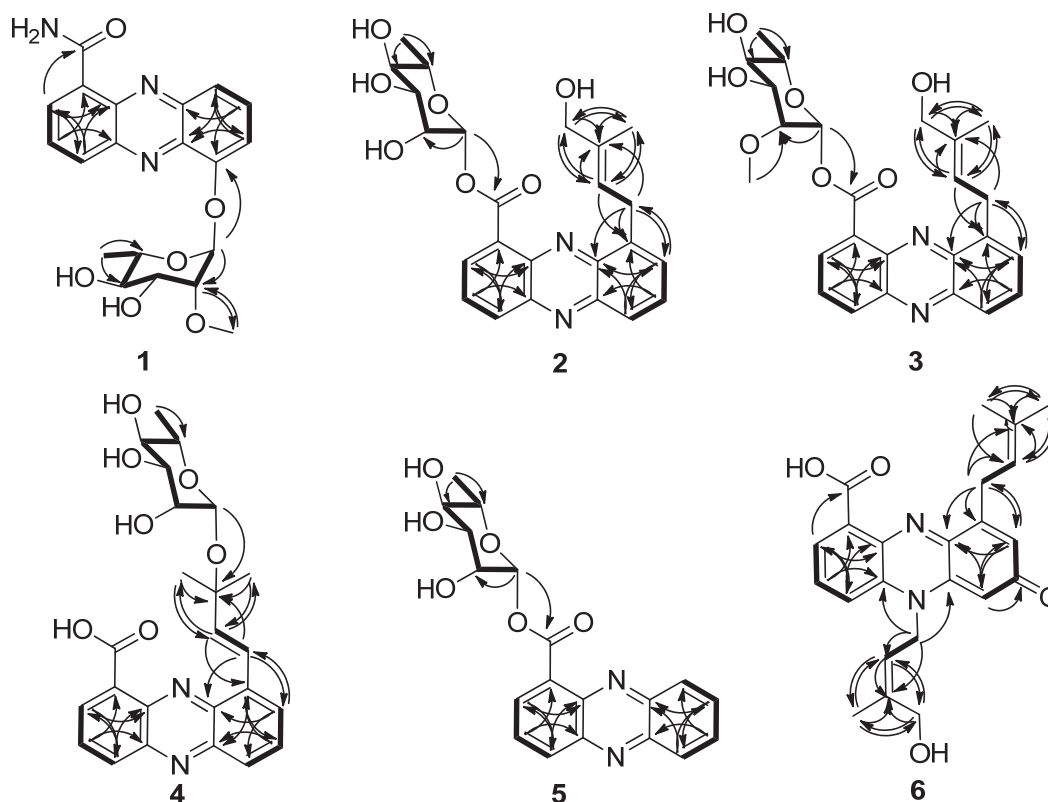


Figure 2. HMBC and COSY correlations of new phenazine derivatives. Key HMBC (↷) and ^1H - ^1H COSY (—) NMR correlations of new phenazine derivatives are presented.

The relative stereochemistry of the 6-deoxyhexopyranose in compounds **1–5** was determined on the basis of NOE observations and coupling constants.^{28,30} The small coupling constant ($J_{1',2'}$ 1.5 Hz) between the anomeric proton and H-2' suggested that H-1' and H-2' were both in equatorial positions and assigned the relative configuration of H-1' to be α , since the NOE was not observed either between H-1' and H-3' or between H-1' and H-5'. The NOE observed between H-3' and H-5' as well as the coupling constant between H-4' and H-5' ($J_{4',5'}$ 9.5 Hz) indicated that they were axially oriented. Based on further comparison with NMR data reported previously²⁸, 6-deoxyhexopyranose in compounds **1–5** was identified as an α -rhamnose (Figure 3). As mentioned above, few glycosylated phenazines have so far been found in nature, most of which are rhamnosides with significant bioactivity.^{28,29} Endophenazines A (**1**) and C (**3**) are the first example for 2'-*O*-methylation of the rhamnose moiety in the phenazine glycosides.

The HMBC spectrum does not allow distinguishing between 1,6- or 1,9- disubstituted phenazine.^{28–30,32,33} Indeed, the prenyl- and the carboxyl-group on the phenazine core of compounds **1–4**, could be either 1,6- or 1,9-disubstituted. However, the presence of a 1,9-disubstituted pattern is deduced from previous analysis of the aglycone moiety³² that is also present in compounds **2–4**. Moreover, close scrutiny of *Kitasatospora* sp. MBT66 by UHPLC-Q-TOF analysis showed that *Kitasatospora* sp. MBT66 could in fact also produce most of the previously described aglycons (endophenazines) produced by *Kitasatospora* sp. HK714, albeit with a low yield (Figure S2 and Table S1). Conversely, the disubstitution pattern in compound **1** was tentatively proposed as [1,6-], which was deduced from the

possible biosynthesis pathway (see below).

Table 1. ^1H NMR data for compounds **1–6**.

NO.	1	2	3	4	5	6
2	8.88, dd (7.2, 1.2)	8.27, dd (7.0, 1.0)	8.29, dd (7.0, 1.0)	8.30, dd (7.0, 1.0)	8.38, dd (7.0, 1.5)	7.51, dd (7.2, 1.2)
3	8.07, dd (9.0, 7.2)	8.00, dd (8.5, 7.0)	8.02, dd (8.5, 7.0)	8.02, dd (7.5, 7.0)	8.02, dd (7.5, 7.0)	7.79, dd (8.4, 7.2)
4	8.54, dd (9.0, 1.2)	8.41, dd (8.5, 1.0)	8.43, dd (8.5, 1.0)	8.42, dd (7.5, 1.0)	8.46, dd (7.5, 1.5)	7.68, dd (8.4, 1.2)
6		8.10, dd (9.0, 1.5)	8.12, dd (9.0, 1.5)	8.15, dd (7.0, 1.0)	8.01, m	6.19, d (2.4)
7	7.69, dd (7.2, 0.6)	7.91, dd (9.0, 7.0)	7.93, dd (9.0, 7.0)	7.98, dd (7.5, 7.0)	8.02, m	
8	7.97, dd (9.0, 7.2)	7.81, dd (6.5, 1.0)	7.83, dd (6.5, 1.0)	8.16, dd (7.5, 1.0)	8.28 (m, 1H)	6.88, d (2.4)
9	8.05, dd (9.0, 0.6)				8.01, m	
1'	6.03, d (1.2)	6.41, d (1.5)	6.56, d (1.5)	6.43, d (1.5)	6.39, d (1.5)	3.83, d (7.8)
2'	4.04, dd (2.4, 1.8)	4.15, dd (3.5, 1.5)	3.81, dd (3.5, 1.5)	4.27, dd (3.5, 1.5)	4.13, dd (3.5, 1.5)	5.49, m
3'	4.28, dd (9.6, 3.6)	3.75, dd (9.5, 3.5)	3.80, dd (9.5, 3.5)	3.79, dd (9.5, 3.5)	4.01, dd (9.5, 3.5)	
4'	3.52, t (9.6)	3.53, t (9.5)	3.48, t (9.5)	3.58, t (9.5)	3.60, t (9.5)	1.78, s
5'	3.81, m	3.83, m	3.80, m	3.82, m	4.21, m	1.73, s
6'	1.26, d (6.0)	1.35, d (6.5)	1.33, d (6.0)	1.33, d (6.5)	1.39, d (6.0)	
7'	3.64, s		3.63, s			
1''		4.25, dd (15.5, 7.5); 4.18, dd (15.5, 2.5)	4.27, dd (15.5, 7.5); 4.20, dd (15.5, 2.5)	7.89, d (16.0)		5.17, d (5.4)
2''		5.86, m	5.87, m	7.12, d (16.0)		5.45, m
4''		1.88, s	1.89, s	1.54, s		1.98, s
5''		4.01, s	4.02, s	1.53, s		4.00, s

Recorded in CD_3OD . ^1H NMR spectra of **2–5** were recorded at 500 MHz, **1**, **6** at 600 MHz. Proton coupling constants (J) in Hz are given in parentheses. All chemical shift assignments were done on the basis of 1D and 2D NMR techniques.

Table 2. ^{13}C NMR data for compounds **1–6**.

NO.	1	2	3	4	5	6
1	130.1	132.7	131.3	131.4	132.2	145.5
2	136.2	132.3	132.9	132.7	133.5	122.7
3	130.6	130.5	130.5	130.7	130.2	134.6
4	135.0	133.3	132.5	133.1	133.9	114.7
4a	143.1	142.9	142.1	142.1	143.5	132.1
5a	137.8	143.3	143.1	143.6	143.6	139.8
6	152.8	127.7	127.1	127.8	132.4	98.6
7	114.8	132.7	132.8	132.7	132.5	185.8
8	133.0	130.6	130.5	128.0	129.7	132.3
9	123.5	142.0	141.3	137.0	132.4	146.6
9a	143.7	144.4	143.6	142.1	144.7	146.8
10a	141.7	140.6	140.2	140.3	141.8	133.4
11	168.7	166.2	166.0	165.6	166.7	175.8
1'	97.9	96.4	93.0	96.2	96.4	29.9
2'	81.5	70.9	80.4	70.5	70.7	122.1
3'	71.8	71.7	71.2	71.3	71.5	135.7
4'	73.7	73.1	72.9	72.7	72.8	25.9
5'	70.9	72.4	72.0	72.2	72.2	18.0
6'	17.7	17.9	17.5	17.7	17.6	
7'	59.3		59.2			
1''		29.5	29.1	121.3		47.1
2''		124.1	123.7	142.7		117.2
3''		137.6	137.1	71.7		142.8
4''		13.7	13.4	29.2		14.3
5''		68.4	68.0	29.2		67.3

Table 3. Antimicrobial activity of purified phenazine derivatives **1–7**.

Compound #	Inhibition zone (mm)	
	<i>Bacillus subtilis</i>	<i>Escherichia coli</i>
1	8	7
2	8	12
3	14	9
4	8	0
5	15	7
6	8	0
7	26	0
AMP	30	21
STR	12	19
NC	0	0

AMP, ampicillin; STR, streptomycin; NC, negative control (methanol)

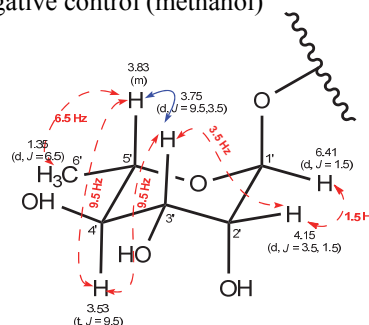


Figure 3. Selected NOE correlations (↔) and coupling constants (↔) in the sugar moiety of compound **2**.

2.2. Antimicrobial activity of the endophenazines

Antimicrobial properties of compounds **1–7** were evaluated by agar diffusion assay, against the Gram-positive indicator strain *Bacillus subtilis* 168 and the Gram-negative *Escherichia coli* K12 (Table 3). Compound **5** showed highest antibiotic activity against both strains, while compounds **4** and **6** only inhibited *B. subtilis* and not *E. coli*. These antimicrobial activities are similar to the data published previously for the aglycones³². Remarkably, the glycosylation enhanced the activity of compounds **2** and **3** against Gram-negative *E. coli*, in contrast to what was reported for the corresponding aglycone. Taken together, our data show that the novel endophenazines described here have significant antimicrobial activity, whereby the glycosylation did not inhibit the bioactivity.

2.3. The phenazine biosynthetic gene cluster

Annotation of the genome sequence of *Kitasatospora* sp. MBT66 (Genbank accession number JAIY000000000²⁶) identified a gene cluster of around 25 kb that is likely responsible for the production of phenazine-type compounds. The gene products of the biosynthetic gene cluster (Table 4) were nearly identical to those of the endophenazine biosynthetic gene cluster of *Kitasatospora* sp. HKI 714 (*epaA-epaU*³²; GenBank accession number: KJ207079).

Table 4. Gene organization of the phenazine biosynthetic gene cluster of *Kitasatospora* sp. MBT66 (Genbank accession number JAIY000000000). Proteins were named according to the gene cluster of *Kitasatospora* sp. HKI 714 (Genbank Accession KJ207079).³² Literature references corresponding to the predicted functions are indicated in superscript.

ORF	Protein	Contig	Length (aa)	Annotation	Homologue in HKI 714	Homology
ORF1	EpaA	182	303	Prenyltransferase	AHW81460.1	100%
ORF2	EpaB	182	345	Mevalonate kinase	AHW81461.1	100%
ORF3	EpaC	182	353	Diphosphomevalonate decarboxylase	AHW81462.1	100%
ORF4	EpaD	182, 329	413	Phosphomevalonate kinase	AHW81463.1	98%
ORF5	EpaE	329	360	Isopentenyl-diphosphate δ -isomerase	AHW81464.1	100%
ORF6	EpaF	329	352	HMG-CoA reductase	AHW81465.1	100%
ORF7	EpaG	329	391	3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) synthase	AHW81466.1	100%
ORF8	EpaH	329	51	Regulatory protein MerR	AHW81467.1	98%
ORF9	EpaI	329	207	Unknown function	AHW81468.1	99%
ORF10	EpaJ	329	160	Phenazine biosynthesis enzyme (PhzA/B) ¹⁹	AHW81469.1	100%
ORF11	EpaK	329	237	FMN-dependent oxidase (PhzG) ¹⁹	AHW81470.1	97%
ORF12	EpaL	329	278	<i>trans</i> -2, 3-dihydro-3-hydroxyanthranilate isomerase (PhzF) ¹⁹	AHW81471.1	99%
ORF13	EpaM	329	666	2-amino-2-desoxy-isochorismate (ADIC) synthase (PhzE) ¹⁹	AHW81472.1	99%
ORF14	EpaN	329	207	2, 3-dihydro-3-hydroxy-anthranilate (DHHA) synthase (PhzD) ¹⁹	AHW81473.1	100%
ORF15	EpaO	329	391	3-deoxy-D- <i>arabino</i> -heptulosonic acid 7-phosphate (DAHP) synthase (PhzC) ³⁴	AHW81474.1	99%
ORF16	EpaP	329	162	enzyme of phenazine biosynthesis (PhzA/B)	AHW81475.1	99%
ORF17	EpaQ	329	421	Rhodocoxin reductase	AHW81476.1	99%
ORF18	EpaR	329	107	Putative rhodocoxin	AHW81477.1	100%
ORF19	EpaS	329	305	Transcriptional regulator	AHW81478.1	99%
ORF20	EpaT	329	431	Cytochrome d ubiquinol oxidase subunit I	AHW81479.1	99%
ORF21	EpaU	329	360	Cytochrome d ubiquinol oxidase subunit II	AHW81480.1	99%
ORF22	<i>orf1</i>	329	598	ABC transporter ATPase	AHW81481.1	100%
ORF23	<i>orf2</i>	329	569	ABC transporter	AHW81482.1	99%

It was proposed that EpaA catalyzes the prenylation at C-9 and/or N-5 positions of phenazine-1-carboxylic acid, similar to the function of enzyme PpzP in *Streptomyces anulatus*,³⁵

which generated the [1,9-] substitution mode on the phenazine nucleus. From the perspective of biosynthesis, the formation of phenazine-1-carboxylic acid not only depends on the enzymes encoded by the gene cluster, but also involves spontaneous reaction without enzymatic catalysis.^{19,36} The carboxyl group derived from chorismate is likely cleaved by molecular oxygen through oxidative decarboxylation, which would be alternatively replaced by a hydroxyl group in the aqueous environment (Figure 4). This may explain why 6-hydroxyphenazine-1-carboxylic acid derivatives were frequently identified,^{28,29,37–39} while 9-hydroxyphenazine-1-carboxylic acid^{40,41} was less often found.

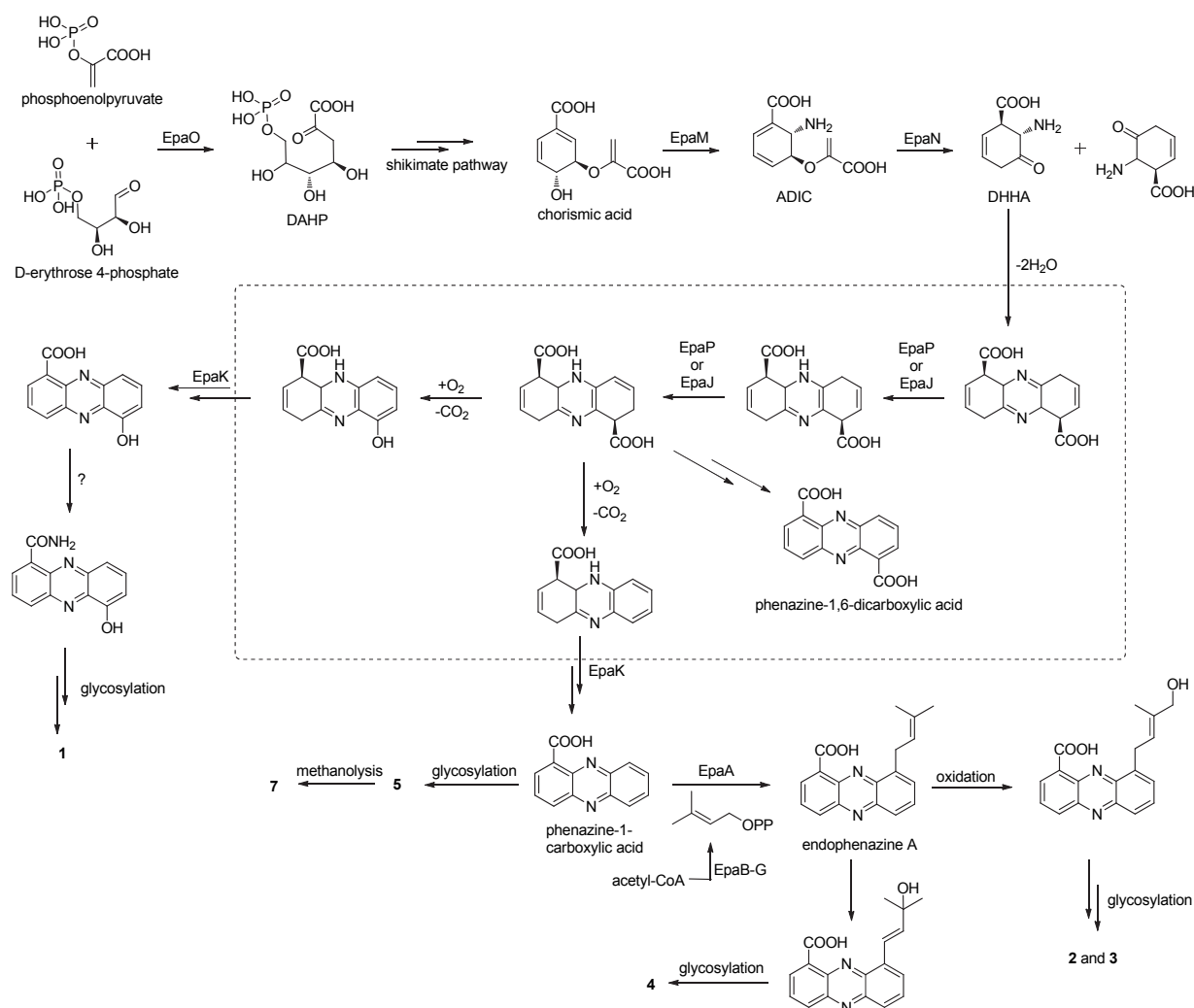


Figure 4. Proposed biogenesis of phenazine derivatives. Enzymes catalyzing the reaction correspond to Table 4. This biosynthesis pathway of phenazines was summarized according to references^{19,20,34,36}. Reactions leading to the formation of the phenazine core are highlighted by a dashed box, which includes spontaneous reaction without enzymatic catalysis.^{19,36} A yet uncharacterized glycosyltransferase and methyltransferase then give rise to the new glycosylated endophenazines (1–5) described in this study.

Besides the glycosylated endophenazines, we were able to show that *Kitasatospora* sp. MBT66 also produces small amounts of the same endophenazines as produced by *Kitasatospora* sp. HKI 714 (Figure S2 and Table S1). We expect that the two strains are closely related kitasatosporae, a genus closely related to *Streptomyces* in the family of *Streptomycetaceae*.⁴² This should be resolved by comparison of the genomes and phenotypic characteristics of the two strains. It is yet unclear which genes are responsible for the

glycosylation of the endophenazine aglycones, as no such genes were found within or in close proximity of the *epa* gene cluster (Table 4). Therefore, the genes responsible for the transfer and methylation of the rhamnose moiety to either the carboxyl or the hydroxyl group of the phenazines should be located elsewhere on the chromosome. This is reminiscent of the biosynthesis of 7-prenylisatin in *Streptomyces* sp. MBT28, whereby the indispensable oxygenases responsible for the oxidation of 7-prenylindole into the 7-prenylisatin were not found within or in close proximity of the *isa* gene cluster (Chapter 4). Research to identify the genes and enzymes responsible for the glycosylation of the endophenazines in *Kitasatospora* sp. MBT66 is described in Chapter 10.

3. CONCLUSION

We isolated a group of novel glycosylated endophenazines A–E (**1–5**), together with *N*-prenylated endophenazine F1 (**6**) from *Kitasatospora* sp. MBT66. The new compounds featured unusual structural properties including a 2'-*O*-methylation of the rhamnosyl moiety, which enriches the structural diversity of phenazine-class antibiotics. All phenazine derivatives showed antimicrobial activity against the Gram-positive *Bacillus subtilis*, while compounds **1–3** and **5** also inhibited growth of the Gram-negative *Escherichia coli*. The gene cluster (*epa*) responsible for the biosynthesis of phenazine core was identified in *Kitasatospora* sp. MBT66, which shows high homology to that of *Kitasatospora* sp. HKI 714. However, the genes responsible for the transfer and methylation of the rhamnose moiety to the phenazines are missing in *epa*, which must be located elsewhere on the chromosome (Chapter 10). We tentatively conclude that a non-clustered or subclustered genome arrangement for secondary metabolites biosynthesis in actinomycetes is more common than previously suggested. Though synthetic biology approaches will be a major trend for future antibiotics discovery (Chapter 1),⁴³ it may be less effective to handle this type of intertwined metabolic pathways, which could be instead overcome by the traditional systematic isolation methodology.

4. EXPERIMENTAL SECTION

4.1. General Experimental Procedures

FT-IR was measured on Perkin-Elmer FT-IR Spectrometer Paragon 1000. UV measurements were performed using a Shimadzu UV mini-1240. NMR spectra were recorded in methanol-*d*₄ on a Bruker DMX 500 MHz NMR and 600 MHz calibrated to a residual methanol-*d*₄ (3.30 ppm).

High resolution mass data (HRESIMS) were collected on an Agilent 1200 series HPLC connected to a LTQ-Orbitrap spectrometer. Semi-preparative HPLC (pHPLC) was performed with a Shimadzu HPLC system and a 5 mL Rheodyne manual injection loop, equipped with a reversed-phase C18 column (Phenomenex Luna C18 (2) 100 Å 5 micron 250 × 10 mm). All the pHPLC experiments used 2 mL/min flow rate and fraction collection based on detected peak. TLC was performed with silica gel 60 (Merck, Darmstadt, Germany) plates using CHCl₃/MeOH 10:1 and visualized with anisaldehyde/sulfuric acid reagent. All solvents and chemicals were of analytical and HPLC grade.

4.2. Microorganisms and Culture Conditions

Kitasatospora sp. MBT66 was isolated from Himalayan mountain soil as described previously.¹⁶ The strain was cultivated on minimal media agar plates (MM)⁴⁴ with glycerol and mannitol (1% w/v) as the carbon sources. Square agar plates (12 cm × 12 cm) were inoculated with 5×10^7 spores from a fresh spore suspension. In total 125 plates were used, and incubated at 30 °C for 6 days. *Escherichia coli* K12 and *Bacillus subtilis* 168 were used for screening of antimicrobial activity.

4.3. Extraction and Isolation

Kitasatospora sp. MBT66 was fermented by stationary cultivation with a total volume of 5 liter agar medium. After fermentation for 6 days, mycelia were combined and extracted with ethyl acetate by soaking in solvent overnight at room temperature. The extraction was evaporated under reduced pressure to obtain 2.0 g of a crude residue. This extract adsorbed by silica gel was first chromatographed on a Macroporous resin Diaion HP-20 from Supelco (Bellefonte, PA, USA) by eluting stepwise from H₂O to MeOH. For monitoring, 20 µL of each fraction was subjected to TLC on premade silica gel plates (Merck, Darmstadt, Germany) using developing solvent system of chloroform and methanol (10:1). The 40% (v/v) methanol eluent showing interesting bands on TLC and moderate antibiotic activity in disc paper agar diffusion test with bacterium *Bacillus subtilis*, was subsequently separated by silica gel (pore size 60 Å, 70–230 mesh, St. Louis, MO, USA) column chromatography employing gradient elution from CHCl₃ to MeOH, to give four subfractions (sfr.1–sfr.4). Sfr.2 was rechromatographed on Sephadex LH-20 coloumn, to purify the compound **2** (2.5 mg) and compound **6** (0.2 mg). Semi-preparative reversed-phase HPLC (Phenomenex Luna C18 (2) 100 A 5 micron 250 × 10 mm) purification of sfr.1, eluting with a gradient of MeOH in H₂O from 80% to 100% in 30 min, gave the pure compound **3** (t_R = 30.25 min, 0.3 mg) and **4** (t_R = 29.10 min, 0.2 mg). Separation of sfr.3 by HPLC using the gradient of MeOH in H₂O from 50% to 80% in 30min resulted in the purification of **5** (t_R = 21.25 min, 1.3 mg), and **7** (t_R = 28.50 min, 0.2 mg). Sfr.4 was similarly subjected to pHPLC employing the mobile phase MeOH in H₂O from 50% to 80% in 30 min and led to the separation of **1** (t_R = 17.7 min, 0.35 mg).

4.4. Antimicrobial activity assays

Antimicrobial activity was determined basically following the methods described in reference.³² The respective compound (**1**–**7**) was dissolved in methonal (2 mg/ml), and 25 µl of the solution was applied on a paper disk (d = 6 mm). The disks were then placed onto an agar plate containing a softagar overlay with either *Bacillus subtilis* or *Escherichia coli*. Ampicillin and streptomycin were used as positive controls, and the solvent methanol as the negative control. After incubation at 37 °C for 18 h, growth inhibition zones (in mm) were recorded as antimicrobial activity.

REFERENCES

- (1) Giske, C. G.; Monnet, D. L.; Cars, O.; Carmeli, Y. *Antimicrob. Agents Chemother.*

- 2008, 52, 813–821.
- (2) Spellberg, B.; Guidos, R.; Gilbert, D.; Bradley, J.; Boucher, H. W.; Scheld, W. M.; Bartlett, J. G.; Edwards, J. *Clin. Infect. Dis.* **2008**, 46, 155–164.
 - (3) Rice, L. B. *J. Infect. Dis.* **2008**, 197, 1079–1081.
 - (4) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat. Rev. Drug Discov.* **2007**, 6, 29–40.
 - (5) Claessen, D.; Rozen, D. E.; Kuipers, O. P.; Søgaaard-Andersen, L.; van Wezel, G. P. *Nat. Rev. Microbiol.* **2014**, 12, 115–124.
 - (6) Flärdh, K.; Buttner, M. J. *Nat. Rev. Microbiol.* **2009**, 7, 36–49.
 - (7) Hopwood, D. A. *Streptomyces in nature and medicine: the antibiotic makers.*; New York, Oxford University Press, 2007.
 - (8) Miyadoh, S. *Actinomycetol* **1993**, 7, 100–106.
 - (9) Challis, G. L.; Hopwood, D. A. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, 100, 14555–14561.
 - (10) Bentley, S. D.; Cerdeño-Tárraga, K. F. C. A.-M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Bentley, S. D.; Harper, D.; Bateman, A.; Brown, S.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Larke, L.; Murphy, L.; Oliver, K.; Rabbinowitsch, E.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Woodward, J.; Barrell, B. G.; Parkhill, J. *Nature* **2002**, 3, 141–147.
 - (11) Ohnishi, Y.; Ishikawa, J.; Hara, H.; Suzuki, H.; Ikenoya, M.; Ikeda, H.; Yamashita, A.; Hattori, M.; Horinouchi, S. *J. Bacteriol.* **2008**, 190, 4050–4060.
 - (12) Cruz-Morales, P.; Vijgenboom, E.; Iruegas-Bocardo, F.; Girard, G.; Yáñez-Guerra, L. A.; Ramos-Aboites, H. E.; Pernodet, J.-L.; Anné, J.; van Wezel, G. P.; Barona-Gómez, F. *Genome Biol. Evol.* **2013**, 5, 1165–1175.
 - (13) Craney, A.; Ozimok, C.; Pimentel-Elardo, S. M.; Capretta, A.; Nodwell, J. R. *Chem. Biol.* **2012**, 19, 1020–1027.
 - (14) van Wezel, G. P.; McKenzie, N. L.; Nodwell, J. R. *Chapter 5. Applying the genetics of secondary metabolism in model actinomycetes to the discovery of new antibiotics.*; 1st ed.; Elsevier Inc., 2009; Vol. 458.
 - (15) Zhu, H.; Sandiford, S. K.; van Wezel, G. P. *J. Ind. Microbiol. Biotechnol.* **2014**, 41, 371–386.
 - (16) Zhu, H.; Swierstra, J.; Wu, C.; Girard, G.; Choi, Y. H.; van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. *Microbiology* **2014**, 160, 1714–1725.
 - (17) Abken, H.; Tietze, M.; Brodersen, J. *J. Bacteriol.* **1998**, 180, 2027–2032.
 - (18) Price-Whelan, A.; Dietrich, L. E. P.; Newman, D. K. *Nat. Chem. Biol.* **2006**, 2, 71–78.
 - (19) Ahuja, E. G.; Janning, P.; Mentel, M.; Graebisch, A.; Breinbauer, R.; Hiller, W.; Costisella, B.; Thomashow, L. S.; Mavrodi, D. V.; Blankenfeldt, W. *J. Am. Chem. Soc.* **2008**, 130, 17053–17061.
 - (20) Laursen, J. B.; Nielsen, J. *Chem. Rev.* **2004**, 104, 1663–1686.
 - (21) Cholo, M. C.; Steel, H. C.; Fourie, P. B.; Germishuizen, W. a; Anderson, R. J. *Antimicrob. Chemother.* **2012**, 67, 290–298.
 - (22) Cook, R.; Thomashow, L. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, 92, 4197–4201.
 - (23) Maria E. Hernandez, A. K. and D. K.; Newman *Appl. Environ. Microbiol.* **2004**, 70,

- 921–928.
- (24) Wang, Y.; Kern, S. E.; Newman, D. K. *J. Bacteriol.* **2010**, *192*, 365–369.
 - (25) Mavrodi, D. V.; Blankenfeldt, W.; Thomashow, L. S. *Annu. Rev. Phytopathol.* **2006**, *44*, 417–445.
 - (26) Girard, G.; Willemse, J.; Zhu, H.; Claessen, D.; Bukarasam, K.; Goodfellow, M.; van Wezel, G. P. *Antonie Van Leeuwenhoek* **2014**, *106*, 365–380.
 - (27) Romer, A. *Org. Magn. Reson.* **1982**, *19*, 66–68.
 - (28) Abdelfattah, M. S.; Toume, K.; Ishibashi, M. *J. Antibiot. (Tokyo)*. **2011**, *64*, 271–275.
 - (29) Rusman, Y.; Oppegard, L. M.; Hiasa, H.; Gelbmann, C.; Salomon, C. E. *J. Nat. Prod.* **2013**, *76*, 91–96.
 - (30) Kato, S.; Shindo, K.; Yamagishi, Y. *J. Antibiot. (Tokyo)*. **1993**, *46*, 1485–1493.
 - (31) Zendah, I.; Riaz, N.; Nasr, H.; Frauendorf, H.; Schüffler, A.; Raies, A.; Laatsch, H. *J. Nat. Prod.* **2012**, *75*, 2–8.
 - (32) Heine, D.; Martin, K.; Hertweck, C. *J. Nat. Prod.* **2014**, *77*, 1083–1087.
 - (33) Krastel, P.; Zeeck, A. *J. Antibiot. (Tokyo)*. **2002**, *55*, 801–806.
 - (34) Schneemann, I.; Wiese, J.; Kunz, A. L.; Imhoff, J. F. *Mar. Drugs* **2011**, *9*, 772–789.
 - (35) Saleh, O.; Gust, B.; Boll, B.; Fiedler, H.-P.; Heide, L. *J. Biol. Chem.* **2009**, *284*, 14439–14447.
 - (36) Rui, Z.; Ye, M.; Wang, S.; Fujikawa, K.; Akerele, B.; Aung, M.; Floss, H. G.; Zhang, W.; Yu, T.-W. *Chem. Biol.* **2012**, *19*, 1116–1125.
 - (37) Ohlendorf, B.; Schulz, D.; Erhard, A.; Nagel, K.; Imhoff, J. F. *J. Nat. Prod.* **2012**, *75*, 1400–1404.
 - (38) Abdelfattah, M. S.; Kazufumi, T.; Ishibashi, M. *J. Nat. Prod.* **2010**, *73*, 1999–2002.
 - (39) Abdelfattah, M. S.; Toume, K.; Ishibashi, M. *Chem. Pharm. Bull. (Tokyo)*. **2011**, *59*, 508–510.
 - (40) Fotso, S.; Santosa, D. A.; Saraswati, R.; Yang, J.; Mahmud, T.; Zabriskie, T. M.; Proteau, P. J. *J. Nat. Prod.* **2010**, *73*, 472–575.
 - (41) Gerber, N. *J. Heterocycl. Chem.* **1969**, *6*, 4–7.
 - (42) Girard, G.; Traag, B. a; Sangal, V.; Mascini, N.; Hoskisson, P. a; Goodfellow, M.; van Wezel, G. P. *Open Biol.* **2013**, *3*, 130073.
 - (43) Smanski, M. J.; Zhou, H.; Claesen, J.; Shen, B.; Fischbach, M. A.; Voigt, C. A. *Nat. Rev. Microbiol.* **2016**, *14*, 135–149.
 - (44) Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. John Innes Foundation, Norwich, UK, 2000.

Chapter 10

Leucanicidin and endophenazines result from methyl-rhamnosylation by the same tailoring enzymes in *Kitasatospora* sp. MBT66

Changsheng Wu ^{1,2}, Marnix H. Medema ³, Rianne M. Läkamp ^{1,4}, Le Zhang ¹, Pieter C. Dorrestein ⁴, Young Hae Choi ², and Gilles P. van Wezel ¹

¹ *Molecular Biotechnology, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, Leiden, The Netherlands*

² *Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, Leiden, The Netherlands*

³ *Bioinformatics Group, Wageningen University, Droevendaalsesteeg 1, 6708PB, Wageningen, The Netherlands*

⁴ *Collaborative Mass Spectrometry Innovation Center, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0751*

ACS Chem. Biol. **2016**, *11*, 478–490

ABSTRACT

The increasing bacterial multi-drug resistance necessitates novel drug-discovery efforts. One way to obtain novel chemistry is glycosylation, which is prevalent in nature, with high diversity in both the sugar moieties and the targeted aglycones. *Kitasatospora sp.* MBT66 produces endophenaside antibiotics, which is a family of (methyl-)rhamnosylated phenazines. Here we show that this strain also produces the plecomacrolide leucanicidin (**1**), which is derived from bafilomycin A₁ by glycosylation with the same methyl-rhamnosyl moiety as present in the endophenasides. Immediately adjacent to the *baf* genes for bafilomycin biosynthesis lie *leuA* and *leuB*, which encode a sugar-*O*-methyltransferase and a glycosyltransferase, respectively. LeuA and LeuB are the only enzymes encoded by the genome of *Kitasatospora sp.* MBT66 that candidate for the methyl-rhamnosylation of natural products, and mutation of *leuB* abolished glycosylation of both families of natural products. Thus, LeuAB mediate the post-PKS methyl-rhamnosylation of bafilomycin A₁ to leucanicidin and of phenazines to endophenasides, showing surprising promiscuity by tolerating both macrolide and phenazine skeletons as the substrates. Detailed metabolic analysis by MS/MS based molecular networking facilitated the characterization of nine novel phenazine glycosides **6—8**, **16**, and **22—26**, whereby compounds **23** and **24** represent an unprecedented tautomeric glyceride phenazine, further enriching the structural diversity of endophenasides.

1. INTRODUCTION

The rapid increase in antimicrobial resistance poses one of the major threats to human health.¹ A particular problem with drug discovery from microbial sources is the high frequency of re-discovery of known compounds, which necessitates new approaches to replenish the antimicrobial drug pipelines.^{2–5} As producers of some two thirds of all known antibiotics and many other medically relevant natural products, actinomycetes are a major source of clinical drugs.^{2,6,7} Sequencing of the genomes of actinomycetes revealed that the natural products producing potential of even the best-studied model organisms has been underestimated.^{8–10} However, many of these gene clusters are poorly expressed in the laboratory.^{11–14} One way of obtaining novel chemistry is by sugar-mediated tailoring, *i.e.* the decoration of molecules by glycosylation. Over 20% of the bacterial natural products (NPs) in the databases is glycosylated, with structurally highly diverse aglycones containing one or more glycosyl groups.¹⁵ Glycosylation can dramatically influence the pharmacological properties of the parent scaffold and directly mediate bioactivity, such as in anthracycline, aureolic acid and enediyne antibiotics.¹⁶ Many microbial glycosides also find their applications in agriculture, such as the insecticide avermectin.¹⁷ Antibiotic glycan alteration (so-called glycorandomization) is a potentially powerful strategy in combating emerging bacterial resistance.^{18–20}

Given the profound biological significance of glycosylation, it is important to harness the biosynthetic machinery for the formation of glycoconjugates,²¹ which will pave the way for the glycodiversification of NPs through genetic engineering approaches.²² The biosynthesis of glycosylated natural products includes (i) assembly of the aglycone; (ii) biosynthesis of an activated form of the sugar moiety, typically a nucleotide diphospho (NDP)-activated sugar; and (iii) transfer of the NDP-sugar to acceptor molecules by glycosyltransferases. Glycosylated macrolides and macrolactams represent the largest allocation in bacterial saccharidic compounds.¹⁵ The plecomacrolides are a family of macrolides that typically feature a 16- or 18-membered macrolactone containing two conjugated diene units connected with a six-membered hemiacetal side chain through a C₃ spacer.²³ Endowed with biologically important intramolecular hydrogen bonding network among the lactone/C₃ linker/hemiacetal structural motif,^{24,25} plecomacrolides exhibit a variety of bioactivities, including antitumor,²⁶ antifungal,²⁷ antiparasitic,²⁸ immunosuppressant,²⁹ and particularly selective vacuolar ATPase (V-ATPase) inhibitors.^{23,30} Additional substituents (mostly on the secondary alcohol of the hemiacetal portion) and/or modification on backbone largely diversified this class of antibiotics, such as bafilomycins,^{31,32,33,34} concanamycins,³⁵ hygrolidin,³⁶ setamycin,³⁷ micromonospolide,³⁸ and formamycin,³⁹ simultaneously affording a variation of bioactivity and toxicity.^{25,40} The genus *Kitasatospora* has a similar life style as *Streptomyces*, and also has a rich arsenal of secondary metabolites.⁴¹ Several plecomacrolide type compounds have been characterized from *Kitasatospora* species, such as bafilomycins A₁ and B₁ from *K. setae*,³⁷ bafilomycins A₁–C₁ and respective amide derivatives from *K. cheerisanensis*.⁴² Bafilomycins are known as specific inhibitors of vacuolar ATPases.³⁰ So far, no detailed reports are available on the gene cluster organization for plecomacrolide scaffolds in *Kitasatospora* species, although the pioneering genome sequencing of *K. setae* predicted the plausible presence of the PKS genetic loci responsible for bafilomycin B₁.⁴¹

We previously characterized the endophenazines in *Kitasatospora* sp. MBT66, which

constitute a family of novel rhamnosylated phenazines.⁴³ Here we show that *Kitasatospora sp.* MBT66 also produces the bafilomycin-derived plecomacrolide antibiotic leucanicidin (**1**). The methylated form of leucanicidin, previously identified as NK155141 (**2**),⁴⁴ is not produced biosynthetically but was derived from reaction with methanol. The biosynthetic gene cluster for leucanicidin was elucidated, which includes genes for a glycosyltransferase (LeuB) and a methyltransferase (LeuA); these likely modify both plecomacrolides and phenazines. The biosynthetic insights together with MS/MS based molecular networking allowed us to identify novel 2'-*O*-methylated and 2'-*O*-unmethylated rhamnosylated endophenazines **6**–**8**, **16**, **22**, **25**, and **26**, as well as an unprecedented tautomer consisting of glyceride phenazines **23** and **24**.

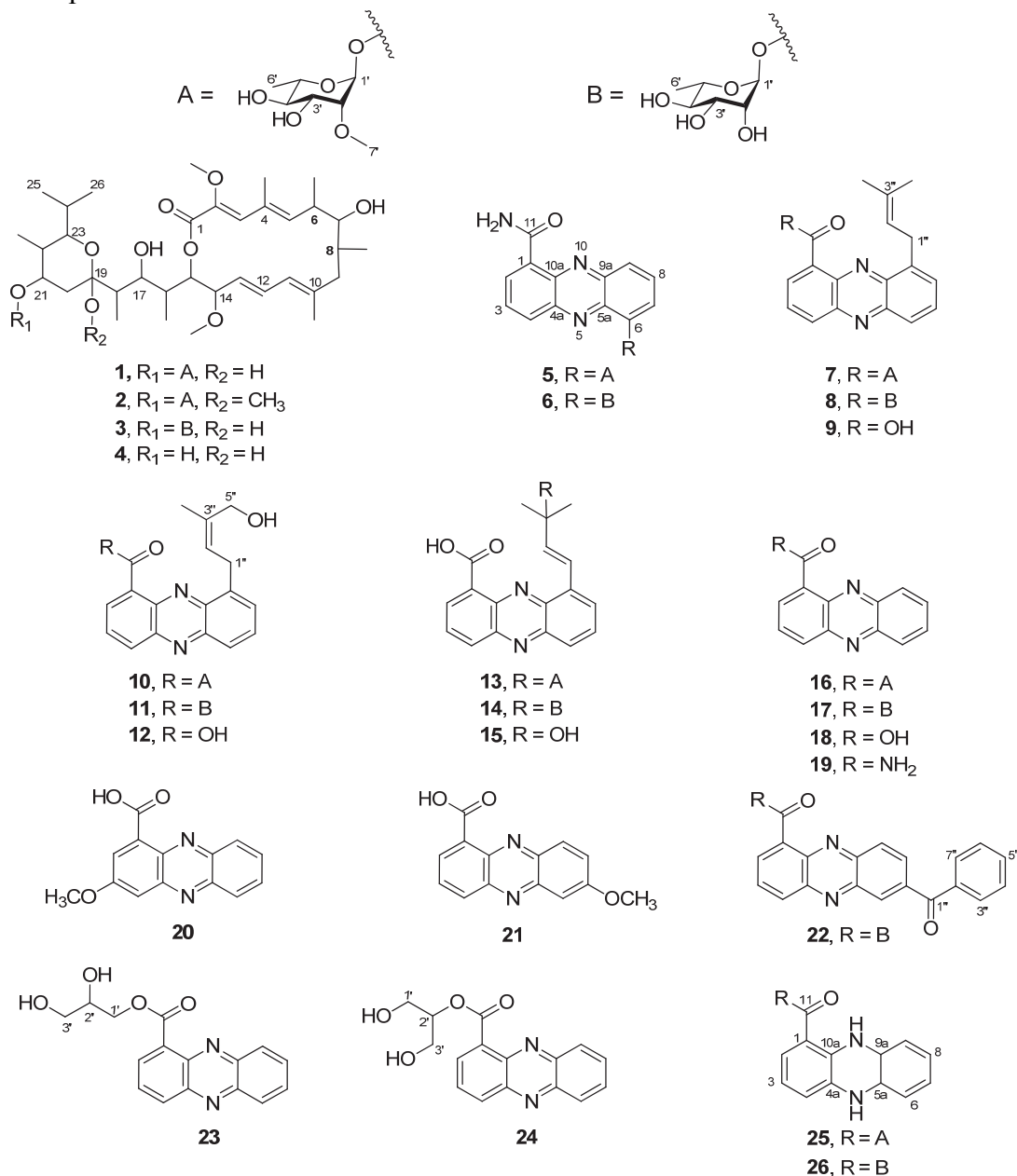


Figure 1. Glycosides from *Kitasatospora sp.* MBT66. Compounds **1**–**26** are classified into two groups according to the aglycones, namely glycosylated plecomacrolides (**1**–**4**) and phenazines (**5**–**26**), which bear the same rhamnosylation. The rhamnosyl substituents in rhamnosides are either 2'-*O*-methylated or unmethylated, except compounds **23** and **24** which are glyceride phenazine tautomers. Phenazine derivatives **6**–**8**, **16**, and **22**–**26** were previously undescribed. The ¹H NMR data assignments for these compounds are summarized in Table S3.

2. RESULTS & DISCUSSION

2.1. Biosynthetic pathway of plecomacrolide glycosides leucanicidin and NK155141

Our previous chemical investigation of *Kitasatospora* sp. MBT66 led to the discovery of five minor rhamnosylated endophenazines A—E, which were isolated from 40% methanol eluent of Macroporous resin Diaion HP-20 column chromatography.⁴³ Ongoing investigation into nonpolar fractions, via 80% methanol eluent, resulted in the purification of two additional compounds. Spectral data interpretation, including NMR, HRMS, and UV, showed that these two compounds were the plecomacrolide glycosides leucanicidin (**1**, Figure 1),⁴⁵ and its methylated derivative NK155141 (**2**).⁴⁴ Moreover, another known plecomacrolide glycoside, bafilomycin A₁-21-*O*-(α -L-rhamnopyranoside) (**3**),⁴⁶ was later identified in MBT66 crude extract by U(H)PLC-UV-TOF analysis, which was judged from comparison of the earlier retention time, the UV spectrum, and high resolution mass with those of leucanicidin. In view of the unusual structural scaffold and important biological properties, plecomacrolide polyketides and their genetics have been studied extensively,^{47,48} because of their complex architecture and corresponding difficulty in total synthesis.^{49,50,51}

Annotation of the genome sequence of *Kitasatospora* sp. MBT66⁵² led to the identification of a biosynthetic gene cluster (BGC) for a PKS responsible for the biosynthesis of bafilomycins, the precursor of leucanicidin (Table 1). The domain organization of the PKS genes (ORF1–ORF13) is consistent with the assembly of macrolactone/C₃ linker/hemiacetal core (Figure S1) of bafilomycin B₁.^{47,48} However, post-PKS tailoring components for the synthesis of leucanicidin were distinct from those for bafilomycin B₁. Genes required for installing the C₅N moiety are absent,^{47,48} and instead two genes, encoding a putative 2'-*O* methyltransferase (ORF15) and a glycosyltransferase (ORF16), were located in the downstream region, which explain the structural differences between leucanicidin and bafilomycin B₁. This is a rather typical genetic configuration, as the glycosylation-associated biosynthetic genes are usually coclustered with those for aglycones in microbial genomes⁵³. This strong linkage was the basis for the development of so-called glycogenomics, an MSⁿ-based genome-mining method for microbial glycosylated molecules.⁵³ Therefore, we reasoned that the macrolactone core (bafilomycin A₁, **4**) of leucanicidin was assembled by the classical type I PKS system. *bafAI*-*AV* encode in total 12 PKS modules, which load isobutyrate as starter unit and subsequently incorporate 11 extender building blocks. The downstream genes *leuA* for a sugar-*O*-methyltransferase and *leuB* for a glycosyltransferase are likely responsible for post-PKS modification by installing a 2-*O*-methylated-rhamnosyl group at the 21-OH position of bafilomycin A₁⁵⁴ (Figure S1). The *leuA* and *leuB* genes are also present adjacent to the homologous *baf* gene clusters in the genomes of *Kitasatospora purpeofuscus* strain NRRL B-1817 (GenBank accession NO. JODS01000000) and *Kitasatospora* sp. NRRL S-495 (GenBank accession NO. JZWY01000000) (both erroneously termed *Streptomyces* in GenBank), but are absent from all bafilomycin gene clusters in *Streptomyces* genomes. This suggests that phylogenetic linkage exists between the methyl-rhamnosylation and the genus *Kitasatospora*.

How then is leucanicidin (**1**) methylated to generate NK155141 (**2**), and can it originate from *LeuA*-mediated *in vivo* enzymatic catalysis? An alternative is that **2** may have arisen from the *in vitro* non-enzymatic reaction with methyl donors during the isolation process, because cyclic hemiketal hydroxyl groups are reactive even in the presence of moderately

nucleophilic reagents.^{55,56} To address this, we optimized leucanicidin production by MBT66. Liquid NMMP,⁵⁷ solid MM,⁵⁷ and R5,⁵⁸ were supplemented with different carbon sources, additional additives, and high alkalinity.^{59,60} Varying growth conditions or the addition of chemical elicitors can be applied to activate the biosynthesis of poorly expressed natural products.⁵⁹ *N*-acetylglucosamine (GlcNAc) has previously been applied for the activation of various BGCs, and acts via the metabolic inactivation of the nutrient-responsive global regulator DasR.^{61–63} Indeed, growth of *Kitasatospora* sp. MBT66 on R5 agar plates with 25 mM GlcNAc effectively increased the production of leucanicidin, thus allowing ready monitoring of leucanicidin production by HPLC-UV profiling.

Crude extracts obtained from mycelia of *Kitasatospora* sp. MBT66 grown on R5 agar with 25mM GlcNAc were dissolved in either methanol or acetonitrile without any additional catalyst. The samples were incubated at room temperature for one week and monitored by HPLC-UV. Methylated leucanicidin (NK155141) was observed exclusively in the methanol solution, whereby the leucanicidin concentration gradually decreased in favor of a time dependent increase in the level of NK155141 (Figure S2, A). However, no NK155141 was detected even after a week of incubation in acetonitrile (Figure S2, B). This *in vitro* experiment provides conclusive evidence that NK155141 is not synthesized *in vivo*, but instead is an artifact resulting from the reaction of the hemiacetal hydroxyl group (19-OH) of leucanicidin with methanol.

Table 1. Organization for the leucanicidin biosynthetic gene cluster of *Kitasatospora* sp. MBT66. The gene cluster architecture for the biosynthesis of bafilomycin A₁ is the same as in *Streptomyces griseus* DSM 2608⁴⁸ and *Streptomyces lohii*.⁴⁷ The genome sequence of *Kitasatospora* sp. MBT66 is available at GenBank with accession number JAIY000000000 and the annotation was submitted to MIBiG with accession number BGC0001232.

ORF	Locus tag	Protein	Length	Annotation	Nearest homologue Homology, Protein, origin	GenPept accession	Comments
ORF1	BI06_RS39075	BafAI	1033	PKS modules 1–4	91%, WP_018955924.1, <i>Streptomyces lohii</i>	N/A #	Gapped sequence
ORF2	BI06_RS39070	BafAII	5019	PKS modules 5–7	90%, ADC79617.1, <i>Streptomyces lohii</i>	WP_043476519.1	
ORF3	BI06_RS39065	BafAIII	3966	PKS modules 8, 9	89%, ADC79618.1, <i>Streptomyces lohii</i>	WP_043476516.1	
ORF4	BI06_RS39060 BI06_RS32440	BafAIV	3453	PKS modules 10, 11	88%, ADC79619.1, <i>Streptomyces lohii</i>	N/A #	Gapped sequence
ORF5	BI06_RS32435	BafAV	2158	PKS module 12+thioesterase	85%, ADC79620.1, <i>Streptomyces lohii</i>	N/A #	
ORF6	BI06_RS32430	BafB	296	glyceryl-ACP oxidase	93%, ADC79621.1, <i>Streptomyces lohii</i>	WP_043474332.1	
ORF7	BI06_RS32425	BafC	115	acyl carrier protein (ACP)	90%, ADC79622.1, <i>Streptomyces lohii</i>	WP_030397161.1	
ORF8	BI06_RS32420	BafD	371	acyl-CoA dehydrogenase	92%, WP_019761696.1, <i>Streptomyces</i> sp. Wigar10	WP_043474112.1	
ORF9	BI06_RS32415	BafE	377	glycerate ACP ligase	93%, WP_018568170.1, <i>Streptomyces</i> sp. PsTaAH-124	WP_051742398.1	
ORF10	BI06_RS32410	BafF	220	<i>O</i> -methyl transferase	90%, ADC79625.1, <i>Streptomyces lohii</i>	WP_043474109.1	
ORF11	BI06_RS32405	BafG	606	AfsR family transcriptional regulator	86%, ADC79626.1, <i>Streptomyces lohii</i>	WP_043474107.1	
ORF12	BI06_RS32400	BafH	253	TEII	93%, WP_019761700.1, <i>Streptomyces</i> sp. Wigar10	WP_030397166.1	
ORF13	BI06_RS32395		126	Putative LuxR_C-like protein	83%, ADC79628.1, <i>Streptomyces lohii</i>	WP_043474105.1	
ORF14			38	malonyl transferase	58%, ADC79629.1, <i>Streptomyces lohii</i>		
ORF15	BI06_RS32390	Leu A	430	sugar <i>O</i> -methyltransferase	59%, WP_005321729.1, <i>Streptomyces pristinaespiralis</i>	WP_051742397.1	
ORF16	BI06_RS32385	Leu B	394	glycosyl transferase	55%, WP_019074879.1, <i>Streptomyces</i> sp. R1-NS-10	WP_043474103.1	

N/A: Not available due to gapped sequence.

2.2. LeuA and LeuB display broad flexibility towards their substrates

We previously reported that the *epa* BGC of *Kitasatospora* sp. MBT66 (on the genome with Genbank accession number JAIY000000000) is responsible for the assembly of the phenazine backbone in endophenazines A—E, but the essential glycosylation genes remained unresolved.⁴³ The characterization of the leucanicidin biosynthetic pathway provided more insights into the rhamnosylation of the endophenazines. Leucanicidin and the endophenazines (*i.e.* glycosylated endophenazines) contain the same α -L-rhamnosyl substituents, and coexist in the same cultures (MM + 0.5% mannitol + 1% glycerol) of *Kitasatospora* sp. MBT66. Moreover, bioinformatic analysis showed that the *leuAB* sub-cluster is the only locus in the entire genome where genes for a natural product-related glycosyltransferase and methyltransferase co-occur within a 20 kb distance. Because of the absence of any other genes for candidate enzymes that may catalyze this reaction, either in the phenazine BGC or elsewhere on the genome, it is likely that LeuA and LeuB not only decorate bafilomycin A₁ to leucanicidin (Figure 2, A), but also modify endophenazines⁵⁴ into endophenazines⁴³ (Figure 2, B).

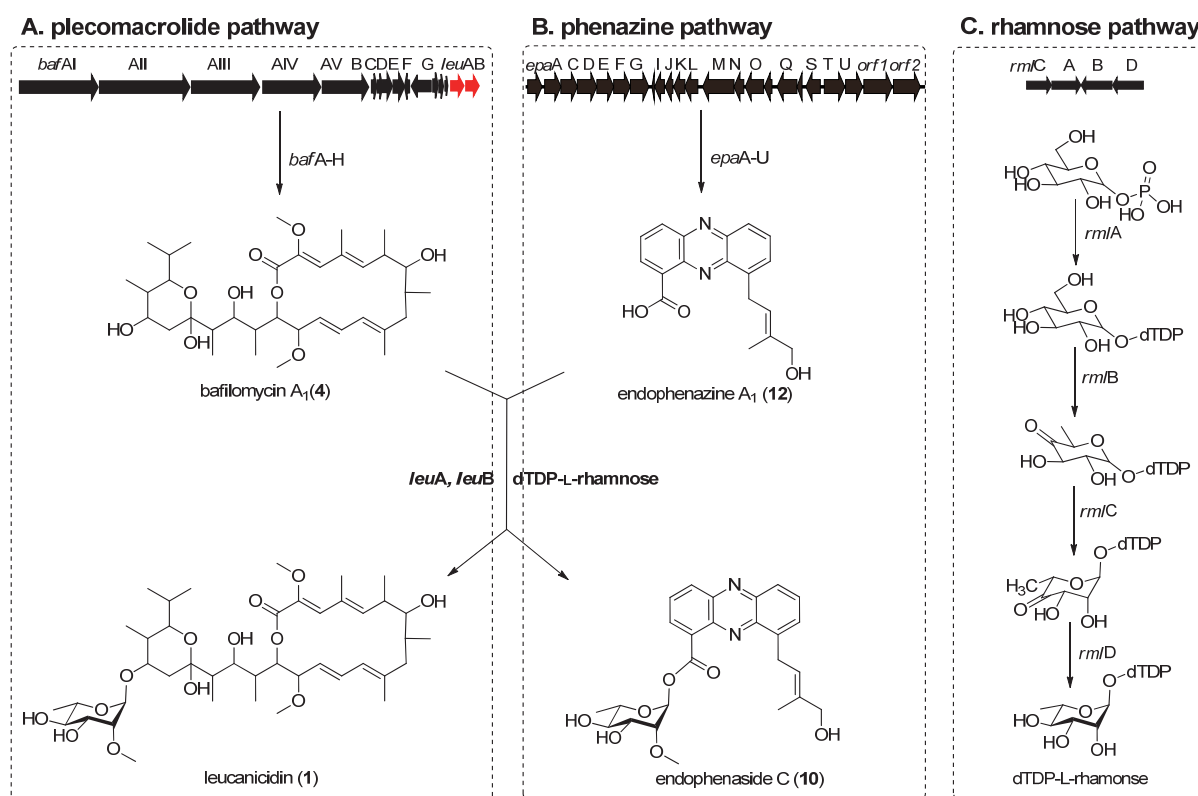


Figure 2. Promiscuous methyl-rhamnosylation of bafilomycin A₁ and endophenazine by methyltransferase LeuA and rhamnosyltransferase LeuB. Biosynthetic “crosstalk” among chromosomally distant clusters enabled *de novo* synthesis of different classes of rhamnosylated natural products. The *bafA*—H (A) and *epaA*—U (B) gene clusters are responsible for plecomacrolide and endophenazine backbone biosynthesis, respectively, while *rmlA*—D (C) supply the dTDP-L-rhamnose building block for rhamnosylation. LeuA and LeuB encoded by the *leuAB* genes immediately adjacent to the *baf* cluster are the tailoring enzymes for the conversion of bafilomycin A₁ to leucanicidin, and of endophenazine A₁ into endophenaside C. The type I PKS bioassembly line for plecomacrolide backbone is detailed in Figure S1, and the shikimate pathway for endophenazine was described previously⁴³.

We first performed a detailed computational genomic analysis of the *leuAB* genes and their homologues. A MultiGeneBlast architecture search with the *leuAB* genes as query on the

full set of 1170 BGCs from the Minimum Information about a Biosynthetic Gene cluster (MIBiG) repository⁶⁴ resulted in seventeen experimentally characterized BGCs that contain homologs of both *leuA* and *leuB*. All hits represented genes involved in the attachment of methylrhannose and (at larger evolutionary distances) related deoxysugars. While a phylogenetic analysis of glycosyltransferase amino acid sequences (Figure 3) showed that the most closely related glycosyltransferases from the set (TiaG1 and TiaG2) are involved in methyl-rhamnosylation of the macrolide tiacumicin B, several other homologous *leuAB*-like sub-clusters are involved in the methyl-rhamnosylation of a wide range of scaffolds, including the indolocarbazole K-252a, the nucleosides A-90289 and caprazamycin, and the anthracyclines elloramycin, steffimycin and aranciamycin. This strongly suggested that this family of glycosyltransferases (exemplified by *LeuB*) has great evolutionary target promiscuity.

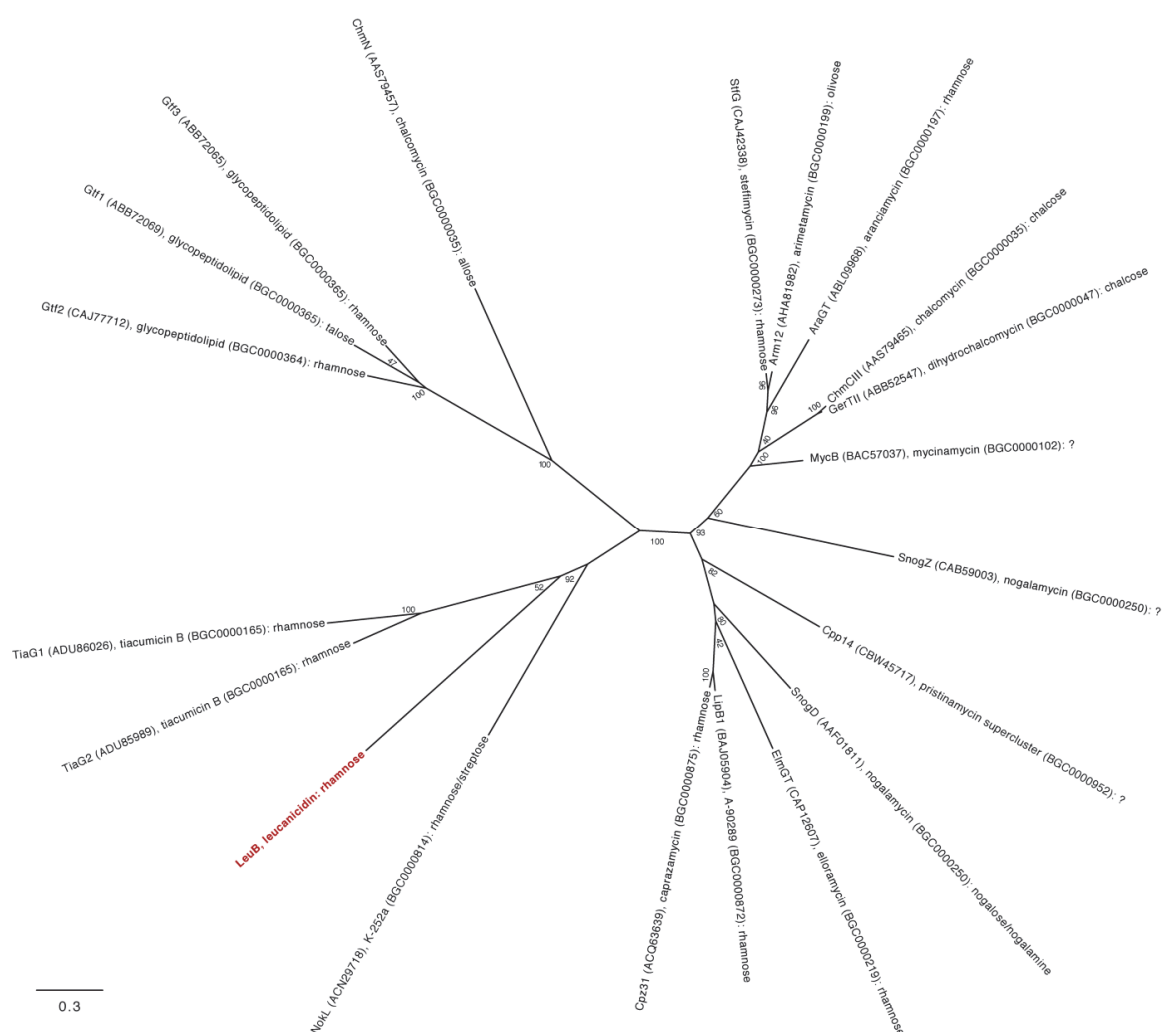


Figure 3. Unrooted maximum likelihood phylogenetic tree of glycosyltransferase *LeuB* and its homologs from homologous sub-clusters. NCBI GenPept accessions, natural product names, MIBiG BGC accessions and substrate specificities (if known) are provided at the tips of each branch.

To create a mutation in *leuB*, we made use of the CRISPR-Cas9 system that was adapted recently for use in actinomycetes.⁶⁵ For this, construct pGWS1002 was introduced into *Kitasatospora* sp. MBT66 by conjugation and ex-conjugants were selected based on their

resistance to apramycin. These ex-conjugants were then propagated to select for loss of the plasmid, and the correct colonies were verified by PCR. After conjugation, single ex-conjugants were streaked onto SFM agar plates containing nalidixic acid and incubated at 30 °C for 3-5 days. Colonies were then grown in liquid TSBS for genomic isolation, followed by PCR using oligonucleotides LeuB_F-370 and LeuB_R+584 (Table S1). PCR products were digested by HindIII, and the desired frame-shift mutants were confirmed by the appearance of 401 bp and 538 bp DNA fragments (Figure S3). The obtained *leuB* frame-shift mutant and its parental strain MBT66 were grown on R5 agar plates with GlcNAc to identify bafilomycins, and on MM agar plates with mannitol and glycerol to identify phenazines. HPLC-UV analysis of R5-grown cultures demonstrated that the production of both leucanicidin (**1**) and bafilomycin A₁ (**4**) was abolished in the *leuB* mutant (Figure 4, A), which indicated that the disruption of *leuB* affected the overall gene expression of the *baf* gene cluster. Importantly, HPLC-UV analysis of cultures grown on MM agar plates showed that the rhamnosylation of the phenazines was also aborted in the *leuB* mutant (Figure 4B), and this was verified by subsequent UHPLC-TOF-MS analysis (Figure 4C). Taken together, analysis of the *leuB* mutant validated the bioinformatics analysis, establishing unequivocally that *leuB* is indeed required for the methyl-rhamnosylation of both types of natural products.

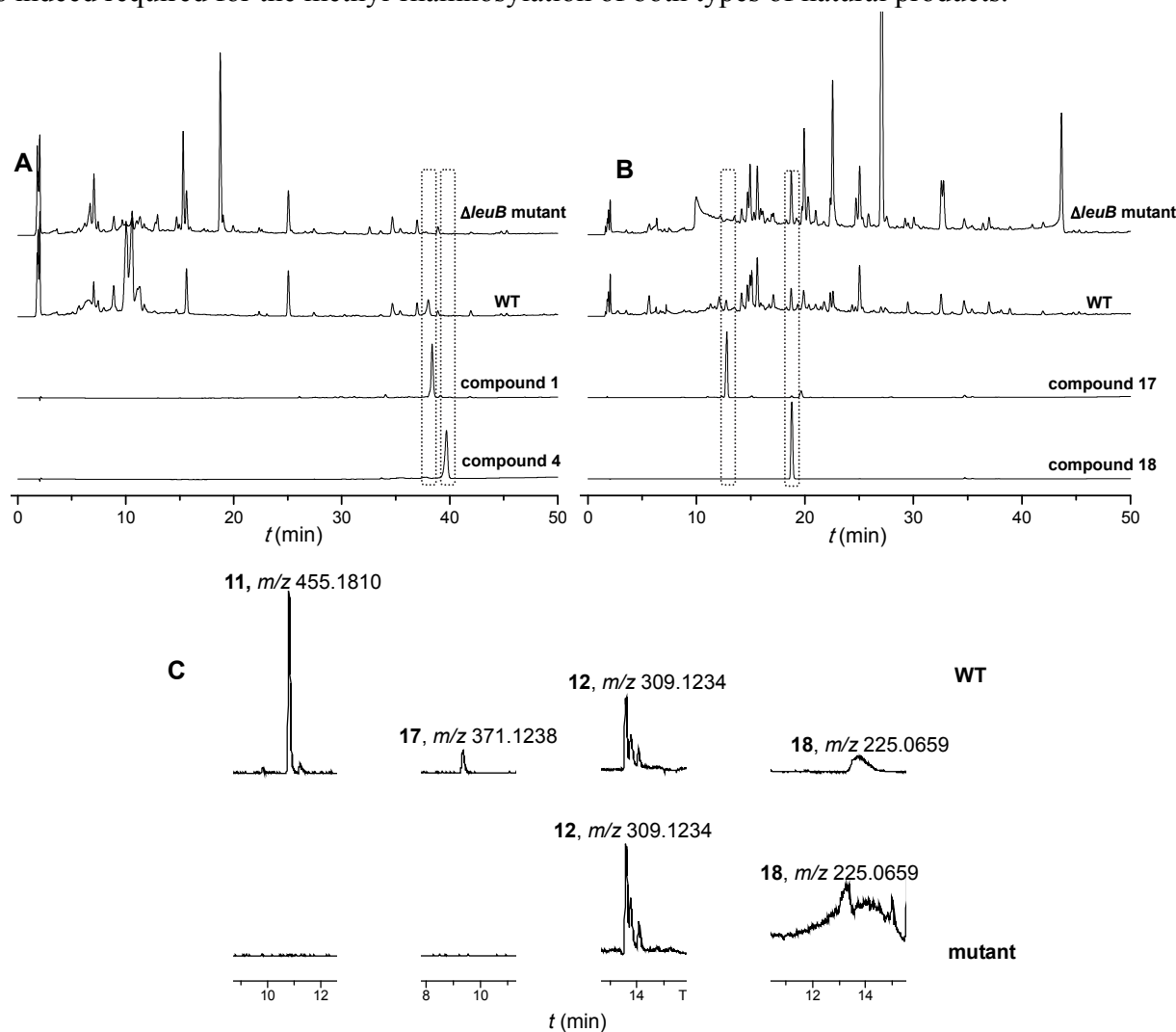


Figure 4. Leucanicidin and endophenaside production by *Kitasatospora* sp. MBT66 and its *leuB* mutant. A) HPLC-UV profile of R5 medium (detected at 254 nm) showing that the production of leucanicidin (**1**) was

abolished in the *leuB* mutant, but without its aglycone bafilomycin A₁ (**4**) accumulation. **B**) HPLC-UV profile of MM medium (detected at 254 nm) showing that the production of endophenaside E (**17**) was abolished in the *leuB* mutant, accompanied by the accumulation of its aglycone phenazine-1-carboxylic acid (**18**). As references we used the chromatographically purified compounds **1**, **17**, and **18** obtained from *Kitasatospora* sp. MBT66, while compound **4** was purchased from Sigma. **C**) Ion chromatography (EIC) of UHPLC-TOF-MS analysis further confirmed that the rhamnosylation of endophenaside B (**11**, *m/z* 455.1810) and E (**17**, *m/z* 371.1238) was indeed abrogated in the *leuB* mutant, while the production of the corresponding aglycones endophenazine A₁ (**12**, *m/z* 309.1234) and phenazine-1-carboxylic acid (**18**, *m/z* 225.0659) was not affected. EIC(s) was pairwise compared between wild-type *Kitasatospora* sp. MBT66 (WT) and its *leuB* mutant.

The promiscuity of LeuAB is not an exception, and indeed glycodiversification of bacterial secondary metabolites may arise from many glycosyltransferases with high substrate promiscuity towards either the (deoxy)sugar donors or the aglycones. For example, the versatile macrolide glycosyltransferase OleD tolerates a wide variety of aglycones including aromatics, coumarins, flavanols and macrolides, remarkably generating three different type of glycosidic bonding (*O*-, *S*- and *N*-glycoside).⁶⁶ The flexible glycosyltransferase GtfE uses variant NDP-sugars to generate glycorandomized vancomycin-analogues that rival vancomycin.¹⁸

The next question to answer was, how is the rhamnosyl substrate for LeuB synthesized? Neither the *baf* nor the *epa* BGC contained components for biosynthesis of NDP-activated rhamnose. Scanning the *Kitasatospora* sp. MBT66 genome identified *rmlABCD* as the likely biosynthetic genes for the activated rhamnose moiety (Table 2), which are required for *de novo* biosynthesis of dTDP-L-rhamnose from D-glucose-1-phosphate (Figure 2, C).^{67,68} This dTDP-L-rhamnose then serves as a substrate for the LeuB-mediated rhamnosyl transfer to the aglycones.^{69,70} When *Kitasatospora* sp. MBT66 was grown in MM supplemented with different carbon sources, glucose was the best carbon source for simultaneous production of endophenasides as well as of leucanicidin, but addition of rhamnose did not improve the levels of these NPs. The latter is consistent with dTDP-L-rhamnose being the substrate for glycosylation.

Table 2. Organization of the *rml* gene cluster for dTDP-L-rhamnose biosynthesis in *Kitasatospora* sp. MBT66.

ORF	Locus tag	Protein	Length	Annotation	Nearest homologue Homology, Protein, origin	GenPept accession
ORF1	BI06_RS21780	RlmC	202	dTDP-4-dehydrorhamnose-3,5-epimerase	98%, WP_045937815.1, <i>Streptomyces</i> sp. NRRL S-495	WP_030393284.1
ORF2	BI06_RS21775	RlmA	291	glucose-1-phosphate thymidyltransferase	99%, WP_045937814.1, <i>Streptomyces</i> sp. NRRL S-495	WP_030393285.1
ORF3	BI06_RS21770	RlmB	321	dTDP-glucose-4,6-dehydratase	94%, WP_030232436.1, <i>Streptomyces</i> sp. NRRL S-350	WP_030393286.1
ORF4	BI06_RS21765	RlmD	311	Putative dTDP-4-keto-L-rhamnose reductase	83%, WP_037899046.1, <i>Streptomyces</i> sp. NRRL S-350	WP_051741708.1

2.3. Molecular networking-driven discovery of new endophenasides

The simultaneous occurrence of both 2'-*O*-methylated and 2'-*O*-unmethylated variants, such as leucanicidin (**1**) and bafilomycin A₁-21-*O*-(α -L-rhamnopyranodise) (**3**), together with endophenaside C (**10**) and endophenaside B (**11**), suggested that either LeuB has a relatively broad substrate specificity by accepting both unmethylated and 2'-*O*-methylated dTDP-L-rhamnose as sugar donors, or that LeuA regioselectively methylates the 2'-OH group after rhamnosyltransfer, regardless of the precise chemical topology of the aglycones. Based on this, we hypothesize that *Kitasatospora* sp. MBT66 may therefore have the potential to produce the corresponding counterparts of endophenasides A (**5**), D (**14**), and E (**17**),⁴³ which

could have been missed in our prior chemical investigation due to intrinsic low yields. To test this, crude extracts of *Kitasatospora* sp. MBT66 grown on MM + 1% glucose were subjected to MS/MS-based molecular networking analysis (Table S2).^{71–74} The fundamental principle is based on the fact that structurally related natural products are typically characterized by similar MS/MS fragmentation patterns. The MS/MS structural relatedness among molecules can be detected in an automated manner, and can subsequently generate a molecular network wherein analogues cluster together. As a result, a network of secondary metabolites produced by *Kitasatospora* sp. MBT66 was created (Figure S4), which contains subnetworks for phenazine-type molecules (Figure 5). This among others identified endophenaside E (**17**) at m/z 371 and its 2'-*O*-methylated congener (**16**) at m/z 385. Two peaks at m/z 469 corresponded to endophenaside C (**10**) and 2'-*O*-methylated-endophenaside D (**13**). Moreover, the networking analysis presented many molecular features that could not be assigned to any of the previously identified endophenazines⁵⁴ or endophenasides,⁴³ strongly suggesting that *Kitasatospora* sp. MBT66 produces many other and likely novel phenazine-type compounds.

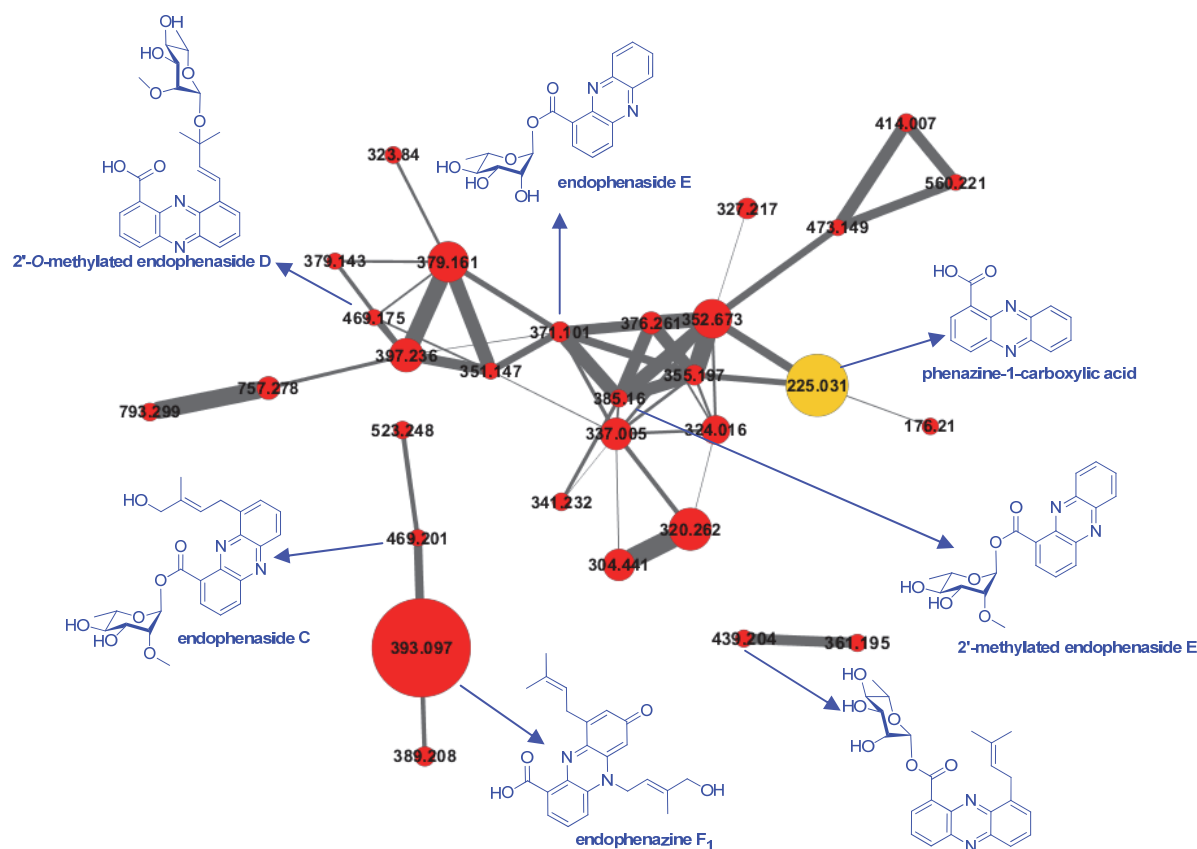


Figure 5. Molecular networking of endophenasides produced by *Kitasatospora* sp. MBT66. The size of the nodes corresponds to the signal intensities of the compounds (see Table S2) and the thickness of the edge between connecting nodes defines the degree of similarity of the MS/MS spectra. The full network of secondary metabolites produced by strain *Kitasatospora* sp. MBT66 is presented in Figure S4.

To characterize these putative new phenazine derivatives, another round of up-scale fermentation of *Kitasatospora* sp. MBT66 followed by compound purification and identification was performed. As previous studies showed that endophenaside E (**17**) was prone to methanolysis, the use of methanol as solvent during isolation was avoided. UV-

and/or MS-guided fractionation indeed resulted in the elucidation of nine new phenazine glycosides **6**–**8**, **16**, and **22**–**26**, as well as their corresponding known endophenazine aglycones **9**, **12**, **15**, **18**, and **19**–**21**. Further purification of **13** was not feasible due to its extremely low abundance. The follow-up ^1H NMR (Table S3) measurements unambiguously confirmed the planar structures of new compounds **6**–**8**, **16**, **25**, and **26**, by comparison with those of known endophenazines A–E⁴³ and with the endophenazines⁵⁴, all of which were confirmed by high resolution mass spectrometry (for details see Supplementary data file). While the ^1H NMR spectrum of **22** revealed features typical of endophenazines, the phenazine core was elucidated as 1,7-disubstituted based on the coupling of H-6 (δ 8.57, d, J = 1.2 Hz), H-8 (δ 8.34, dd, J = 9.0, 1.2 Hz), and H-9 (δ 8.60, d, J = 9.0 Hz). In addition, the substituent at C-7 was further identified as a benzoyl group and confirmed by HRMS at m/z 475.1482 (calculated for $\text{C}_{26}\text{H}_{23}\text{N}_2\text{O}_7$ 475.1500). Exceptionally, the NMR spectrum of a mixture of **23** and **24** presented two sets of characteristic glycerol signals instead of the usual rhamnose in the δ 3.5–5.0 region. The different ester linkage was distinguished by the downfield shift of H-1' (δ 4.63 and 4.57) in **23** and H-2' (δ 5.43) in **24**, originated from the shielding effect of carbonyl group (C-11). Though these isomers were chromatographically separable (Figure 6, A), they were spontaneously interconverted into one another, with a fast asymmetric equilibrium of **23** around 7.5 times that of **24**. The tautomerization likely originated from the spatially vicinal –OH group on the glycerol side chain, which could serve as an alcoholic reagent for ester bond, while the phenazine-1-carboxyl performed as an anchor (or carrier) for self-refresh esterification of three hydroxyls (Figure 6, B).

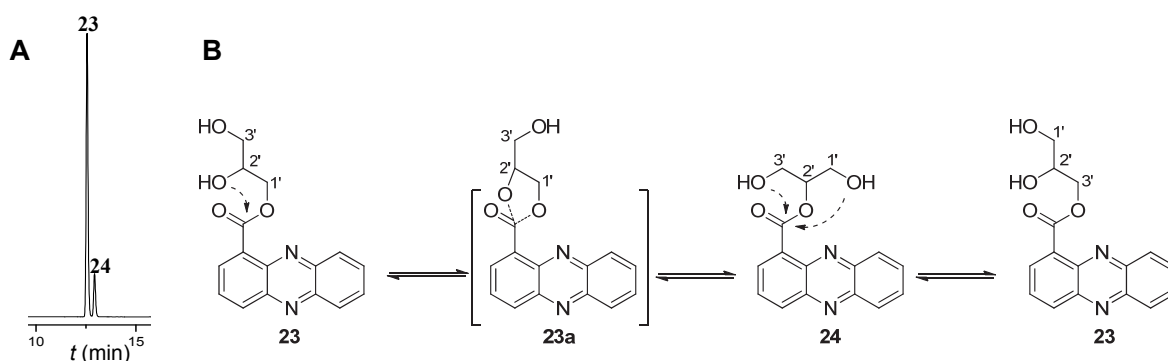


Figure 6. Diagram for the tautomerism between glyceride phenazines **23 and **24**.** HPLC analysis revealed the spontaneous interconversion of **23** and **24**. When the equilibrium was reached, compound **23** was 7.5 times **24** (A). An explanatory mechanism for this phenomenon is that all the three hydroxyl groups can form ester bonds with phenazine-1-carboxylic acid. The plausible transition intermediate **23a** with a relatively stable five-member ring system probably mediated intramolecular exchange of glycerol esterification (B).

To get an idea of the antimicrobial activity of the newly isolated endophenazines, antimicrobial assays were performed using compounds **8**, **16**, and **23**–**25** that represented substitute variation of rhamnosylated, glycerolated, and prenylated phenazines, respectively (Table 3). The assays were done according to the method that was also used for endophenazines A–E.⁴³ All glycosylated phenazines that have been tested inhibited growth of the Gram-negative bacteria *Escherichia coli* K12 and *Pseudomonas aeruginosa* PAO1, which was in agreement with the data obtained for endophenazines A–E.⁴³ The compounds, and in particular phenazines **23** and **24**, also inhibited growth of *Bacillus subtilis* 168, while conversely, the compounds had a negligible efficacy against the Gram-positive bacterium

Staphylococcus aureus CECT976.**Table 3.** Antimicrobial activity of representative new endophenazines **8**, **16**, and **23–25**. For each compound, 25 µl was spotted of a 2 mg/ml solution in methanol.

Compound NO.	Inhibition zone (mm)			
	<i>Bacillus subtilis</i>	<i>Escherichia coli</i>	<i>Staphylococcus aureus</i>	<i>Pseudomonas aeruginosa</i>
8	10	10	0	10
16	14	9	0	9
23, 24	20	10	0	11
25	11	9	0	9
AMP ^a	32	23	35	0
STR	13	18	17	15
NC	0	0	0	0

^a AMP, ampicillin; STR, streptomycin; NC, negative control (methanol)

3. CONCLUSION

Characterization of the genetic basis for the glycosylation is of utmost significance, because it can expedite the downstream biochemical investigation to refresh the chemistry of natural products and accordingly optimize their pharmaceutical properties. In this study, we characterized a type I PKS gene cluster that encodes the biosynthesis of the rhamnosylated plecomacrolide antibiotic leucanicidin. However, the previously described NK155141 (19-methyl-leucanicidin) was shown not to be produced *in vivo*, but an artifact from the reaction of leucanicidin with the solvent methanol. The gene cluster includes the *leuAB* genes for the methyl-rhamnosylation of bafilomycin A₁ to leucanicidin. Besides plecomacrolides, *Kitasatospora* sp. MBT66 also produces a range of phenazines, including endophenazines that are phenazines decorated with the same methyl-rhamnosyl group as leucanicidin. MS/MS-based molecular networking guided the further identification of nine new phenazine-type antibiotics, including a pair of interconverting glyceride phenazines. Since LeuA and LeuB are the only enzymes for the methyl-rhamnosylation of natural products encoded by the genome of *Kitasatospora* sp. MBT66, it is likely that these enzymes use both plecomacrolides and phenazines as the substrate. Such promiscuity is surprising but at the same time not unprecedented. In view of the urgent need for new antimicrobials and the challenge of discovering molecules with a novel chemical scaffold, modification of known structures is an attractive alternative to obtain molecules with novel bioactivities and pharmacokinetic properties. The inherent flexibility of the LeuA and LeuB enzymes described in this work may be applicable for the glycorandomization of a broad range of natural products and the same may well be true for other glycosyltransferases. The application of such promiscuous natural product tailoring for drug discovery is currently under investigation in our laboratory.

4. EXPERIMENTAL SECTION

4.1. Strains and culturing conditions

Kitasatospora sp. MBT66 was described previously.^{43,52} As growth media we used liquid minimal media (NMMP),⁵⁷ minimal media agar plates (MM),⁵⁷ solid R5 agar plates,⁵⁸ and soy flower medium (SFM) agar plates. Culture media were supplemented with different carbon sources (glycerol, mannitol, glucose, rhamnose, *N*-acetylglucosamine), additional

additives (yeast extract, potato extract, peptone, starch, soy flower), or high alkalinity (pH 10). Agar plates (12 cm × 12 cm petri dishes) were inoculated with 5×10^7 spores from a fresh spore suspension, and incubated at 30 °C for 7 days. For liquid-grown cultures, 50 ml of NMMP media with additives were inoculated with 5×10^7 spores in 250 ml flasks equipped with a spring, and grown at 30 °C with constant shaking at 220 rpm for 7 days.

4.2. Extraction and isolation of metabolites

The extraction of metabolites produced by *Kitasatospora* sp. MBT66 basically followed our method published previously.⁴³ Briefly, after 7 days of incubation, agar plates were cut into pieces and soaked in ethyl acetate (EtOAc) overnight at room temperature. The EtOAc was removed under vacuum at 40 °C and the residue was dissolved in methanol (MeOH) or acetonitrile (ACN) for HPLC-UV and/or UHPLC-TOF-MS analysis.

The first round of isolation was done to enable chemical investigation of *Kitasatospora* sp. MBT66.⁴³ After 6 days of incubation of MBT66 in 5 liter of MM supplemented with 1% glycerol and 0.5% mannitol (w/v), the EtOAc-soluble component (2.0 g) was fractionated on a Macroporous resin Diaion HP-20 from Supelco (Bellefonte, PA, USA) by eluting stepwise from H₂O to MeOH, to give 20%, 40%, 60%, and 80% MeOH fractions. Previous investigation of the 40% (v/v) MeOH fraction identified the novel endophenazines A—E.⁴³ Here, the 80% (v/v) MeOH fraction was further separated by silica gel (pore size 60 Å, 70–230 mesh, St. Louis, MO, USA) column chromatography employing gradient elution from CHCl₃ to MeOH, to give 14 subfractions (sfr.1–sfr.14). Sfr.11 was purified by semi-preparative reversed-phase HPLC (Phenomenex Luna C18 (2) 100 Å 5 micron 250 × 10 mm) on a Shimadzu HPLC system and a 5 ml Rheodyne manual injection loop, eluting with a gradient of MeOH in H₂O from 80% to 100%, to isolate compound **1** (t_R = 25.75 min, 0.65 mg) and **2** (t_R = 28.10 min, 0.50 mg). The 60% (v/v) MeOH fraction was defatted with *n*-hexane, which was further separated by semi-preparative reversed-phase HPLC (Phenomenex Luna C18 (2) 100 Å 5 micron 250 × 10 mm) on an Agilent 1200 series HPLC apparatus (Agilent technologies Inc, Santa Clara, CA, USA), eluting with a gradient of ACN in H₂O from 20% to 100% at flow rate 2 ml/min in 40 min. The peaks detected in the HPLC chromatogram at 254 nm, were manually collected, which gave semi-pure **7** (t_R = 30.12 min, 0.58 mg), **8** (t_R = 25.48 min, 1.19 mg), semi-pure **9** (t_R = 38.16 min, 0.46 mg), **10** (t_R = 18.90 min, 0.82 mg), a mixture of **12** and **15** (t_R = 20.48 min, 0.77 mg), **14** (t_R = 17.24 min, 0.70 mg), **19** (t_R = 19.61 min, 1.50 mg), and impure **22** (t_R = 22.22 min, 0.20 mg).

The second round of isolation was performed on 1 liter of *Kitasatospora* sp. MBT66 culture grown under the same conditions as above, but using different fractionation methods. Specifically, after EtOAc extraction, 0.5 g of extract was first separated using silica gel, employing a gradient elution by acetone/*n*-hexane as the solvent. All fractions were pooled into seven fractions based on TLC (Silica gel 60 F₂₅₄, Merck, Darmstadt, Germany) detection under UV light 254 nm. These fractions (Fr.1—Fr.7) dissolved in ACN, were analyzed in parallel by TLC, HPLC-UV, and UHPLC-TOF-MS (see below). Fr.3 and Fr.4 that contained compounds that displayed a UV spectrum and mass typical of endophenazines or endophenazines^{43,54} were purified further. Fr.3 was separated by semi-preparative reversed-phase HPLC (Phenomenex Luna C18 (2) 100 Å 5 micron 250 × 10 mm) on an Agilent 1200 series HPLC (Agilent technologies Inc, Santa Clara, CA, USA), eluting with a

gradient of ACN in H₂O adjusted with 0.1% TFA from 5% to 36%. HPLC peaks were manually collected, resulting in the isolation of compound **6** (semi-pure; t_R = 28.55 min, 0.25 mg), **21** (semi-pure; t_R = 30.49 min, 0.79 mg), **26** (t_R = 32.48 min, 0.32 mg), **20** (semi-pure; t_R = 33.74 min, 0.82 mg), **19** (semi-pure; t_R = 35.44 min, 0.73 mg), **25** (t_R = 37.97 min, 0.63 mg), and **18** (t_R = 43.40 min, 2.08 mg). Fr.4 was further subjected to preparative TLC (PLC Silica gel 60 F₂₅₄, 1 mm, Merck, Darmstadt, Germany), migrated with solvent system *n*-hexane/acetone (1:1) and detected under UV light 254 nm. Two closely migrating dark bands were scraped off and rinsed with acetone, and identified as semi-pure compound **17** (1.62 mg) and coded as Fr.4-1, respectively. Subsequently, Fr.4-1 was separated using reversed-phase chromatography on an Agilent 1200 HPLC, which was separated into **23** (t_R = 18.89 min, 2.25 mg; convertible into **24**), **24** (t_R = 19.68 min, 0.35 mg; convertible into **23**), **16** (t_R = 21.30 min, 0.82 mg), and **18** (semi-pure; t_R = 30.31 min, 0.20 mg), eluting with a gradient of ACN in H₂O from 30% to 100%.

4.3. Antimicrobial activity assays

Antimicrobial activity of purified new compounds was determined according to a disc diffusion method as described.⁴³ 25 μ l of the novel endophenaside compounds **8**, **16**, and **23–25** (2 mg/ml in methanol) was spotted onto paper discs (6 mm diameter) placed on agar plates containing a soft agar overlay with indicator bacteria. Indicator bacteria were *Bacillus subtilis* 168, *Escherichia coli* K12, *Staphylococcus aureus* CECT976, or *Pseudomonas aeruginosa* PAO1. Ampicillin and streptomycin were used as positive controls, and the solvent methanol as the negative control. After incubation at 37 °C for 18 h, growth inhibition zones (in mm) were recorded as antimicrobial activity.

4.4. NMR measurements

NMR sample preparation and measurements were performed according to a protocol that was published previously.^{75,76} Briefly, 500 μ l of methanol-*d*₄ were added to freeze-dried samples, and the resultant mixtures were vortexed for 10 sec and sonicated for 20 min at 42 kHz using an Ultrasonicator 5510E-MT (Branson, Danbury, CT, USA), followed by centrifugation at 16,000x *g* at room temperature for 5 min. The supernatant (300 μ l) was transferred to a 3 mm micro NMR tube and analyzed. The ¹H NMR spectra were recorded at 25 °C on a 600 MHz Bruker DMX-600 spectrometer (Bruker, Karlsruhe, Germany) operating at a proton NMR frequency of 600.13 MHz. Deuterated methanol was used as the internal lock. Each ¹H NMR spectrum consisted of 128 scans using the following parameters: 0.16 Hz/point, pulse width (PW) = 30 (11.3 s) and relaxation delay (RD) = 1.5 s. Free induction decays (FIDs) were Fourier transformed with a line broadening (LB) = 0.3 Hz. The resulting spectra were manually phased and baseline corrected, and calibrated to residual methanol-*d*₄ at 3.30 ppm, using MestReNova 8.1.

4.5. HPLC-UV analysis

HPLC analysis was performed with an Agilent 1200 series HPLC apparatus (Agilent technologies Inc, Santa Clara, CA, USA), using a 150 $\times$ 4.6 mm Luna 5 micron C18 (2) 100 Å column equipped with a guard column containing C18 4 $\times$ 3 mm cartridges (Phenomenex Inc, Torrance, CA, USA). The mobile phase consisted of water (A) and acetonitrile (B, HPLC

grade) in a linear gradient program from 10% B to 100% B in 50 minutes at a flow rate of 1.0 ml/min. Chromatograms were recorded at 210 nm, 254 nm, and 280 nm. The injection volume was 10 μ l.

4.6. UHPLC-ToF-MS analysis

Mass analysis was performed on an Agilent 1290 UHPLC coupled to a Bruker Daltonics microToF-QII equipped with standard electrospray source. The instrument was fitted with a Kinetex C18 column (50cm x 2.1 mm, 100 Å). A linear gradient analysis from 5% B to 100% B was performed over 25 minutes with mobile phase A (H₂O with 0.1% formic acid) and mobile phase B (acetonitrile with 0.1% formic acid). For each sample 10 μ l was injected onto the column at a flow rate of 0.5 ml/min. The mass spectrometer was programmed to acquire MS/MS in a data-dependent manner, acquiring 5 MS/MS scans following each precursor MS¹ scan.

4.7. MS/MS-based molecular networking

Mass spectral networks were assembled as described in reference.⁷⁷ Tandem MS spectra were clustered using MS-Clustering⁷⁸ that builds consensus spectra for repeatedly observed ions (this was performed using the natural product analysis infrastructure at <http://gnps.ucsd.edu>). The MS² spectra were scored based on their similarity; a cosine score of 1 indicates identical spectra, while a cosine score of 0 indicates no similarity. The cosine score threshold to make a match was set to 0.7 and the minimum matched peak was 6. The algorithm assumed a parent peak mass tolerance of 2.0 Da and an MS² peak tolerance of 0.5 Da. The networks were visualized with Cytoscape software, whereby consensus spectra are represented as nodes connected by edges to aligning nodes. The thickness of the edge indicates the level of similarity between the nodes. The FM3 layout was used to organize and align nodes within the network. The data is available as MSV000079139 and MSV000079279. at <http://gnps.ucsd.edu>.

4.8. Bioinformatics

Genome sequencing and annotation of *Kitasatospora* sp. MBT66 was described previously,⁵² and is available under Genbank accession number JAIY00000000. MultiGeneBlast architecture searches to exhaustively scan the *Kitasatospora* sp. MBT66 genome for loci with glycosyltransferases and methyltransferases were performed using default parameters, using a dataset of glycosyltransferases (accessions AAN65238, AAN65243, AAG29794, AAG29803, AAK83176, AAL06683, ABY66020, ABY66027, AAM77991, AAM77984, CAC93718, ABC02795, ABC02796, BAC55213, BAC55218, CCD33145, BAA84598, AAG23269, AAG23270, AAG23272, AAG23280, ABI22137, ABI22145, AAS79443, and AAU93810) and methyltransferases (accessions AAN65229, AAG29785, AAC01731, CBA11567, AAK83182, AAK83192, AAK83193, AAM77992, CAC93713, BAC55209, CAA76551, CAA76552, CAA76553, CCD33143, CCD33144, AAG23268, and AAZ94402) known to be involved in natural product methylglycosylation. MultiGeneBlast searches of the *leuAB* sub-cluster against MIBiG BGCs⁶⁴ were performed with a 15% sequence identity threshold and otherwise default parameters. Multiple sequence alignment of glycosyltransferases was performed using Muscle v3.8.31,⁷⁹ using default parameters. Phylogenetic trees were

calculated in MEGA 6.06,⁸⁰ using the maximum likelihood method, with 100 bootstrap replicates.

4.9. Construction of a *leuB* frame-shift mutant using the CRISPR/Cas9 system

All oligonucleotides used in this work are listed in Table S1. The *leuB* spacer insert was generated by annealing oligonucleotides LeuB_spacer_For and LeuB_spacer_Rev. By using Golden Gate assembly,⁸¹ the generated *leuB* spacer insert was cloned into plasmid pCRISPomyces-2,⁶⁵ which was obtained from Addgene (Plasmid #61737), to generate construct pGWS1001. The approximately 1 kb left and right flanking regions of the *leuB* editing template were amplified by PCR from the genomic DNA of *Kitasatospora* sp. MBT66 using primer pairs LeuB_LF-976_EX + LeuB_LR+33_H and LeuB_FS_RF+47_H + LeuB_RR+1032_EX, respectively. PCRs were done as described.⁸² Fragments were then digested with XbaI and HindIII, and ligated into pGWS1001 to generate construct pGWS1002. In the *leuB* editing template, the +34/+46 part of *leuB* (whereby +1 refers to the first nt of the translational start codon) was erased and replaced by a HindIII site, so as to introduce a frame-shift in *leuB*. The correct construct assembly was confirmed by DNA sequencing (BaseClear B.V., Leiden, the Netherlands).

Construct pGWS1002 was then introduced into the parental strain HM125 by conjugation as described.⁵⁷ *E. coli* ET12567/pUZ8002 containing pGWS1002 were incubated to OD₆₀₀ of 0.4-0.6, and mixed with 10⁸ spores. Ex-conjugants were selected by overlaying each plate with water containing apramycin (1 mg) and nalidixic acid (500 µg). After conjugation, single ex-conjugants were streaked onto SFM agar plates containing nalidixic acid and incubated at 30 °C for 3-5 days. Colonies were then grown in liquid TSBS for genomic isolation, followed by PCR using oligonucleotides LeuB_F-370 and LeuB_R+584. PCR products were digested by HindIII, whereby the correct mutants were identified by bands of 401bp and 538bp, followed by DNA sequencing.

REFERENCES

- (1) WHO Geneva, Switzerland, 2014.
- (2) Genilloud, O.; González, I.; Salazar, O.; Martín, J.; Tormo, J. R.; Vicente, F. J. *Ind. Microbiol. Biotechnol.* **2011**, 38, 375–389.
- (3) Baltz, R. H. *Curr. Opin. Pharmacol.* **2008**, 8, 557–563.
- (4) Cooper, M. A.; Shlaes, D. *Nature* **2011**, 472, 32.
- (5) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat. Rev. Drug Discov.* **2007**, 6, 29–40.
- (6) Barka, E. A.; Vatsa, P.; Sanchez, L.; Gaveau-vaillant, N.; Jacquard, C.; Klenk, H.; Clément, C.; Ouhdouch, Y.; van Wezel, Gilles, P. *Microbiol. Mol. Biol. Rev.* **2016**, 80, 1–43.
- (7) Hopwood, D. A. *Streptomyces in nature and medicine: the antibiotic makers*; Oxford

- University Press: New York, 2007.
- (8) Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S. *Nat. Biotechnol.* **2003**, *21*, 526–531.
 - (9) Cruz-Morales, P.; Vijgenboom, E.; Iruegas-Bocardo, F.; Girard, G.; Yáñez-Guerra, L. A.; Ramos-Aboites, H. E.; Pernodet, J.-L.; Anné, J.; van Wezel, G. P.; Barona-Gómez, F. *Genome Biol. Evol.* **2013**, *5*, 1165–1175.
 - (10) Bentley, S. D.; Cerdeño-Tárraga, K. F. C. A.-M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Bentley, S. D.; Harper, D.; Bateman, A.; Brown, S.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Larke, L.; Murphy, L.; Oliver, K.; Rabbinowitsch, E.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Woodward, J.; Barrell, B. G.; Parkhill, J. *Nature* **2002**, *3*, 141–147.
 - (11) van Wezel, G. P.; McKenzie, N. L.; Nodwell, J. R. *Chapter 5. Applying the genetics of secondary metabolism in model actinomycetes to the discovery of new antibiotics.*; 1st ed.; Elsevier Inc., 2009; Vol. 458.
 - (12) Nett, M.; Ikeda, H.; Moore, B. S. *Nat. Prod. Rep.* **2009**, *26*, 1362–1384.
 - (13) Medema, M. H.; Breitling, R.; Bovenberg, R.; Takano, E. *Nat. Rev. Microbiol.* **2011**, *9*, 131–137.
 - (14) Gross, H. *Curr. Opin. Drug Discov. Devel.* **2009**, *12*, 207–219.
 - (15) Elshahawi, S. I.; Shaaban, K. A.; Kharel, M. K.; Thorson, J. S. *Chem. Soc. Rev.* **2015**, *44*, 7591–7697.
 - (16) La Ferla, B.; Airoidi, C.; Zona, C.; Orsato, A.; Cardona, F.; Merlo, S.; Sironi, E.; D’Orazio, G.; Nicotra, F. *Nat. Prod. Rep.* **2011**, *28*, 630–648.
 - (17) Ikeda, H.; Nonomiya, T.; Usami, M.; Ohta, T.; Omura, S. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9509–9514.
 - (18) Fu, X.; Albermann, C.; Jiang, J.; Liao, J.; Zhang, C.; Thorson, J. S. *Nat. Biotechnol.* **2003**, *21*, 1467–1469.
 - (19) Yang, M.; Proctor, M. R.; Bolam, D. N.; Errey, J. C.; Field, R. A.; Gilbert, H. J.; Davis, B. G. *J. Am. Chem. Soc.* **2005**, *127*, 9336–9337.
 - (20) Barton, W. A.; Biggins, J. B.; Jiang, J.; Thorson, J. S.; Nikolov, D. B. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 13397–13402.
 - (21) Gantt, R. W.; Peltier-Pain, P.; Thorson, J. S. *Nat. Prod. Rep.* **2011**, *28*, 1811–1853.
 - (22) Blanchard, S.; Thorson, J. S. *Curr. Opin. Chem. Biol.* **2006**, *10*, 263–271.
 - (23) Bindseil, K. U.; Zeeck, A. *Liebigs Ann. Chem.* **1994**, 305–312.
 - (24) Dröse, S.; Bindseil, K. U.; Bowman, E. J.; Siebers, A.; Zeeck, A.; Altendorf, K. *Biochemistry* **1993**, *32*, 3902–3906.
 - (25) Yoshimoto, Y.; Jyojima, T.; Arita, T.; Ueda, M.; Imoto, M.; Matsumura, S.; Toshima, K. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3525–3528.
 - (26) Wilton, J. H.; Hokanson, G. C.; French, J. C. *J. Antibiot. (Tokyo)*. **1985**, *38*, 1449–1452.
 - (27) Poeta, M. Del; Cruz, M. C.; Cardenas, M. E.; Perfect, J. R.; Heitman, J. *Antimicrob. Agents Chemother.* **2000**, *44*, 739–746.
 - (28) Goetz, M. A.; McCormick, P. A.; Monaghan, R. L.; Ostlind, D. A. *J. Antibiot. (Tokyo)*. **1985**, *38*, 161–168.
 - (29) Vani, Z. K.; Curdova, E. *Folia Microbiol* **1991**, *36*, 99–111.

- (30) Bowman, E. J.; Siebers, A.; Altendorf, K. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 7972–7976.
- (31) Werner, G.; Hagenmaier, H. *J. Antibiot. (Tokyo)*. **1984**, *37*, 110–117.
- (32) Li, J.; Lu, C.; Shen, Y. *J. Antibiot. (Tokyo)*. **2010**, *63*, 595–599.
- (33) Yim, S.; Yu, Z.; Zhao, L.; Jiang, C.; Duan, Y.; Wong, L.; Carver, K. C.; Schuler, L. A.; Shen, B. *J. Antibiot. (Tokyo)*. **2010**, *64*, 159–162.
- (34) Zhang, D.; Wei, G.; Wang, Y.; Si, C.; Tian, L.; Tao, L.; Li, Y. *J. Antibiot. (Tokyo)*. **2011**, *64*, 391–393.
- (35) Haydock, S. F.; Appleyard, A. N.; Mironenko, T.; Lester, J.; Scott, N.; Leadlay, P. F. *Microbiology* **2005**, *151*, 3161–3169.
- (36) Seto, H.; Tajima, I.; Akao, H.; Furihata, K.; Otake, N. *J. Antibiot. (Tokyo)*. **1984**, *37*, 610–613.
- (37) Otoguro, K.; Nakagawa, A.; Omura, S. *J. Antibiot. (Tokyo)*. **1988**, *41*, 250–252.
- (38) Ohta, E.; Kubota, N. K.; Ohta, S.; Suzuki, M.; Ogawa, T.; Yamasaki, A.; Ikegami, S. *Tetrahedron* **2001**, *57*, 8463–8467.
- (39) Igarashi, M.; Nakamura, H.; Takeuchi, T. *J. Antibiot. (Tokyo)*. **1997**, *50*, 932–936.
- (40) Dröse, S.; Boddien, C.; Gassel, M. *Biochemistry* **2001**, *40*, 2816–2825.
- (41) Ichikawa, N.; Oguchi, A.; Ikeda, H.; Ishikawa, J.; Kitani, S.; Watanabe, Y.; Nakamura, S.; Katano, Y.; Kishi, E.; Sasagawa, M.; Ankai, A.; Fukui, S.; Hashimoto, Y.; Kamata, S.; Otoguro, M.; Tanikawa, S.; Nihira, T.; Horinouchi, S.; Ohnishi, Y.; Hayakawa, M.; Kuzuyama, T.; Arisawa, A.; Nomoto, F.; Miura, H.; Takahashi, Y.; Fujita, N. *DNA Res.* **2010**, *17*, 393–406.
- (42) Moon, S.S.; Hwang, W. H. *J. Antibiot. (Tokyo)*. **2003**, *56*, 856–861.
- (43) Wu, C.; van Wezel, G. P.; Choi, Y. H. *J. Antibiot. (Tokyo)*. **2015**, *68*, 445–452.
- (44) Takaaki, N.; Mutsuko, K.; Tadashi, I.; Hisao, Y.; Takashi, H.; Seiichi, S.; Nobuyoshi, S. Novel antibiotic substance NK155141 and its production, 1991.
- (45) Isogai, A.; Sakuda, S.; Matsumoto, S.; Ogura, M.; Furihata, K.; Seto, H.; Suzuki, A. *Agric. Biol. Chem.* **1984**, *48*, 1379–1381.
- (46) Shea, M. G. O.; Rickards, R. W.; Rothschild, J. M.; Lacey, E. *J. Antibiot. (Tokyo)*. **1997**, *50*, 1073–1077.
- (47) Zhang, W.; Fortman, J. L.; Carlson, J. C.; Yan, J.; Liu, Y.; Bai, F.; Guan, W.; Jia, J.; Matainaho, T.; Sherman, D. H.; Li, S. *Chembiochem* **2013**, *14*, 301–306.
- (48) Hwang, J. Y.; Kim, H. S.; Kim, S. H.; Oh, H. R.; Nam, D. H. *AMB Express* **2013**, *3*, 24–32.
- (49) Toshima, K.; Jyojima, T.; Yamaguchi, H.; Noguchi, Y.; Yoshida, T.; Murase, H.; Nakata, M.; Matsumura, S. *J. Org. Chem.* **1997**, *62*, 3271–3284.
- (50) Scheidt, K. A.; Bannister, T. D.; Tasaka, A.; Wendt, M. D.; Savall, B. M.; Fegley, G. J.; Roush, W. R. *J. Am. Chem. Soc.* **2002**, *124*, 6981–6990.
- (51) Durham, T. B.; Blanchard, N.; Savall, B. M.; Powell, N. A.; Roush, W. R. *J. Am. Chem. Soc.* **2004**, *126*, 9307–9317.
- (52) Girard, G.; Willemse, J.; Zhu, H.; Claessen, D.; Bukarasam, K.; Goodfellow, M.; van Wezel, G. P. *Antonie Van Leeuwenhoek* **2014**, *106*, 365–380.
- (53) Kersten, R. D.; Ziemert, N.; Gonzalez, D. J.; Duggan, B. M.; Nizet, V.; Dorrestein, P. C.; Moore, B. S. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 4407–4416.

- (54) Heine, D.; Martin, K.; Hertweck, C. *J. Nat. Prod.* **2014**, *77*, 1083–1087.
- (55) Deeg, M.; Hagenmaier, H.; Kretschmer, A. *J. Antibiot. (Tokyo)*. **1987**, *40*, 320–328.
- (56) Wu, C.; Tan, Y.; Gan, M.; Wang, Y.; Guan, Y.; Hu, X.; Zhou, H.; Shang, X.; You, X.; Yang, Z.; Xiao, C. *J. Nat. Prod.* **2013**, *76*, 2153–2157.
- (57) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F. & Hopwood, D. A. John Innes Foundation, Norwich, UK, 2000.
- (58) Hopwood, D. A.; Bibb, M. J.; Chater, K. F.; Kieser, T.; Bruton, C. J.; Kieser, H. M.; Lydiate, D. J.; Smith, C. P.; Ward, J. M. and Schrempf, H. . The John Innes Foundation: Norwich, UK, 1985.
- (59) Zhu, H.; Swierstra, J.; Wu, C.; Girard, G.; Choi, Y. H.; van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. *Microbiology* **2014**, *160*, 1714–1725.
- (60) van Wezel, G. P.; McDowall, K. J. *Nat. Prod. Rep.* **2011**, *28*, 1311–1333.
- (61) Rigali, S.; Titgemeyer, F.; Barends, S.; Mulder, S.; Thomae, A. W.; Hopwood, D. A.; van Wezel, G. P. *EMBO Rep.* **2008**, *9*, 670–675.
- (62) Craig, M.; Lambert, S.; Jourdan, S.; Tenconi, E.; Colson, S.; Maciejewska, M.; Ongena, M.; Martin, J. F.; van Wezel, G. P.; Rigali, S. *Environ. Microbiol. Rep.* **2012**, *4*, 512–521.
- (63) Nothaft, H.; Rigali, S.; Boomsma, B.; Swiatek, M.; McDowall, K. J.; Van Wezel, G. P.; Titgemeyer, F. *Mol. Microbiol.* **2010**, *75*, 1133–1144.
- (64) Medema, M. H.; Kottmann, R.; Yilmaz, P.; Cummings, M.; Biggins, J. B.; Blin, K.; de Bruijn, I.; Chooi, Y. H.; Claesen, J.; Coates, R. C.; Cruz-Morales, P.; Duddela, S.; Düsterhus, S.; Edwards, D. J.; Fewer, D. P.; Garg, N.; Geiger, C.; Gomez-Escribano, J. P.; Greule, A.; Hadjithomas, M.; Haines, A. S.; Helfrich, E. J. N.; Hillwig, M. L.; Ishida, K.; Jones, A. C.; Jones, C. S.; Jungmann, K.; Kegler, C.; Kim, H. U.; Kötter, P.; Krug, D.; Masschelein, J.; Melnik, A. V.; Mantovani, S. M.; Monroe, E. a; Moore, M.; Moss, N.; Nützmann, H.-W.; Pan, G.; Pati, A.; Petras, D.; Reen, F. J.; Rosconi, F.; Rui, Z.; Tian, Z.; Tobias, N. J.; Tsunematsu, Y.; Wiemann, P.; Wyckoff, E.; Yan, X.; Yim, G.; Yu, F.; Xie, Y.; Aigle, B.; Apel, A. K.; Balibar, C. J.; Balskus, E. P.; Barona-Gómez, F.; Bechthold, A.; Bode, H. B.; Borriss, R.; Brady, S. F.; Brakhage, A. a; Caffrey, P.; Cheng, Y.-Q.; Clardy, J.; Cox, R. J.; De Mot, R.; Donadio, S.; Donia, M. S.; van der Donk, W. a; Dorrestein, P. C.; Doyle, S.; Driessen, A. J. M.; Ehling-Schulz, M.; Entian, K.-D.; Fischbach, M. a; Gerwick, L.; Gerwick, W. H.; Gross, H.; Gust, B.; Hertweck, C.; Höfte, M.; Jensen, S. E.; Ju, J.; Katz, L.; Kaysser, L.; Klassen, J. L.; Keller, N. P.; Kormanec, J.; Kuipers, O. P.; Kuzuyama, T.; Kypides, N. C.; Kwon, H.-J.; Lautru, S.; Lavigne, R.; Lee, C. Y.; Linquan, B.; Liu, X.; Liu, W.; Luzhetskyy, A.; Mahmud, T.; Mast, Y.; Méndez, C.; Metsä-Ketelä, M.; Micklefield, J.; Mitchell, D. a; Moore, B. S.; Moreira, L. M.; Müller, R.; Neilan, B. a; Nett, M.; Nielsen, J.; O’Gara, F.; Oikawa, H.; Osbourn, A.; Osburne, M. S.; Ostash, B.; Payne, S. M.; Pernodet, J.-L.; Petricek, M.; Piel, J.; Ploux, O.; Raaijmakers, J. M.; Salas, J. a; Schmitt, E. K.; Scott, B.; Seipke, R. F.; Shen, B.; Sherman, D. H.; Sivonen, K.; Smanski, M. J.; Sosio, M.; Stegmann, E.; Süssmuth, R. D.; Tahlan, K.; Thomas, C. M.; Tang, Y.; Truman, A. W.; Viaud, M.; Walton, J. D.; Walsh, C. T.; Weber, T.; van Wezel, G. P.; Wilkinson, B.; Willey, J. M.; Wohlleben, W.; Wright, G. D.; Ziemert, N.; Zhang, C.; Zotchev, S. B.; Breitling, R.; Takano, E.; Glöckner, F. O. *Nat. Chem. Biol.* **2015**, *11*, 625–631.

-
- (65) Cobb, R. E.; Wang, Y.; Zhao, H. *ACS Synth. Biol.* **2015**, *4*, 723–728.
- (66) Gantt, R. W.; Goff, R. D.; Williams, G. J.; Thorson, J. S. *Angew. Chemie - Int. Ed.* **2008**, *47*, 8889–8892.
- (67) Tsukioka, Y.; Yamashita, Y.; Nakano, Y.; Oho, T.; Koga, T. *J. Bacteriol.* **1997**, *179*, 4411–4414.
- (68) Martinez, V.; Ingwers, M.; Smith, J.; Glushka, J.; Yang, T.; Bar-Peled, M. *J. Biol. Chem.* **2012**, *287*, 879–892.
- (69) Barber, G. A.; Neufeld, E. F. *Biochem. Biophys. Res. Commun.* **1961**, *6*, 44–48.
- (70) Barber, G. A. *Biochemistry* **1962**, *1*, 463–468.
- (71) Yang, J. Y.; Sanchez, L. M.; Rath, C. M.; Liu, X.; Boudreau, P. D.; Bruns, N.; Glukhov, E.; Wodtke, A.; De Felicio, R.; Fenner, A.; Wong, W. R.; Linington, R. G.; Zhang, L.; Debonsi, H. M.; Gerwick, W. H.; Dorrestein, P. C. *J. Nat. Prod.* **2013**, *76*, 1686–1699.
- (72) Nguyen, D. D.; Wu, C.-H.; Moree, W. J.; Lamsa, A.; Medema, M. H.; Zhao, X.; Gavilan, R. G.; Aparicio, M.; Atencio, L.; Jackson, C.; Ballesteros, J.; Sanchez, J.; Watrous, J. D.; Phelan, V. V.; van de Wiel, C.; Kersten, R. D.; Mehnaz, S.; De Mot, R.; Shank, E. A.; Charusanti, P.; Nagarajan, H.; Duggan, B. M.; Moore, B. S.; Bandeira, N.; Palsson, B. Ø.; Pogliano, K.; Gutiérrez, M.; Dorrestein, P. C. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, E2611–2620.
- (73) Phelan, V. V.; Liu, W.-T.; Pogliano, K.; Dorrestein, P. C. *Nat. Chem. Biol.* **2012**, *8*, 26–35.
- (74) Liu, W.-T.; Lamsa, A.; Wong, W. R.; Boudreau, P. D.; Kersten, R.; Peng, Y.; Moree, W. J.; Duggan, B. M.; Moore, B. S.; Gerwick, W. H.; Linington, R. G.; Pogliano, K.; Dorrestein, P. C. *J. Antibiot. (Tokyo)*. **2014**, *67*, 99–104.
- (75) Wu, C.; Zacchetti, B.; Ram, A. F. J.; van Wezel, G. P.; Claessen, D.; Choi, Y. H. *Sci. Rep.* **2015**, *5*, 10868.
- (76) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Nat. Protoc.* **2010**, *5*, 536–549.
- (77) Watrous, J.; Roach, P.; Alexandrov, T.; Heath, B. S.; Yang, J. Y.; Kersten, R. D.; van der Voort, M.; Pogliano, K.; Gross, H.; Raaijmakers, J. M.; Moore, B. S.; Laskin, J.; Bandeira, N.; Dorrestein, P. C. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1743–1752.
- (78) Frank, A. M.; Bandeira, N.; Shen, Z.; Tanner, S.; Briggs, S. P.; Smith, R. D.; Pevzner, P. A. *J. proteome Res.* **2008**, 113–122.
- (79) McWilliam, H.; Li, W.; Uludag, M.; Squizzato, S.; Park, Y. M.; Buso, N.; Cowley, A. P.; Lopez, R. *Nucleic Acids Res.* **2013**, *41*, 597–600.
- (80) Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739.
- (81) Engler, C.; Kandzia, R.; Marillonnet, S. *PLoS One* **2008**, *3*, e3647.
- (82) Świątek, M. A.; Tenconi, E.; Rigali, S.; van Wezel, G. P. *J. Bacteriol.* **2012**, *194*, 1136–1144.

Chapter 11

Utilization of cellulose as a solid sorbent for antibiotic extraction

Changsheng Wu ^{1,2}, Robert Verpoorte ¹, Gilles P. van Wezel ², and Young Hae Choi ¹

¹ *Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands;*

² *Molecular Biotechnology, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, The Netherlands;*

Manuscript in preparation

ABSTRACT

Rational phytochemical analysis plays an important role in drug discovery approaches. Here, we exploited the natural polymer cellulose as a solid adsorbent of antibiotics in an aqueous matrix. Cellulose was shown to selectively bind several antibiotics with different chemical scaffolds, with preference towards those rich in hydrogen bonding such as aminoglycosides and glycopeptides. The binding selectivity of cellulose was also seen in the practical extraction of antibiotics specialized by actinomycetes. Cellulose preferentially absorbed polar compounds that were missed in conventional organic solvent fractionation such as using ethyl acetate extraction. This allowed the characterization of a group of structurally novel catechol-type siderophores (**12—14**) in *Streptomyces* sp. MBT76, which were exclusively extracted by cellulose. In summary, cellulose extraction can be considered as a complementary approach to conventional liquid-liquid extraction (LLE) to cover the widest range of secondary metabolites produced by actinomycetes, especially for the easily-overlooked water-soluble antibiotics.

1. INTRODUCTION

During the isolation of natural products from cultures, a rational extraction procedure is very important in terms of the removal of interfering components, such as from the culture medium, and to concentrate the widest range of specialized secondary metabolites with diverse polarities. Conventionally, the extraction of secondary metabolites from microbial cultures (predominantly in aqueous broth) is conducted with the organic solvents immiscible with water, such as ethyl acetate, chloroform. However, this conventional extraction method does not make use of all the metabolites contained in actinomycetes, since microbial NPs cover a tremendous chemical diversity with diverse chemical polarity. Particularly, many water-soluble polar antibiotics, such as aminoglycosides, are not amenable to extraction with ethyl acetate, which instead results in the extraction of a considerable amount of apolar lipids. On the other hand, extraction specificity in some cases is needed to simplify the separation procedure of compounds with specific structural scaffolds or chemical functionalities,¹ which is challenging to achieve in liquid-liquid extraction (LLE) that adversely consume large amounts of environment-unfriendly organics.

Various types of solid sorbents (e.g. C18 silica, polymeric resins, etc) have been employed to extract microbial NPs, among them, Diaion HP-20 resin is the most used, but it shows relatively low recovery for polar compounds. It is cost-ineffective to manufacture these artificial materials, and what is worse that they impose potential adverse effect on the environment. Alternatively, there has been an explosion of interest in the use of green biomass as a renewable resource for extraction purpose. Because of the omnipresence of interacting surface hydroxyl groups, the hydrophilic polymer cellulose is saturated with water when immersed in water or conditioned in moist atmosphere, which is the basis of the paper thin-layer chromatography. The retention of small molecules by cellulose is also observed in daily life, such as the irreversible adsorption of pigments by white filter paper after use and vegetables obtaining the flavors of meat when cooked together. Furthermore, Wagman and co-workers experimentally confirmed filtration materials and feces adsorb aminoglycoside-type antibiotics.^{2,3} All of these facts reveal that cellulose might have great potential as a solid adsorbent for extraction of some suitable antibiotics. However, to the best of our knowledge, there have been no reports on this topic.

Herein, the applicability of cellulose as trapping adsorbent for antibiotics was investigated. Eight commercial antibiotics with diverse chemical skeletons were first compared for their interaction with cellulose. Then, cellulose was employed for the extraction of polar metabolites produced by three actinomycetes, in comparison with the more traditional ethyl acetate extraction.

2. RESULTS AND DISCUSSION

2.1. Cellulose selectively binds antibiotics with different chemical scaffolds

An examination of the degree of physical association of eight commercial antibiotics with cellulose was investigated, namely the aminoglycosides hygromycin B, spectinomycin and streptomycin, the (glyco-)peptides vancomycin, capreomycin, and actinomycin D and the polyketides oxytetracyclin and chloramphenicol. All of these antibiotics except actinomycin D bound well to cellulose to give nonspecific cellulose-antibiotic complexes. The binding

efficiency of each of these antibiotics was manifested by the quantity of the cellulose-trapped molecules, in a setup where 10 mg of compound was added to 1 g of absorbent. Cellulose efficiently (>10%) absorbed aminoglycosides (Figure 1), with hygromycin B (22.6%) binding particularly well, in agreement with previous reports.^{2,3} In addition to the aminoglycosides, the peptide antibiotic vancomycin (14.8%) bound well too, with an efficiency comparable to those of the aminoglycosides streptomycin (17.5%) and spectinomycin (13.7%). Of the capreomycin, 8.1% was retained by the cellulose. Oxytetracyclin (2.5%) and chloramphenicol (0.5%) are deficient in H-bonding and were not captured well by cellulose. In this sense, cellulose could be reconsidered as a natural material for solid phase extraction and/or isolation of antibiotics.

The observed selectivity in Figure 1 prompted us to hypothesize a plausible binding mechanism of cellulose. Hydrogen bonding should be the principal factor, which explains aminoglycosides rich in –OH, and –NH₂ much more captured by residual –OH on cellulose surface than polyketide oxytetracyclin. Besides, size effects might also be involved. Cellulose is derived from D-glucose units condensed through $\beta(1\rightarrow4)$ -glycosidic bonds, to form a straight-chain polymer. The multiple hydroxyl groups on the glucose residues from one chain form hydrogen bonds with oxygen atoms on the same or on a neighbor chain, holding the chains firmly together side-by-side. The complicated three-dimensional hydrogen-bonding network presumably brings about an additional “microcage” among the glucose units. The size of this type of pores fits a six-membered glucose ring. Therefore, if a suitable ring or residue is rich in hydrogen-bonding groups, it would fall into the “microcage” and thus be trapped by –OH on the cellulose. This theory would explain why vancomycin with two glycosyl moieties was extracted much more efficiently by cellulose than actinomycin D, though they are both cyclic peptides. In light of these findings, it is recommended that caution prevail in filtration of aminoglycosides that are tenaciously bound to cellulose, especially when low concentrations of these antibiotics are used, large losses may occur.³

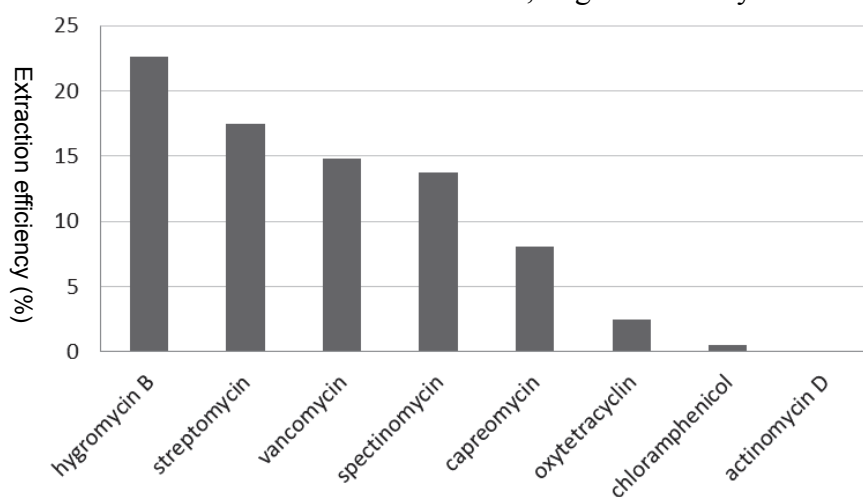


Figure 1. Cellulose selectively binds antibiotics with different chemical scaffolds. The extraction efficiency (in %) of the eight tested antibiotics is displayed in descending order.

2.2. Cellulose extracts antibiotics from the culture broth of actinomycetes

We then wondered if cellulose could be used to expedite antibiotic discovery from actinomycetes, especially those easily-neglected polar compounds? *Streptomyces* sp. MBT76 was previously identified as a prolific producer of antibiotics,⁴ and recent NMR-based

metabolomics revealed it could produce a diversity of secondary metabolites, such as isocoumarins, prodiginines, acetyltryptamine, and fervenulin, *etc.*⁵ In view of the intensive chemical investigation, this strain thus formed an ideal model for cellulose extraction of metabolites produced by a natural isolate. After 5 days of fermentation in perturbed NMMP medium, the culture broth of MBT76 was subjected to cellulose extraction and ethyl acetate (EtOAc) liquid-liquid partition, respectively. The extracts were compared by NMR profiling⁶ to compare these two extraction methods, whereby marked differences were observed (Figure 2). The compounds **1**–**14** (Figure 3) were identified with ¹H NMR technique and UHPLC-ToF-MS analysis (positive mode). EtOAc resulted in the extraction of a wide range of compounds (**1**–**9**) with diverse scaffolds and polarities, the identification of which were effortless due to the prior phytochemical knowledge of MBT76⁵ and the common components, via *di*-*tert*-butylphenols **8** and **9** in routine NMMP culture medium.^{7,8} In contrast, cellulose preferentially extracted polar chemicals from the culture broth of MBT76, *e.g.* a considerable amount of water-soluble mannitol (**10**), and glycerol (**11**), which avoided the apolar compounds **8** and **9** that oftentimes interfere with the purification of the desired natural products. Furthermore, the ¹H NMR spectrum of the cellulose extracts presented characteristic aromatic signals for catechol-type siderophores, including a suite of resonances at δ 7.34 (brd, J = 7.8 Hz), 7.03 (brd, J = 7.8 Hz), and 6.77 (t, J = 7.8 Hz) for 2,3-dihydroxybenzoic acid residue, a multitude of protons in the region δ 3.10–4.70 for *O*- or *NH*-substituted methylene or methine, and highly overlapped signals in the saturated region δ 1.70–2.10. UHPLC-ToF-MS analysis of cellulose crude extract presented a molecular feature at m/z 777.3157 for a formula of C₃₂H₄₅N₁₀O₁₃ (calculated for 777.3162), which confirmed the existence of divanchrobactin (**12**). Another mass peak at m/z 770.3408 established the molecular formula of C₃₁H₄₇N₉O₁₄ (calculated for 770.3315) for **14**, with one more unsaturation compared with that of compound **13**. Further MS/MS fragmentation indicated the unsaturation located in the hOrn-5 moiety, and was tentatively assigned as a double bond as depicted (Figure 4).

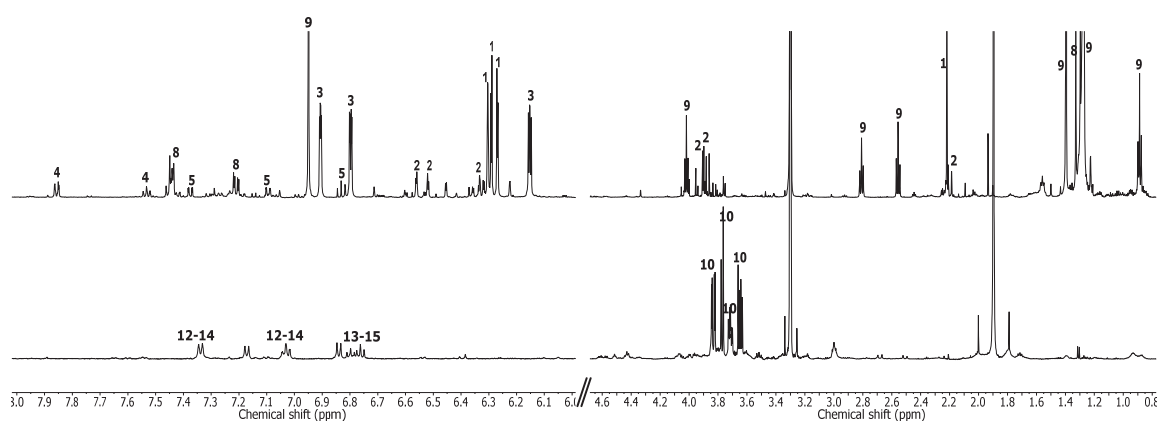
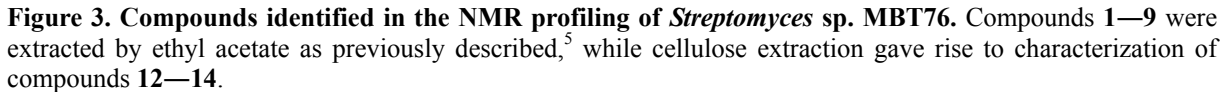


Figure 2. NMR comparison of extracts obtained from the culture broth *Streptomyces* sp. MBT76 using EtOAc (top) or cellulose (bottom). Proton signals were assigned to compounds numbered in Figure 3. EtOAc extracted a wide range of chemicals (**1**–**9**) with different scaffolds and diverse polarities, while cellulose showed specificity towards siderophores (**12**–**14**). Though the extraction of lipids **8** and **9** in EtOAc liquid partition could be avoided, a considerable amount of supplemented mannitol (**10**) in the NMMP culture medium was bound by cellulose.



164

amount of mannitol and glycerol, while EtOAc extracted a broad spectrum of compounds, one of which was identified as 9-methylstreptimidone (Figure S1). The extraction selectivity of cellulose was also seen for *Streptomyces* sp. MBT73,⁴ whereby more yellow pigment resistomycin than the apolar red pigment undecylprodigine was bound, as visually detected by thin-layer chromatography (TLC) analysis (Figure S2).

The large genomes of actinomycetes (>8 Mb) contain a plethora of gene clusters for natural products with diverse polarities.⁹ The classical liquid-liquid extraction (LLE) employing ethyl acetate tends to miss water-soluble antibiotics, and lack the extraction selectivity towards specialized metabolites, especially for valuable chemicals produced in trace titers. Besides the comparative advantage in green chemistry, cellulose effectively avoids the metabolic complexity and unwanted lipids in LLE (Figure 2), but instead shows preference towards specific polar metabolites. In view of the notorious difficulty to purify water-soluble compounds,¹⁰ ¹H NMR profiling tool affording structural information in an early stage, can rationalize the downstream isolation workflow of compounds extracted by cellulose. For instance, NMR-revealed catechol-type siderophores (**12–14**) in the cellulose extraction of *Streptomyces* sp. MBT76 could be potentially separated from the abundant sugars (mannitol and glycerol) by using Fe₃O₄ magnetic nanoparticles that selectively bond salicylic acid moiety¹.

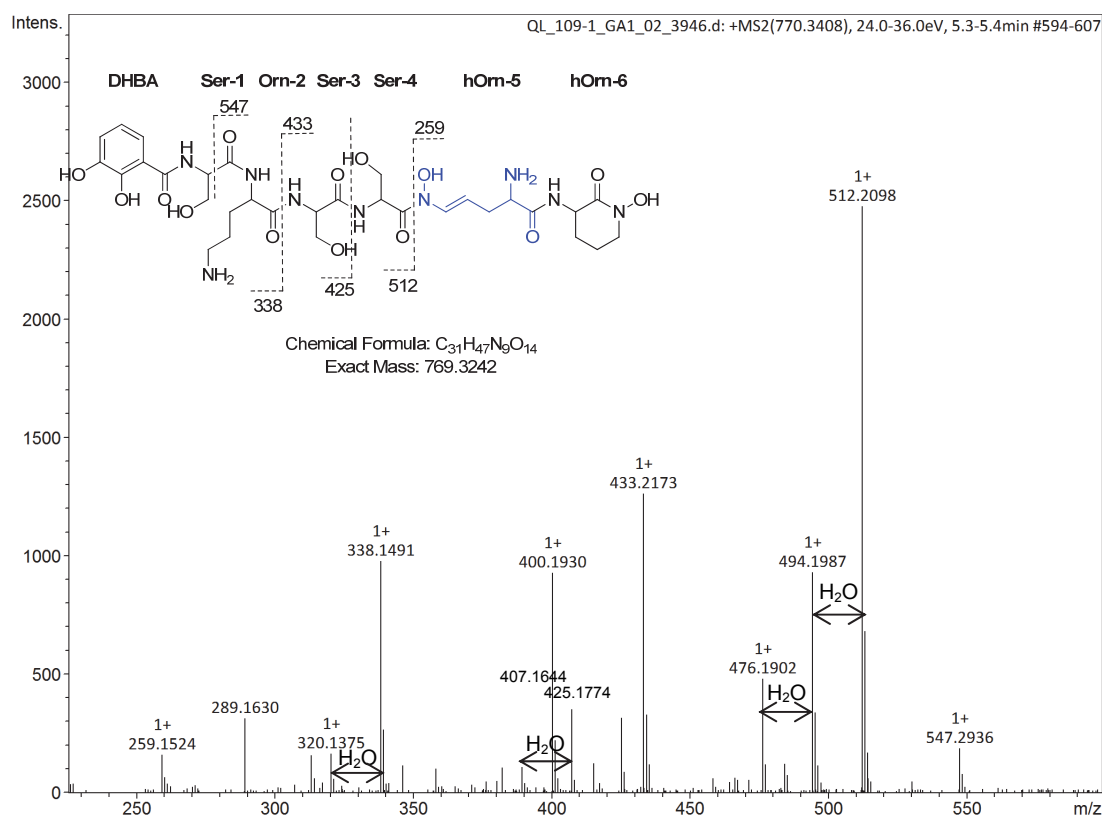


Figure 4. ESI-tandem mass spectrum (MS²) of the siderophore **14** in the cellulose extraction of *Streptomyces* sp. MBT76.

3. CONCLUSION

Our work shows that cellulose solid-phase extraction (CL-SPE) is complementary to conventional LLE, thus giving a broader coverage of metabolites to be retrieved from cultures

of actinomycetes, and in particular aiding in isolating water-soluble metabolites that readily escape EtOAc extraction. Therefore, the binding selectivity of cellulose in combination with NMR-based profiling will expedite the discovery of some novel chemical scaffolds and low-abundant metabolites from actinomycetes because it does not require prior knowledge of biosynthetic potential. It is important to note that cellulose is the most abundant natural polymer on the planet, which has far reaching implications for the ecological role of this polysaccharide. For example, we envision that natural products will accumulate around cellulose as they do in our extraction procedure and this may for example explain why water-soluble antibiotics are retained in the soil and not easily washed away. Thus, the antibiotic-binding nature of cellulose could provide novel ecological insights with implications for among others soil microbiology. Studies in this direction are currently underway in our laboratory.

4. EXPERIMENTAL SECTION

4.1. General Experimental Procedures

Cellulose was used as microcrystalline powder (Sigma, St. Louis, MO, USA). Silica gel 60 F254 (Merck, Darmstadt, Germany) was used for analytical TLC. Antibiotics were: spectinomycin dihydrochloride pentahydrate (Duchefa, Haarlem, Netherland), streptomycin sulfate (Boehringer Mannheim GmbH, Germany), vancomycin hydrochloride (Duchefa, Haarlem, Netherland), oxytetracyclin hydrochloride (Duchefa, Haarlem, Netherland), capreomycin sulfate (Sigma), hygromycin B (Duchefa), chloramphenicol (Sigma), actinomycin D (Sigma). All organic solvents were of analytical or HPLC grade, depending on the experiment.

4.2. Cellulose extraction of commercially obtained antibiotics

10 mg of hygromycin B, spectinomycin, streptomycin, vancomycin, capreomycin, oxytetracyclin, and chloramphenicol were dissolved in 25 ml of distilled water, except actinomycin D which was dissolved in ethyl acetate. After adding 1 g of cellulose powder, the mixture was stirred under room temperature for 2 h. The suspension was centrifuged and the supernatant was discarded. The cellulose in each sample was then rinsed three times with 20 ml of fresh solvents. Desorption of antibiotics from cellulose was conducted with three cycles of 15 ml acidified water (1% acetic acid), by stirring for 1 h. The combined acid extract was lyophilized using a freeze-dryer (Edwards Ltd., Crawley, UK). The obtained residue was dissolved in a mixture of $\text{CH}_3\text{OH}-d_4$ (0.4 ml) and KH_2PO_4 buffer in D_2O (pH 6.0, 0.4 ml) containing 0.01% (w/v) 3-(trimethylsilyl)propionic-2,2,3,3- d_4 acid sodium (TMSP) as internal standard for ^1H NMR analysis. The extraction efficiency of cellulose was a ratio that the amount of acid-recovered antibiotics was divided by initially added amount (10 mg). Quantitative NMR was used to calculate the quantity of antibiotics recovered from cellulose, based on the relative ratio of intensity of the well-separated NMR peak to the known amount of TMSP as reference.¹¹

4.3. Cellulose extraction of metabolites from culture broth

Streptomyces sp. MBT76, MBT73, and MBT18 from the actinomycete collection⁴ of

Molecular Biotechnology, IBL, Leiden University was maintained in 20% glycerol at -20 °C. The liquid culture medium employed for fermentation was Minimal Liquid Medium (NMMP),¹² containing (NH₄)₂SO₄ (2 g), Casaminoacid (5.0 g), MgSO₄·7H₂O (0.6 g), minor element solution (1 mL), Na⁺-K⁺ phosphate buffer pH 6.8 (150 mL), 10% mannitol solution (50 mL), 20% glycerol solution (50 mL), peptone (8 g),^{4,5} and distilled water up to one liter. Additional 0.8% peptone was supplemented for MBT76, and 0.5% yeast extract for MBT73 according to previous study.⁴ Spores suspension was inoculated into 50 mL culture medium, and the culture broth was continuously shaken at 220 g and 30 °C for 5 days.

Cultures were harvested by centrifugation at 4000xg for 10 min. The culture broth (50 ml) was divided into two equal portions, one of which was extracted three times with 10 ml of ethyl acetate (EtOAc). The organic phases were pooled and washed with 30 ml of water and subsequently dried with 5 g of anhydrous Na₂SO₄. Finally, the EtOAc was removed under vacuum at 40 °C and the residue dissolved in 2.0 ml of EtOAc in a microtube (Eppendorf type-5415C, Hamburg, Germany). The solvent was then evaporated at room temperature under nitrogen gas, and subsequently dipped into liquid nitrogen and lyophilized using a freeze dryer (Edwards Ltd., Crawley, England). Cellulose (1 g) extraction of another one portion of microbial culture broth (25 ml) followed the same procedure for commercial antibiotics as described above. The obtained residue was dissolved in 0.3 ml of CH₃OH-*d*₄ for ¹H NMR analysis. All experiments were conducted in triplicate.

4.4. NMR measurement

NMR sample preparation and measurements were done according to our previously published protocol.⁶ Briefly, 500 µl of methanol-*d*₄ was added to the freeze-dried sample. Subsequently, the mixture was vortexed for 10 s and sonicated for 20 min at 42 kHz using an Ultrasonicator 5510E-MT (Branson, Danbury, CT, USA), followed by centrifugation at 13,000 g at room temperature for 5 min. The supernatant (300 µl) was transferred to a 5 mm micro NMR tube and analyzed. The ¹H NMR spectra were recorded at 25 °C on a 600 MHz Bruker DMX-600 spectrometer (Bruker, Karlsruhe, Germany) operating at a proton NMR frequency of 600.13 MHz. Deuterated methanol was used as the internal lock. Each ¹H NMR spectrum consisted of 128 scans using the following parameters: 0.16 Hz/point, pulse width (PW) = 30 (11.3 µs) and relaxation delay (RD) = 1.5 s. Free induction decays (FIDs) were Fourier transformed with a line broadening (LB) = 0.3 Hz. The resulting spectra were manually phased and baseline corrected, and calibrated to residual methanol-*d*₄ at 3.30 ppm, using MestReNova 8.1.

4.5. UHPLC-ToF-MS analysis

UHPLC-ToF-MS analyses were performed on an UHPLC system (Ultimate 3000, ThermoScientific, Germany) coupled to an ESI-IIQ-ToF spectrometer (microTOF-QII, Bruker Daltonics, Germany) in the positive mode. The chromatographic separation was done using a Kinetex C₁₈ UPLC 2.6 µm particle size column 150 × 2.0 mm (Phenomenex, Torrance, CA, USA) at a flow rate of 0.3 ml/min and a column temperature of 30 °C. Samples (3 µl) were eluted using a gradient of solvent A (water) and B (acetonitrile), both with 0.1% formic acid (v/v). The initial percentage of B was 5%, which was linearly increased to 90% in 19.5 min,

followed by a 2 min isocratic period and, then re-equilibrated with original conditions in 2 min. Nitrogen was used as drying and nebulizing gas. The gas flow was set at 10.0 l/min at 250 °C and the nebulizer pressure was 2.0 bar. The MS data were acquired over m/z range of 100–1000. The capillary voltage was 3.5 kV. For internal calibration, a 10 mM solution of sodium formate (Fluka, Steinheim, Germany) was infused. Formic acid, water and acetonitrile were LCMS grade, Optima (Fisher Scientific, Waltham, MA, USA).

REFERENCES

- (1) Li, R.; Shen, Y.; Zhang, X.; Chen, B.; Beek, T. A. *J. Nat. Prod.* **2014**, *77*, 571–575.
- (2) Wagman, G. H.; Bailey, J. V.; Weinstein, M. J. *Antimicrob. Agents Chemother.* **1974**, *6*, 415–417.
- (3) Wagman, G. H.; Bailey, J. V.; Weinstein, M. J. *Antimicrob. Agents Chemother.* **1975**, *7*, 316–319.
- (4) Zhu, H.; Swierstra, J.; Wu, C.; Girard, G.; Choi, Y. H.; van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. *Microbiology* **2014**, *160*, 1714–1725.
- (5) Wu, C.; Zhu, H.; van Wezel, G. P.; Choi, Y. H. *Metabolomics* **2016**, DOI 10.1007/s11306-016-1025-6
- (6) Kim, H. K.; Choi, Y. H.; Verpoorte, R. *Nat. Protoc.* **2010**, *5*, 536–549.
- (7) Wu, C.; Du, C.; Gubbens, J.; Choi, Y. H.; van Wezel, G. P. *J. Nat. Prod.* **2015**, *78*, 2355–2363.
- (8) Wu, C.; Zacchetti, B.; Ram, A. F. J.; van Wezel, G. P.; Claessen, D.; Choi, Y. H. *Sci. Rep.* **2015**, *5*, 10868.
- (9) Bérdy, J. *J. Antibiot. (Tokyo)*. **2005**, *58*, 1–26.
- (10) Shimizu, Y. *J. Nat. Prod.* **1985**, *48*, 223–235.
- (11) Kim, H. K.; Choi, Y. H.; Lujendijk, T. J. C.; Rocha, R. A. V.; Verpoorte, R. *Phytochem. Anal.* **2004**, *15*, 257–261.
- (12) Hodgson, D. J. *Gen. Microbiol.* **1982**, *3*, 2417–2430.

Chapter 12

Summary and Discussion

1. Summary and discussion of the experimental results

Since Fleming's ground-breaking discovery of penicillin,¹ antibiotic usage has increased enormously. This has led to a concomitant sharp rise in infectious diseases associated with drug-resistant pathogens since the 1980s, which rapidly spread across the globe.² Many multi-drug resistant (MDR) bacteria have developed resistance to most if not all antibiotics discovered to date, which imposes a massive challenge on society.³ Soon after new anti-infectives enter the market, including those representing our last line of defense (e.g. vancomycin), resistance rapidly emerges, which has led to the return of diseases once considered nearly eradicated, with tuberculosis as clear example. Nosocomial pathogens collectively known as the ESKAPE pathogens,⁴ namely *Pseudomonas aeruginosa*, *Enterobacteriaceae*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Staphylococcus aureus* and *Enterococcus spp.* are also a constant threat to human health care.⁵ The increase in resistance comes at a time where drug-discovery pipelines are drying out,⁶ primarily caused by the fact that finding new drugs is time consuming and hence expensive.^{7,8} New ideas and insights should be brought in to revitalize natural products research, and improve our chances of finding new antimicrobial drugs, particularly those with novel modes of action and/or unique chemical skeleton are of paramount importance to human health.

Should we lose faith that we can overcome these challenges? An old Chinese philosophy is "things in the world are allelopathy". If so, what can curb the raging tiny bugs? As always, the answer should come from Nature. Billions of years of evolution has resulted in an extraordinary and diverse library of chemical structures with huge functional variety. In other words, it is like a textbook encoded in the language of chemistry. Modern analytical tools now enable us to understand how structure relates to bio- or physiological function. With the advances made in synthetic biology and chemistry,⁹ we can not only make copies of naturally occurring molecules, but also design new bio-inspired compounds, some with the functions unprecedented in the natural world - compounds that will influence, if not shape, every facet of our life.¹⁰

To come back to antibiotics, where can we find new molecules to replenish the pipelines of the pharma industry? Although scientists have been trying to tap the huge repository of microorganism in nature, the biosynthetic potential of microbes is still largely underestimated, and probably just a tip of iceberg is discovered.^{11,12} Several causes can be discriminated. Firstly, while the number of filamentous microbes in soil and marine environments is enormous, the bulk of them resists cultivation in the laboratory and thus escape screening and exploitation.¹³ Recent work has shown that enabling growth of these 'uncultivable' microorganisms might open up a new area of the chemical space of natural products.^{14,15} Secondly, many natural product gene clusters of cultivable microbes remain silent under standard laboratory cultivation. As an example, genome sequencing of actinomycetes revealed that the producing capacity of even the best-studied model organisms has been grossly underestimated.^{16,17} And thirdly, those gene clusters that are expressed, in particular if expression is relatively low compared to other bioactive compounds produced by the same organism, often escape discovery due to the very demanding phytochemical analysis or chromatographic purification that is required.¹⁸

Actinomycetes are industrially and medicinally important microorganisms because of their ability to produce a plethora of secondary metabolites. Indeed, more than two-thirds of

known antibiotics are produced by these filamentous bacteria, the majority of which by *Streptomyces* species. The advent of next-generation sequencing (NGS) technologies has uncovered that even the best-studied model actinomycete *Streptomyces coelicolor* A3(2) still possesses many yet underexplored resources for NPs discovery.¹⁶ New approaches are therefore required to exploit the unexplored reservoir of bioactive natural products, and to rejuvenate the drug-discovery pipelines.¹⁹

In this thesis, approaches to aid the discovery of novel antibiotics from actinomycetes are presented, focusing on new approaches and ideas from the perspectives of both chemistry and biology. Many efforts aim to elevate the chances of finding new lead compounds with novel modes of action, whereby the emphasis lies on unlocking cryptic biosynthetic pathways in microbes, followed by efficiently prioritizing the elicited molecules. Various proofs of concept are provided that NMR-based metabolomics integrated with different antibiotics-eliciting strategies is a very efficient method to facilitate the discovery of new antibiotics from actinomycetes.

In recent years, the implementation of the microbial co-cultivation strategy attracted considerable interest and application, because it proved to be an effective way to harvest unique structures with pronounced biological activities.²⁰ In Chapter 4, it is shown that co-cultivation of *Streptomyces coelicolor* and the filamentous fungus *Aspergillus niger* has a major impact on their respective secondary metabolism. NMR-based metabolomics revealed several compounds that correlated specifically to co-cultures, including the cyclic dipeptide cyclo(Phe-Phe), and 2-hydroxyphenylacetic acid. These compounds were produced by *A. niger* in response to the chemicals produced by *S. coelicolor*. Inspired by this co-culture metabolomics study, biotransformation of hydroxycinnamic acids resulted in the specific production of the novel molecule (2*E*,4*E*)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid by the co-cultures microorganisms.

Selection for streptomycin-resistant mutants of actinomycetes is another pleiotropic approach to enhance the production of antibiotics.²¹ In Chapter 5, a strategy based on NMR-based metabolomics combined with the introduction of streptomycin resistance in *Streptomyces* sp. MBT28, led to the identification of the new compound 7-prenylisatin that showed antimicrobial activity against *Bacillus subtilis*. Furthermore, the metabolite-guided genome mining combined with Natural Product Proteomining platform²² identified an orphan gene cluster (*isa*) with an indole prenyltransferase that transforms tryptophan into 7-prenylindole. However, the biosynthesis genes for the final product 7-prenylisatin are not confined to a single cluster in the genome of *Streptomyces* sp. MBT28, as indispensable oxygenases responsible for the oxidation of 7-prenylindole into the 7-prenylisatin were not found within or in close proximity of the *isa* gene cluster. This study underlines that combination of various ‘omics’ approaches can facilitate the discovery of novel molecules, whilst offering a straightforward way to characterize the gene cluster responsible for the identified compounds.

Secondary metabolites are often produced in a growth phase-dependent manner by actinomycetes, and finding the optimal harvesting time is crucial for successful isolation of the desired bioactive metabolites. In Chapter 6, the growth phase-dependence of secondary metabolite production by *Streptomyces* sp. MBT76 was analyzed using NMR-based metabolomics. A group of methoxylated isocoumarins were characterized, including a

tetra-methoxylated compound 5,6,7,8-tetramethoxyl-3-methyl-isocoumarin. Multivariate data analysis indicated that methoxylation led to an increase in the antimicrobial activity of isocoumarin-type antibiotics, which was confirmed by antimicrobial assays using the methoxylated and non-methoxylated isocoumarins. Furthermore, metabolomics-guided biotransformation of isoflavone genistein resulted in a suite of methoxylated isoflavones, undergoing exactly the same modification pattern of isocoumarin. Interestingly, while isocoumarin-type metabolites are frequently isolated from actinomycetes,^{23,24} the underlying genetic basis for this type of polyketide is still poorly understood. In consideration of the structural similarity between 6,8-dihydroxy-3-methyl-isocoumarin and genistein, it is conceivable that the same *O*-methyltransferase is involved in the methylation of both substrates, and this putative *O*-methyltransferase seems to have strong methylation capacity. Therefore, it will make sense to characterize the genetic basis for isocoumarin biosynthesis in *Streptomyces* sp. MBT76. This should shed light on the nature of the methyltransferase responsible for the post-PKS modification, since the genes for PKS assembly and post-PKS decoration are always clustered in the genome of actinomycetes. Furthermore, this effort will also benefit further genetic engineering of this type of important antibiotics, such as combinatorial biosynthesis.

In Chapter 7, an approach was followed whereby the pathway-specific activator of a biosynthetic gene cluster (BGC) was over-expressed, leading to the activation of a poorly expressed type II PKS gene cluster (*qin*) in *Streptomyces* sp. MBT76. NMR-based metabolic profiling then efficiently linked the chemotype to the genotype of interest in the recombinant strain, whereby we shed first light on the intriguing architecture of qinimycins A-C, a group of novel pyranonaphthoquinones with structural features of 8-glycosylation, 5,14-epoxidation, and 13-hydroxylation. The methodology exemplified in Chapter 7 is particularly useful when information pertaining to the structural novelty of the products of particular BGCs can be gleaned from predictive bioinformatics analysis. NMR profiling of the molecules produced in highly complex matrices thereby allows rapid linkage of chemotype to genotype, greatly streamlining the downstream isolation and structure elucidation. However, there are also intrinsic limitations in this pathway-specific strategy, such as lack of high throughput, in some cases difficulty in recognition of pathway-specific activators, and the fact that many wild isolates are not genetically tractable. Selected novel compounds discovered in the course of the NMR-based metabolomics studies are displayed in Figure 1.

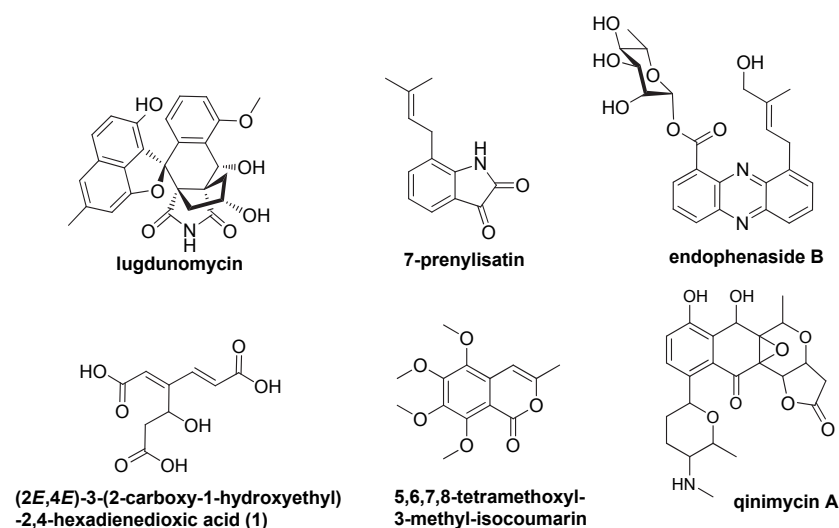


Figure 1. Representative new compounds described in this thesis.

Especially, lugdunomycin is a angucycline-derived with unprecedented chemical scaffold.

Metabolomics-driven drug discovery offers advantages over the routine bioassay-guided screening. The major shortcoming in the traditional method is lack of prior structural information on the bioactivity of interest prior to the labor-intensive and time-consuming compound isolation. It should be noted that deficiencies exist in NMR-based metabolomics for handling microbial extracts. Routinely in this thesis, actinomycetes were fermented in the liquid or solid minimal media. Many components from growth media are inevitably extracted by ethyl acetate, and in most cases are more abundant relative to the true secondary metabolites produced by actinomycetes. As NMR spectroscopy is a quantitative but relatively insensitive technique as compared to Mass Spectrometry (MS), the proton signals belonging to the compound of interest are often in low abundance and therefore obscured by the signals assignable to culture media. Therefore, NMR-based metabolomics works better when combined with an efficient antibiotics-eliciting approach. Alternatively, one can remove the unwanted interfering compounds in the initial step and accordingly highlight the signals of interest. For this purpose, exploitation of a comprehensive extraction method other than single organic solvent (ethyl acetate) liquid-liquid partition awaits further investigation, whereby solid-phase extraction seems promising (see below), because a wide range of solid sorbents with various physiochemical properties are available. Moreover, the disadvantage of the insensitivity and congestion problems in NMR-based metabolic profiling could potentially be overcome by HPLC-NMR hyphenation techniques.²⁵ The major advantages of on-flow HPLC-NMR lie in the real-time access to structure information of chromatographic peaks by using NMR spectrometers as detector.²⁶ To leverage the inherent lower sensitivity of NMR spectra and the limited injection volume onto the HPLC unit, solid-phase-extraction (SPE) is coupled to trap the analytes onto an identical SPE cartridge when repeated HPLC analysis, followed by final elution with a deuterated solvent into miniaturization tube for NMR analysis.²⁷ By using this LC-SPE-NMR approach,²⁸ comprehensive structural information about chemical composition of crude extracts can be obtained online, thus enabling early and prospective assessment of the valuable components of microbial extract or fraction.²⁹ Furthermore, when integrating MS into the system, like LC-DAD-SPE-MS-NMR instrumentation,³⁰ the architecture of novel compounds can be elucidated *de novo* and online without laborious purification of the analyte(s) from the highly complex matrix.

There are perhaps no strictly silent biosynthetic gene clusters. The compounds will undoubtedly be produced under some conditions in the original habitat where the bacteria live, as else the BGCs would have been lost during the course of evolution, and most likely would never have evolved in the first place. However, many microbial secondary metabolites are produced in trace amounts, which are out of the detection limit of NMR technique, and thus missed in the NMR-based metabolomics strategy. Logically, minor compounds are more likely to have been missed by routine screening efforts in the past, and hence have a better chance of being novel over those expressed abundantly under normal laboratory culturing conditions. A logical approach to unveil these minor compounds is up-scale fermentation to accumulate enough material for follow-up repeated chromatography isolation and structure elucidation. Though time- and labor- extensive, it has been fruitful to discover novel chemical skeletons by isolation of all the detectable compounds in any biological sample, in other words a systematic isolation method, which is prevalent in the field of traditional natural products chemistry. Actually, it is common that a single microbial source produces multiple

secondary metabolites,³¹ and systematic isolation is unambiguously the most efficient way to bring light the chemical diversity of a given actinomycete. Besides the discovery of new compounds, systematic isolation often aids our understanding of the biosynthetic pathway of a given natural product with complex architecture, because the structures of the co-isolated intermediates or side products could provide clues to elucidate the bioassembly of final target product. Furthermore, the fact that changes in the cultivation conditions can completely shift the metabolic profile of many microorganisms, resulted in the OSMAC (one strain many compounds) strategy for drug discovery from microbial sources.³² A microorganism never produces its entire arsenal of compounds at the same time under one specific environmental condition, which would be energetically and metabolically too costly. Rather, each compound is produced when needed, ensuring a competitive advantage when environmental conditions change. Consequently, varying culturing conditions is a logical approach to attempt to induce or optimize the production of secondary metabolites. In this context, systematic alteration of cultivation parameters (media composition, aeration, pH, temperature, addition of chemicals, etc.) of a given microorganism greatly increases the likelihood of finding new metabolites from a single strain. Of particular note is the static cultivation for diversification of secondary metabolites in the laboratory, because this is more consistent with their natural survival environment. It was recognized previously that the coproduction by actinomycetes in nature of two or more secondary metabolites is necessary to synergistically or contingently beat biological competitors, and that synergy and contingency may be common driving forces for the evolution of multiple secondary metabolite production by these sessile saprophytes.³¹ The OSMAC strategy was applied to discover the natural products produced by *Streptomyces* sp. QL37 (Chapter 8) and *Kitasatospora* sp. MBT66 (Chapters 9-10). Perturbed liquid NMMP media and solid MM, supplemented with different carbon sources (glycerol, mannitol, glucose, rhamnose, *N*-acetylglucosamine), additional additives (yeast extract, potato extract, peptone, starch, soy flower), or high alkalinity (pH 10), were compared, and solid fermentation was found better for secondary metabolites production in these two actinomycetes.

Systematic isolation of *Streptomyces* sp. QL37 cultured in MM supplemented with mannitol and glycerol as the carbon sources, led to the production of the lugdunomycins, which are natural products with a completely new chemical scaffold (Figure 1). Their chemical structure indicates that the lugdunomycins are derived from angucyclines, but the basic structure is then modified to generate a highly complex molecule with an unprecedented three dimensional structure. Lugdunomycins possess a benzaza[4,3,3]propellane skeleton adorned with a spirocyclic 2*H*-naphtho[1,8-*bc*]furan moiety and two all-carbon quaternary centers embedded within five contiguous stereogenic carbons. The striking backbone of benzaza[4,3,3]propellane-6-spiro-2'-2*H*-naphtho[1,8-*bc*]furan has a so far unprecedented chemistry. Further fermentation of *Streptomyces* sp. QL37 in R5 agar plates with peptone followed by systematic isolation, resulted in the characterization of many additional rearranged angucyclines (Chapter 8). As a result, in total nearly 30 angucycline derivatives were characterized, and roughly classified into four different chemical backbones, including the seven-ring lugdunomycins, three-ring anthraquinones, four-ring benz[*a*]nthaquinones, and the five-ring limamycins. The metabolite-guided genome mining identified a type II PKS (*lug*) gene cluster for angucycline production in *Streptomyces* sp. QL37, which was confirmed

by genetic disruption of minimal PKS genes. However, it was possible that a single *lug* is insufficient for the synthesis of all the discovered angucyclines, whereby outside genes (clusters) were needed to construct the complex architecture of lugdunomycin. More intriguingly, all the characterized rearranged and/or unrearranged angucycline derivatives in QL37 involve an oxidative C-C cleavage in the post-PKS stage to substantially change the basic backbone of benz[a]nthaquinone. It seems likely that more than one oxygenase is involved in the post-PKS modification in the lugdunomycin pathway, because multiple oxidations were observed in the isolated angucyclines. In particular, the oxygenase responsible for the quinone ring opening via Bayer-Villiger oxidation was of utmost importance for lugdunomycin biosynthesis, which has the potential to modify other polyketide antibiotics, like tetracycline.

In Chapter 9, systematic investigation of *Kitasatospora* sp. MBT66 enabled the discovery of five novel rhamnosylated phenazines (termed as endophenazines A—E),³³ as well as the rhamnosylated plecomacrolide leucanicidin. Metabolite-guided genome mining of MBT66 identified the endophenazine gene cluster that is nearly identical to those of the endophenazine biosynthetic gene cluster (*epa*) of *Kitasatospora* sp. HKI 714.³⁴ However, genes responsible for the glycosylation of the endophenazine aglycones was unresolved in Chapter 9, as no such genes were found within or in close proximity of the *epa* gene cluster. The genetic basis for this rhamnosylation was subsequently resolved in Chapter 10.³⁵ Interestingly, the *de novo* rhamnosylation of both types of structure cores in *Kitasatospora* sp. MBT66, via phenazine and plecomacrolide, originated from decoration by the same glycosyl transferase and methyltransferase, LeuAB. This inherent substrate flexibility of LeuAB evoked us to reinvestigate the structural diversity of endophenazines in MBT66, whereby MS/MS based molecular networking facilitated the characterization of nine new endophenazines by the OSMAC strategy, including an unprecedented tautomeric glyceride phenazine (Chapter 10). Molecular networking tool is an effective tool for chemical dereplication strategy and a way of prioritizing novel compounds,³⁶ and its fundamental principle is based on the fact that structurally related natural products are characterized by similar MS/MS fragmentation patterns. In view of the glycosylation has profound impact on pharmaceutical properties, the promiscuity of LeuAB (or other glycosyltransferases) by accepting totally different chemical skeletons as substrates can be exploited to glycorandomize a broad range of natural products. This offers an alternative way to develop new antibiotics based on post-biosynthetic modification strategies.

Genes for secondary metabolite production in actinomycetes are typically clustered in the genomes, which is the basis of bioinformatics-guided discovery of natural products. However, the crosstalk of the *epa*, *baf*, and *rml* gene clusters for endophenazines production in *Kitasatospora* sp. MBT66 (Chapter 10), the joint contribution of *isa* and putative oxygenase for 7-prenylisatin production in *Streptomyces* sp. MBT28 (Chapter 5), as well as the genetics for lugdunomycin production in *Streptomyces* sp. QL37 (Chapter 8), is intriguing and led us conclude that a non-clustered or subclustered genome arrangement for secondary metabolites biosynthesis in actinomycetes is more common than previously suggested. The combinational biosynthesis of a single compound employing several distant regions of the genome is also seen for the biosynthesis of sioxanthin³⁷ in the actinomycete genus *Salinispora*, of prodiginines in *Streptomyces coelicolor*³⁸, and pyrrolamides in *Streptomyces netropsis*.³⁹ In

other cases, single one gene cluster tends to deliver a multitude of structurally diverse products. For instance, the type II PKS *act* gene cluster in *S. coelicolor* synthesize different aryl scaffolds, including pyranonaphthoquinone,^{40,41} anthraquinones^{42,43}, naphthoquinones^{22,44}, and mutactin.⁴⁵ The metabolic profile of a recombinant *Streptomyces* strain expressing the *S. coelicolor whiE* minimal PKS led to the production of ~30 polyketides with varying chain lengths and cyclization patterns.⁴⁶ In this thesis, the *epa* gene cluster of *Kitasatospora* sp. MBT66 encodes the synthesis of more than 20 phenazines (Chapters 9 and 10), while the *lug* cluster of *Streptomyces* sp. QL37 results in the production of at least four kinds of chemical skeletons, with several tens of different metabolites (Chapter 8). Taken together, though synthetic biology approaches will be a major trend for future antibiotics discovery (Chapter 1),⁴⁷ it may be less effective for the handling of this type of intertwined metabolic pathways, and thus traditional systematic isolation should not be abandoned.

Whatever pipeline is developed for the discovery of new antibiotics, a common feature is the eventual phytochemical analysis of microbial cultures to characterize the molecules of interest. A rational extraction procedure is thus very important to remove the interfering components (such as components of the culture media), and to concentrate the widest possible range of specialized secondary metabolites with diverse polarities. In Chapters 4—10, the organic solvent-ethyl acetate was uniformly used to extract metabolites produced by all the investigated actinomycetes. As mentioned above, many media components are inevitably extracted by ethyl acetate, and in many cases these are more abundant relative to the true secondary metabolites produced by actinomycetes. In particular, a group of apolar chemicals from our routinely used culture media, including 2,5-di-tert-butylphenol, 3,4-di-tert-butylphenol, and (3,5-di-tert-butyl-4-hydroxyphenyl)-propionic acid octadecyl ester, were extracted by ethyl acetate in large amounts, which tended to be “common impurities or contaminants” in almost all our isolated compounds, and severely affected the final structure elucidation and bioactivity tests when the compounds of interest were in low yield. Another drawback of ethyl acetate is that it is less appropriate for the extraction of polar compounds, such as water-soluble aminoglycosides. In Chapter 11, we therefore explored the application of the natural polymer cellulose as a solid adsorbent of antibiotics in an aqueous matrix. This work showed that indeed cellulose-based solid-phase extraction can be considered as a complementary approach to conventional liquid-liquid extraction to uncover a wide range of secondary metabolites produced by actinomycetes, especially for the easily-overlooked water-soluble antibiotics.

2. Perspectives

With the rapid increase in bacterial resistance, it is of ultimate importance to develop new antibiotics. Historically, antibiotic discovery or other natural product research, has originally been dominated by chemistry, focusing on the isolation of pure compounds, structural characterization, and total or semi- synthesis. However, as connecting natural products to their biosynthetic genes becomes more important to exploit the full potential of genome and metagenome sequencing, the biology plays an increasingly important role in natural product research in the current (post-)genomics era. The recent explosion of publications dealing with genomics approaches to discover novel natural products has demonstrated that the

“genotype-to-chemotype” roadmap is a rewarding approach that has already revolutionized traditional natural product research.^{6,48,49} This field of research now merges multiple disciplines such as whole genome sequencing, bioinformatics, transcriptomics, proteomics, synthetic biology, systems biology, chemical ecology, and others.⁵⁰ Synthetic biology has a great potential to reinvigorate drug discovery through engineering biosynthetic pathways.⁹ As summarized by Jon Clardy and co-workers,⁵¹ with known compounds like artemisinin, synthetic biology supports natural products by improving supply; with ‘known unknowns’ like the products of cryptic biosynthetic gene clusters identified in the bacterial genomes, synthetic biology will likely support natural products by enabling their biosynthesis and elucidation; and with ‘unknown unknowns’, natural products can provide new insights into the biosynthetic knowledge and new tools for synthetic biology. The advances in the modern genetics tools,⁵² such as transformation-associated recombination (TAR),⁵³ Red/ET recombination,⁵⁴ and CRISPR-Cas9 genome editing^{55,56}, will accelerate synthetic biology and hence natural product research. In this context, it is imperative that a pharmacognosist committed to find antibiotics, has the appropriate biological knowledge and practical genetic tools. Extensive collaboration between scientific experts from various disciplines is the best way forward to accelerate the discovery of the new antibiotics.

Microorganisms occupy almost all the observed ecological niches, playing a pivotal role in maintaining the sustainable run of macroscopic or microscopic ecosystems. Many ecological phenomena are directly derived from natural roles of microbial natural products, such as rhizoxin causing rice seedling blight.⁵⁷ It might be fruitful to follow “ecology-inspired” approaches for antibiotic discovery, in view of the tremendous biodiversity on the earth. In this sense, nature is always generous to offer inspiration. What is important is to direct our attention to the interesting natural phenomena in life, consider the underlying principles, link this to the hidden microbes, and use various “OMICS” tools for bioprospecting. Auguste Rodin already said: “It's not the lack of beauty in life, but the lack of the eyes to find the beauty”.

The central dogma of molecular biology—the transmittance of information from DNA to RNA to protein—has now been extended to encompass the metabolome.^{58,59} There is corresponding “OMICS” developed in every step of the extended central dogma to facilitate natural products discovery (Figure 2). The fact that all natural products have a genetic basis suggests a new concept that I would like to refer to as “Genometabolomics”, which unifies the universally available knowledge of biosynthetic genes and biosynthetic protein reactivity, and can be exploited to predict and also characterize new natural products — and *vice versa* — by harnessing biosynthetic logic.

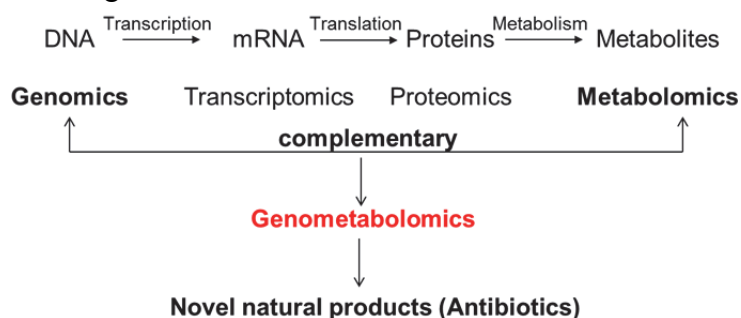


Figure 2. Genometabolomics for natural product research in actinomycetes.

REFERENCES

- (1) Fleming, A. *Br. J. Exp. Pathol.* **1929**, *10*, 226–236.
- (2) Woodford, N.; Turton, J. F.; Livermore, D. M. *FEMS Microbiol. Rev.* **2011**, *35*, 736–55.
- (3) Zhu, H.; Sandiford, S.; Wezel, G. van *J. Ind. Microbiol. ...* **2013**.
- (4) Rice, L. B. *J. Infect. Dis.* **2008**, *197*, 1079–1081.
- (5) Bode, H. B. *Curr. Opin. Chem. Biol.* **2009**, *13*, 224–230.
- (6) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat. Rev. Drug Discov.* **2007**, *6*, 29–40.
- (7) Nicolaou, K. C.; Chen, J. S.; Edmonds, D. J.; Estrada, A. a *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 660–719.
- (8) von Nussbaum, F.; Brands, M.; Hinzen, B.; Weigand, S.; Häbich, D. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 5072–129.
- (9) Kim, E.; Moore, B. S.; Yoon, Y. J. *Nat. Chem. Biol.* **2015**, *11*, 649–659.
- (10) Wender, P. a; Miller, B. L. *Nature* **2009**, *460*, 197–201.
- (11) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2007**, *70*, 461–77.
- (12) Bérdy, J. *J. Antibiot. (Tokyo)*. **2005**, *58*, 1–26.
- (13) Kaeberlein, T.; Lewis, K.; Epstein, S. S. *Science* **2002**, *296*, 1127–1129.
- (14) Ling, L. L.; Schneider, T.; Peoples, A. J.; Spoering, A. L.; Engels, I.; Conlon, B. P.; Mueller, A.; Hughes, D. E.; Epstein, S.; Jones, M.; Lazarides, L.; Steadman, V. a; Cohen, D. R.; Felix, C. R.; Fetterman, K. A.; Millett, W. P.; Nitti, A. G.; Zullo, A. M.; Chen, C.; Lewis, K. *Nature* **2015**, *517*, 455–459.
- (15) Wilson, M. C.; Mori, T.; Rückert, C.; Uriä, A. R.; Helf, M. J.; Takada, K.; Gernert, C.; Steffens, U. a E.; Heycke, N.; Schmitt, S.; Rinke, C.; Helfrich, E. J. N.; Brachmann, A. O.; Gurgui, C.; Wakimoto, T.; Kracht, M.; Crüsemann, M.; Hentschel, U.; Abe, I.; Matsunaga, S.; Kalinowski, J.; Takeyama, H.; Piel, J. *Nature* **2014**, *506*, 58–62.
- (16) Bentley, S. D.; Cerdeño-Tárraga, K. F. C. A.-M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Bentley, S. D.; Harper, D.; Bateman, A.; Brown, S.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Larke, L.; Murphy, L.; Oliver, K.; Rabbinowitsch, E.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Woodward, J.; Barrell, B. G.; Parkhill, J. *Nature* **2002**, *3*, 141–147.
- (17) Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S. *Nat. Biotechnol.* **2003**, *21*, 526–531.
- (18) Pauli, G. F.; Chen, S. N.; Friesen, J. B.; McAlpine, J. B.; Jaki, B. U. *J. Nat. Prod.* **2012**, *75*, 1243–1255.
- (19) Cooper, M. A.; Shlaes, D. *Nature* **2011**, *472*, 32.
- (20) Bertrand, S.; Bohni, N.; Schnee, S.; Schumpp, O.; Gindro, K.; Wolfender, J.-L. *Biotechnol. Adv.* **2014**, *32*, 1180–1204.
- (21) Hosaka, T.; Ohnishi-Kameyama, M.; Muramatsu, H.; Murakami, K.; Tsurumi, Y.; Kodani, S.; Yoshida, M.; Fujie, A.; Ochi, K. *Nat. Biotechnol.* **2009**, *27*, 462–4.
- (22) Gubbens, J.; Zhu, H.; Girard, G.; Song, L.; Florea, B. I.; Aston, P.; Ichinose, K.; Filippov, D. V.; Choi, Y. H.; Overkleeft, H. S.; Challis, G. L.; van Wezel, G. P. *Chem.*

- Biol.* **2014**, *21*, 707–718.
- (23) Zinad, D. S.; Shaaban, K. A.; Abdalla, M. A.; Islam, M. T.; Schüffler, A.; Laatsch, H. *Nat. Prod. Commun.* **2011**, *6*, 45–48.
 - (24) Singh, B.; Parshad, R.; Khajuria, R. K.; Guru, S. K.; Pathania, A. S.; Sharma, R.; Chib, R.; Aravinda, S.; Gupta, V. K.; Khan, I. a.; Bhushan, S.; Bharate, S. B.; Vishwakarma, R. A. *Tetrahedron Lett.* **2013**, *54*, 6695–6699.
 - (25) Cox, D. G.; Oh, J.; Keasling, A.; Colson, K.; Hamann, M. T. *Biochim. Biophys. Acta* **2014**, *1840*, 3460–3474.
 - (26) Månsson, M.; Phipps, R. K.; Gram, L.; Munro, M. H. G.; Larsen, T. O.; Nielsen, K. F. *J. Nat. Prod.* **2010**, *73*, 1126–1132.
 - (27) Seger, C.; Sturm, S.; Stuppner, H. *Nat. Prod. Rep.* **2013**, *30*, 970–987.
 - (28) Lambert, M.; Stærk, D.; Hansen, S. H.; Sairafianpour, M.; Jaroszewski, J. W. *J. Nat. Prod.* **2005**, *68*, 1500–1509.
 - (29) Johansen, K. T.; Wubshet, S. G.; Nyberg, N. T.; Jaroszewski, J. W. *J. Nat. Prod.* **2011**, *74*, 2454–61.
 - (30) Exarchou, V.; Godejohann, M.; Beek, T. A. Van; Gerothanassis, I. P.; Vervoort, J. *J. Nat. Prod.* **2003**, *75*, 6288–6294.
 - (31) Challis, G. L.; Hopwood, D. A. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 14555–14561.
 - (32) Bode, H. B.; Bethe, B.; H⁺fs, R.; Zeeck, A. *ChemBioChem* **2002**, *3*, 619–627.
 - (33) Wu, C.; van Wezel, G. P.; Choi, Y. H. *J. Antibiot. (Tokyo)*. **2015**, *68*, 445–452.
 - (34) Heine, D.; Martin, K.; Hertweck, C. *J. Nat. Prod.* **2014**, *77*, 1083–1087.
 - (35) Wu, C.; Medema, M. H.; Läkamp, R. M.; Zhang, L.; Dorrestein, P. C.; Choi, Y. H.; van Wezel, G. P. *ACS Chem. Biol.* **2016**, *11*, 478–490.
 - (36) Yang, J. Y.; Sanchez, L. M.; Rath, C. M.; Liu, X.; Boudreau, P. D.; Bruns, N.; Glukhov, E.; Wodtke, A.; De Felicio, R.; Fenner, A.; Wong, W. R.; Linington, R. G.; Zhang, L.; Debonsi, H. M.; Gerwick, W. H.; Dorrestein, P. C. *J. Nat. Prod.* **2013**, *76*, 1686–1699.
 - (37) Richter, T. K. S.; Hughes, C. C.; Moore, B. S. *Environ. Microbiol.* **2014**, *17*, 2158–2171.
 - (38) Singh, R.; Reynolds, K. A. *J. Nat. Prod.* **2016**, *79*, 240–243.
 - (39) Vingadassalon, A.; Lorieux, F.; Juguet, M.; Gof, G. Le; Gerbaud, C.; Pernodet, J.-L.; Lautru, S. *ACS Chem. Biol.* **2015**, *10*, 601–610.
 - (40) Okamoto, S.; Taguchi, T.; Ochi, K.; Ichinose, K. *Chem. Biol.* **2009**, *16*, 226–236.
 - (41) Metsä-Ketelä, M.; Oja, T.; Taguchi, T.; Okamoto, S.; Ichinose, K. *Curr. Opin. Chem. Biol.* **2013**, *17*, 562–570.
 - (42) Taguchi, T.; Itou, K.; Ebizuka, Y.; Malpartida, F.; Hopwood, D. A.; Surti, C. M.; Booker-Milburn, K. I.; Stephenson, G. R.; Ichinose, K. *J. Antibiot. (Tokyo)*. **2000**, *53*, 144–152.
 - (43) Bartel, P. L.; Zhu, C. B.; Lampel, J. S.; Dosch, D. C.; Connors, N. C.; Strohl, W. R.; Beale, J. M.; Floss, H. G. *J. Bacteriol.* **1990**, *172*, 4816–4826.
 - (44) Taguchi, T.; Ebizuka, Y.; Hopwood, D. A.; Ichinose, K. *Tetrahedron Lett.* **2000**, *41*, 5253–5256.
 - (45) McDaniel, R.; Ebert-Khosla, S.; Fu, H.; Hopwood, D. A.; Khosla, C. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 11542–11546.
 - (46) Shen, Y.; Yoon, P.; Yu, T. W.; Floss, H. G.; Hopwood, D.; Moore, B. S. *Proc. Natl.*

- Acad. Sci. U. S. A.* **1999**, 96, 3622–3627.
- (47) Smanski, M. J.; Zhou, H.; Claesen, J.; Shen, B.; Fischbach, M. A.; Voigt, C. A. *Nat. Rev. Microbiol.* **2016**, 14, 135–149.
- (48) Winter, J. M.; Behnken, S.; Hertweck, C. *Curr. Opin. Chem. Biol.* **2011**, 15, 22–31.
- (49) Zazopoulos, E.; Huang, K.; Staffa, A.; Liu, W.; Bachmann, B. O.; Nonaka, K.; Ahlert, J.; Thorson, J. S.; Shen, B.; Farnet, C. M. *Nat. Biotechnol.* **2003**, 21, 187–190.
- (50) Walsh, C. T.; Fischbach, M. A. *J. Am. Chem. Soc.* **2010**, 132, 2469–2493.
- (51) Seyedsayamdost, M. R.; Clardy, J. *ACS Synth. Biol.* **2014**, 3, 745–747.
- (52) Weber, T.; Charusanti, P.; Musiol-Kroll, E. M.; Jiang, X.; Tong, Y.; Kim, H. U.; Lee, S. Y. *Trends Biotechnol.* **2015**, 33, 15–26.
- (53) Ongley, S. E.; Bian, X.; Neilan, B. a; Müller, R. *Nat. Prod. Rep.* **2013**, 30, 1121–1138.
- (54) Fu, J.; Bian, X.; Hu, S.; Wang, H.; Huang, F.; Seibert, P. M.; Plaza, A.; Xia, L.; Müller, R.; Stewart, a F.; Zhang, Y. *Nat. Biotechnol.* **2012**, 30, 440–446.
- (55) Sander, J. D.; Joung, J. K. *Nat. Biotechnol.* **2014**, 32, 347–355.
- (56) Cobb, R. E.; Wang, Y.; Zhao, H. *ACS Synth. Biol.* **2015**, 4, 723–728.
- (57) Partida-Martinez, L. P.; Hertweck, C. *Nature* **2005**, 437, 884–888.
- (58) Bachmann, B. O.; Van Lanen, S. G.; Baltz, R. H. *J. Ind. Microbiol. Biotechnol.* **2014**, 41, 175–184.
- (59) Tietz, J. I.; Mitchell, D. *Curr. Top. Med. Chem.* **2015**, 16, 1–1.

SAMENVATTING EN DISCUSSIE

1. Samenvatting en discussie van experimentele resultaten

Sinds de baanbrekende ontdekking van penicilline door Fleming is het gebruik van antibiotica explosief toegenomen. Dit heeft een scherpe stijging in multiresistente infectieziekten met zich meegebracht die zich al sinds de jaren 80 van de vorige eeuw zich over de wereld verspreiden. Veel ‘multi-drug resistente’ (MDR) bacteriën zijn inmiddels resistent tegen veel, zo niet alle, antibiotica die in de kliniek worden gebruikt. Dit brengt een grote maatschappelijke uitdaging met zich mee. Zodra een nieuw antibioticum op de markt verschijnt duurt het niet lang totdat resistentie ook verschijnt, dit geldt ook voor de laatste linie middelen zoals vancomycine. Dit heeft er toe geleid dat ziekten die ooit als bijna uitgeroeid beschouwd werden nu weer terugkomen, met tuberculose als duidelijk voorbeeld. De zogenaamde ESKAPE pathogenen (*Enterobacteriaceae*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, en *Enterococcus spp.*) vormen echter wellicht de grootste bedreiging. De toename in resistentie komt tegelijkertijd met een afname van ‘drug-discovery pipelines’. Deze afname is vooral te wijten aan het feit dat het ontdekken van nieuwe antibiotica veel tijd kost en daardoor erg duur is. Nieuwe ideeën en inzichten moeten worden geïntroduceerd om het onderzoek naar antibiotica te revitaliseren en om de kans om nieuwe antimicrobiële middelen te vinden te vergroten. Zeker verbindingen met een nieuwe werking en/of unieke chemische structuren zijn van groot belang voor de gezondheidszorg.

Moeten wij nu de hoop om deze obstakels te overkomen verliezen? Een oude Chinese filosofie zegt “dingen in de wereld zijn allelopatie”. Als dit zo is, wat kan dan deze minuscule organismen stoppen? Zoals altijd moet het antwoord in de natuur worden gevonden. Miljarden jaren van evolutie hebben geresulteerd in een buitengewoon diversiteit aan chemische verbindingen met zeer diverse functionaliteiten. Met andere woorden, het is als een tekstboek geschreven in de taal van chemie. Moderne analytische technieken maken het mogelijk om structuren aan bio- en fysiologische functies te koppelen. Met de grote stappen die de synthetische biologie en chemie maken is het niet alleen mogelijk om kopieën te maken van natuurlijke moleculen maar ook om geheel nieuwe biologisch geïnspireerde verbindingen te maken. Deze stoffen kunnen geheel nieuwe functies bezitten die hun gelijke niet hebben in de natuur en zullen ieder opzicht van ons leven kunnen beïnvloeden, zo niet vormen.

Om terug te komen op antibiotica, waar kunnen wij nieuwe moleculen vinden om de pijplijn naar de Parma-industrie opnieuw te bevoorraden? Ook al hebben wetenschappers uitputtend onderzoek gedaan op zoek naar nieuwe moleculen die kunnen functioneren als onze toekomstige medicijnen, toch wordt de biosynthese potentieel van microben nog onderschat en is waarschijnlijk slechts het topje van de ijsberg ontdekt. Hier zijn een aantal redenen voor aan te geven. Ten eerste is ondanks de grote hoeveelheid filamenteuze micro-organismen die zich in de bodem en in de oceanen bevinden slechts een kleine deel in het lab te kweken. Recentelijk is aangetoond dat methoden om "onkweekbare" bacteriën toch op te groeien mogelijk een nieuw gebied in de chemie van ‘natural products’ kan openen. Ten

tweede zijn veel biosynthese clusters van de kweekbare microben niet actief in laboratoriumomstandigheden. Zo heeft bijvoorbeeld het bepalen van de genomsequentie van actinomyceten aangetoond dat zelfs bij de best bestudeerde modelorganismen de productiecapaciteit volledig onderschat is. Ten derde is het zo dat als een gencluster tot expressie komt dat het product hiervan vaak niet ontdekt wordt vanwege de intensieve fytochemische analyse of chromatografische zuivering die nodig is. Dit is zeker het geval als het deze expressie veel lager is dan die van andere bioactieve verbindingen binnen hetzelfde organisme.

Actinomyceten zijn belangrijke micro-organismen voor de (medische) industrie vanwege hun vermogen om een breed scala aan secundaire metabolieten te produceren. Meer dan tweederde van alle ons bekende antibiotica wordt door deze filamenteuze bacteriën gemaakt, de meerderheid hiervan door *Streptomyces*. De komst van Next Generation Sequencing (NGS) heeft aangetoond dat zelfs de best bestudeerde modelactinomyceet *Streptomyces coelicolor* A3(2) nog veel onbenutte bronnen heeft voor natuurstoffen. Dit toont aan dat er een nieuwe aanpak nodig is om gebruik te kunnen maken van dit reservoir van bioactieve verbindingen om de drug-discovery pijplijn te vullen.

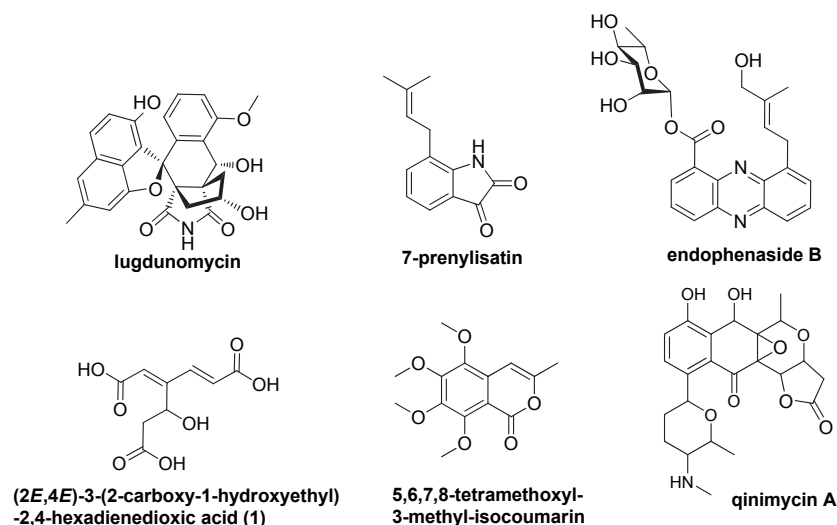
In dit proefschrift worden nieuwe methodieken die de ontdekking van nieuwe antibiotica bewerkstelligen gepresenteerd. Er wordt nadruk gelegd op methoden en ideeën uit zowel chemisch als biologisch perspectief. Veel van deze inspanningen hebben als doel om nieuwe verbindingen te vinden met nieuwe werkingsmethoden, met nadruk op het ontsluiten van slapende of cryptische biosyntheseroutes in bacteriën en het vervolgens karakteriseren van de moleculen die worden geproduceerd. Er wordt ‘proof of concept’ geleverd dat het integreren van NMR metabolomics met verschillende strategieën voor het induceren van antibiotica een efficiënte methode is om de ontdekking van nieuwe antibiotica uit actinomyceten te bewerkstelligen.

De co-cultivatatie van verschillende micro-organismen heeft de afgelopen jaren veel interesse en toepassingen gevonden, vooral omdat het zich bewezen heeft als effectieve strategie voor het verkrijgen van unieke structuren met nieuwe biologische activiteit. In Hoofdstuk 4 van dit proefschrift wordt aangetoond dat co-cultivatatie van *Streptomyces coelicolor* met de draadschimmel *Aspergillus niger* grote verandering teweeg brengt in de secundaire metabolieten die ze produceren. NMR metabolomics laat hier zien dat een aantal verbindingen, waaronder het cyclische dipeptide cyclo(Phe-Phe) en 2-hydroxyphenylacetic acid, specifiek worden gemaakt tijdens co-cultivatatie. Deze verbindingen worden door *A. niger* geproduceerd als response aan de chemicaliën geproduceerd door *S. coelicolor*. Geïnspireerd door deze co-cultivatatie studie is ook de biotransformatie van hydroxycinnamic acid onderzocht, wat liet zien dat dit tijdens co-cultivatatie werd omgezet in de nieuwe verbinding (2*E*,4*E*)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid.

Selectie van mutanten die resistent zijn tegen het antibioticum streptomycine is een andere breed toepasbare methode om de productie van antibiotica te bevorderen. In hoofdstuk 5 wordt middels NMR metabolomics aangetoond hoe het selecteren voor streptomycine resistente variant van *Streptomyces* sp. MBT28 leidt tot de productie van een nieuw molecuul, namelijk 7-prenylindole, met antimicrobiële activiteit tegen *Bacillus subtilis*. Verdere analyse middels genome mining gecombineerd met *Natural Product Proteomining* leidde vervolgens tot identificatie van een gencluster (*isa*), welke codeert voor onder meer de isoprenyl

transferase die phenylalanine verandert in 7-prenylindole. Dit onderzoek onderstreept dat gecombineerde ‘omics-technieken de ontdekking van nieuwe moleculen kan vergemakkelijken en tegelijkertijd een duidelijke manier biedt om het verantwoordelijke gen-cluster te identificeren. Toch zijn de biosynthesegenen voor het eindproduct 7-prenylisatin niet in één enkel gencluster te vinden in het genoom van *Streptomyces* sp. MBT28: de oxygenases die verantwoordelijk zijn voor de oxidatie van 7-prenylindole naar 7-prenylisatin zijn namelijk niet binnen of nabij de *isa* gencluster gevonden.

Secundaire metabolieten worden vaak in een groeifase afhankelijke manier geproduceerd door actinomyceten, dit betekent dat het cruciaal is om de optimale oogst tijd te vinden om succesvol bioactieve metabolieten isoleren. In hoofdstuk 6 wordt de groei-fase-afhankelijkheid van secundaire metabolieten productie van *Streptomyces* sp. MBT76 geanalyseerd met behulp van metabolomics op basis van NMR. Er zijn hierbij een groep gemethoxyleerde isocoumarines gekarakteriseerd, waaronder een *tetra*-gemethoxyleerde verbinding: 5,6,7,8-tetramethoxyl-3-methyl-isocoumarin. Multivariate data analyse heeft aangegeven dat methoxylering leidde tot een toename in de antimicrobiële activiteit van isocoumarines, zowel van gemethoxyleerde als van niet-gemethoxyleerde isocoumarines. Verder leidde biotransformatie van de isoflavone genistein tot een hele reeks aan gemethoxyleerde isoflavones welke exact dezelfde modificaties ondergaan als isocoumarine. Hoewel isocoumarine-achtige metabolieten veel uit actinomyceten worden geïsoleerd is de onderliggende genetische basis voor dit type polyketide nog nagenoeg onbekend. Als men de gelijkenissen van de structuur van 6,8-dihydroxy-3-methyl-isocoumarine en genistein bekijkt is het niet ondenkbaar dat dezelfde *O*-methyltransferase gebruikt wordt voor de methylering van beide substraten. Met dit in gedachte is het zinnig om de genetische basis van isocoumarine biosynthese in *Streptomyces* sp. MBT76 te karakteriseren. Dit zou nieuw licht kunnen laten schijnen op de methyltransferase die verantwoordelijk is voor de post-PKS modificaties omdat deze genen voor PKS assemblage en post-PKS modificatie zich altijd clusteren in de genomen van actinomyceten. Verder zal dit ook inzichten kunnen verschaffen die gebruikt kan worden voor verdere genetische modificatie, zoals combinatorieke biosynthese, van dit type antibiotica.



Figuur 1. representatieve nieuwe moleculen die in dit proefschrift worden beschreven. In het bijzonder Lugdunomycine, een angucycline-derivaat met een totaal nieuwe chemische structuur.

In hoofdstuk 7 wordt een methode beschreven die gebruik maakt van de overexpressie van een biosynthetische gencluster (BGC). Dit leidde tot de activering van een BGC die normaal gesproken nauwelijks tot expressie komt, namelijk de *qin* cluster voor een type II PKS in *Streptomyces* sp. MBT76. Het chemotype en de genotype van deze recombinante stam werden door NMR-gebaseerde metabolieten profiling op efficiënte wijze gelinkt. Hierbij laten wij als eerste de intrigerende architectuur van qinimycines A-C zien. Dit is een groep van nieuwe pyranonaphtoquinonen die kenmerken hebben zoals 8-glycosylering, 5,14-epoxydatie en 13-hydroxylering. De methode die in hoofdstuk 7 uiteen word gezet is vooral toepasbaar als er vanuit voorspellende bioinformatica analyse interessante nieuwe structuren uit BGCs worden opgedaan. NMR-profilering van de moleculen die in een zeer complexe matrix worden geproduceerd maakt het hierbij mogelijk om zeer snel het chemotype aan het genotype te koppelen. Dit kan de downstream isolatie en structuur bepaling stroomlijnen. Er zijn echter wel intrinsieke beperkingen aan deze syntheseroute-specifieke strategie. Zo is het niet mogelijk dit high-throughput te doen en zijn in sommige gevallen de syntheseroute-specifieke activatoren niet gemakkelijk te identificeren. Verder is het ook zo dat veel van de uit de natuur geïsoleerde stammen genetisch niet toegankelijk. In Figuur 1 zijn enkele moleculen weergegeven die tijdens NMR-gebaseerde metabolomics studies zijn geïdentificeerd.

Metabolomics-gestuurde drug discovery biedt ten opzichte van de gewoonlijke bioactiviteit-gestuurde screening veel voordelen. Het grootste bezwaar van de meer traditionele methoden is dat de chemische structuur en bioactiviteit van het target molecuul onbekend zijn voordat het intensieve en tijdrovende werk van isolatie begint. Wel moet worden opgemerkt dat er ook enkele nadelen zijn met betrekking tot het werken met microbiële extracten in NMR metabolomics. Voor dit proefschrift zijn gewoonlijk actinomyceten gekweekt in vloeibare en vaste minimale media. Veel componenten uit deze media worden onlosmakelijk mee-geëxtraheerd met ethylacetaat, en zijn in veel gevallen aanwezig in hogere concentraties dan de secundaire metabolieten uit de actinomyceten die worden gezocht. Ondanks dat NMR spectroscopie een kwantitatieve techniek is, blijft het relatief ongevoelig ten opzichte van massa spectrometrie (MS). De signalen van de protonen in NMR van de gezochte verbinding zijn vaak in kleine hoeveelheden aanwezig, en vele malen minder dan de signalen van de mediacomponenten. Hierdoor zijn zij vaak verborgen door de sterkere signalen van de media. Dit heeft als gevolg dat metabolomics op basis van NMR het beste werkt als het gecombineerd wordt met een antibiotica-inducerende methode. Het is daarentegen ook mogelijk om de ongewenste verbindingen bij de eerste stap te verwijderen om zodoende de gewenste signalen te accentueren. Om dit bereiken zal een allesomvattende extractiemethode moeten worden gebruikt in plaats van het enkele organische oplosmiddel (ethylacetaat); deze vloeistof-vloeistof partitie zal nog verder moeten worden onderzocht. Een techniek die veelbelovend lijkt is solid-phase extractie (SPE), omdat er een breed scala aan adsorberende materialen met verschillende fysicochemische eigenschappen beschikbaar is. Behalve dat zou het nadeel van ongevoeligheid in metabole profiling met behulp van NMR kunnen worden voorkomen door gebruik te maken van HPLC-NMR technieken. Het grootste voordeel van het gebruik van on-flow HPLC-NMR is directe informatie over de chemische structuur van de pieken uit de HPLC door de NMR spectrometer als detector te gebruiken. Om de inherent lagere gevoeligheid van NMR

spectrometrie en het gelimiteerde injectievolume van HPLC chromatografie te omzeilen kan SPE aan het systeem gekoppeld worden, om zo bij herhaalde injecties de analyten ‘te vangen’ op hetzelfde SPE patroon. Deze geadsorbeerde analyten kunnen dan van de SPE kolom worden geëluëerd met gedeutereerde oplosmiddelen voor NMR analyse. Door gebruik te maken van deze LC-SPE-NMR methode is het mogelijk om omvangrijke structurele informatie en chemische samenstelling van ruwe extracten online te verkrijgen. Hierdoor kunnen we sneller voorspellen of er mogelijk interessante moleculen aanwezig zijn in een bepaald extract. Als men dan verder MS in het systeem integreert, zoals bij LC-DAD-SPE-MS-NMR instrumenten, kan de structuur van nieuwe verbindingen *de novo* online bepaald worden zonder dat er veel tijdrovende zuiveringsstappen van het analyt uit de zeer complexe matrix nodig zijn.

Er zijn waarschijnlijk geen biosynthetische genclusters niet nooit tot expressie komen. De verbindingen zullen ongetwijfeld onder een bepaalde conditie geproduceerd worden in de oorspronkelijke leefomgeving. Als dit niet het geval zou zijn dan zouden veel van deze clusters verloren zijn gegaan tijdens de evolutie, of zouden überhaupt nooit zijn ontstaan. Het is wel zo dat veel microbiële secundaire metabolieten in slechts geringe hoeveelheden worden geproduceerd, vaak onder de detectiegrens van NMR. Hierdoor kunnen zij gemist worden in NMR-gebaseerde metabolomics. Logischerwijs is het ook zo dat componenten die een kleinere hoeveelheid tot expressie komen ook veel gemist gaan worden in traditionele screeningstechnieken, en is het dus ook waarschijnlijker dat zij nog niet eerder geïsoleerd c.q. gekarakteriseerd zijn, in tegenstelling tot de verbindingen die onder routine laboratoriumcondities wel hoog tot expressie komen. Een logische methode om verbindingen die laag tot expressie komen toch te vinden is schaalvergroting. Hoewel deze aanpak tijdrovend en arbeidsintensief is heeft deze techniek wel vruchten afgeworpen. Het bepalen van de chemische skeletten door alle waarneembare verbindingen te isoleren uit een biologisch monster met een systematische zuiveringsmethode komt veel voor in het veld van traditionele natuurstoffen. Het is eigenlijk normaal dat een enkele microbiële bron meerdere secundaire metabolieten produceert; en systematische isolatie is ondubbelzinnig de meest efficiënte manier om meer licht te schijnen op de chemische diversiteit van een bepaalde actinomyceet. Naast het ontdekken van nieuwe chemische verbindingen geeft systematische zuivering vaak ook een beter inzicht in de biosynthetische routes van een complexe natuurstof. Dit komt omdat de structuren van meegeïsoleerde tussen- of zijproducten inzicht kunnen geven in de bioassemblage van het eindproduct. Verder is het ook zo dat de verandering van groeicondities kan leiden tot een complete verandering in het metabole profiel van micro-organismen, wat heeft geleid tot de OSMAC (one strain many compounds) strategie voor microbiële drug discovery. Een micro-organisme zal nooit zijn volledige arsenaal aan antimicrobiële verbindingen tegelijk in een bepaalde groeiconditie aanmaken, dit zou namelijk te kostbaar zijn omdat het produceren van complexe moleculen veel energie en dure bouwstenen kost. Daarom worden veel verbindingen alleen aangemaakt wanneer deze nodig zijn, hetgeen competitief voordeel geeft wanneer de omgevingscondities veranderen. Daarom is het variëren van de groeicondities een logische aanpak voor de inductie, c.q. optimalisatie, van de productie van secundaire metabolieten. In deze context vergroot de systematische verandering van groeicondities (medium samenstelling, beluchting, pH, temperatuur, toevoeging van chemicaliën, etc.) voor een micro-organisme de kans op het vinden van

secundaire metabolieten uit een enkele bacteriestam. De statische cultivatie van micro-organismen is van bijzonder belang voor de diversificatie van secundaire metabolieten in het laboratorium, dit komt doordat deze methode het meest overeenstemt met de natuurlijke groeicondities. Het is in het verleden al aangetoond dat de coproductie door actinomyceten in de natuur van twee of meer secundaire metabolieten nodig is om concurrenten de baas te blijven. Het is ook voorgesteld dat synergie en toeval samen de drijvende kracht voor de productie van meerdere secundaire metabolieten door deze onbeweeglijke saprofyt is. De OSMAC strategie was toegepast in de ontdekking van nieuwe natural products van *Streptomyces* sp. QL37 (hoofdstuk 8) en *Kitasatospora* sp. MBT66 (hoofdstukken 9 en 10). Er zijn daarbij verschillende koolstofbronnen (glycerol, mannitol, glucose, rhamnose, *N*-acetylglucosamine) en verschillende toevoegingen (gist extract, aardappel extract, pepton, zetmeel, sojameel), of basische (pH 10) condities getest in vloeibare NMMP media in schuldkolven, net als in vaste MM media; hierbij bleken vaste voedingsbodems beter dan vloeibare voor productie van secundaire metabolieten door deze twee actinomyceten.

Systematische isolatie van *Streptomyces* sp. QL37 gegroeid in MM met mannitol en glycerol leidde tot de productie van lugdunomycines. Dit zijn natuurstoffen met een volledig nieuwe chemische structuur (Figuur 1). Hun chemische structuur geeft aan dat de lugdunomycines zijn afgeleid van angucyclines, echter is hun basisstructuur gemodificeerd om een zeer complex molecuul te geven met een ongehoorde drie-dimensionele structuur. Lugdunomycines bevatten een benzaza[4,3,3]propellane skelet met daaraan spirocyclische 2*H*-naphthol[1,8,bc]furan en twee volledig koolstof quaternaire centra met daarin 5 aangrenzende stereogene koolstoffen. De opvallende ruggengraat van benzaza[4,3,3]propellane-6-spiro-2'-2*H*-naphthol[1,8,bc]furan is tot dusver ongehoorde scheikunde. Verdere fermentatie van *Streptomyces* sp. QL37 op R5 agarplaten met peptone gevolgd door systematische isolatie resulteerde in de karakterisatie van velen anderen herrangschikte angucyclines (hoofdstuk 8). Dit had als gevolg dat in totaal bijna 30 angucycline derivaten zijn gekarakteriseerd welke in vier verschillende klassen onderverdeeld konden worden: de zeven-ring lugdunomycines, drie-ring anthraquinonen, vier-ring benz[a]nthraquinonen, en vijf-ring limamycines. Metaboliet gerichte genome mining identificeerde een type II PKS (*lug*) gencluster welke verantwoordelijk is voor de angucyclineproductie in *Streptomyces* sp. QL37. Dit werd bevestigd door mutatie-analyse van de minimale PKS genen. Het is echter mogelijk dat de *lug* cluster niet voldoende is voor de synthese van alle gevonden angucyclines; met andere woorden, op dit moment kan niet worden uitgesloten dat genen van buiten het cluster eveneens nodig zijn om de complexe structuur van lugdunomycine te vormen. Wat verder ook nog moet worden onderzocht is dat alle gevonden angucycline derivaten een oxidatieve C-C splitsing hebben in de post-PKS stap die de basisstructuur van de benz[a]nthraquinone vormt. Het is aannemelijk dat er meerdere oxygenases nodig zijn voor de modificatie van lugdunomycine omdat er meerdere verschillende oxidaties plaatsvinden tijdens de synthese. De oxygenase die verantwoordelijk is voor de opening van de quinon ring via Baeyer-Villiger oxidatie is van groot belang voor de biosynthese van lugdunomycine en zou mogelijk ook andere polyketide antibiotica kunnen modifieren zoals tetracycline.

Hoofdstuk 9 beschrijft de systematische analyse van *Kitasatospora* sp. MBT66 en de ontdekking van zowel vijf nieuwe gerhamnosyleerde phenazines (genaamd endophenazines

A-E), als de eveneens gerhamnosyleerde plecomacrolide leucanicidine. Met behulp van metaboliet gerichte genome mining van MBT66 werd een endophenaside gencluster geïdentificeerd welke bijna identiek is aan de endophenaside BGC (*epa*) van *Kitasatospora* sp. HKI 714. De genen die verantwoordelijk zijn voor de glycosylering van de endophenazine aglycons worden beschreven in hoofdstuk 10. Het is interessant om te zien dat de *de novo* rhamnosylering van beiden type centrale structuur in *Kitasatospora* sp. MBT66, phenazine en plecomacrolide, beiden hun oorsprong vinden in de modificatie via dezelfde glycosyltransferase en methyltransferase - LeuAB. Deze opvallende substraat flexibiliteit van LeuAB was de aanleiding om de structurele diversiteit van endophenasides in MBT66 opnieuw te onderzoeken. Hierbij maakte moleculaire netwerken middels MS/MS fragmentatie de karakterisering van negen nieuwe endophenasides mogelijk, waaronder een nog niet eerder gevonden tautomerische glyceride phenazine (hoofdstuk 10). Moleculaire netwerken zijn een effectief gereedschap voor chemische dereplicatie en een manier om nieuwe verbindingen op prioriteit te rangschikken. Het fundamentele principe achter deze netwerken is dat natuurstoffen met vergelijkbare structuur ook vergelijkbare MS/MS fragmentatie patronen zullen geven. De lage selectiviteit van LeuB voor het substraat kan mogelijk worden gebruikt om een breed scala aan natural products te 'glycorandomisen'. Dit biedt een alternatieve methode om nieuwe antibiotica te ontwikkelen met een post-biosynthese modificatie strategie.

In actinomyceten liggen de genen voor enzymen van de biosyntheseroutes doorgaans geclusterd op het chromosoom, en hiervan kunnen we gebruik maken in de bioinformatica bij het zoeken naar kandidaatgenclusters voor een bepaald antibioticum. Het is echter geen zekerheid dat alle genen bij elkaar liggen, gezien de cross-talk tussen de *epa*, *baf* en *rml* genclusters voor endophenaside productie in *Kitasatospora* sp. MBT66 (hoofdstuk 10), en de samenwerking van de *isa* cluster en een elders gelegen gen voor een oxygenase die leiden tot productie van 7-prenylisatine in *Streptomyces* sp. MBT28 (hoofdstuk 5), en wellicht ook de genetica voor lugdunomycine productie in *Streptomyces* sp. QL37 (hoofdstuk 8). Daarom kunnen we concluderen dat samenwerking tussen twee of wellicht zelfs meer genclusters bij de vorming van secundaire metabolieten vaker voorkomt in actinomyceten dan tot nu toe wordt aangenomen. Dit is eerder waargenomen in actinomyceten van het genus *Salinispora* bij de biosynthese van sioxanthin. Ook bij de biosynthese van prodiginines in *Streptomyces coelicolor* en pyrrolamiden in *Streptomyces netropsis* zijn verschillende regio's van het chromosoom betrokken. In andere gevallen kan één gencluster leiden tot meerdere producten, zoals bij de *act* (type II PKS) cluster van *S. coelicolor* die verschillende aryl-verbindingen kan synthetiseren, zoals pyranonaphthoquinonen, anthraquinonen, naphthoquinonen en mutacine. Het metabole profiel van een recombinante *Streptomyces* stam die het minimale PKS van de *whiE* gencluster uit *S. coelicolor* tot expressie bracht resulteerde in zo'n 30 polyketiden van verschillende lengtes en cyclisatiepatronen. Enkele voorbeelden staan in dit proefschrift, namelijk productie van meer dan 20 phenazines in *Kitasatospora* sp. MBT66 (hoofdstuk 9 en 10), of de productie van tenminste vier verschillende chemische skeletten met tientallen verschillende metabolieten door de *lug* BGC in *Streptomyces* sp. QL37 (hoofdstuk 8). Met dit in het achterhoofd is het niet ondenkbaar dat bij de voortgaande trend om BGC's te synthetiseren middels synthetische biologie als strategie (hoofdstuk 1), men het risico loopt dat niet alle genen die nodig zijn voor de productie daadwerkelijk worden meegenomen.

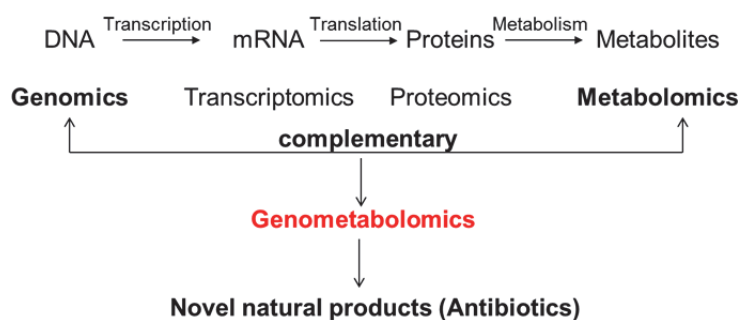
Ongeacht de pijplijnen die ontwikkeld worden voor de ontdekking van nieuwe antibiotica, zal de fytochemische analyse van microbiële culturen om de moleculen van interesse te profileren een gemeenschappelijke factor zijn in elk proces. Een rationele extractieprocedure zal dus van groot belang zijn om de ongewenste componenten (zoals uit de kweekmedia) te verwijderen en zoveel mogelijk secundaire metabolieten met diverse polariteiten te isoleren. In hoofdstukken 4-10 is ethylacetaat gebruikt om metabolieten die door de actinomyceten werden geproduceerd te extraheren. Zoals eerder besproken zijn veel mediacomponenten meegeëxtraheerd, in veel gevallen zijn deze relatief meer prominent aanwezig dan de gezochte (nieuwe) secundaire metabolieten. Dit waren vooral apolaire stoffen als 2,5-di-tert-butylphenol, 3,4-di-tert-butylphenol en (3,5-di-tert-butyl-4-hydroxyphenyl)-propionzuur octadecyl ester. Dit heeft een verregaand effect op de structuurbepaling en bioactiviteitsassays, zeker als de gezochte verbindingen in lage concentraties aanwezig zijn. Een ander nadeel van ethylacetaat is dat het minder geschikt is voor de extractie van polaire stoffen, zoals bijvoorbeeld de wateroplosbare aminoglycosiden. In hoofdstuk 11 is daarom de toepassing van het natuurlijke polymeer cellulose onderzocht als middel voor de absorptie van antibiotica uit een polaire matrix. Dit onderzoek heeft aangetoond dat solid-phase extractie op basis van cellulose als complementaire methode kan worden beschouwd naast vloeibaar-vloeibaar extractie om een breed scala aan secundaire metabolieten uit actinomyceten te ontdekken, zeker voor wateroplosbare antibiotica die bij veel extractiemethoden worden gemist.

2. Vooruitzichten

Met de snel groeiende toename in antibioticumresistentie is het van groot belang om nieuwe antibiotica te ontwikkelen. Historisch gezien is de ontdekking van antibiotica en andere natuurstoffen het domein van de scheikunde geweest met een nadruk op de isolatie van zuivere producten, structuurbepaling en hun (semi-)totaalsynthese. De biologie gaat ook een steeds grotere rol spelen in het (post-)genomics tijdperk. De recentelijke explosie aan publicaties over genomics methoden voor de ontdekking van nieuwe antibiotica laat zien dat de ‘genotype-to-chemotype’ roadmap een veelbelovende strategie is die het traditionele onderzoek sterk vernieuwt. Dit onderzoeksveld begint zich nu te mengen met meerdere disciplines zoals genome sequencing, bioinformatica, transcriptomics, proteomics, synthetische biologie, systeembioïologie en chemische ecologie. Synthetische biologie heeft daarbij grote potentie om drug discovery een nieuw leven in te blazen door de voorraad aan natuurstoffen aan te vullen. Daarbij kan synthetische biologie de exploitatie van cryptische BGC's en de opheldering van de structuur van het eindproduct mogelijk maken. De vooruitgang in moderne genetische gereedschappen, zoals transformatie-afhankelijke recombinatie (TAR), Red/ET recombinatie, en CRISPR-Cas9 genoom editing, zal de synthetische biologie verder kunnen versnellen en daarmee ook het onderzoek naar nieuwe antibiotica en andere natuurstoffen. In deze context is het van groot belang dat een wetenschapper die zich toelegt op de zoektocht naar nieuwe antibiotica ook de biologische kennis en praktische genetische gereedschappen tot zijn beschikking heeft. Verregaande samenwerking tussen wetenschappelijke experts van verschillende disciplines is dan ook de beste manier om de ontdekking van nieuwe antibiotica verder te brengen.

Micro-organismen leven in bijna alle ecologische niches, en hier spelen zij een doorslaggevende rol in het onderhoud van duurzame macroscopische en microscopische ecosystemen. Veel ecologische fenomenen zijn direct te herleiden tot de natuurlijke functie van natuurstoffen, zoals bij rhizoxin dat rijst zaailingziekte veroorzaakt. Het zou daarom heel goed zijn om ‘ecologisch geïnspireerde’ methodes te volgen voor de ontdekking van antibiotica, zeker met de onvoorstelbare biodiversiteit van de aarde. In die zin kan de natuur altijd als een belangrijke inspiratiebron dienen. Het is vooral belangrijk om onze aandacht te richten op de interessante natuurlijke fenomenen en na te denken over de onderliggende principes, dit te koppelen aan nog niet ontdekte microben en dan verschillende ('omics)-tools te gebruiken bij het screenen naar nieuwe biomoleculen. Zoals Auguste Rodin al zei: “het is niet een tekort aan schoonheid in het leven, maar een tekort aan ogen om de schoonheid te vinden”.

Het centrale dogma van de moleculaire biologie – de overdracht van informatie van DNA via RNA naar eiwit – is nu uitgebreid met het metabooloom. Daarbij zijn diverse 'omics technieken ontwikkeld voor iedere stap in dit bredere centrale dogma om de ontdekking van natuurstoffen mogelijk te maken (Figuur 2). Het gegeven dat alle biomoleculen een genetische basis hebben suggereert een nieuw concept wat ik “Genometabolomics” zou willen noemen. Dit verenigt de universeel beschikbare kennis van biosynthesegenen en biosynthetische enzymen en reacties wat vervolgens kan worden gebruikt om de aanwezigheid van nog onbekende metabolieten te voorspellen en te karakteriseren – en ook andersom.



Figuur 2. Genometabolomics voor onderzoek naar nieuwe biomoleculen in actinomyceten.

APPENDIX CONTENTS

Appendix I	Supplemental data belonging to Chapter 4	193
Appendix II	Supplemental data belonging to Chapter 5	205
Appendix III	Supplemental data belonging to Chapter 6	217
Appendix IV	Supplemental data belonging to Chapter 7	221
Appendix V	Supplemental data belonging to Chapter 8	229
Appendix VI	Supplemental data belonging to Chapter 9	239
Appendix VII	Supplemental data belonging to Chapter 10	269
Appendix VIII	Supplemental data belonging to Chapter 11	295

Appendix I: Supplementary Information belonging to Chapter 4



Figure S1 Spontaneously generated streptomycin-resistant colonies of *Streptomyces* sp. MBT28 (left), *Streptomyces lividans* (middle), and wild type *Streptomyces* sp. MBT28 (right). Spores suspension of wild-type were spread on SFM agar plates containing 10 µg/mL streptomycin, and incubated at 30°C for nine days.

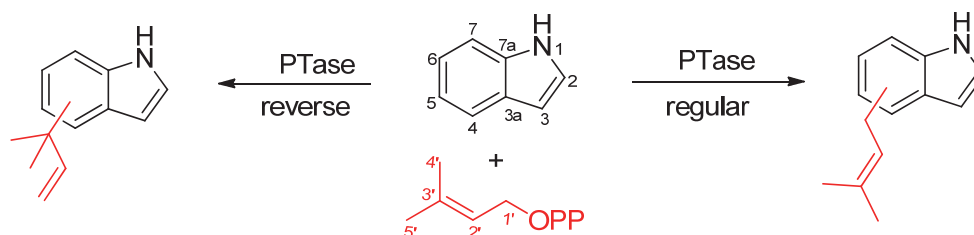


Figure S2. Schematic diagram of indole prenylation by prenyltransferases (PTase). All possible positions of the indole ring (N-1, C-2, C-3, C-4, C-5, C-6, and C-7) could be prenylated, except the two bridgeheads (C-3a, and C-7a). The prenylations may take place in a “regular” sense where the primary carbon of the allylic moiety (C-1’) adds to the indole ring, or in a “reverse” sense where the tertiary carbon (C-3’) adds.^{1,2} Some PTs responsible for prenylation of indole ring are summarized in Table S5.

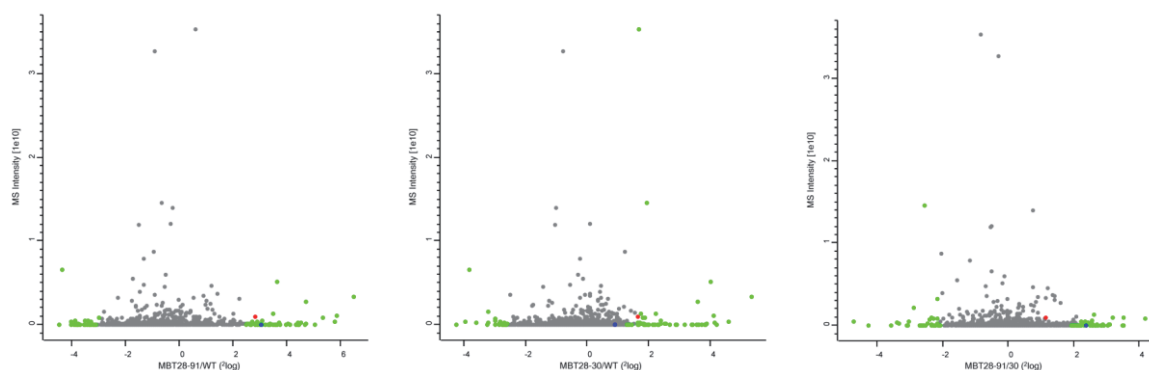


Figure S3. Volcano plot analysis of proteomics data. X axis shows the $^2\log$ ratio of protein expression, Y axis shows the MS intensity of the protein. Green dots represent proteins with the $^2\log$ expression ratio significant (p-value less than 0.05). Blue dot represent IsaA and red dot represent IsaB.

Table S1 ^1H NMR chemical shift (δ) and coupling constants (Hz) assignments for the compounds identified during the metabolomics study of *Streptomyces* sp. MBT28.

NO	Compounds	Chemical shifts
1	borrelidin	6.60 (dd, $J = 15.6, 11.4$ Hz), 6.90 (d, $J = 11.4$ Hz), and 6.31 (ddd, $J = 15.6, 10.2, 4.8$ Hz)
2	<i>p</i> -cresol	6.68 (d, $J = 8.4$ Hz); 7.07 (d, $J = 8.4$ Hz); 2.56 (s)
3	4-hydroxybenzoic acid	6.62 (d, $J = 9.0$ Hz); 7.73 (d, $J = 9.0$ Hz)
4	caffeic acid	6.27 (d, $J = 15.6$ Hz); 7.61 (d, $J = 15.6$ Hz); 6.71 (d, $J = 8.4$ Hz); 7.16 (d, $J = 2.4$ Hz); 7.09 (dd, $J = 8.4, 2.4$ Hz)
5	1 <i>H</i> -pyrrole-2-carboxylic acid	7.70 (dd, $J = 6.0, 3.0$ Hz); 7.60 (dd, $J = 6.0, 3.0$ Hz); 6.28 (t, $J = 6.0$)
6	syringic acid	7.38 (s)
7	salicylic acid	7.75 (dd, $J = 7.2, 2.4$ Hz); 7.28 (t, $J = 7.2$ Hz); 7.10 (m); 7.09 (m)
8	succinic acid	2.59 (s)
9	sucrose	5.62 (d, $J = 4.2$ Hz)
10	2,5-di- <i>tert</i> -butylphenol	7.39 (d, $J = 9.0$ Hz); 7.21 (dd, $J = 9.0, 2.4$ Hz); 7.38 (d, $J = 2.4$ Hz)
11	(3,5-di- <i>tert</i> -butyl-4-hydroxyphenyl)-propionic acid octadecyl ester	6.95 (s); 2.81 (t, $J = 7.2$ Hz); 2.64 (t, $J = 7.8$ Hz); 1.40 (s)

Table S2. Type A indole prenyltransferase gene cluster in *Streptomyces* sp. MBT28.³ This type is proposed to be involved in 7-prenylisatin biosynthesis in current study. The putative function was indicated in Figure 5 of manuscript.

orf	Protein products	Amino acid number	Putative function	Nearest (enzyme, origin)	homologue	Homology	Accession number
1		986	Transcriptional regulator	<i>Streptomyces</i> sp. WC-3795	NRRL	100%	WP_031020590.1
2	IsaA	341	aromatic prenyltransferase	<i>Streptomyces albogriseolus</i>		84%	AHD24368.1
3	IsaB	502	tryptophanase	<i>Streptomyces</i> sp. F-4835	NRRL	99%	WP_030978677.1

Table S3. Type B indole prenyltransferase gene cluster in *Streptomyces* sp. MBT28. This type like SCO7467 in *Streptomyces coelicolor* was reported to be involved in the biosynthesis of 5-dimethylallylindole-3-acetonitrile.³

orf	Amino acid number	Putative function	Nearest (enzyme, origin)	homologue	Homology	Accession number
1	276	ABC transporter ATPase	<i>Streptomyces</i> WC-3795	sp. NRRL	99%	WP_031018365.1
2	228	trypsin hydrolase	<i>Streptomyces</i> WC-3795	sp. NRRL	99%	WP_031018366.1
3	255	membrane protein	<i>Streptomyces</i> WC-3795	sp. NRRL	99%	WP_031018369.1
4	382	histidine kinase	[<i>Streptomyces</i>]		100%	WP_029392892.1
5	133	hypothetical protein	<i>Streptomyces violaceorubidus</i>		97%	WP_030183598.1
6	125	hypothetical protein	<i>Streptomyces</i> ATCC 23877	<i>ambofaciens</i>	94%	CAJ89643.1
7	191	ATP/GTP-binding protein	<i>Streptomyces</i> ATCC 23877	<i>ambofaciens</i>	99%	CAJ89642.1
8	449	cytochrome P450	<i>Streptomyces</i> WC-3795	sp. NRRL	99%	WP_031018371.1
9	373	indole prenyltransferase	<i>Streptomyces</i> WC-3795	sp. NRRL	99%	WP_031018373.1
10	428	flavin-binding monooxygenase	<i>Streptomyces</i> WC-3795	sp. NRRL	99%	WP_031018379.1

Table S4. Significantly up or down regulated (with p-value less than 0.05) quantified proteins in at least one comparison

Protein IDs	MBT28-30/WT (2log)	p-val ue	MBT28-91/WT (2log)	p-val ue	MBT28-91/30 (2log)	p-val ue	predicted function
NODE_001_gen e_001	0.2	0.604	3.1	0.020	2.8	0.008	1 4-alpha-glucan branching enzyme GlgB 2
NODE_001_gen e_012	-1.9	0.174	-3.1	0.033	-1.4	0.184	6-phosphofructokinase 2
NODE_001_gen e_037	-2.4	0.075	-3.3	0.025	-1.0	0.384	Methylmalonyl-CoA epimerase
NODE_005_gen e_007	1.1	0.077	-1.6	0.577	-2.6	0.010	Alanine--tRNA ligase
NODE_009_gen e_004	-0.3	0.994	-2.3	0.150	-2.2	0.045	Secreted protease
NODE_012_gen e_017	-1.5	0.479	-3.5	0.021	-2.3	0.029	ATP-dependent Clp protease proteolytic subunit 1
NODE_016_gen e_001	3.6	0.000	4.3	0.002	0.6	0.576	Secreted protein
NODE_016_gen e_019	-1.2	0.658	-3.8	0.010	-2.5	0.014	Ribose-phosphate pyrophosphokinase
NODE_020_gen e_023	0.2	0.581	2.8	0.030	2.4	0.026	Secreted protein
NODE_020_gen e_034	0.8	0.263	2.8	0.035	2.0	0.064	Flp pilus assembly protein CpaB
NODE_021_gen e_028	-2.6	0.068	-3.9	0.007	-1.2	0.272	Cytochrome oxidase subunit I
NODE_021_gen e_047	2.4	0.004	3.4	0.012	0.9	0.388	Secreted protein
NODE_021_gen e_048	2.5	0.003	4.7	0.001	2.7	0.012	Nuclear export factor GLE1
NODE_025_gen e_009	-2.9	0.039	-1.9	0.400	1.0	0.268	Putative PadR-like family transcriptional regulator
NODE_025_gen e_015	-1.0	0.532	-3.2	0.029	-2.4	0.029	Ectoine/hydroxyectoine ABC transporter permease EhuC
NODE_025_gen e_017	1.3	0.097	3.7	0.006	2.6	0.017	Methyltransferase
NODE_025_gen e_022	1.7	0.037	0.6	0.467	-0.9	0.439	Nucleotide-binding protein
NODE_025_gen e_023	1.7	0.033	4.5	0.001	2.5	0.021	Peptidoglycan-binding domain 1 protein
NODE_026_gen e_007	1.9	0.009	3.7	0.007	1.7	0.072	Secreted protein
NODE_027_gen e_001	2.3	0.003	3.4	0.013	1.6	0.088	Conserved hypothetical secreted protein
NODE_029_gen e_014	3.2	0.000	2.5	0.051	-0.9	0.411	Secreted protein
NODE_032_gen e_004	-1.9	0.174	-3.3	0.025	-1.2	0.254	Membrane protein
NODE_032_gen e_009	0.9	0.205	-2.5	0.117	-3.1	0.004	Starvation-induced DNA protecting protein
NODE_033_gen e_057	-2.6	0.048	-1.9	0.266	0.7	0.522	Cysteine--tRNA ligase
NODE_035_gen e_011	-2.5	0.062	-3.4	0.018	-1.0	0.368	Glycerol-3-phosphate dehydrogenase
NODE_035_gen e_012	-2.4	0.073	0.5	0.507	3.2	0.003	Glycerol kinase 1
NODE_035_gen e_015	-0.3	0.707	3.0	0.024	2.6	0.007	Methionine synthase
NODE_035_gen e_017	3.4	0.000	3.0	0.025	-0.7	0.532	Lipoprotein oligopeptide binding protein
NODE_036_gen e_009	0.1	0.429	2.2	0.073	2.1	0.030	Oxidoreductase
NODE_042_gen e_025	-1.6	0.262	-3.6	0.013	-1.8	0.089	Nucleotide sugar-1-phosphate transferase
NODE_047_gen e_002	-3.8	0.003	-4.4	0.002	-0.5	0.644	Alkylhydroperoxidase like protein AhpD family
NODE_047_gen e_007	-3.0	0.022	-2.9	0.056	-0.1	0.958	Alkylhydroperoxidase like protein AhpD family
NODE_048_gen e_013	0.9	0.111	3.5	0.011	2.2	0.028	Carboxypeptidase
NODE_048_gen e_015	3.9	0.000	3.0	0.025	-0.8	0.493	Metallopeptidase
NODE_051_gen e_011	-2.1	0.133	0.2	0.628	2.2	0.039	Monooxygenase
NODE_057_gen e_030	-3.3	0.015	-2.5	0.169	0.5	0.585	Putative oxidoreductase
NODE_057_gen e_031	-2.7	0.040	-1.6	0.388	1.7	0.105	Putative 6-phospho-3-hexuloisomerase
NODE_057_gen e_032	-2.8	0.031	-0.4	0.923	2.3	0.029	Putative triosephosphate isomerase
NODE_058_gen e_031	-3.6	0.005	-1.5	0.410	0.8	0.465	Alkylhydroperoxidase domain-containing protein AhpD family core
NODE_062_gen e_005	0.0	0.471	-2.4	0.207	-2.3	0.026	DNA-binding protein
NODE_063_gen e_013	-3.2	0.013	-2.8	0.064	0.9	0.393	Glutamine synthetase 2
NODE_063_gen e_031	2.2	0.007	2.0	0.102	-0.4	0.758	Maltose-binding protein
NODE_064_gen e_033	2.5	0.002	1.5	0.202	-1.2	0.286	Secreted protein

Appendix I: Supplementary Information belong to Chapter 4

NODE_065_gen e_019	-2.7	0.040	-2.7	0.078	0.3	0.745	Carboxylesterase
NODE_065_gen e_022	1.0	0.163	4.0	0.003	3.1	0.004	Cold shock domain-containing protein CspD
NODE_066_gen e_007	-4.0	0.002	-4.0	0.005	-0.2	0.872	Secreted protein
NODE_069_gen e_017	4.2	0.000	2.9	0.027	-1.1	0.323	Sugar transporter sugar binding protein
NODE_070_gen e_012	2.9	0.000	4.1	0.004	1.0	0.271	Oxidoreductase
NODE_071_gen e_024	0.2	0.347	-2.4	0.197	-2.7	0.009	Integral membrane protein
NODE_072_gen e_025	1.1	0.134	-0.8	0.798	-2.2	0.043	Succinate-semialdehyde dehydrogenase
NODE_075_gen e_003	-0.1	0.555	2.0	0.092	2.7	0.007	Putative methyltransferase
NODE_078_gen e_010	-2.6	0.049	-3.8	0.008	-1.0	0.376	Putative 3-oxoacyl-ACP synthase II
NODE_079_gen e_007	1.0	0.177	-1.8	0.318	-2.3	0.035	Chloride peroxidase
NODE_079_gen e_008	0.7	0.290	-1.5	0.444	-2.4	0.029	Alpha/beta hydrolase
NODE_082_gen e_003	2.2	0.004	0.5	0.380	-1.3	0.211	Uncharacterized protein
NODE_089_gen e_031	-0.5	0.822	2.7	0.038	3.0	0.002	SPFH domain/Band 7 family protein
NODE_090_gen e_004	1.8	0.024	0.9	0.369	-1.1	0.333	Integral membrane protein
NODE_091_gen e_005	1.2	0.057	-2.2	0.257	-3.6	0.000	Hypothetical cytosolic protein
NODE_092_gen e_010	-0.1	0.526	1.9	0.095	1.9	0.046	Redoxin
NODE_092_gen e_011	-1.3	0.587	-3.4	0.024	-2.1	0.042	3-phosphoserine phosphatase
NODE_099_gen e_033	-0.8	0.665	2.3	0.066	3.1	0.004	NDP-4-keto-6-deoxy-L-hexose 2 3-reductase
NODE_102_gen e_017	2.3	0.003	2.1	0.076	-0.1	0.956	ABC-type cobalt transport system
NODE_110_gen e_015	-0.4	0.918	2.1	0.096	2.2	0.039	Pseudouridine-5'-phosphate glycosidase
NODE_110_gen e_016	0.4	0.453	2.0	0.107	2.3	0.032	Dioxygenase
NODE_118_gen e_013	1.9	0.011	4.0	0.005	2.0	0.041	Predicted protein
NODE_119_gen e_005	0.6	0.347	2.6	0.045	1.8	0.085	Putative secreted extracellular small neutral protease
NODE_126_gen e_032	1.0	0.179	-1.4	0.492	-2.3	0.030	Putative calcium binding protein
NODE_131_gen e_004	1.5	0.056	2.5	0.050	1.0	0.360	lipoprotein
NODE_131_gen e_035	-1.9	0.260	-3.2	0.044	-1.2	0.265	Uncharacterized protein
NODE_136_gen e_010	-0.9	0.880	2.5	0.047	3.5	0.000	AdpA
NODE_136_gen e_016	0.1	0.397	2.1	0.076	1.9	0.045	B-N-acetylhexosaminidase
NODE_137_gen e_015	-2.6	0.048	-1.9	0.258	0.9	0.392	Flavoprotein reductase
NODE_139_gen e_026	-1.1	0.726	0.8	0.309	2.5	0.011	Putative uncharacterized protein
NODE_143_gen e_007	-1.4	0.322	-3.3	0.025	-1.5	0.176	Succinyl-CoA ligase [ADP-forming] subunit beta-1
NODE_143_gen e_008	-4.3	0.001	-3.4	0.028	0.5	0.587	Succinyl-CoA ligase [ADP-forming] subunit alpha
NODE_144_gen e_030	2.9	0.000	3.6	0.009	0.5	0.541	Iron transport lipoprotein
NODE_146_gen e_013	0.3	0.329	-2.9	0.076	-2.6	0.011	Uncharacterized protein
NODE_146_gen e_033	1.6	0.041	-1.2	0.590	-2.6	0.014	Dehydrogenase
NODE_152_gen e_033	2.1	0.005	4.4	0.002	2.2	0.023	Amino acid transport integral membrane protein
NODE_153_gen e_002	2.7	0.001	3.3	0.014	0.4	0.595	Solute-binding lipoprotein
NODE_154_gen e_010	2.1	0.005	3.3	0.015	1.3	0.165	Putative secreted protein
NODE_154_gen e_060	NaN	NaN	2.8	0.030	2.8	0.004	Aminotransferase
NODE_154_gen e_064	-0.3	0.710	1.9	0.104	2.3	0.018	Oxidoreductase
NODE_155_gen e_014	4.6	0.000	5.8	0.000	1.2	0.244	Substrate binding protein
NODE_155_gen e_032	1.0	0.161	-1.6	0.372	-2.7	0.011	Mesaconyl-CoA hydratase
NODE_155_gen e_040	0.8	0.138	-1.7	0.530	-2.4	0.022	Secreted peptidase
NODE_155_gen e_052	-1.2	0.627	0.8	0.299	2.0	0.041	Uncharacterized protein
NODE_159_gen e_015	3.8	0.000	4.5	0.002	0.6	0.496	Putative membrane protein
NODE_168_gen e_002	0.9	0.190	-1.3	0.538	-2.2	0.045	6-pyruvoyl tetrahydropterin synthase

Appendix I: Supplementary Information belong to Chapter 4

NODE_171_gen e_005	-1.9	0.162	-4.0	0.004	-2.4	0.029	Sugar phosphotransferase
NODE_173_gen e_018	1.8	0.022	0.8	0.408	-1.1	0.312	Peptidoglycan-binding domain 1 protein
NODE_175_gen e_006	1.0	0.180	-1.5	0.432	-2.3	0.032	Putative lysozyme
NODE_180_gen e_024	0.9	0.116	3.0	0.023	2.4	0.016	Aromatic prenyltransferase DMATS type
NODE_180_gen e_025	1.6	0.040	2.8	0.031	1.1	0.286	Aromatic amino acid beta-eliminating lyase/threonine aldolase
NODE_188_gen e_004	2.4	0.004	-2.4	0.137	-4.7	0.000	Iron sulfur protein (Secreted protein)
NODE_191_gen e_023	0.1	0.680	2.0	0.103	2.5	0.018	3-oxoacyl-[acyl-carrier-protein] synthase 2
NODE_196_gen e_002	0.6	0.330	3.5	0.010	2.6	0.015	Secreted protein
NODE_198_gen e_010	-2.9	0.039	-2.7	0.129	0.5	0.587	Aminoacylase
NODE_207_gen e_013	0.9	0.103	3.5	0.011	2.4	0.015	Neutral zinc metalloprotease
NODE_207_gen e_015	-1.7	0.226	1.5	0.191	3.5	0.001	Phosphorylase
NODE_207_gen e_016	0.6	0.323	3.1	0.020	2.3	0.030	Alpha-amylase
NODE_211_gen e_006	3.9	0.000	2.2	0.072	-1.9	0.070	ABC transport system integral membrane protein BldKC
NODE_212_gen e_021	-0.1	0.897	-4.5	0.001	-4.3	0.000	Putative signal transduction protein with EFhand domain
NODE_226_gen e_007	-1.5	0.434	-3.7	0.012	-2.2	0.030	Exporter
NODE_227_gen e_010	1.8	0.028	1.5	0.186	-0.2	0.869	Proteinase
NODE_230_gen e_006	-3.0	0.028	-2.2	0.258	0.4	0.637	GCN5-related N-acetyltransferase
NODE_231_gen e_019	-2.9	0.030	-3.9	0.006	-0.9	0.393	Peptide transporter
NODE_231_gen e_043	-2.7	0.044	-3.1	0.035	-0.7	0.523	FAD-dependent pyridine nucleotide-disulfide oxidoreductase
NODE_244_gen e_025	-1.9	0.260	-3.3	0.033	-1.2	0.266	Phosphodiesterase
NODE_245_gen e_011	-2.2	0.143	-3.2	0.044	-0.9	0.391	Mycothioli conjugate amidase Mca
NODE_248_gen e_003	0.9	0.209	-1.5	0.441	-2.5	0.021	Peptide transport system secreted peptide-binding protein
NODE_249_gen e_004	-0.3	1.000	-3.5	0.015	-3.0	0.005	Urease subunit alpha 1
NODE_250_gen e_002	-1.3	0.558	1.3	0.188	2.6	0.009	Putative transmembrane transport protein
NODE_250_gen e_012	1.4	0.034	2.1	0.076	0.7	0.429	Cationic amino acid transporter
NODE_250_gen e_016	-2.2	0.166	-3.2	0.045	-1.1	0.312	Thioredoxin
NODE_251_gen e_002	-1.4	0.503	0.8	0.312	2.1	0.033	Sigma 54 modulation protein/SSU ribosomal protein S30P
NODE_259_gen e_010	1.7	0.035	1.3	0.242	-0.5	0.659	Subtilisin-like protease
NODE_267_gen e_023	3.5	0.000	3.6	0.009	0.1	0.872	Secreted alkaline phosphatase
NODE_271_gen e_008	-3.3	0.012	-2.1	0.208	0.1	0.886	UDP-glucose 4-epimerase
NODE_272_gen e_009	-0.5	0.854	1.3	0.197	2.0	0.043	WD40 repeat-containing protein
NODE_279_gen e_009	1.2	0.113	-1.6	0.378	-2.9	0.007	Putative lipoprotein
NODE_279_gen e_082	1.9	0.017	-0.6	0.919	-2.6	0.017	Glyceraldehyde 3-phosphate dehydrogenase
NODE_279_gen e_107	0.5	0.236	3.3	0.015	2.7	0.005	Putative cysteine desulphurases SufS
NODE_281_gen e_006	-0.1	0.825	-2.6	0.094	-2.6	0.014	3-hydroxyacyl-CoA dehydrogenase PaaC
NODE_282_gen e_002	-2.7	0.041	-2.6	0.096	0.1	0.881	Phosphocarrier protein HPr
NODE_286_gen e_026	0.7	0.302	-1.7	0.322	-2.4	0.025	Catalase
NODE_288_gen e_011	0.7	0.169	-1.6	0.580	-2.1	0.045	Uncharacterized protein
NODE_299_gen e_002	-2.8	0.049	-1.4	0.744	1.5	0.127	Bifunctional uroporphyrinogen-III synthetase/response regulator domain protein
NODE_299_gen e_014	1.9	0.011	2.2	0.073	0.4	0.616	Sugar transporter sugar-binding protein
NODE_300_gen e_002	-1.3	0.389	1.0	0.333	2.3	0.031	Adenylosuccinate synthetase
NODE_309_gen e_008	-2.1	0.177	-3.1	0.049	-0.7	0.545	Orotate phosphoribosyltransferase
NODE_316_gen e_021	1.3	0.050	-0.6	0.821	-1.4	0.172	Secreted hydrolase
NODE_320_gen e_035	0.0	0.502	2.1	0.075	2.2	0.025	3-ketosteroid-delta-1-dehydrogenase (Fragment)
NODE_321_gen e_026	-3.2	0.013	-2.7	0.081	0.5	0.647	Phosphoribosylaminoimidazole-succinocarboxamide synthase
NODE_330_gen e_015	-3.0	0.022	-3.4	0.019	-0.3	0.790	Peptidase M48

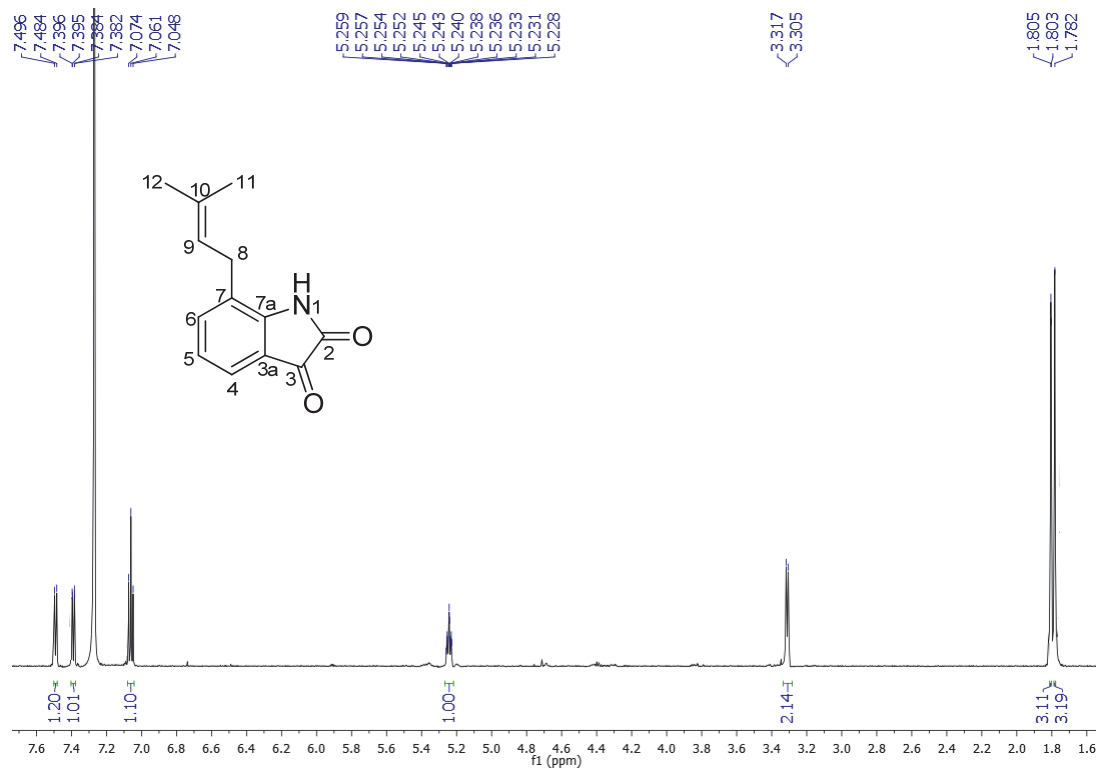
Appendix I: Supplementary Information belong to Chapter 4

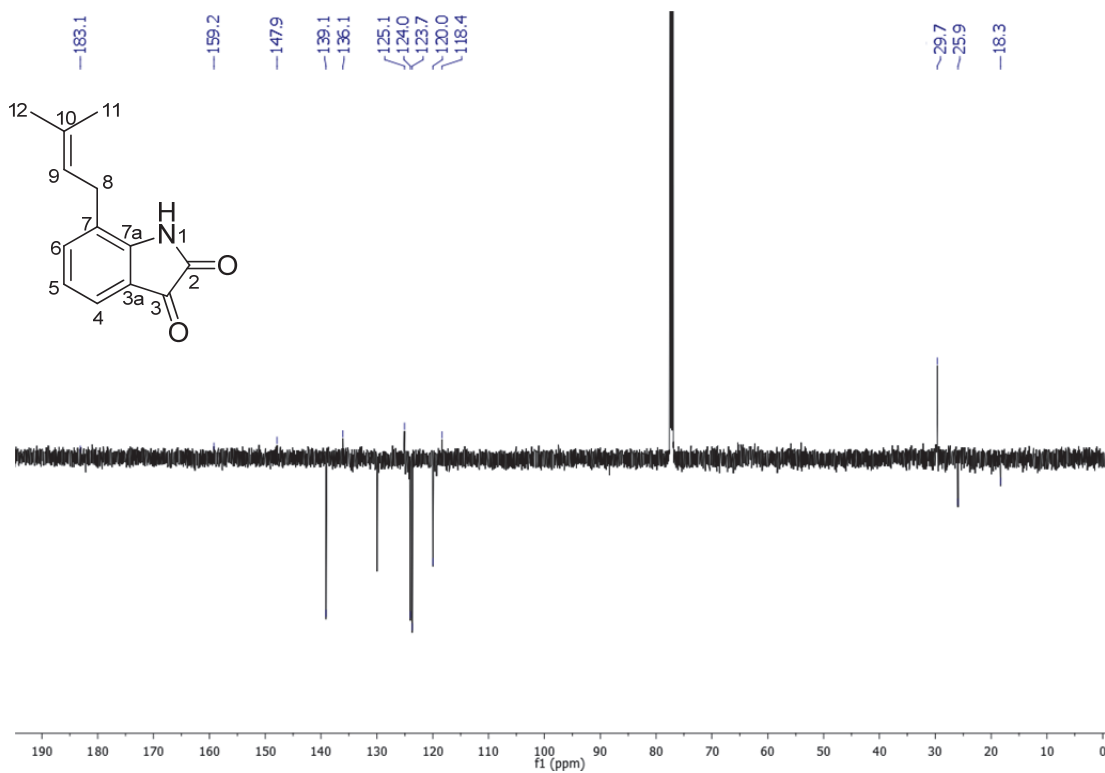
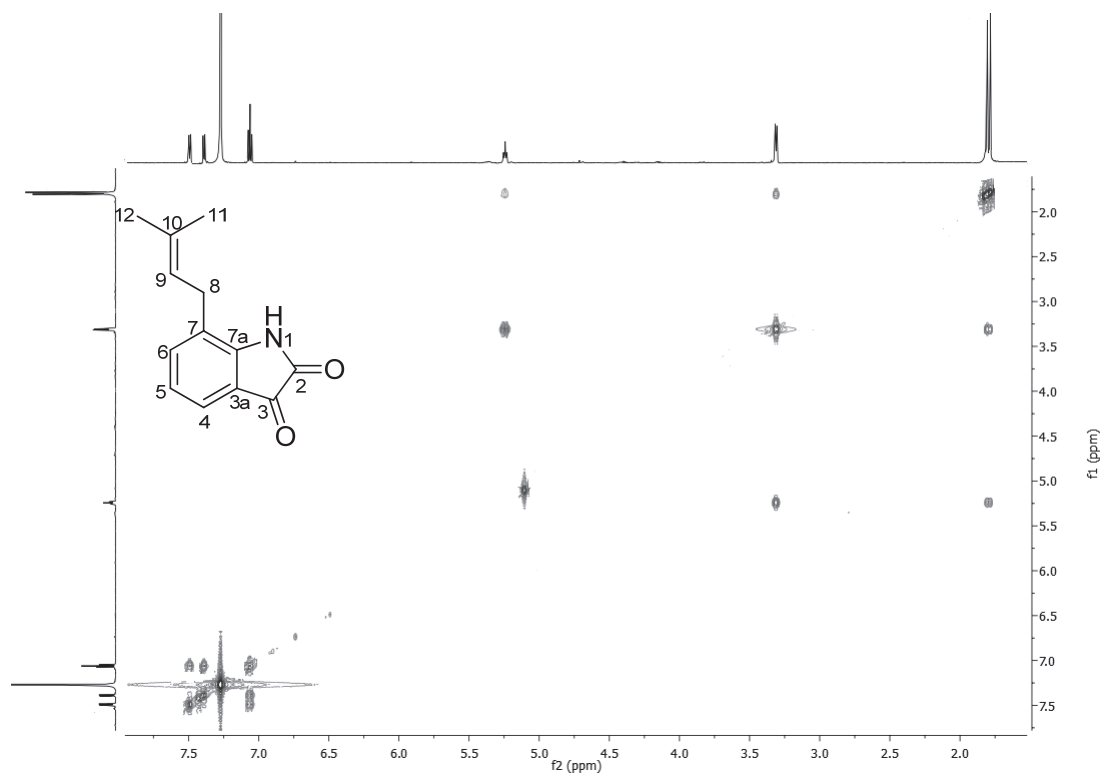
NODE_337_gen e_001	-2.5	0.093	-3.8	0.009	-0.3	0.830	Putative oxygenase
NODE_340_gen e_005	-0.8	0.937	0.9	0.272	1.9	0.049	Pyruvate phosphate dikinase
NODE_346_gen e_006	-0.6	0.899	1.1	0.238	2.0	0.038	Translation initiation factor IF-3
NODE_347_gen e_004	4.1	0.000	4.4	0.001	0.5	0.642	Lipoprotein
NODE_359_gen e_003	-0.9	0.576	-3.0	0.044	-2.0	0.066	Putative secreted protein
NODE_363_gen e_001	1.8	0.025	5.3	0.000	4.2	0.000	Putative iron transport lipoprotein
NODE_366_gen e_016	4.0	0.000	3.7	0.007	-0.2	0.882	Glutamate binding protein
NODE_366_gen e_028	-0.4	0.744	-3.1	0.056	-2.5	0.013	Predicted protein
NODE_366_gen e_037	3.4	0.000	3.6	0.010	0.3	0.714	Secreted protein
NODE_395_gen e_033	-1.8	0.311	-3.4	0.025	-1.6	0.131	HesB/YadR/YfhF family protein
NODE_395_gen e_039	-2.2	0.098	-3.8	0.008	-1.6	0.145	Uncharacterized protein
NODE_399_gen e_018	0.8	0.133	-1.7	0.554	-2.1	0.038	Uncharacterized protein
NODE_410_gen e_001	1.6	0.041	2.8	0.031	1.1	0.282	Secreted protein
NODE_416_gen e_006	1.7	0.030	1.3	0.243	-0.3	0.822	Germacradienol/geosmin synthase
NODE_431_gen e_010	0.3	0.545	2.8	0.030	2.3	0.032	Uncharacterized protein
NODE_434_gen e_008	-2.0	0.210	-3.9	0.008	-1.2	0.244	Putative polyketide cyclase
NODE_451_gen e_001	-1.1	0.474	0.9	0.361	2.4	0.026	NB-ARC domain-containing protein
NODE_453_gen e_003	-2.7	0.038	-2.8	0.060	0.4	0.705	Methylmalonyl-CoA mutase
NODE_461_gen e_007	1.7	0.038	-1.6	0.397	-3.4	0.002	Protease
NODE_463_gen e_011	2.0	0.008	3.0	0.025	1.3	0.173	ABC-type Fe3+-siderophore transporter ATP-binding protein
NODE_463_gen e_013	4.1	0.000	5.9	0.000	2.1	0.054	ABC-type Fe3+-siderophore transporter substrate-binding protein
NODE_464_gen e_004	0.4	0.282	2.6	0.039	2.1	0.032	Cobalt transport integral membrane protein
NODE_464_gen e_005	1.8	0.024	2.5	0.055	1.2	0.248	Putative ABC transporter ATP-binding protein SCO5958
NODE_487_gen e_009	1.4	0.042	4.7	0.001	3.5	0.000	Muramoyl-pentapeptide carboxypeptidase
NODE_487_gen e_101	3.6	0.000	4.7	0.001	1.6	0.135	Lipoprotein
NODE_487_gen e_102	3.7	0.000	5.1	0.001	2.0	0.041	Solute-binding lipoprotein
NODE_487_gen e_103	5.3	0.000	6.5	0.000	1.1	0.282	Secreted protein
NODE_491_gen e_002	1.8	0.013	1.4	0.168	-0.3	0.854	Secreted alkaline phosphatase
NODE_510_gen e_026	-0.3	0.664	-2.1	0.330	-2.1	0.041	Membrane protein
NODE_514_gen e_002	-2.2	0.144	-3.8	0.010	-1.6	0.127	Secreted penicillin binding protein
NODE_526_gen e_007	-2.8	0.048	-3.3	0.032	-0.3	0.865	Lipoprotein
NODE_528_gen e_003	0.8	0.135	3.2	0.018	2.5	0.010	Carrier protein membrane protein (Fragment)
NODE_532_gen e_001	1.3	0.096	2.9	0.028	1.9	0.074	Alpha-1 4-glucan:maltose-1-phosphate maltosyltransferase 1
NODE_534_gen e_001	-1.7	0.342	1.3	0.196	2.9	0.003	Alpha-1 4-glucan:maltose-1-phosphate maltosyltransferase 1
NODE_585_gen e_004	-0.9	0.607	1.6	0.172	2.4	0.023	DUF364 domain-containing protein
NODE_633_gen e_002	0.7	0.284	2.7	0.037	1.7	0.116	Nitrogen regulatory protein P-II
NODE_644_gen e_008	2.0	0.008	-1.4	0.749	-3.1	0.002	hydrolase
NODE_659_gen e_009	-0.6	0.763	2.2	0.080	2.9	0.008	Amidase
NODE_679_gen e_001	-1.9	0.169	-3.9	0.006	-2.1	0.052	Phage tail sheath protein
NODE_746_gen e_022	0.9	0.202	-2.5	0.113	-3.4	0.002	Lipoprotein
NODE_746_gen e_044	0.8	0.263	-1.3	0.543	-2.1	0.050	Putative glyoxalase/bleomycin resistance family protein
NODE_746_gen e_047	0.5	0.230	-2.7	0.119	-3.1	0.003	Phytanoyl-CoA dioxygenase

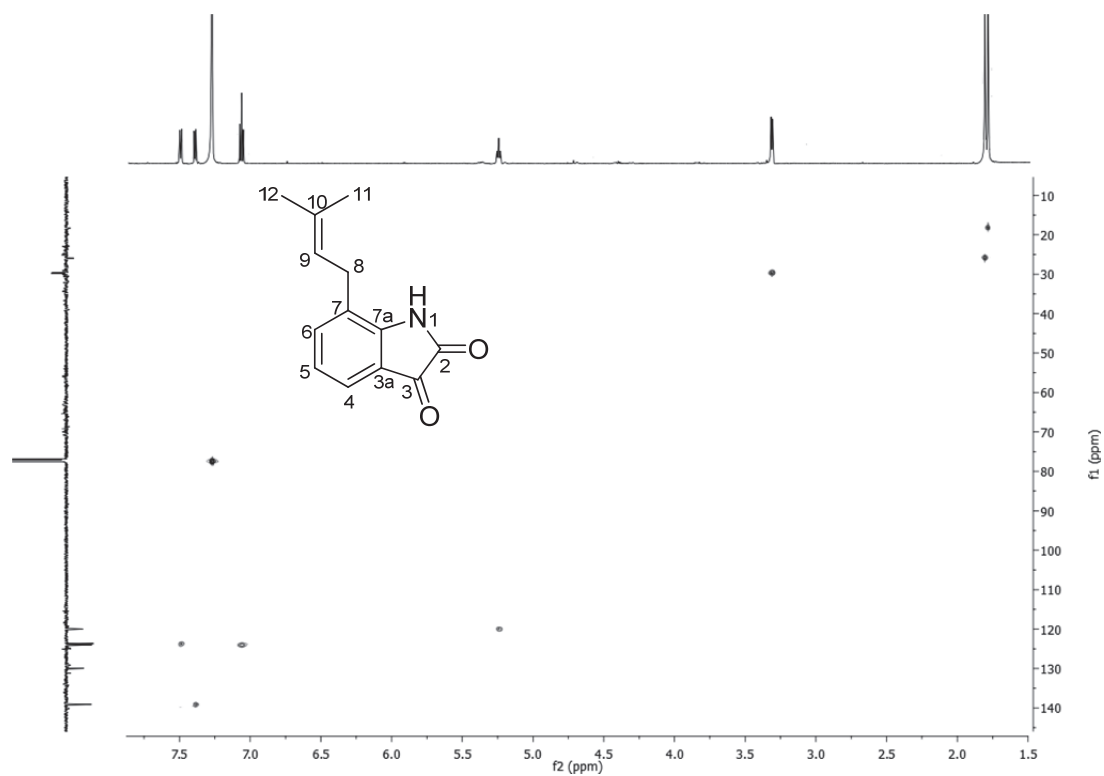
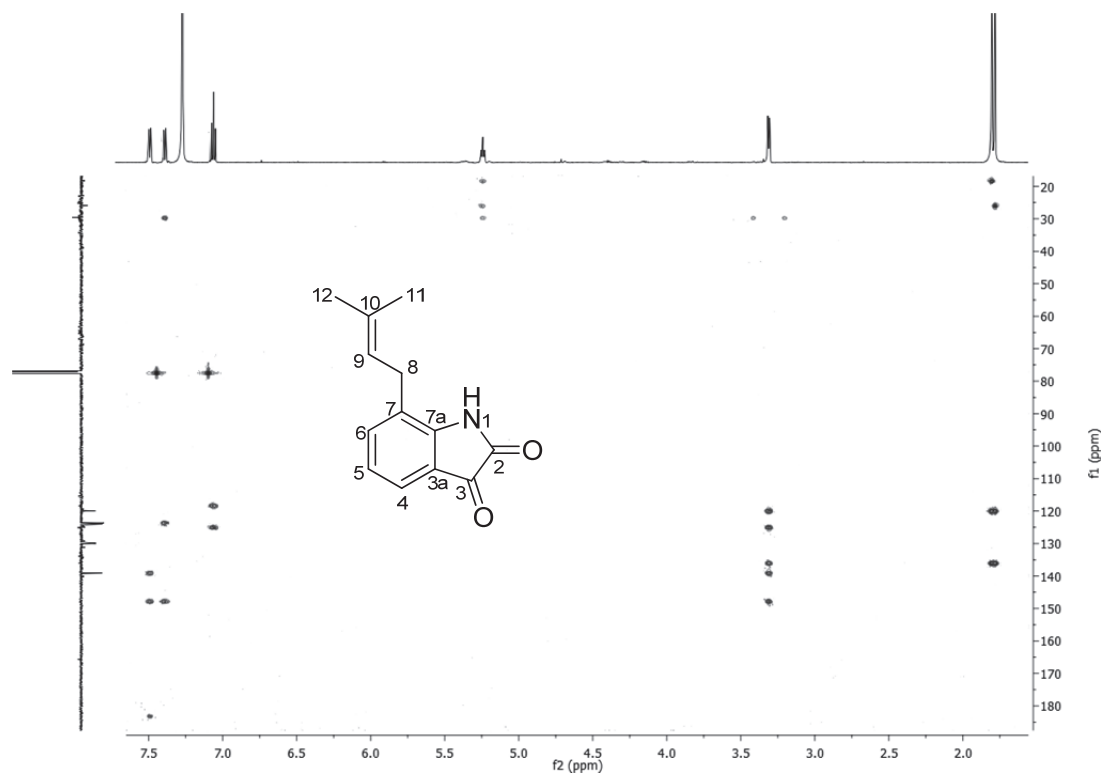
Supporting data file. Spectra of 7-prenylisatin (**1**), including NMR, MS, and UV.

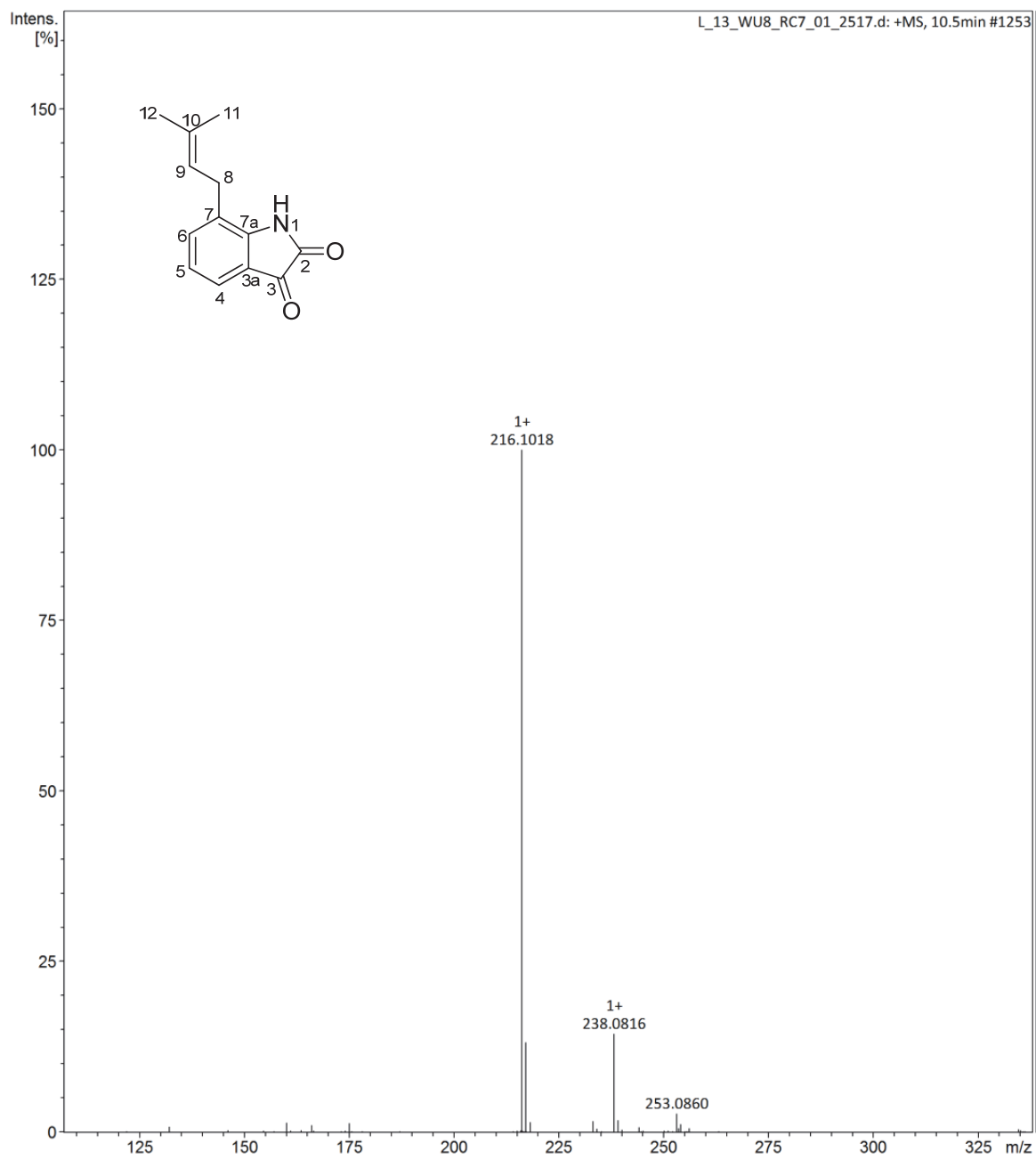
- S1. ^1H NMR spectrum of 7-prenylisatin (**1**) in CDCl_3 .
- S2. APT spectrum of 7-prenylisatin (**1**) in CDCl_3 .
- S3. ^1H - ^1H COSY spectrum of 7-prenylisatin (**1**) in CDCl_3 .
- S4. HSQC spectrum of 7-prenylisatin (**1**) in CDCl_3 .
- S5. HMBC spectrum of 7-prenylisatin (**1**) in CDCl_3 .
- S6. HRESIMS of 7-prenylisatin (**1**).
- S7. UV spectrum of 7-prenylisatin (**1**).

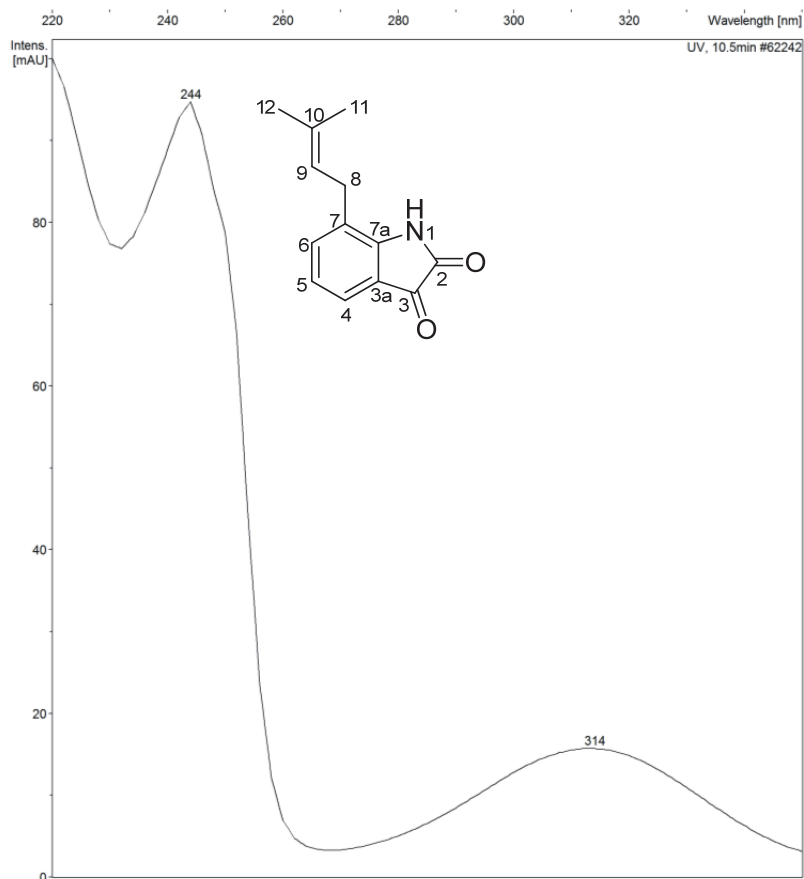
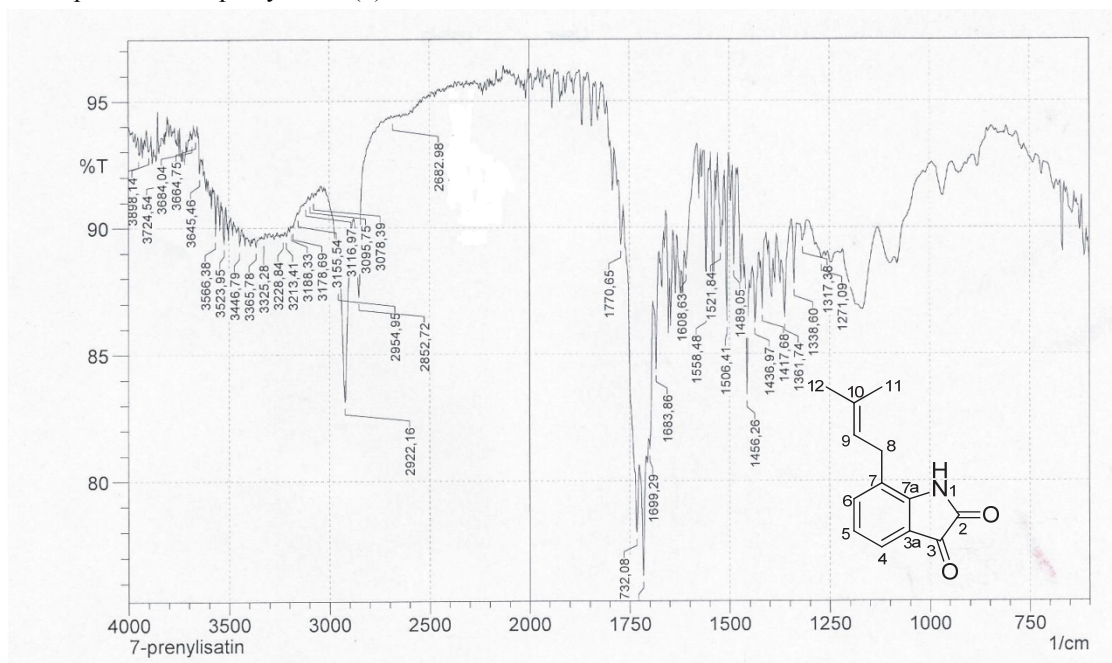
S1. ^1H NMR spectrum (600 MHz) of 7-prenylisatin (**1**) in CDCl_3 .



S2. APT spectrum (600 MHz) of 7-prenylisatin (**1**) in CDCl₃.S3. ¹H-¹H COSY spectrum (600 MHz) of 7-prenylisatin (**1**) in CDCl₃.

S4. HSQC spectrum of 7-prenylisatin (**1**) in CDCl₃.S5. HMBC spectrum of 7-prenylisatin (**1**) in CDCl₃.

S6. HRESIMS of 7-prenylisatin (**1**).

S7. UV spectrum of 7-prenylisatin (**1**).S8. IR spectrum of 7-prenylisatin (**1**).

References

- (1) Tanner, M. E. *Nat. Prod. Rep.* **2015**, 32, 88–101.
- (2) Walsh, C. T. *ACS Chem. Biol.* **2014**, 9, 2718–2728.
- (3) Ozaki, T.; Nishiyama, M.; Kuzuyama, T. *J. Biol. Chem.* **2013**, 288, 9946–56.

Appendix II: Supplementary Information belonging to Chapter 5

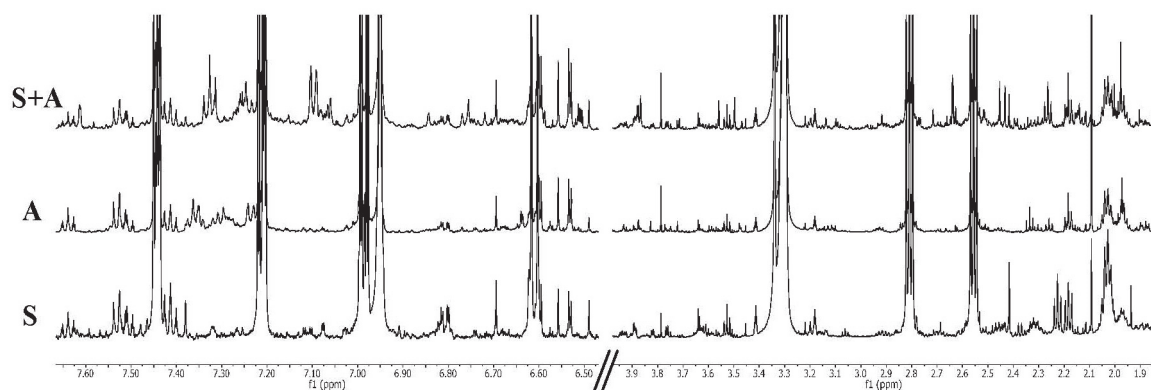


Figure S1. ¹H NMR spectra in the region of δ 6.50—7.60, and 1.90—3.90 of *Streptomyces coelicolor* single culture (S), *Aspergillus niger* single culture (A), and coculture of *Streptomyces coelicolor* with *Aspergillus niger* (S+A). Major discriminators for PCA separation (Figure 2) are summarized in Table 1.

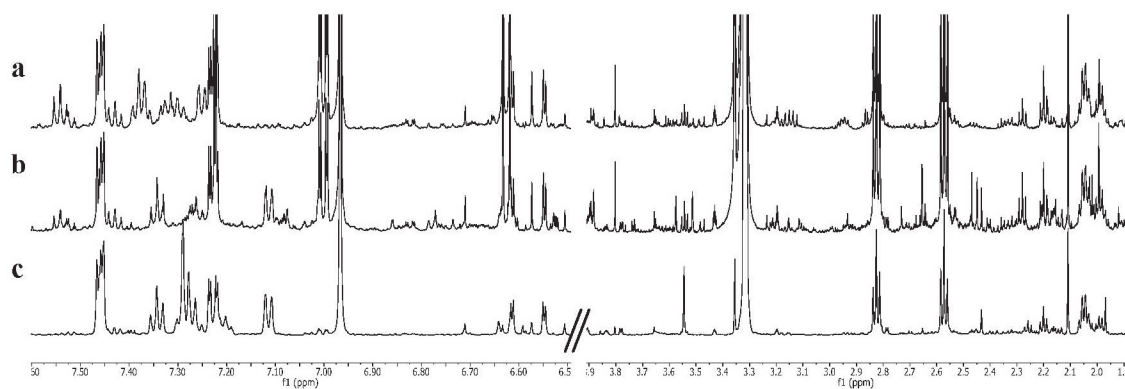


Figure S2. ¹H NMR spectra in the region of δ 6.50—7.60 and 1.90—3.90 of *Aspergillus niger* single culture (a), *Aspergillus niger* cocultured with *Streptomyces coelicolor* (b), and *Aspergillus niger* cultured in a cell-free extract of *Streptomyces coelicolor* (c). Cell-free extract of *S. coelicolor* was sufficient for eliciting *Aspergillus niger* to produce cyclo-(Phe-Phe) and phenylacetic acid

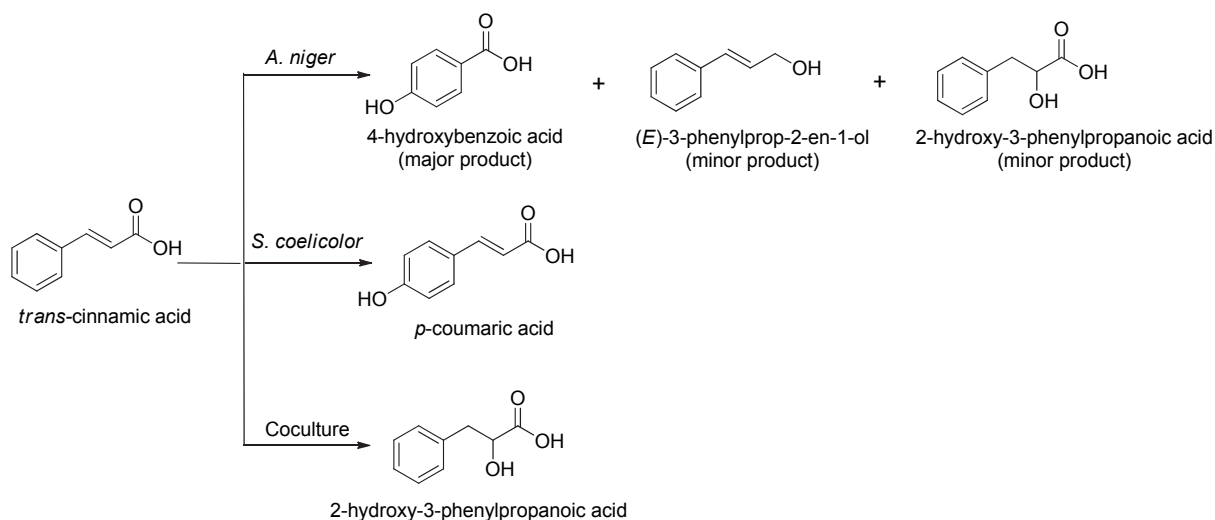


Figure S3. Biotransformation products of *trans*-cinnamic acid by *S. coelicolor* monoculture, *A. niger* monoculture and their coculture.

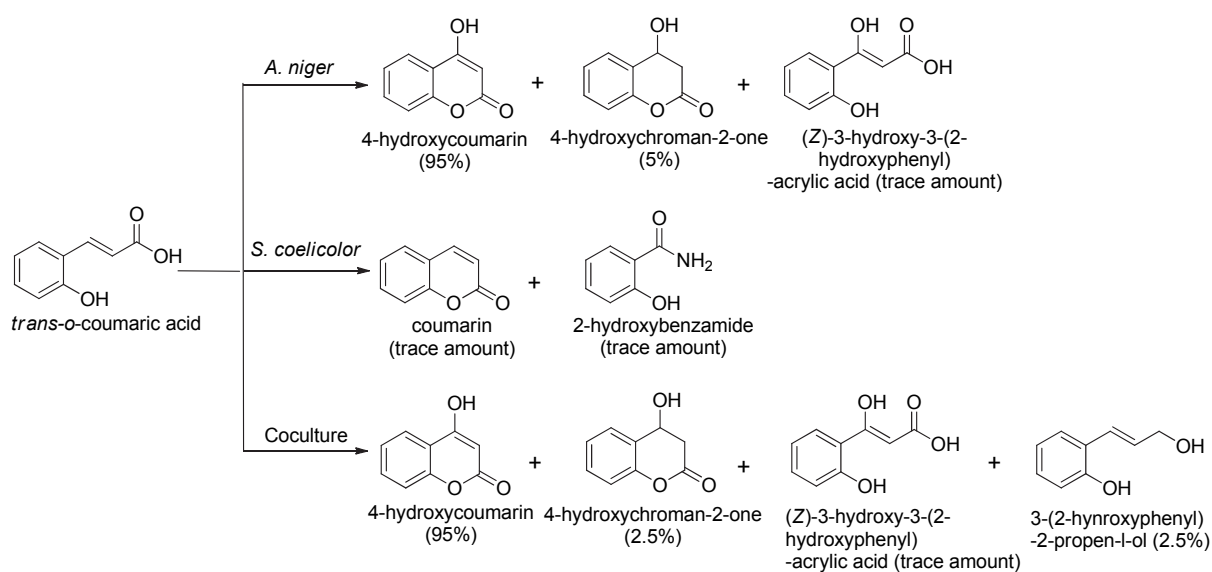


Figure S4. Biotransformation products of *trans*-*o*-coumaric acid by *S. coelicolor* monoculture, *A. niger* monoculture and their coculture.

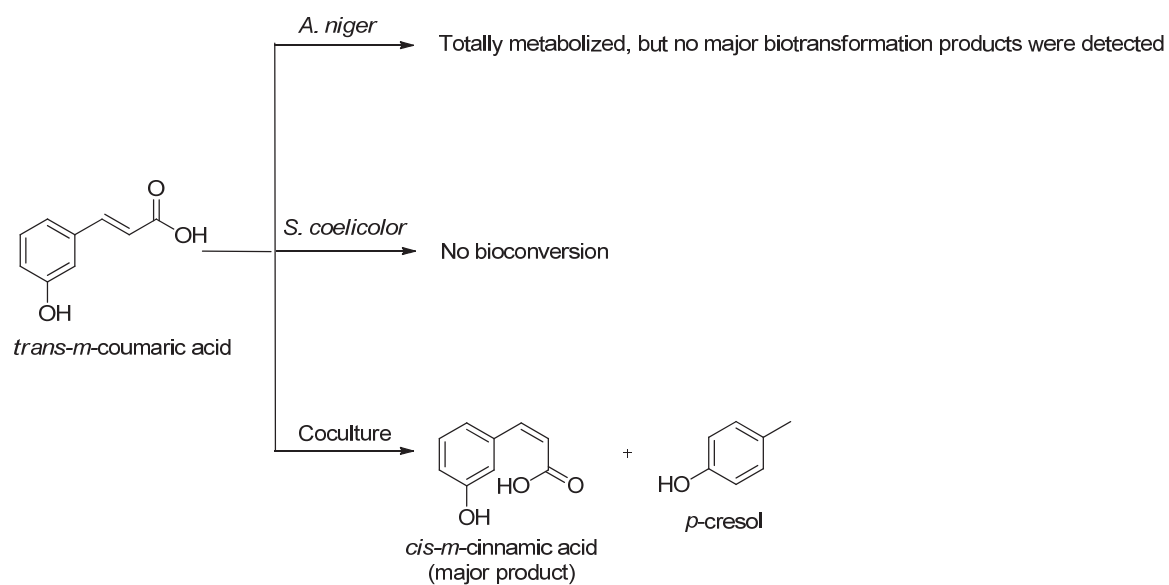


Figure S5. Biotransformation products of *trans-m-coumaric acid* by *S. coelicolor* monoculture, *A. niger* monoculture and their coculture.

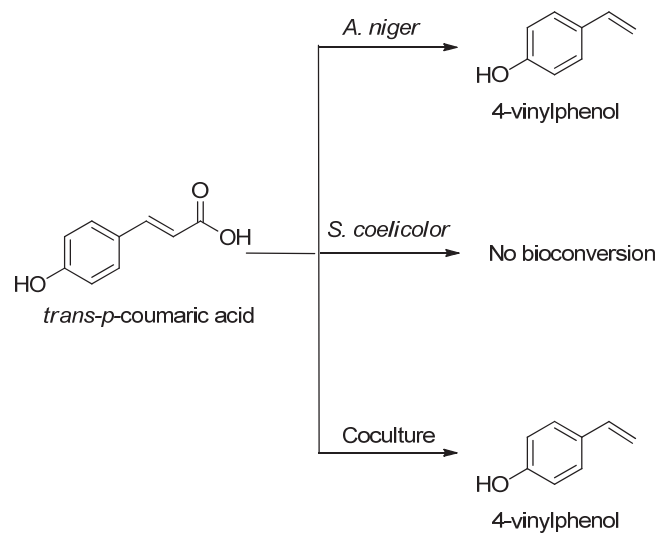


Figure S6. Biotransformation products of *trans-p-coumaric acid* by *S. coelicolor* monoculture, *A. niger* monoculture and their coculture.

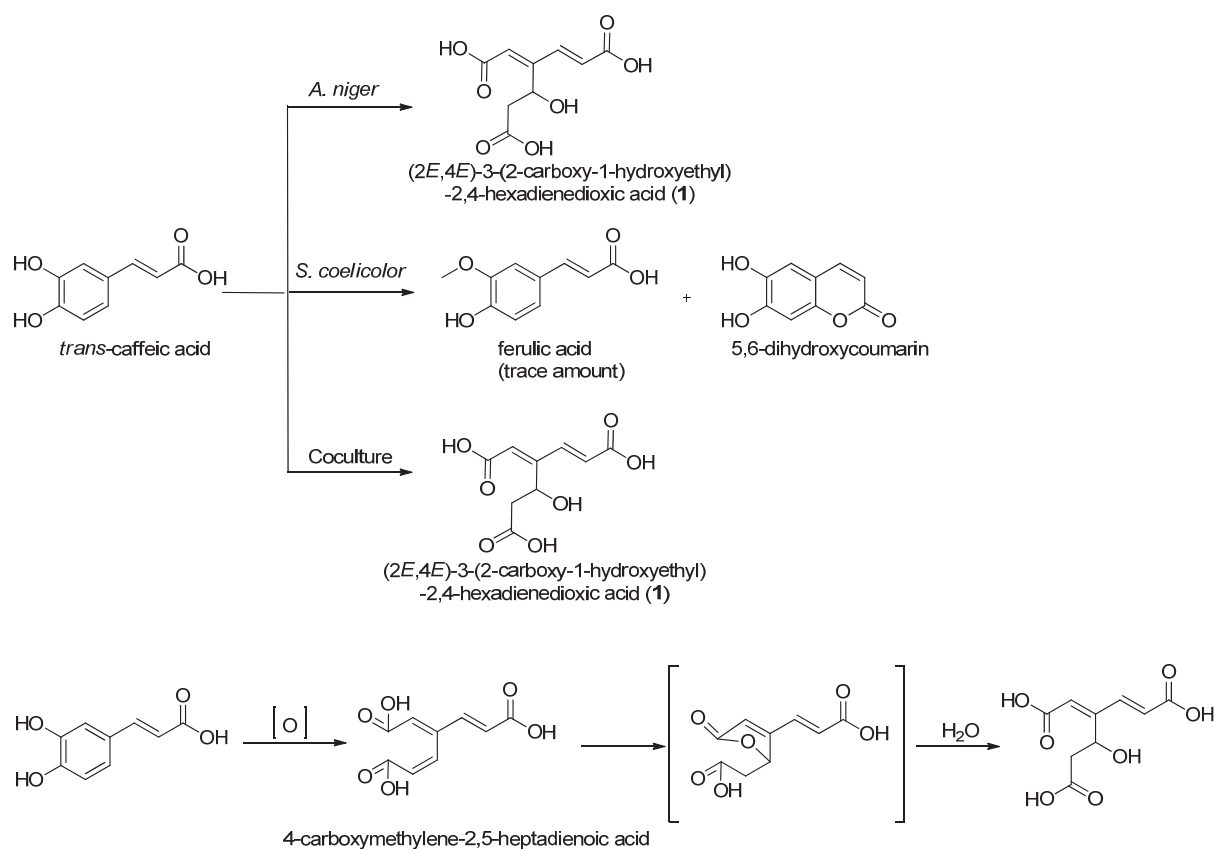


Figure S7. Biotransformation products of *trans*-caffeic acid by *S. coelicolor* monoculture, *A. niger* monoculture, and coculture.

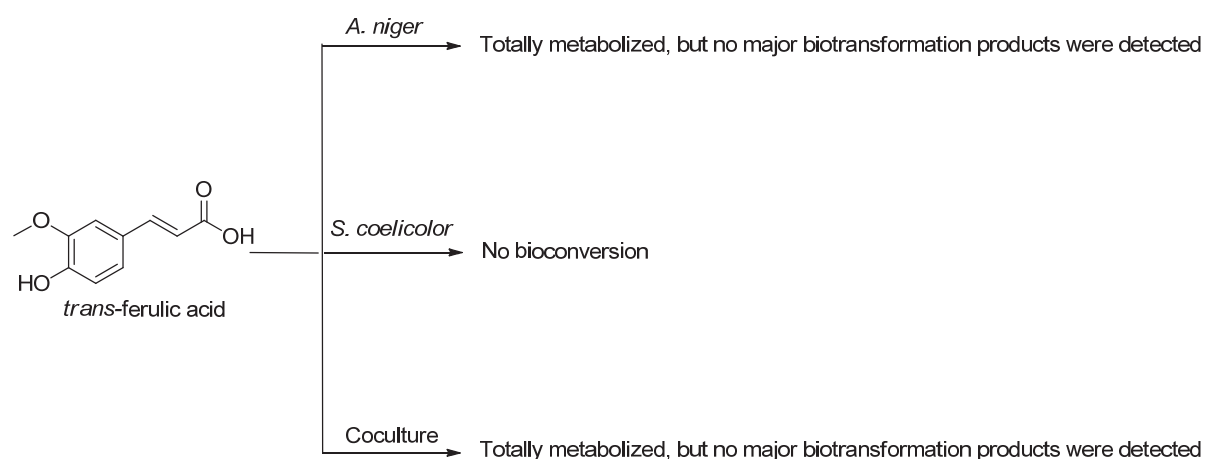


Figure S8. Biotransformation products of *trans*-ferulic acid by *S. coelicolor* monoculture, *A. niger* monoculture, and coculture.

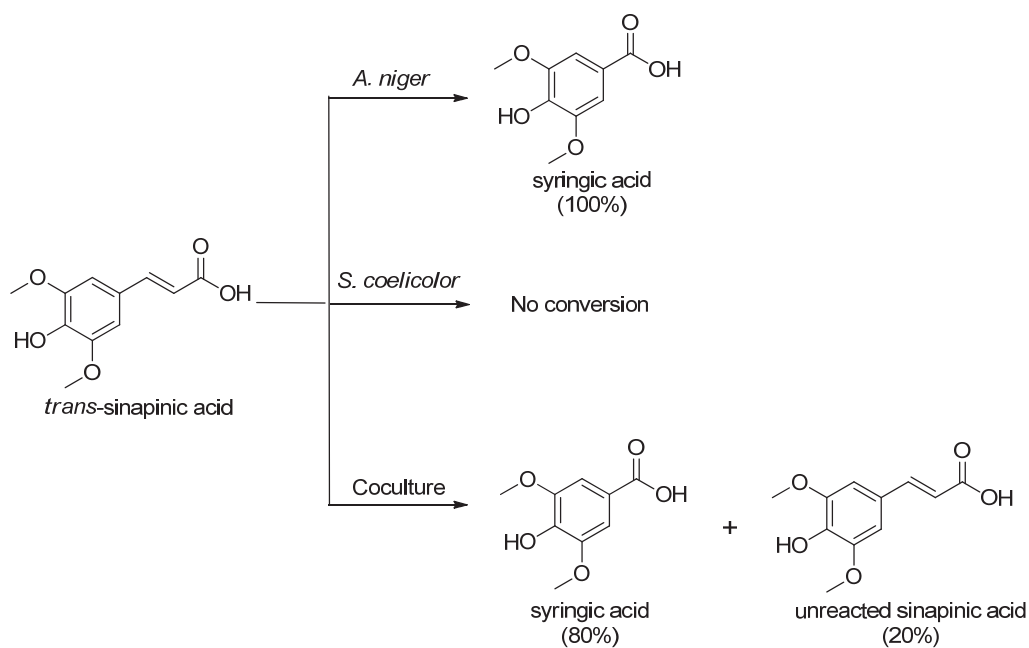


Figure S9. Biotransformation products of *trans*-sinapinic acid by *S. coelicolor* monoculture, *A. niger* monoculture and their coculture.

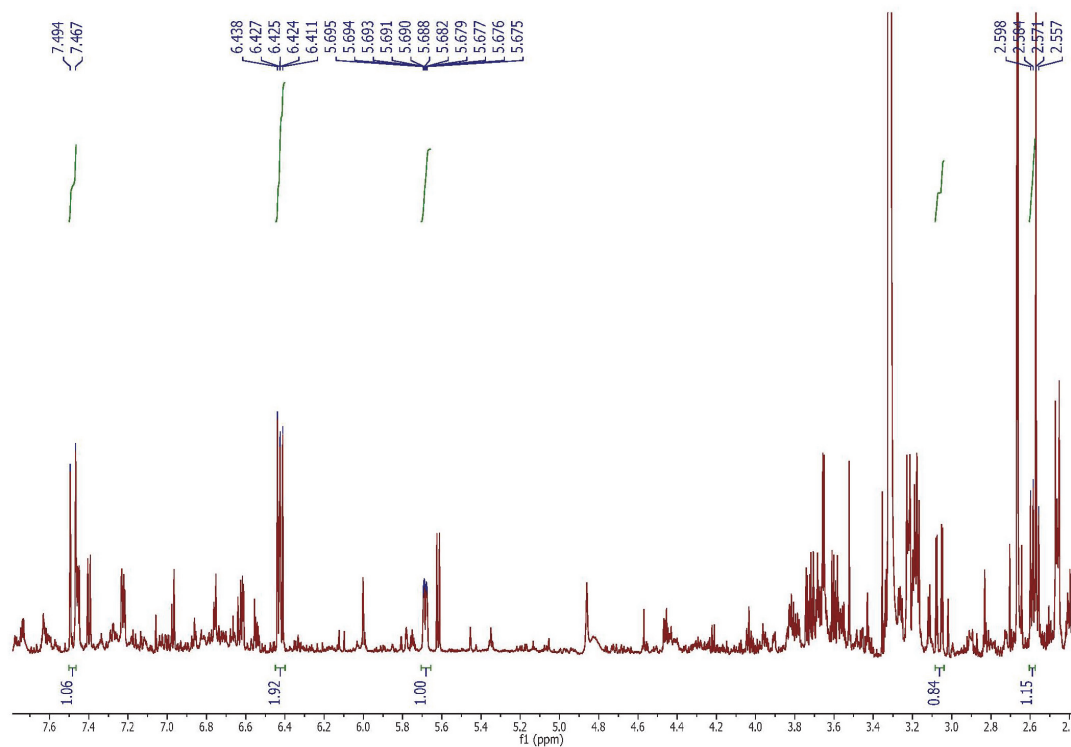


Figure S10. ¹H NMR spectrum of caffeic acid biotransformation by coculture of *S. coelicolor* with *A. niger*. Integrated signals belong to (2*E*,4*E*)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid (**1**)

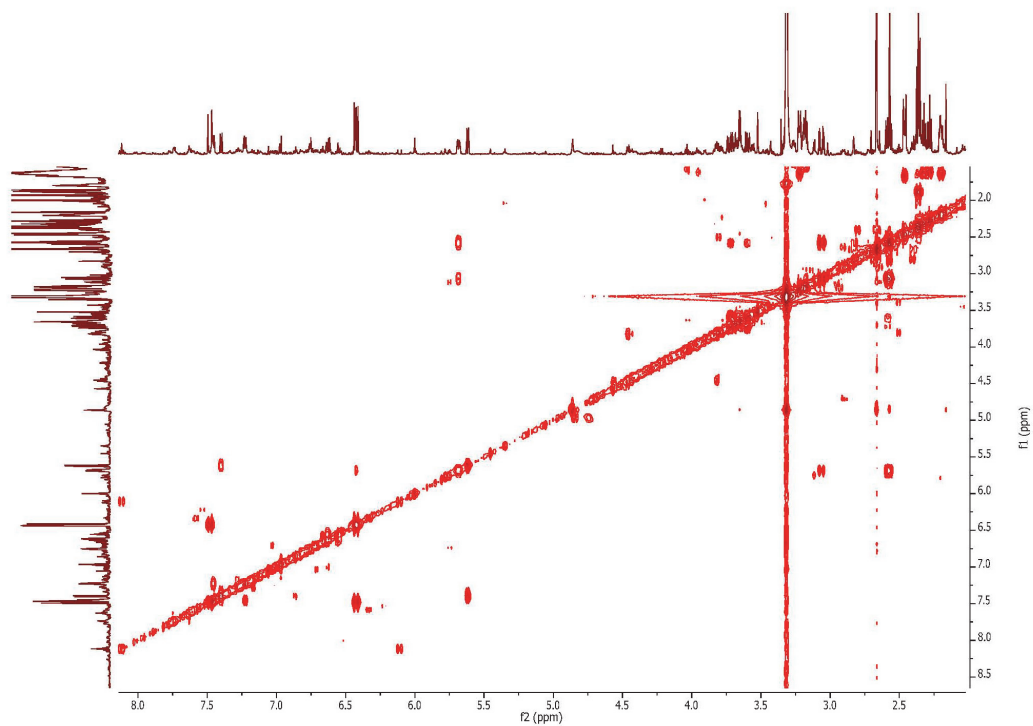


Figure S11. ^1H - ^1H COSY of caffeic acid biotransformation by coculture of *S. coelicolor* with *A. niger*.

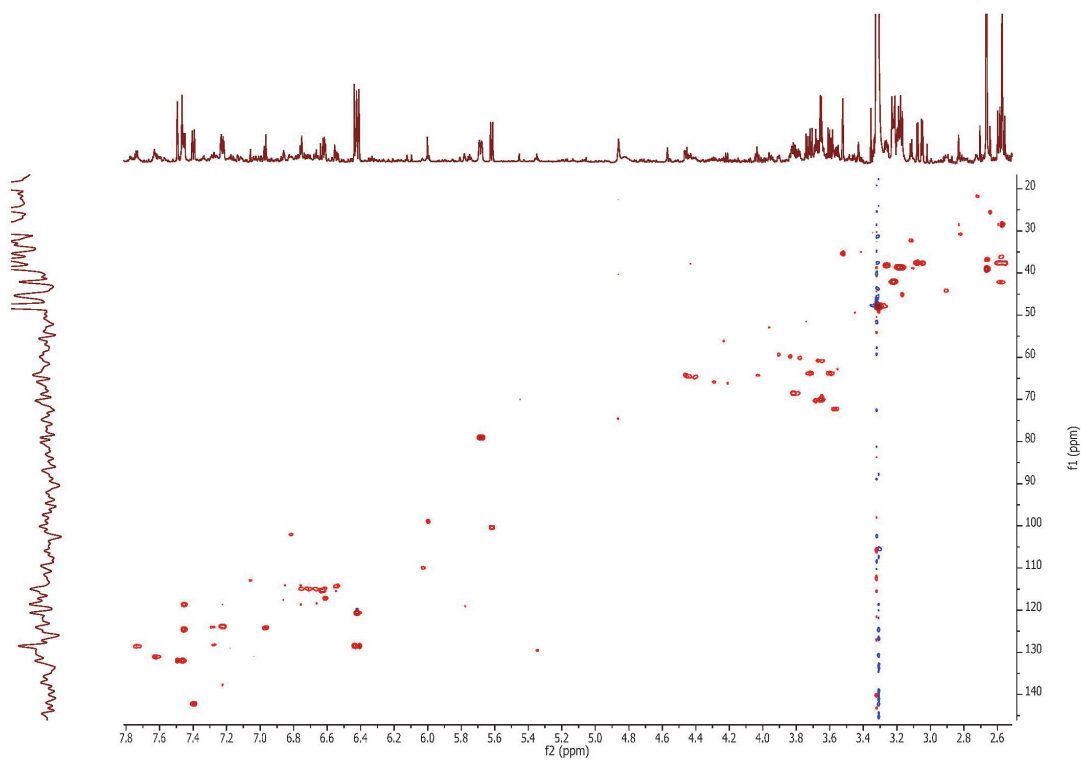


Figure S12. HSQC spectrum of caffeic acid biotransformation by coculture of *S. coelicolor* with *A. niger*.

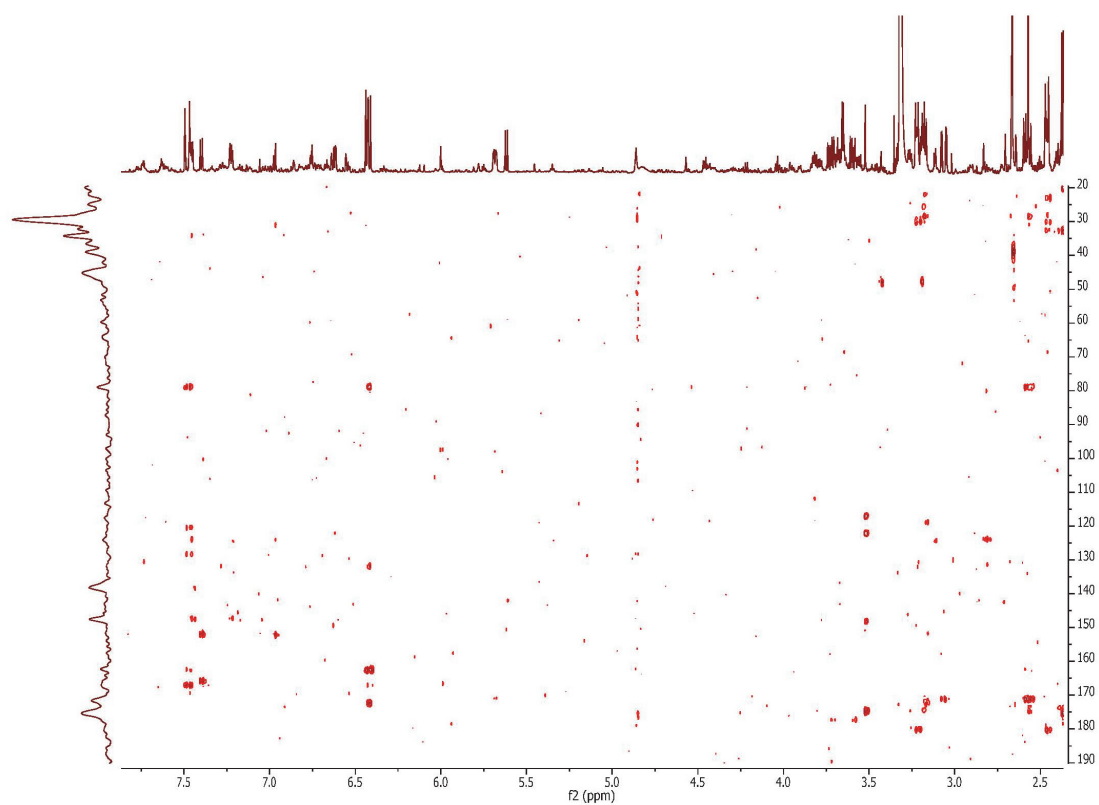


Figure S13. HMBC spectrum of caffeic acid biotransformation by coculture of *S. coelicolor* with *A. niger*.

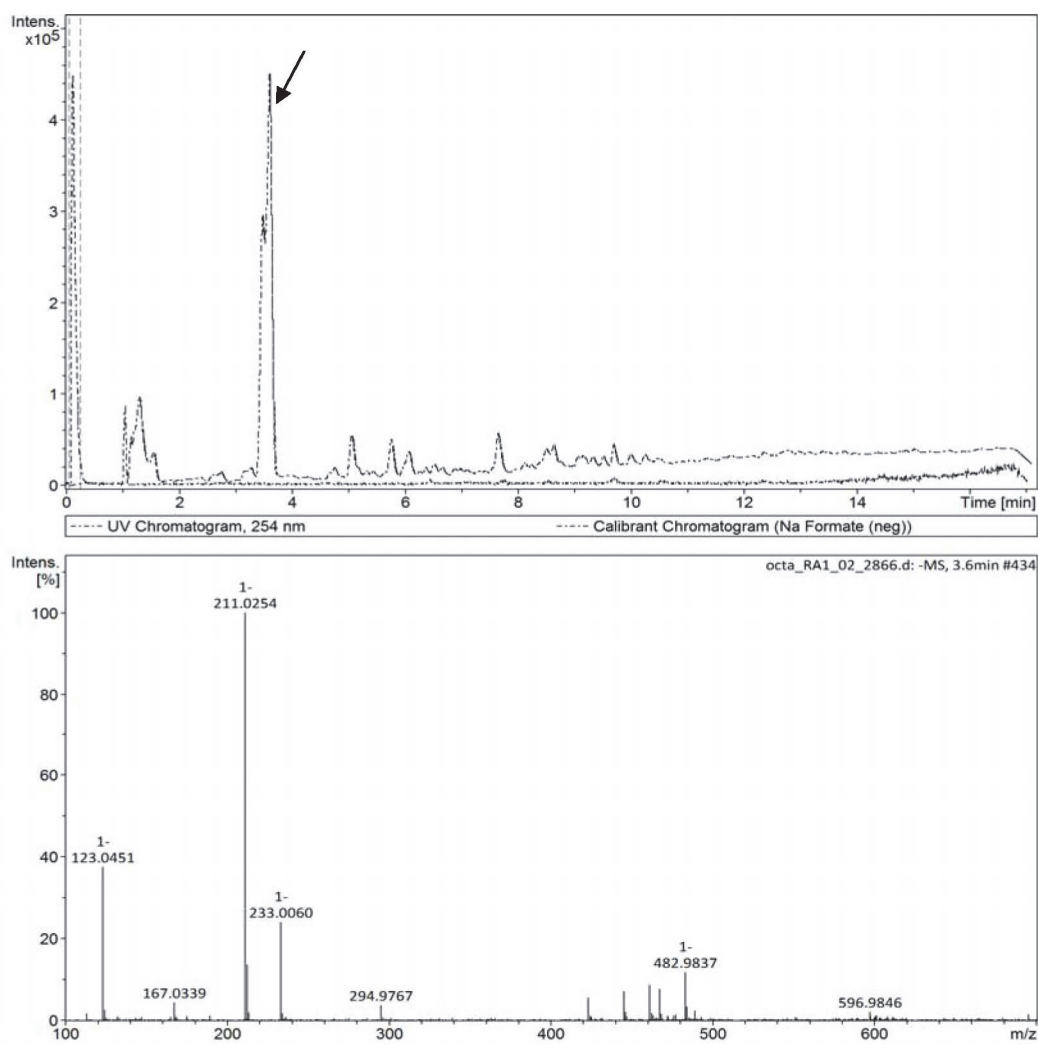


Figure S14. U(H)PLC-Q-TOF analysis of caffeic acid biotransformation by coculture of *S. coelicolor* with *A. niger*. Arrow points at the MS peak at 211.0254.

Table S1. ¹H NMR and HRMS assignment for the biotransformation products of *trans*-cinnamic acid by *S. coelicolor* monoculture, *A. niger* monoculture, and their coculture.

Origin	Compounds	Characteristic ¹ H NMR	UHPLC-Q-TOF-MS
Substrate	<i>trans</i> -cinnamic acid	7.68, d (16.2); 6.48, d (16.2); 7.59, m; 7.41, m	149.0604 [M + H] ⁺ ; 131.0507 [M – H ₂ O + H] ⁺ ; 147.0451 [M – H] ⁺
<i>S. coelicolor</i>	<i>p</i> -coumaric acid	6.28, d (6.2); 6.81, d (8.4)	165.0538 [M + H] ⁺ ; 147.0438 [M – H ₂ O + H] ⁺ ; 163.0397 [M – H ₂ O – H] ⁺
<i>A. niger</i>	4-hydroxybenzoic acid	7.88, d (9.0); 6.82, d (9.0)	139.0390 [M + H] ⁺ ; 121.0280 [M – H ₂ O + H] ⁺ ; 137.0238 [M – H] ⁺
	(<i>E</i>)-3-phenylprop-2-en-1-ol	7.41, m; 7.39, m; 6.37, dt (16.2, 6.0); 4.23, dd (6.0, 1.8)	117.0707 [M – H ₂ O + H] ⁺
	2-hydroxy-3-phenylpropanoic acid	7.27, m; 4.31, dd (7.8, 4.2); 3.10, dd (7.8, 4.2)	165.0553 [M – H] ⁺ ; 147.0461 [M – H ₂ O + H] ⁺
Coculture	2-hydroxy-3-phenylpropanoic acid	7.27, m; 4.31, dd (7.8, 4.2); 3.10, dd (7.8, 4.2)	165.0553 [M – H] ⁺ ; 147.0461 [M – H ₂ O + H] ⁺

Table S2. ¹H NMR and HRMS assignment for the biotransformation products of *trans*-*o*-coumaric acid by *S. coelicolor* monoculture, *A. niger* monoculture and their coculture.

Origin	Compounds	Characteristic ¹ H NMR	UHPLC-Q-TOF-MS
Substrate	<i>trans</i> - <i>o</i> -coumaric acid	7.97, d (16.2); 6.55, d (16.2); 7.48, dd (8.4, 1.8); 7.21, td (8.4, 1.8); 6.84, td (8.4, 1.8); 6.85, dd (8.4, 1.8)	165.0549 [M + H] ⁺ ; 187.0358 [M + Na] ⁺ ; 147.0441 [M – H ₂ O + H] ⁺ ; 163.0408 [M – H] ⁺
<i>S. coelicolor</i>	coumarin	6.43, d (9.6); 7.60, td (8.4, 1.8); 7.29, m	147.0431 [M + H] ⁺
	2-hydroxybenzamide	7.81, dd (8.4, 1.8); 7.23, td (8.4, 1.8); 6.73, dd (8.4, 1.8); 6.57, td (8.4, 1.8)	138.0548 [M + H] ⁺ ; 120.0441 [M – H ₂ O + H] ⁺ ; 136.0405 [M – H] ⁺
	(<i>E</i>)-methyl 3-(2-hydroxyphenyl)acrylate	3.60, s	179.0703 [M + H] ⁺ ; 177.0552 [M – H] ⁺
<i>A. niger</i>	4-dihydroxycoumarin	7.91, dd (7.8, 1.8); 7.64, td (8.4, 1.2); 7.35, td (8.4, 1.2); 7.34, dd (7.8, 1.8); 5.65, s	163.0375 [M + H] ⁺ ; 185.0207 [M + Na] ⁺ ; 161.0274 [M – H] ⁺
	4-hydroxychroman-2-one	7.22, dd (8.4, 2.4); 7.08, td (7.8, 1.2); 6.82, td (7.8, 1.2); 6.76, dd (8.4, 1.2); 5.39, dd (9.6, 3.6); 2.79, dd (15.6, 3.6); 2.59, dd (15.6, 9.6)	165.0547 [M + H] ⁺ ; 147.0439 [M – H ₂ O + H] ⁺ ; 163.0396 [M – H] ⁺
	(<i>Z</i>)-3-hydroxy-3-(2-hydroxyphenyl)acrylic acid	7.96, dd (7.8, 1.8); 7.52, td (8.4, 1.2); 7.27, td (8.4, 1.2); 7.26, dd (7.8, 1.8); 5.90, s	181.0490 [M + H] ⁺ ; 203.0331 [M + Na] ⁺ ; 163.0391 [M – H ₂ O + H] ⁺ ; 179.0346 [M – H] ⁺
	2-allylphenol	6.33, m; 5.15, dt (10.2, 1.8); 5.12, dt (16.8, 1.8)	135.0796 [M + H] ⁺
Coculture	4-dihydroxycoumarin	7.91, dd (7.8, 1.8); 7.64, td (8.4, 1.2); 7.35, td (8.4, 1.2); 7.34, dd (7.8, 1.8); 5.65, s	163.0375 [M + H] ⁺ ; 185.0207 [M + Na] ⁺ ; 161.0274 [M – H] ⁺
	4-hydroxychroman-2-one	7.22, dd (8.4, 2.4); 7.08, td (7.8, 1.2); 6.82, td (7.8, 1.2); 6.76, dd (8.4, 1.2); 5.39, dd (9.6, 3.6); 2.79, dd (15.6, 3.6); 2.59, dd (15.6, 9.6)	165.0547 [M + H] ⁺ ; 147.0439 [M – H ₂ O + H] ⁺ ; 163.0396 [M – H] ⁺
	3-(2-hydroxyphenyl)-2-propen-1-ol	7.36, td (7.8, 1.2); 7.04, td (7.8, 1.8); 6.76, dd (7.8, 1.2); 6.88, dt (16.2, 1.8); 6.36, dt (16.2, 6.0); 4.22, dd (6.0, 1.8)	133.0649 [M – H ₂ O + H] ⁺ ; 149.0615 [M – H] ⁺ ; 131.0502 [M – H ₂ O – H] ⁺
	(<i>Z</i>)-3-hydroxy-3-(2-hydroxyphenyl)acrylic acid	7.96, dd (7.8, 1.8); 7.52, td (8.4, 1.2); 7.27, td (8.4, 1.2); 7.26, dd (7.8, 1.8); 5.90, s	181.0490 [M + H] ⁺ ; 203.0331 [M + Na] ⁺ ; 163.0391 [M – H ₂ O + H] ⁺ ; 179.0346 [M – H] ⁺
	2-allylphenol	6.33, m; 5.15, dt (10.2, 1.8); 5.12, dt (16.8, 1.8);	135.0796 [M + H] ⁺

Table S3. ^1H NMR and HRMS assignment for the biotransformation products of *trans-m*-coumaric acid by *S. coelicolor* monoculture, *A. niger* monoculture and their coculture.

Origin	Compounds	Characteristic ^1H NMR	UHPLC-Q-TOF-MS
Substrate	<i>trans-m</i> -coumaric acid	7.59, d (16.2); 6.41, d (16.2); 7.22, t (7.8); 7.05, brd (7.2); 7.0, t (1.8); 6.84, ddd (8.4, 2.4, 1.2)	165.0547 $[\text{M} + \text{H}]^+$; 187.0361 $[\text{M} + \text{Na}]^+$; 147.0442 $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$; 163.0417 $[\text{M} - \text{H}]^-$
<i>S. coelicolor</i>	No products		
<i>A. niger</i>	No major biotransformation products		
Coculture	<i>cis-m</i> -coumaric acid	6.87, d (12.6); 5.93, d (12.6); 7.15, t (7.8); 7.06, t (1.8); 7.01, ddd (8.4, 2.4, 1.2)	147.0432 $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$; 163.0405 $[\text{M} - \text{H}]^-$
	<i>p</i> -cresol	6.73, d (8.4); 7.09, d (8.4)	107.0501 $[\text{M} - \text{H}]^-$

Table S4. ^1H NMR and HRMS assignment for the biotransformation products of *trans-p*-coumaric acid by *S. coelicolor* monoculture, *A. niger* monoculture and their coculture.

Origin	Compounds	Characteristic ^1H NMR	UHPLC-Q-TOF-MS
Substrate	<i>trans-p</i> -coumaric acid	7.61, d (16.2); 6.29, d (16.2); 7.45, d (8.4); 6.81, d (8.4)	165.0548 $[\text{M} + \text{H}]^+$; 187.0361 $[\text{M} + \text{Na}]^+$; 147.0446 $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$; 163.0417 $[\text{M} - \text{H}]^-$
<i>S. coelicolor</i>	No products		
<i>A. niger</i>	4-vinylphenol	6.63, dd (17.4, 10.8); 5.56, dd (17.4, 1.2); 5.03, dd (10.8, 1.2); 6.73, d (8.4); 7.26, d (8.4)	121.0649 $[\text{M} + \text{H}]^+$; 119.0497 $[\text{M} - \text{H}]^-$
Coculture	4-vinylphenol	6.63, dd (17.4, 10.8); 5.56, dd (17.4, 1.2); 5.03, dd (10.8, 1.2); 6.73, d (8.4); 7.26, d (8.4)	121.0649 $[\text{M} + \text{H}]^+$; 119.0497 $[\text{M} - \text{H}]^-$

Table S5. ^1H NMR and HRMS assignment for the biotransformation products of *trans*-caffeic acid by *S. coelicolor* monoculture, *A. niger* monoculture and their coculture.

Origin	Compounds	Characteristic ^1H NMR	UHPLC-Q-TOF-MS
Substrate	<i>trans</i> -caffeic acid	7.53, d (16.2); 6.22, d (16.2); 7.04, d (1.8); 6.94, dd (8.4, 1.8); 6.79, d (8.4)	181.0492 $[\text{M} + \text{H}]^+$; 203.0312 $[\text{M} + \text{Na}]^+$; 163.0393 $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$; 179.0374 $[\text{M} - \text{H}]^-$
<i>S. coelicolor</i>	<i>trans</i> -ferulic acid	7.60, d (16.2); 6.32, d (16.2); 7.18, d (1.8); 7.07, dd (8.4, 1.8); 6.82, d (8.4); 3.90, s	195.0657 $[\text{M} + \text{H}]^+$; 177.0540 $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$; 193.0514 $[\text{M} - \text{H}]^-$
	5,6-dihydroxycoumarin	7.79, d (9.6); 6.19, d (9.6); 6.76, s	179.0335 $[\text{M} + \text{H}]^+$; 177.0200 $[\text{M} - \text{H}]^-$
<i>A. niger</i>	(2 <i>E</i> ,4 <i>E</i>)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid (1)	see Table 2	211.0254 $[\text{M} - \text{H}_2\text{O} - \text{H}]^-$
Coculture	(2 <i>E</i> ,4 <i>E</i>)-3-(2-carboxy-1-hydroxyethyl)-2,4-hexadienedioic acid (1)	see Table 2	211.0254 $[\text{M} - \text{H}_2\text{O} - \text{H}]^-$

Table S6. ^1H NMR and HRMS assignment for the biotransformation products of *trans*-ferulic acid by *S. coelicolor* monoculture, *A. niger* monoculture and their coculture.

Origin	Compounds	Characteristic ^1H NMR	UHPLC-Q-TOF-MS
Substrate	<i>trans</i> -ferulic acid	7.60, d (16.2); 6.32, d (16.2); 7.18, d (2.4); 7.07, dd (8.4, 2.4); 6.82, d (8.4); 3.90, s	195.0661 $[\text{M} + \text{H}]^+$; 217.0447 $[\text{M} + \text{Na}]^+$; 177.0563 $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$; 193.0511 $[\text{M} - \text{H}]^-$
<i>S. coelicolor</i>	No products		
<i>A. niger</i>	No products		
Coculture	No products		

Table S7. ^1H NMR and HRMS assignment for the biotransformation products of *trans*-sinapinic acid by *S. coelicolor* monoculture, *A. niger* monoculture and their coculture.

Origin	Compounds	Characteristic ^1H NMR	UHPLC-Q-TOF-MS
Substrate	<i>trans</i> -sinapinic acid	7.60, d (16.2); 6.34, d (16.2); 6.90, s	225.0779 $[\text{M} + \text{H}]^+$; 247.0590 $[\text{M} + \text{Na}]^+$; 207.0683 $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$; 223.0633 $[\text{M} - \text{H}]^-$
<i>S. coelicolor</i>	No conversion		
<i>A. coelicolor</i>	4-hydroxy-3,5-dimethoxybenzoic acid	7.33, s; 3.89, s	199.0608 $[\text{M} + \text{H}]^+$; 221.0423 $[\text{M} + \text{Na}]^+$; 181.0493 $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$; 197.0468 $[\text{M} - \text{H}]^-$
Coculture	<i>trans</i> -sinapinic acid	7.60, d (16.2); 6.34, d (16.2); 6.90, s	225.0779 $[\text{M} + \text{H}]^+$; 247.0590 $[\text{M} + \text{Na}]^+$; 207.0683 $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$; 223.0633 $[\text{M} - \text{H}]^-$
	4-hydroxy-3,5-dimethoxybenzoic acid	7.33, s; 3.89, s	199.0608 $[\text{M} + \text{H}]^+$; 221.0423 $[\text{M} + \text{Na}]^+$; 181.0493 $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$; 197.0468 $[\text{M} - \text{H}]^-$

Appendix III: Supplementary Information belonging to Chapter 6

Table S1. Gene organization of type I iterative non-reducing PKS gene cluster (*icm*) for isocoumarins biosynthesis in *Streptomyces* sp. MBT76.

ORF	Product	Accession in GeneBank	Length	Putative function	Homologue in <i>Streptomyces roseovirentis</i>	Homology
1	<i>icmA</i>	WP_058043254.1	943	LuxR family transcriptional regulator	WP_030365549.1	84%
2	<i>icmB</i>	WP_058043255.1	224	TetR family transcriptional regulator	WP_043188202.1	92%
3	<i>icmC</i>	WP_058043256.1	496	MFS transporter	WP_043188355.1	84%
4	<i>icmD</i>	WP_058043257.1	200	NADPH-dependent FMN reductase	WP_030365552.1	88%
5	<i>icmE</i>	WP_058043258.1	246	ToxA protein	WP_030365553.1	87%
6	<i>icmF</i>	WP_058043259.1	416	oxidoreductase	WP_052392495.1	71%
7	<i>icmG</i>	WP_058043260.1	376	diaminohydroxyphosphoribosylaminopyrimidine deaminase/5-amino-6-(5-phosphoribosylamino)uracil reductase	WP_030365554.1	88%
8	<i>icmH</i>	WP_058043261.1	330	Serine/threonine kinase	WP_030365555.1	90%
9	<i>icmI</i>	WP_058043262.1	251	methyltransferase type 12	WP_043188204.1	81%
10	<i>icmJ</i>	WP_058043263.1	570	WD repeat-containing protein	WP_030365558.1	95%
11	<i>icmK</i>	WP_058043264.1	208	GTP cyclohydrolase II	WP_043188207.1	98%
12	<i>icmL</i>	WP_058043265.1	130	Hypothetical protein	WP_030365560.1	88%
13	<i>icmM</i>	WP_058043266.1	1239	iterative type I polyketide synthase	WP_030365561.1	91%
14	<i>icmN</i>	WP_058043267.1	397	AMP-dependent synthetase/ligase	WP_030365562.1	95%
15	<i>icmO</i>	WP_058043268.1	434	MFS transporter	WP_053688342.1	81%
16	<i>icmP</i>	WP_058043269.1	381	hypothetical protein	WP_030365564.1	93%
17	<i>icmQ</i>	WP_058043270.1	388	acyl-protein synthase	WP_030365565.1	96%
18	<i>icmR</i>	WP_058043271.1	834	hypothetical protein	WP_030365566.1	94%
19	<i>icmS</i>	WP_058043272.1	303	phosphotriesterase	WP_052392497.1	88%
20	<i>icmT</i>	WP_058043273.1	188	Zn-ribbon-like motif containing protein	WP_030365568.1	88%

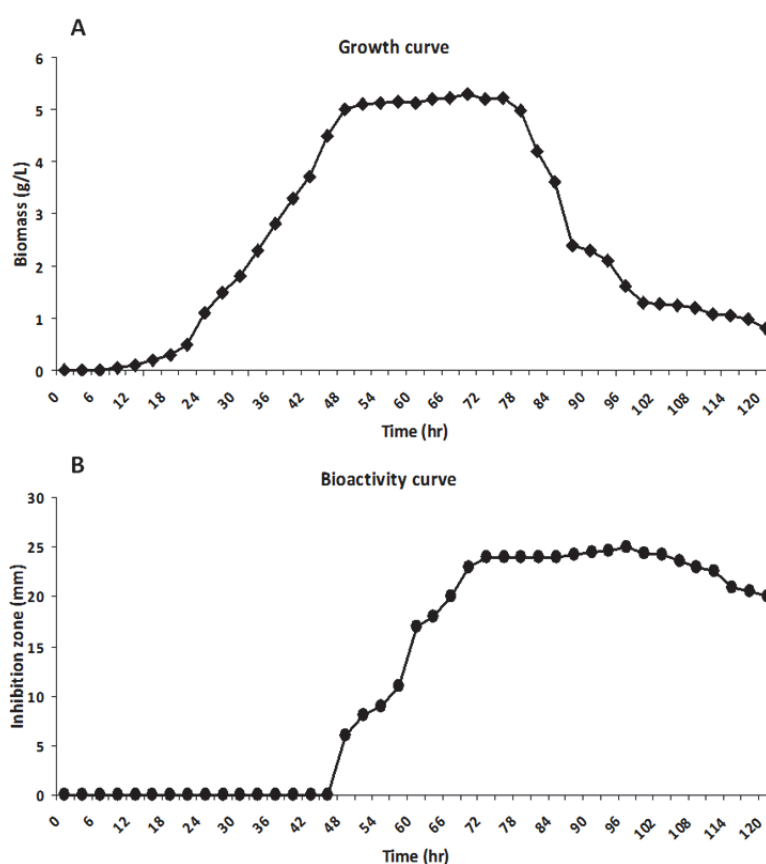


Figure S1. Growth curve of *Streptomyces* sp. MBT76. (A) and corresponding antimicrobial activity against *Bacillus subtilis* (B). Some 1.5 ml culture was collected every two hours during the daytime from the first to the fifth inoculation day. Separated supernatant from the pellet by centrifuging and kept them independently. The pellet was heated at 70 °C for overnight to determine the dry weight, while the supernatant was accessed for antimicrobial activity against *Bacillus subtilis* 168. The data were means of three replicates.

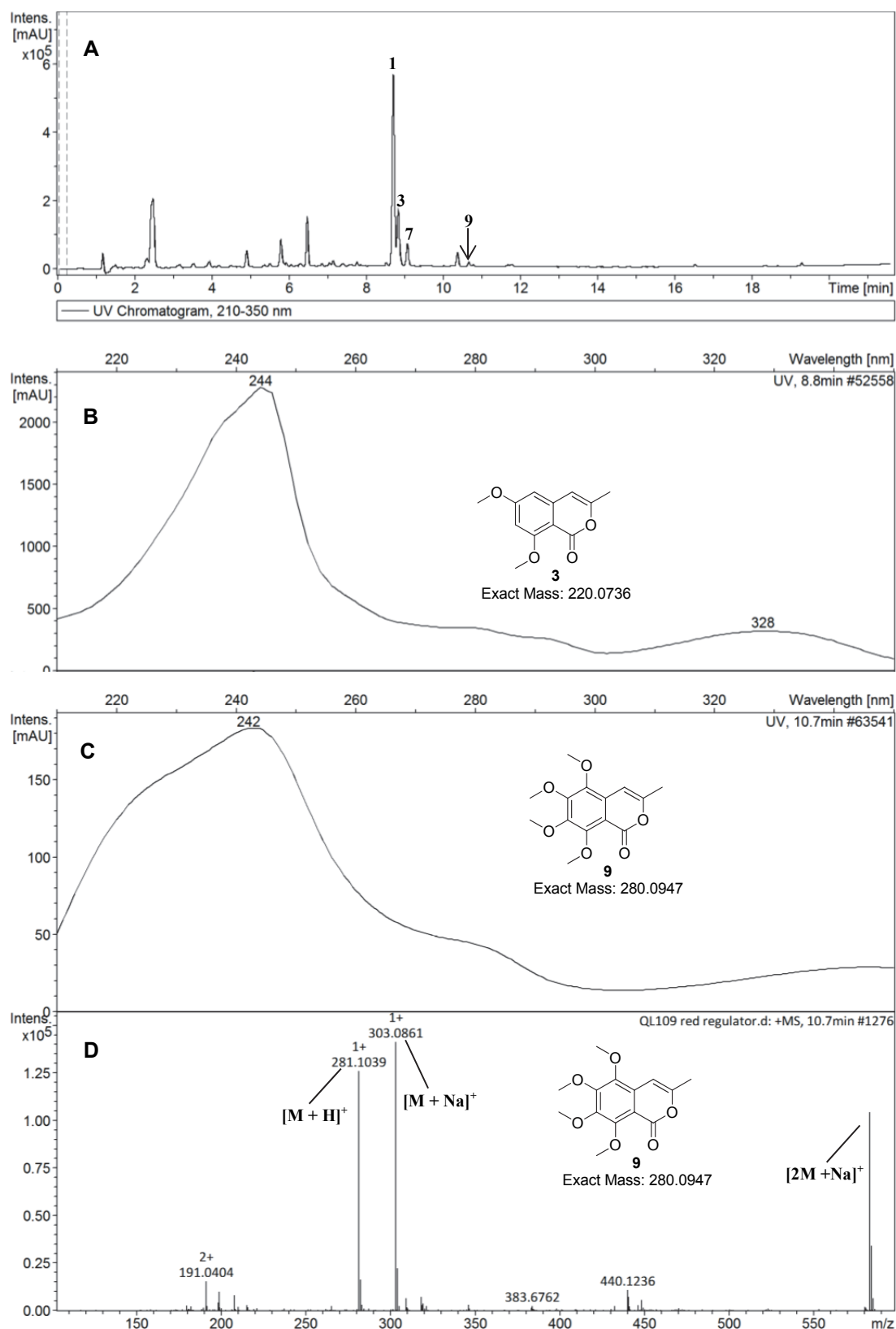


Figure S2. U(H)PLC-UV-ToF-HRMS analysis of *Streptomyces* sp. MBT76 crude extract. A) UV chromatogram of *Streptomyces* sp. MBT76 crude extract, whereby the isocoumarins 1, 3, 7, 9 were labelled; B) UV spectrum of 6,8-dimethoxy-3-methyl-isocoumarin (3); C) UV spectrum of 5,6,7,8-tetramethoxy-3-methyl-isocoumarin (9); D) HRESIMS spectrum of 5,6,7,8-tetramethoxy-3-methyl-isocoumarin (9); Electrospray (ESI) detection was achieved in positive ion mode.

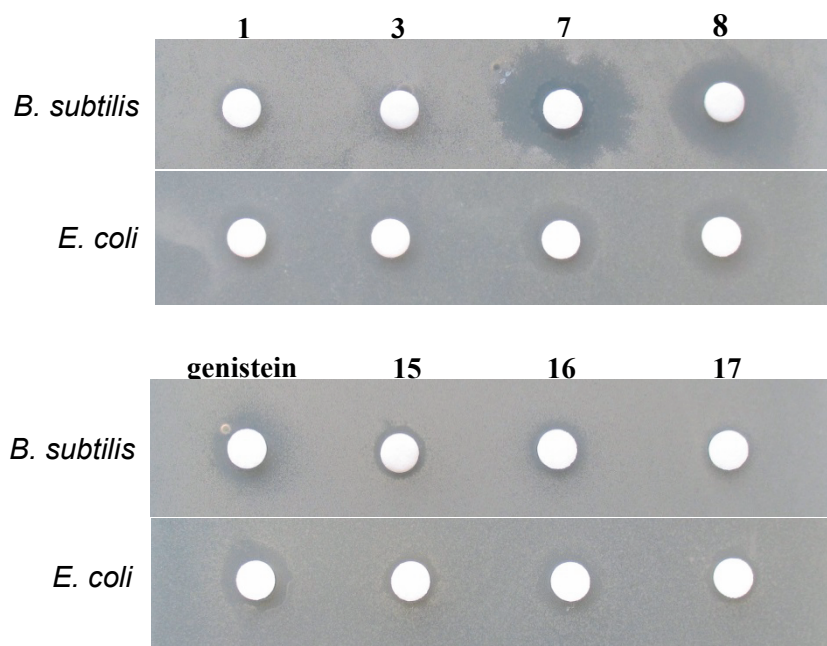


Figure S3. Antimicrobial activity of selected isocoumarins and isoflavones. The methoxylation enhanced the antimicrobial potency of isocoumarins against both Gram-positive *Bacillus subtilis* 168 and Gram-negative *Escherichia coli* K12, which in turn validated the metabolomics OPLS analysis in Figure 4. On the contrary, methylation decreased the antimicrobial activity of isoflavones.

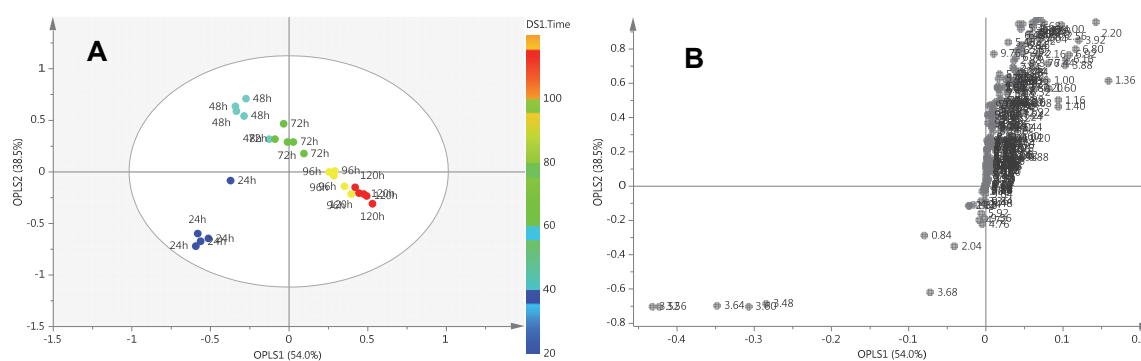


Figure S4. OPLS analysis of the NMR spectra by using time as the Y-variable. (A) score plot, and (B) the corresponding loading S-plot.

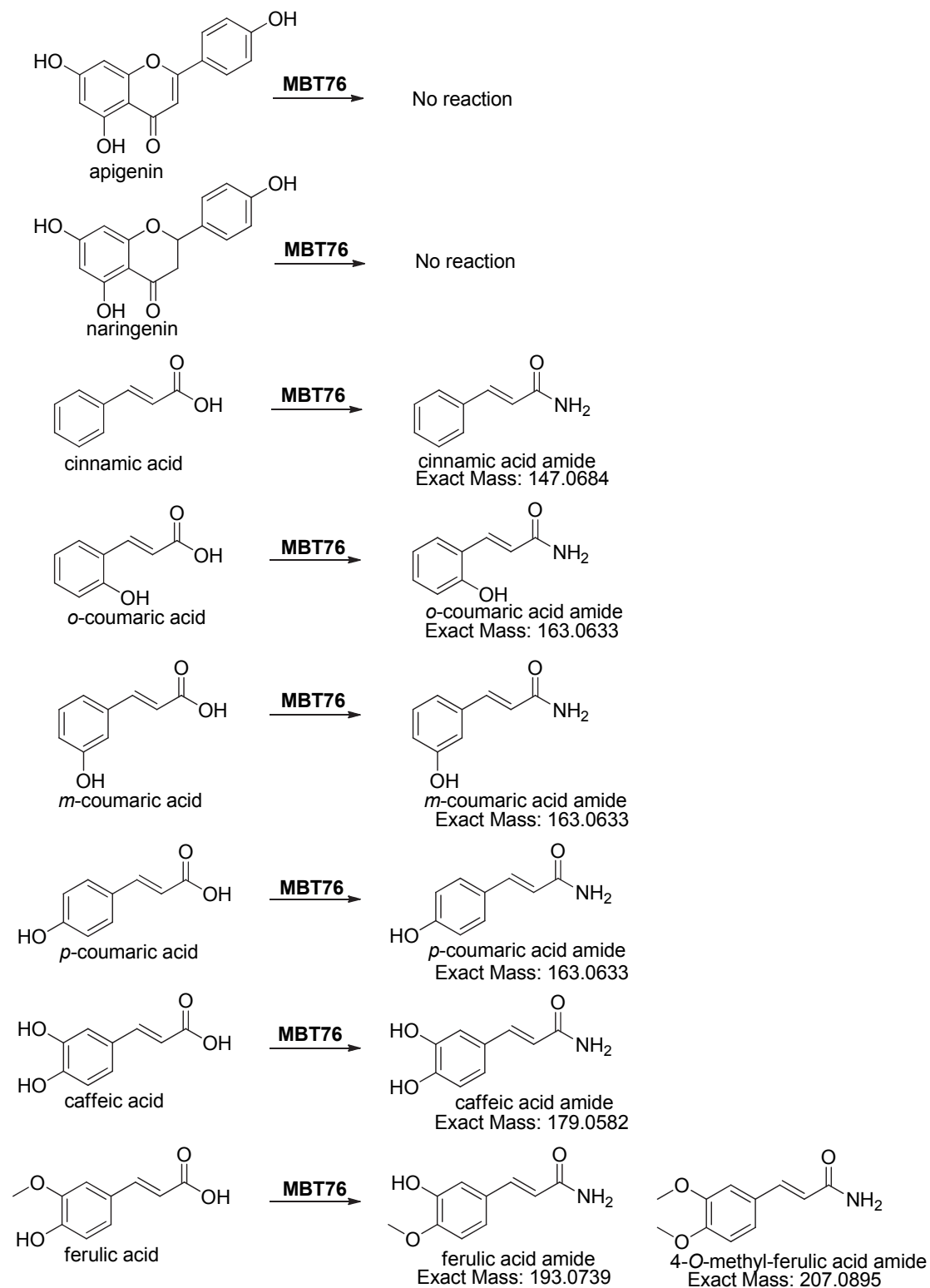
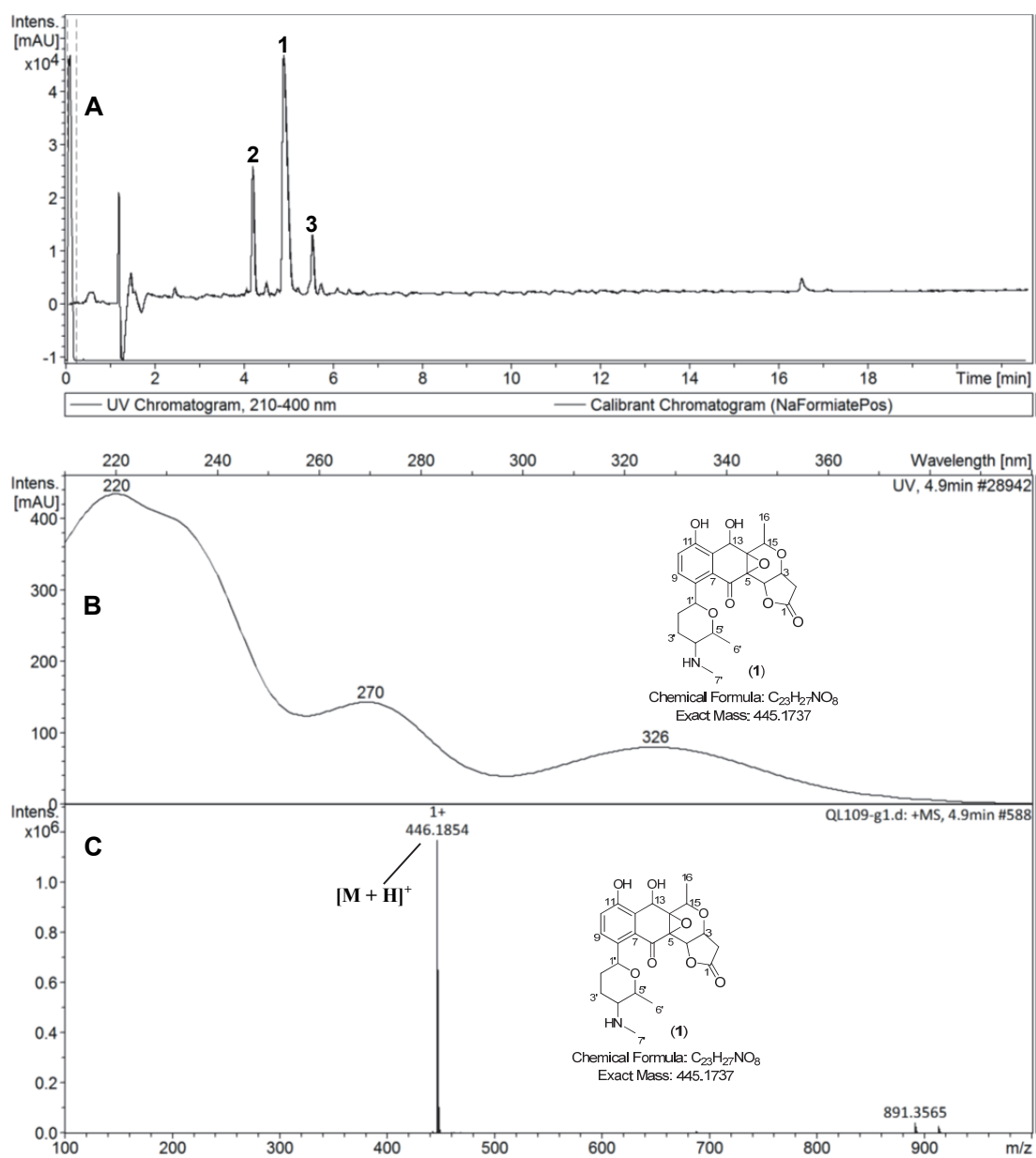


Figure S5. Biotransformation study in *Streptomyces* sp. MBT76 by using apigenin, naringenin, and (hydroxyl)cinnamic acids as substrates. Products were identified on the basis of ^1H NMR and/or UHPLC-TOF-MS high resolution mass.

Appendix IV: Supplementary Information belonging to Chapter 7

Table S1. Spectral data assignments for the compounds **4**—**15** shown in Figure 2. Compound identification was based on ^1H NMR, high resolution mass spectrometry, and UV absorption spectra. Proton coupling constants (J in Hz) are given in parentheses.

NO.	Name	Chemical shifts (ppm)	HRMS (Da)	UV (λ_{max} nm)	Ref
4	BE-54238A	7.90, d (8.4); 6.47, d (8.4); 5.26, m; 4.51, m; 4.31, m; 3.32, m; 3.18, dd (10.8, 9.0); 2.96, dd (16.2, 10.8); 2.77, dd (8.4, 4.2); 2.73, dd (15.6, 4.2); 2.62, dd (15.6, 8.4); 1.63, d (6.6); 1.32, d (6.6)	398.1630 $[\text{M} + \text{H}]^+$	276, 414	¹
5	BE-54238B	8.01, d (9.0); 6.56, d (9.0)	396.1459 $[\text{M} + \text{H}]^+$	278, 423	¹
6	kalafungfin		301.0710 $[\text{M} + \text{H}]^+$	216, 256, 425	²
7	dihydrokalafungfin		303.0856 $[\text{M} + \text{H}]^+$	222, 278	³
8	11- <i>O</i> -methyl-aloesaponarin II	7.80, dd (7.8, 1.2); 7.70, t (7.8); 7.50, brd (8.4); 7.43, d (1.8); 7.43, d (1.8); 6.99, d (1.8); 3.98, s; 2.71, s	269.0825 $[\text{M} + \text{H}]^+$; 291.0638 $[\text{M} + \text{Na}]^+$	218, 274, 390	⁴
9	3,8-dihydroxy-1-methyl-anthraquinone-carboxylic acid	7.70, d (7.8); 7.64, t (7.8); 7.27, d (7.8); 7.60, s	299.0555 $[\text{M} + \text{H}]^+$	222, 280, 410	⁴
10	GTTRI-02	6.62, s; 4.24, m; 3.15, dd (16.2, 3.6); 2.91, dd (16.2, 7.2); 2.85, dd (16.2, 3.6); 2.60, dd (16.2, 7.2); 2.45, s; 2.42, s	235.0956 $[\text{M} + \text{H}]^+$; 257.0776 $[\text{M} + \text{Na}]^+$; 217.0850 $[\text{M} - \text{H}_2\text{O} + \text{H}]^+$	216, 278	⁵
11	2-acetyl-3-hydroxy-furan	7.92, d (6.0); 6.38, d (6.0); 2.34, s	127.0428 $[\text{M} + \text{H}]^+$	214, 276	⁶
12	2-acetyl-3-methoxy-furan	7.95, d (6.0); 6.39, d (6.0); 3.78, s; 2.35, s	141.0599 $[\text{M} + \text{H}]^+$; 163.0416 $[\text{M} + \text{Na}]^+$	212, 260	⁶
13	6,8-dihydroxy-3-methyl-isocoumarin	6.29, d (1.2); 6.24, d (2.4); 6.27, d (2.4); 2.21, d (1.2)	193.0511 $[\text{M} + \text{H}]^+$; 215.0312 $[\text{M} + \text{Na}]^+$	244, 328	⁷
14	phenylacetamide	7.29, m; 7.23, m; 3.50, s	136.0759 $[\text{M} + \text{H}]^+$	218, 260	⁸



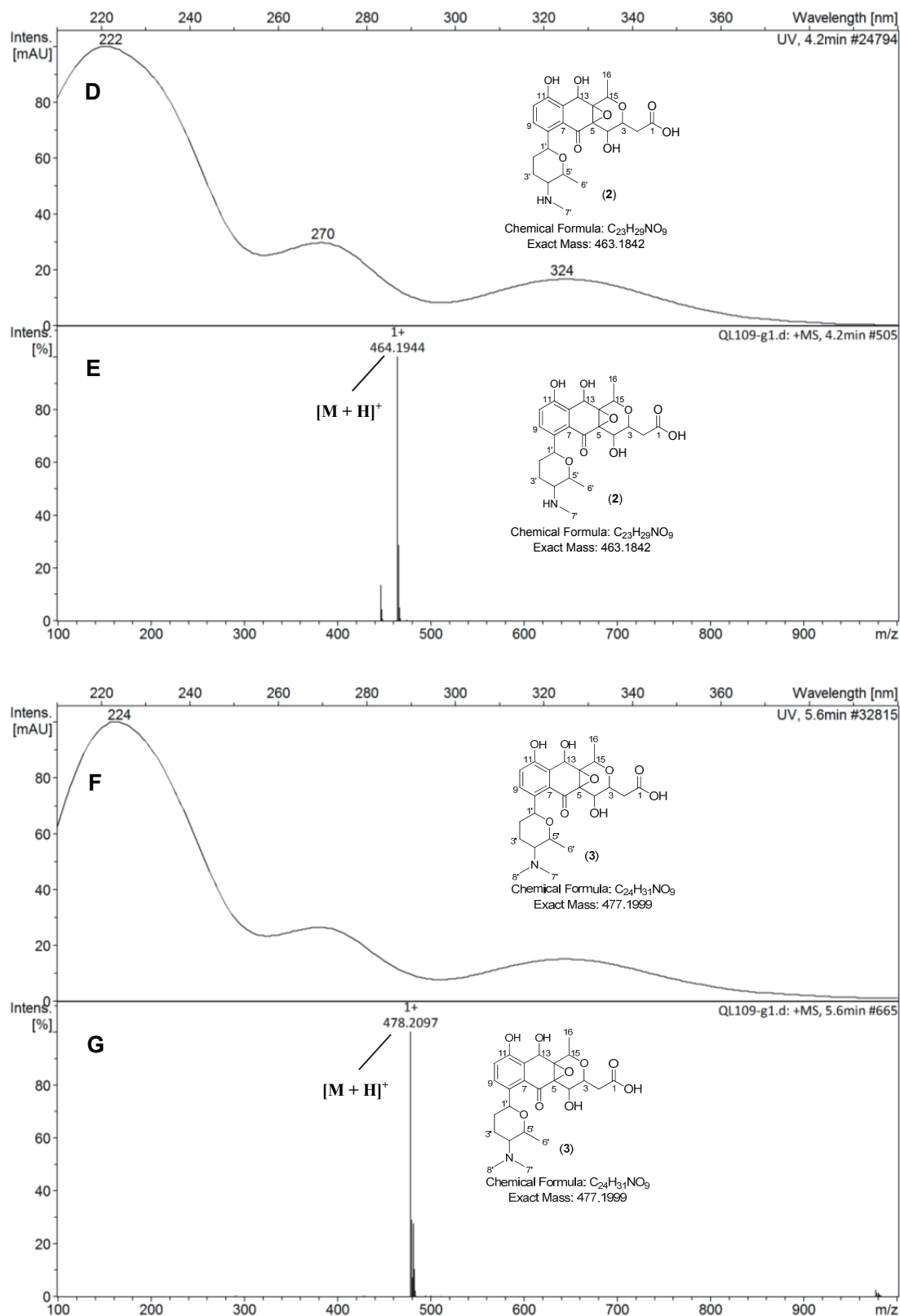


Figure S1. U(H)PLC-UV-TOF-HRMS analysis of Fr1. A) UV chromatogram (detected at 210 nm) of Fr1 presented three major compounds **1–3** with quantity ratio around 6:2:1; B) UV spectrum of compound **1**; C) HRESIMS spectrum of compound **1**; D) UV spectrum of compound **2**; E) HRESIMS spectrum of compound **2**; F) UV spectrum of compound **3**; G) HRESIMS spectrum of compound **3**. Electrospray (ESI) detection was achieved in positive ion mode. The respective inserts are structures of compounds **1–3** with chemical formula and theoretical mass calculated by software ChemDraw Ultra 12.0.

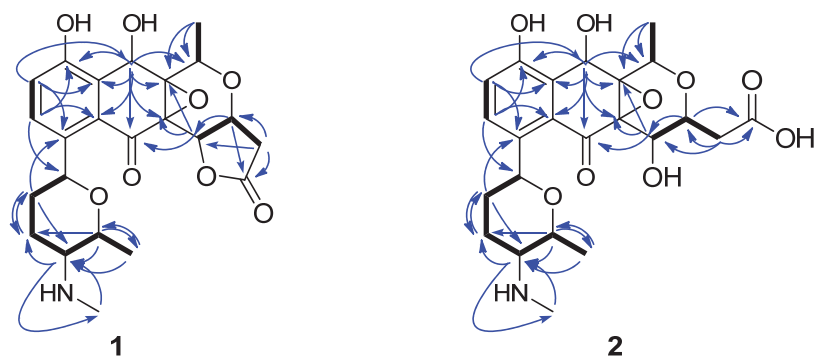
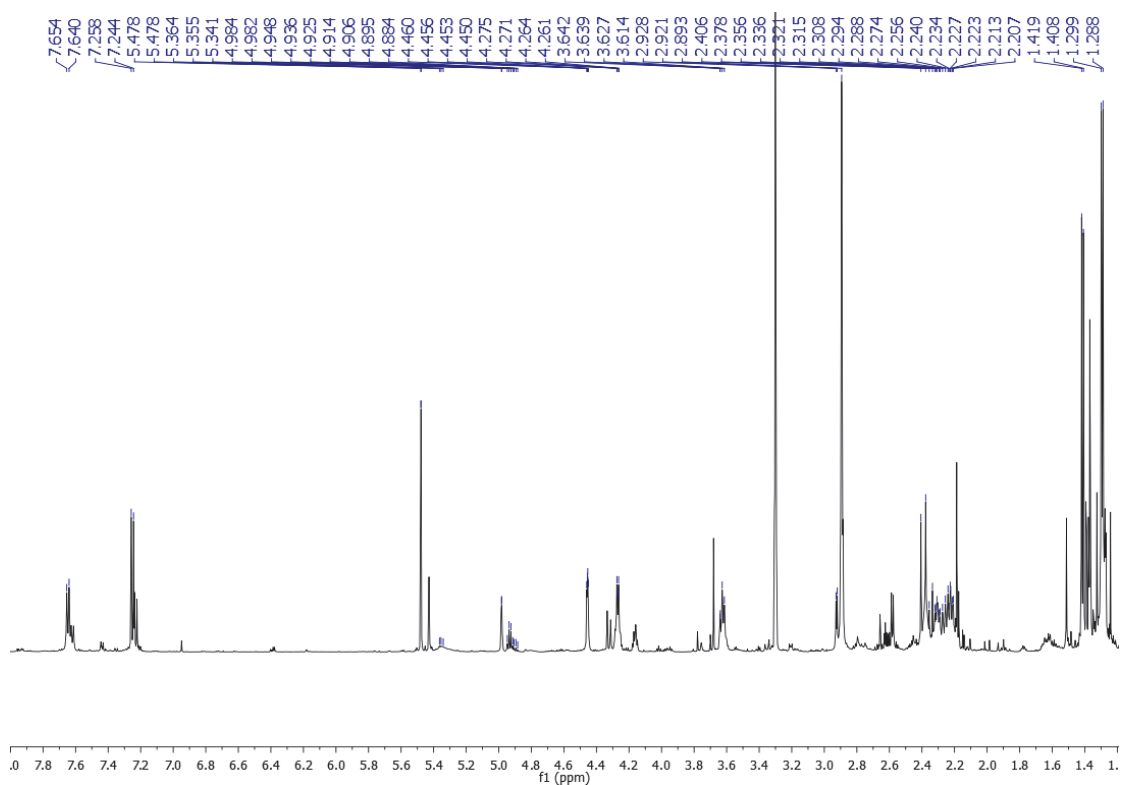


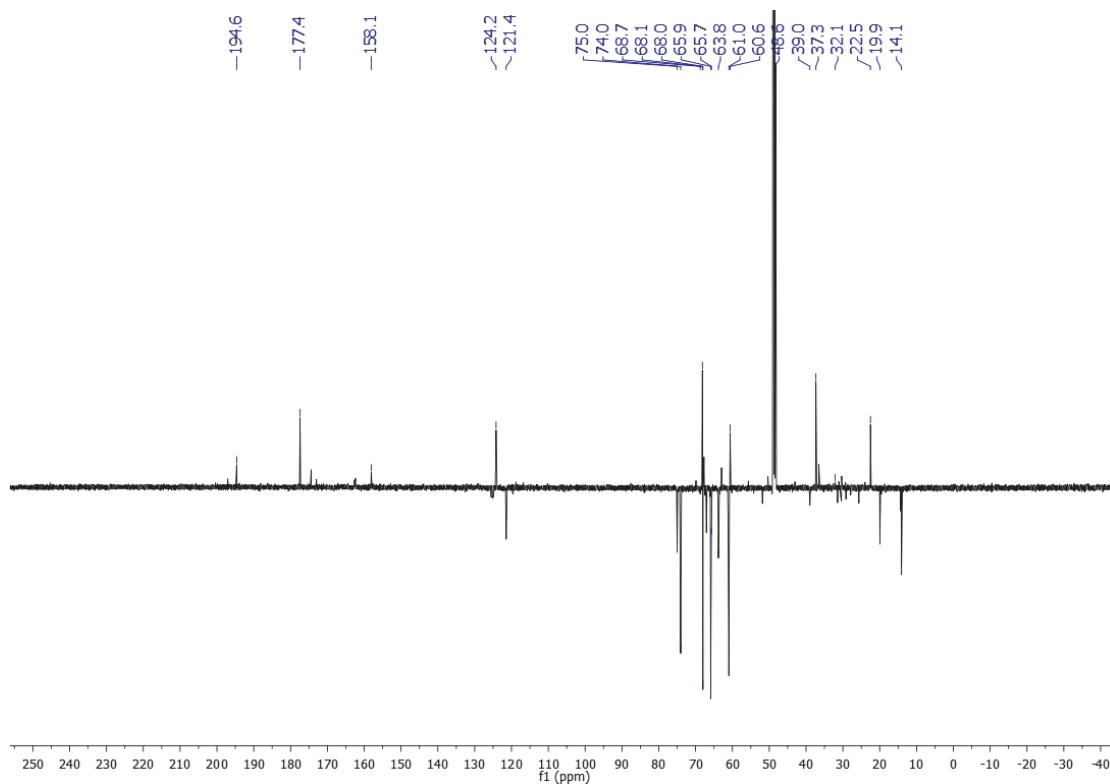
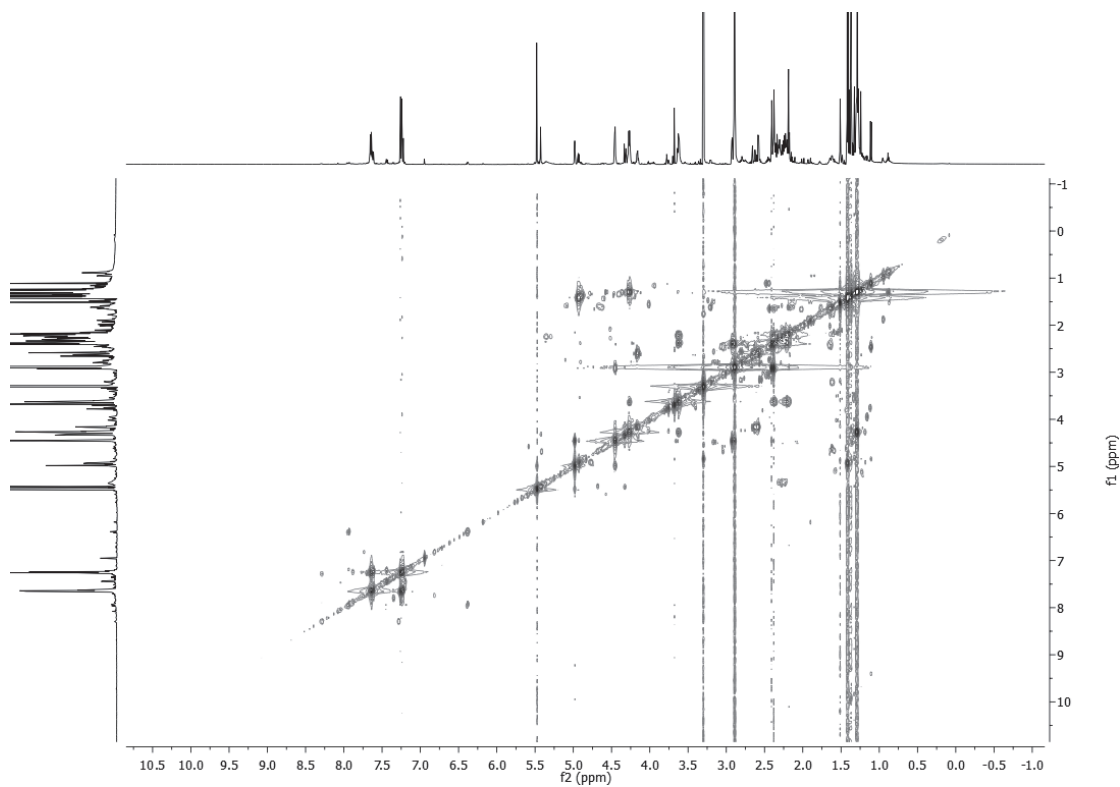
Figure S2. Observed HMBC (↷) and COSY (—) correlations of compounds 1 and 2.

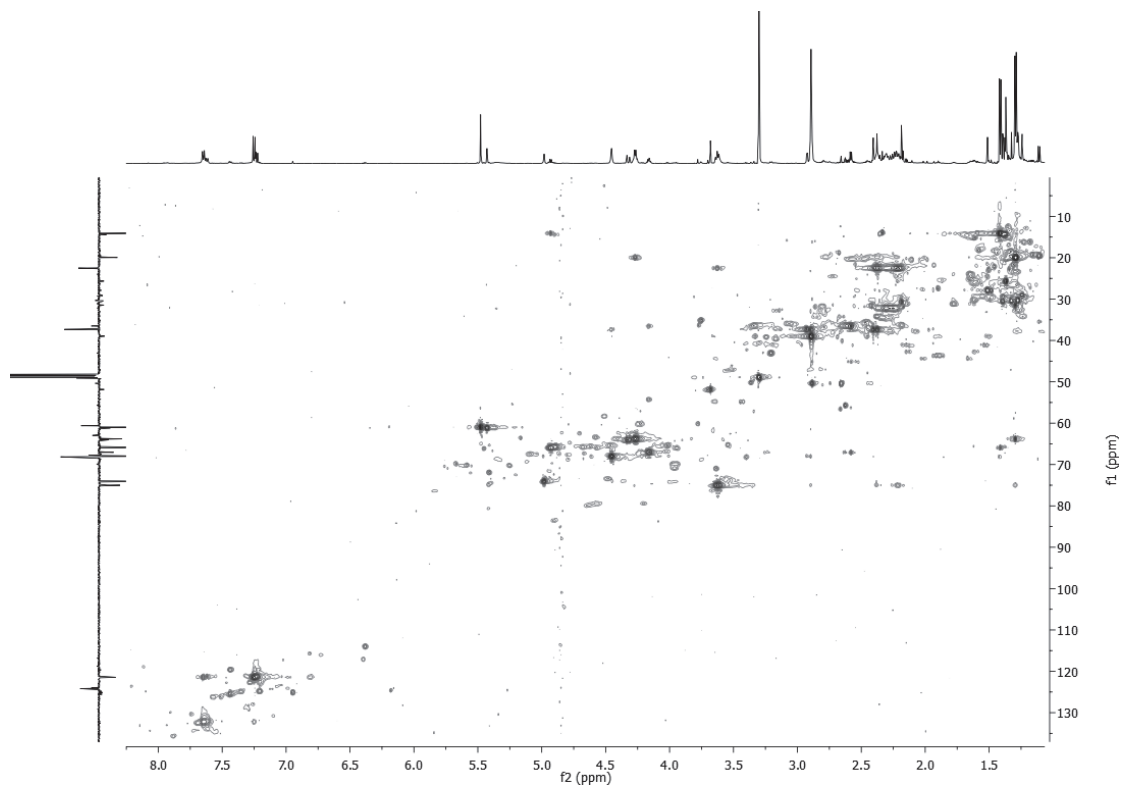
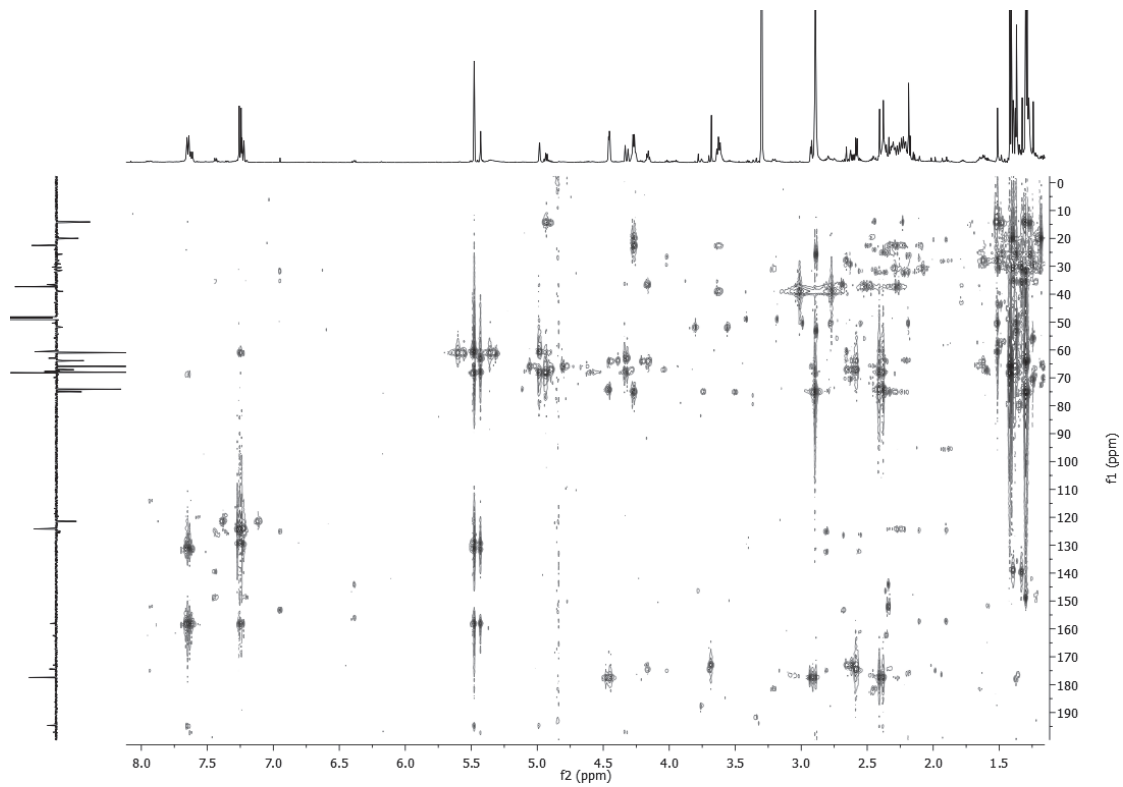
Supporting data file. NMR Spectra of Fr1 mainly consisting of compounds 1—3.

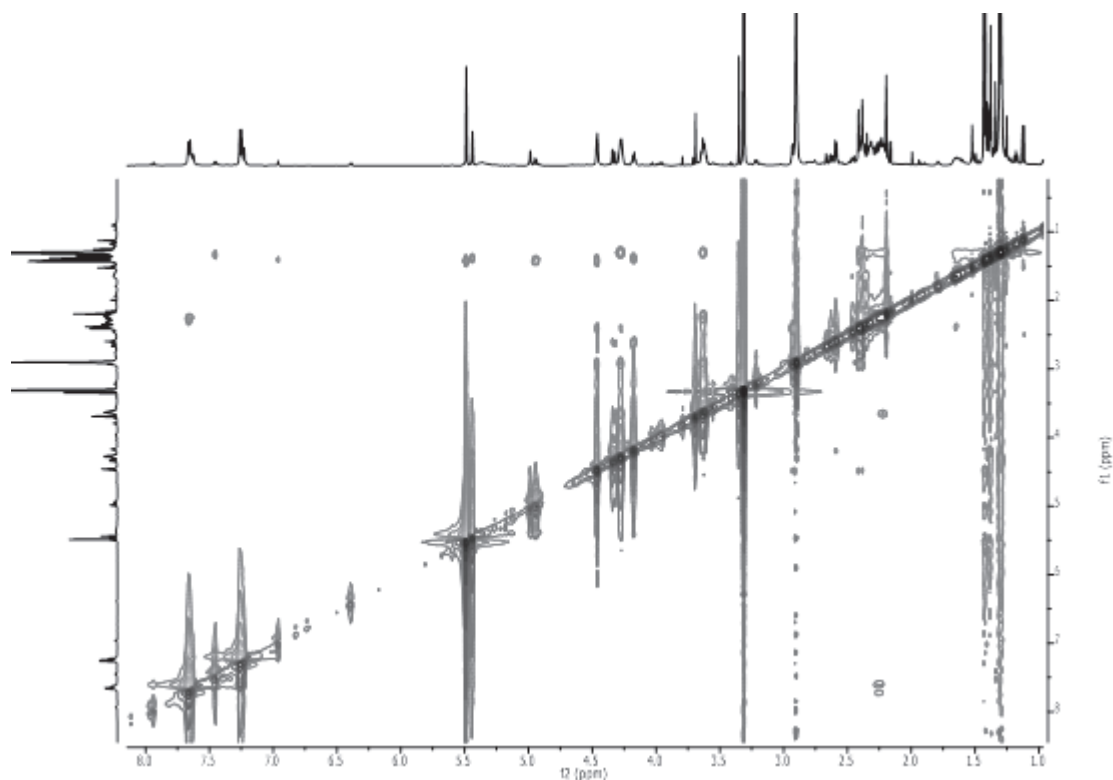
- S1. ^1H NMR spectrum of Fr1 in methanol- d_4 .
- S2. APT spectrum of Fr1 in methanol- d_4 .
- S3. ^1H - ^1H COSY spectrum of Fr1 in methanol- d_4 .
- S4. HSQC spectrum of Fr1 in methanol- d_4 .
- S5. HMBC spectrum of Fr1 in methanol- d_4 .
- S6. NOESY spectrum of Fr1 in methanol- d_4 .

S1. ^1H NMR spectrum of Fr1 in methanol- d_4 .



S2. APT spectrum of Fr1 in methanol- d_4 .S3. ^1H - ^1H COSY spectrum of Fr1 in methanol- d_4 .

S4. HSQC spectrum of Fr1 in methanol- d_4 .S5. HMBC spectrum of Fr1 in methanol- d_4 .

S6. NOESY spectrum of Fr1 in methanol- d_4 .

References

- (1) Tsukamoto, M.; Nakajima, S.; Murooka, K.; Hirayama, M.; Hirano, K.; Yoshida, S.; Kojiri, K.; Suda, H. *J. Antibiot. (Tokyo)*. **2000**, *53*, 26–32.
- (2) Bergy, M. *J. Antibiot. (Tokyo)*. **1968**, *21*, 454–457.
- (3) Kakinuma, S.; Ikeda, H.; Takada, Y.; Tanaka, H.; Hopwood, D. A.; Ōmura, S. *J. Antibiot. (Tokyo)*. **1995**, *48*, 484–487.
- (4) Taguchi, T.; Itou, K.; Ebizuka, Y.; Malpartida, F.; Hopwood, D. A.; Surti, C. M.; Booker-Milburn, K. I.; Stephenson, G. R.; Ichinose, K. *J. Antibiot. (Tokyo)*. **2000**, *53*, 144–152.
- (5) Yeo, W.; Yun, B.-S.; Kim, S.-S.; Park, E.-K.; Kim, Y.-H.; Yoo, I.-D.; Yu, S.-H. *J. Antibiot. (Tokyo)*. **1998**, *51*, 952–953.
- (6) Sun, J. L.; Wang, J. L.; Wang, L. Y.; Tao, L. M. *Chem. Bioeng.* **2008**, *25*, 43–51.
- (7) Zinad, D. S.; Shaaban, K. A.; Abdalla, M. A.; Islam, M. T.; Schöffler, A.; Laatsch, H. *Nat. Prod. Commun.* **2011**, *6*, 45–48.
- (8) Frohwein, Y. Z.; Dafni, Z.; Friedman, M.; Mateles, R. I. *Agric. Biol. Chem.* **1973**, *37*, 679–680.

Appendix V: Supplementary Information belonging to Chapter 8

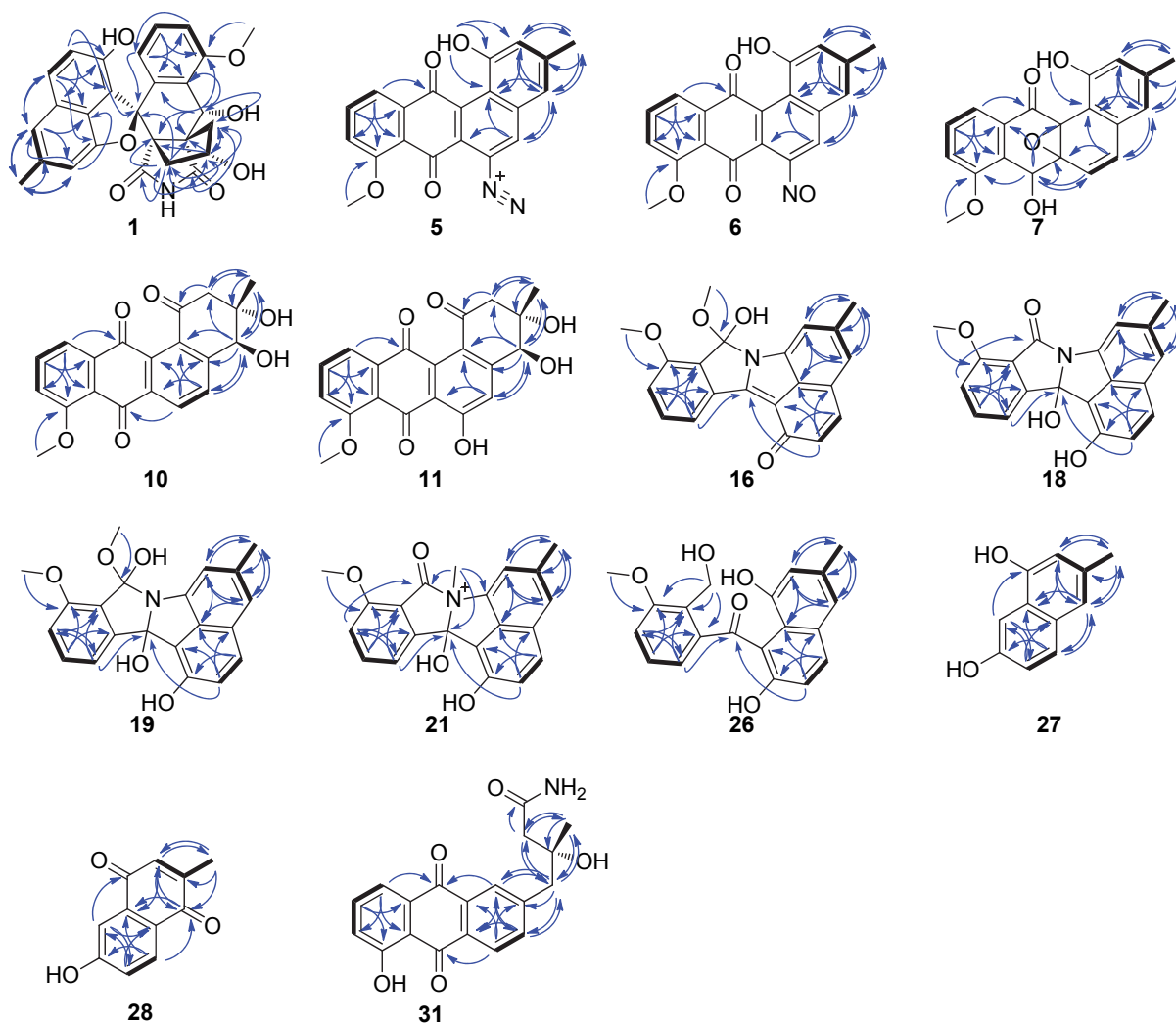


Figure S1. 2D NMR correlations of lugdunomycin (1), and selected new angucyclines 5–7, 10, 11, 16, 18, 19, 22, 27, and 31. Displayed are HMBC ($\curvearrowright$) and COSY (—) correlations for determination of planar structure.

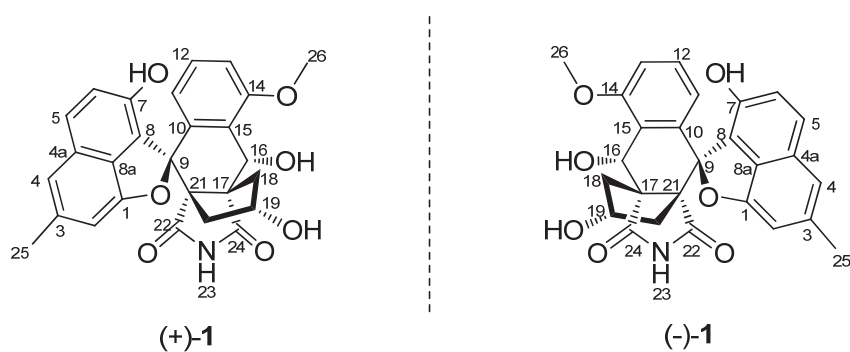


Figure S2. Structure of lugdunomycin (+)-1 and (-)-1

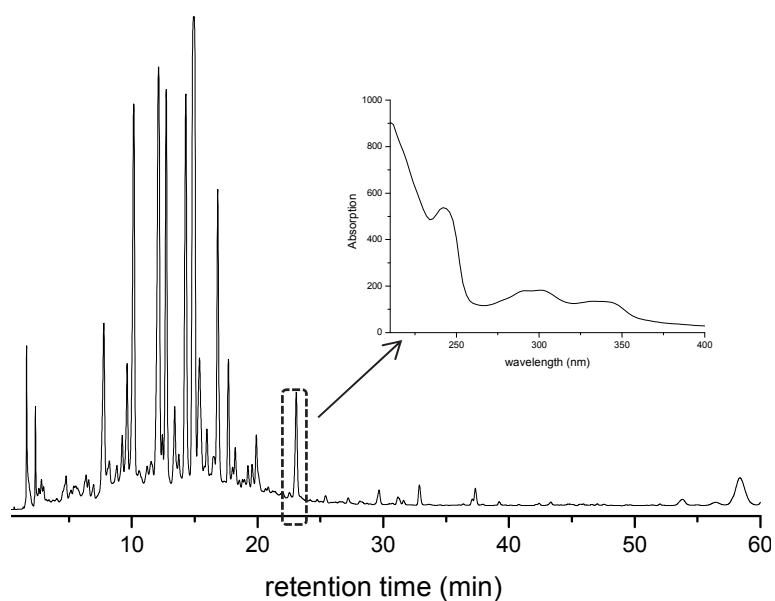


Figure S3. HPLC-DAD analysis (254 nm) of *Streptomyces sp.* QL37 grown on R5 + 0.8% peptone + 1% mannitol. The compound in dashed box gives a UV spectrum comparable to that of lugdunomycin (see S9).

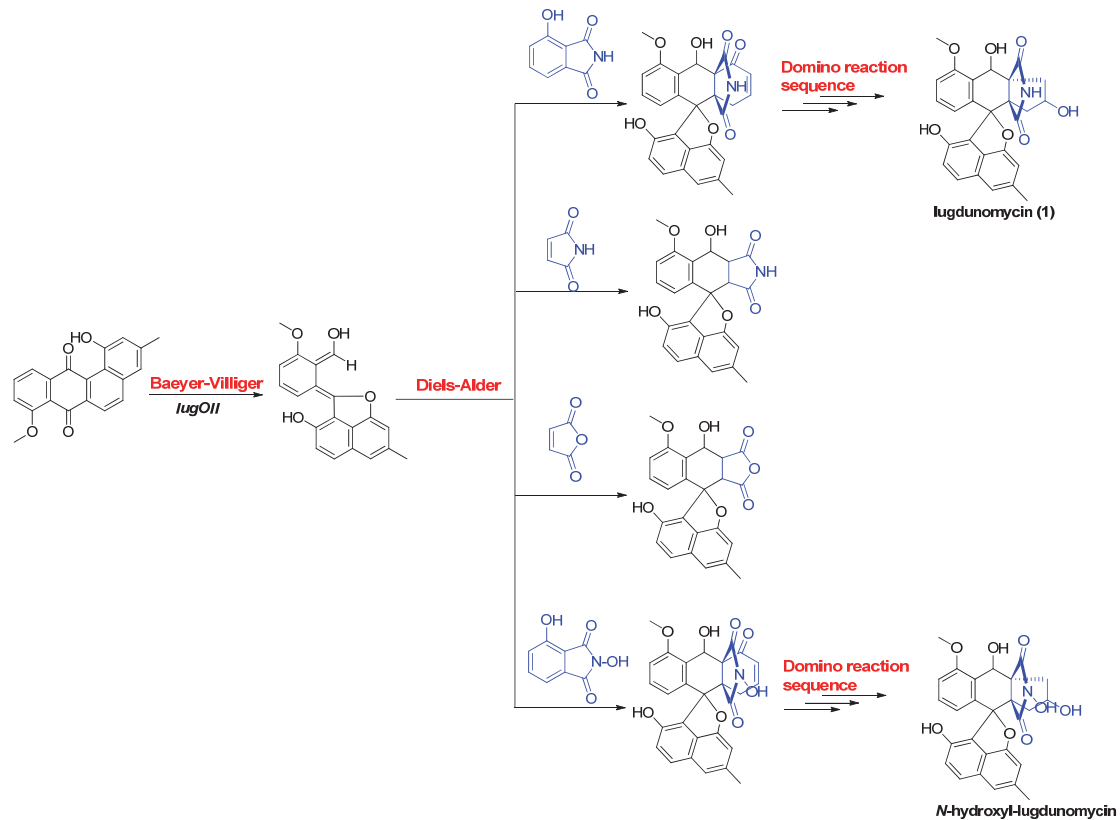


Figure S4. Schematic illumination of potential chemoenzymatic approach to synthesize lugdunomycin variants.

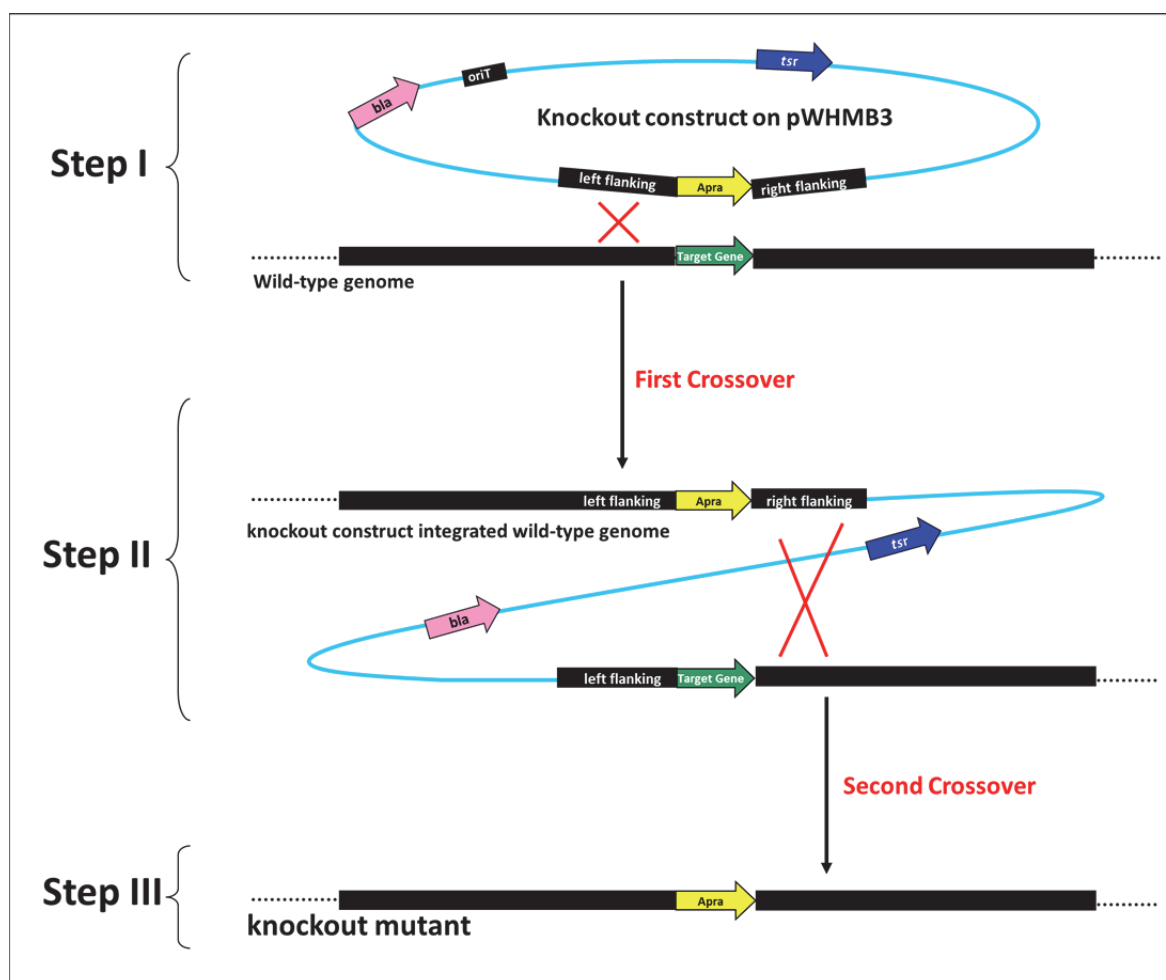


Figure S5. Schematic explanation of deletion mutants for genes *lugA–lugC*, *lugOII*, *lugOIV*, and *lugS*. Plasmid pWHM3 engineered with *oriT* fragment was used for conjugation .

Table S1. Crystal Data for lugdunomycin (1)

Identification code	Lugdunomycin	F(000)	1984
Empirical formula	C ₂₇ H ₂₃ NO ₇	Crystal size	0.7 x 0.5 x 0.5 mm ³
Formula weight	473.46	Theta range for data collection	1.241 to 25.340°
Temperature	100(2) K	Index ranges	-27<=h<=27, -27<=k<=27, -9<=l<=9
Wavelength	0.71073 Å	Reflections collected	31098
Crystal system	Tetragonal	Independent reflections	4053 [<i>R</i> _{int} = 0.1666]
Space group	P 42/n	Completeness to theta = 25.242°	100.0 %
Unit cell dimensions	a = 23.204(2) Å, α = 90°.	Refinement method	Full-matrix least-squares on F ²
	b = 23.204(2) Å, β = 90°.	Data/restraints/parameters	4053 / 330 / 321
	c = 8.2073(9) Å, γ = 90°.	Goodness-of-fit on F²	1.036
Volume	4419.1(10) Å ³	Final R indices [<i>I</i>>2σ(<i>I</i>)]	<i>R</i> ₁ = 0.0615, w <i>R</i> ₂ = 0.1134
Z	8	R indices (all data)	<i>R</i> ₁ = 0.1245, w <i>R</i> ₂ = 0.1379
Density (calculated)	1.423 Mg/m ³	Extinction coefficient	n/a
Absorption coefficient	0.104 mm ⁻¹	Largest diff. peak and hole	0.287 and -0.236 e.Å ⁻³

Table S2. Culture media with different components combination for lugdunomycin production.

No.	Media Combination	No.	Media Combination
1	NMMP + 0.2% L-asparagine + 50 mM TES buffer + 2% glycerol	40	MM + 0.5% yeast extract + 50mM GluNAc
2	NMMP + 0.2% L-asparagine + 50 mM TES buffer + 1% glucose	41	MM + 0.8% peptone + 1% mannitol
3	NMMP + 0.2% L-asparagine + 50 mM TES buffer + 1% mannitol	42	MM + 0.8% peptone + 1% glucose
4	NMMP + 0.2% L-asparagine + 50 mM TES buffer + 1% xylose	43	MM + 0.8% peptone + 1% glycerol
5	NMMP + 0.2% L-asparagine + 50 mM TES buffer + 1% GluNAc	44	MM + 0.8% peptone + 50mM GluNAc
6	NMMP + 0.2% (NH ₄) ₂ SO ₄ + 50 mM TES buffer + 1% glycerol	45	MM + 1% mannitol
7	NMMP + 0.2% (NH ₄) ₂ SO ₄ + 50 mM TES buffer + 1% glucose	46	MM + 1% glucose
8	NMMP + 0.2% (NH ₄) ₂ SO ₄ + 50 mM TES buffer + 1% mannitol	47	MM + 1% glycerol
9	NMMP + 0.2% (NH ₄) ₂ SO ₄ + 50 mM TES buffer + 1% xylose	48	MM + 50mM GluNAc
10	NMMP + 0.2% (NH ₄) ₂ SO ₄ + 50 mM TES buffer + 1% GluNAc	49	MM + 25mM NaBu + 1% mannitol
11	NMMP + 0.2% L-glutamine + 50 mM TES buffer + 1% glycerol	50	MM + 25mM NaBu + 1% glucose
12	NMMP + 0.2% L-glutamine + 50 mM TES buffer + 1% glucose	51	MM + 25mM NaBu + 1% glycerol
13	NMMP + 0.2% L-glutamine + 50 mM TES buffer + 1% mannitol	52	MM + 25mM NaBu + 50mM GluNAc
14	NMMP + 0.2% L-glutamine + 50 mM TES buffer + 1% xylose	53	R5+ proline + NaOH + CaCl ₂ + 1% glucose
15	NMMP + 0.2% L-glutamine + 50 mM TES buffer + 1% GluNAc	54	R5+ proline + NaOH + CaCl ₂ + 1% glycerol
16	NMMP + 0.2% L-glutamine + 50 mM TES buffer + 1% fructose	55	R5+ proline + NaOH + CaCl ₂ + 1% mannitol
17	NMMP + 0.2% L-glutamine + 50 mM TES buffer + 1% maltose	56	R5+ proline + NaOH + CaCl ₂ + 1% 50mM GluNAc
18	NMMP + 0.2% L-arginine + 50 mM TES buffer + 1% glycerol	57	R5+ proline + NaOH + CaCl ₂ + without sugar
19	NMMP + 0.2% L-arginine + 50 mM TES buffer + 1% glucose	58	R5 + 1% glucose
20	NMMP + 0.2% L-arginine + 50 mM TES buffer + 1% mannitol	59	R5+ 1% glycerol
21	NMMP + 0.2% L-arginine + 50 mM TES buffer + 1% xylose	60	R5+ 1% mannitol
22	NMMP + 0.2% L-arginine + 50 mM TES buffer + 1% GluNAc	61	R5 + 50mM GluNAc
23	NMMP + 0.2% L-arginine + 50 mM TES buffer + 1% fructose	62	R5 + without sugar
24	NMMP + 0.2% L-arginine + 50 mM TES buffer + 1% maltose	63	R5 + 200mM NaCl+ 1% glucose
25	NMMP + 0.2% L-proline + 50 mM TES buffer + 1% glycerol	64	R5 + 200mM NaCl+ 1% glycerol
26	NMMP + 0.2% L-proline + 50 mM TES buffer + 1% glucose	65	R5 + 200mM NaCl+ 1% mannitol
27	NMMP + 0.2% L-proline + 50 mM TES buffer + 1% mannitol	66	R5 + 200mM NaCl + 50mM GluNAc
28	NMMP + 0.2% L-proline + 50 mM TES buffer + 1% xylose	67	R5 + 200mM NaCl + without sugar
29	NMMP + 0.2% L-proline + 50 mM TES buffer + 1% GluNAc	68	R5 + 0.8% peptone+ 1% glucose
30	NMMP + 0.2% L-proline + 50 mM TES buffer + 1% fructose	69	R5 + 0.8% peptone + 1% glycerol
31	NMMP + 0.2% L-proline + 50 mM TES buffer + 1% maltose	70	R5 + 0.8% peptone + 1% mannitol
32	NMMP + 0.2% (NH ₄) ₂ SO ₄ + phosphate buffer + 1% glycerol + 0.5% mannitol	71	R5 + 0.8% peptone + 50mM GluNAc
33	NMMP + 0.2% (NH ₄) ₂ SO ₄ + phosphate buffer + 1% glycerol + 0.5% mannitol + 1% SFM	72	R5 + 0.8% peptone + without sugar
34	NMMP + 0.2% (NH ₄) ₂ SO ₄ + phosphate buffer + 1% glycerol + 0.5% mannitol + 0.5% yeast extract	73	R5 + 0.5% Yeast extract+ 1% glucose
35	NMMP + 0.2% (NH ₄) ₂ SO ₄ + phosphate buffer + 1% glycerol + 0.5% mannitol + 1% peptone	74	R5 + 0.5% Yeast extract + 1% glycerol
36	NMMP + 0.2% (NH ₄) ₂ SO ₄ + phosphate buffer + 1% glycerol + 0.5% mannitol + pH 10	75	R5 + 0.5% Yeast extract + 1% mannitol
37	MM + 0.5% yeast extract + 1% mannitol	76	R5 + 0.5% Yeast extract + 50mM GluNAc
38	MM + 0.5% yeast extract + 1% glucose	77	R5 + 0.5% Yeast extract + without sugar
39	MM + 0.5% yeast extract + 1% glycerol		

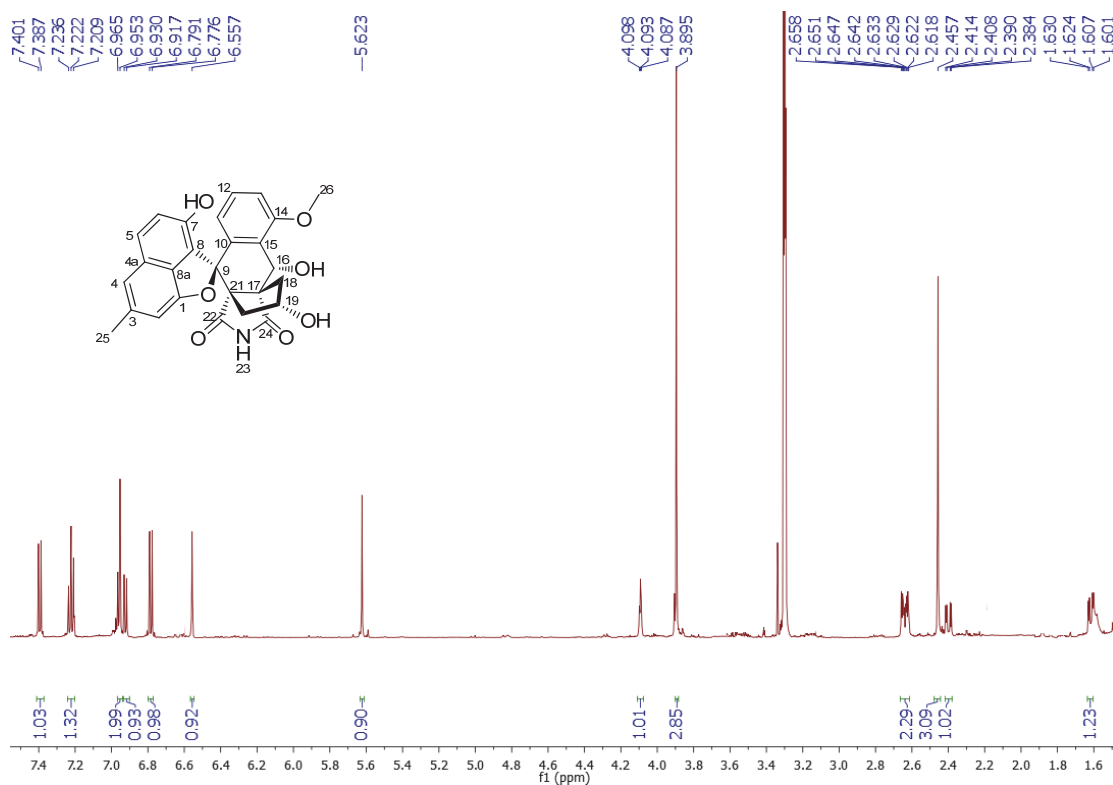
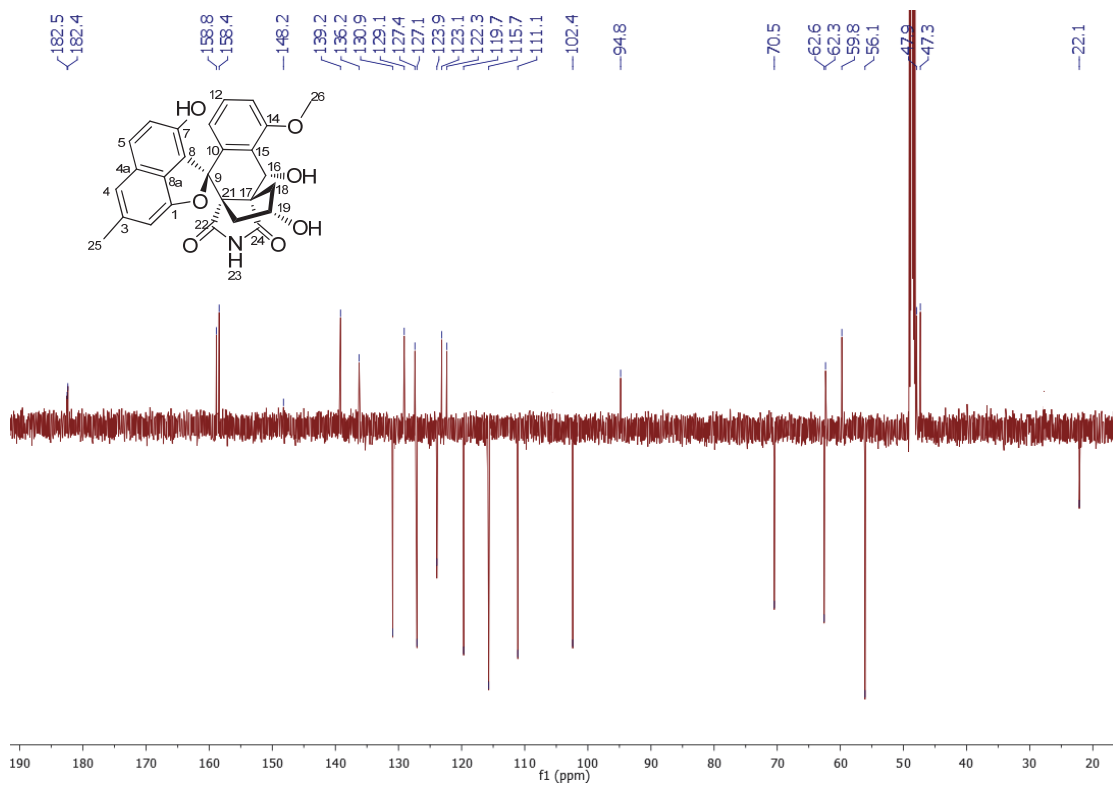
Table S3. Primers used for gene knockout of genes *lugA*–*lugC*, *lugOII*, *lugOIV*, and *lugS*

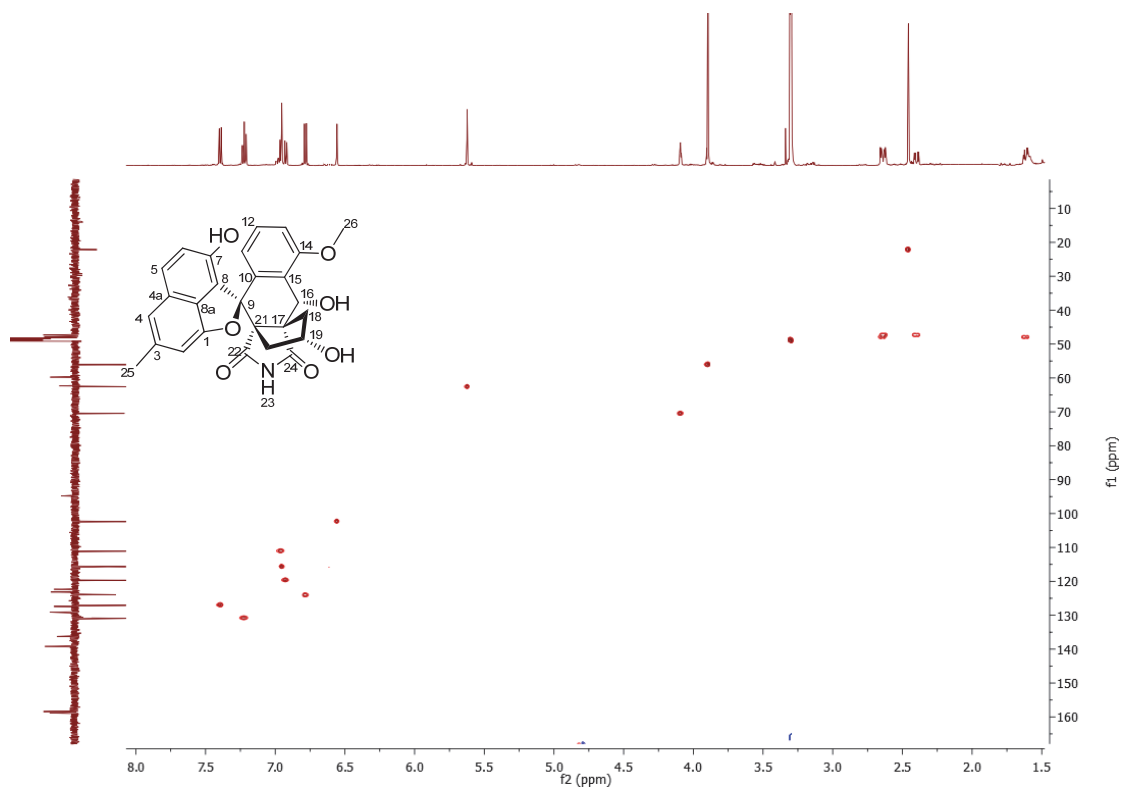
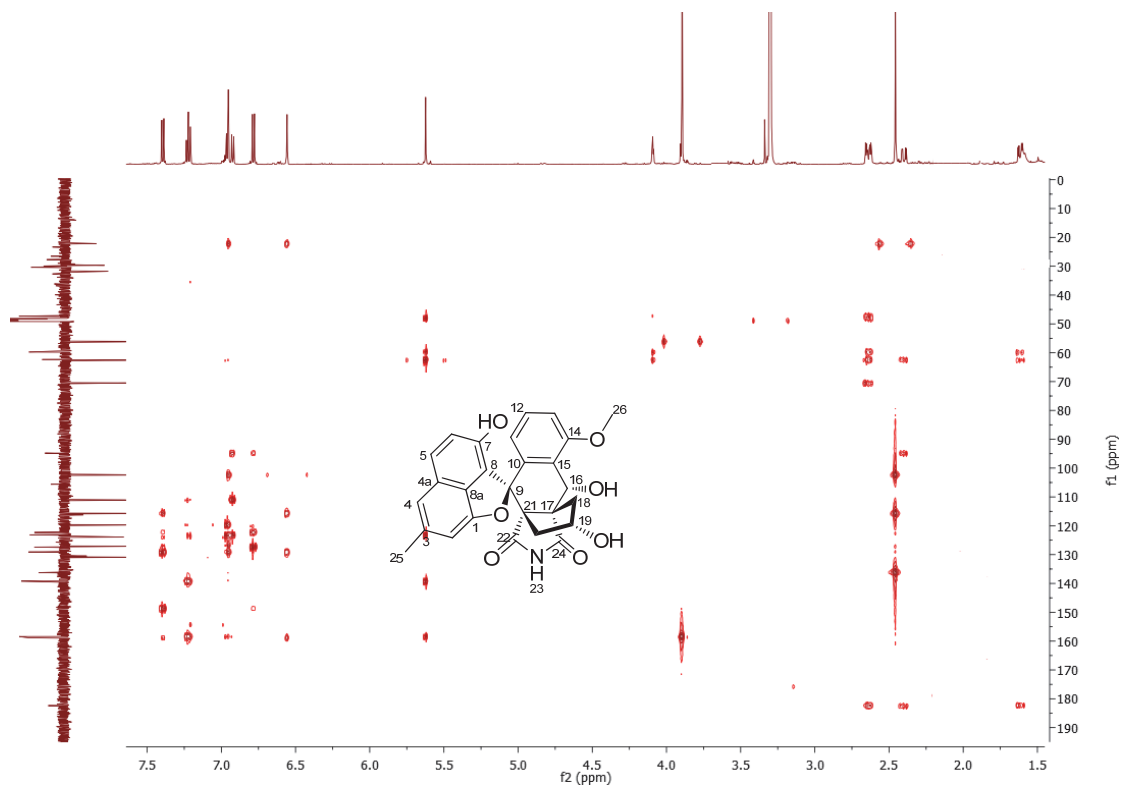
primer	sequence (5'→3') [#]	Function
MinPKS_LF_Fw	ctag <u>GAATTC</u> CCCGCCACCACCGAGCTCTTC	Forward primer for PCR of left flanking region of <i>lugA-C</i>
MinPKS_LF_Rv	GAAGTTATCCATCACCT <u>TCTAGA</u> GATACCGGTGATGACGACCC	Reverse primer for PCR of left flanking region of <i>lugA-C</i>
MinPKS_RF_Fw	GAAGTTATCGCGCATCT <u>TCTAGA</u> GCCGAGCAGCTCGACCGTTAC	Forward primer for PCR of right flanking region of <i>lugA-C</i>
MinPKS_RF_Rv	ctag <u>GGATCC</u> CTGCCCTTGTCTGAGAAGCAGTG	Reverse primer for PCR of right flanking region of <i>lugA-C</i>
LugOII_LF_Fw	cgat <u>AAGCTT</u> GGTACCGAGCTGTGGCTGGACGTGATCAAC	Forward primer for PCR of left flanking region of <i>lugOII</i>
LugOII_LF_Rv	cgat <u>TCTAGA</u> CGTGGTGCCACGGGTCAGC	Reverse primer for PCR of left flanking region of <i>lugOII</i>
LugOII_RF_Fw	cgat <u>TCTAGA</u> GGCGGGACGCTCCTCGGATG	Forward primer for PCR of right flanking region of <i>lugOII</i>
LugOII_RF_Rv	cgat <u>GAATTC</u> GGTGAGGGCCGGCGAGTAGG	Reverse primer for PCR of right flanking region of <i>lugOII</i>
LugOII_Chk_Fw	TCGACCCACACCGGAATC	Forward primer for PCR check of <i>lugOII</i> knockout mutant
LugOII_Chk_Rv	TTCCTCGCCGATGTGCTTGG	Reverse primer for PCR check of <i>lugOII</i> knockout mutant
LugOIV_LF_Fw	ctag <u>GAATTC</u> CACCGGGTGCCGAGGTAGC	Forward primer for PCR of left flanking region of <i>lugOIV</i>
LugOIV_LF_Rv	ctag <u>TCTAGA</u> TCGACGAGCTGACCCGAGAAC	Reverse primer for PCR of left flanking region of <i>lugOIV</i>
LugOIV_RF_Fw	ctag <u>TCTAGA</u> TTCCGACGCGGTGCGGATGAG	Forward primer for PCR of right flanking region of <i>lugOIV</i>
LugOIV_RF_Rv	ctag <u>GGATCC</u> GGCCGCCAGATGAAGGGTG	Reverse primer for PCR of right flanking region of <i>lugOIV</i>
LugOIV_Chk_Fw	GTATCCGGCCCGCATGCCTC	Forward primer for PCR check of <i>lugOIV</i> knockout mutant
LugOIV_Chk_Rv	CCCTCGCCGACGCACACTTC	Reverse primer for PCR check of <i>lugOIV</i> knockout mutant
LugS_LF_Fw	cgatAAGCTTGAATTCCCGCTACATGGACCACTAC	Forward primer for PCR of left flanking region of <i>lugS</i>
LugS_LF_Rv	cgat <u>TCTAGA</u> CATGCGGGTGCTCCATGTTT	Reverse primer for PCR of left flanking region of <i>lugS</i>
LugS_RF_Fw	cgat <u>TCTAGA</u> CGCGTCTGATCCGCTCTCC	Forward primer for PCR of right flanking region of <i>lugS</i>
LugS_RF_Rv	cgat <u>GAATTC</u> CCCGCGTACCAAGTTGGACGAC	Reverse primer for PCR of right flanking region of <i>lugS</i>
LugS_Chk_Fw	GAAGATCTCCGGCATGGAGG	Forward primer for PCR check of <i>lugS</i> knockout mutant
LugS_Chk_Rv	GGGAGAGGTCTGGATCTGGG	Reverse primer for PCR check of <i>lugS</i> knockout mutant

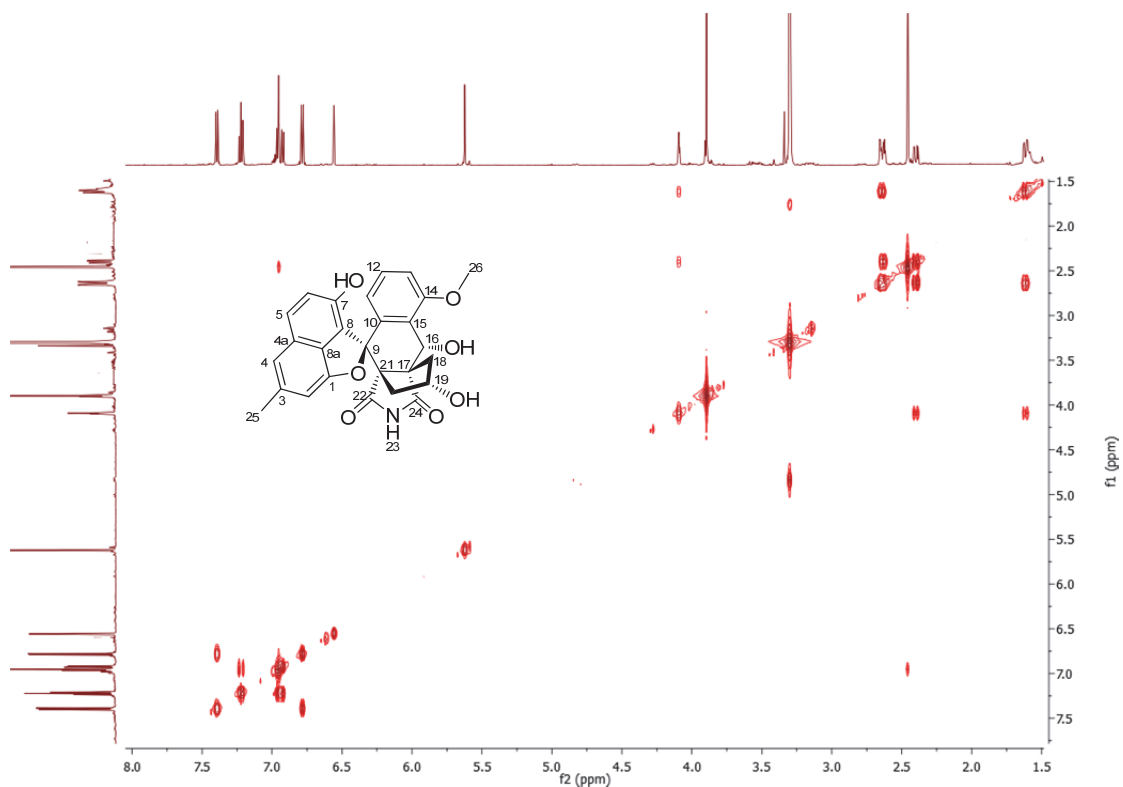
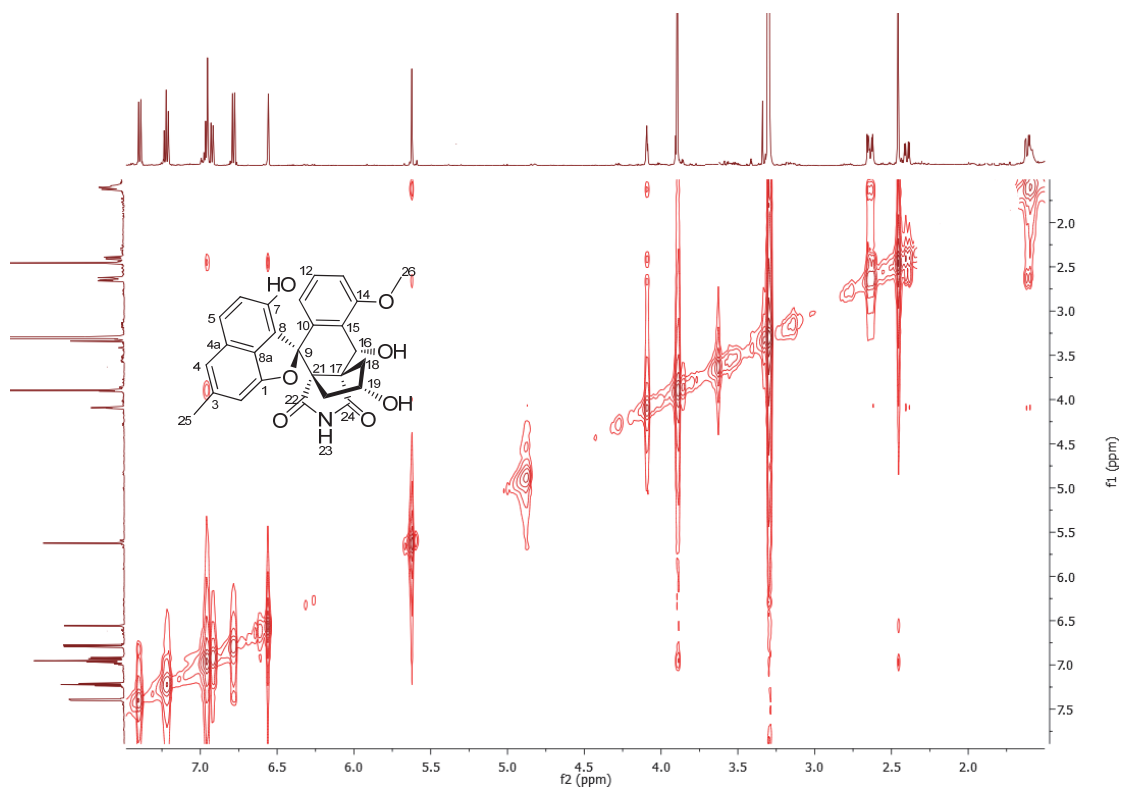
[#]Restriction sites used for cloning are underlined and in bold face. GAATTC, EcoRI; TCTAGA, XbaI; AAGCTT, HindIII; GGATCC, BamHI.

Spectra list of lugdunomycin (1).

- S1. ¹H NMR spectrum (600 MHz) of lugdunomycin (1) in CD₃OD.
- S2. APT spectrum (600 MHz) of lugdunomycin (1) in CD₃OD.
- S3. HSQC spectrum (600 MHz) of lugdunomycin (1) in CD₃OD.
- S4. HMBC spectrum (600 MHz) of lugdunomycin (1) in CD₃OD.
- S5. ¹H-¹H COSY spectrum (600 MHz) of lugdunomycin (1) in CD₃OD.
- S6. ¹H-¹H NOSEY spectrum (600 MHz) of lugdunomycin (1) in CD₃OD.
- S7. HRESIMS spectrum of lugdunomycin (1).
- S8. IR spectrum of lugdunomycin (1).
- S9. UV spectrum of lugdunomycin (1).
- S10. CD spectrum of lugdunomycin (1).

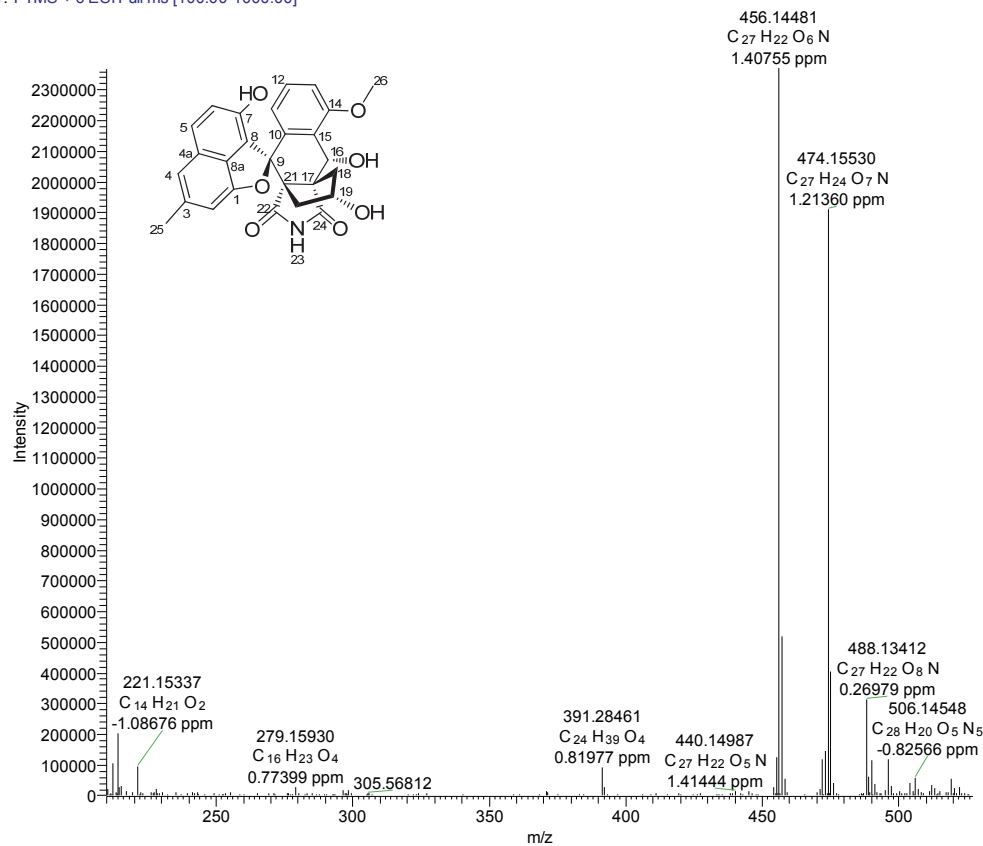
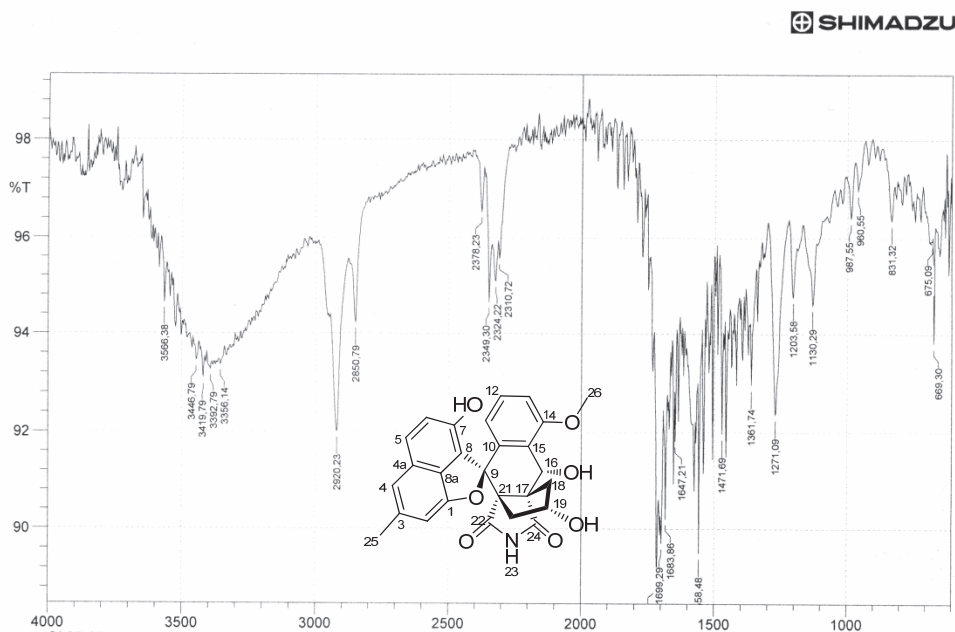
S1. ^1H NMR spectrum (600 MHz) of lugdunomycin (**1**) in CD_3OD .S2. APT spectrum (150 MHz) of lugdunomycin (**1**) in CD_3OD .

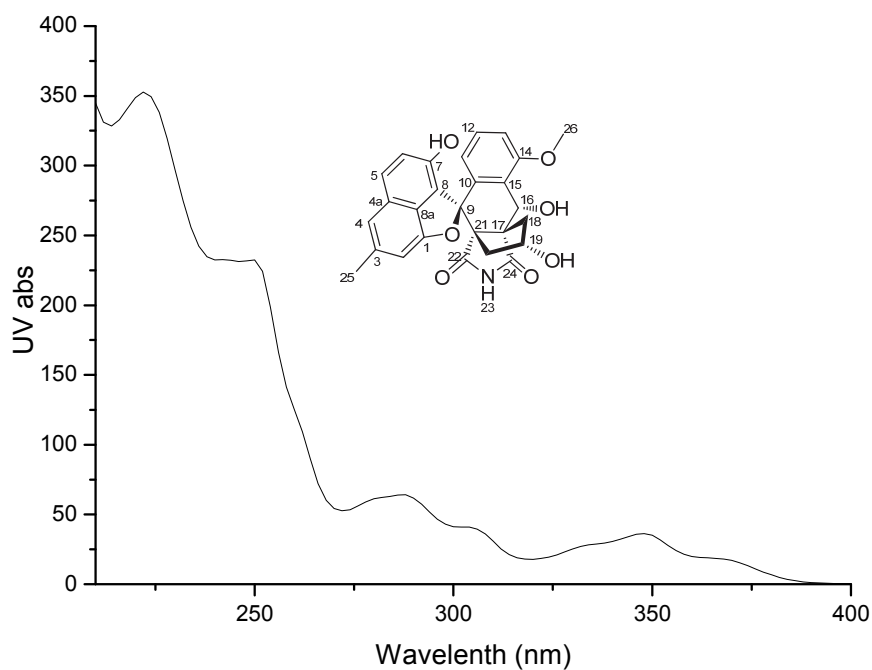
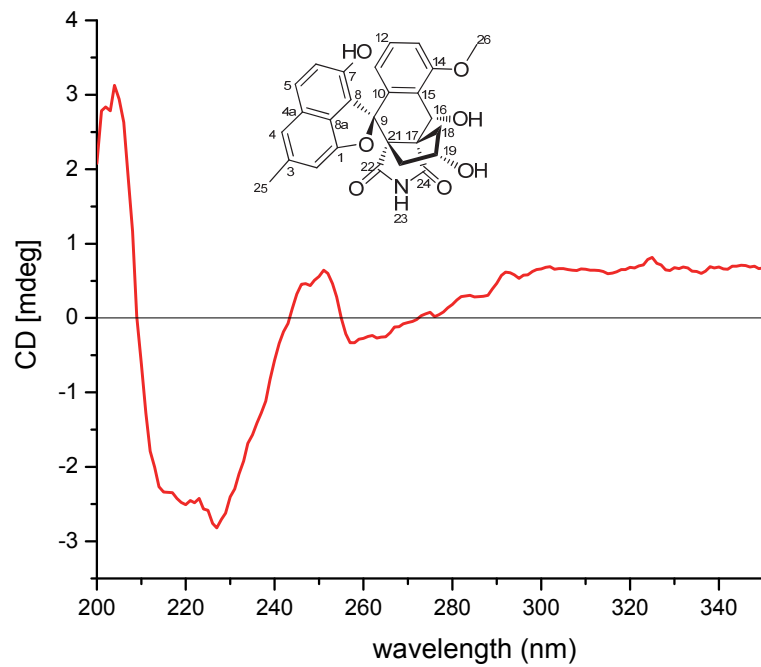
S3. HSQC spectrum (600 MHz) of lugdunomycin (**1**) in CD₃OD.S4. HMBC spectrum (600 MHz) of lugdunomycin (**1**) in CD₃OD.

S5. ^1H - ^1H COSY spectrum (600 MHz) of lugdunomycin (**1**) in CD_3OD .S6. ^1H - ^1H NOESY spectrum (600 MHz) of lugdunomycin (**1**) in CD_3OD .

S7. HRESIMS spectrum of lugdunomycin (**1**).

QL37-15 #741 RT: 12.82 AV: 1 NL: 2.37E6
T: FTMS + c ESI Full ms [100.00-1000.00]

S8. IR spectrum of lugdunomycin (**1**).

S9. UV spectrum of lugdunomycin (**1**).S10. CD spectrum of lugdunomycin (**1**).

Appendix VI: Supplementary Information belonging to Chapter 9

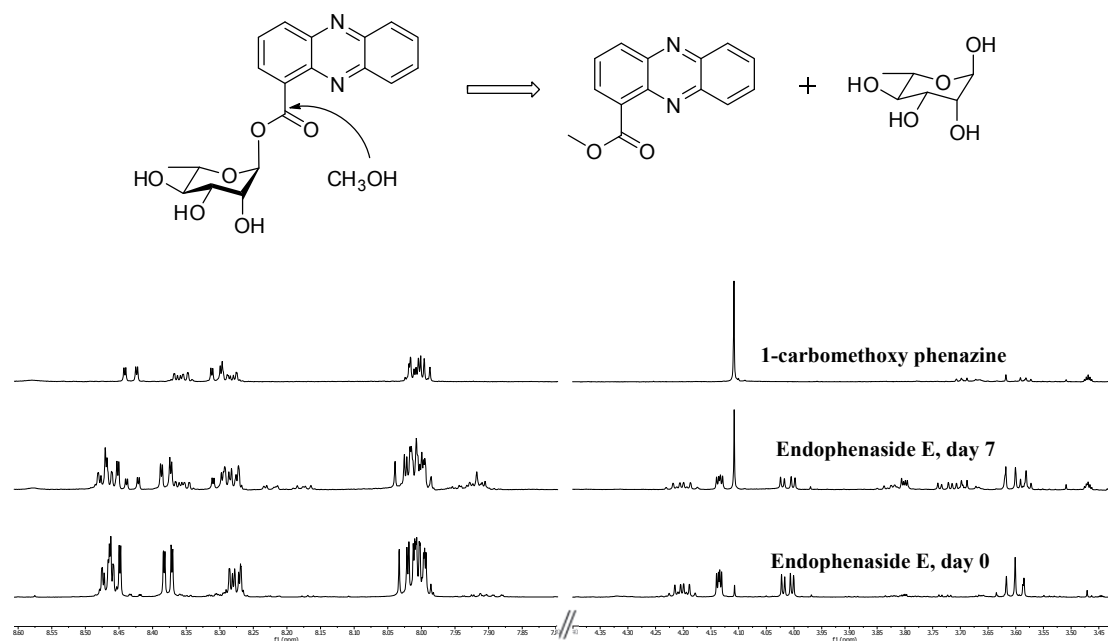


Figure S1. Instability of endophenaside E in methanol. Endophenaside E was unstable in methanol, which was prone to nucleophilic reaction (methanolysis) and degraded into 1-carbomethoxy phenazine and rhamnose. The reaction was observed in deuterated methanol when measuring NMR data in one week. The product identification was done in comparison with the ¹H NMR of purified 1-carbomethoxy phenazine.

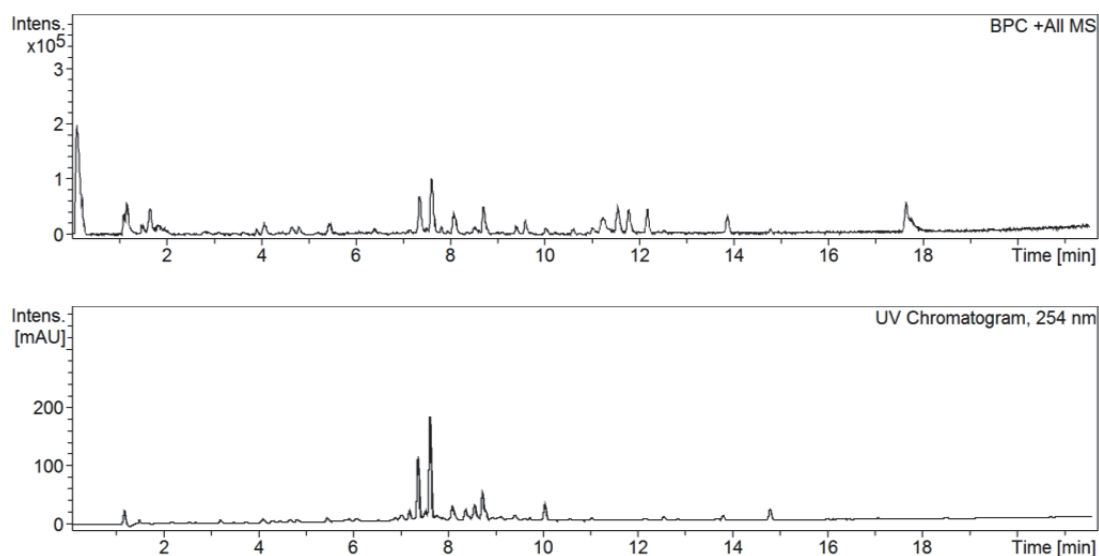


Figure S2. UHPLC-Q-TOF analysis of *Kitasatospora* sp. MBT66. UHPLC-Q-TOF profiling of secondary metabolites produced by *kitasatospora* sp. MBT66. A) mass detection, and B) UV detection at 254 nm.

Table S1. Extracted peaks and ions of UHPLC-Q-TOF chromatogram for the confirmation of previously identified endophenazines.

Retention time	Detected mass (m/z)	Deduced chemical formula	Corresponding compound described by Heine D. <i>et al</i>
8.5 min	285.0879 [M + H] ⁺	C ₁₅ H ₁₃ N ₂ O ₄	5-(2-Hydroxyacetyl)-5,10-dihydrophenazine-1-carboxylic acid (2)
8.8 min	309.1231 [M + H] ⁺	C ₁₈ H ₁₇ N ₂ O ₃	(<i>E</i>)-9-(4-Hydroxy-3-methylbut-2-en-1-yl)phenazine-1-carboxylic acid (3)
9.4 min	309.1234 [M + H] ⁺	C ₁₈ H ₁₇ N ₂ O ₃	(<i>E</i>)-6-(3-Hydroxy-3-methylbut-1-en-1-yl)phenazine-1-carboxylic acid (5)
14.1 min	377.1853 [M + H] ⁺	C ₂₃ H ₂₅ N ₂ O ₃	5,9-bis(3-Methylbut-2-en-1-yl)-7-oxo-5,7-dihydrophenazine-1-carboxylic acid (4)

Physical properties of the new compounds

Endophenaside A (**1**): yellowish, amorphous powder, UV (MeOH) $\lambda_{\max}$ (log ϵ) 251 (4.79), 366 (4.17) nm; IR 3352, 2926, 2855, 1674, 1597, 1560, 1456, 1204, 16, 1067, 958, 748 $\nu_{\max}$ cm⁻¹; ¹H and ¹³C NMR data, see Table 1 and 2; HRMS (positive mode) m/z 400.15076 [M + H]⁺ (calcd for C₂₀H₂₂N₃O₆ 400.15031), 422.13339 [M + Na]⁺ (calcd for C₂₀H₂₁N₃O₆Na 422.13281).

Endophenaside B (**2**): yellowish, amorphous powder, UV (MeOH) $\lambda_{\max}$ (log ϵ) 252 (4.72), 365 (4.10) nm; IR $\nu_{\max}$ 3350, 2920, 1732, 1672, 1531, 85, 1263, 1123, 1061, 978, 947, 760 cm⁻¹; ¹H and ¹³C NMR data, see Table 1 and 2; HRMS (positive mode) m/z 455.18187 [M + H]⁺ (calcd for C₂₄H₂₇N₂O₇ 455.18128), 477.16375 [M + Na]⁺ (calcd for C₂₄H₂₇N₂O₇Na 477.16377).

Endophenaside C (**3**): yellowish, amorphous powder, UV (MeOH) $\lambda_{\max}$ (log ϵ) 251 (4.50), 367 (3.97) nm; IR $\nu_{\max}$ 3341, 2926, 2853, 1682, 1456, 1418, 1271, 1203, 18, 1067, 841, 764 cm⁻¹; ¹H and ¹³C NMR data, see Table 1 and 2; HRMS (positive mode) m/z 469.19641 [M + H]⁺ (calcd for C₂₅H₂₉N₂O₇ 469.19693), 491.18021 [M + Na]⁺ (calcd for C₂₅H₂₈N₂O₇Na 491.17942).

Endophenaside D (**4**): yellowish, amorphous powder, UV (MeOH) $\lambda_{\max}$ (log ϵ) 251 (4.58), 368 (4.05) nm; IR $\nu_{\max}$ 3400, 2926, 2855, 1734, 1717, 1682, 1456, 77, 1277, 1126, 1069, 976, 762 cm⁻¹; ¹H and ¹³C NMR data, see Table 1 and 2; HRMS (positive mode) m/z 455.18243 [M + H]⁺ (calcd for C₂₄H₂₇N₂O₇ 455.18128), 477.16458 [M + Na]⁺ (calcd for C₂₄H₂₆N₂O₇Na 477.16377), 493.13879 [M + K]⁺ (calcd for C₂₄H₂₆N₂O₇K 493.13771).

Endophenaside E (**5**): yellowish, amorphous powder, UV (MeOH) $\lambda_{\max}$ (log ϵ) 248 (5.53), 365 (3.91) nm; IR $\nu_{\max}$ 3358, 2924, 2853, 1721, 1670, 1522, 1422, 1277, 1125, 1059, 941, 831, 750 cm⁻¹; ¹H and ¹³C NMR data, see Table 1 and 2; HRMS (positive mode) m/z 371.1233 [M + H]⁺ (calcd for C₁₉H₁₉N₂O₆ 371.1238), 393.1062 [M + Na]⁺ (calcd for C₁₉H₁₈N₂O₆Na 393.1063), 225.0657 [M + H - sugar]⁺ (calcd for C₁₃H₉N₂O₂ 225.0659).

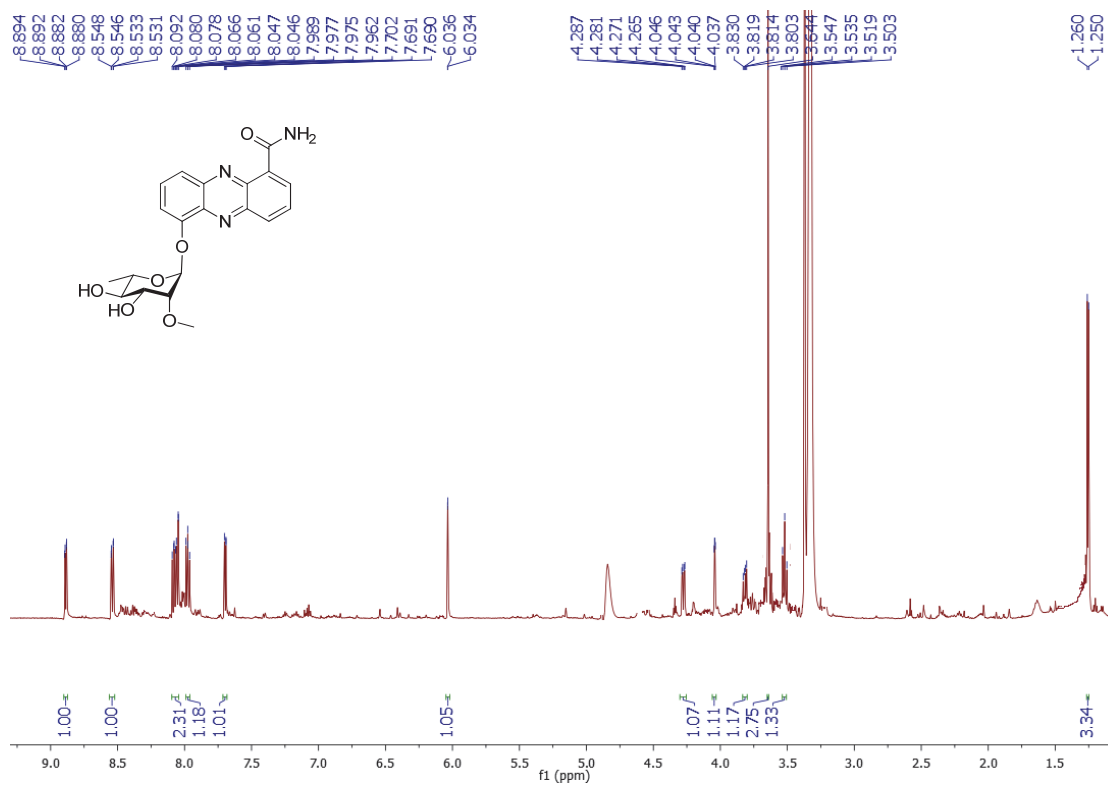
Endophenazine F1 (**6**): purple, amorphous powder, UV/vis (MeOH) $\lambda_{\max}$ (log ϵ) 283 (4.33), 376 (3.83), 518 (3.79) nm; IR $\nu_{\max}$ 3348, 2920, 2855, 1732, 1717, 1684, 1624, 1539, 1456, 96, 1233, 1206, 745 cm⁻¹; ¹H and ¹³C NMR data, see Table 1 and 2; HRMS (positive mode) m/z 393.18115 [M + H]⁺ (calcd for C₂₃H₂₅N₂O₄ 393.18088).

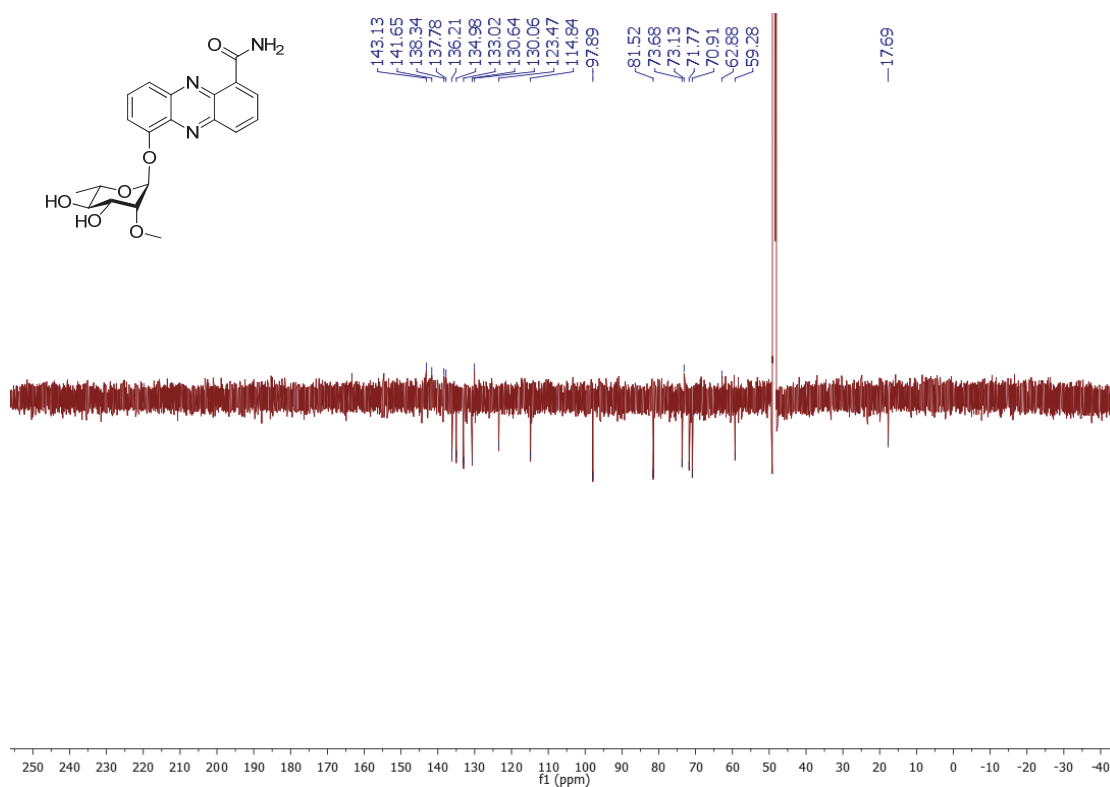
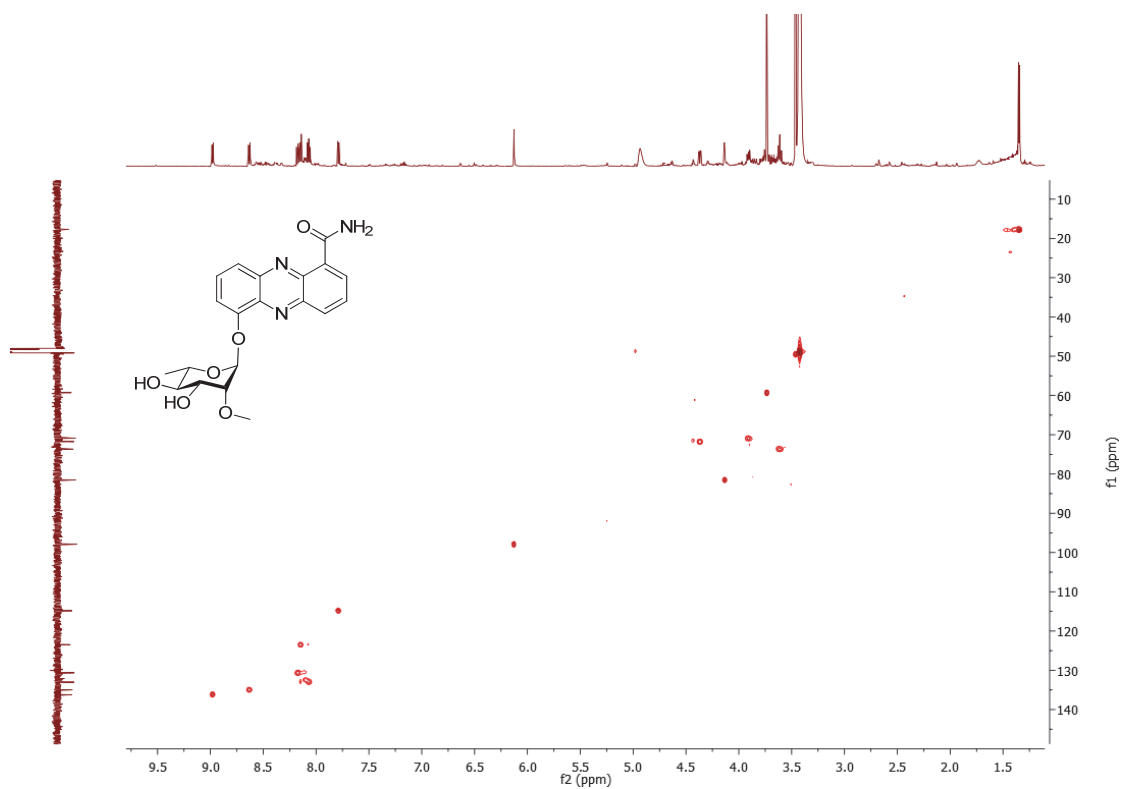
Spectral file. Spectra list of new compounds

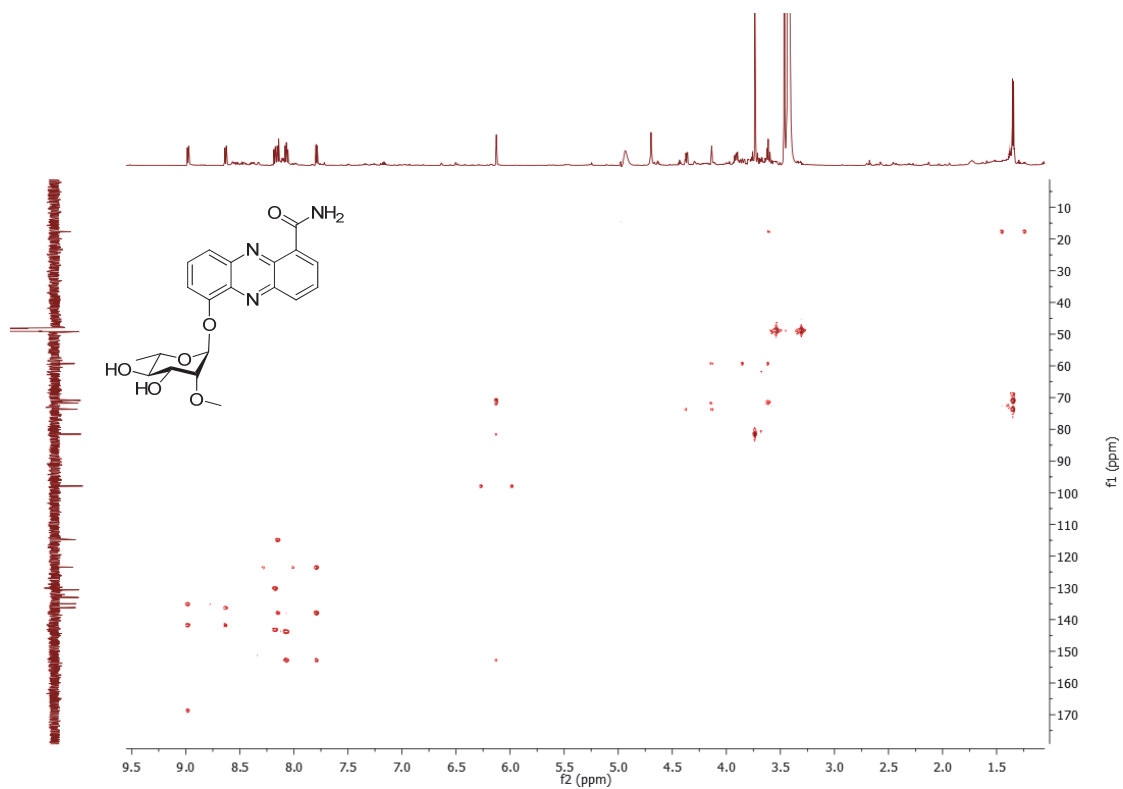
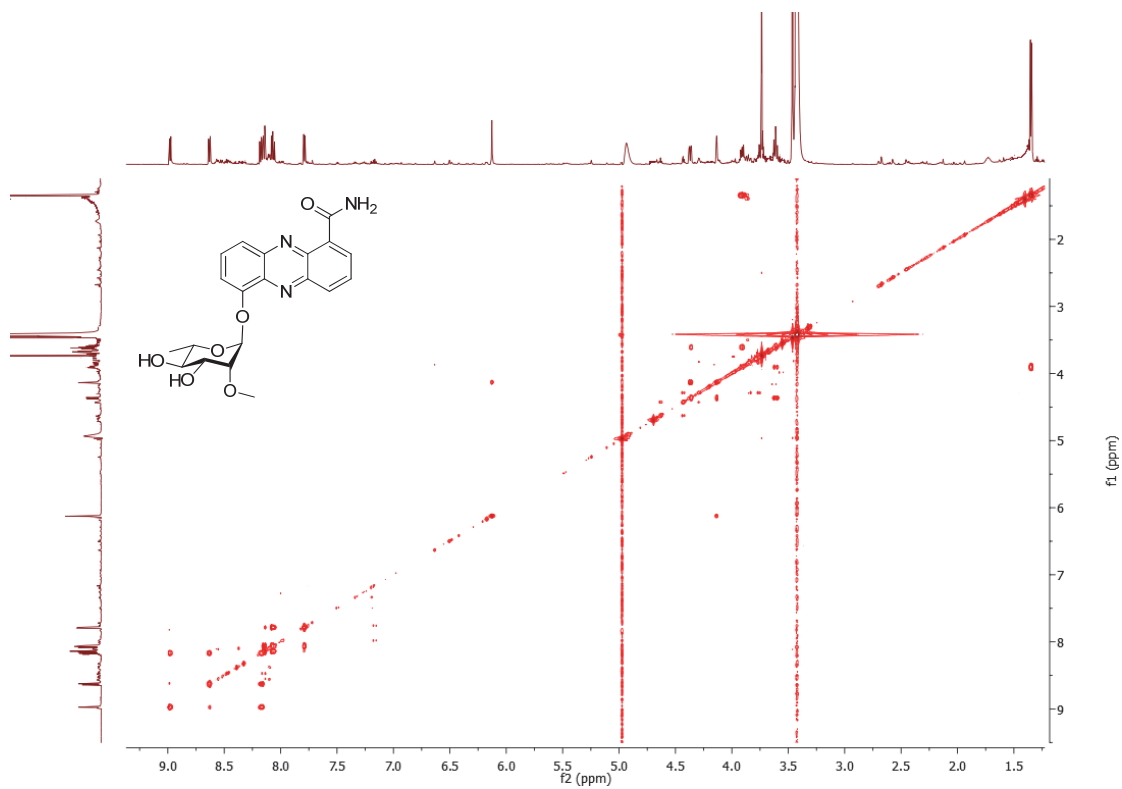
- S1. ^1H NMR spectrum (600 MHz) of endophenaside A (**1**) in CD_3OD .
- S2. APT spectrum of endophenaside A (**1**) in CD_3OD .
- S3. HSQC spectrum of endophenaside A (**1**) in CD_3OD .
- S4. HMBC spectrum of endophenaside A (**1**) in CD_3OD .
- S5. ^1H - ^1H COSY spectrum of endophenaside A (**1**) in CD_3OD .
- S6. ^1H - ^1H NOSEY spectrum of endophenaside A (**1**) in acetone- d_6 .
- S7. HRMS spectrum of endophenaside A (**1**).
- S8. IR spectrum of endophenaside A (**1**).
- S9. UV spectrum of endophenaside A (**1**).
- S10. ^1H NMR spectrum (500 MHz) of endophenaside B (**2**) in CD_3OD .
- S11. APT spectrum of endophenaside B (**2**) in CD_3OD .
- S12. HSQC spectrum of endophenaside B (**2**) in CD_3OD .
- S13. HMBC spectrum of endophenaside B (**2**) in CD_3OD .
- S14. ^1H - ^1H COSY spectrum of endophenaside B (**2**) in CD_3OD .
- S15. ^1H - ^1H NOSEY spectrum of endophenaside B (**2**) in acetone- d_6 .
- S16. HRMS spectrum of endophenaside B (**2**).
- S17. IR spectrum of endophenaside B (**2**).
- S18. UV spectrum of endophenaside B (**2**).
- S19. ^1H NMR spectrum (500 MHz) of endophenaside C (**3**) in CD_3OD .
- S20. APT spectrum of endophenaside C (**3**) in CD_3OD .
- S21. HSQC spectrum of endophenaside C (**3**) in CD_3OD .
- S22. HMBC spectrum of endophenaside C (**3**) in CD_3OD .
- S23. ^1H - ^1H COSY spectrum of endophenaside C (**3**) in CD_3OD .
- S24. HRMS spectrum of endophenaside C (**3**).
- S25. IR spectrum of endophenaside C (**3**).
- S26. UV spectrum of endophenaside C (**3**).
- S27. ^1H NMR spectrum (500 MHz) of endophenaside D (**4**) in CD_3OD .
- S28. APT spectrum of endophenaside D (**4**) in CD_3OD .
- S29. HSQC spectrum of endophenaside D (**4**) in CD_3OD .
- S30. HMBC spectrum of endophenaside D (**4**) in CD_3OD .
- S31. ^1H - ^1H COSY spectrum of endophenaside D (**4**) in CD_3OD .
- S32. HRMS spectrum of endophenaside D (**4**).
- S33. IR spectrum of endophenaside D (**4**).
- S34. UV spectrum of endophenaside D (**4**).
- S35. ^1H NMR spectrum (500 MHz) of endophenaside E (**5**) in CD_3OD .
- S36. HSQC spectrum of endophenaside E (**5**) in CD_3OD .
- S37. HMBC spectrum of endophenaside E (**5**) in CD_3OD .
- S38. ^1H - ^1H COSY spectrum of endophenaside E (**5**) in CD_3OD .
- S39. HRMS spectrum of endophenaside E (**5**).
- S40. IR spectrum of endophenaside E (**5**).
- S41. UV spectrum of endophenaside E (**5**).
- S42. ^1H NMR spectrum (600 MHz) of endophenazine F1(**6**) in CD_3OD .
- S43. APT spectrum of endophenazine F1(**6**) in CD_3OD .
- S44. HSQC spectrum of endophenazine F1(**6**) in CD_3OD .

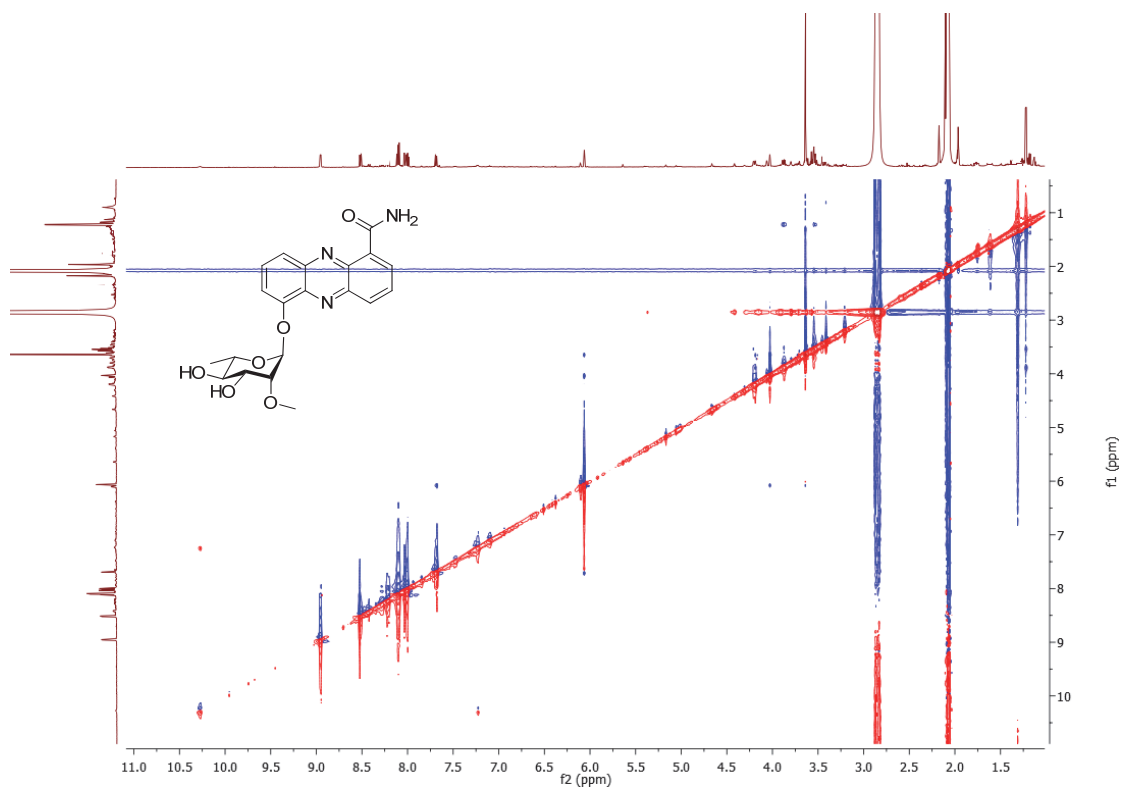
- S45. HMBC spectrum of endophenazine F1(6) in CD₃OD.
S46. ¹H-¹H COSY spectrum of endophenazine F1 (6) in CD₃OD.
S47. ¹H-¹H NOSEY spectrum of endophenazine F1 (6) acetone-*d*₆..
S48. HRMS spectrum of endophenazine F1(6).
S49. IR spectrum of endophenazine F1(6).
S50. UV spectrum of endophenazine F1(6).

S1. ¹H NMR spectrum (600 MHz) of endophenaside A (1) in CD₃OD.

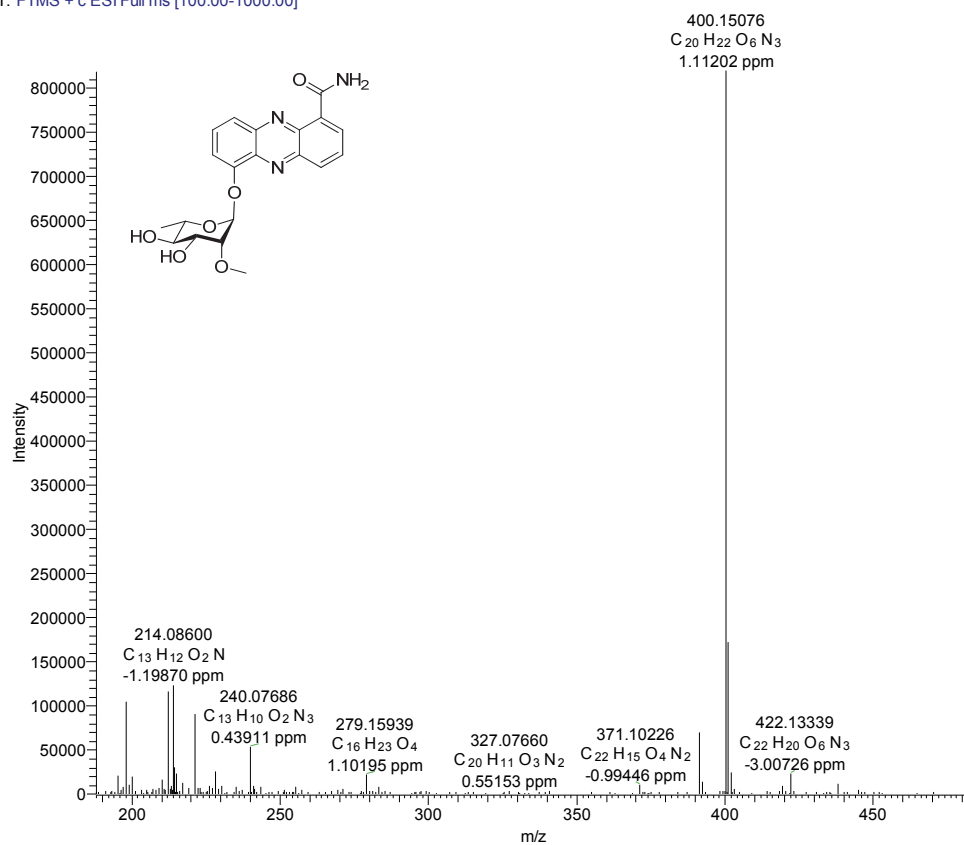


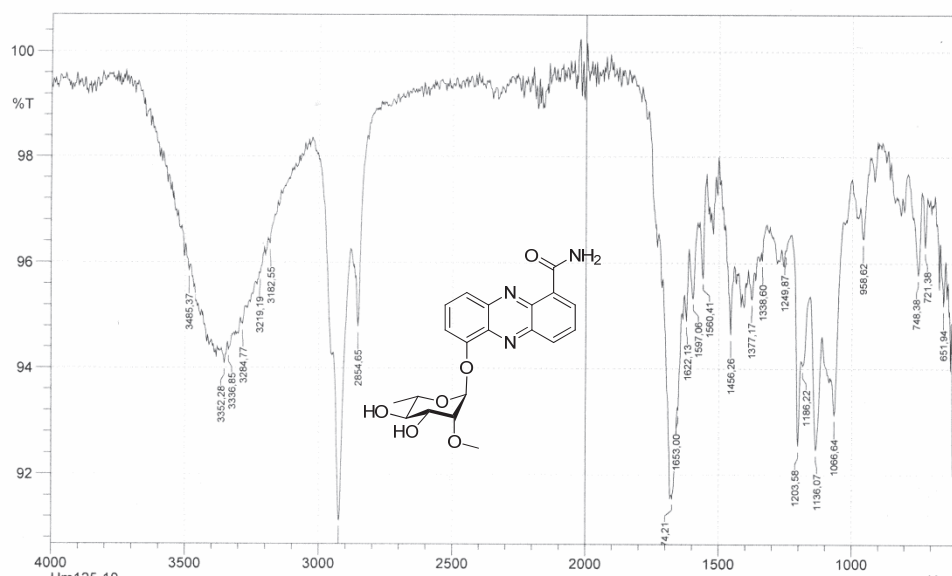
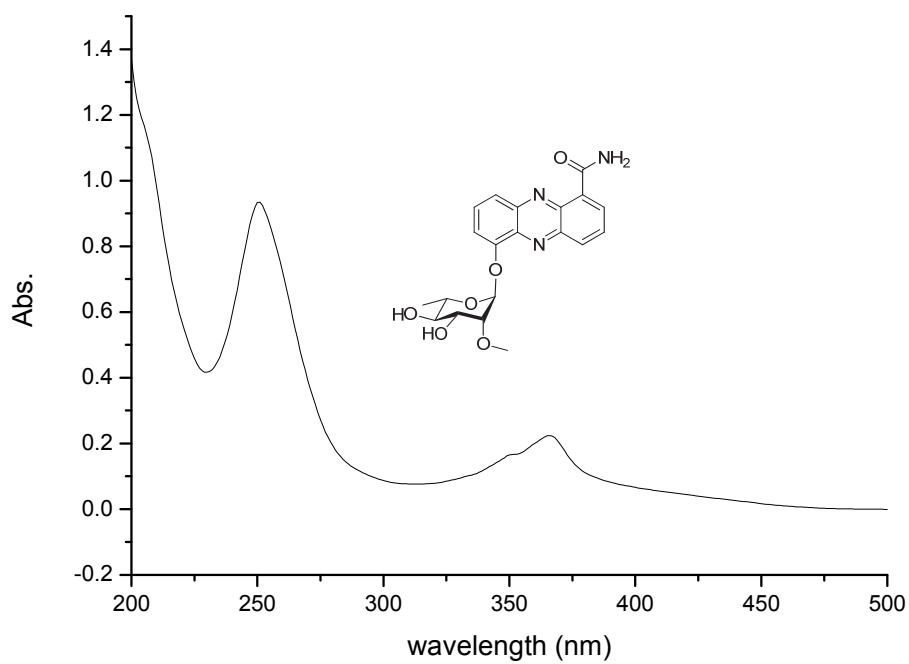
S2. APT spectrum of endophenaside A (**1**) in CD₃OD.S3. HSQC spectrum of endophenaside A (**1**) in CD₃OD.

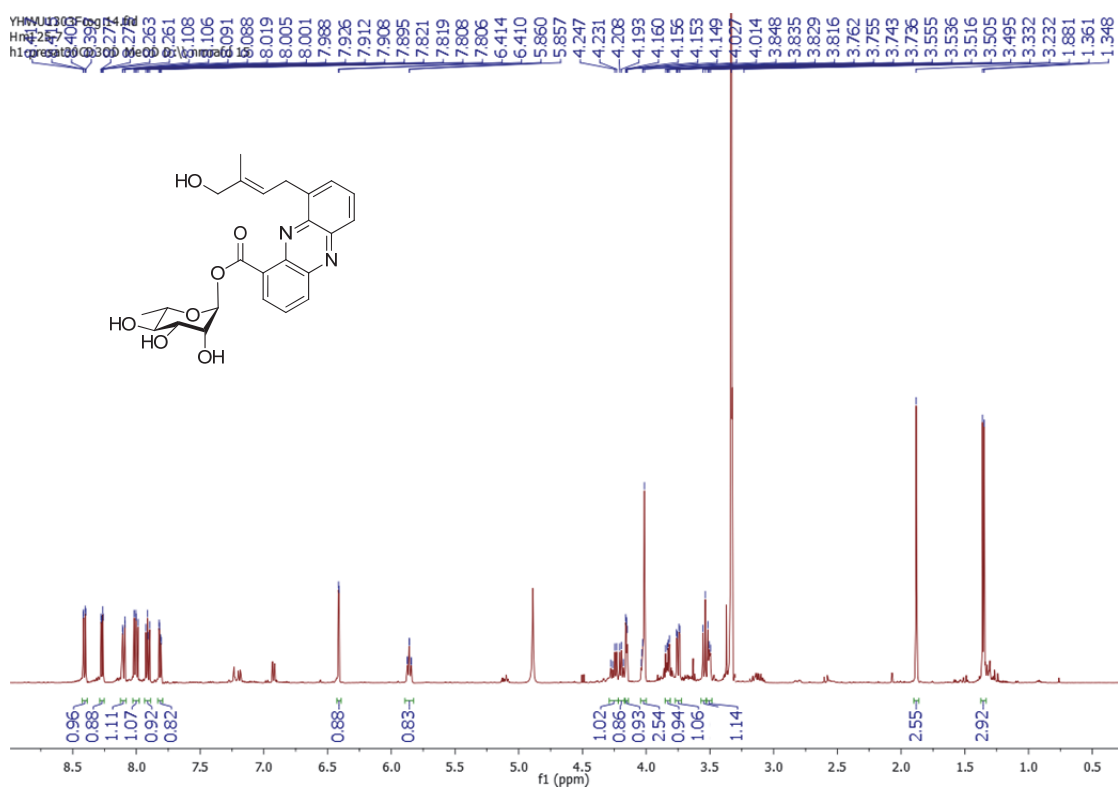
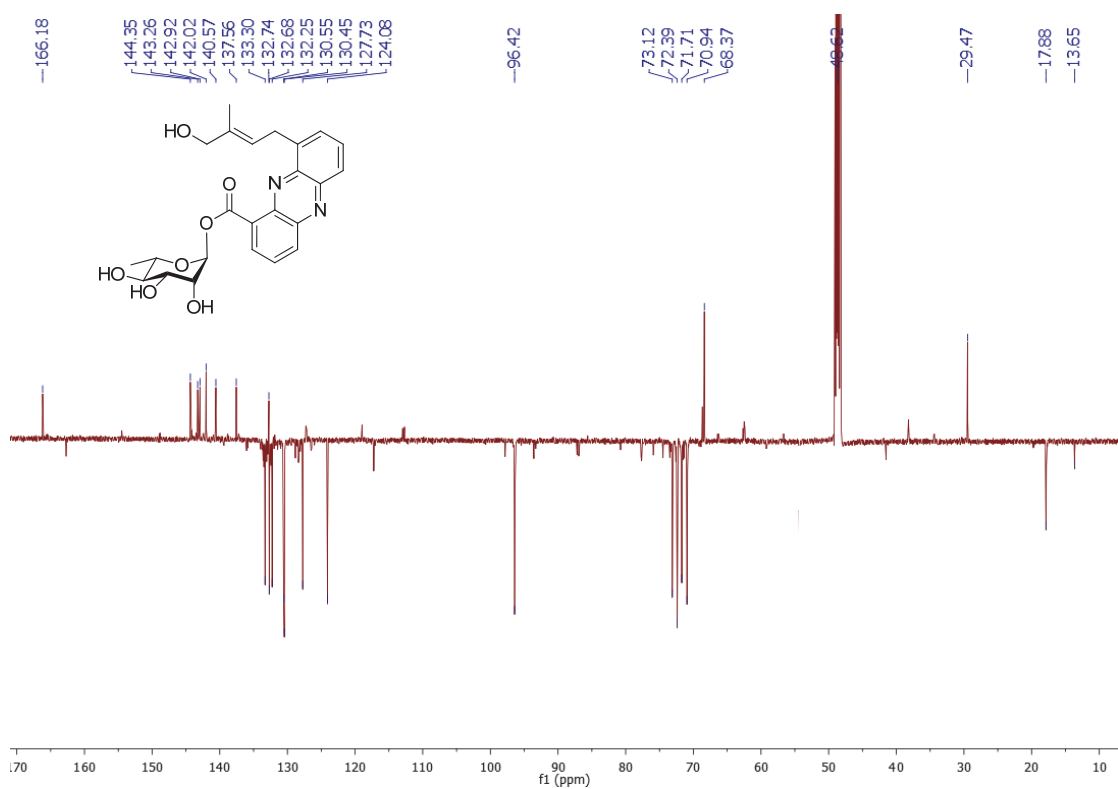
S4. HMBC spectrum of endophenaside A (**1**) in CD₃OD.S5. ¹H-¹H COSY spectrum of endophenaside A (**1**) in CD₃OD.

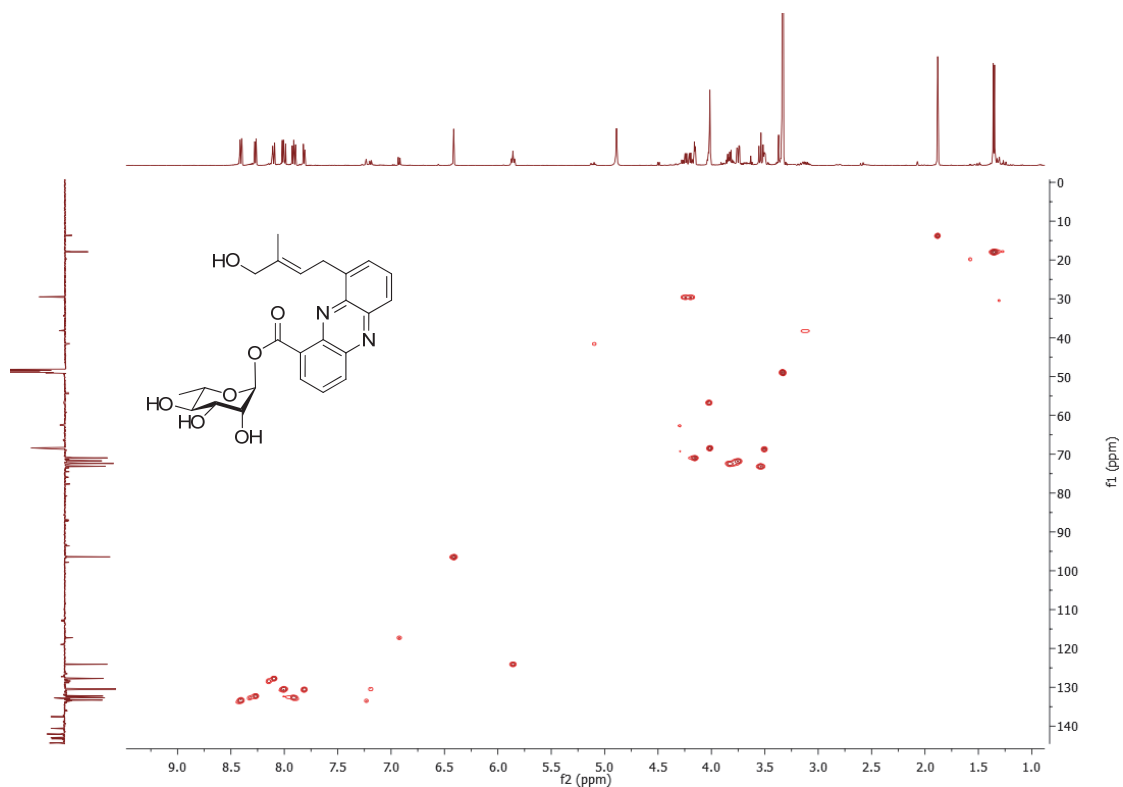
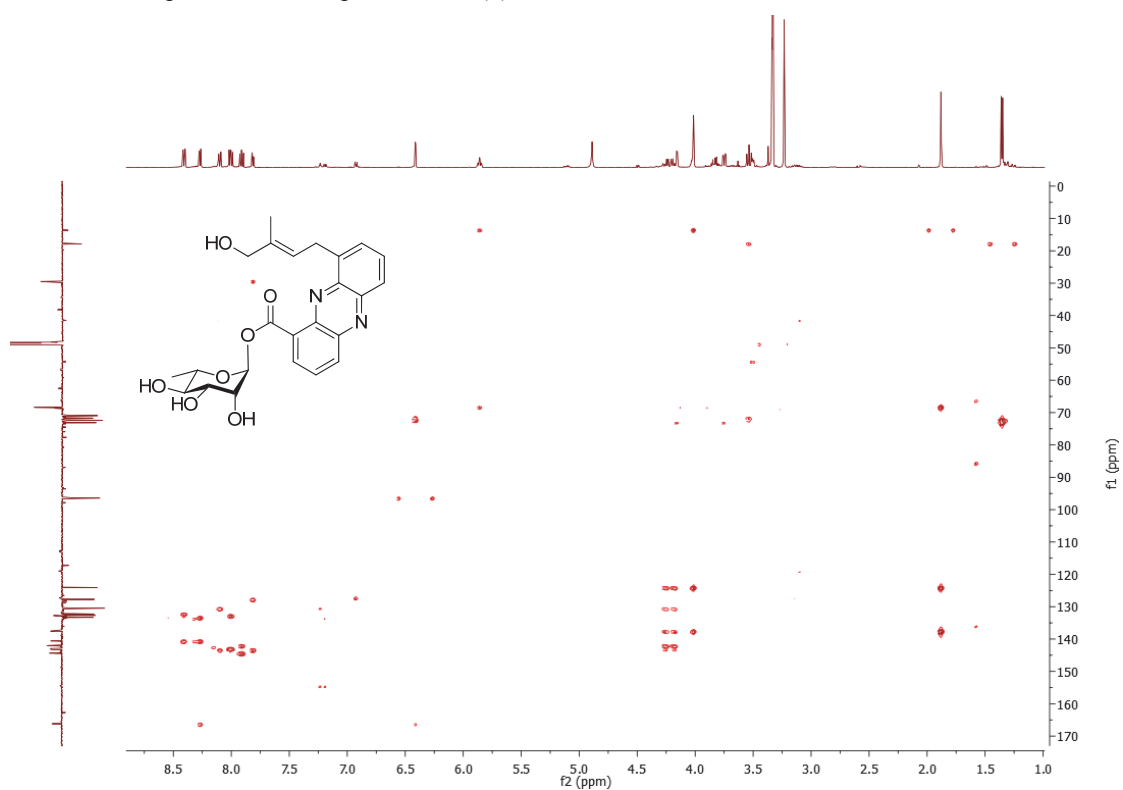
S6. ^1H - ^1H NOSEY spectrum of endophenaside A (**1**) in acetone- d_6 .S7. HRMS spectrum of endophenaside A (**1**).

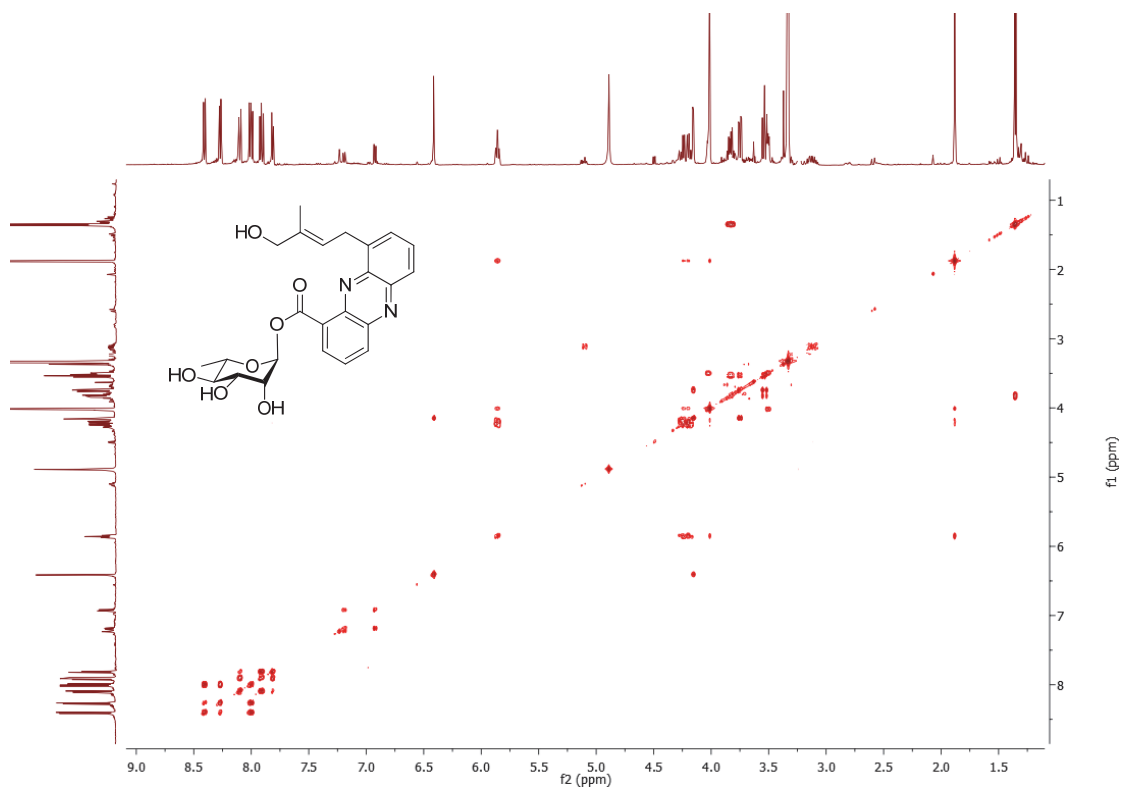
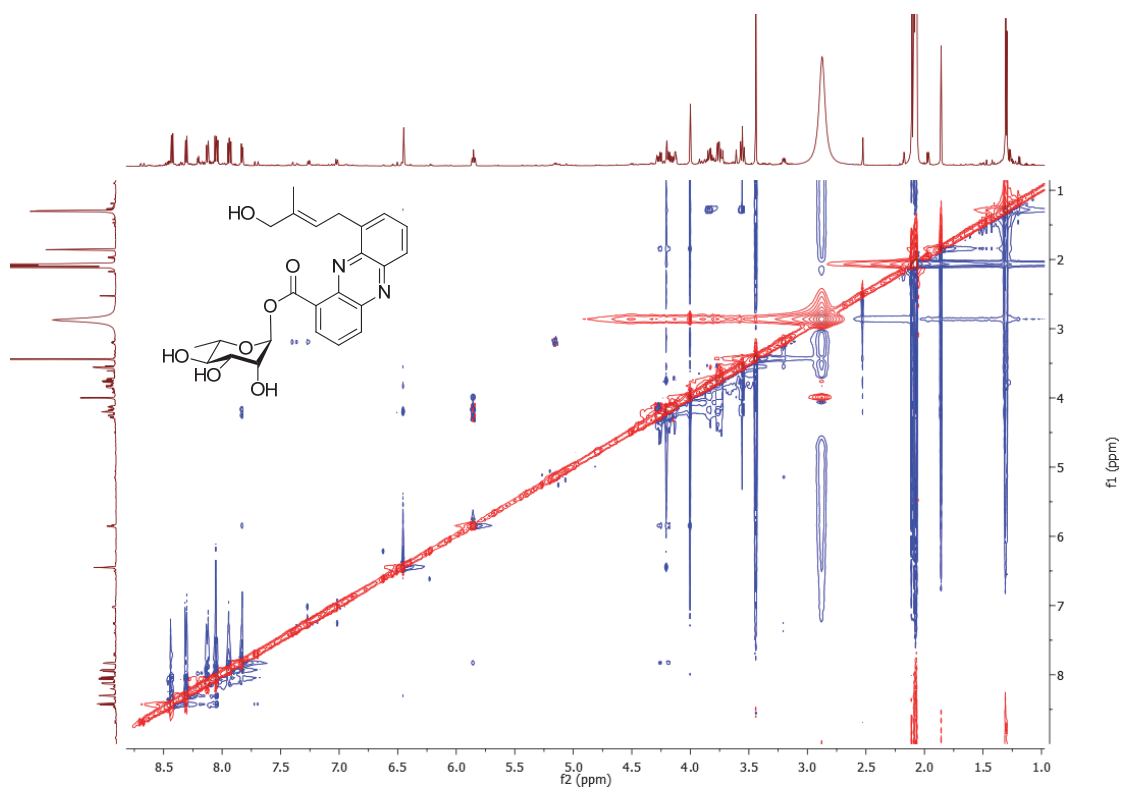
Hm125-10 #464 RT: 8.02 AV: 1 NL: 8.18E5
T: FTMS + c ESI Full ms [100.00-1000.00]



S8. IR spectrum of endophenaside A (**1**).S9. UV spectrum of endophenaside A (**1**).

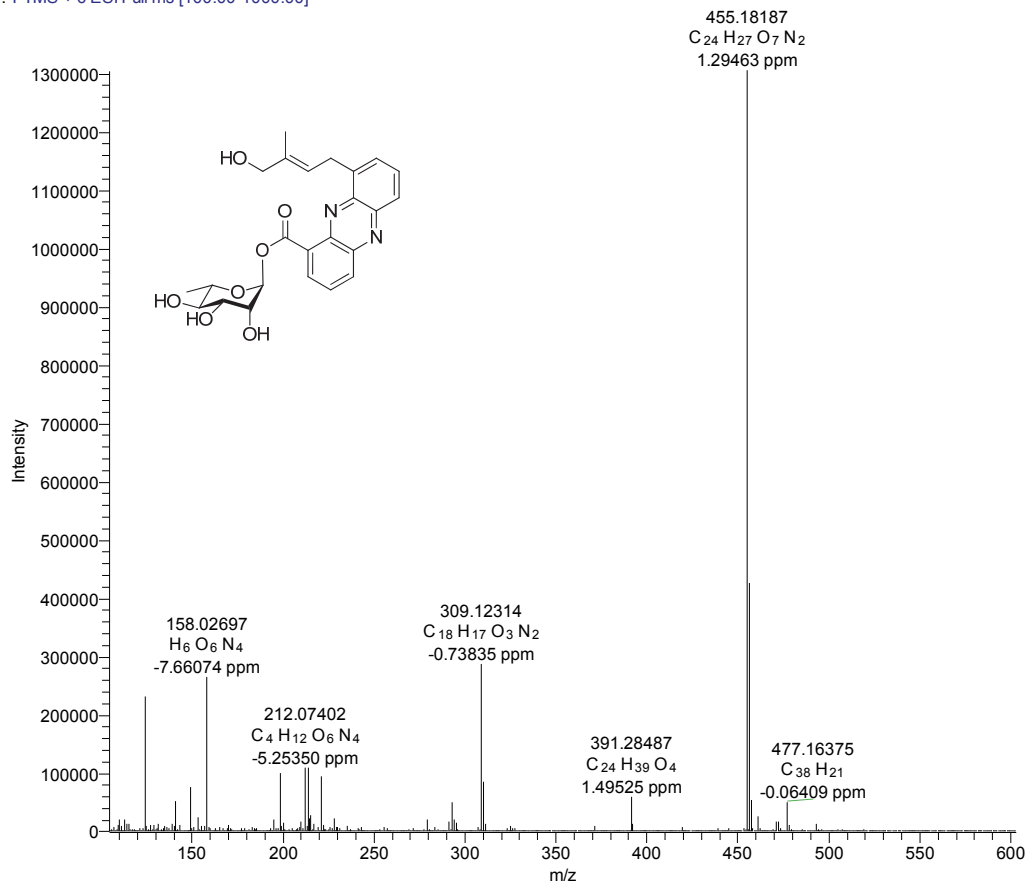
S10. ^1H NMR spectrum (500 MHz) of endophenaside B (**2**) in CD_3OD .S11. APT spectrum of endophenaside B (**2**) in CD_3OD .

S12. HSQC spectrum of endophenaside B (**2**) in CD₃OD.S13. HMBC spectrum of endophenaside B (**2**) in CD₃OD.

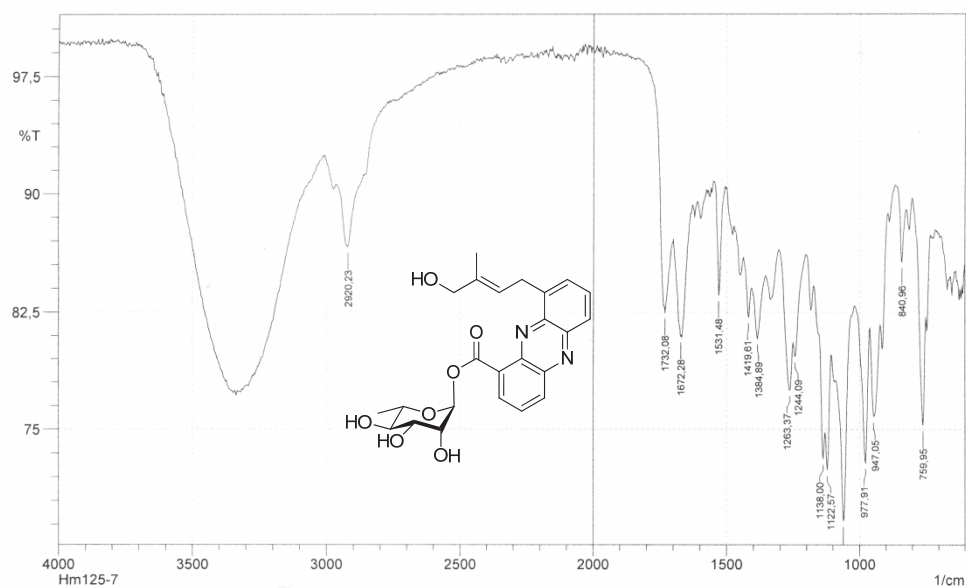
S14. ^1H - ^1H COSY spectrum of endophenaside B (**2**) in CD_3OD .S15. ^1H - ^1H NOSEY spectrum of endophenaside B (**2**) in acetone- d_6 .

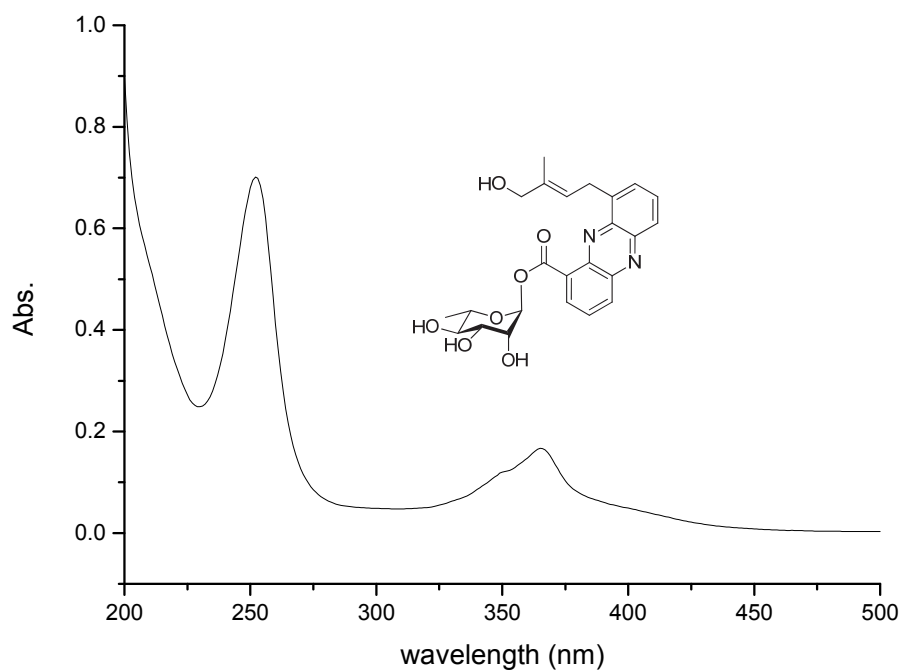
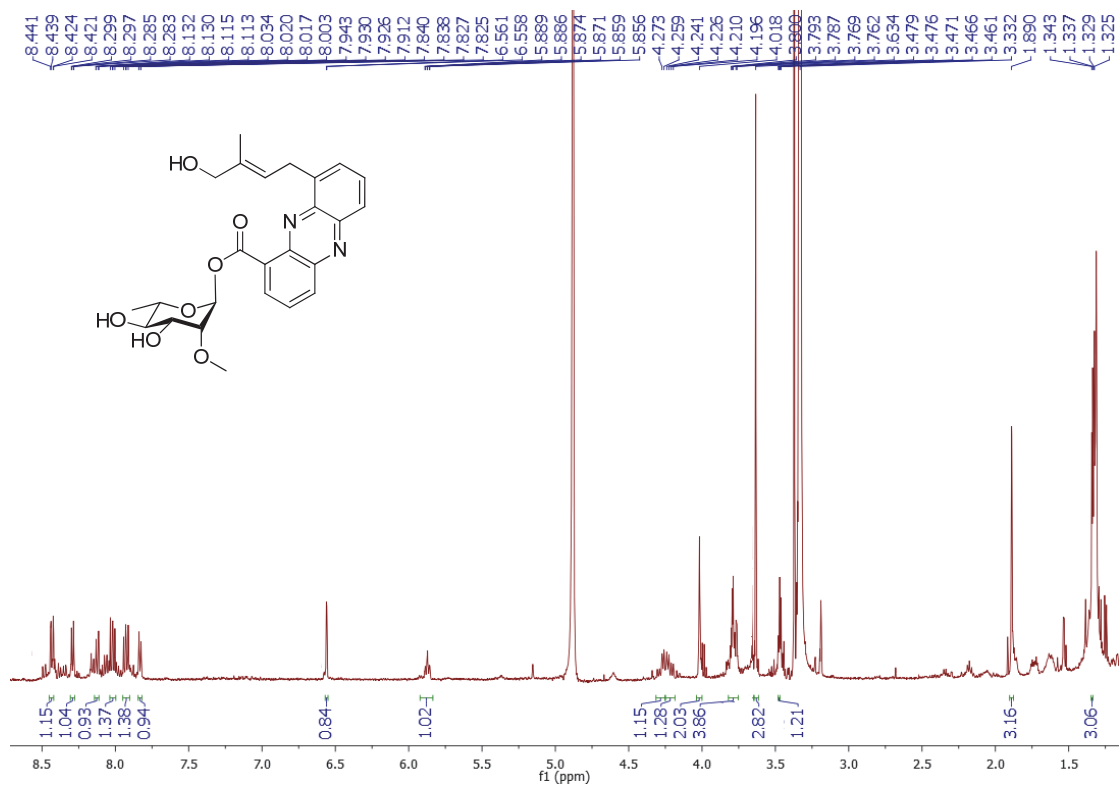
S16. HRMS spectrum of endophenaside B (2).

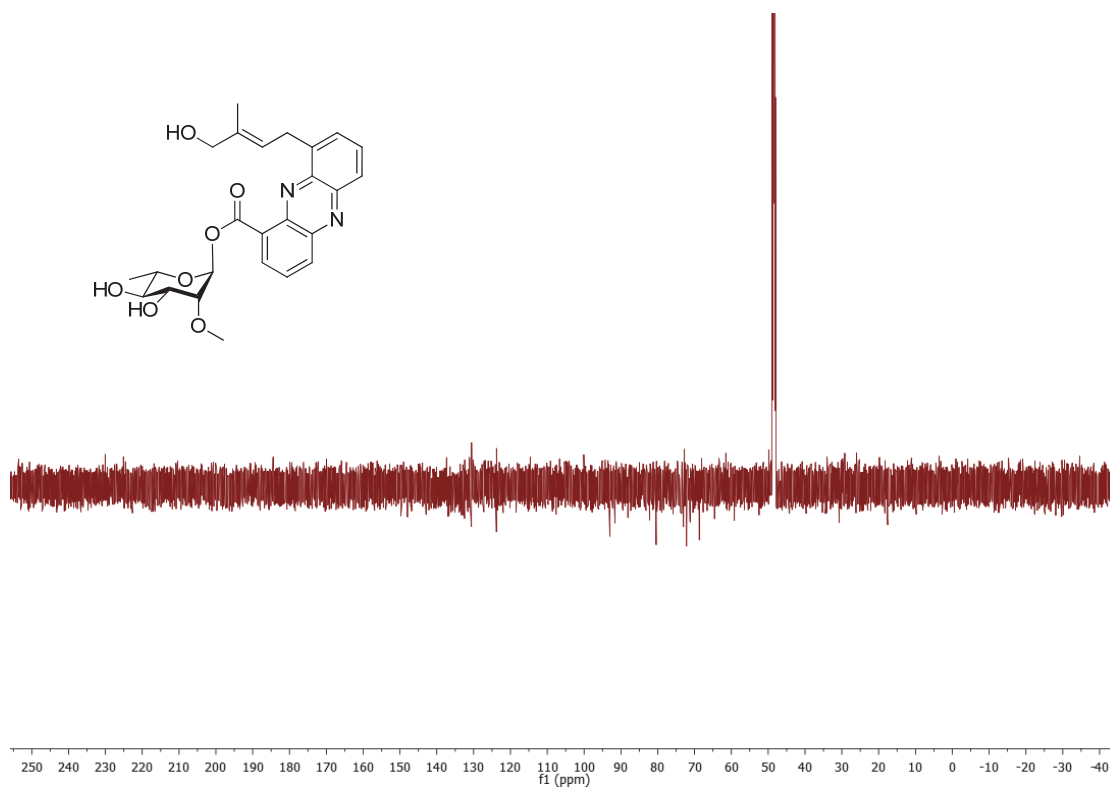
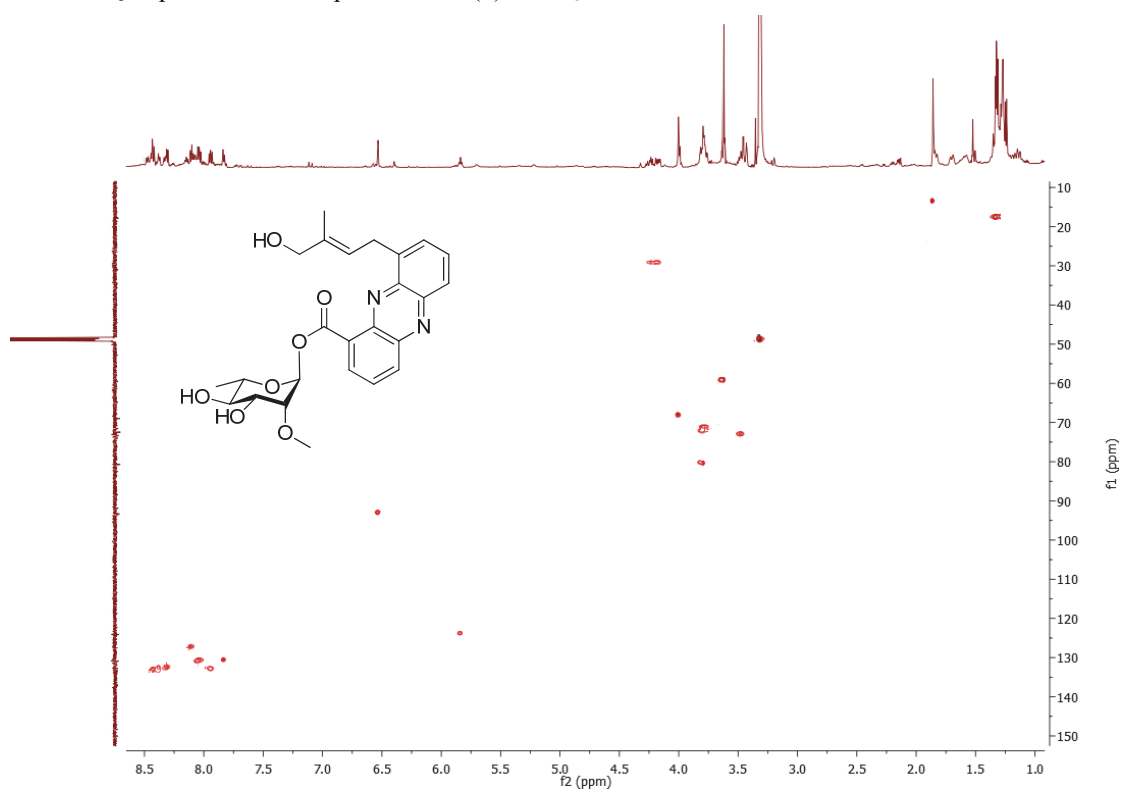
Hm125-7 #555-592 RT: 9.61-10.24 AV: 38 NL: 1.30E6
T: FTMS + c ESI Full ms [100.00-1000.00]

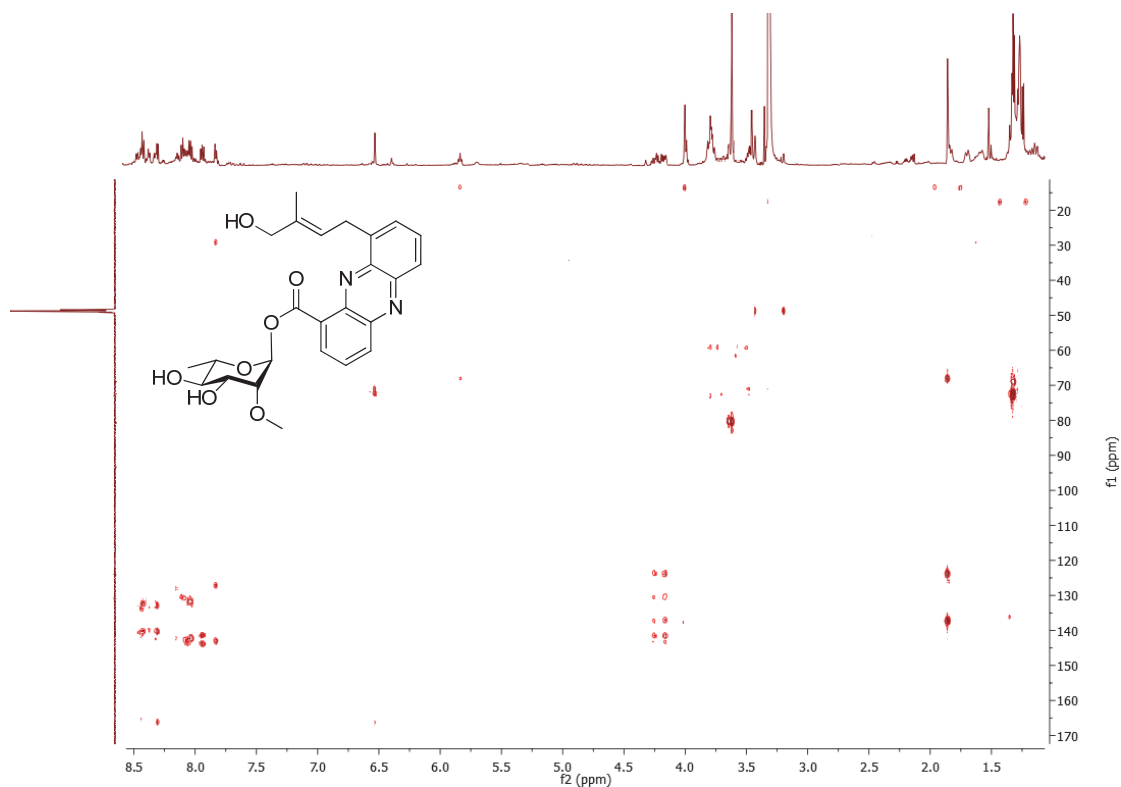
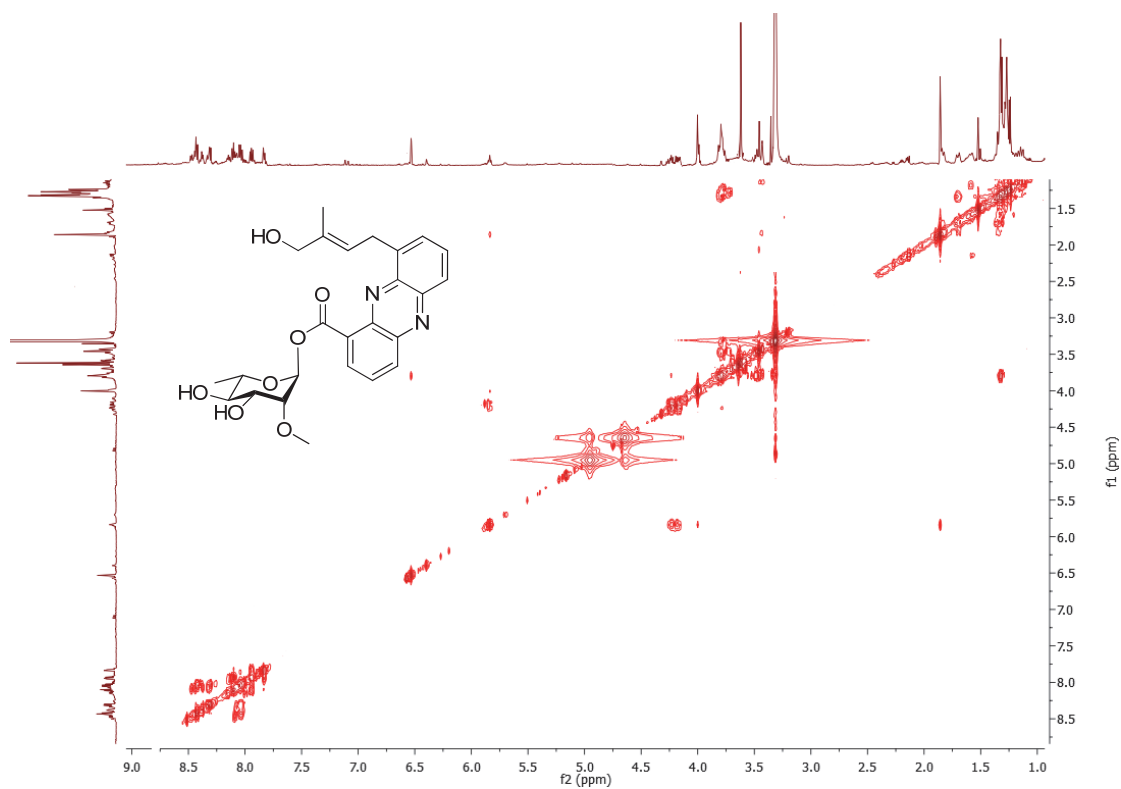


S17. IR spectrum of endophenaside B (2).



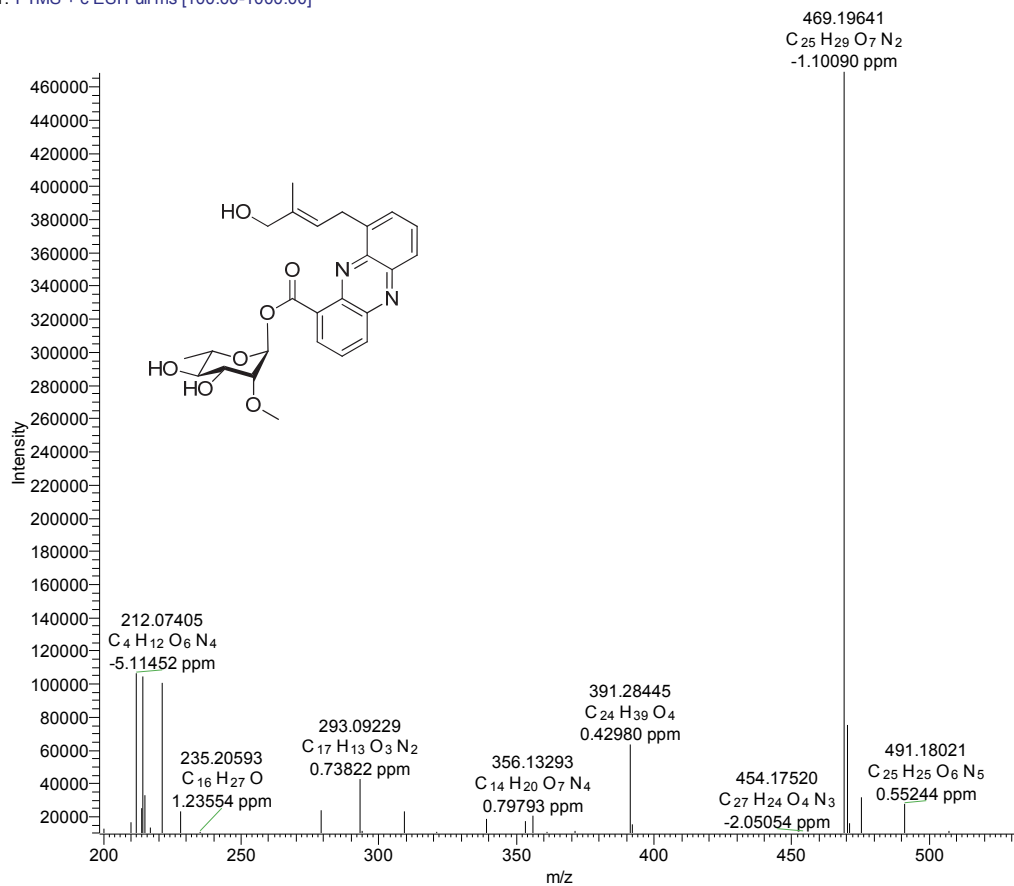
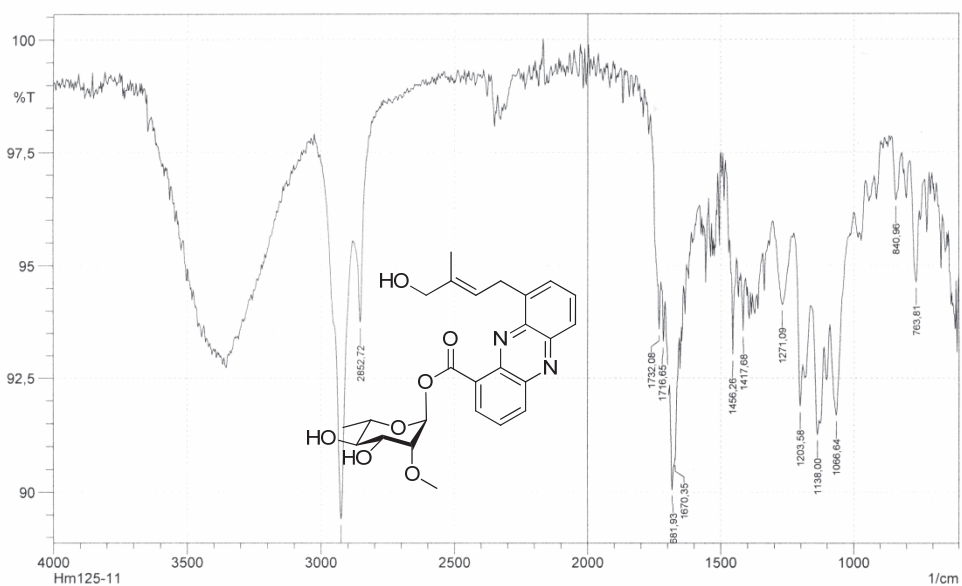
S18. UV spectrum of endophenaside B (**2**).S19. ^1H NMR spectrum (500 MHz) of endophenaside C (**3**) in CD_3OD .

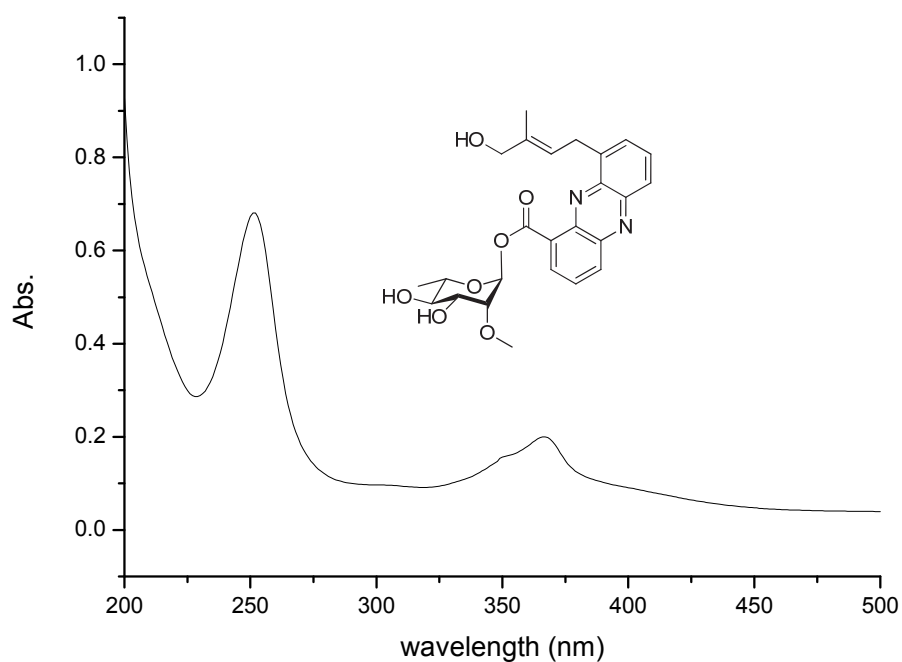
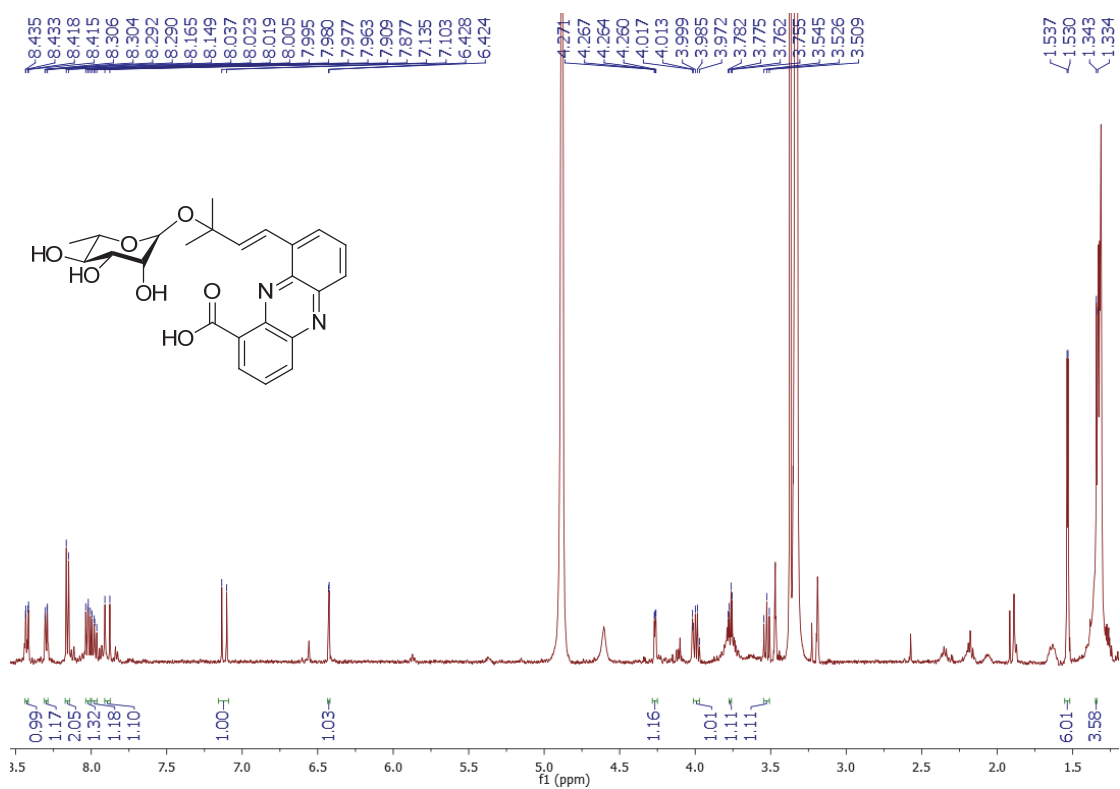
S20. APT spectrum of endophenaside C (**3**) in CD₃OD.S21. HSQC spectrum of endophenaside C (**3**) in CD₃OD.

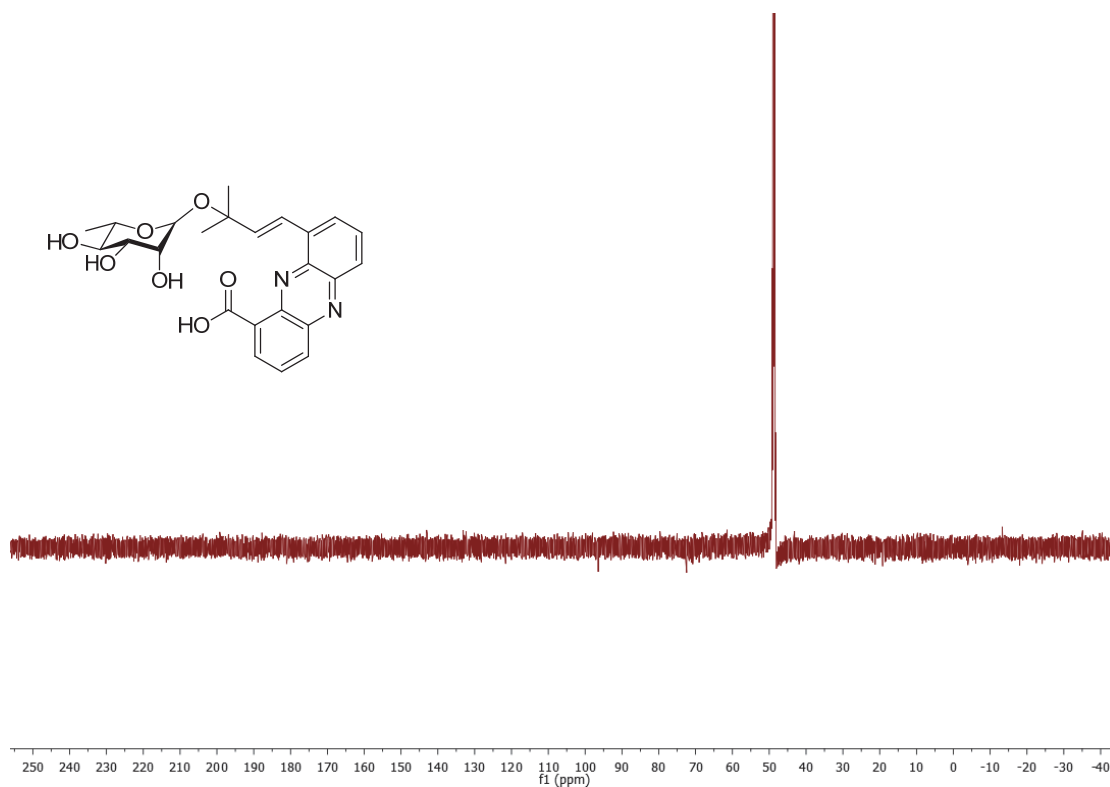
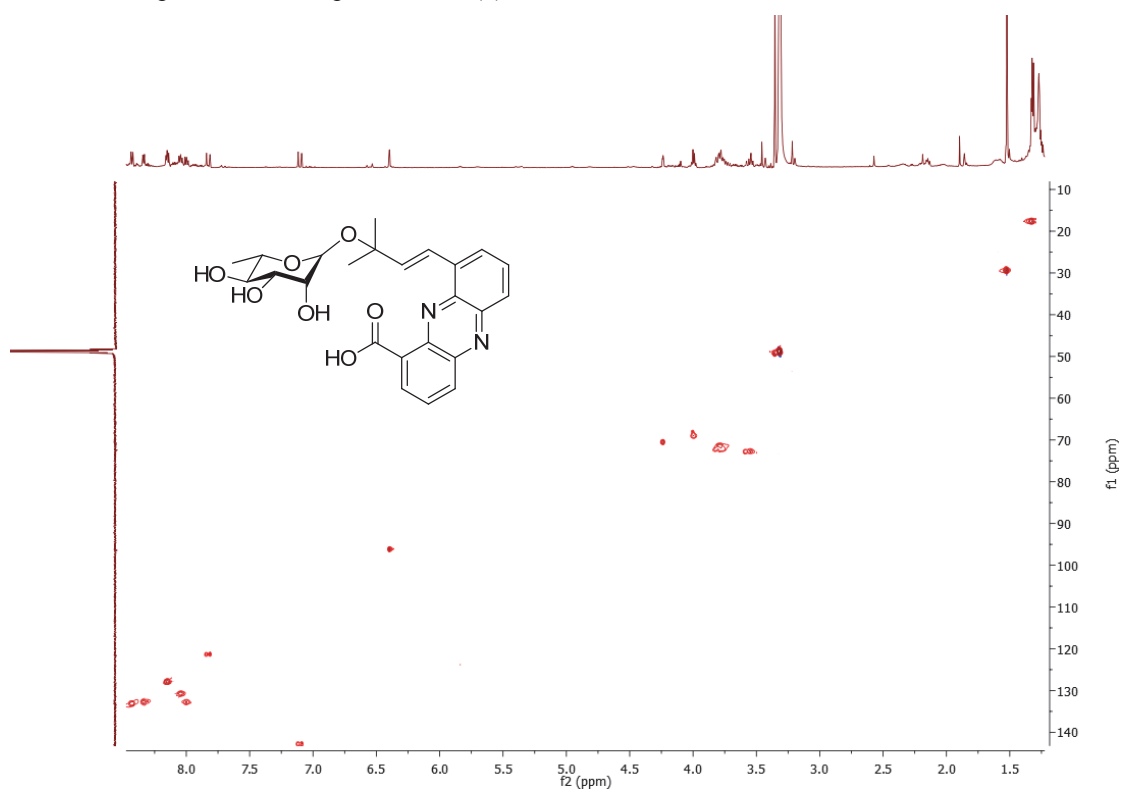
S22. HMBC spectrum of endophenaside C (**3**) in CD₃OD.S23. ¹H-¹H COSY spectrum of endophenaside C (**3**) in CD₃OD.

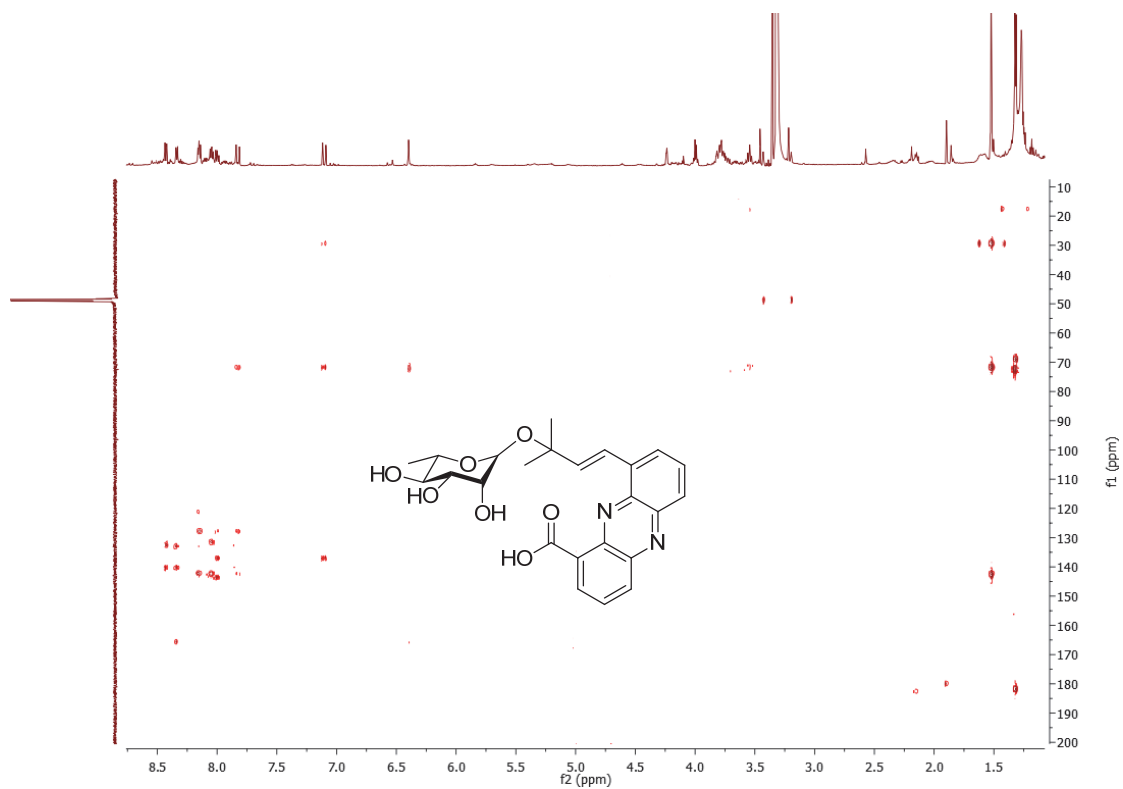
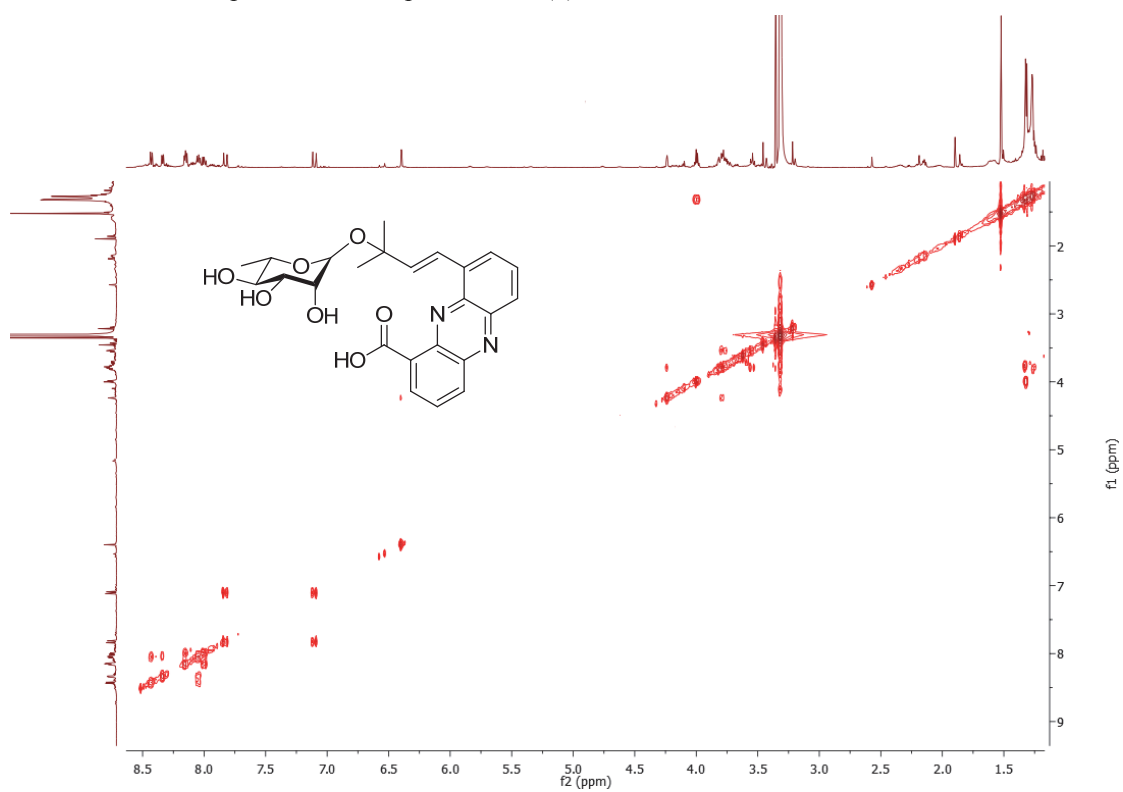
S24. HRMS spectrum of endophenaside C (**3**).

Hm125-11 #588 RT: 10.17 AV: 1 NL: 4.68E5
T: FTMS + c ESI Full ms [100.00-1000.00]

S25. IR spectrum of endophenaside C (**3**).

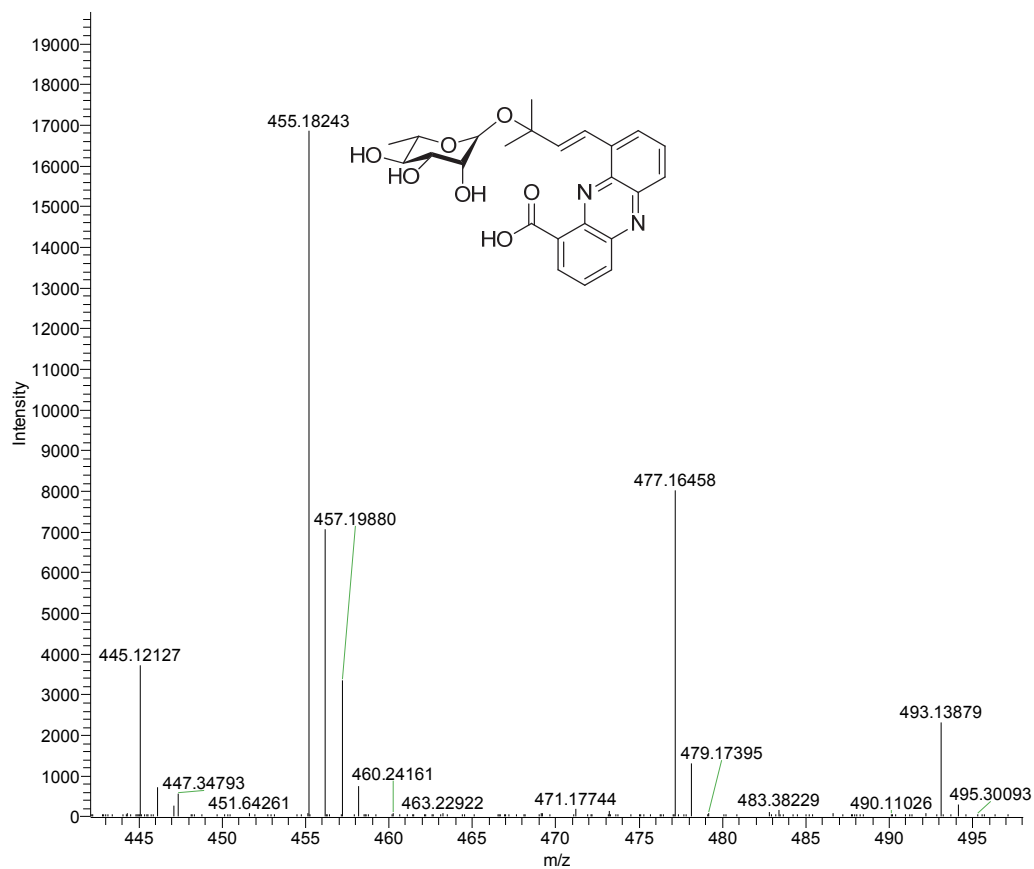
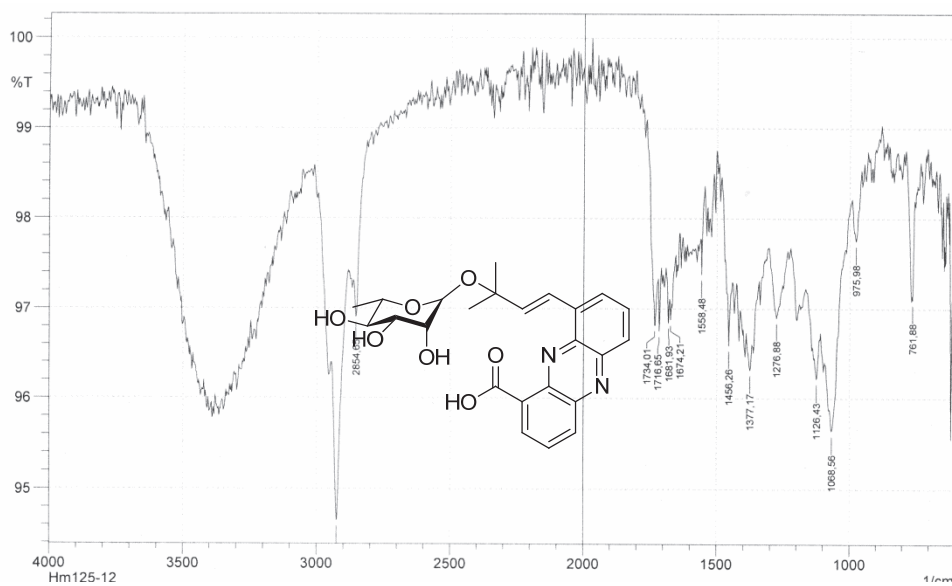
S26. UV spectrum of endophenaside C (**3**).S27. ^1H NMR spectrum (500 MHz) of endophenaside D (**4**) in CD_3OD .

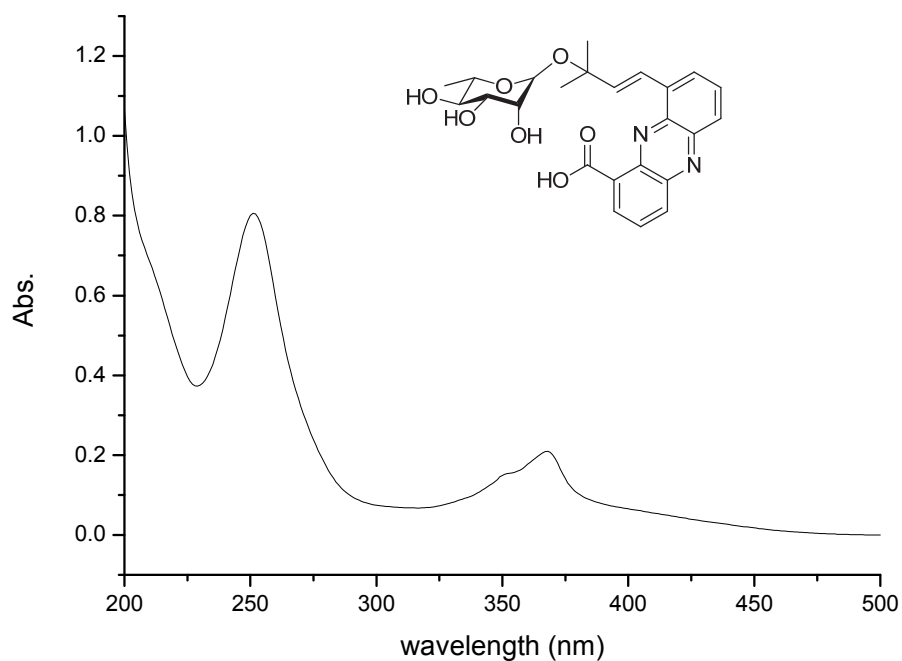
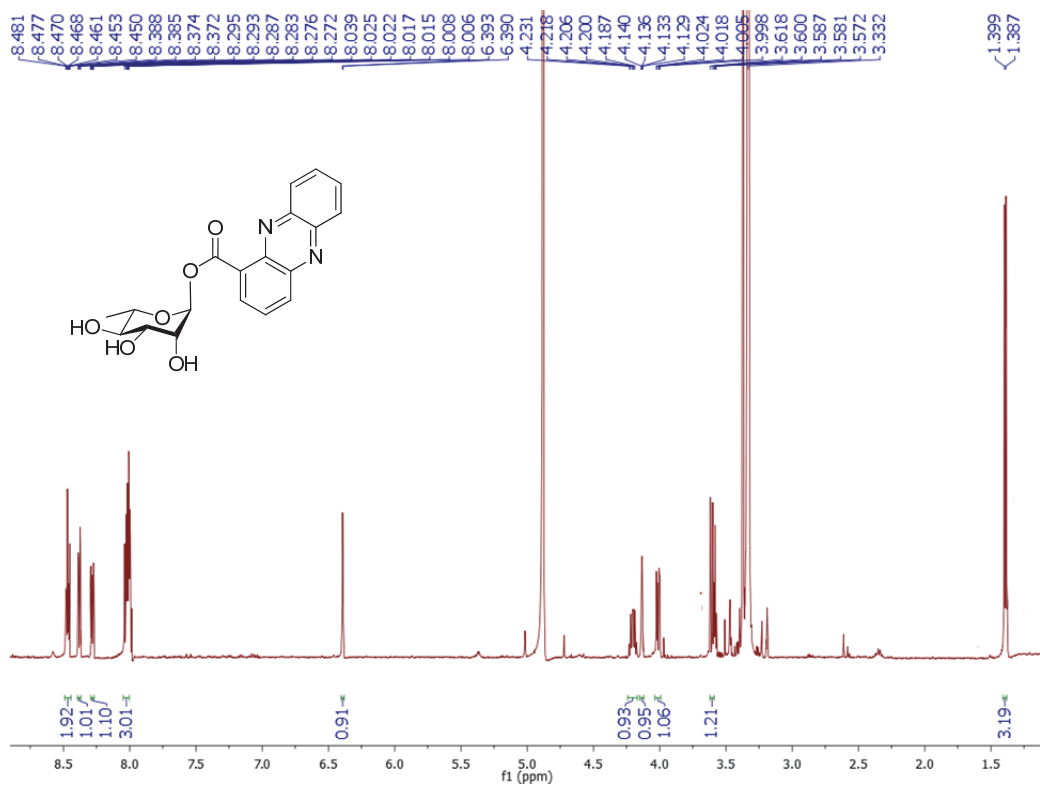
S28. APT spectrum of endophenaside D (**4**) in CD₃OD.S29. HSQC spectrum of endophenaside D (**4**) in CD₃OD.

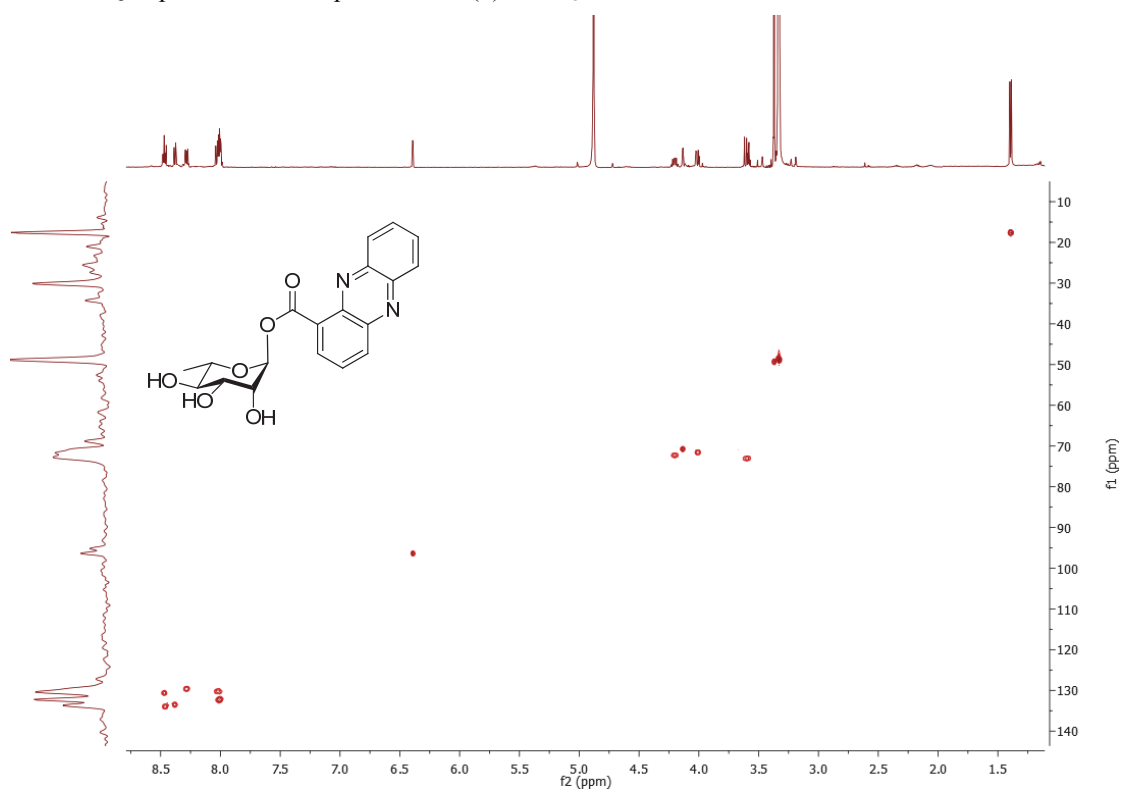
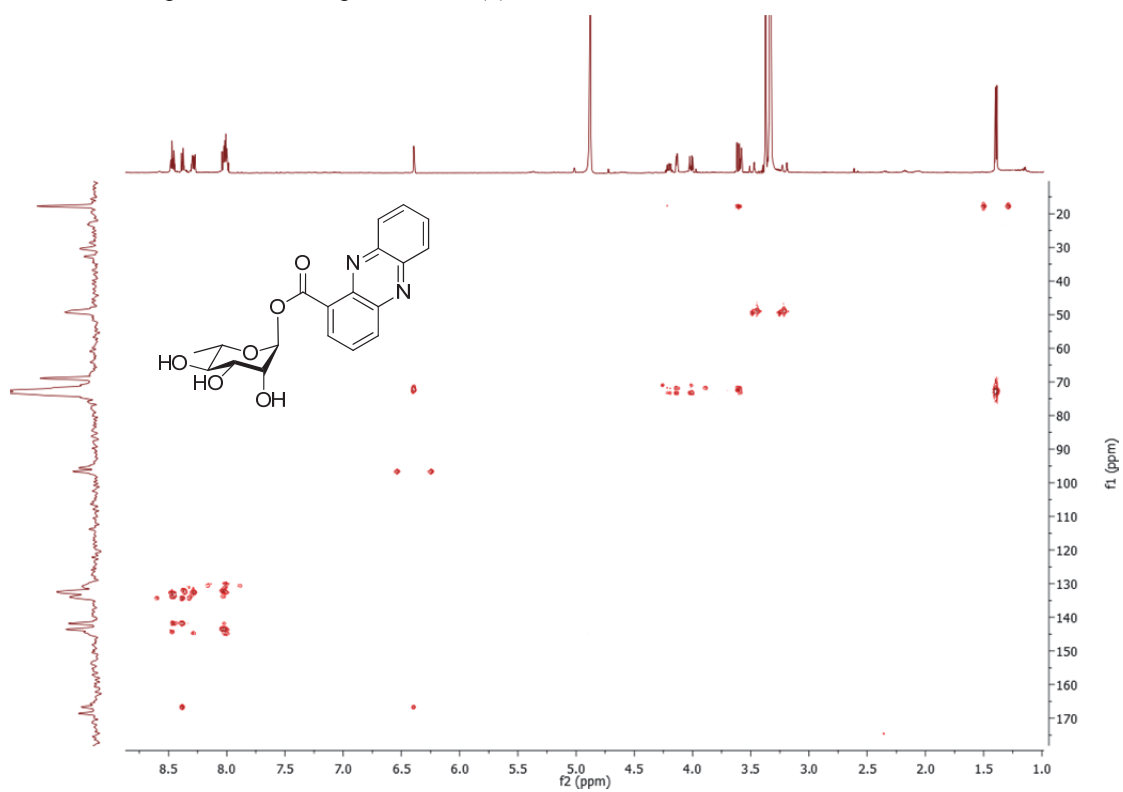
S30. HMBC spectrum of endophenaside D (**4**) in CD₃OD.S31. ¹H-¹H COSY spectrum of endophenaside D (**4**) in CD₃OD.

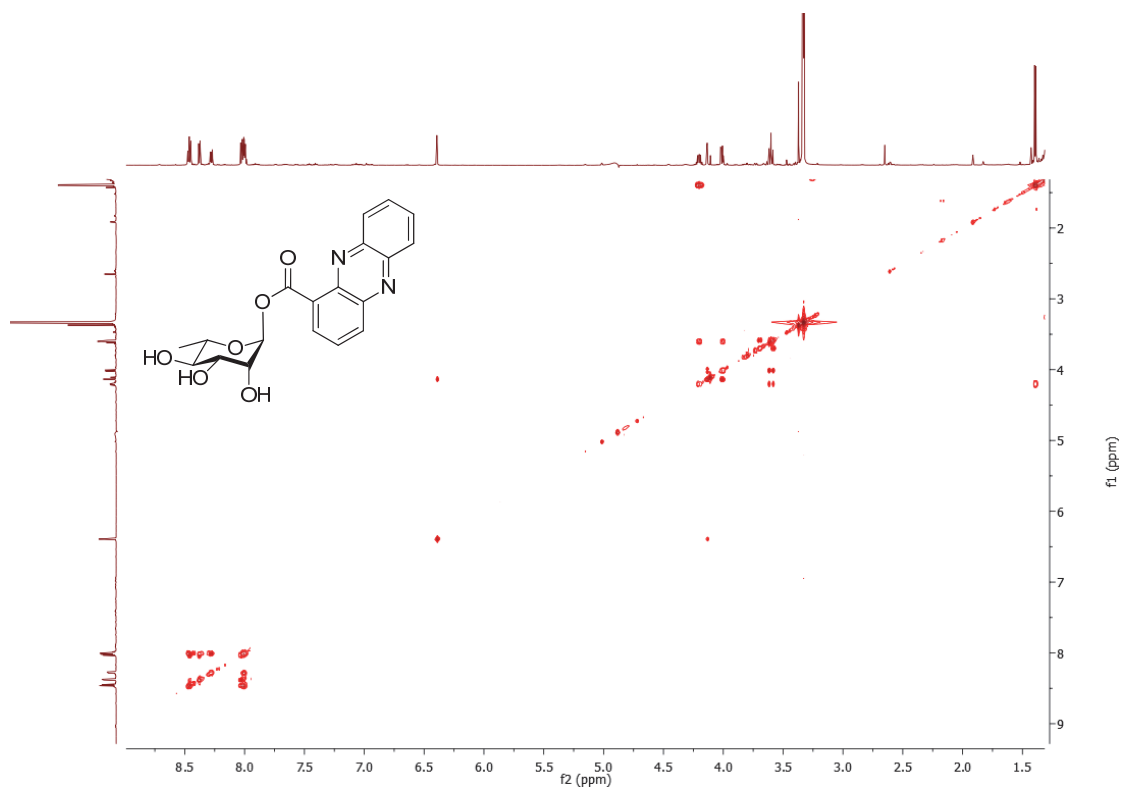
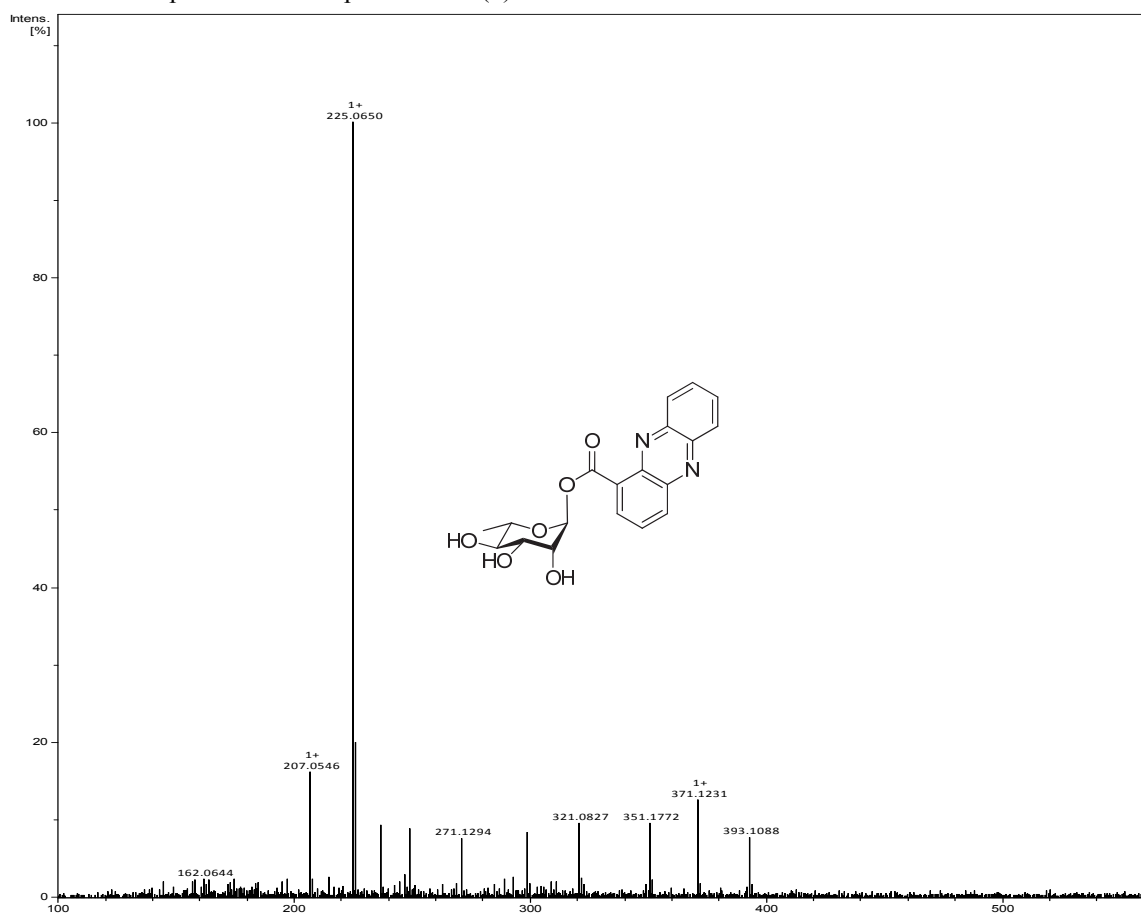
S32. HRMS spectrum of endophenaside D (**4**).

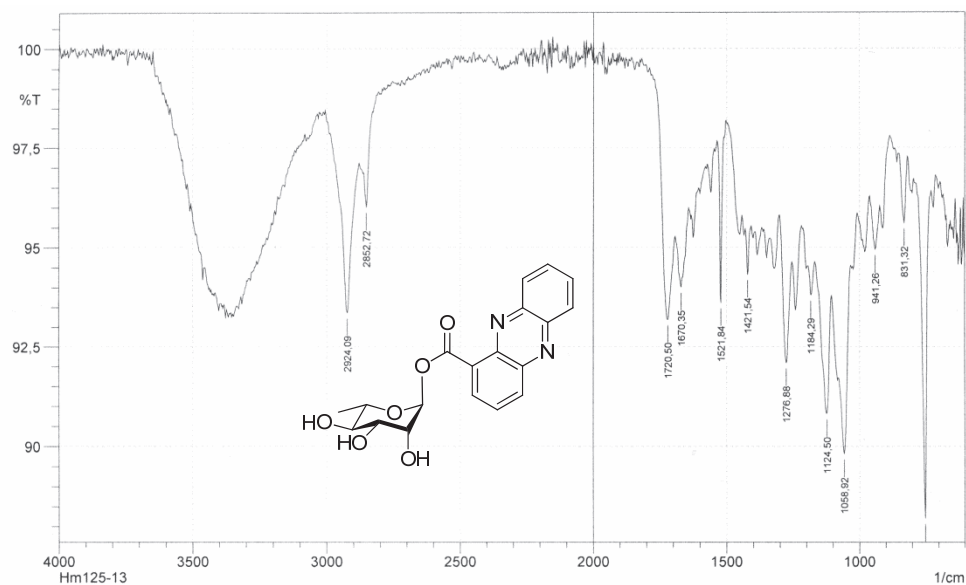
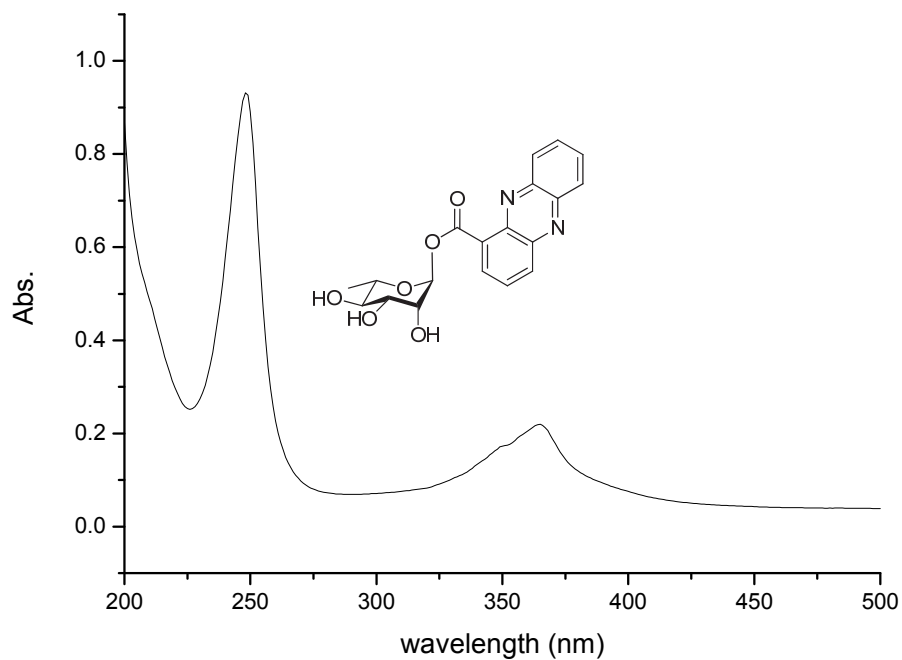
Hm125-12 #554-578 RT: 9.62-10.03 AV: 25 NL: 1.68E4
T: FTMS + c ESI Full ms [100.00-1000.00]

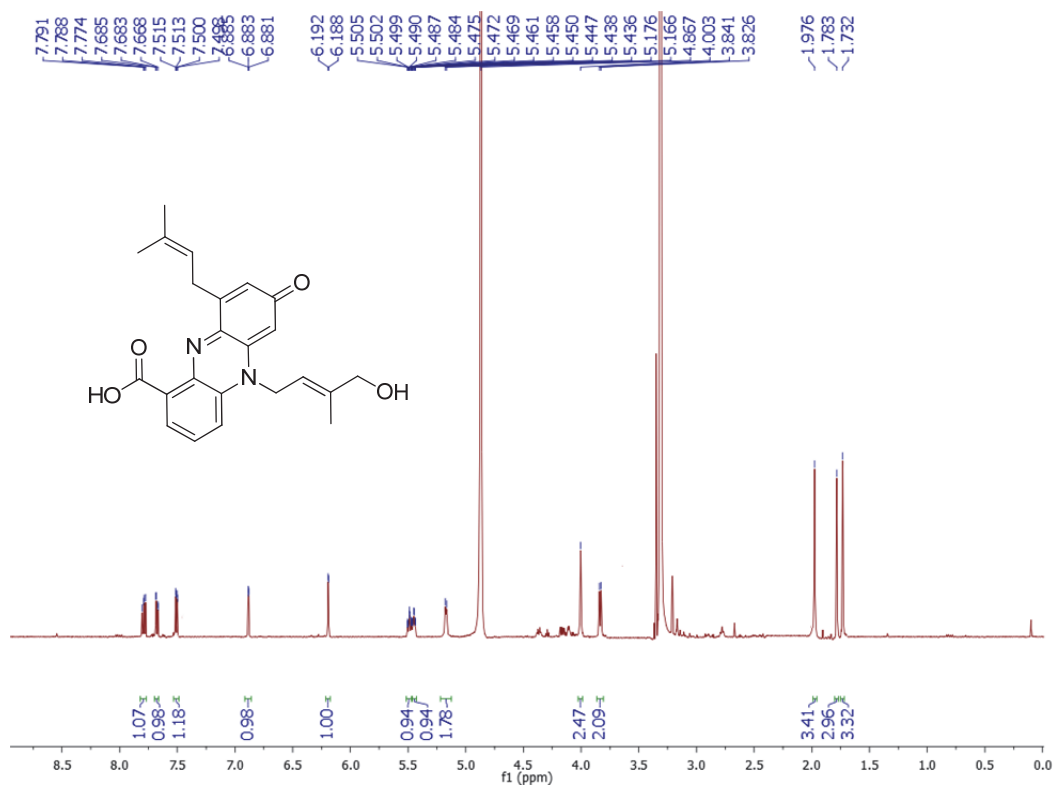
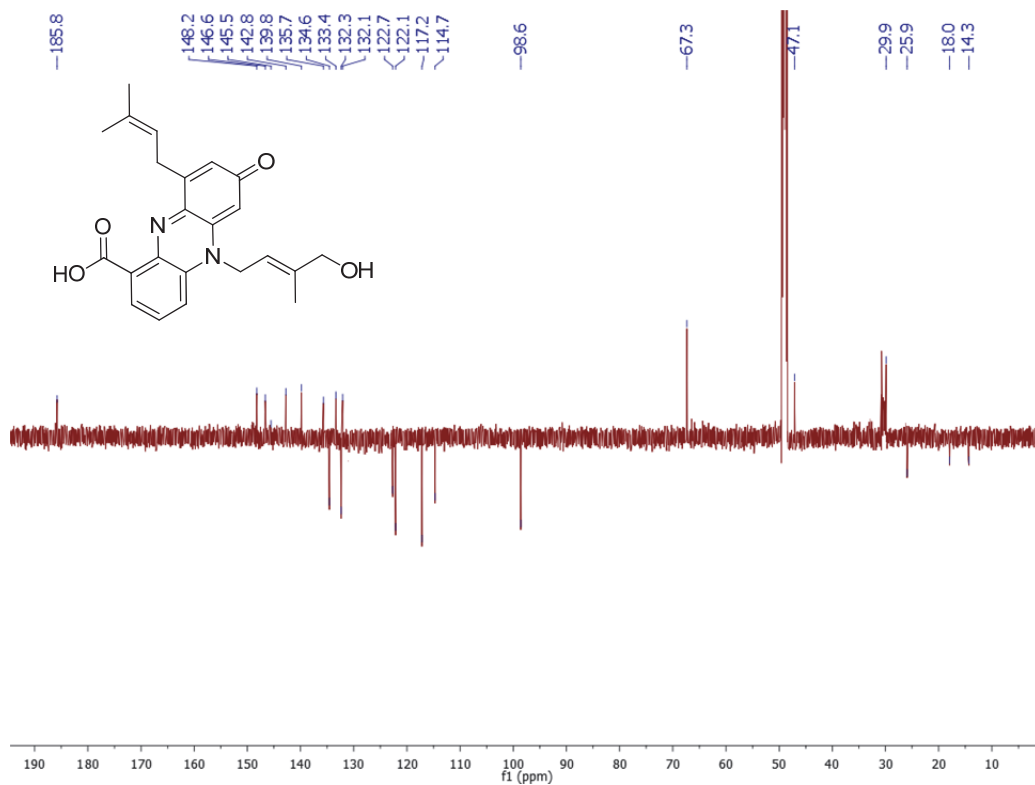
S33. IR spectrum of endophenaside D (**4**).

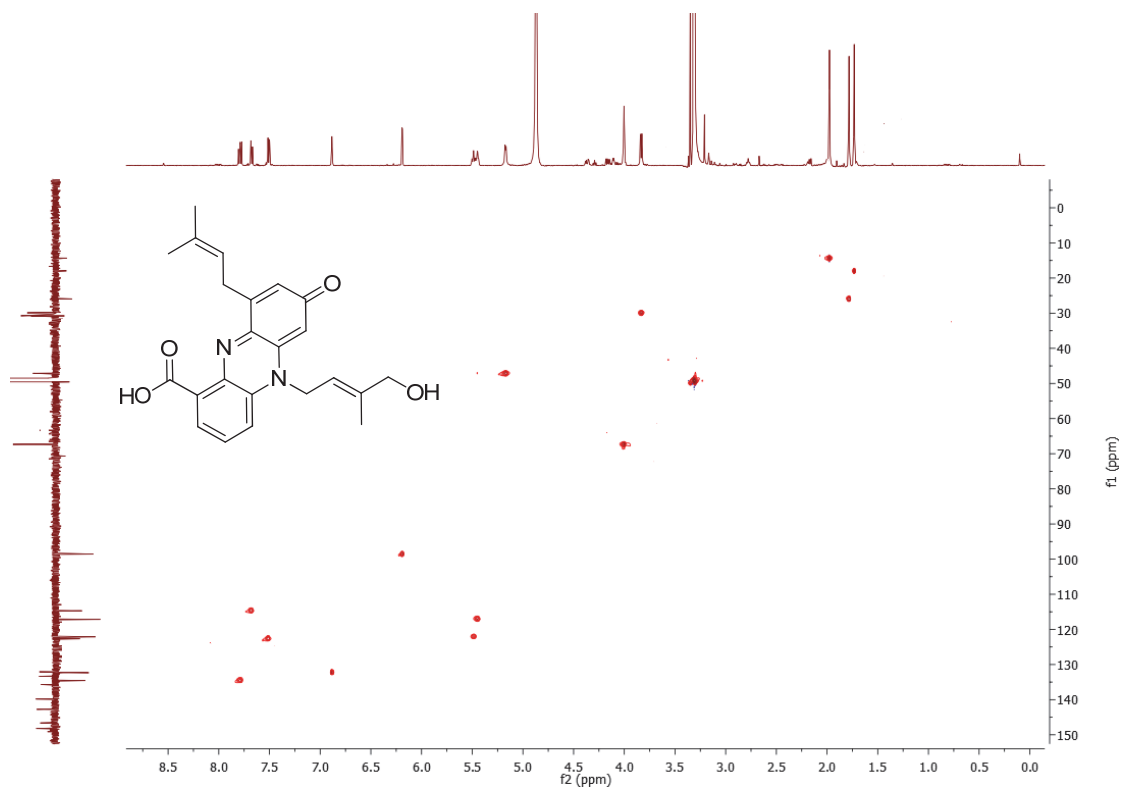
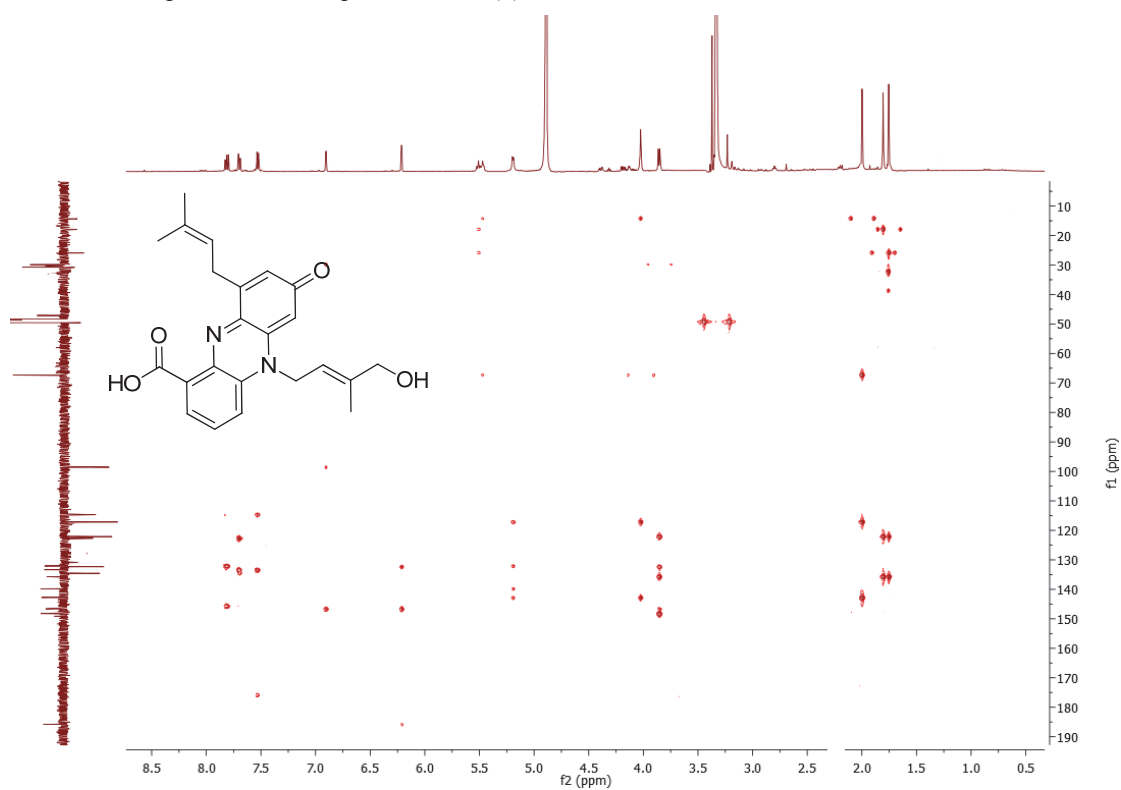
S34. UV spectrum of endophenaside D (**4**).S35. ^1H NMR spectrum (500 MHz) of endophenaside E (**5**) in CD_3OD .

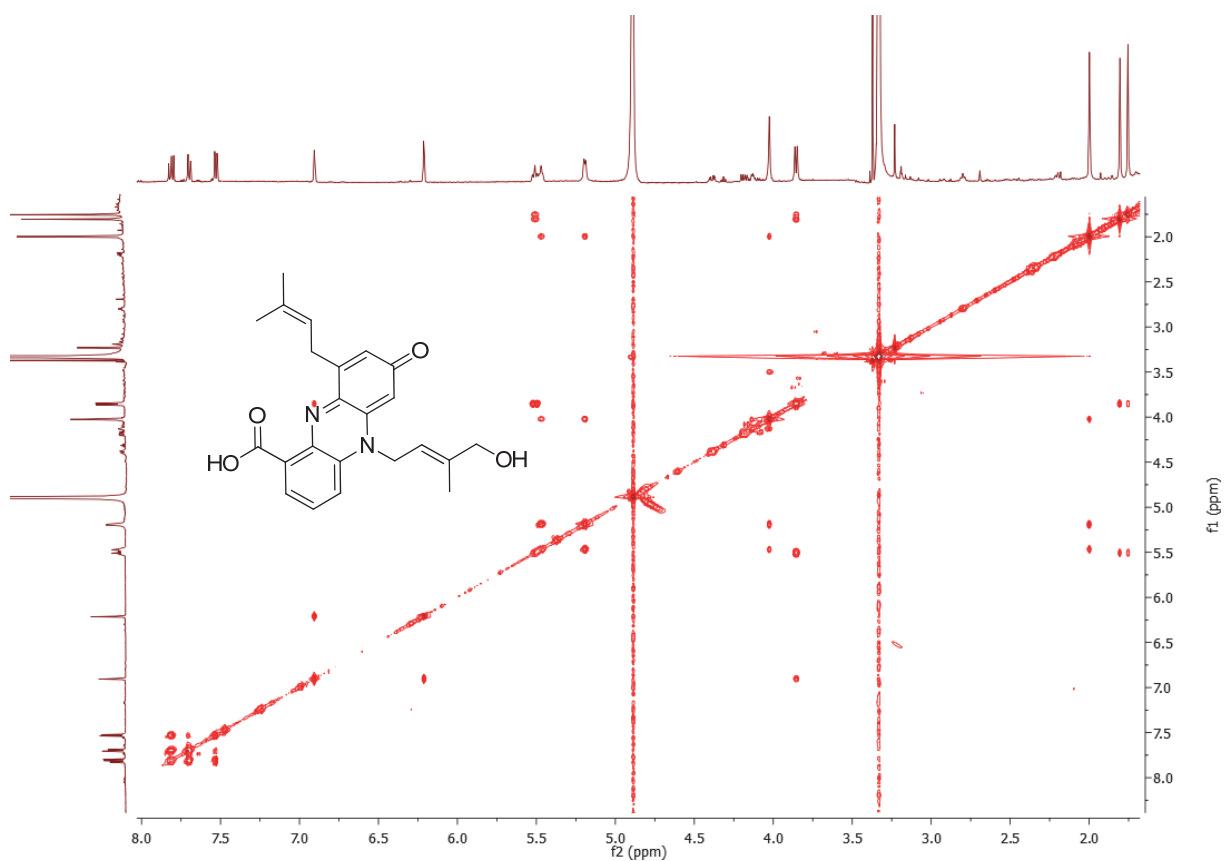
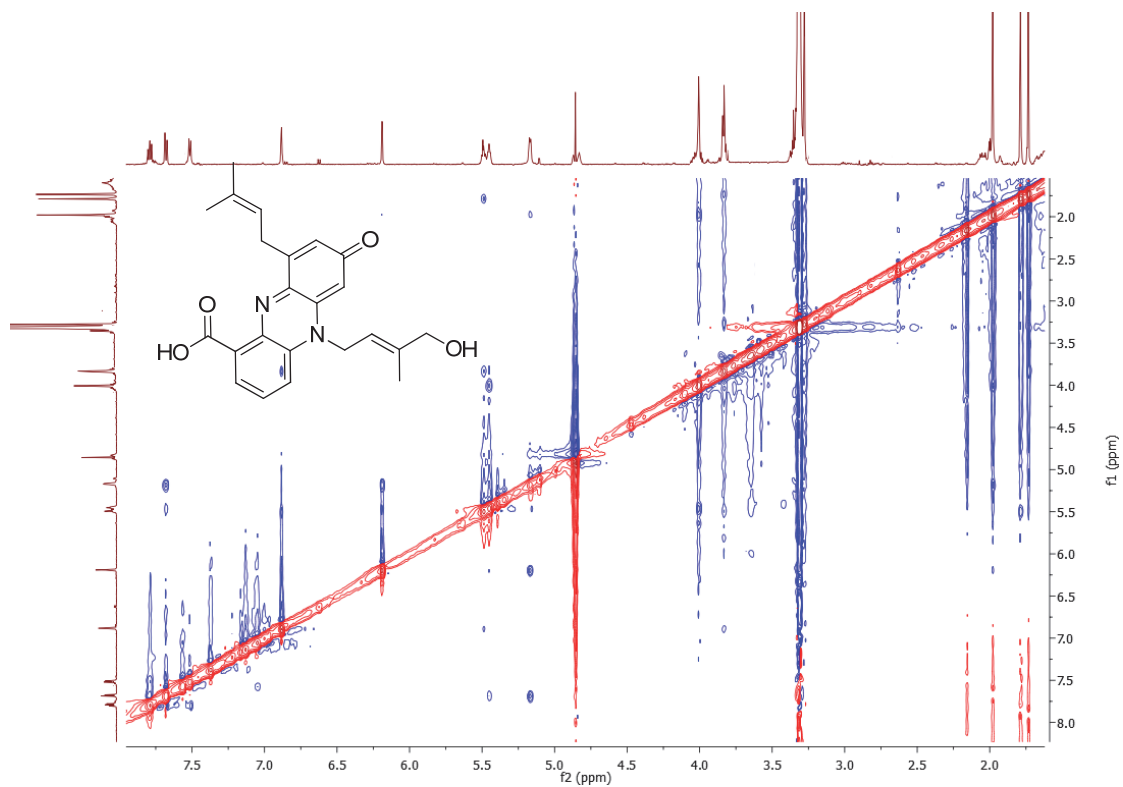
S36. HSQC spectrum of endophenaside E (**5**) in CD₃OD.S37. HMBC spectrum of endophenaside E (**5**) in CD₃OD.

S38. ^1H - ^1H COSY spectrum of endophenaside E (**5**) in CD_3OD .S39. HRMS spectrum of endophenaside E (**5**).

S40. IR spectrum of endophenaside E (**5**).S41. UV spectrum of endophenaside E (**5**).

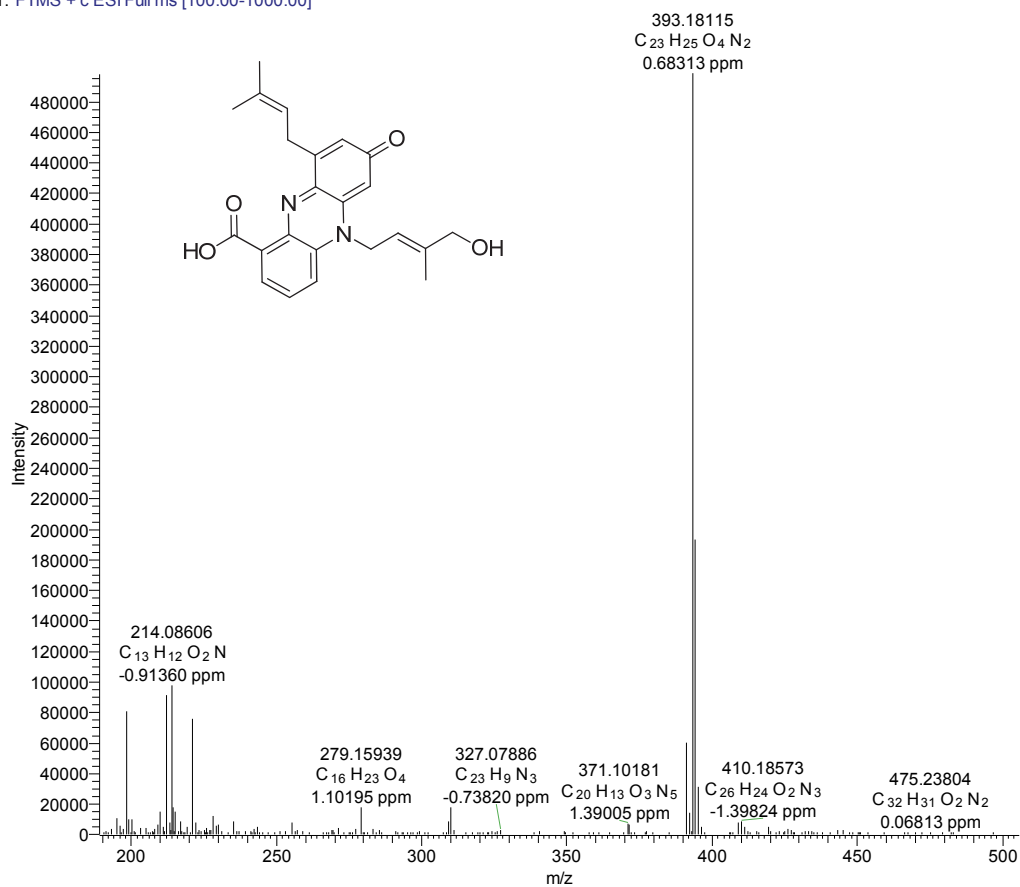
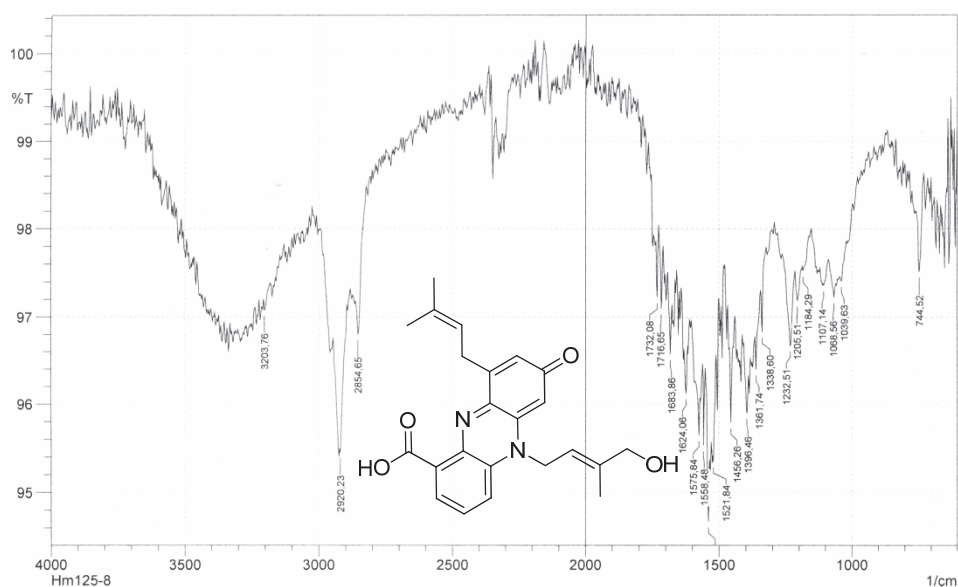
S42. ^1H NMR spectrum (600 MHz) of endophenazine F1 (**6**) in CD_3OD .S43. APT spectrum of endophenazine F1 (**6**) in CD_3OD .

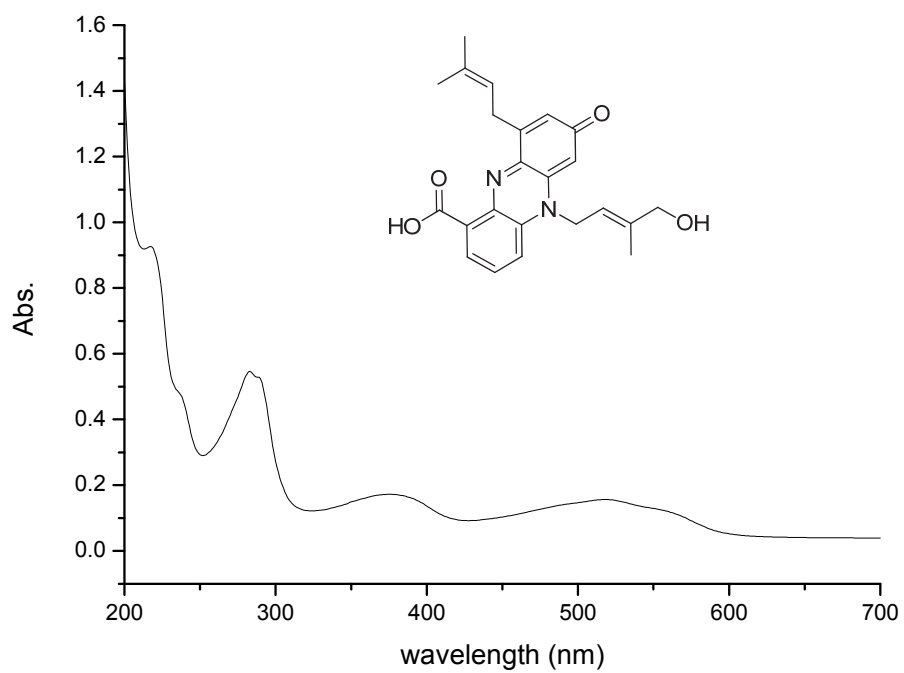
S44. HSQC spectrum of endophenazine F1 (**6**) in CD₃OD.S45. HMBC spectrum of endophenazine F1 (**6**) in CD₃OD.

S46. ^1H - ^1H COSY spectrum of endophenazine F1 (**6**) in CD_3OD .S47. ^1H - ^1H NOSEY spectrum of endophenazine F1 (**6**) in CD_3OD .

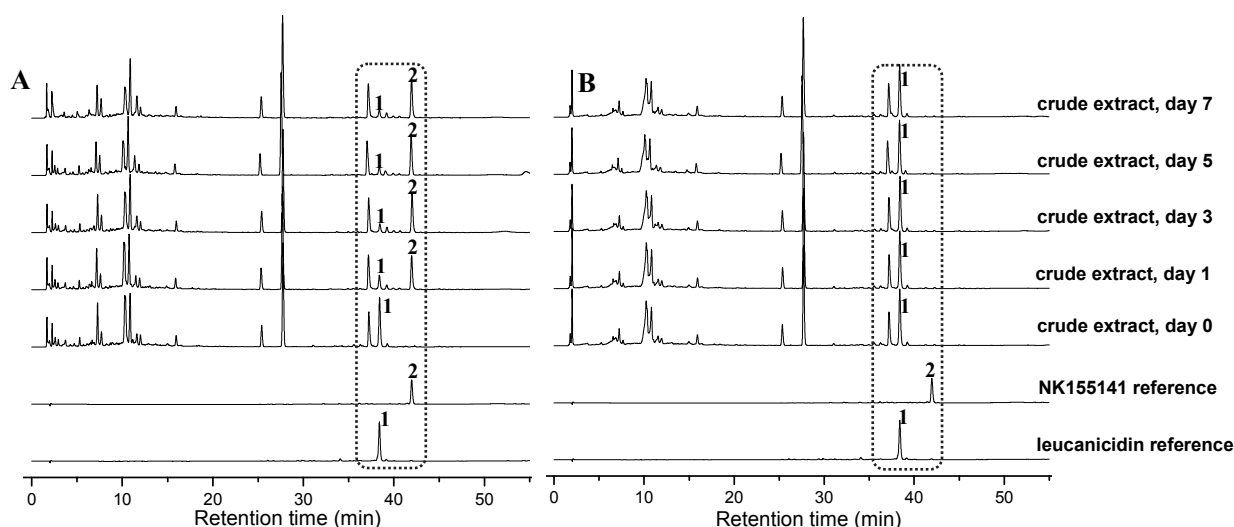
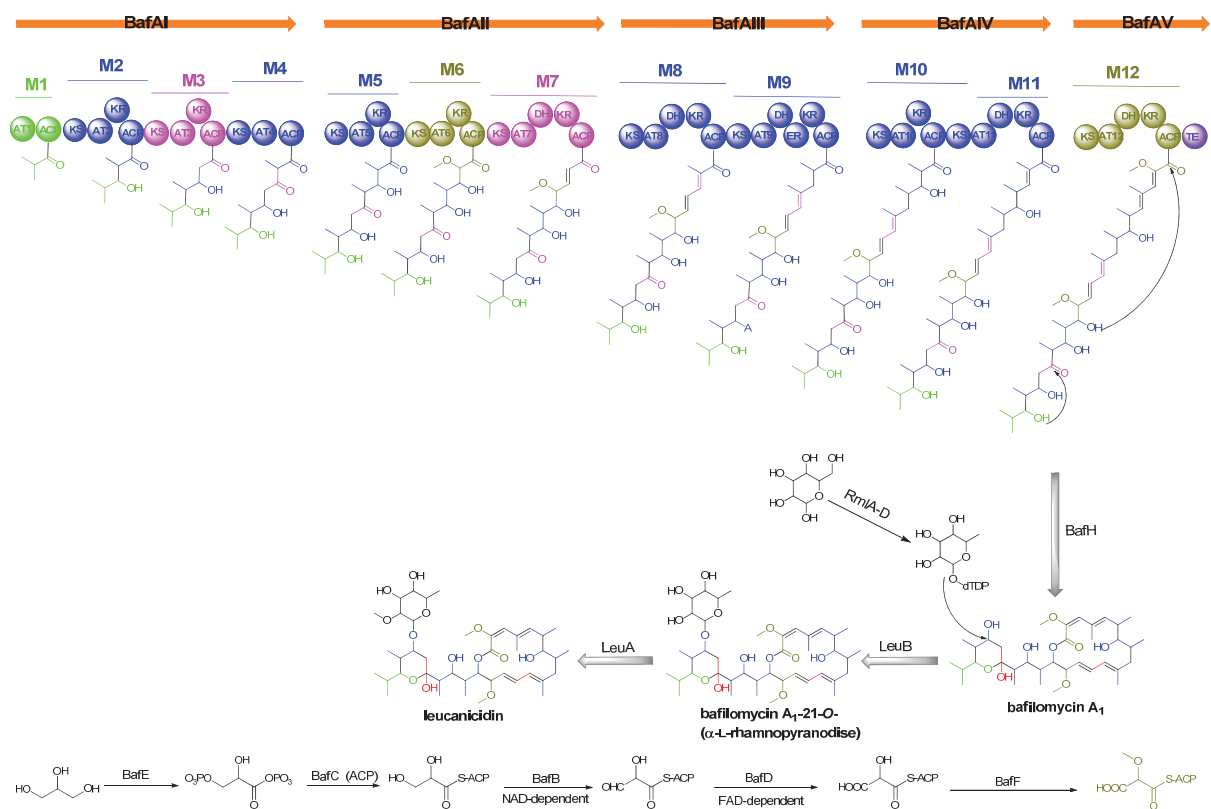
S48. HRMS spectrum of endophenazine F1 (**6**).

Hm125-8 #673 RT: 11.64 AV: 1 NL: 4.98E5
T: FTMS + c ESI Full ms [100.00-1000.00]

S49. IR spectrum of endophenazine F1 (**6**).

S50. UV spectrum of endophenazine F1 (**6**).

Appendix VII: Supplementary Information belonging to Chapter 10



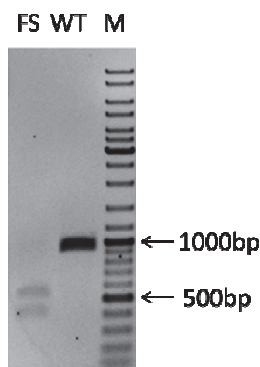


Figure S3. Verification of the frame-shift mutation in *leuB* by PCR. The -370/+584 region of *leuB* was amplified by PCR and separated electrophoretically on a 1.2% agarose gel in 1xTAE buffer. FS, PCR-amplified DNA fragment from one of ex-conjugants; WT, PCR-amplified *leuB* obtained from wild-type *Kitasatospora* sp. MBT66; M, DNA marker. 500 bp and 1000 bp bands of the DNA marker are indicated.

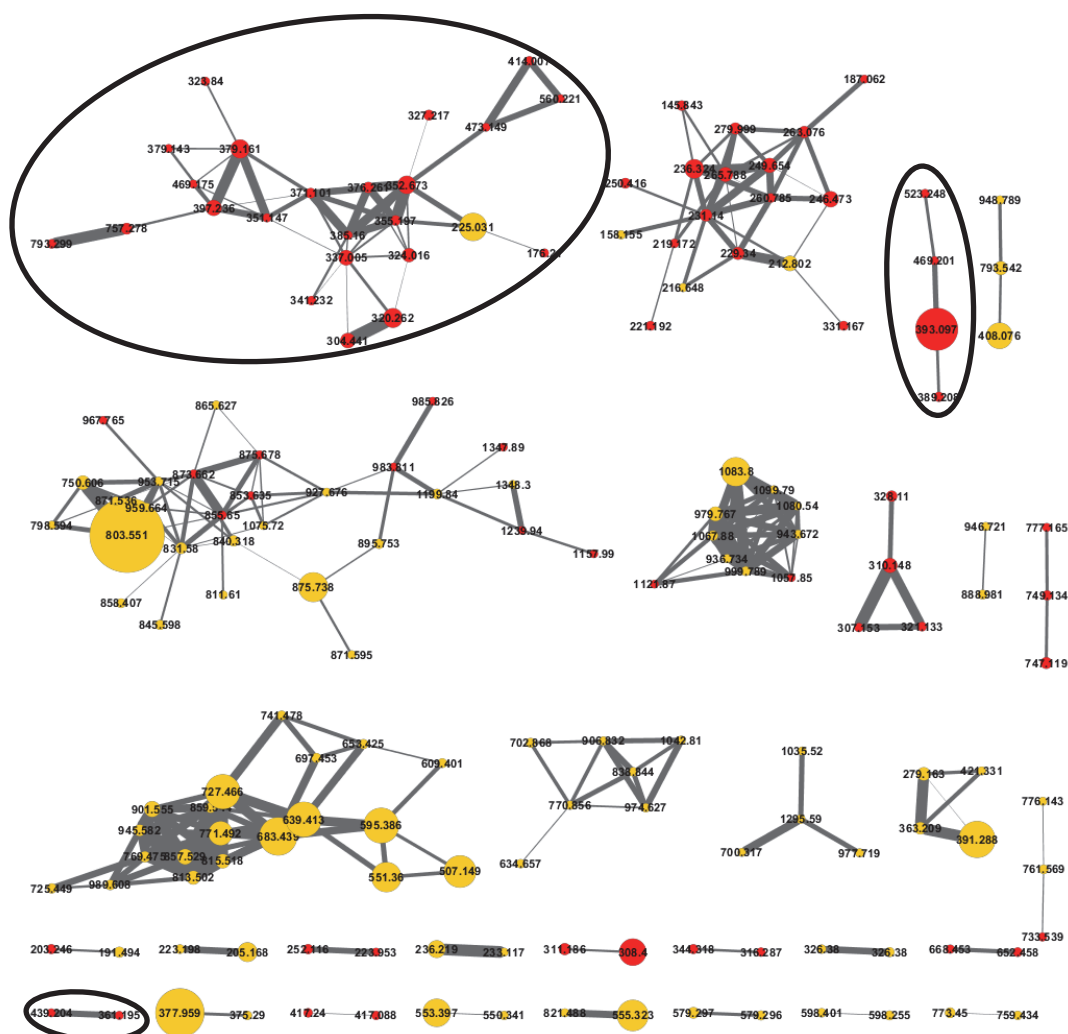


Figure S4. Molecular networking for all secondary metabolites harvested from *Kitasatospora* sp. MBT66 grown on MM + 1% glucose. Yellow nodes represent compounds that were simultaneously observed in the control experiment (blank culture medium) and bacterial extracts, while red nodes were exclusively found in the MBT66 extract. Circled are subnetworks for endophenazines and/or endophenazines.

Table S1. Oligonucleotides used in construction of *leuB* mutants using the CRISPR/Cas9 system

primer	sequence (5'→ 3') [#]
LeuB_spacer_For	ACGCTTCACCGACGACCGCCAGAG
LeuB_spacer_Rev	AAACCTCTGGCGGTCGTCGGTGAA
LeuB_LF-976_EX	CAGT <u>GAATTCTCTAGA</u> ATCACCTGGCACGGCATCGAG
LeuB_LR+33_H	CTGACATGAG <u>AAGCTT</u> GAAGACATGGCCGTGGGCCGGGAT
LeuB_FS_RF+47_H	CTGACAGTAG <u>AAGCTT</u> TCGTCGGTGAACTGGTACGCCG
LeuB_RR+1032_EX	CTGAG <u>GAATTCTCTAGA</u> CAGCAGGCTCACGTCGACCTCC
LeuB_F-370	CAACTGCTCGACGGCCTCAAC
LeuB_R+584	TCGTCGAAGCTCTCGCCGTAC

[#]Restriction sites used for cloning are underlined and in bold face. GAATTC, EcoRI; TCTAGA, XbaI; AAGCTT, HindIII.

Table S2. Signal intensities for the masses of the secondary metabolites harvested from *Kitasatospora* sp. MBT66 grown on MM + 1% glucose.

Mass	Intensity	Mass	Intensity	Mass	Intensity	Mass	Intensity
145.843	611720	331.167	390584	634.657	213764	865.627	204844
158.155	149132	337.005	2254280	639.413	8876750	871.536	58668
176.21	113348	341.232	281596	652.458	327564	871.595	105668
187.062	55008	344.318	458592	653.425	442716	873.662	197120
191.494	777504	351.147	131356	668.453	197540	875.678	181480
203.246	442216	352.673	3224250	683.439	9768180	875.738	6980100
205.168	3631080	355.197	638224	697.453	458052	888.981	397808
212.802	2256640	361.195	90196	700.317	48700	895.753	495132
216.648	103164	363.209	1276040	702.868	383580	901.555	2490700
219.172	265696	371.101	633928	725.449	508188	906.832	209588
221.192	383768	375.29	206656	727.466	8745590	927.676	83500
223.198	331496	376.261	912612	733.539	144124	936.734	171248
223.953	70096	377.959	1.33E+07	741.478	395956	943.672	191484
225.031	6403020	379.143	66380	747.119	1072710	945.582	1080510
229.34	1403060	379.161	3433840	749.134	518052	946.721	120304
231.14	1317040	385.16	249796	750.606	1882480	948.789	201756
233.117	24360	389.208	506440	757.278	1005690	953.715	269248
236.219	3001340	391.288	9285400	759.434	201568	959.664	150388
236.324	3589910	393.097	1.12E+07	761.569	309696	967.765	19080
246.473	2493900	397.236	2480790	769.475	869976	974.627	194080
249.654	2102040	408.076	5840700	770.856	215160	977.719	111076
250.416	690616	414.007	318992	771.492	5210540	979.767	1996310
252.116	499196	417.088	206896	773.45	644072	983.811	121616
260.785	368800	417.24	267976	776.143	65944	985.826	22652
263.076	1118620	421.331	73448	777.165	84452	989.608	596864
265.788	2099370	439.204	56852	793.299	329708	999.789	402352

279.163	1632550	469.175	65140	793.542	1721760	1035.52	62748
279.999	1153380	469.201	128208	798.594	85912	1042.81	116364
304.441	2211310	473.149	66048	803.551	2.19E+07	1057.85	38840
307.153	249204	507.149	7845160	811.61	58148	1067.88	606872
308.4	6145340	523.248	216752	813.502	1785100	1075.72	49324
310.148	2129790	550.341	160000	815.518	1924600	1080.54	24352
311.186	1143310	551.36	7095300	821.488	85840	1083.8	6809140
316.287	227768	553.397	6451200	831.58	1028170	1099.79	38856
320.262	3680420	555.323	7501280	838.844	242556	1121.87	28828
321.133	360704	560.221	45760	840.318	822044	1157.99	32684
323.84	162548	579.296	411056	845.598	470628	1199.84	579312
324.016	1797830	579.297	975100	853.635	67244	1239.94	49932
326.38	232984	595.386	8902080	855.65	309252	1295.59	211464
326.38	430364	598.255	387284	857.529	2381860	1347.89	26764
327.217	704036	598.401	261604	858.407	335136	1348.3	49268
328.11	874660	609.401	363288	859.544	616224		

Table S3. ¹H NMR data assignments for new endophenazines **6—8**, **16**, and **22—26**.

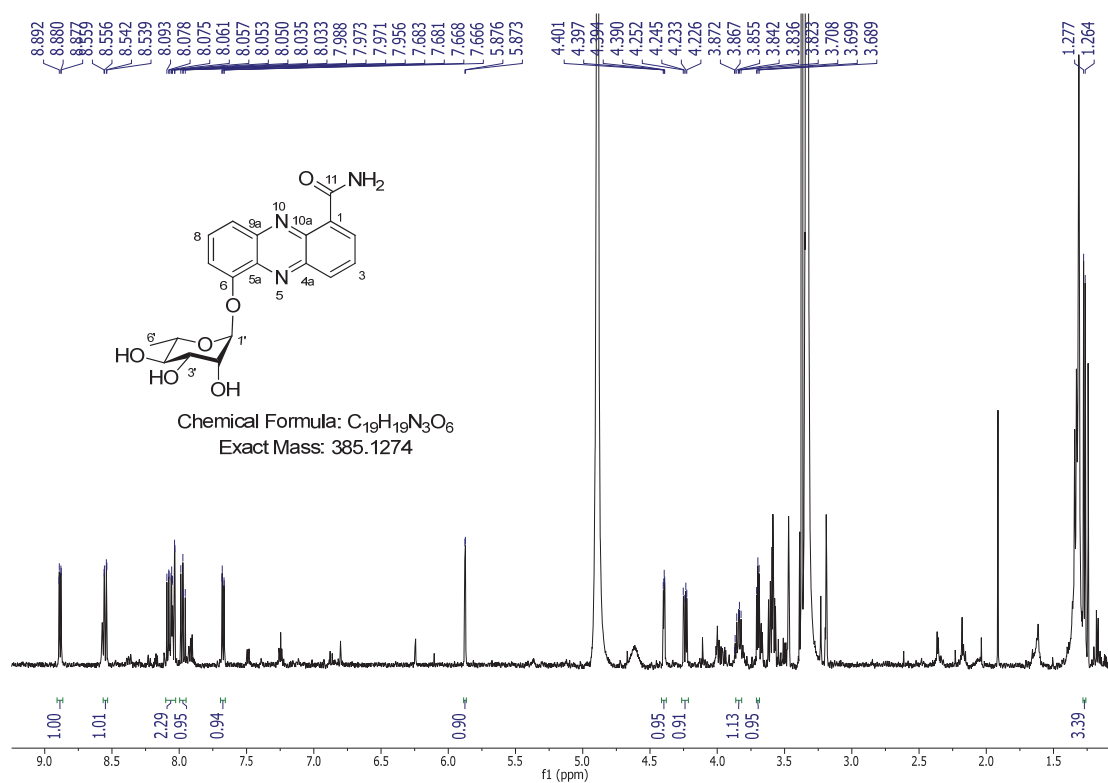
NO.	6	7	8	16	22	23	24	25	26
2	8.89, dd (7.5, 1.5)	8.26, dd (7.2, 1.2)	8.25, dd (7.2, 1.2)	8.38, dd (7.2, 1.2)	8.48, dd (9.0, 1.2)	8.38, dd (7.2, 1.2)	8.38, dd (7.2, 1.2)	7.87, dd (8.4, 1.2)	7.88, dd (7.8, 1.2)
3	8.08, dd (9.0, 7.5)	8.00, dd (9.0, 7.2)	7.99, dd (9.0, 7.2)	8.02, dd (8.4, 7.2)	8.05, dd (9.0, 7.2)	8.02, dd (8.4, 7.2)	8.02, dd (8.4, 7.2)	7.23, m	7.23, m
4	8.55, dd (9.0, 1.5)	8.41, dd (9.0, 0.6)	8.40, dd (9.0, 0.6)	8.46, dd (8.4, 1.2)	8.48, dd (7.2, 1.2)	8.46, dd (8.4, 1.2)	8.46, dd (8.4, 1.2)	7.65, brd (7.8)	7.65, dd (7.8, 1.2)
6		8.09, brd (9.0)	8.08, brd (9.0)	8.28, m	8.57, d (1.2)	8.28, m	8.28, m	7.39, brd (8.4)	7.39, brd (7.8, 1.2)
7	7.68, dd (7.5, 1.0)	7.91, dd (9.0, 7.0)	7.90, dd (9.0, 7.0)	8.02, m		8.02, m	8.02, m	7.06, m	7.06, m
8	7.97, dd (8.5, 7.5)	7.77, brd (6.6)	7.76, brd (6.6)	8.01, m	8.34, dd (9.0, 1.2)	8.01, m	8.01, m	7.07, m	7.07, m
9	8.04, dd (8.5, 1.0)			8.01, m	8.60, d (9.0)	8.01, m	8.01, m	7.05, m	7.05, m
1'	5.87, d (1.5)	6.55, d (1.8)	6.40, d (1.2)	6.54, d (1.8)	6.39, d (1.8)	4.63, d (10.8, 4.8); 4.57, d (10.8, 6.0)	3.96, d (10.8, 4.8); 3.92, d (10.8, 6.0)	6.40, d (1.8)	6.25, d (1.8)
2'	4.40, dd (3.5, 2.0)	3.78, dd (3.6, 1.8)	4.13, dd (3.6, 1.8)	3.78, dd (3.0, 1.8)	4.13, dd (3.0, 1.8)	4.12, m	5.43, m	3.64, dd (3.0, 1.8)	3.99, dd (3.0, 1.8)
3'	4.24, dd (9.5, 3.5)	3.75, dd (9.6, 3.6)	3.72, dd (9.6, 3.0)	4.08, dd (9.6, 3.0)	4.05, dd (9.6, 3.0)	3.82, d (11.4, 5.4); 3.78, d (11.4, 5.4)	3.96, d (10.8, 4.8); 3.92, d (10.8, 6.0)	3.88, dd (9.6, 3.0)	3.85, dd (9.6, 3.0)
4'	3.59, t (9.5)	3.44, t (9.6)	3.52, t (9.6)	3.54, t (9.6)	3.60, t (9.6)			3.46, t (9.6)	3.52, t (9.6)
5'	3.84, m	3.78, m	3.82, m	4.16, m	4.24, m			3.75, m	3.77, m
6'	1.27, d (6.5)	1.30~1.35 d (6.6)	1.34, d (6.6)	1.33, d (6.6)	1.38, d (6.6)			1.25~1.35, d (6.6)	1.25~1.35, d (6.6)
7'		3.62, s		3.64, s				3.57, s	
1''		4.17, brd (7.2)	4.15, brd (7.2)						
2''		5.58, m	5.57, m					4.31, s	4.31, s
3''		1.83, s	1.82, s		7.96, dd (8.4, 1.2)				
4''		1.79, s	1.78, s		7.63, t (7.8)				
5''					7.75, m				
6''					7.63, t (7.8)				
7''					7.96, dd (8.4, 1.2)				

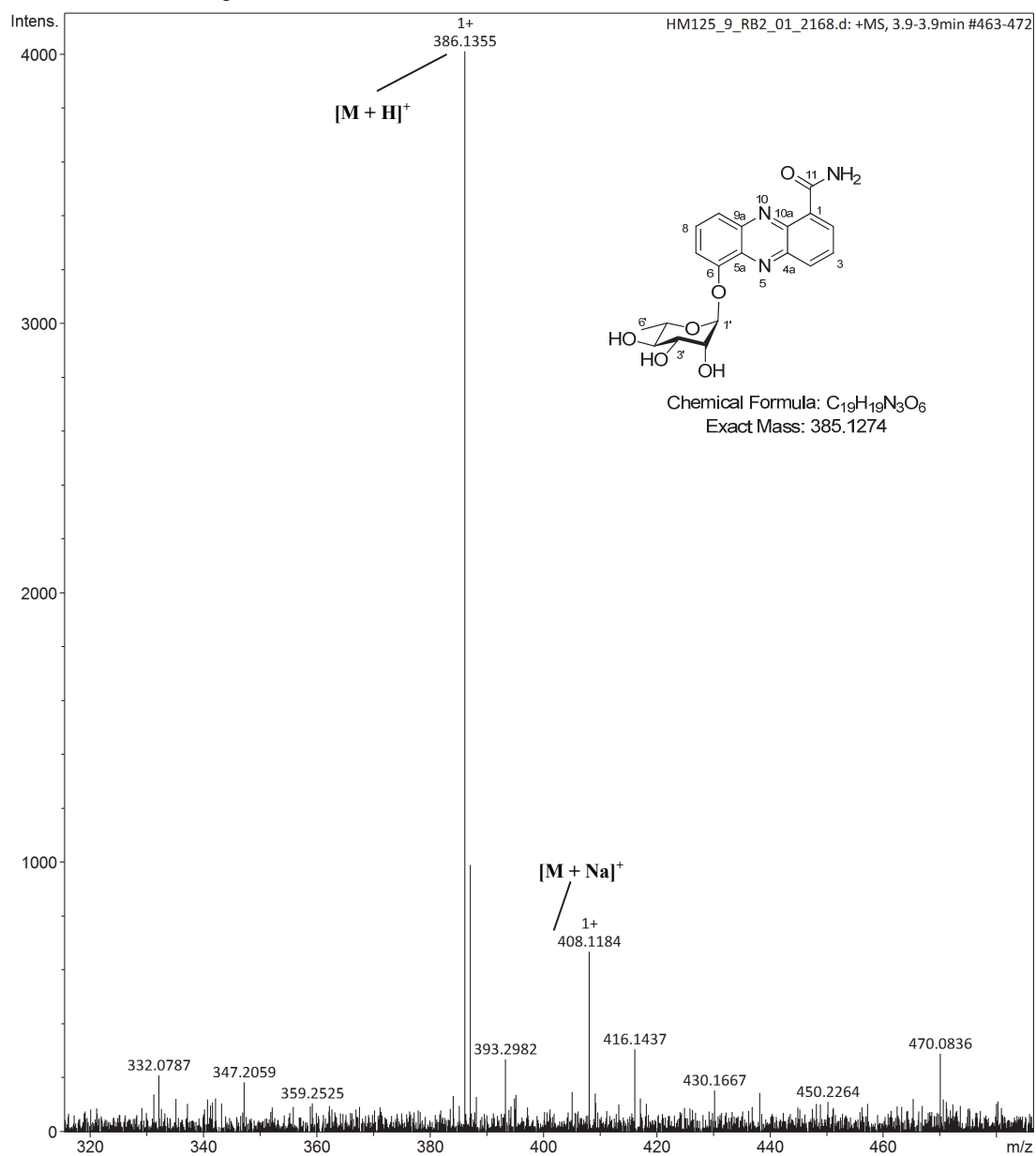
Recorded in CD₃OD. All ¹H NMR spectra were recorded at 600 MHz, except **6** at 500 MHz. Proton coupling constants (*J*) are in Hz.

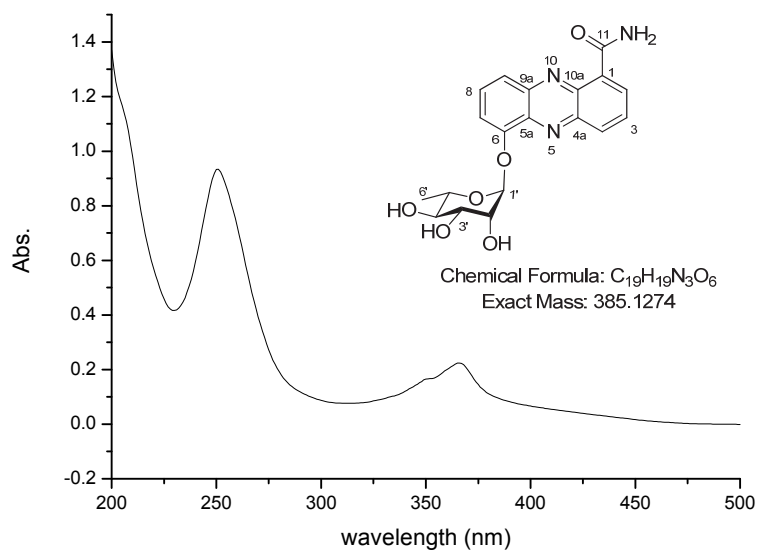
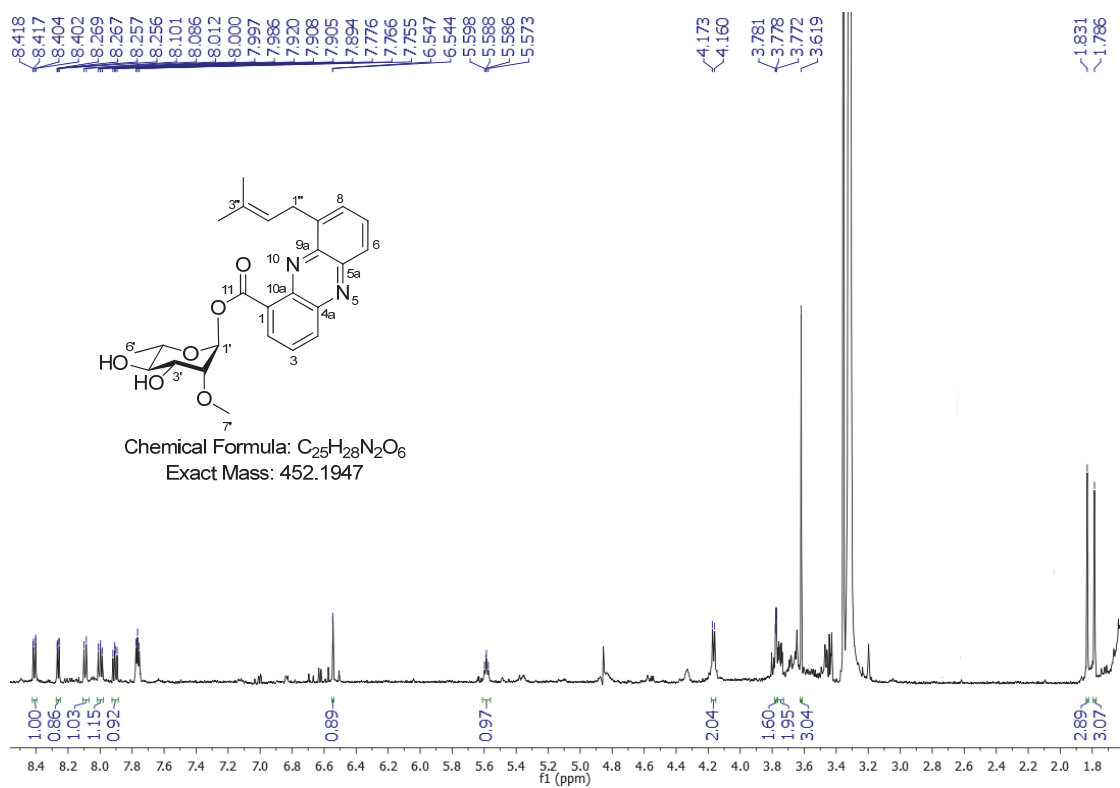
Spectral file. Spectra of nine new phenazine glycosides **6—8**, **16**, and **22—26**, including ^1H NMR, MS, and UV.

- S1. ^1H NMR spectrum of compound **6** in CD_3OD .
- S2. HRESIMS of compound **6**.
- S3. UV spectrum of compound **6**.
- S4. ^1H NMR spectrum of compound **7** in CD_3OD .
- S5. HRESIMS of compound **7**
- S6. UV spectrum of compound **7**
- S7. ^1H NMR spectrum of compound **8** in CD_3OD .
- S8. HRESIMS of compound **8**
- S9. UV spectrum of compound **8**
- S10. ^1H NMR spectrum of compound **16** in CD_3OD .
- S11. HRESIMS of compound **16**.
- S12. UV spectrum of compound **16**.
- S13. ^1H NMR spectrum of compound **22** in CD_3OD .
- S14. HRESIMS of compound **22**.
- S15. UV spectrum of compound **22**.
- S16. ^1H NMR spectrum of compounds **23** and **24** in CD_3OD .
- S17. APT spectrum of compounds **23** and **24** in CD_3OD .
- S18. ^1H - ^1H COSY spectrum of compounds **23** and **24** in CD_3OD .
- S19. HSQC spectrum of compounds **23** and **24** in CD_3OD .
- S20. HMBC spectrum of compounds **23** and **24** in CD_3OD .
- S21. HRESIMS of compounds **23** and **24**.
- S22. UV spectrum of compounds **23** and **24**.
- S23. ^1H NMR spectrum of compound **25** in CD_3OD .
- S24. HRESIMS of compound **25**.
- S25. UV spectrum of compound **25**.
- S26. ^1H NMR spectrum of compound **26** in CD_3OD .
- S27. HRESIMS of compound **26**.
- S28. UV spectrum of compound **26**.

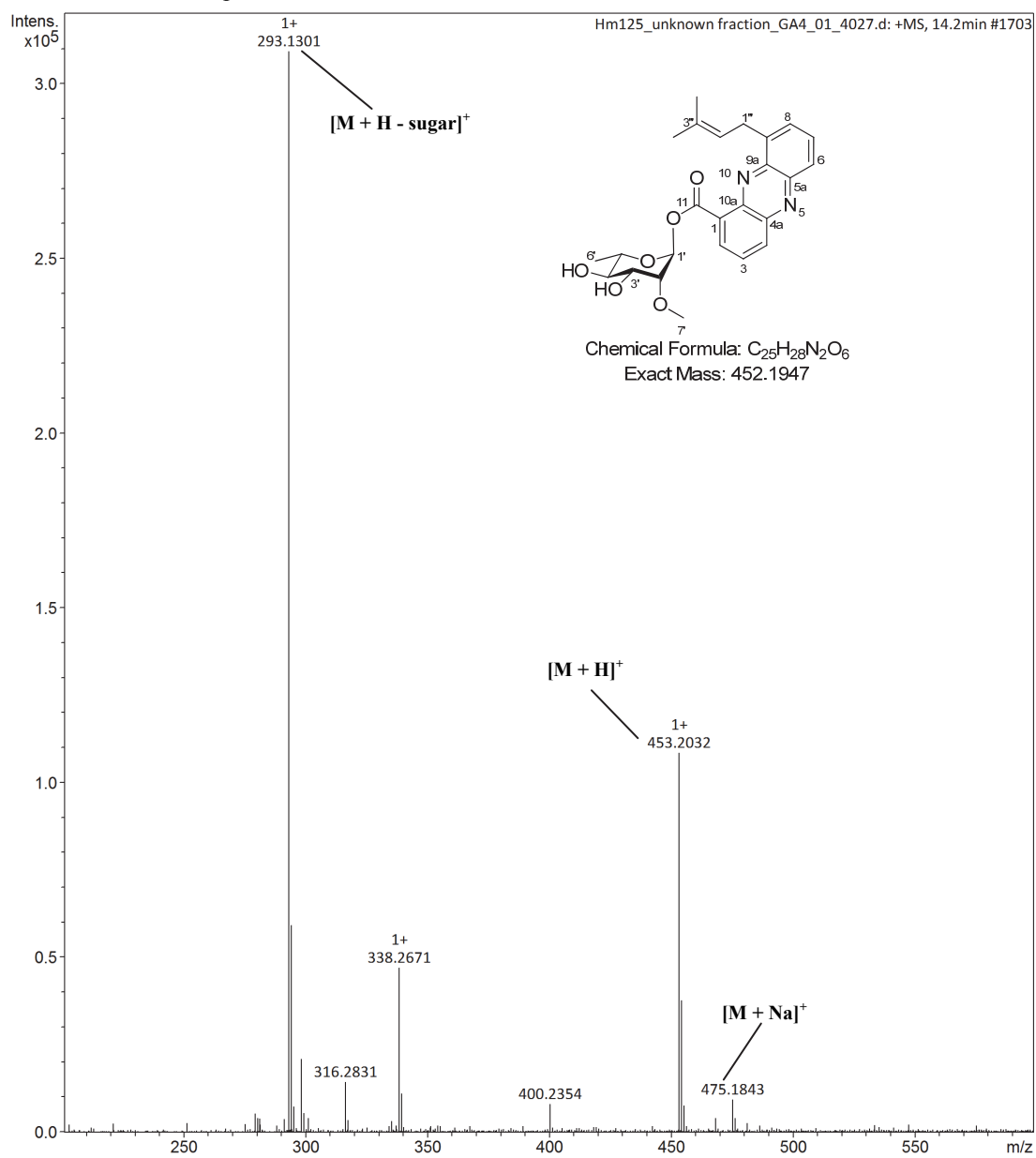
S1. ^1H NMR spectrum of compound **6** in CD_3OD .

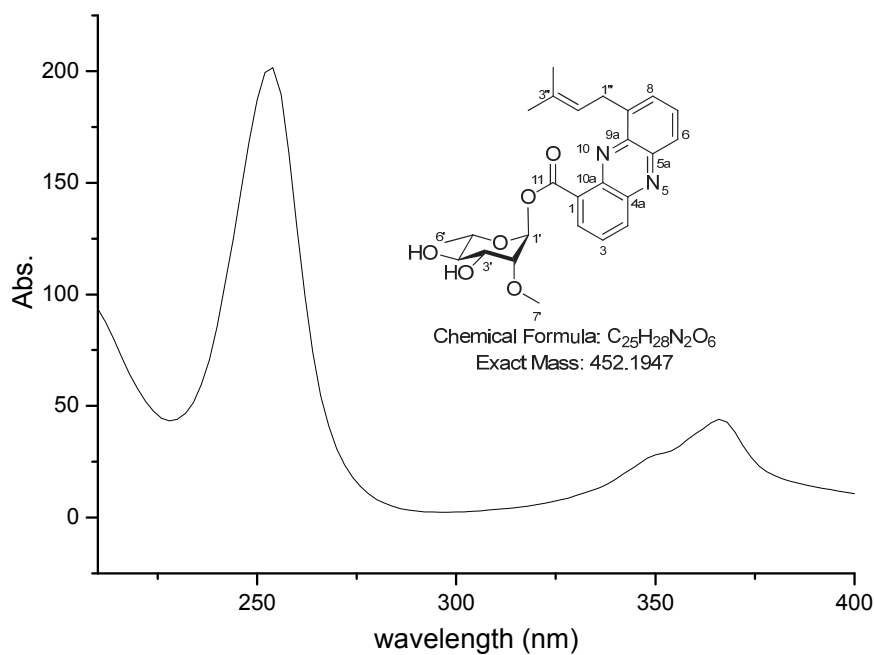
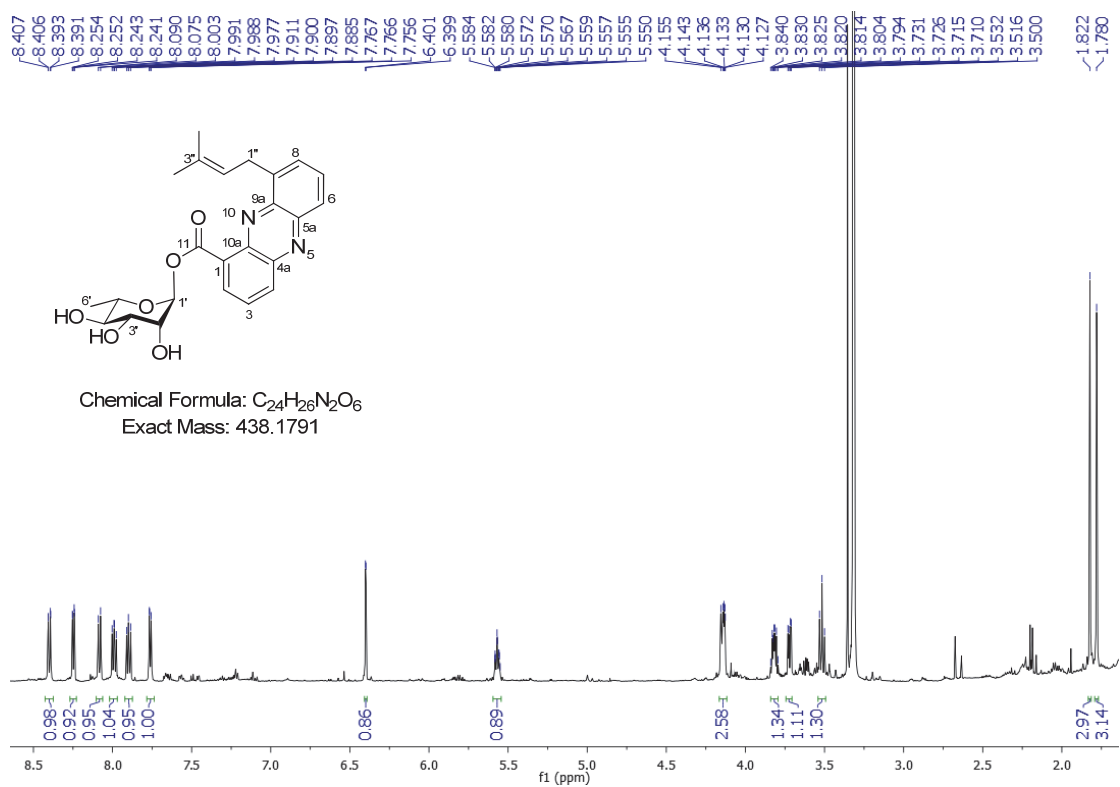


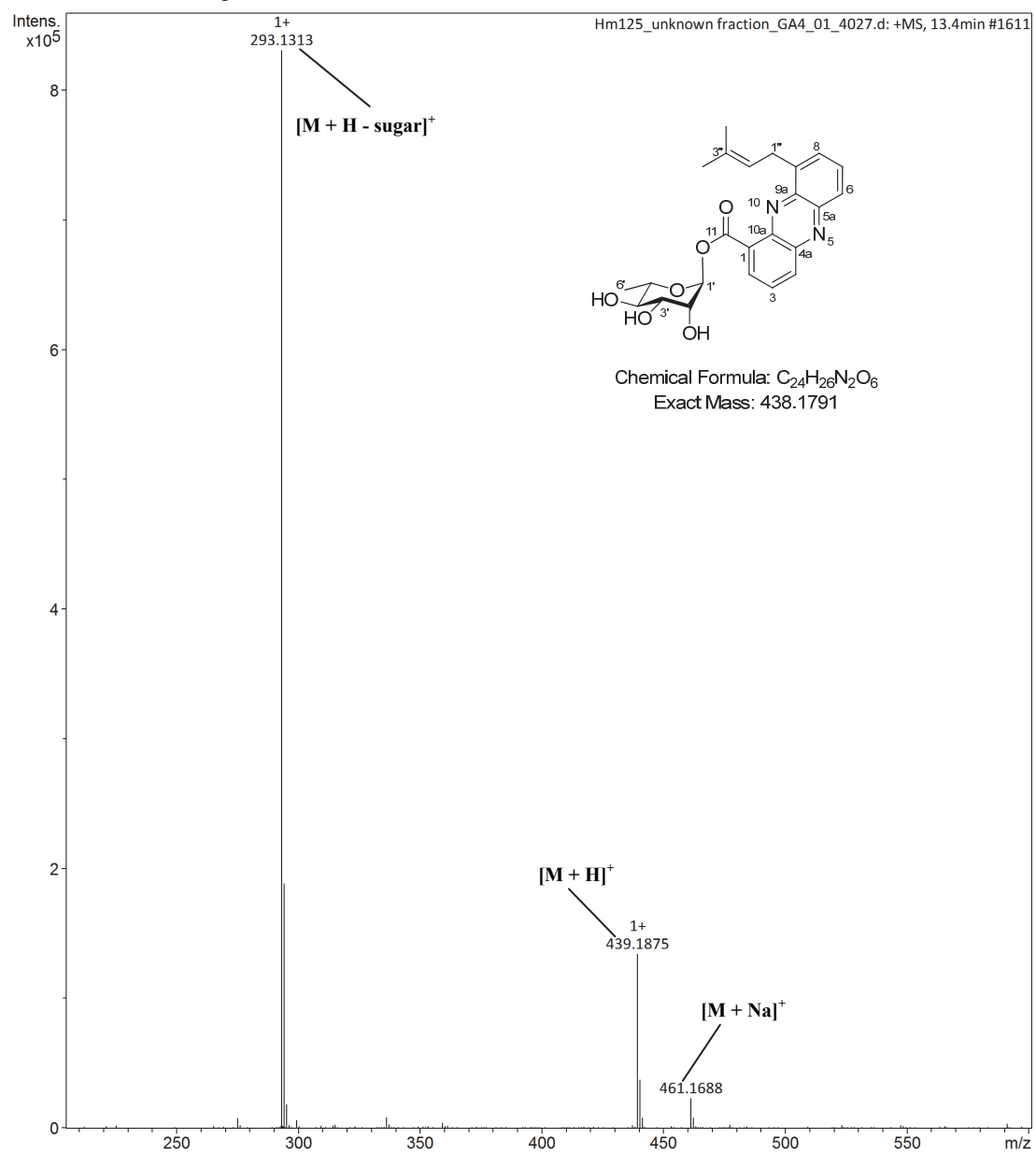
S2. HRESIMS of compound **6**.

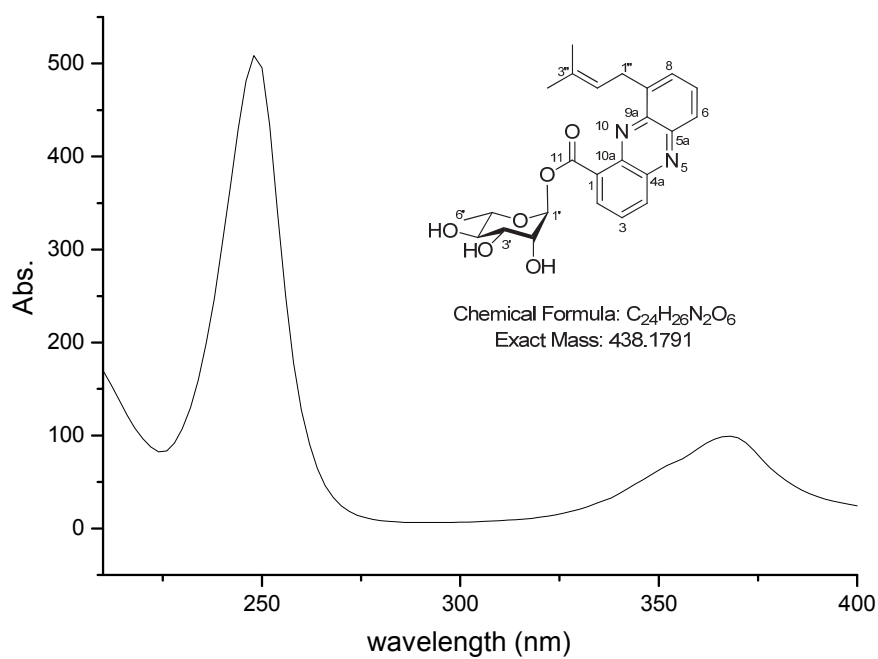
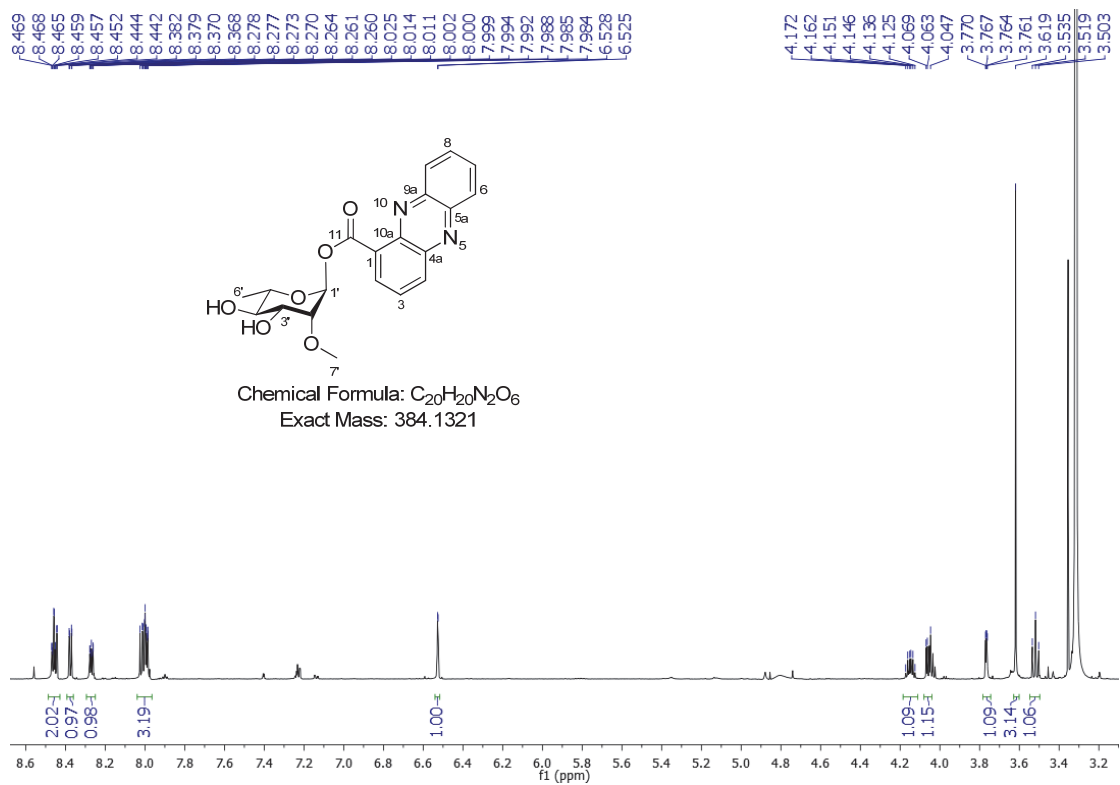
S3. UV spectrum of compound **6**.S4. 1H NMR spectrum of compound **7** in CD_3OD .

S5. HRESIMS of compound 7

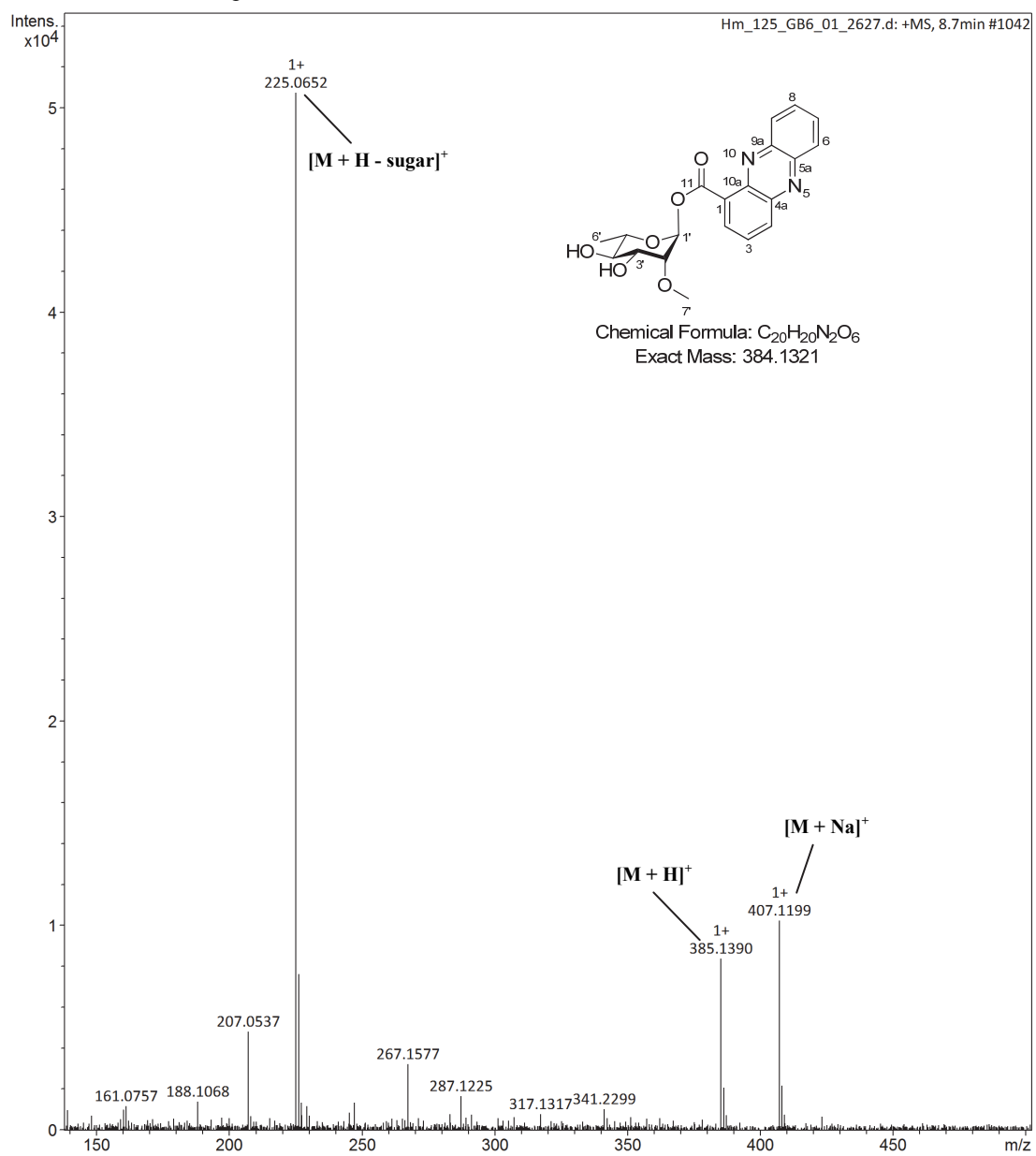


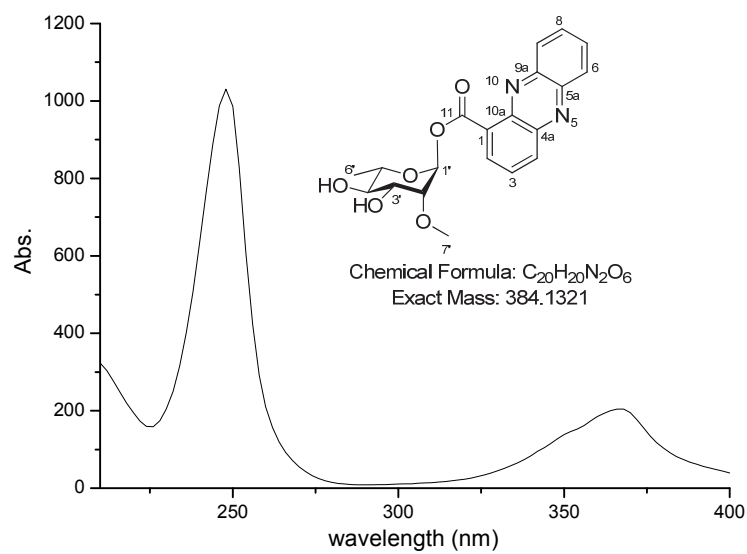
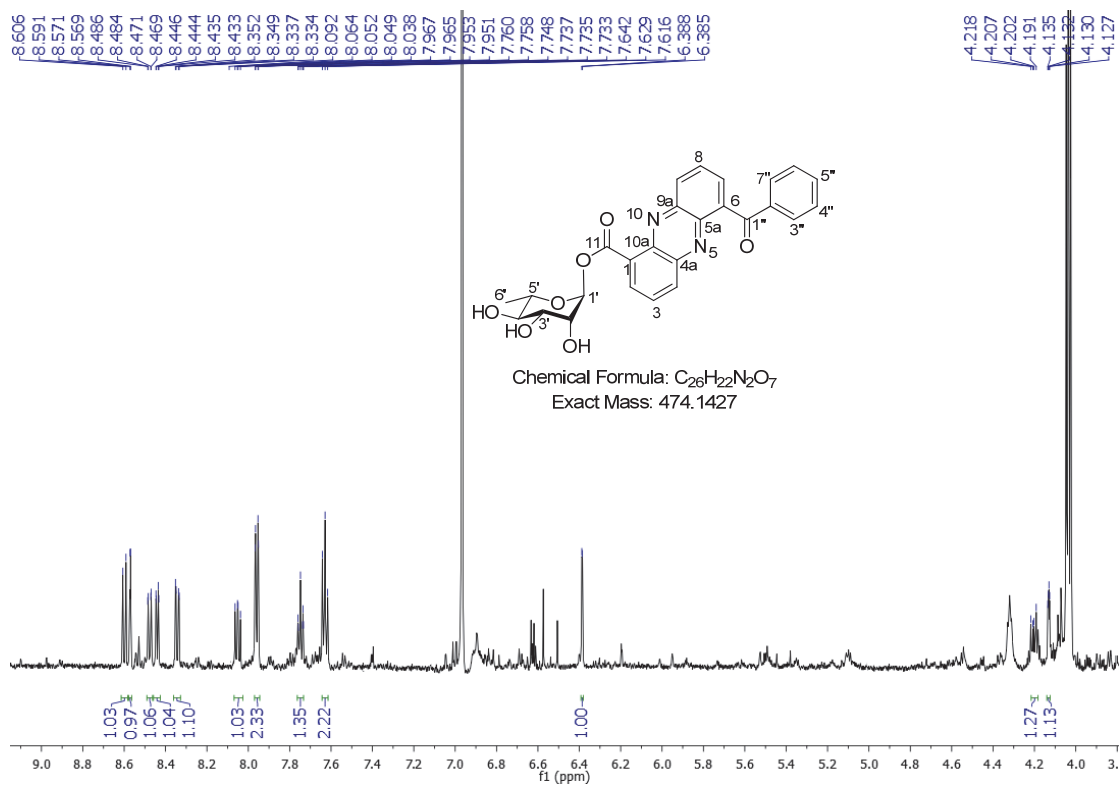
S6. UV spectrum of compound **7**S7. 1H NMR spectrum of compound **8** in CD_3OD .

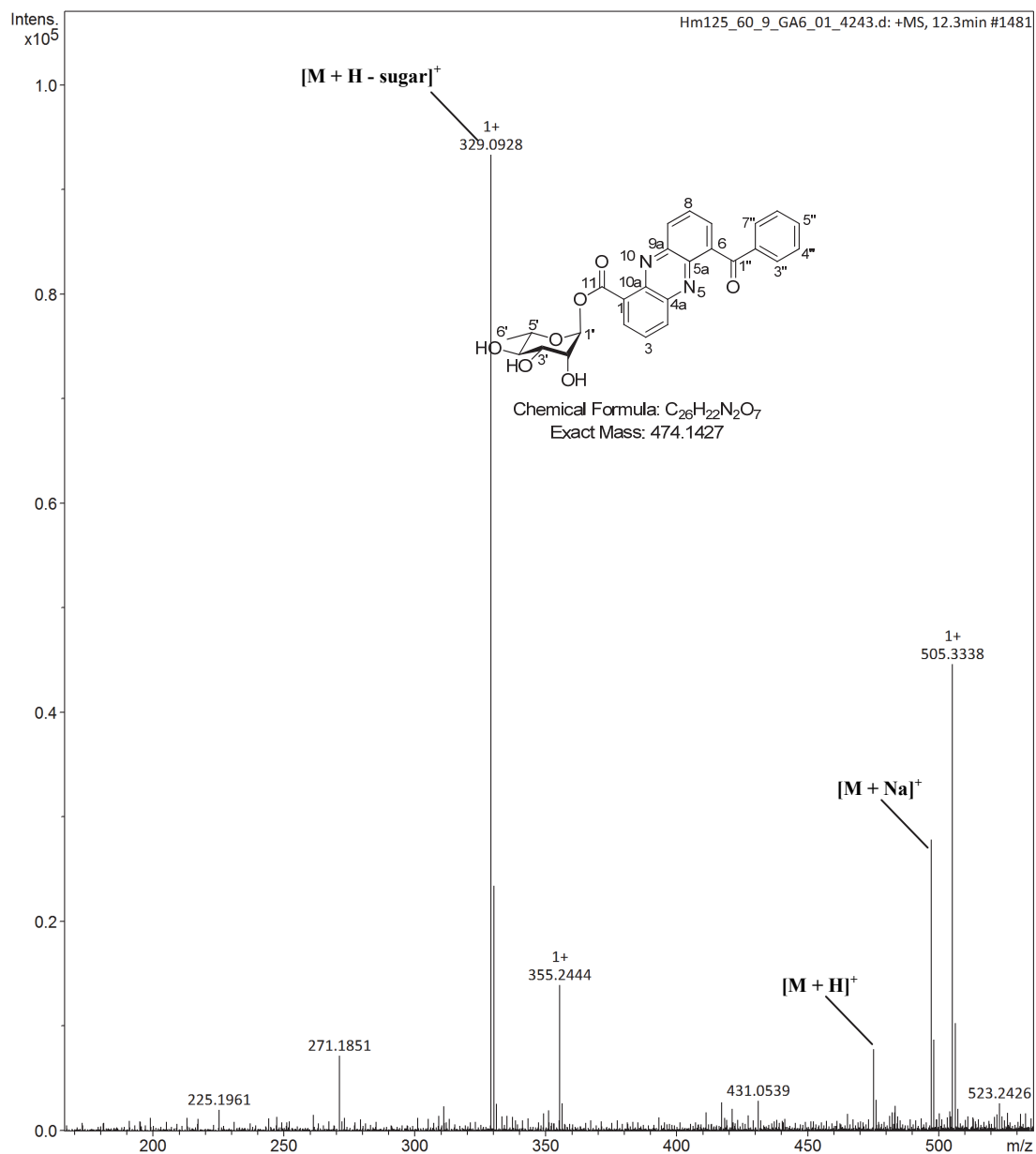
S8. HRESIMS of compound **8**

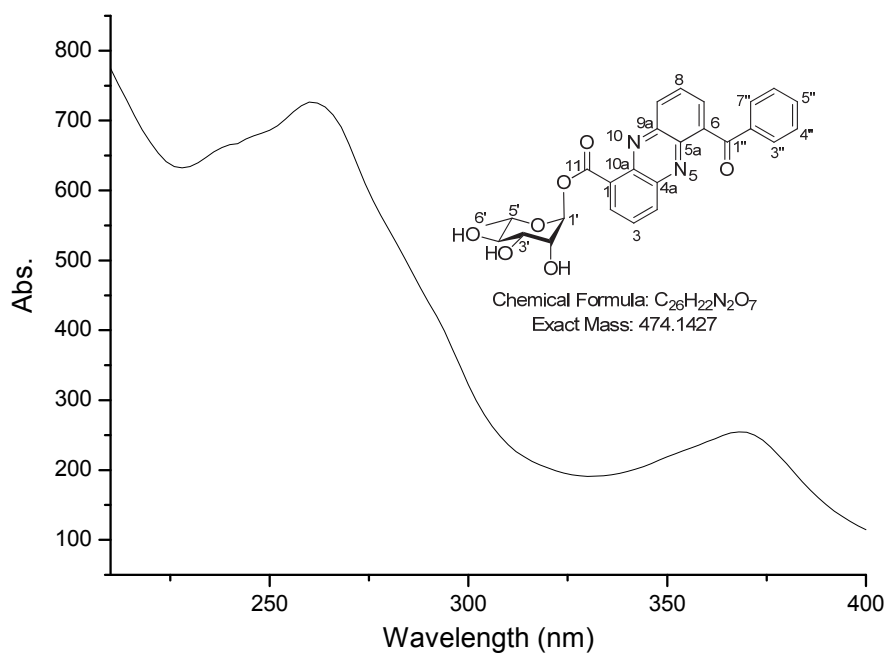
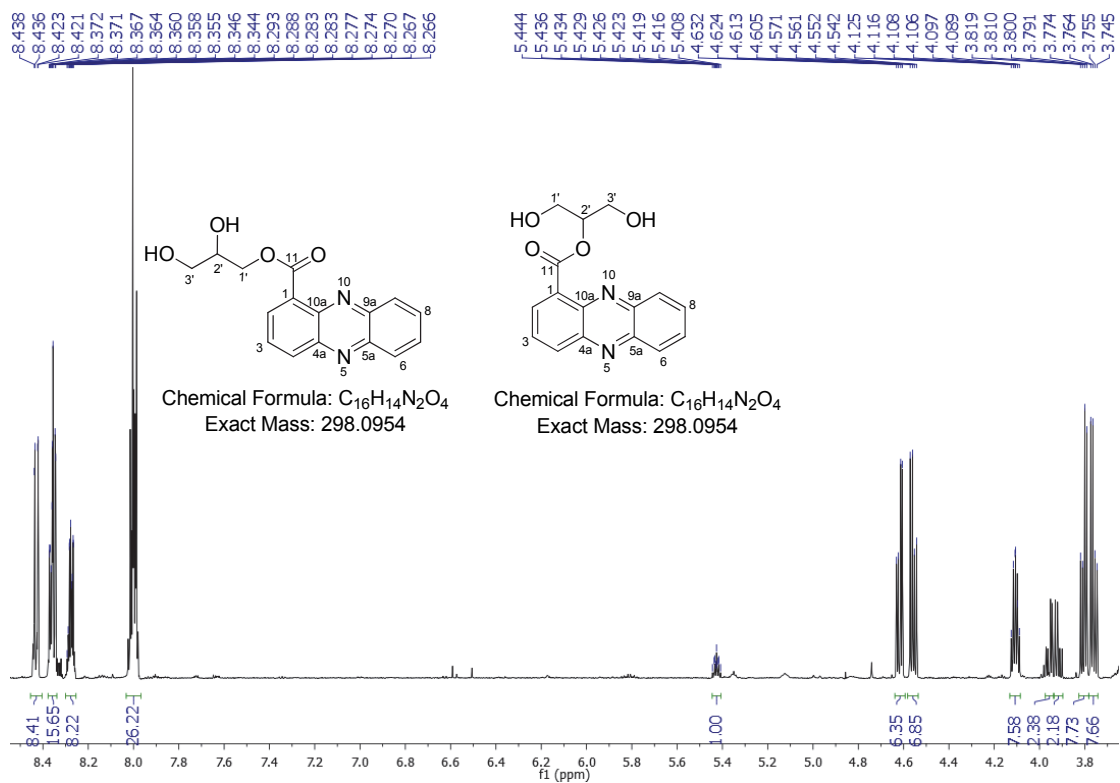
S9. UV spectrum of compound **8**S10. 1H NMR spectrum of compound **16** in CD_3OD .

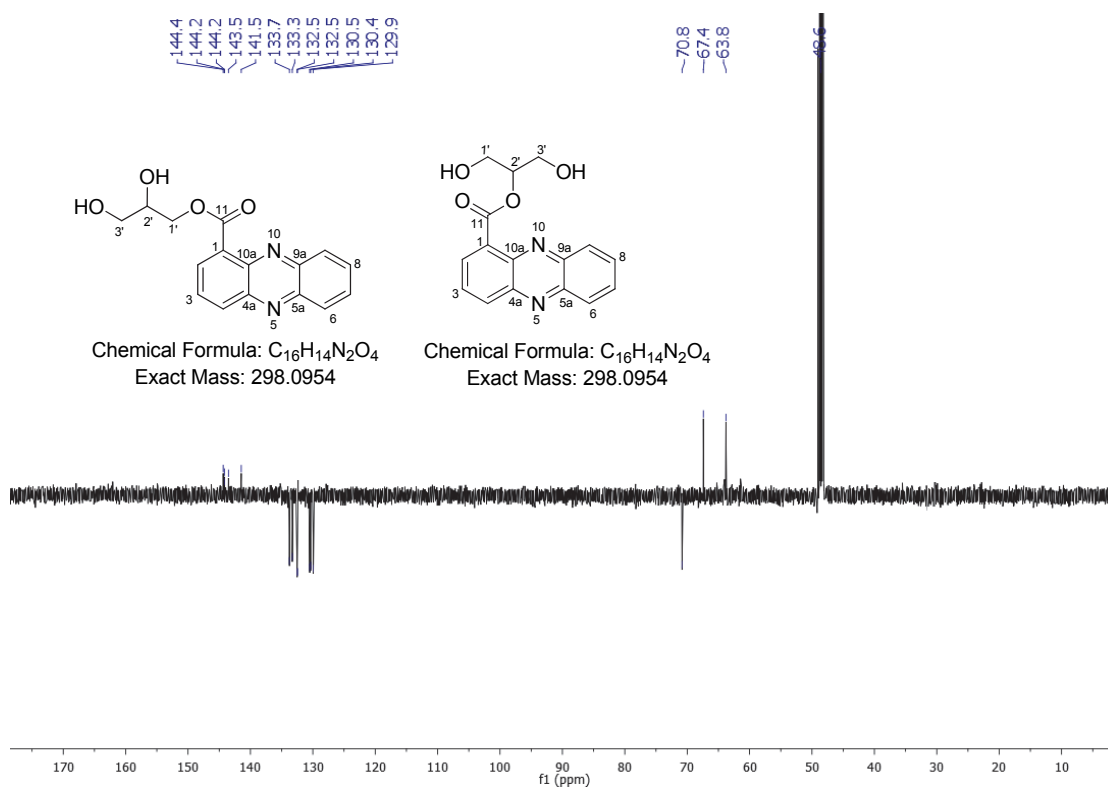
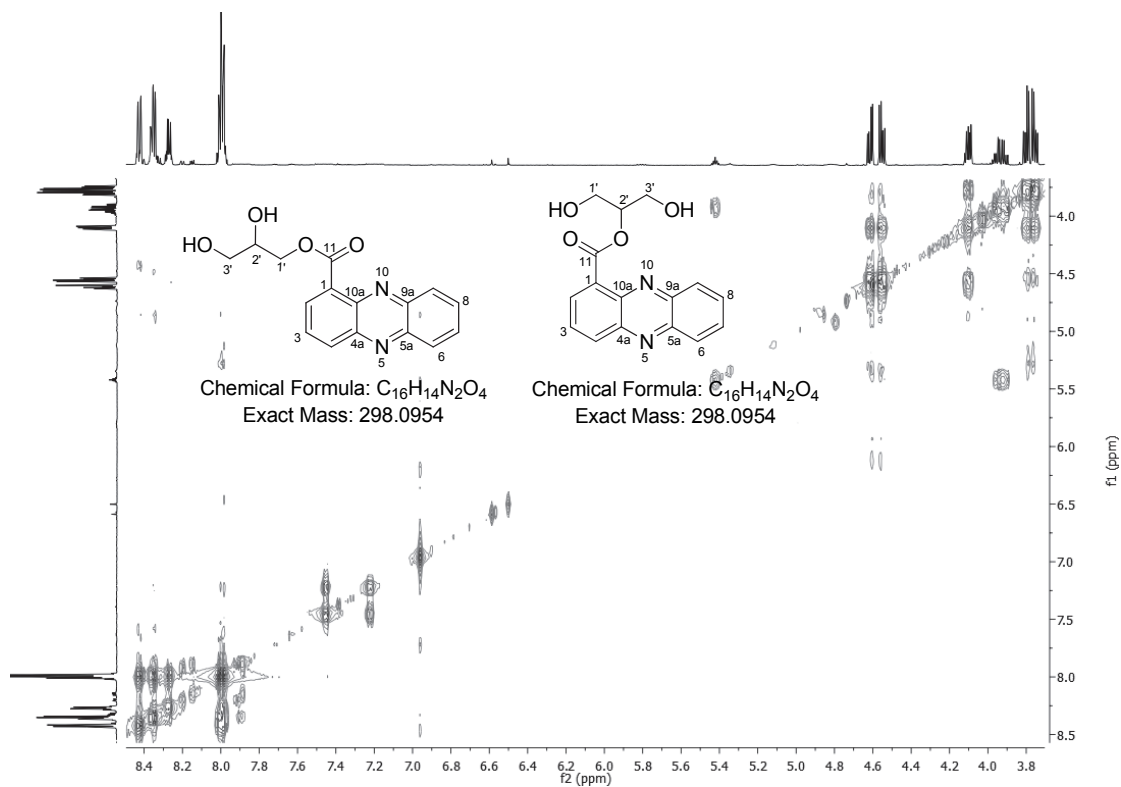
S11. HRESIMS of compound 16.

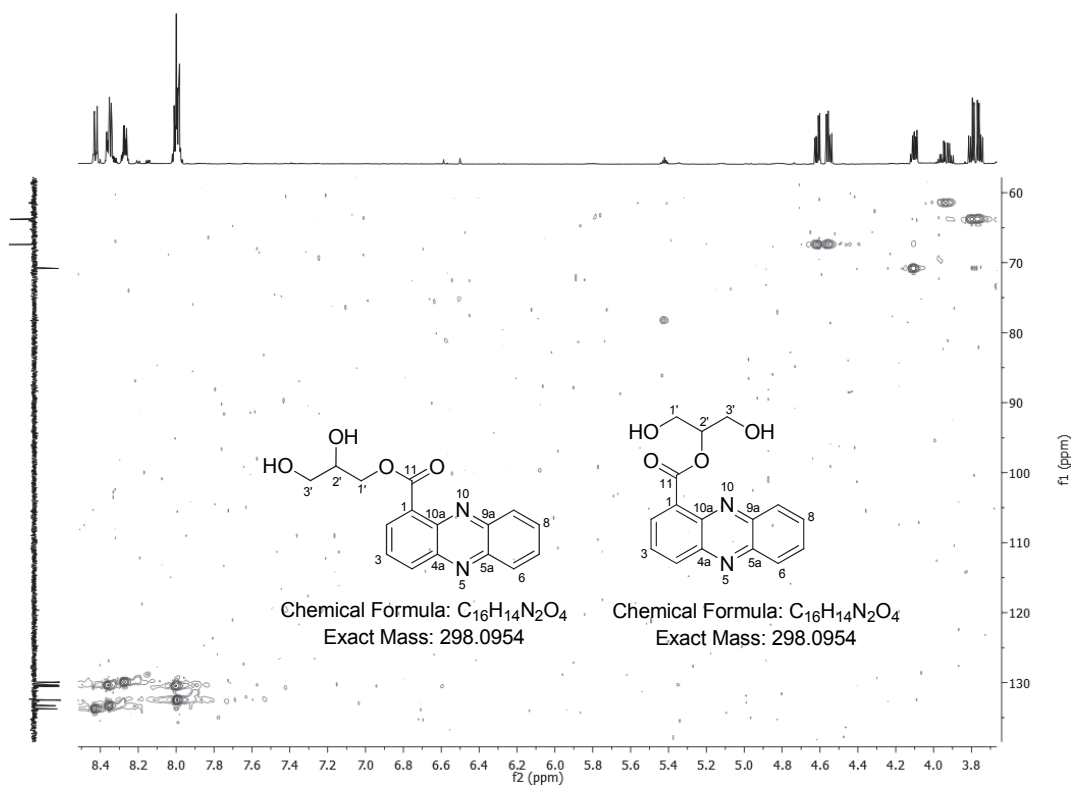
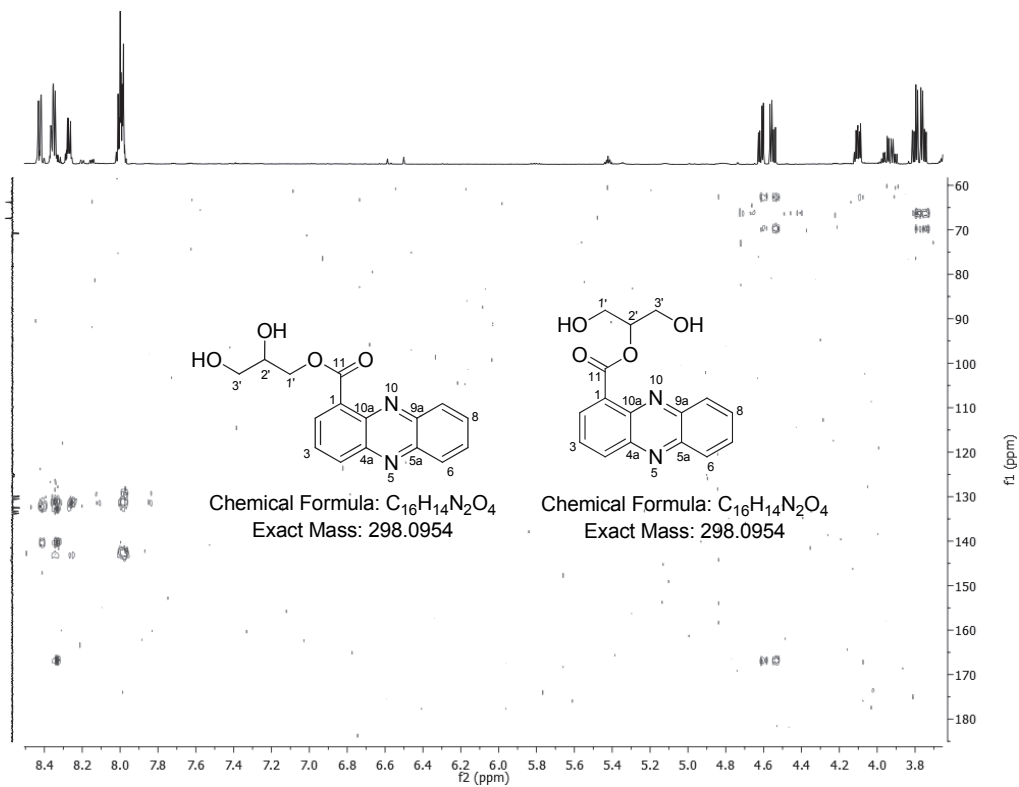


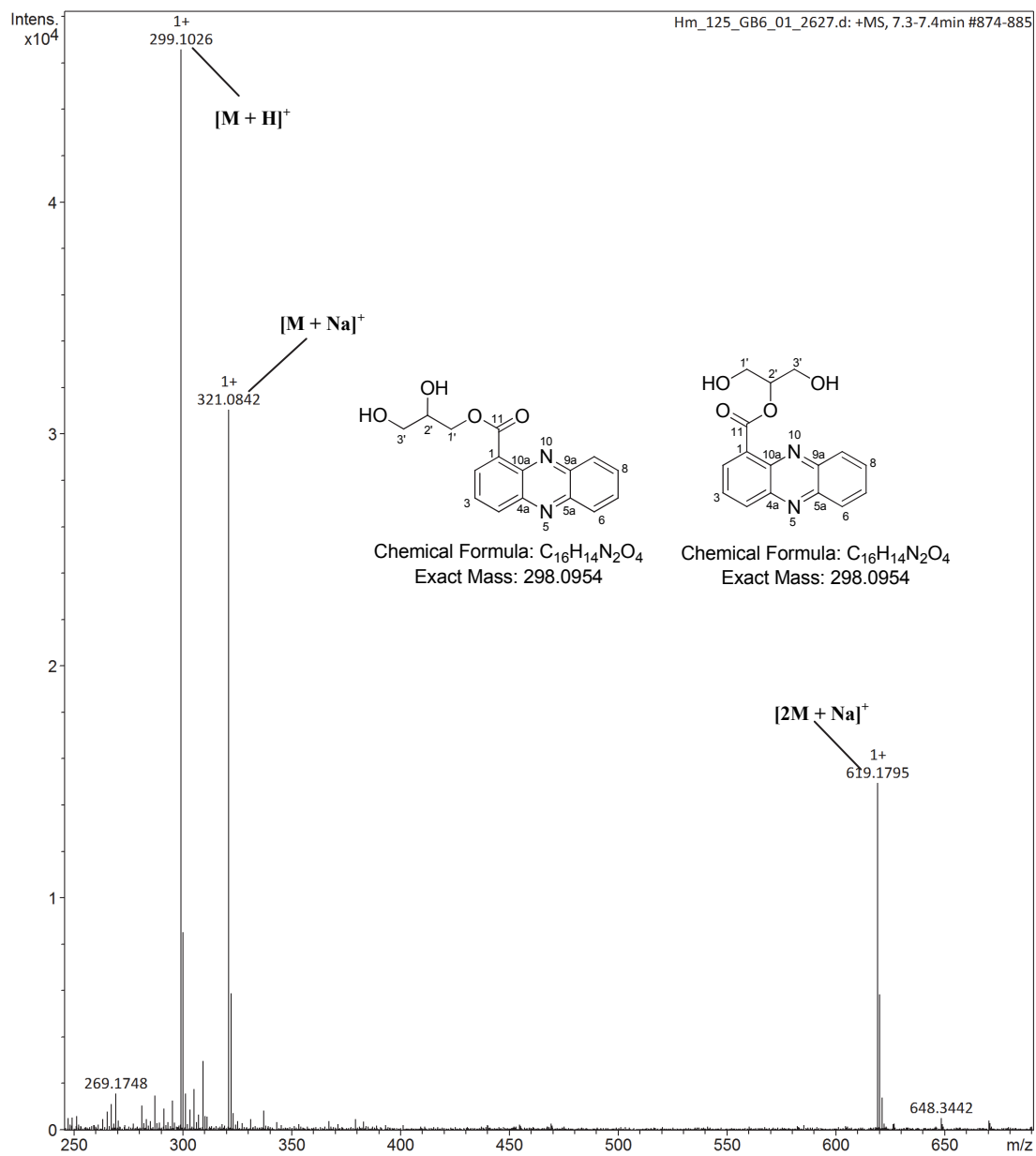
S12. UV spectrum of compound **16**.S13. 1H NMR spectrum of compound **22** in CD_3OD .

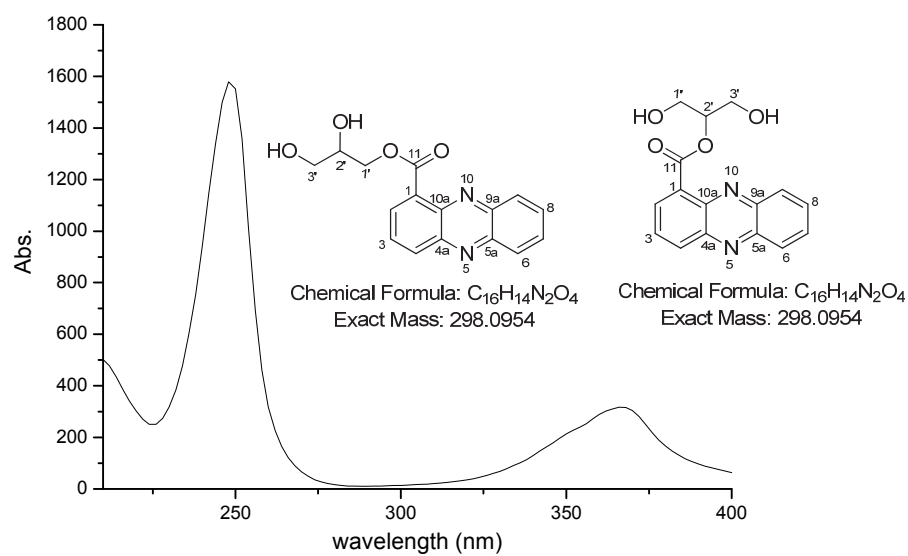
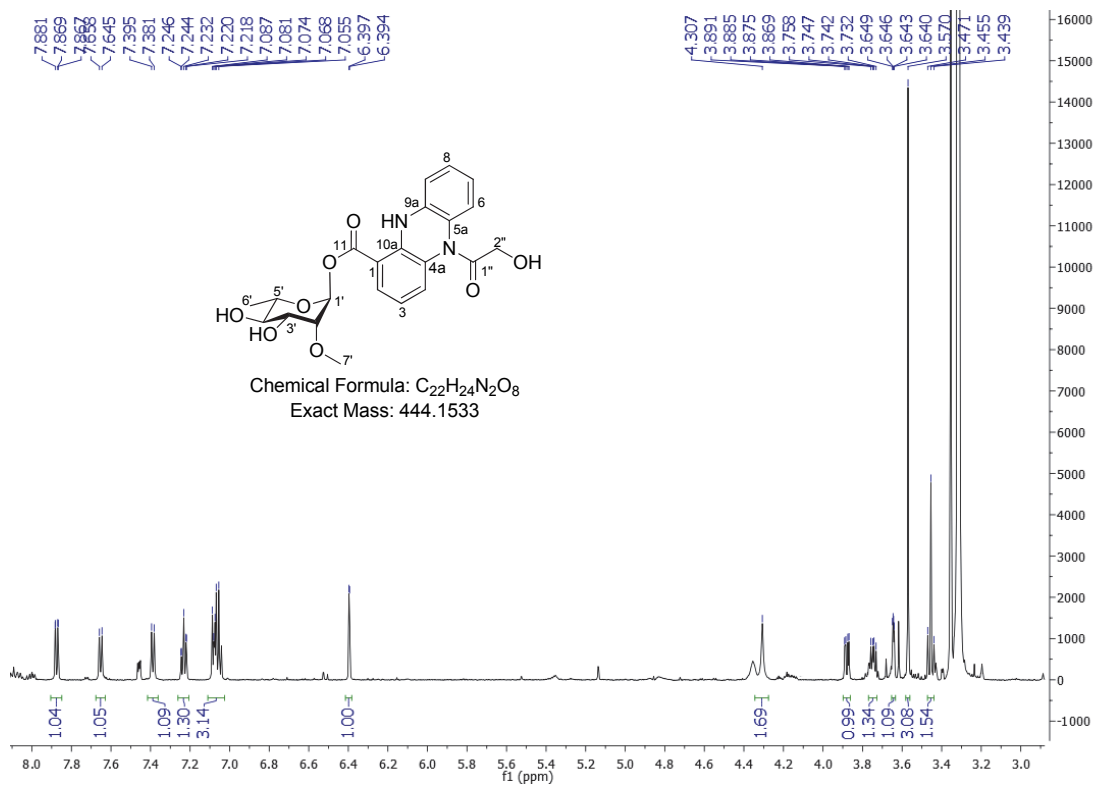
S14. HRESIMS of compound **22**.

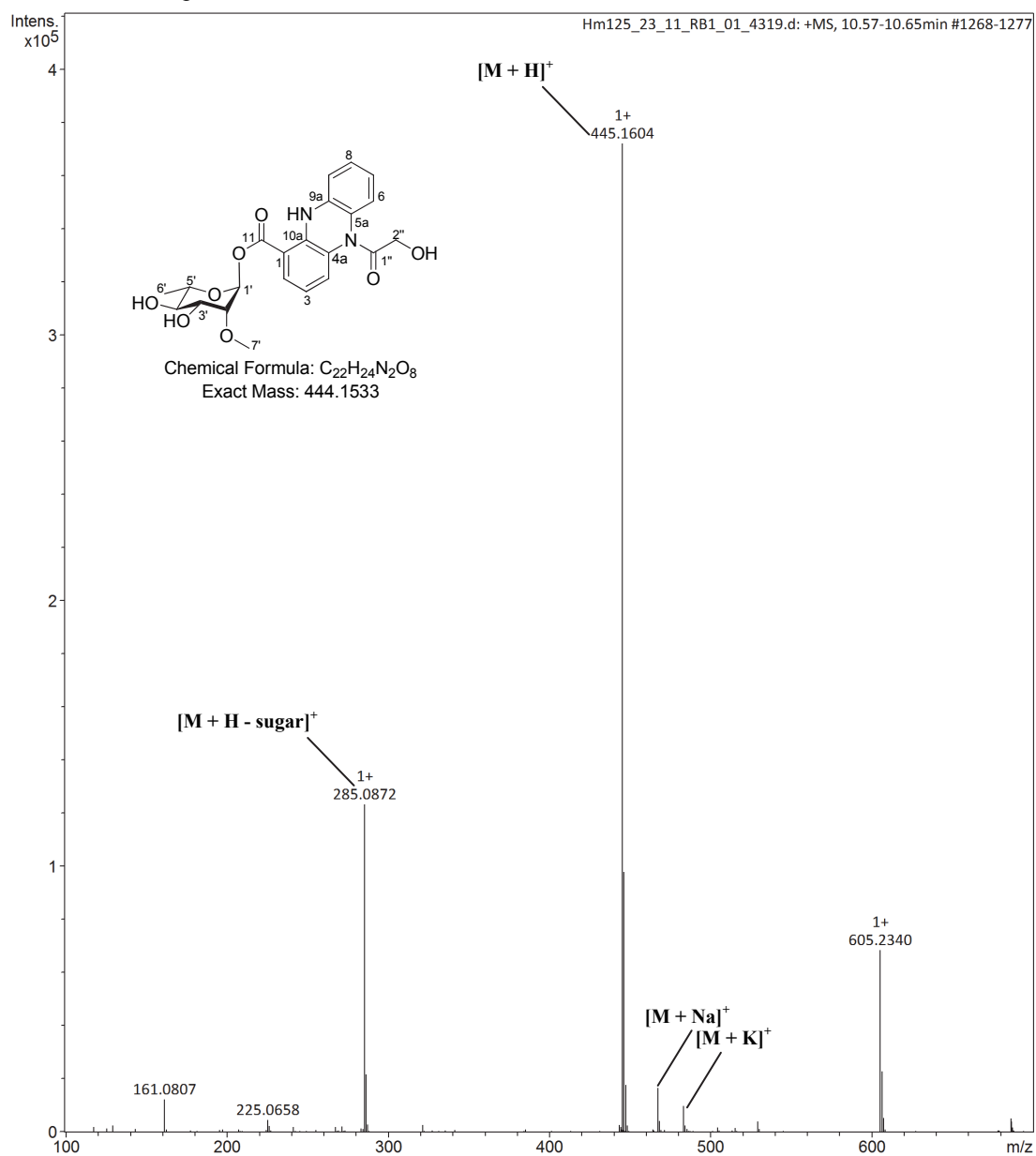
S15. UV spectrum of compound **22**.S16. 1H NMR spectrum of compounds **23** and **24** in CD_3OD .

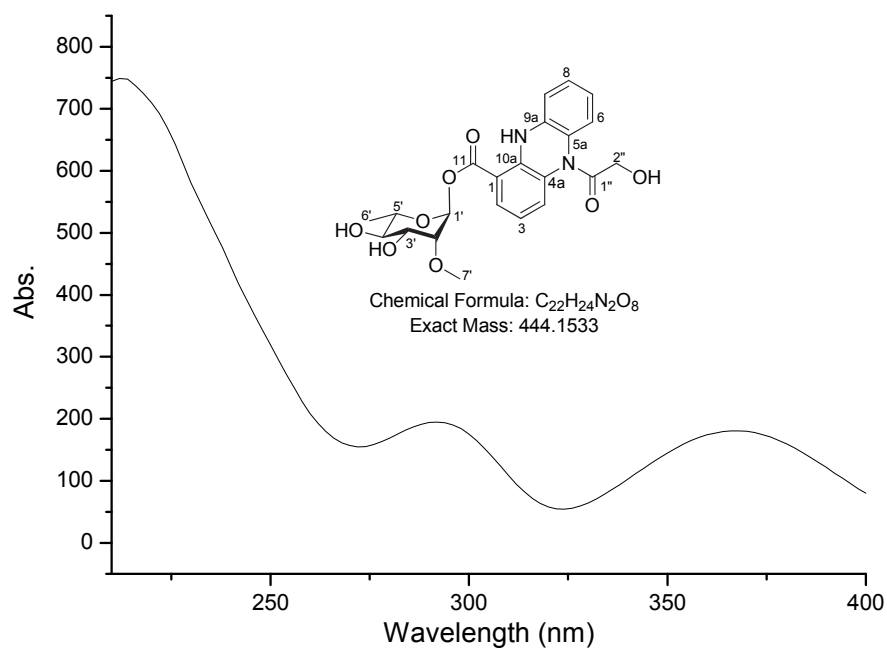
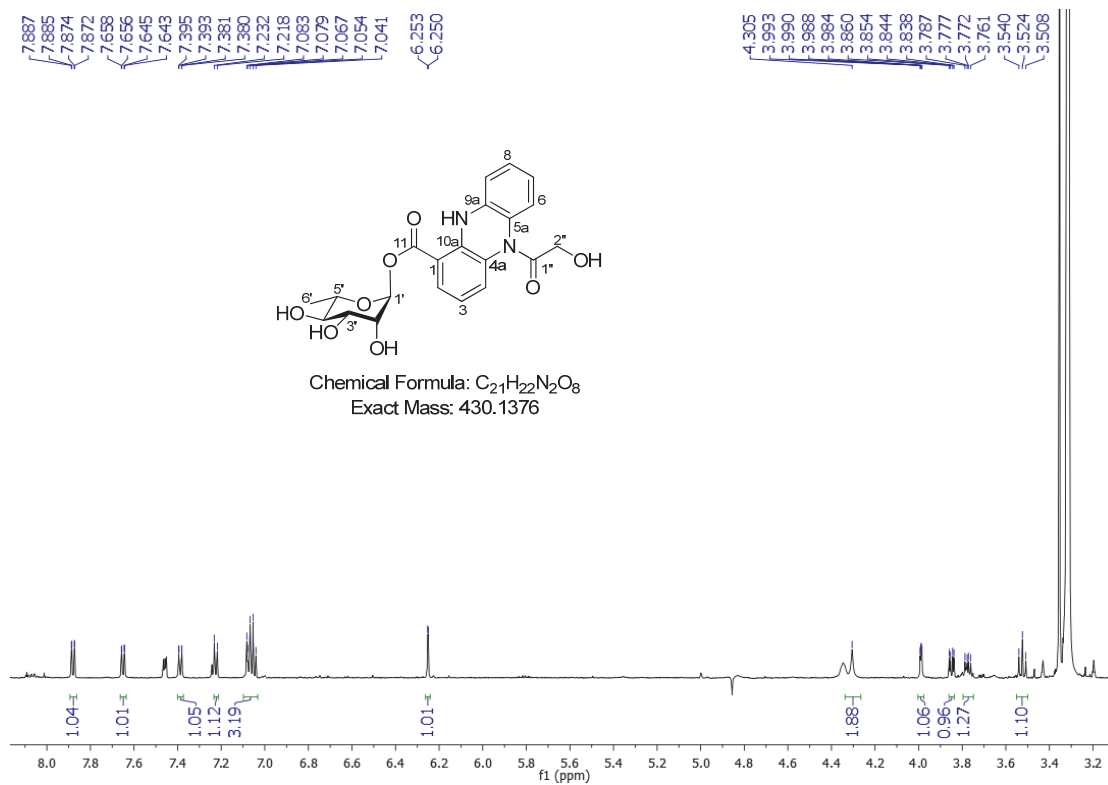
S17. APT spectrum of compounds **23** and **24** in CD₃OD.S18. ¹H-¹H COSY spectrum of compounds **23** and **24** in CD₃OD.

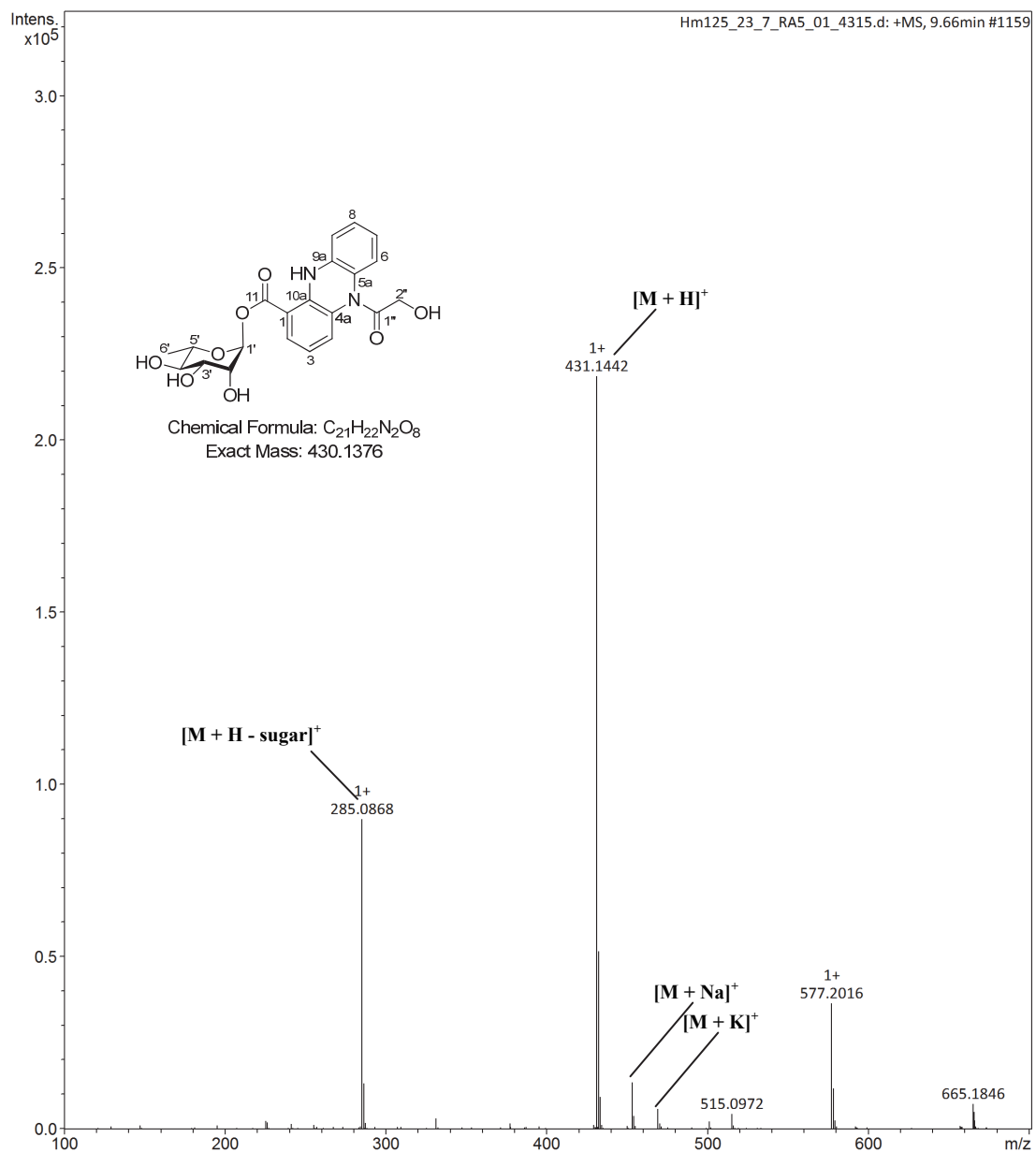
S19. HSQC spectrum of compounds **23** and **24** in CD₃OD.S20. HMBC spectrum of compounds **23** and **24** in CD₃OD.

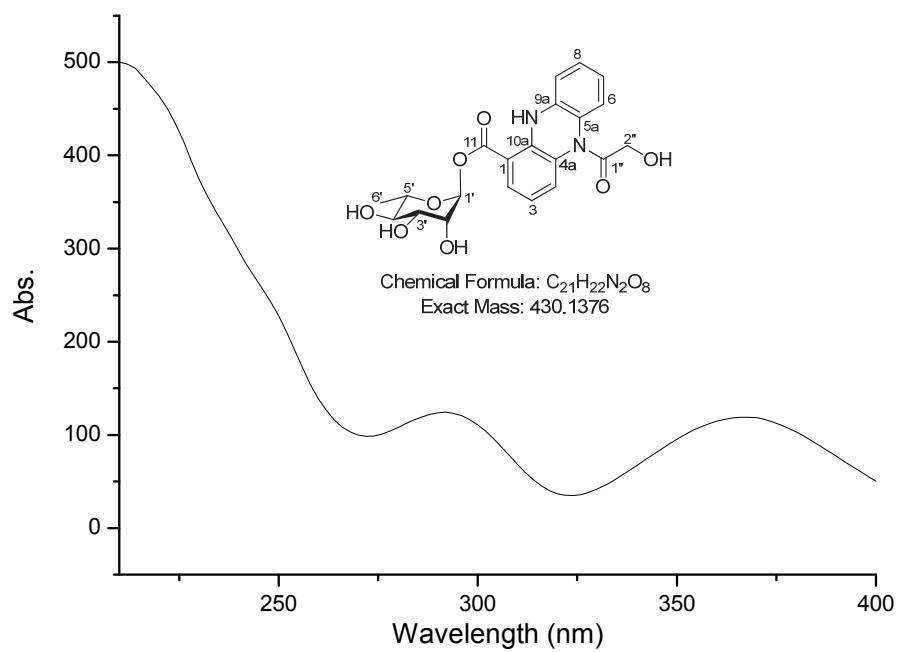
S21. HRESIMS of compounds **23** and **24**.

S22. UV spectrum of compounds **23** and **24**.S23. 1H NMR spectrum of compound **25** in CD_3OD .

S24. HRESIMS of compound **25**.

S25. UV spectrum of compound **25**.S26. 1H NMR spectrum of compound **26** in CD_3OD .

S27. HRESIMS of compound **26**

S28. UV spectrum of compound **26**.

Appendix VIII: Supplementary Information belonging to Chapter 11

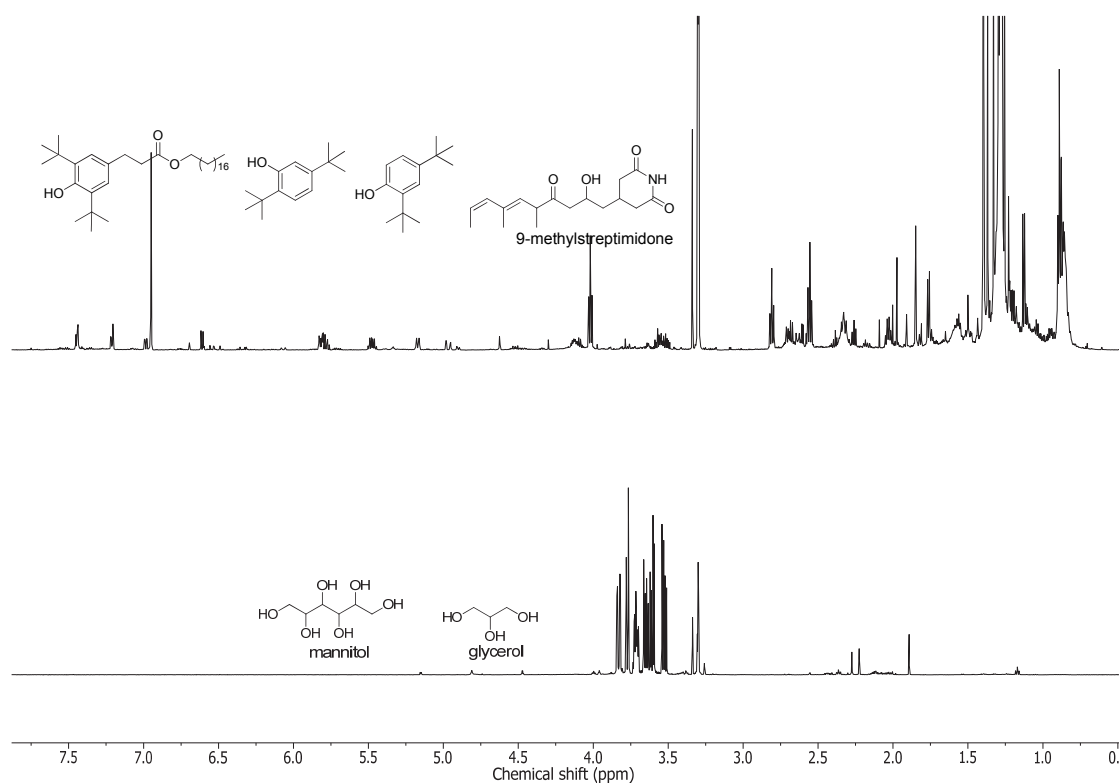


Figure S1. NMR comparison of cellulose (down) and EtOAc (up) extraction of culture broth *Streptomyces* sp. MBT73. EtOAc extracted apolar *di*-tert-butylphenols in culture medium and secondary metabolite 9-methylstreptimidone, while cellulose exclusively bonded a considerable amount of supplemented mannitol and glycerol.

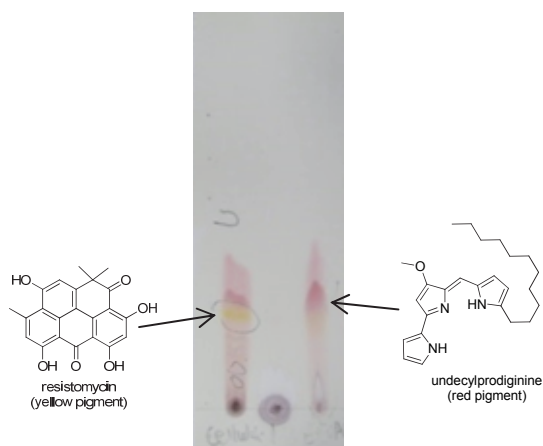


Figure S2. TLC comparison of cellulose (left) and EtOAc (right) extraction of culture broth *Streptomyces* sp. MBT73. Cellulose selectively extracted more yellow pigment (resistomycin) than the red pigment (undecylprodiginine).

PUBLICATION LIST

- (1) **Wu, C.**; van der Heul, H.; Choi, Y. H.; van Wezel, G. P. Lugdunomycin, an antibiotic with unprecedented chemical scaffold. **2016**. *In preparation*
- (2) **Wu, C.**; Du, C.; Ichinose, K.; Choi, Y. H.; van Wezel, G. P. The cryptic *qin* gene cluster of *Streptomyces* sp. MBT76 specifies C-glycosylpyranonaphthoquinones with new chemistry. *J. Nat. Prod.* **2016**, *Submitted*.
- (3) **Wu, C.**; Choi, Y. H.; van Wezel, G. P. Metabolic profiling as a tool for prioritizing antimicrobial compounds. *J. Ind. Microbiol. Biotechnol.* **2016**, *43*, 299–312.
- (4) **Wu, C.**; Du, C.; Gubbens, J.; Choi, Y. H.; van Wezel, G. P. Metabolomics-Driven Discovery of a Prenylated Isatin Antibiotic Produced by *Streptomyces* Species MBT28. *J. Nat. Prod.* **2015**, *78*, 2355–2363.
- (5) **Wu, C.**; Kim, H. K.; van Wezel, G. P.; Choi, Y. H. Metabolomics in the natural products field – a gateway to novel antibiotics. *Drug Discov. Today Technol.* **2015**, *13*, 11–17.
- (6) **Wu, C.**; Medema, M. H.; Läkamp, R. M.; Zhang, L.; Dorrestein, P. C.; Choi, Y. H.; van Wezel, G. P. Leucanicidin and Endophenazines Result from Methyl-Rhamnosylation by the Same Tailoring Enzymes in *Kitasatospora* sp. MBT66. *ACS Chem. Biol.* **2016**, *11*, 478–490.
- (7) **Wu, C.**; van Wezel, G. P.; Choi, Y. H. Identification of novel endophenazine antibiotics produced by *Kitasatospora* sp. MBT66. *J. Antibiot. (Tokyo)*. **2015**, *68*, 445–452.
- (8) **Wu, C.**; Zacchetti, B.; Ram, A. F. J.; van Wezel, G. P.; Claessen, D.; Choi, Y. H. Expanding the chemical space for natural products by *Aspergillus-Streptomyces* co-cultivation and biotransformation. *Sci. Rep.* **2015**, *5*, 10868.
- (9) **Wu, C.**; Zhu, H.; van Wezel, G. P.; Choi, Y. H. Metabolomics-guided analysis of isocoumarin production by *Streptomyces* species MBT76 and biotransformation of flavonoids and phenylpropanoids. *Metabolomics* **2016**, *12*:90.
- (10) **Wu, C.**; Lin, Z.; Wang, L.; Guo, D. Phenolic Compounds with NF- κ B Inhibitory Effects from the Fungus *Phellinus baumii*. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 3261–3267.

- (11) **Wu, C.**; Choi, Y. H.; van Wezel, G. P. Novel antibiotics, methods of use and preparation. *Patent application PCT/NL2016/050398*.
- (12) Yuan, C.; Li, G.; **Wu, C.**; Wang, H.; Zhao, Z.; Lou, H. A New Fatty Acid from the Endolichenic Fungus *Massarina* sp. *Chem. Nat. Compd.* **2015**, *51*, 415–417.
- (13) Yuan, C.; Wang, H.; **Wu, C.**; Jiao, Y.; Li, M.; Wang, Y.; Wang, S.; Zhao, Z.; Lou, H. Austdiol, fulvic acid and citromycetin derivatives from an endolichenic fungus, *Myxotrichum* sp. *Phytochem. Lett.* **2013**, *6*, 662–666.
- (14) Zheng, L.; Wang, M.; Ibarra-Estrada, E.; **Wu, C.**; Wilson, E.; Verpoorte, R.; Klinkhamer, P.; Choi, Y. H. Investigation of Chemomarkers of *Astragali Radix* of Different Ages and Geographical Origin by NMR Profiling. *Molecules* **2015**, *20*, 3389–3405.
- (15) Zhu, H.; Swierstra, J.; **Wu, C.**; Girard, G.; Choi, Y. H.; Van Wamel, W.; Sandiford, S. K.; van Wezel, G. P. Eliciting antibiotics active against the ESKAPE pathogens in a collection of actinomycetes isolated from mountain soils. *Microbiol. (United Kingdom)* **2014**, *160*, 1714–1726.
- (16) Wang, L.; Guo, D.; Wang, S.; **Wu, C.**; Rehman, M. U.; Lou, H. Phenolic Glycosides from the Chinese Liverwort *Reboulia hemisphaerica*. *Helv. Chim. Acta* **2011**, *94*, 1146–1152.

CURRICULUM VITAE

Changsheng Wu was born on March 13th, 1985, in Nancheng, Jiangxi province, China. After finishing high school in July 2004, he entered Shandong University (Jinan, China) to study pharmacy, where he three times won the “Annual Merit Scholarship”. He got trained in multiple disciplines involving both chemistry and biology, and obtained his bachelor degree in July of 2008. He accepted a recommendation to be exempted from the admission exam to his Master course in the Department of Natural Product Chemistry, School of Pharmaceutical Sciences, Shandong University. In July 2011 he obtained his Master’s degree. His research project on the phytochemical investigation of the fungus *Phellinus baumii* was awarded an ‘Excellent Graduate’ award in the department. In September 2011, he received a scholarship from the China Scholarship Council (CSC) to carry out a PhD study at the Institute of Biology at Leiden University. The project was carried out under the supervision of Prof. dr. Gilles P. van Wezel (Molecular Biotechnology) and Dr. Young Hae Choi (Natural Product Laboratory), with research theme as the discovery of novel antibiotics from actinomycetes. Currently, he is continuing his work in the laboratory of Prof. van Wezel as a post-doctoral researcher.

