

Exploration of bioresources enables detection and characterization of natural products with antibacterial activities and elucidation of the corresponding biosyntheses for derivative generation

Dissertation

Zur Erlangung des Doktorgrades (Dr. rer. Nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

Vorgelegt von

Nils Böhringer

aus

Troisdorf

Bonn, 2020

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms Universität Bonn

Erstgutachterin: Prof. Dr. Gabriele M. König

Zweitgutachter: Prof. Dr. Till F. Schäberle

Tag der Promotion: 04.10.2021

Erscheinungsjahr: 2021

Vorabpublikationen aus dieser Arbeit

Folgende Auszüge aus dieser Arbeit wurden mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät, repräsentiert durch die Erstgutachterin dieser Arbeit, vorab publiziert:

Böhringer, N., Patras, M. A., & Schäberle, T. F. (2021). Heterologous Expression of Pseudouridimycin and Description of the Corresponding Minimal Biosynthetic Gene Cluster. *Molecules (Basel, Switzerland)*, 26(2), 510. <https://doi.org/10.3390/molecules26020510>

Imai, Y., Meyer, K. J., Iinishi, A., Favre-Godal, Q., Green, R., Manuse, S., Caboni, M., Mori, M., Niles, S., Ghiglieri, M., Honrao, C., Ma, X., Guo, J. J., Makriyannis, A., Linares-Otoya, L., **Böhringer, N.**, Wuisan, Z. G., Kaur, H., Wu, R., Mateus, A., Typas, A., Savitski M.M., Espinoza, J.L., O'Rourke, A., Nelson, K.E., Hiller, S., Noinaj, N., Schäberle, T.F., D'Onofrio, A., Lewis, K. (2019). **A new antibiotic selectively kills Gram-negative pathogens.** *Nature*, 576(7787), 459–464. <https://doi.org/10.1038/s41586-019-1791-1>

Böhringer, N., Fisch, K. M., Schillo, D., Bara, R., Hertzner, C., Grein, F., Eisenbarth, J. H., Kaligis, F., Schneider, T., Wägele, H., König, G. M., & Schäberle, T. F. (2017). Antimicrobial Potential of Bacteria Associated with Marine Sea Slugs from North Sulawesi, Indonesia. *Frontiers in microbiology*, 8, 1092. <https://doi.org/10.3389/fmicb.2017.01092>

Zusammenfassung

Die Entdeckung und Markteinführung von Antibiotika zu Anfang des 20. Jahrhundert war der Beginn einer medizinischen Revolution. Zum ersten Mal wurde die Menschheit in die Lage versetzt bakterielle Infektionskrankheiten effektiv zu bekämpfen. Trotz des immensen Erfolgs ging die Entdeckung und Entwicklung neuer Antibiotika stark zurück während sich gleichzeitig Resistenzen gegen alle klinisch relevanten Antibiotika entwickelten und verbreiteten. Dies muss mit großer Sorge betrachtet werden, da dadurch die erfolgreiche Behandlung von bakteriellen Infektionen stark erschwert ist. Insbesondere Gram- Bakterien, die Resistenzen gegen klinisch relevante Antibiotika entwickelt haben, stellen aktuell die größte bakterielle Bedrohung für die Menschheit dar. Daher hat die der Entdeckung und Entwicklung neuer effektiver Antibiotika gegen diese Erreger eine immense Bedeutung. In dieser Arbeit werden drei Einzelprojekte beschrieben die einen Beitrag zur Entdeckung und Erforschung neuer antibakterieller Naturstoffe darstellen:

Durch die scheinbare Erschöpfung von terrestrischen Bioressourcen, insbesondere aktinobakterieller Umweltisolate, rückten andere Bioressourcen, wie zum Beispiel marine Habitats, in den Fokus. Kapitel I dieser Arbeit beschreibt die Erforschung von *Nudibranchia*- (marine Nacktschnecken)-assoziierten Bakterien aus dem Indopazifik um die Insel Bunaken, Indonesien im Hinblick auf ihre Fähigkeit zur Biosynthese von antibakteriellen Naturstoffen. Hierbei konnten 49 Bakterienstämme isoliert und ihr Potenzial zur Biosynthese antibiotischer Naturstoffe untersucht werden.

In Kapitel II wurde die Biosynthese des Streptomycetalen antibakteriellen Naturstoffs Pseudouridimycin eingehend untersucht. Basierend auf einer vorausgegangenen Arbeit von Sosio *et al.*, wurde dazu das Biosynthesegencluster rearrangiert und unter der Kontrolle von starken Promotoren heterolog in *Streptomyces coelicolor* M1146 exprimiert. Insgesamt wurden drei Versionen des Genclusters kloniert, was die Identifizierung der minimalen genetischen Ausstattung zur Biosynthese von PUM ermöglichte.

Kapitel III beschreibt die Erforschung der Biosynthese der neu entdeckten antibakteriellen Substanz Darobactin. Durch seinen neuartigen Wirkmechanismus, bei dem mit BamA ein Protein in der äußeren Membran gehemmt wird, greift Darobactin ausschließlich Gram- Bakterien an, was insbesondere im Hinblick auf die zuvor genannte Bedrohung durch resistente Gram- Pathogene interessant ist. Hierbei wurde das zu Grunde liegende Biosynthesegencluster identifiziert und dessen Beteiligung an der Darobactin Biosynthese bestätigt. Dieser Einblick in die Darobactin Biosynthese ermöglichte zusätzlich die Identifikation und initiale Charakterisierung von weiteren Darobactin Derivaten.

Acknowledgements

I wish to express my sincere gratitude to Prof. Dr. Gabriele M. König for giving me the great opportunity to conduct my Doctoral studies in her workgroup at the Institute for Pharmaceutical Biology. Her kind support and her encouragement were of great help during my work.

I am especially grateful to Prof. Dr. Till F. Schäberle, my supervisor, who supported me with his expert guidance and his knowledge. I have to thank him for the great opportunity to work on such interesting and rewarding projects. His contribution to this work therefore deserves particular appreciation.

I would like to thank Prof. Dr. Gerd Bendas for his kind participation in my examination committee and Prof. Dr. Heike Wägele, also for her kind participation in my examination committee and for being the best diving instructor one can hope for.

My sincere gratitude goes to Prof. Dr. Kim Lewis for his kind invitation to contribute to a most fascinating and ultimately most rewarding project.

My gratitude also goes to Dr. Max Crüsemann, Dr. Maria Patras, Dr. Stefan Kehraus and Christoph Hartwig for thoroughly training me in analytical chemistry as well as always having an open door for questions and discussion.

A special thanks goes to the members of the AG Schäberle “Natural Product Research wFI”, the members of the Fraunhofer Bioresources group and the members Prof. Dr. Königs workgroup of the Institute of Pharmaceutical Biology for the great, friendly and supportive atmosphere. Regardless of whether it concerned scientific issues or social events, it was always a pleasure to be with you!

Finally, I want to thank my friends and my family for their support and unwavering trust, which helped me to pursue my goals and to maintain focus during my studies. Above all, my gratitude goes to my parents for the unconditional and unlimited love and support I have received. This work is dedicated to you!

TABLE OF CONTENT

I.	INTRODUCTION	11
I.1.	Antibiotic Natural Products	15
I.1.1.	Non-ribosomal peptides (NRPs).....	15
I.1.2.	Polyketides (PKs)	16
I.1.3.	Aminoglycosides.....	17
I.2.	Antimicrobial activity	17
I.2.1.	Cell wall.....	18
I.2.2.	Outer membrane.....	19
I.2.3.	Intracellular targets.....	20
I.3.	Natural product discovery – Past, present and future?.....	20
I.4.	Scope of this thesis.....	24
II.	ANTIMICROBIAL POTENTIAL OF BACTERIA ASSOCIATED WITH MARINE SEA SLUGS FROM NORTH SULAWESI, INDONESIA.....	26
II.1.	INTRODUCTION	26
II.2.	MATERIAL AND METHODS	27
II.2.1.	Environmental Samples	27
II.2.2.	Identification of Sea Slugs and Sponges	28
II.2.3.	Isolation of Bacteria	28
II.2.4.	Generation of Extracts and Activity Screening.....	29
II.2.5.	Detection of PKS and NRPS Biosynthetic Genes	29
II.2.6.	Phylogenetic Analysis.....	30
II.3.	RESULTS	31
II.3.1.	Isolation and Identification of Sea Slug Associated Bacteria.....	31
II.3.2.	Antimicrobial Activity Assays.....	32
II.3.3.	Phylogenetic Analysis.....	33
II.3.4.	Detection of PKS and NRPS Biosynthetic Genes	34
II.3.5.	DISCUSSION	35
III.	HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS	38
III.1.	Introduction.....	38
III.2.	Material and Methods	43
III.2.1.	Bacterial strains used in this study	43
III.2.2.	Media and additives used in this study	45

III.2.3.	Cultivation of bacteria.....	47
III.2.4.	Primers used in this study	48
III.2.5.	Molecular biology techniques.....	49
III.2.6.	Construction of pGEM-teasy_Apra-ermE* and pGEM-teasy_Apra-tcp830.....	50
III.2.7.	Construction of the integrative <i>Streptomyces</i> PUM expression plasmids pCAP03-PUM Δ FAH_ermE*/tcp830, pCAP03-PUM Δ F_ermE*/tcp830 and pCAP03-PUM_ermE*/tcp830 51	51
III.2.8.	Conjugation of PUM expression plasmids <i>S. coelicolor</i> M1146, <i>S. lividans</i> TK24 and <i>S. albidoflavus</i> J1074	52
III.2.9.	Enrichment of PUM from <i>Streptomyces</i> sp. DSM26212 culture broth	52
III.2.10.	Heterologous expression of the PUM BGC.....	53
III.2.11.	LC/MS analysis of PUM	53
III.2.12.	Matrix assisted laser desorption/ionisation mass spectrometry (MALDI-MS).....	53
III.3.	Results.....	55
III.3.1.	Construction of pGEM-teasy_Apra-ermE* and pGEM-teasy_Apra-tcp830.....	55
III.3.2.	Cloning of the PUM BGC.....	55
III.3.3.	PUM production in <i>Streptomyces</i> sp. DSM26212	57
III.3.4.	Heterologous expression of the PUM Δ FAH BGC in <i>S. coelicolor</i> M1146 on agar plate using MALDI-MS	58
III.3.5.	Heterologous expression of the PUM BGCs in liquid media	59
III.4.	Discussion	61
III.4.1.	Cloning the PUM BGC.....	61
III.4.2.	Analysis of PUM production in <i>Streptomyces</i> sp. DSM26212	61
III.4.3.	Heterologous expression of three streamlined versions of the PUM BGC	63
IV.	DAROBACTIN	65
IV.1.	Introduction.....	65
IV.1.1.	Ribosomally synthesised post-translationally modified peptides (RiPPs)	68
IV.1.2.	Scope of this study	70
IV.2.	Material and Methods	71
IV.2.1.	Bacterial strains used in this study	71
IV.2.2.	Media and additives used in this study	73
IV.2.3.	Primers used in this study	75
IV.2.4.	Cultivation of bacteria.....	77
IV.2.5.	Bioinformatic analysis of Darobactin BGCs	77
IV.2.6.	General molecular biology techniques.....	78

IV.2.7.	LCMS analysis of Darobactin and Darobactin analogue containing samples.....	79
IV.2.8.	Interspecific conjugation between <i>E. coli</i> WM3064 and <i>Photorhabdus khanii</i> DSM3369.....	80
IV.2.9.	Genetic inactivation of DarE in <i>Photorhabdus khanii</i> DSM3369	80
IV.2.10.	Markerless deletion of <i>darABCDE</i> in <i>Photorhabdus khanii</i> DSM3369	81
IV.2.11.	Construction of darA-E expression plasmids pNB03 – darA-E and pNBDaro.....	82
IV.2.12.	Homologous and heterologous overexpression of the Darobactin BGC in <i>P. khanii</i> DSM3369 Δ <i>darABCDE</i> and <i>E. coli</i> BW25113	83
IV.2.13.	Identification of the minimal Darobactin BGC	83
IV.2.14.	Mutasynthetic production of Darobactin analogues B,C,D,E,F by modification of WT Darobactin A BGC.....	84
IV.2.15.	MS/MS based characterisation of Darobactin analogues	85
IV.2.16.	Isolation of Darobactin and its derivatives	86
IV.2.17.	Detection of natural Darobactin analogues Marinodarobactin A and B and Didehydrodarobactin A from <i>Pseudoalteromonas luteoviolacea</i> H33 and H33s	86
IV.2.18.	Determination of MIC values of Darobactin A and Darobactin analogues	87
IV.3.	Results.....	88
IV.3.1.	Identification of the Darobactin BGC.....	88
IV.3.2.	Genetic inactivation of darE and knock out of darABCDE in <i>Photorhabdus khanii</i> DSM3369.....	92
IV.3.3.	Homologous and heterologous overexpression of the Darobactin A BGC in <i>Photorhabdus khanii</i> DSM3369 Δ <i>darABCDE</i> and <i>E. coli</i> BW25113.....	94
IV.3.4.	Identification of the Darobactin minimal BGC	95
IV.3.5.	Mutasynthetic production of Darobactin analogues B-F.....	97
IV.3.6.	Discovery of Dehydrodarobactin A and Marinodarobactin A and B from <i>Pseudoalteromonas luteoviolaceae</i> H33 and H33S	103
IV.3.7.	MIC values of Darobactin analogues	109
IV.4.	Discussion	110
IV.4.1.	Identification and corroboration of the Darobactin minimal BGC.....	110
IV.4.2.	Mutasynthetic production of Darobactin analogues	114
IV.4.3.	Identification of natural Darobactin A analogues from <i>Pseudoalteromonas luteoviolacea</i> H33 and H33S	118
V.	SUMMARY AND OUTLOOK	122
VI.	REFERENCES.....	125
VII.	SUPPLEMENTARY MATERIAL	133
VIII.	Publications	139

Abbreviations

AA	amino acid
ACP	acyl carrier protein
BGC	biosynthetic gene cluster
BLAST	basic local alignment tool
bp	basepair
CoA	coenzymeA
Da	dalton
DNA	deoxyribonucleic acid
EIC	extracted ion chromatogram
FA	formic acid
Flash	medium pressure liquid chromatography
HPLC	high performance liquid chromatography
kb	kilobasepair
KO	knock-out
LCMS	liquid chromatography with mass spectrometer detector
<i>m/z</i>	mass to charge ratio
MeCN	acetonitrile
MeOH	methanol
MS	mass spectrometry
NP	natural product
NRP	non-ribosomal peptide
NRPS	non-ribosomal peptide synthetase
nt	nucleotide(s)
OD	optical density
ORF	open reading frame
PCP	peptidyl carrier protein
PK	polyketide
PKS	polyketide synthase
PUM	pseudouridimycin
RaS	radical S-adenosylmethionine (enzyme)
rDNA	DNA coding for ribosomal RNA
RiPP	ribosomally synthesised and posttranslationally modified peptide
RNA	ribonucleic acid
rSAM	radical S-adenosylmethionine (enzyme)
sp.	species
SPE	solid phase extraction
TOF	time of flight
UPLC	ultra performance liquid chromatography
WT	wild-type

I. INTRODUCTION

The discovery of antibiotics in the early 20th century sparked a medicinal revolution, which for the first time in history enabled humankind to effectively combat bacterial infections. To this date, even opportunistic infections of small wounds could have fatal effect to the individuum. In this context, the outbreak of diseases by more aggressive bacterial pathogens could not be handled and epi- and pandemics could not be contained effectively. This is well exemplified by outbreaks of the plague caused by the Enterobacterium *Yersinia pestis* between the 14th and 19th century which led to significant decrease in world population. Estimations conclude that during the first outbreak between 1331 and 1351 alone, 30 % to 60 % of the European population perished (Benedictow, 2006). Only historically very recent advances in the field of optics, biology and medicine allowed identification of bacterial pathogens and putting into place effective countermeasures. The immense success of sulfonamides and especially the newly discovered penicillin during the 2nd world war fuelled the post war discovery and development of new antibacterial compounds, many of which bacterial or fungal metabolites i.e. microbial natural products (NPs). This discovery not only caused rapid development of the antibiotics research field but kickstarted the whole discipline of NP chemistry and the plethora of antibiotic compounds and compound classes still in use today originate from that time. This “Golden Age of Antibiotics” lasted to the 1960s and was accompanied by the illusory hope to eradicate all infectious diseases globally. This hope faded quick and the “Golden Age” was followed by a massive gap in antibiotic discovery, lasting to this day (Fig. 1). Over time, pharmaceutical companies lost “interest” in the development of new antibiotics due to high development costs paired with low revenue of the final product, hence this field of research faced massive disinvestment. In contrast to the decline in discovery and development, the use of antibiotics immediately sparked the occurrence and rapid spread of bacterial pathogens resistant to one or more antibiotics. Today, this situation raises a lot of concern among the medicinal society, not only from the perspective of the plethora of primary bacterial diseases such as the (in developed countries rather exotic) plague and especially food and livestock associated pathogens like enterohaemorrhagic *Escherichia coli* (EHEC), *Listeria monocytogenes*, *Salmonella* species and methicillin resistant *Staphylococcus aureus* (MRSA). All the mentioned food and livestock associated pathogens are of great concern for the food and feed industry and outbreaks are common also in developed countries (Forsythe, 2020). While damage on level of the infected individuum is already high, usually more than one person is affected thus amassing to a much larger macroeconomic damage. Also, opportunistic nosocomial and secondary infections become harder to treat with drugs on the market. Textbook example is MRSA, causing hard to treat wound infections, which usually are associated with surgical procedures and lesions that require hospital treatment. However, more

recent investigations put Gram- opportunistic pathogens such as *Pseudomonas aeruginosa* and *Acinetobacter baumannii* into focus, since they become increasingly harder to treat due to rapid acquisition of resistances and are harder to contain by hygienic measures due to their broad spectrum of ecological niches and environmental persistence. Currently, the most important nosocomial resistant bacterial pathogens are subsumed under the term ESKAPE-pathogens describing *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* und *Enterobacter* species.

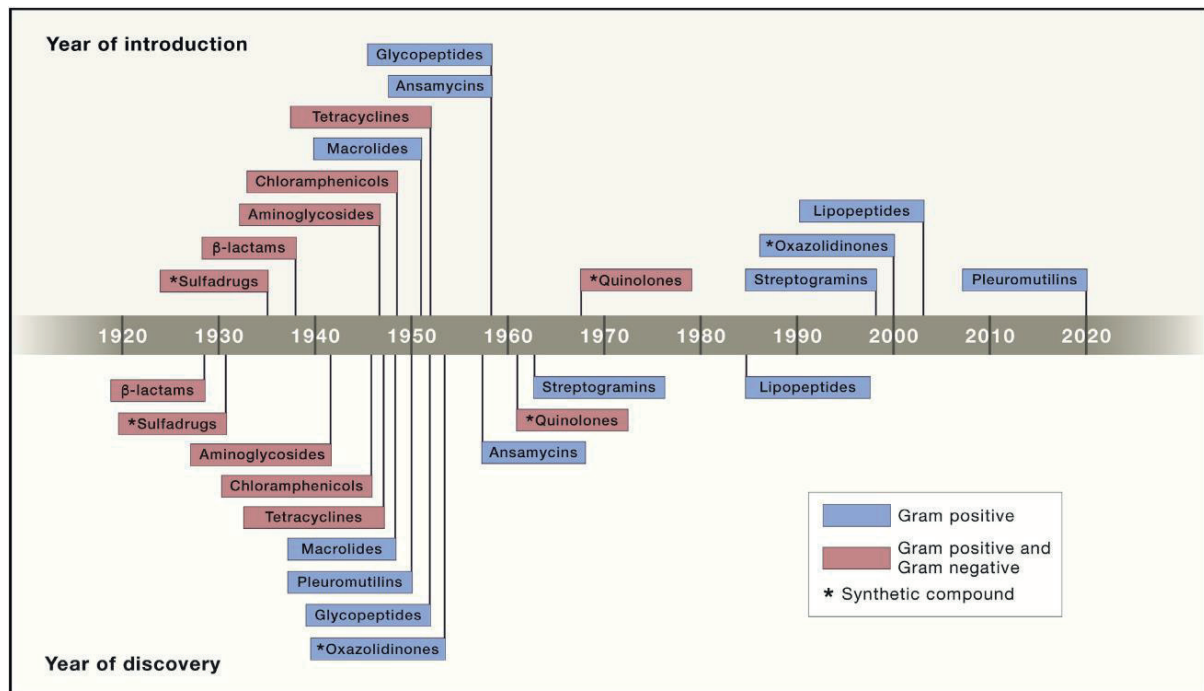


Figure 1: Overview of discovery and market introduction of antibiotic classes over the course of 100 years starting in 1920 (Lewis, 2020).

Especially in the scope of the current (late 2020) COVID-19 pandemic, antimicrobial resistance becomes a hard to ignore issue. Historic examples of viral pandemics like the Spanish Influenza or the Hong Kong Influenza in the 20th century were extensively studied and a common factor contributing to overall high mortality was the coinfection of the lung with bacterial pathogens such as *Haemophilus influenzae* and *Streptococcus pneumoniae* (Morris *et al.*, 2017). Literature suggests that during the Spanish Influenza pandemic in 1918 alone, secondary bacterial pneumonia was responsible for a majority (> 90 %) of deaths in this first wave of the pandemic (Morens *et al.*, 2008). While later examples like the Hong Kong Influenza pandemic had a significantly lower death toll than the Spanish Influenza and antibiotics were available at that time, reports clearly show that bacterial coinfection was still a key factor, determining the severity of progression and ultimately fatality of the initial viral infection (Lindsay *et al.*, 1970). Hence, lack of effective antibiotics also affects the

ability to mitigate the impact of viral epi- and pandemics like the current one. This illustrates well that antibiotics are not just a weapon to combat bacterial infections, but an invaluable tool of modern medicine *per se*. Many surgical procedures naturally rely on antibiotic treatment to not trade risk of the cause of surgery for the risk of post-surgery bacterial infection. Despite all draw- and setbacks associated with antibiotic treatment and the emerging threat of antimicrobial resistance, the history of antibiotics is an immense success story for humanity. Development of life expectancy and standard of living during the 20th century cannot be envisioned without the discovery and development of antibiotics and thus antibiotics are shaping the world we live in, to this day.

Despite this immense success of antibiotics, resistances against all clinically relevant antibiotics arose over the past decades, leaving doctors with few to no options for successful treatment of certain bacterial infections. Often termed “antimicrobial resistance (AMR) crisis”, the WHO identifies this situation as one major threat to global human health. To monitor this threat, the WHO started curating and evaluating data from all member states and since 2016 also curates a priority list of the most threatening antibiotic resistant bacterial pathogens, ranked among criteria like treatability, mortality, preventability and more. While the first iteration of this monitoring process in 2014 identified well known pathogens such as MRSA, antibiotic resistant *Mycobacterium tuberculosis* (both Gram+), *E. coli* and *Klebsiella pneumoniae* (both Gram-) as major threats, this first investigation was centred around resistant *E. coli*, *K. pneumoniae* and MRSA and their impact on public health (WHO, 2014). This changed drastically within the next 4 years of investigation where in the first iteration of the WHO priority list 25 bacterial species were evaluated and ranked among aforementioned criteria (Tacconelli *et al.*, 2018). This clearly illustrates that especially resistant Gram- pathogens pose the biggest threat today (Fig. 2).

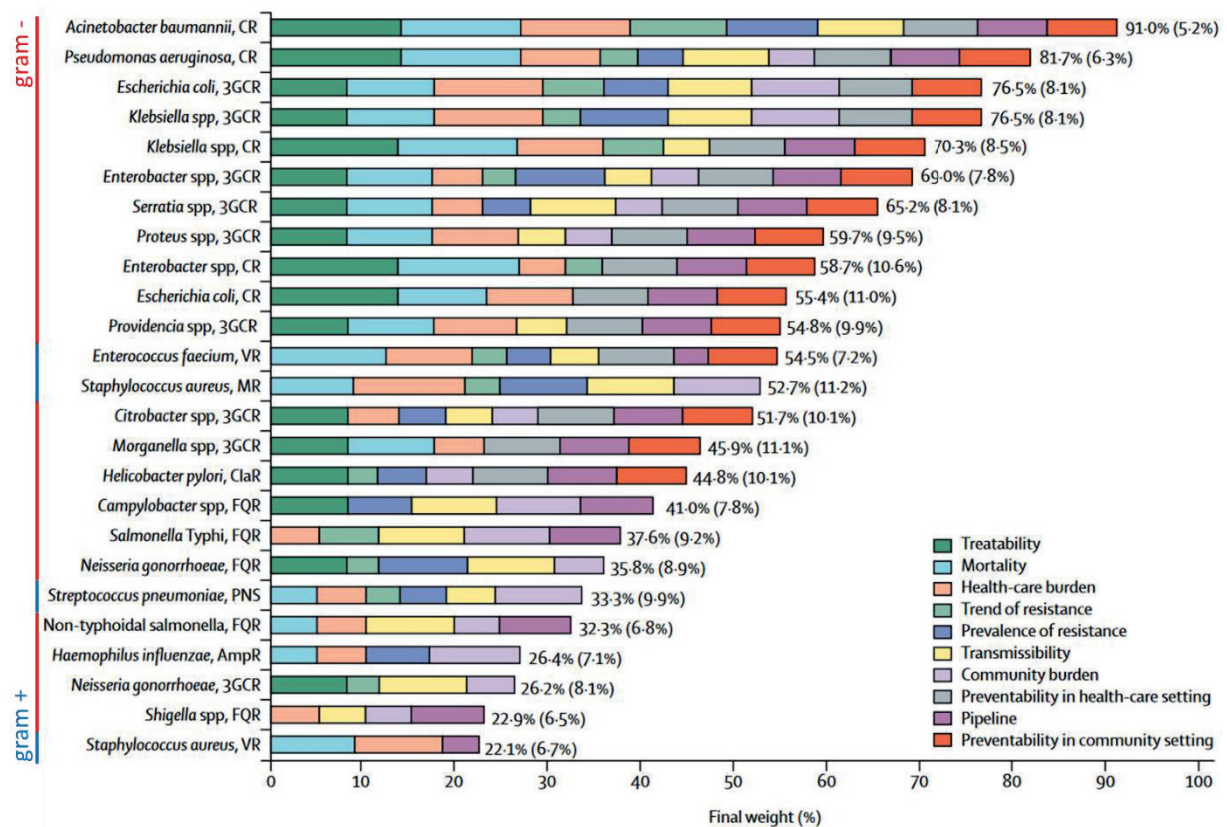


Figure 2: WHO priority list of antibiotic resistant bacterial pathogens, ranked among factors like treatability, mortality and healthcare burden with Gram classification indicated by a coloured bar on the left. (Tacconelli *et al.*, 2018, modified).

Nearly the full top half of this priority list is composed of resistant Gram- pathogens and only 4 of the 25 bacterial strains are Gram+. This list is led by *Acinetobacter baumannii*, followed by *Pseudomonas aeruginosa*, *Klebsiella pneumoniae* and *Escherichia coli*, with the first Gram+ pathogen *Enterococcus faecium* only at position 12. Unlike Gram+ bacteria, Gram- bacteria possess an outer membrane, which vastly restricts accessibility of targets for drugs with intracellular modes of action. Furthermore, horizontal gene transfer on small genetic elements, e.g. plasmids is common among Gram- bacteria, facilitating the rapid spread of newly evolved resistance mechanisms. This illustrates drastically, that new antibiotics that are active against these highly threatening pathogens need to be discovered and developed now. At the same time, this threat must be tackled from all angles, from close monitoring of resistance development to optimisation of hygiene concepts in the food industry, public health services and more.

I.1. Antibiotic Natural Products

The bulk of all antibiotics in medical use are NPs or derivatives thereof (Newman and Cragg, 2016) with soil-derived *Actinomycetes* and filamentous fungi being the predominant group of producers (Newman and Cragg, 2012), illustrating the importance of NP research. Despite considerable efforts to develop antibiotic discovery platforms based on chemical synthesis, it was not possible to replace NPs as the dominant source of antibiotics. While creation and high-throughput screening of synthetic compound libraries is cheaper than NP research, synthetic compounds usually lack the necessary chemical complexity (size, number of stereocenters etc.) for interference with the respective target and even if they show good interaction, they often fail to permeate lipid bilayers to access their cellular targets *in-vivo* (Lewis, 2020). Antibiotic NPs do not represent a uniform class of compounds. However, most antimicrobial NPs can be assigned to one of only few biosynthetic pathways which are covered in the following. Due to their importance for this work, ribosomally synthesised and post-translationally modified peptides (RiPPs) will be covered separately in detail in chapter III of this thesis.

I.1.1. Non-ribosomal peptides (NRPs)

Penicillins, representing the first discovered NP antibiotics are the product of a non-ribosomal peptide synthetase (NRPS), linking α -amino adipic acid, cysteine and valine to form a α -amino adipoyl-cysteine-valine tripeptide. This tripeptide subsequently undergoes modification to form the characteristic β -lactam (eponymous for this antibiotic class) moiety. NRP synthetases are large multidomain enzymes that resemble an assembly line architecture, with each module consisting of a condensation (C) domain, an adenylation (A) domain and a peptidyl carrier protein (PCP) domain. Commonly, individual modules are expanded to contain epimerisation domains, changing the stereochemistry of the incorporated building block. The first (or starter) module does not contain a C-domain, instead the A-domain links the first building block e.g. a specific free amino acid (AA) to a phosphopantetheinyl cofactor, bound to the PCP. The A domain of the second module loads the specific second substrate AA to its module's PCP and the C domain condenses the AA from the starter module to the AA located at the subsequent module forming a peptide bond. This process can be repeated by the third and any number of subsequent NRPS modules, generating a growing peptide chain with defined AA composition. On the last module, a thioesterase (TE) domain cleaves the thioester connecting the peptide chain to the enzyme, thus releasing the formed peptide chain. Upon cleavage, TE domains can release the linear AA chain or mediate intramolecular

cyclisation, changing the topology of the final molecule (Finking and Marahiel, 2004). Subsequently, tailoring enzymes can derivatise the peptide backbone further, e.g. by forming intramolecular rings or addition of halogens or glycosides. This assembly line logic (termed co-linearity rule) allows for a certain predictability of the final shape of the product, e.g. by enabling prediction of the order of building blocks. In contrast to a ribosomal peptide, NRPs are not limited to the 23 (+ 2 non-canonical) proteinogenic L-AAs. NRP synthetases are shown to specifically incorporate a variety of different D - and L – AAs, such as the aforementioned α -aminoadipic acid and also other organic acids such as β -hydroxylated fatty acids as is the case in lipopeptide biosynthesis. NRPs represent a large and diverse group of NPs, comprising many of the clinically relevant antibiotic in use today. Aside from the already mentioned β -lactams (Penicillins, Cephalosporines), glycopeptides (Vancomycin, Teicoplanin) and lipopeptides (Polymyxins, Daptomycin) are biosynthesised by NRPS pathways.

I.1.2. Polyketides (PKs)

Biosynthesis of polyketides is evolutionary related to the fatty acid biosynthesis and they share a common reaction mechanism for carbon chain elongation. Polyketide synthases (PKS) share the same large multidomain assembly line architecture, subdivided into modules responsible for a specific chemistry, with NRPS systems. Here, minimal modules consist of a ketosynthase (KS) Domain, an acyltransferase (AT) domain and an acyl carrier protein (ACP) domain. Individual modules often carry further domains, diversifying the chemistry, such as ketoreductase (KR) domains, dehydratase (DH) domains and enoylreductase (ER) domains. In contrast to NRPS systems, PKS systems do not load AAs but small building blocks such as malonyl-(CoA)-CoenzymeA or methylmalonyl-CoA. These building blocks are decarboxylated and fused by Claisen condensation to form a long carbon chain. This strongly resembles the fatty acid synthesis, in which malonyl-CoA building blocks are decarboxylated and the resulting acetate unit is incorporated into the growing carbon chain. This results in formation of a keto group in β -position for every incorporated acetate unit which is subsequently reduced to saturation of the carbon chain. In contrast to fatty acids, in PKS synthesis not all positions are fully saturated *per se*, but varying degrees of reduction (keto-, hydroxy and enol functions) at each incorporated building block can be observed. After being processed by all modules of this “assembly line”, the final carbon chain is also cleaved by a TE-domain, which again can form intramolecular cyclisation (Hertweck, 2009). In contrast to NRPs however, the residual reactivity of the highly functionalised carbon chain can also form further intramolecular reactions such as intramolecular ring formation and aromatisation in Tetracycline

biosynthesis, thus also contributing to the shape of the final molecule. Like NRP backbones, the final PK can be further augmented with other building blocks such as halogens or glycosides by tailoring enzymes. PKS gene clusters can be subdivided into three distinct types (Type I, II and III) and more subclasses (Shen, 2003). However, while this affects the domain organisation of these types of BGCs and to a degree the shape of the produced molecules, the underlying chemistry remains untouched. PKs also comprise clinically important antibiotics and other drugs such as the tetracyclines, macrolide antibiotics (Erythromycins, Fixadomycin) and immunomodulatory agents like Rapamycin. Furthermore, NRPS/PKS hybrid BGCs and the cognate compounds (e.g. the antibiotically active Antimycins) are commonly found in nature.

I.1.3. Aminoglycosides

Aminoglycosides are a class of NPs that are built from sugar derived monomers and unlike most other NP classes are exclusively produced by *Actinobacteria*. Initial biosynthetic step is the rearrangement of common monosaccharides (e.g. glucose) to the respective precursor aminosugars or monosaccharides, such as L-streptose, which are then connected to form di-, tri- or even larger polysaccharides. Aminoglycosides inhibit protein biosynthesis by interfering with the ribosome during protein translation. This can lead dose dependently to metabolic arrest of the target cell (bacteriostatic activity) or in higher dosage to cell death (bactericidal activity). Members of this class are among others the medically relevant antibiotics Streptomycin, Kanamycin and Neomycin. Furthermore, aminoglycoside antibiotics are commonly used in research laboratories to serve as selective marker for cloning procedures.

I.2. Antimicrobial activity

While upon discovery of the first antibiotics such as the sulfonamides or penicillins lack of knowledge of bacterial physiology impeded elucidation of their precise mode of action, progress in the fields of microbiology, biochemistry and the advent of genetics helped to characterise antibiotic - target interaction. Not only did this generate knowledge about bacterial physiology, but also valuable information about antibiotic resistance mechanisms. Most antibiotics interfere with conserved bacterial targets such as the cell wall or protein biosynthesis. Depending on the target and its evolutionary conservation, antibiotics can be active against a multitude of bacterial families such as the β -lactams (broad spectrum) or very specific against a bacterial family (narrow spectrum)

such as Murepavidin. Murepavidin targets the outer membrane protein LptD and is only active against Pseudomonads (Srinivas *et al.*, 2010). Identification of cellular targets also enabled design of specific screening platforms against respective targets, which can assist discovery of novel antibiotics specifically active against this target. Despite the fact that great knowledge of antibiotic modes of action exists today, this does not imply that the list of antibiotic targets can be concluded and unprecedented modes of actions can still be discovered. In contrast to the immense progress in studying antibiotic modes of action, for some molecules like Daptomycin the mode of action remained poorly understood for decades. Daptomycin has complex mode of actions, which historically led to proposal of a multitude of targets affected by it. Only recently, one specific mode of action that incorporates previous in part contradictory findings could be established (Grein *et al.*, 2020). Such observations often lead to the term “cryptic mode of action”, i.e. the mode of action of one compound cannot be explained with specific well-investigated compound-target interactions. This however does not necessarily point towards novel mode of actions since a combination of different modes of action with synergistic effect can also be exhibited by one molecule. This impedes elucidation, especially if several targets are hit where it remains unclear how inhibition of one target affects the function of another target. However, elucidation of the mode of action is crucial for clarification of structure activity relationship and fine tuning of antibiotic screening platforms.

To date, a multitude of antibiotic modes of action were elucidated in detail. In the following, targets for the most familiar antibiotic compounds are covered.

I.2.1. Cell wall

Most familiar to many people is the mode of action of the first discovered NP antibiotics, the Penicillins. Despite the early discovery, the first biochemical mode of action was proposed in 1965 (Tipper and Strominger, 1965) and was completely elucidated in 1980. Penicillin acts as an analogue to acyl-D-alanyl-D-alanine. The reactive β -lactam moiety irreversibly binds to the active site serine of the D-alanyl-D-alanine-carboxypeptidase, involved in peptidoglycan crosslinking (Yocum *et al.*, 1980). Lack of this crosslinking between the linear peptidoglycan strands significantly weakens the bacterial cell wall and inhibits bacterial growth, thus β -lactam antibiotics exhibit bacteriostatic activities. Being the second antibiotic in application and the most widely used antibiotics today, β -lactam antibiotics provoked occurrence of resistant bacterial pathogens. Bacteria developed two distinct methods to overcome inhibition by β -lactam antibiotics. First method is the modification of the target protein to allow higher affinity to their physiological substrates and lower affinity to β -

lactam antibiotics, second method is the expression of β -lactamases, cleaving the β -lactam warhead, hence inactivating the molecule. While the first strategy is rather uncommon, the expression of β -lactamases can be commonly observed in clinical pathogens due to the easy spread on small genetic elements like plasmids. Glycopeptides such as Vancomycin also interfere with cell wall biosynthesis similar to the β -lactam antibiotics, however in contrast to the β -lactams they bind the D-alanyl-D-alanine residue of the peptidoglycan instead of the active site of the crosslinking enzyme (Watanakunakorn, 1984). Due to their bulkiness they cannot permeate Gram- bacterial outer membranes, hence they only exhibit anti Gram+ activity. More recently, lipoteichoic acids (LTA) were also identified as antibiotic target. These molecules augment the peptidoglycan cell wall in Gram+ bacteria and are essential for their growth (Richter *et al.*, 2013). While so-far no antibiotics targeting LTAs reached the market, many compounds were identified to interfere with LTA formation *in-vitro* and *in-vivo*, making LTAs an attractive new antibiotic target (Pasquina *et al.*, 2013).

I.2.2. Outer membrane

Bacterial outer membranes (OM, exclusively Gram-) represent a strong barrier to overcome for antibacterial drugs. However, the OM itself is a highly functionalised part of bacterial physiology, hence also plays a role in a multitude of processes important for survival. This makes the OM and proteins located on the OM attractive targets for antibiotics. Lipopeptides such as the Polymyxin antibiotics in general have a destabilising effect on lipid bilayers (Velkov *et al.*, 2013), making them highly potent molecules against Gram- bacteria. This general membrane-destabilising mode of action however causes severe side effects. Hence, these types of compounds were only pushed into application due to the lack of other effective antibiotics. Destabilising the OM simultaneously disturbs all functions associated with it, such as the maturation of signalling peptides or efflux pumps, the generation of lipopolysaccharide (LPS), an important virulence factor of Gram- bacteria. Destabilisation of the OM also allows influx of a variety of molecules into the periplasm, which impedes H⁺ driven generation of energy in the respiratory chain. Specific protein targets on the OM are LptD, which exports LPS to the surface of the cell and BamA, a chaperone that inserts proteins into the OM (Lewis, 2020). LptD is targeted by Murepavidin, a peptidomimetic antibiotic (Srinivas *et al.*, 2010). Due to the importance of BamA for this work, its function will be further elaborated in detail in chapter III.

I.2.3. Intracellular targets

To access intracellular targets, the molecules have to either permeate the bacterial lipid bilayer(s) and the cell wall or exploit the cells active transport mechanisms. Due to bacterial lack of compartmentation, metabolically critical functions are not further protected, once a substance enters the cytoplasm. Especially protein biosynthesis is targeted on every level by different antibiotics. Among others, α -pyronines such as Corallopyronine A, Rifamycin and Lipiarmycins bind the bacterial RNA-polymerase (RNAP), hence blocking the transcription of new mRNA. Aminoglycosides and Tetracyclines bind the 30S ribosomal subunit and macrolide antibiotics, such as Erythromycin bind the 50S ribosomal subunit, hence disturbing the translation of mRNA into a corresponding AA chain, which has significant effect on the viability of the cell. While arrest of protein biosynthesis results in bacteriostatic activities, some antibiotics like the Aminoglycosides dose-dependently cause mistranslation or translation of truncated proteins, causing more perturbation to the metabolism, ultimately resulting in cell death. Synthetic (Fluoro)quinolone antibiotics, such as Nalidixic acid, inhibit the ligase activity of the gyrase which supercoils the bacterial DNA, critical for successful mitosis. Lack of this ligase activity also results in DNA strand breaks, which subsequently also leads to cell death.

I.3. Natural product discovery – Past, present and future?

Historically, the bioprospecting or “classical” approach dominated the discovery of new NPs and NP antibiotics. Talented natural product producers, such as *Actinomycetes* or filamentous fungi are sampled from the environment and cultivated in axenic cultures. The resulting cultures are extracted by maceration with organic solvents or solid phase extraction, generating extracts that undergo screening for activity against well studied test organisms or systems. Once activity is detected, this crude extract is fractionated to find the active compound that is further characterised and may be developed into a drug. While this basic and (relatively) inexpensive method was very successful in finding new antibiotics in the first decades of antibiotic discovery, its limits are obvious. With increasing efforts, the chance of rediscovery of already known molecules and the need for quick dereplication rises. This led to the seeming depletion of soil dwelling *Actinomycetes* as source for novel NP scaffolds. To compensate for this, the classical approach was expanded to other talented producer genera such as *Myxococcales* (Moghaddam *et al.*, 2018, Hoffmann *et al.*, 2018) or *Bacilli* (Stein, 2005), the producers of e.g. Corallopyronines and Polymyxines or to talented producers from different ecological niches like the marine environment (Bowling *et al.*, 2007). The ability to expand

these techniques to new source organisms is highly dependent on nature's biodiversity, since this biodiversity is the basis for the chemical diversity we find in NPs. (Micro)organisms evolve and optimise NPs to adapt to their natural habitat, e.g. to weaken or kill prey or gain upper hand in the combat for scarce resources. Thus, optimising the effectivity of their NPs for their specific task. Loss of this biodiversity in turn causes loss of the highly optimised chemistry that not only brought some of the most important drugs in use today but also the ability to find more chemical structures that have benefit for humankind from natural resources. Protection of biodiversity therefore is more than just the idea of conserving healthy ecosystems (which we are a part of) e.g. to sequester CO₂ or to exploit as food source, but also to conserve the resource that like few others contributed to modern medicine. Monitoring biodiversity and exploration of so far underinvestigated habitats naturally involves sampling hence, NP discovery goes hand in hand with ecological research. Aside from the rediscovery of known compounds, bioprospecting suffers from the poor recovery rate of so far unknown microbes from environmental samples. This phenomenon is referred to as the microbial dark matter and this term describes the fact that to-date we can only cultivate 0.1 % to 1 % of all microbial life under laboratory conditions (D'Onofrio *et al.* 2010). While we see the remaining ~ 99 % under microscopes and even in metagenome 16S rDNA sequencing, a vast majority of microbes remains hidden from investigation. At the same time, modern sequencing methods and projects clearly show that these hard to cultivate or even seemingly unculturable strains also have potential for NP production (Wilson *et al.*, 2014). Considerable improvements in isolation of so far uncultured bacteria, as was demonstrated with the discovery of the potent anti-Gram+ antibiotic Teixobactin using the iChip cultivation method, now also contribute to filling the antibiotic pipeline (Ling *et al.*, 2015). At the same time, it also showcases that exploring the microbial dark matter remains a difficult and labour-intensive task.

Improvements (and more so reduction in cost) of chemical analytical methods and equipment like NMR spectroscopy and especially mass spectrometry facilitate the fast identification of known molecules in early stages of investigation. Another revolution that changed NP research into a very contemporary science was the ever-increasing availability of genome sequences. With this at hand, the biosynthetic origin of most NPs in use could be elucidated and common biosynthetic routes such as the polyketide synthesis by polyketide synthases (PKS) or assembly of non-ribosomal peptides by non-ribosomal peptide synthetases (NRPS) could be defined (Jenke-Kodama and Dittmann, 2009). This knowledge enabled identification of natural product biosynthetic pathways based on sequence data alone and substantial breakthroughs in this field made it possible to quickly grasp the biosynthesis of NPs on a genetic level. Many tools were developed to provide information of the secondary metabolism of bacteria with only the genome sequence available. Among them most

widely known is the antiSMASH platform which allows identification of NP biosynthetic gene clusters (BGCs) based on the genome sequence alone and is constantly updated to incorporate new findings (Medema *et al.*, 2011). With this understanding of the underlying biosynthetic genes, identification and prioritisation of talented producer genera based on genome information is much faster and easier than relying on published chemical data. Furthermore, genetic information can drastically facilitate the elucidation of complex chemical structures, e.g. large non-ribosomal peptides. This accelerates NP research drastically and at the same time allows targeted reinvestigation of already well investigated NP producers for so far undiscovered structures.

Identification of NP BGCs also sparked the idea of heterologous expression of NP biosynthetic pathways in suited host organisms. While historically, this has been a challenge, modern cloning techniques make it possible to express even large BGCs in well studied surrogate hosts (Hashimoto *et al.*, 2018). Heterologous expression has many advantages for production of NPs, from providing a cleaner chemical background for isolation to the ability to overproduce NPs with well characterised regulatory elements. While heterologous expression is no guaranteed success, it became a valuable tool to study biosynthetic pathways, due to relative ease of DNA manipulation in an artificial system. The expression of BGCs without a known associated NP to find the corresponding NP so far remains challenging at best and success rates are low. However, successful heterologous expression provides a platform to further study the underlying biosynthetic machinery and hence is the go-to method for in depth exploration of NP biosynthetic pathways. This is specifically important if the natural producer organism is inaccessible for genetic manipulation. The improvement of cloning techniques also gave rise to the idea that targeted manipulation of BGCs might result in formation of new structures related to the respective NP. Simplest iteration of this idea is the feeding of synthetic precursors while genetically eliminating the pathway of the natural precursor leading to accumulation of an “unnatural” NP as exemplified by the mutasynthetic creation of Erythromycin and Avermectin analogues like Erythromycin D or Doramectin (Weist and Süssmuth, 2005). While mutasynthetic modification of a NP does not necessarily improve activity, combined with subsequent semi-synthetic derivation it may result in higher activity, broader activity spectrum, more stable products or in general products with favourable pharmacochemical properties. Notable example for application of mutasynthesis is the production of semisynthetic β -lactams, where engineering of *Penicillium chrysogenum* shifted production from Penicillin G to adipylyl-7-aminodeacetoxycephalosporanic acid (adipyl-7ADCA). Adipyl-7ADCA can be converted in one enzymatic step to 7ADCA, precursor to Cephalexin and Cephadroxil, avoiding energy and pollution intensive chemical modification of Penicillin G to 7ADCA (Wegman *et al.*, 2001). With deeper understanding of biosynthetic pathways, complete reprogramming of BGCs now becomes possible,

as demonstrated by reprogramming of the Antimycin type NRPS/PKS assembly line, leading to creation of Antimycin analogues with contracted and expanded ring size and exchange of chemical modules (Awakawa *et al.*, 2018). A most intriguing breakthrough was the identification of NRPS docking domains, governing the acceptance of NRPS A-domains towards the products of their upstream modules. This allowed free reconfiguration of NRPS assembly lines and thus free design of non-ribosomal peptides to be synthesised by respectively stitched artificial BGCs (Cai *et al.*, 2019). Both profited from the availability of modern cloning techniques and well-studied heterologous hosts for expression of the respective BGCs and the reprogrammed derivatives thereof.

While prediction of the future is always vague, certain recent publications give insight into how the future of natural product discovery might look like. Microbial growth was long considered a black box, where choice of nutrients is the input and chemistry produced by the organism the output information. Hence, diversifying the cultivation conditions (One strain many compounds - OSMAC) for screening talented producer strains is a go to method to induce production of different NPs, since the underlying processes triggering production of different NPs is still poorly understood. Here, transcriptomics-based approaches can give insight into which condition is activating a pathway of interest and ease linking a BGC to its respective NP (Amos *et al.*, 2017). Natural succession to this technique could be the screening of small molecule libraries and their effect on NPs produced by one strain (rather than testing 10 different media). The feasibility of this was demonstrated by the “High throughput elicitor screening” (HiTES) platform that enables simultaneous screening of ~ 500 small molecules from a pharmaceutical compound library (synthetic and NPs) to induce production of NPs like the Vancomycin analogue Keratinimicin in *Amycolatopsis keratiniphila* (Xu *et al.*, 2019). This technology at the same time can be augmented with activity assays or other reporter methods, thus delivering a powerful and versatile platform for rapid NP discovery. Furthermore, advances in understanding and thorough investigation of the biosynthesis of NP groups coupled with the ever-growing number of sequenced microbial genomes enables targeted genome mining for specific pathways that encode a wide variety of NPs. Intriguingly this works very well, even with NP pathways that so far could not be tackled by rule-based detection, e.g. the ribosomally synthesised and post-translationally modified peptides (RiPPs) (Bhushan *et al.*, 2019, Kloosterman *et al.*, 2020). Combining this with the increasing availability of machine learning tools can facilitate the screening of large genome datasets, point towards new NP pathways and allow identification of so far undiscovered structures. Recently, a neural network approach was already successfully applied to facilitate NMR based structure elucidation (Reher *et al.*, 2020), showing that this technology can contribute to NP chemistry, especially on the level of data interpretation. Furthermore, the discovery of Halicin as a broad-spectrum antibiotic by a neural network driven screen of a synthetic

compound library demonstrates that this technology has the potential to accelerate progress in antibiotic discovery. While Halicin is a synthetic compound and only synthetic compounds were screened, this work heavily relied on NPs and their activity to train the neural network (Stokes *et al.*, 2020). This illustrates that while not every NP will become a drug candidate, the knowledge generated from NP research over the past decades is useful and in manyfold ways contributes to drug research.

These advances more and more put the idea of microbes as a black box into question. First breakthrough was the ability to sequence genomes of strains of interest, basically revealing the hardwiring of all cellular processes. With advance of modern whole RNA sequencing (deep sequencing) and proteome profiling technology, it becomes possible to also get information on the processes that steer the cell under different conditions. With these tools combined, it will become possible to dissect the process of life close to wholistic understanding. Hence, it also will enable a thorough understanding of NP metabolism, not only from the perspective of NP production but also from the perspective of the recipient that is exposed to them (Lewis, 2020). This will drastically accelerate all stages of NP discovery, from the initial dereplication of NPs produced to the elucidation of cryptic modes of action and targets, possibly even allowing the determination of the function of NPs in the natural habitat.

I.4. Scope of this thesis

In this work, three different projects in the field of antibiotic research are presented. Each chapter will contain separate short introduction, material and methods, results and discussion parts. References and supplementary material for all chapters will be in one separate section at the end of this thesis.

In Chapter I “Antimicrobial Potential of Bacteria Associated with Marine Sea Slugs from North Sulawesi, Indonesia”, bioprospecting of the Celebes sea around Manado and Bunaken Island for isolation of talented bacterial NP producers and their initial characterisation for antibiotic production is elaborated.

Chapter II describes the identification and heterologous expression of the minimal Pseudouridimycin biosynthetic gene cluster.

Chapter III covers the identification and proof of the Darobactin BGC, as well as the discovery and mutasynthetic production of Darobactin analogues with altered AA composition. Furthermore, the

discovery of three additional Darobactin analogues carrying two distinct modifications to the Darobactin A core structure in the natural Darobactin producers *Pseudoalteromonas luteoviolacea* H33 and H33S is described. All but one of the novel Darobactin analogues were tested against a panel of clinically important Gram– pathogens.

II. ANTIMICROBIAL POTENTIAL OF BACTERIA ASSOCIATED WITH MARINE SEA SLUGS FROM NORTH SULAWESI, INDONESIA

II.1. INTRODUCTION

A promising strategy to identify novel biologically active compounds is to explore new habitats. Compared to the terrestrial environment, the oceans are still underinvestigated, even though they make up two third of the planet's surface, and harbor high biodiversity with an immanent high potential for discovery of new metabolites (Donia and Hamann, 2003). The rationale underlying the idea of targeting the defined biological niche of marine sea slugs, their food source and associated microorganisms is the highly competitive nature of this environment. The slugs must protect themselves against various predators, e.g. fish and crabs, and against settlement of pathogenic microorganisms on their epidermis. To reach this, chemical defence mediated by natural products plays a major role. It was already shown that the slugs incorporate and enrich bioactive metabolites from their food source (Bogdanov *et al.*, 2014, Bogdanov *et al.*, 2016). However, it has to be kept in mind that compounds present in the food source may be rather produced by associated microorganisms, e.g. bacteria. The antitumor agent Dolastatin 10 was first described from the anaspidean *Dolabella auricularia* (Anaspidea). This compound was thought to be sequestered from algae, the food source of the heterobranchs *Sacoglossa* and *Anaspidea*. However, more recent findings of Dolastatin 10 in cyanobacteria impose an origin in bacteria associated with the algal food (Cimino and Gavagnin, 2006, Luesch *et al.*, 2001, Gerwick and Fenner, 2013). A similar relationship between bacteria and higher organisms can also be expected for many compounds originally described from sponges (Mehbub *et al.*, 2014, Wilson *et al.*, 2014). In this article, a culture-dependent approach to analyse the potential of the slug-associated microbiome is described. The isolated bacteria were tested for the presence of polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NRPSs) coding genes, since a majority of pharmaceutical interesting compounds is produced by these biosynthetic pathways. Further, the available extracts from axenic bacterial cultures were tested for antimicrobial activity against relevant pathogens.

II.2. MATERIAL AND METHODS

II.2.1. Environmental Samples

Sampling was performed at the Bunaken National Park in August 2015, located near Manado, North Sulawesi, Indonesia. In total, 550 Samples were obtained from different sites around the island by snorkeling and scuba diving (Fig.: 3). Samples were either directly processed or cooled until further processing in the lab.

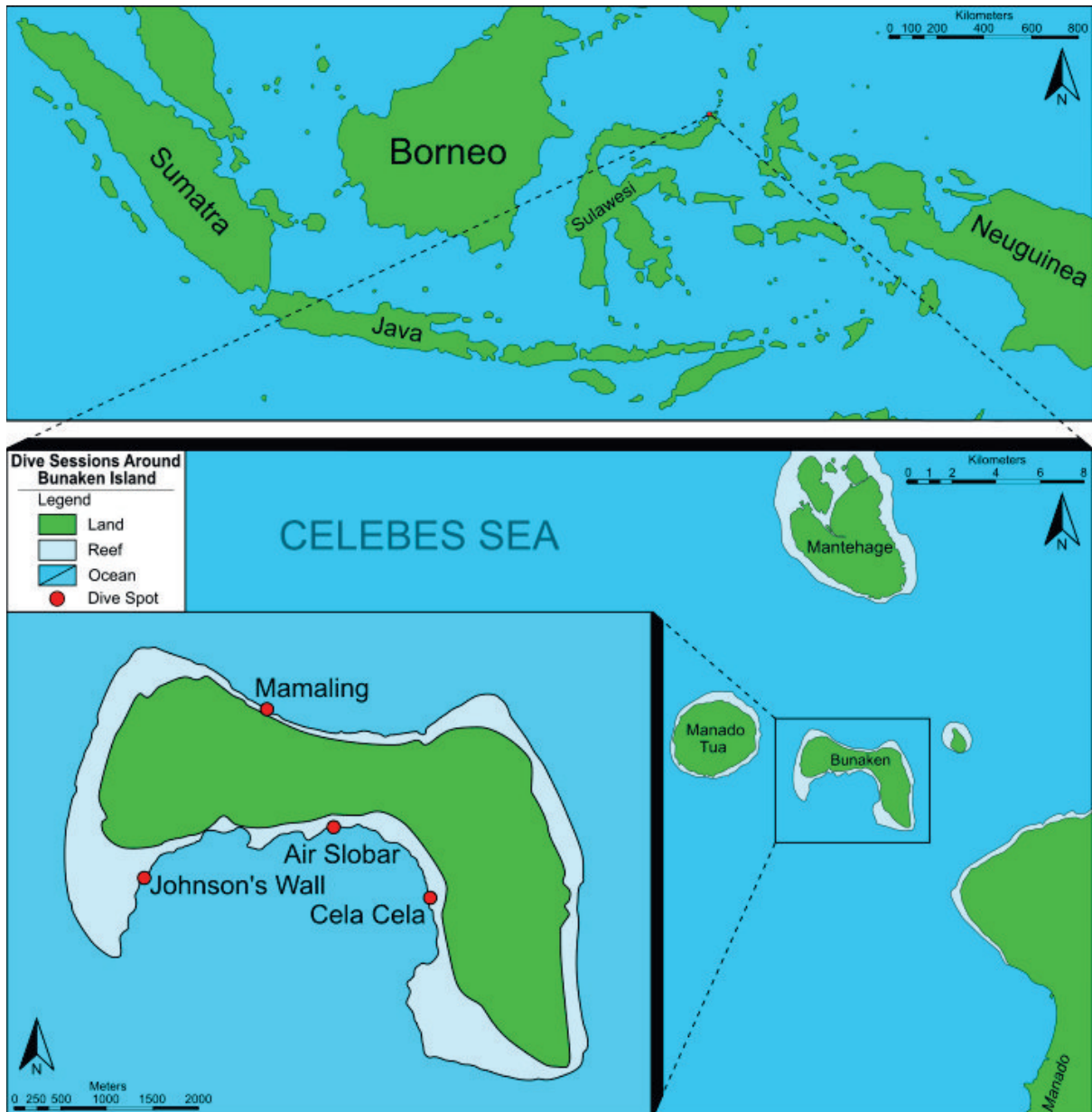


Figure 3: Map of the sampling sites around Bunaken Island, Bay of Manado, Sulawesi, Indonesia.

II.2.2. Identification of Sea Slugs and Sponges

Preliminary identification of slugs and egg masses was performed based on Gosliner et al., 2008 and Gosliner *et al.*, 2015, identification of sponges based on Atlas of living Australia (<https://www.ala.org.au/>). For barcoding, a small portion of the sampled slugs, egg masses and sponges was stored in 96% EtOH and the DNA was isolated using QIAgen DNeasy Blood and Tissue-Kit (Qiagen, Hilden, Germany). Partial 16S rDNA sequences and partial CO1 gene sequences were amplified using 16Sar-5' and 16Sbr-3' primers (Palumbi, 1996) and LCO1490-JJ and HCO2198-JJ primers (Astrin and Stüben, 2008) and sent to Macrogen Inc. (Seoul, Korea). The specimens' identities were determined by subsequent blastn analysis of the obtained sequences.

II.2.3. Isolation of Bacteria

Prior to the isolation of bacteria, the surface of the samples (in the following the identification code of each sample is used; further information can be found via the internet portal of Diversity Workbench in the module Diversity Collection: diversityworkbench.net/Portal/DiversityCollection). Chan15Bu-2 (*Chromodoris annae*), Chdi15Bu-3, Chdi15Bu-55 and Chdi15Bu-38 (*Chromodoris diana*), Chsp3015Bu-4 (*Chromodoris* sp. 30), Chwi15Bu-2 (*Chromodoris williani*), Chan15Bu-11E (*Chromodoris annae* egg mass), Glst15Bu-1 (*Doriprismatica stellata*), Glst15Bu-1E (*Doriprismatica stellata* egg mass), Glst15Bu-1P (*Doriprismatica stellata* sponge), Hesa15Bu-1 and Hesa15Bu-2 (*Hexabranhus sanguineus* egg mass) and Phpu15Bu-1 (*Phyllidiella* cf *pustulosa*) were washed with sterile H₂O, disinfected with 70 % ethanol and washed again with sterile deionised H₂O (Milli-Q, Merck, Germany) to remove contaminations from sampling and documentation. The samples were frozen in liquid nitrogen and homogenized using a Potter S Homogenizer (Sartorius, Germany). The homogenisate was brought onto Marine Broth agar plates (37.9 g/L Marine Broth (Difco), 15 g/L agar) or ISP2 agar plates containing 2 % NaCl (4 g/L yeast extract, 10 g/L malt extract, 4 g/L glucose, 20 g/L NaCl, 15 g/L agar) using an inoculation loop and the 13 streaks technique to obtain axenic colonies. Axenic colonies were transferred to new plates. For cryo-conservation, axenic cultures were grown in the respective liquid medium, mixed 50:50 with sterile glycerol and stored at -80°C.

II.2.4. Generation of Extracts and Activity Screening

Bacteria were grown in 30 mL of the respective medium. Therefrom, 25 mL were transferred to a 50 mL reaction tube, 25 mL ethyl acetate were added and the tube was shaken thoroughly. The organic phase was carefully removed with a pipette and evaporated. The extracts were dissolved to 1 mg/mL in methanol for HR-LCMS and to 10 mg/mL in DMSO for antibiotic testing.

2.5 Antibacterial Assays

For antibacterial assays, test strains were grown in 200 μ L cultures in flat bottom 96 well plates at room temperature. The initial OD₆₀₀ was set to 0.1 and the extracts were added to yield a final concentration of 0.5 μ g/ μ L per culture for the first well and subsequently diluted 1:1 in the following wells to the lowest concentration of 4 ng/ μ L. 0.5 μ g/ μ L ampicillin served as positive control, pure DMSO as negative vehicle control. The tests were evaluated when the negative control reached an OD₆₀₀ of 1.0. Extracts were tested against *Arthrobacter psychrolactophilustolerans* and *Escherichia coli* XL1Blue. For the preparation of *Staphylococcus aureus* COL and *Escherichia coli* O-19592 assay plates, cation adjusted Müller Hinton agar was autoclaved and allowed to cool down to 50-55°C. 10⁷ cells/ml of the test strains were seeded into the agar and after solidification, 3 μ L of the ethyl acetate extracts were spotted onto the plates. Ampicillin (0.1 μ g/ml) and ethyl acetate served as positive and negative controls, respectively. Plates were incubated at 37°C for 18 hours and checked for inhibition zones.

II.2.5. Detection of PKS and NRPS Biosynthetic Genes

Genomic DNA of the different bacterial strains was isolated using the GenElute Bacterial Genomic DNA Kit (Sigma) according to the manufacturer protocol. The genomic DNA was subjected to several PCR screening rounds using degenerated primer (see Suppl. Table 2). Primers were based on literature or designed based on conserved regions in ketosynthase (KS) domains of PKS and adenylation domains (A-domains) of NRPS. Specific PCR for serine activating A-domains was performed as nested PCR, i.e. the PCR products of a general A-domain PCR with degenerated primers were gel purified and used in a subsequent PCR with AnSerin/JS002 as primer pair. The amplified products of approximately 700bp (KS domains), 700bp (A3-A7 region of A-domains), 1000bp (A2-A8 region of A-domains) and 450bp (serine specific nested PCR) were gel purified and cloned into pGEM-T vector (Promega, Madison, USA), transformed into chemical competent *E. coli* Alpha Silver Select Efficiency (Bioline, Luckenwalde, Germany) and sequenced. Sequences were analysed using blastx and blastn and closest protein sequences of were analysed for amino acid specificity using antiSMASH 3.0

(Weber *et al.*, 2015) and submitted to Genbank. The accession numbers for this project are KY671135 - KY671183 (16S rDNA), KY671132 - KY671134 (KS) and KY671184 - KY671199 (A domains).

II.2.6. Phylogenetic Analysis

The bacterial species were determined by 16S rDNA sequencing. Therefore, genomic DNA was isolated from 2 mL of liquid culture using the GenElute Bacterial Genomic DNA Kit (Sigma) and the 16S rDNA region was amplified by PCR using rD1 and fD1 primers (Weisburg *et al.*, 1991). The resulting amplificate was cloned into pGEM-T vector (Promega, Madison, USA), introduced into *E. coli* Alpha Silver Select Efficiency cells (Bioline, Luckenwalde, Germany). The plasmids were isolated using the PureYield Miniprep kit (Promega, Madison, USA) and the insert was sequenced from both sides using T7 and SP6 primers (GATC, Konstanz, Germany). Vector fragments in the sequencing result were removed using VecScreen-Blast and the full 16S rDNA was assembled. BLASTn search of the 16S rDNA sequences revealed the closest relatives, thereby enabling determination of the species, due to the fact that identity to known type strains was in the range of 98% to 100%. The phylogenetic tree of all obtained bacterial isolates was created by aligning the 16S rDNA sequences of the respective type strains using MAFFT (Katoh *et al.* 2002) and constructing the tree using raxML (Stamatakis, 2014).

II.3. RESULTS

II.3.1. Isolation and Identification of Sea Slug Associated Bacteria

Preliminary identification of the slugs and egg masses were confirmed by barcoding. Metadata of the collected macroorganisms will be available via the internet portal of Diversity Workbench within the modul DiversityCollection: <http://diversityworkbench.net/Portal/DiversityCollection>. 49 bacterial strains were isolated as axenic culture starting from 12 different samples. The highest proportion, of strains (20% each) was isolated from the egg masses Hesa15Bu_1 and Hesa15Bu_2 of *Hexabranhus sanguineus* whereby only 2% (equates to 1 strain) were isolated from the processed sponge sample. Most of the strains were recovered using Marine Broth standard medium, while 10 strains were recovered by using ISP2 medium optimized for actinobacteria (see Material and Methods). For a full table of strains and their respective source see Suppl. Table 1.

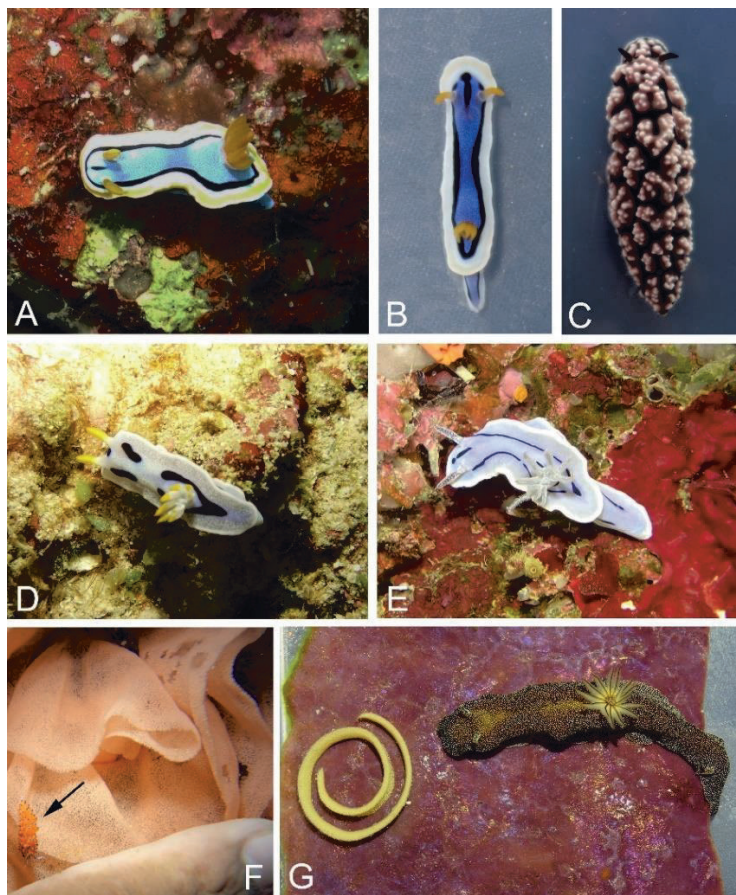


Figure 4: Higher organisms used in this study: A) *Chromodoris annae* (Chan15Bu-2) in its natural environment. B) *Chromodoris annae* (Chsp30_15Bu-4) from same locality but with different coloration. C) *Phyllidiella* cf. *pustulosa* (Phpu15Bu-1). D) *Chromodoris diana* (Chdi15Bu-55) in its natural environment. E) *Chromodoris willani* (Chwi15Bu-2) in its natural environment. F) *Hexabranhus sanguineus* egg mass (Hesa15Bu-2) in its natural environment and with a nudibranch, *Favorinus tsuruganus* feeding on the eggs (arrow). G) *Dorioprismatica stellata* (Glst15Bu-1) with its egg mass (Glst15Bu-1E) on the typical sponge *Phyllospongia* cf. *lamellosa* (Glst15Bu-1P).

II.3.2. Antimicrobial Activity Assays

Axenic bacterial strains were cultivated in liquid medium and ethyl acetate crude extracts were prepared. These extracts were first tested against the Gram+ test strain *Arthrobacter psychrolactophilustolerans* and the Gram- test strain *E. coli* XL1Blue. 35 strains were tested positive in this assay. However, to prioritize the samples, the 10 most active extracts were selected for further testing against relevant pathogens, i.e. *Staphylococcus aureus* COL (MRSA) and *E. coli* O-19592 (EHEC). In this assay, three extracts showed anti-MRSA activity and one extract showed anti-EHEC activity (Fig.: 5). It can be seen that especially isolates of *Pseudoalteromonas* strains showed considerable activity. However, the strain showing activity against the Gram- EHEC strain was a *Marinomonas* species.

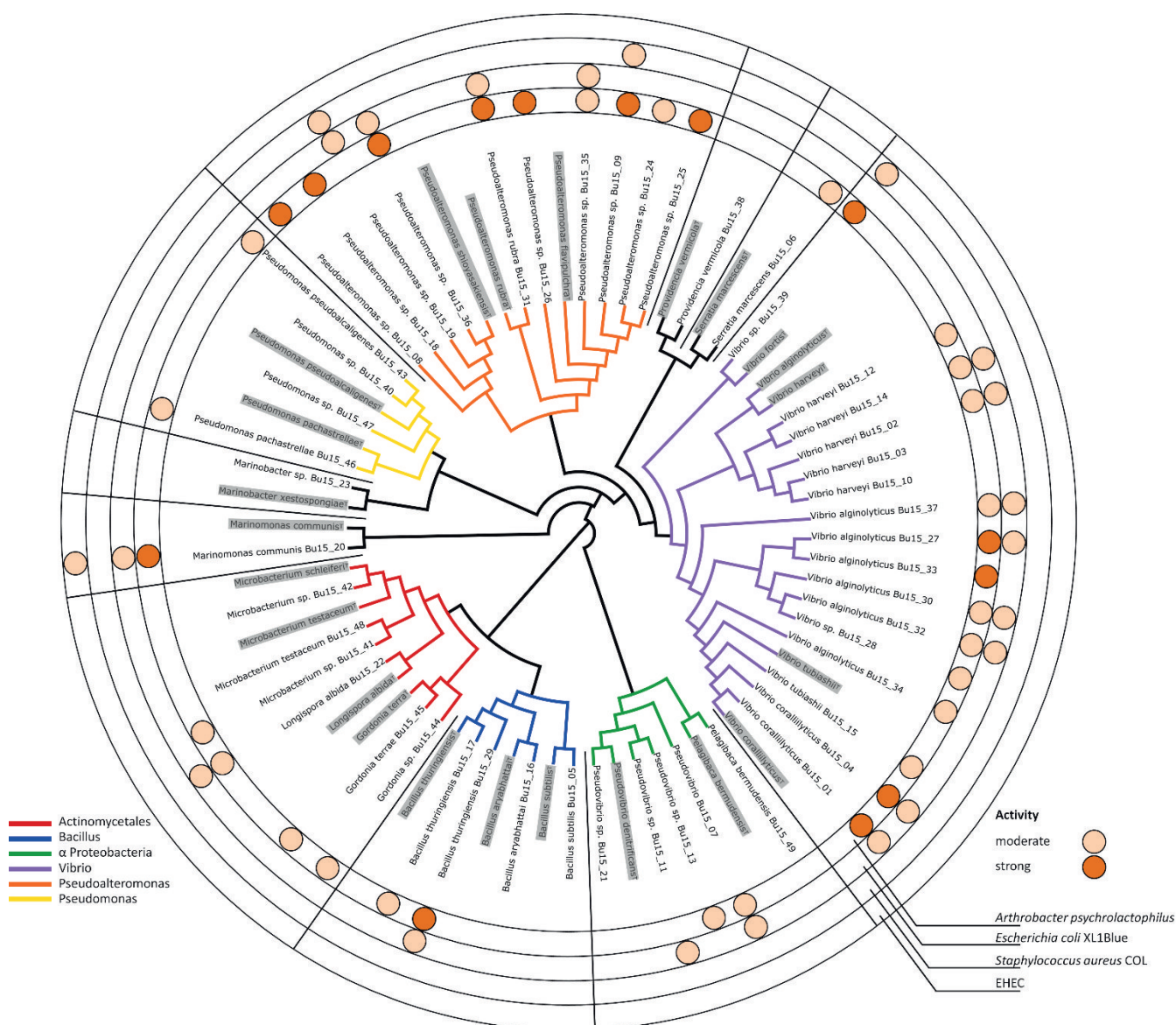


Figure 5: Phylogenetic tree of the isolated axenic bacterial strains and the respective closest type strain (marked in grey). Detected activities towards Gram+ and Gram- test strains, thereunder clinically relevant MRSA and EHEC isolates, are indicated with dark orange (strong activity) and mildly orange (weak activity) circles.

II.3.3. Phylogenetic Analysis

For all isolated axenic bacteria the 16S rDNA gene was amplified using rD1 and fD1 primers, designed for the specific and long range amplification of eubacterial 16S rDNA. All isolated strains belong to known genera and show high homology (98-100% identity) to sequences stored in the database. The closest homologue type strain of each isolated strain was included in an alignment to obtain a phylogenetic tree (Fig.: 3). From the obtained isolates, 35 out of 49 strains (equal to 71%)

showed antibacterial activity. All of the 10 isolated *Pseudoalteromonas* strains showed activity, and 7 thereof revealed strong activity against the Gram+ test strain. Further, 2 strains also inhibited methicillin resistant *Staphylococcus aureus* (MRSA). Another bacterial group with a high proportion of strains showing antibacterial effects was the *Vibrio* group. Therein, 14 out of 16 isolated strains showed at least slight activity, 5 strains exhibited strong activity against the Gram+ test strain, and 1 strain showed anti-MRSA activity. One isolate even showed activity against enterohemorrhagic *Escherichia coli* (EHEC). This strain was the only isolate from the *Marinomonas* clade.

II.3.4. Detection of PKS and NRPS Biosynthetic Genes

All 49 strains were tested for the presence of DNA sequences known to be associated with biosynthetic gene clusters (BGCs) contributing to the production of specialized metabolites, i.e. degenerated primers were used to screen the genomic DNA for the presence of NRPS sequences. Primers were adapted from literature to target conserved regions of A-domains. Despite many trials, the number of sequences obtained by this approach was very low, i.e. 10 A-domain sequences (6 different from *Gordonia terrae* Bu15_45, 2 different from *Pseudoalteromonas rubra* Bu15_31, 1 from *Pseudoalteromonas* sp. Bu15_09 and 1 from *Gordonia* sp. Bu15_44). Since oxazole natural compounds are amongst valuable antibiotics, such as Streptogramin A, a two-step nested PCR approach for serine specific A-domains was projected by designing a primer between the conserved regions A3 and A7 with specificity to serine incorporating A-domains (AnSerin see Suppl. Table 1). This enabled us to obtain further 6 sequences (1 from *Vibrio coralliilyticus* Bu15_04, 3 different ones from *Pseudoalteromonas* sp. Bu15_09, 1 from *Pseudoalteromonas* sp. Bu15_26 and 1 from *Bacillus thuringiensis* Bu15_29). All except one proved to be partial A-domain sequences with first blastx hits of A-domains predicted to incorporate serine. *Gordonia* sp. Bu15_44, *Pseudoalteromonas rubra* Bu15_31, *Vibrio coralliilyticus* Bu15_01 and *Pseudoalteromonas* sp. Bu15_09 were tested for the presence of DNA sequences associated with PKS BGCs, i.e. degenerated primer (see Suppl. Table 1) amplifying partial KS fragments were used. Only *Gordonia* sp. Bu15_44 yielded 3 distinct KS fragments.

II.3.5. DISCUSSION

The isolation of *Vibrio* and *Pseudoalteromonas* strains from marine sea slugs seemed likely, since these bacteria are known as abundant species in the marine habitat, and many strains were already isolated by using commercially available complex media with a concentration of ions mimicking seawater. In this screening effort, many *Pseudoalteromonas* strains were isolated and all of them showed at least slight antibacterial activity in the tests used. The genus *Pseudoalteromonas* currently consists of 41 species, thereunder 16 are known as producers of antimicrobials. The various compounds isolated to date were recently reviewed (Offret *et al.*, 2016.), and reported natural products are alkaloids, polyketides and peptides. The bacteria are frequently associated with macroorganisms and therefrom it can be expected that the chemical molecules fulfil an important role as signaling and/or defense molecules. These bacteria are mostly associated with healthy animals and plants, only few reports exist that these strains act opportunistic or pathogenic (Choudhury *et al.*, 2015). Therefore, a positive effect on the implementation of the microbiome for the macroorganism can be expected, making them a key partner in marine holobionts. The latter is composed of the host and its microbiome. A regulation of the microbiome is desirable for invertebrates since they have to protect their epidermis from unwanted settlement with pathogenic biofilm members. The analysis of these talented antimicrobial producers as probiotics in aquaculture and the analysis of produced molecules as potential lead structures for medicinal drugs will be subject of future studies. In contrast to the *Pseudoalteromonaceae*, the dominant marine bacterial family, i.e. *Vibrionaceae*, consists of 126 species and only 6 of them are known to produce antimicrobials (Mansson *et al.*, 2011). In the here described bioprospecting approach several *Vibrio* strains showed no antimicrobial activity, but one strain possessed anti-MRSA activity. This exemplifies that the pangenome of the *Vibrionaceae* indeed carries the potential to synthesize antimicrobials, and that this must be evaluated individually for the respective strains. The bacterial strain showing anti-EHEC activity belongs to the genus *Marinomonas*. Members of this genus were isolated from most diverse habitats throughout the ocean. However, up to now they were not noticeable as prolific producers of natural products. *Marinomonas mediterranea* was described as the producer of marinocine, a broad-spectrum antibacterial protein (Luca-Elio *et al.*, 2005). Further, the natural product Indole-3-carboxaldehyde biosynthesized by *Marinomonas* sp. was recently described as an anti-biofilm agent against *Vibrio cholerae* O1 (Rajalaxmi *et al.*, 2016). The molecule(s) responsible for the here observed activity will be subject of future studies. The efforts to get members of the sponge microbiome growing in vitro is a good example for the plate-anomaly, i.e. the number of bacterial cells which form colonies on an agar plate is much less than bacterial cells present in the sample. Many natural products originally isolated from sponges can be

contributed to bacteria as the real producers (Mehbub *et al.*, 2014). However, despite big efforts of the scientific community, it remains most challenging to bring these bacteria into culture (Waters *et al.*, 2014). This might explain why in the here described research only one strain, i.e. *Providencia vermicola*, was isolated starting with sponge material. Another most interesting fact is that certain bacteria seem to be associated with the sea slug itself as well as the respective egg masses. Since a large portion of bacteria with antimicrobial activity was isolated from egg masses, it can be speculated that this points towards a symbiotic relationship between sea slugs and bioactive compounds producing bacteria. Beside the sea slug, which is constantly producing mucus to clean the epidermal surface also the egg masses must be protected, even though the offspring, with a high content of lipids and proteins, represents a highly desirable food source. A capsule membrane and various mucus layers surround these eggs. However, to ensure the stability and protective properties of the whole egg mass for the time necessary before veliger or juveniles hatch, quick degradation through bacteria and other organisms has to be prevented. Thus, it can be assumed that biologically active compounds play an essential role in egg mass conservation. This was shown for egg masses of *Hexabranhus sanguineus* that is protected by the macrolide Kabiramide C with antifouling properties (Matsunaga *et al.*, 1986). In the meantime, more Kabiramides, up to Kabiramide L were identified (Sirirak *et al.*, 2013). The effect of such natural antifouling agents towards the protection of egg masses is hardly investigated yet, but the finding of bacteria in the distal part of the nidamental glands in the tropical sea slug *Dendrodoris nigra* is highly interesting. *D. nigra* lives in the upper sublittoral of coral reefs and could provide its eggs with bacterial metabolites before the egg mass is subsequently released from the oviduct (Brodie *et al.*, 1999). Further, bacteria inhabiting the surface of the mucus layers of *Siphonaria* egg masses were identified which inhibit growth of *Vibrio harveyi* (Peters *et al.*, 2011.). Investigation of the interrelation between associated bacteria and desired protection of egg masses opens up a topic for future research. To tackle the threat by antibiotic-resistant bacteria, new treatment options and new bioactive compounds must be identified. The last decades showed that the target-based screening approaches were not as successful as expected. Instead, the “classical” bioactivity-based approach of identifying new sources and the isolation of natural products with the targeted activity is still most rewarding. This is proven by the fact that about 80% of all antimicrobial drugs in current use are natural products or are based on their structures (Newman and Cragg 2012). To get a hand on novel compounds, so far underinvestigated biological niches are now explored. Bioprospecting anaerobic bacteria (Akita *et al.*, 2015), resulted in the isolation of a novel producer strain, enabling the description of the highly active Clostrubin (Pidot *et al.*, 2014). In addition, methods were developed to enable *in-vitro* growth of so far unculturable bacteria. This approach resulted in the identification

of Teixobactin, an antibiotic without detectable resistance (Ling *et al.*, 2015). Another approach is reinvestigation of known antibiotic compounds, which were not developed further in the Golden Age of antibiotic research where enough treatment options had been available. A promising example is Griselimycin that is reinvestigated with the goal to develop it into a clinical drug (Kling *et al.*, 2015). Hence, it became clear that nature is still the most promising resource for novel compounds with antibiotic activity. In this context the presence of PKS- and NRPS-coding sequences in the isolated strains provides further evidence that biosynthesis of specialized metabolites with interesting biological activities is encoded in these bacteria. Therefore, the potential of the sea slug associated microbiome for production of natural products harboring the inherent potential to be developed into drug leads must be judged as very high.

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

III.1. Introduction

Pseudouridimycin (PUM, Fig.: 6) is a novel bacterial RNA polymerase (RNAP) inhibitor that was discovered and characterised by Maffioli *et al.* in 2017 and its biosynthesis was elucidated by Sosio *et al.*, 2018. In contrast to other characterised RNAP inhibitors like Rifamycins, Lipiarmycins and the α -pyrone antibiotics Corallopyronine and Myxopyronine which are blocking transcription by binding to RNAP, PUM competes with uracil-triphosphate (UTP) for access to the active site cavity of the RNA polymerase. Thereby, blocking the incorporation of (UTP) nucleotides to the growing RNA chain. While Rifamycins, Lipiarmycins and α -pyrones are classical type I polyketides, PUM is not the product of a type I polyketide synthase but a C-nucleoside-analogue, structurally resembling synthetic antiviral reverse transcriptase inhibitors like Fozivudin and Fosaltudin (Fig.: 6). Due to the uracil containing nucleoside moiety, PUM shares similarity to other uracil nucleoside containing natural products like Tunicamycin, Muraymycin, Mureidomycins, Liposidomycins, Pacidamycin, Polyoxins, Nikkomycins (All N – nucleoside analogues) (McCarthy *et al.*, 1985), and Malayamycin (C – nucleoside analogue) (Li *et al.*, 2008), all produced by members of the *Streptomyces* bacterial family and exhibiting antifungal or antibacterial activity.

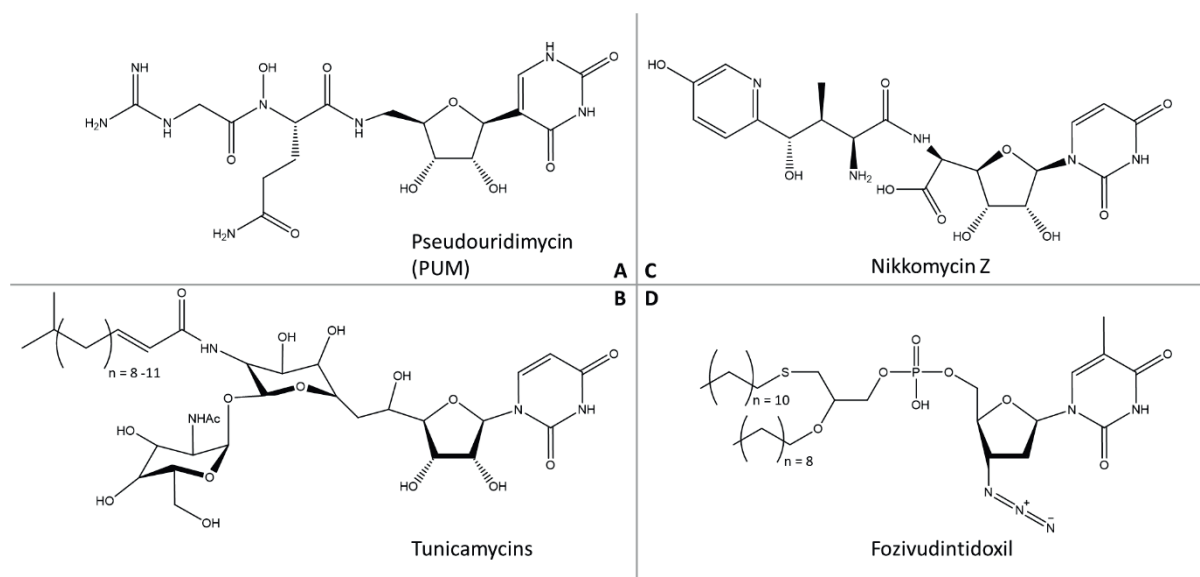


Figure 6: Chemical formulae of natural and non-natural nucleoside analogues with biological activities: A) Pseudouridimycin (PUM) B) Tunicamycins C) Nikkomycin Z D) Fozivudintidoxil.

The previously reported antibacterial Strepturidin, also carrying a pseudouridine unit (Pesic *et al.*, 2014) was recently reinvestigated and its structural assignment was found to be wrong. It could be shown that Strepturidine and Pseudouridimycin indeed are the same compound, based on the identical activities, the presence of PUM in culture extract and on the presence of the PUM biosynthetic gene cluster in the Strepturidin producer strain *Streptomyces albus* DSM40763 (Rosenqvist *et al.*, 2019).

In contrast to the other aforementioned uracil nucleoside containing natural products exhibiting antifungal or antibacterial activities by interfering with the respective cell wall biosynthesis, PUM is a non-incorporated competitor of the RNA exclusive uridine triphosphate (UTP) within the RNAP nucleotide addition site (Maffioli *et al.*, 2017). This makes it the first antibacterial nucleoside analogue targeting the RNA polymerase. Pseudouridimycin is a natural product produced by several members of the *Streptomyces* species with good RNAP inhibitory activity in Gram+ and Gram- bacteria. It exhibits medium to good *in-vitro* antibacterial activity against antibiotic sensitive and (multi-)resistant *Streptococcus pyogenes* and *Staphylococcus aureus* strains, as well as antibiotic sensitive *Moraxella catarrhalis* and *Haemophilus influenzae* strains. Furthermore, it exhibits good *in-vivo* activity against *Streptococcus pyogenes* in a mouse peritonitis model. In contrast, PUM did not exhibit toxicity against human cell lines (Maffioli *et al.*, 2017). Its medium to good activities, low toxicity and at least in antibacterial antibiotics terms novel mode of action make PUM an interesting compound for further drug development and possible lead nomination.

Compared to the N-nucleoside analogs Tunicamycin, Mureidomycin A, Liposidomycin B, Pacidamycin, Nikkomycin Z and Polyoxin, which incorporate uridine, in C-nucleosides like Malayamycin and PUM the ribose sugar unit is connected to uracil at position C - 6, creating the characteristic pseudouridine moiety. Subsequently, nucleosides are further derivatised to the respective natural product by addition of lipidic, glycosidic, peptidic or other moieties. Essential homologous biosynthetic enzymes involved in the formation of the C-nucleoside Malayamycin and PUM are belonging to the TruD-like tRNA pseudouridine synthase family of enzymes. This enzyme family is widespread among the tree of life, with homologs in bacteria, archaea and eukaryota. In primary metabolism, these enzymes are responsible for the modification of position 13 of glutamic acid tRNA into pseudouridine, directly acting on RNA as a substrate (Kaya & Ofengand, 2003). In contrast to TruD-like enzymes from primary metabolism, TruD-like enzymes involved in C-nucleoside analog secondary metabolites biosynthesis are most likely acting on free substrates from the nucleic base metabolism. Thereby, converting the uridine nucleoside into pseudouridine.

The biosynthetic pathway of PUM and a corresponding biosynthetic gene cluster (BGC) was identified and characterised by knock-out (KO) experiments by Sosio *et al.*, 2018. Nine catalytic steps were proposed (Fig.: 7). Thereby, in a first step, uridine nucleotides are converted to Pseudouridine nucleotides and subsequently become dephosphorylated by the catalytic action of the Pseudouridine synthase PumJ and the phosphatase PumD to yield free Pseudouridine (PU). Via intermediate creation of 5'-oxo-PU and subsequent transamination, 5' – amino-PU (APU) is formed by PumI, catalysing oxidation of the 5' hydroxyl group to a keto group and PumG, catalysing transamination at the 5' position with asparagine as NH₃ donor. Then, APU is linked to glutamine by PumK, yielding Gln-APU. In parallel, glycine is converted to guanidinoacetic acid (GAA) by transguanylation (PumN-catalysed), which is subsequently fused to Gln-APU by PumM, resulting in formation of deoxy-PUM. The final biosynthetic step is hydroxylation of the peptide bond nitrogen between Gln and GAA by PumE. The identified BGC consists of 15 genes (*pumA-pumO*). For seven of these genes (i.e. *pumE*, *G*, *I*, *J*, *K*, *M*, *N*) a function could be assigned by experimental data. For five genes (i.e. *pumB*, *C*, *D*, *H*, *L*) a function was proposed based on sequence homology. Hence, each of the nine proposed steps was at least in silico matched with a corresponding enzyme from the PUM BGC [5]. The gene *pumH* that is situated within the BGC was proposed to function as an adenylate kinase, based on sequence homology. However, no biosynthetic step could be readily assigned since it seemed to not play a direct role in PUM biosynthesis. Deleting the gene from the WT producer strain resulted in a significant production drop; however, not in a complete abolition of PUM production. Additionally, one regulator and two transporter genes (i.e. *pumB*, *pumL*) could be readily identified based on homology. In the first reports, no function was assigned to the three genes *pumA*, *pumF* and *pumO* in PUM biosynthesis. However, the *pumF* analogue *mur33* is present in the muraymycin C1 BGC, the latter an uridine nucleoside itself.

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

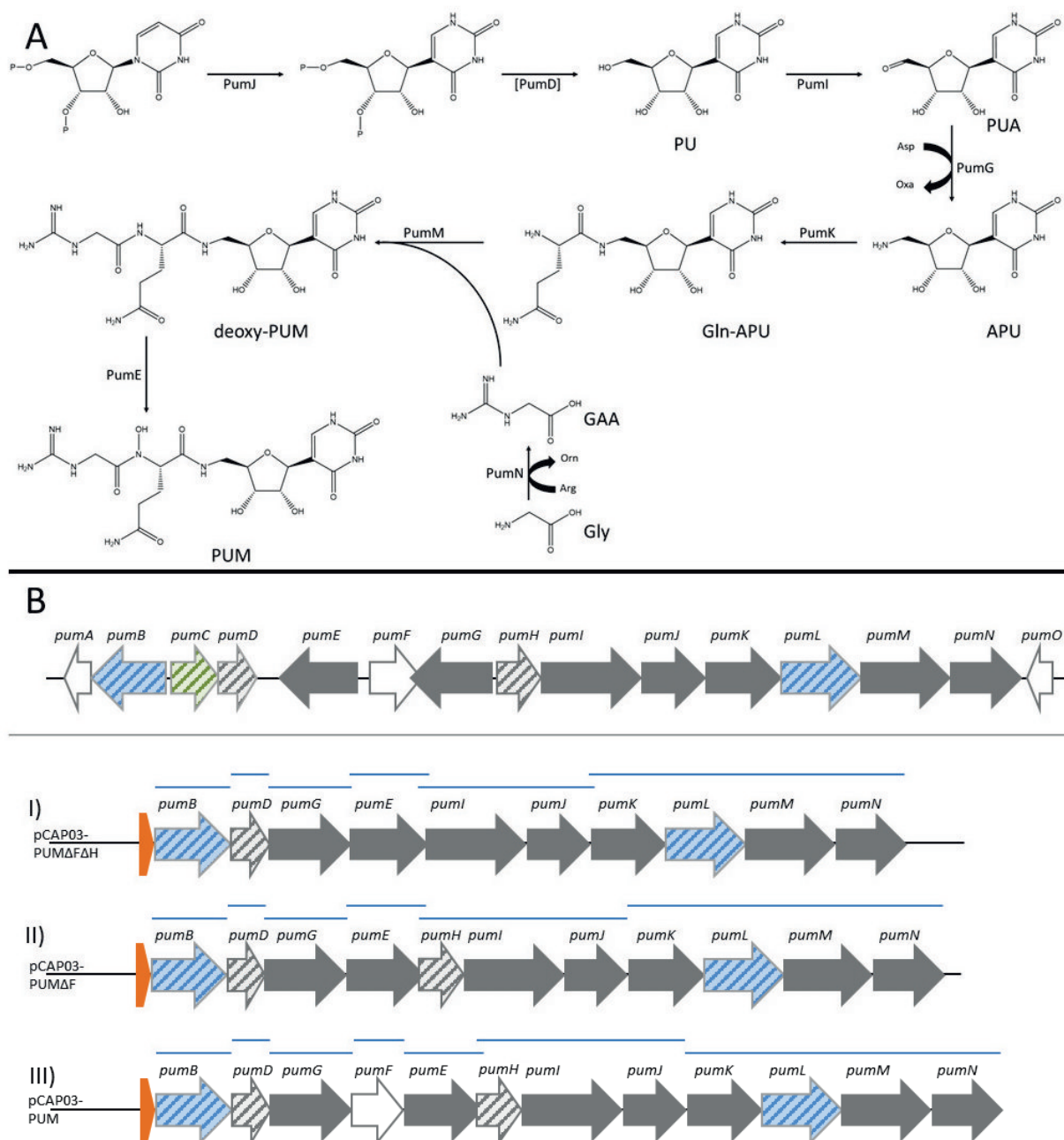


Figure 7: Schematic image of PUM biosynthesis and PUM BGC: A) Biosynthetic pathway of PUM from uridine nucleotides, glutamine, glycine and arginine precursors (Sosio *et al.*, 2018). P = phosphate substituents, PU = pseudouridine, PUA = 5'-oxo-PU, APU = 5'-amino-PU, Gln-APU = glutamine-APU, Gly = glycine, GAA = guanidinoacetic acid B) top: native configuration of the PUM BGC in *Streptomyces* sp. DSM26212, bottom: Rearranged PUM BGCs for heterologous expression in *Streptomyces* hosts. All genes are orientated in the same direction and a promoter was added upstream of the first gene *pumB*. Version I carries the depicted genes, in version II *pumH* was added and in version III *pumF* was added as well. PCR fragments used for assembly are displayed as blue lines. Blue colored genes correspond to transport related genes, green corresponds to regulatory elements and white to genes that were not assigned any function in PUM biosynthesis yet (Sosio *et al.*, 2018). Orange colored elements represent artificial promoters, either constitutive *ermE** or inducible *tcp830*. Genes without experimentally proven function (Sosio *et al.*, 2018) are striped.

Here, the heterologous expression of the PUM BGC in commonly used Streptomyces surrogate hosts is described. Therefore, the BGC was rearranged and several constructs were generated (Fig.: 7) to identify the minimal BGC necessary for heterologous production of the nucleoside antibiotic.

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

Furthermore, it was envisioned to create a PCR and isothermal assembly-based plug and play heterologous expression system to selectively enrich PUM precursors for further studies and semisynthetic derivation of PUM.

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

III.2. Material and Methods

III.2.1. Bacterial strains used in this study

Table 1: Bacterial strains used in this study

Strain	Genotype	Reference
<i>Streptomyces</i> sp. DSM26212	WT	Maffioli <i>et al.</i> , 2017
<i>Streptomyces coelicolor</i> M1146	$\Delta act \Delta red \Delta cpk \Delta cda$	Escribano & Bibb, 2010
<i>Streptomyces lividans</i> TK24	WT	
<i>Streptomyces albidoflavus</i> J1074	WT	
<i>E. coli</i> Top10	F- mcrA $\Delta(mrr-hsdRMS-mcrBC)$ $\phi 80lacZ\Delta M15 \Delta lacX74$ nupG recA1 araD139 $\Delta(ara-leu)7697$ galE15 galK16 rpsL(StrR) endA1 λ -	commercially available (invitrogen)
<i>E. coli</i> ET12567	dam-13::Tn9, dcm-6, hsdM	Flett <i>et al.</i> , 1997
<i>E. coli</i> ET12567 + pUB307	dam-13::Tn9, dcm-6, hsdM + pUB307	Flett <i>et al.</i> , 1997
<i>E. coli</i> BW25113 + pKD46	lacI+rrnBT14 $\Delta lacZ\Delta WJ16$ hsdR514 $\Delta araBADAH33 \Delta rhaBADLD78$ rph-1 $\Delta(araB-D)567 \Delta(rhaD-B)568$ $\Delta lacZ4787(::rrnB-3)$ hsdR514 rph-1 + pKD46	Datsenko & Wanner, 2000
<i>E. coli</i> XL1 Blue + pIJ773	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac + pIJ773	Gust <i>et al.</i> , 2003
<i>E. coli</i> XL1 Blue + pCAP03	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac + pCAP03	Tang <i>et al.</i> , 2015
<i>E. coli</i> XL1 Blue + pGEM-t easy-Apra-ermE*	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac + pGEM-teasy-Apra-ermE*	this work
<i>E. coli</i> XL1 Blue + pGEM-t easy-Apra-tcp830	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac + pGEM-teasy-Apra-tcp830	this work
<i>E. coli</i> Top10 + pCAP03-PUM Δ H Δ F	F- mcrA $\Delta(mrr-hsdRMS-mcrBC)$ $\phi 80lacZ\Delta M15 \Delta lacX74$ nupG recA1 araD139 $\Delta(ara-leu)7697$ galE15 galK16 rpsL(StrR) endA1 λ - + pCAP03-PUM Δ H Δ F	this work
<i>E. coli</i> Top10 + pCAP03-PUM Δ F	F- mcrA $\Delta(mrr-hsdRMS-mcrBC)$ $\phi 80lacZ\Delta M15 \Delta lacX74$ nupG recA1 araD139 $\Delta(ara-leu)7697$ galE15 galK16 rpsL(StrR) endA1 λ - + pCAP03-PUM Δ F	this work
<i>E. coli</i> Top10 + pCAP03-PUM	F- mcrA $\Delta(mrr-hsdRMS-mcrBC)$ $\phi 80lacZ\Delta M15 \Delta lacX74$ nupG recA1 araD139 $\Delta(ara-leu)7697$ galE15 galK16 rpsL(StrR) endA1 λ - + pCAP03-PUM	this work
<i>E. coli</i> BW25113 + pKD46 + pCAP03-PUM Δ H Δ F	lacI+rrnBT14 $\Delta lacZ\Delta WJ16$ hsdR514 $\Delta araBADAH33 \Delta rhaBADLD78$ rph-1 $\Delta(araB-D)567 \Delta(rhaD-B)568$ $\Delta lacZ4787(::rrnB-3)$ hsdR514 rph-1 + pKD46 + pCAP03-PUM Δ H Δ F	this work
<i>E. coli</i> BW25113 + pKD46 + pCAP03-PUM Δ F	lacI+rrnBT14 $\Delta lacZ\Delta WJ16$ hsdR514 $\Delta araBADAH33 \Delta rhaBADLD78$ rph-1 $\Delta(araB-D)567 \Delta(rhaD-B)568$ $\Delta lacZ4787(::rrnB-3)$ hsdR514 rph-1 + pKD46 + pCAP03-PUM Δ F	this work

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

<i>E. coli</i> BW25113 + pKD46 + pCAP03-PUM	lacI+rrnBT14 ΔlacZWI16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pKD46 + pCAP03-PUM	this work
<i>E. coli</i> BW25113 + pCAP03-PUMΔHΔF_ermE*	lacI+rrnBT14 ΔlacZWI16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1- + pCAP03-PUMΔHΔF_ermE*	this work
<i>E. coli</i> BW25113 + pCAP03-PUMΔF_ermE*	lacI+rrnBT14 ΔlacZWI16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1- + pCAP03-PUMΔF_ermE*	this work
<i>E. coli</i> BW25113 + pCAP03-PUM_ermE*	lacI+rrnBT14 ΔlacZWI16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1- + pCAP03-PUM_ermE*	this work
<i>E. coli</i> BW25113 + pCAP03-PUMΔHΔF_tcp830	lacI+rrnBT14 ΔlacZWI16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1- + pCAP03-PUMΔHΔF_tcp830	this work
<i>E. coli</i> BW25113 + pCAP03-PUMΔF_tcp830	lacI+rrnBT14 ΔlacZWI16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1- + pCAP03-PUMΔF_tcp830	this work
<i>E. coli</i> BW25113 + pCAP03-PUM_tcp830	lacI+rrnBT14 ΔlacZWI16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1- + pCAP03-PUM_tcp830	this work
<i>E. coli</i> ET12567 + pCAP03	dam-13::Tn9, dcm-6, hsdM + pCAP03	this work
<i>E. coli</i> ET12567 + pCAP03-PUMΔHΔF_ermE*	dam-13::Tn9, dcm-6, hsdM + pCAP03-PUMΔH_ermE*	this work
<i>E. coli</i> ET12567 + pCAP03-PUMΔF_ermE*	dam-13::Tn9, dcm-6, hsdM + pCAP03-PUMΔF_ermE*	this work
<i>E. coli</i> ET12567 + pCAP03-PUM_ermE*	dam-13::Tn9, dcm-6, hsdM + pCAP03-PUM_ermE*	this work
<i>E. coli</i> ET12567 + pCAP03-PUMΔHΔF_tcp830	dam-13::Tn9, dcm-6, hsdM + pCAP03-PUMΔF_tcp830	this work
<i>E. coli</i> ET12567 + pCAP03-PUMΔF_tcp830	dam-13::Tn9, dcm-6, hsdM + pCAP03-PUMΔF_tcp830	this work
<i>E. coli</i> ET12567 + pCAP03-PUM_tcp830	dam-13::Tn9, dcm-6, hsdM + pCAP03-PUM_tcp830	this work
<i>Streptomyces coelicolor</i> M1146 pCAP03-PUMΔFΔH_ermE*	Δact Δred Δcpk Δcda + pCAP03-ermE*-pumB,D,G,E,I,J,K,L,M,N	this work
<i>Streptomyces coelicolor</i> M1146 pCAP03-PUMΔF_ermE*	Δact Δred Δcpk Δcda + pCAP03-ermE*-pumB,D,G,E,H,I,J,K,L,M,N	this work
<i>Streptomyces coelicolor</i> M1146 pCAP03-PUM_ermE*	Δact Δred Δcpk Δcda + pCAP03-ermE*-pumB,D,G,E,F,H,I,J,K,L,M,N	this work
<i>Streptomyces coelicolor</i> M1146 pCAP03-PUMΔFΔH_tcp830	Δact Δred Δcpk Δcda + pCAP03-tcp830-pumB,D,G,E,I,J,K,L,M,N	this work
<i>Streptomyces coelicolor</i> M1146	Δact Δred Δcpk Δcda +	this work

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

pCAP03-PUMΔF_tcp830	pCAP03-tcp830- pumB,D,G,E,H,I,J,K,L,M,N	
<i>Streptomyces coelicolor</i> M1146 pCAP03-PUM_tcp830	Δact Δred Δcpk Δcda + pCAP03-tcp830- pumB,D,G,E,F,H,I,J,K,L,M,N	this work
<i>Streptomyces lividans</i> TK24 pCAP03	WT + pCAP03	this work
<i>Streptomyces lividans</i> TK24 pCAP03- PUMΔF_ermE*	WT + pCAP03- PUMΔF_ermE*	this work
<i>Streptomyces lividans</i> TK24 pCAP03- PUMΔF_tcp830	WT + pCAP03- PUMΔF_tcp830	this work
<i>Streptomyces albidoflavus</i> J1074 + pCAP03	WT + pCAP03	this work
<i>Streptomyces albidoflavus</i> J1074 + pCAP03- PUMΔF_ermE*	WT + pCAP03- PUMΔF_ermE*	this work

III.2.2. Media and additives used in this study

Lysogenic broth (LB)

Peptone from casein	10 g/L
Yeast extract	5 g/L
NaCl	5 g/L
(Agar-agar	20 g/L)
p.H.	7.0

Super optimal broth without MgCl₂

Tryptone	20 g/L
Yeast extract	5 g/L
NaCl	0.5 g/L
KCl	0.2 g/L
p.H.	7.0

Mannitol soy (MS) agar

Soybean flour	20 g/L
D-mannitol	20 g/L
Agar-agar	20 g/L
p.H.	7.5

International Streptomyces project medium 2 (ISP2)

Malt extract	10 g/L
Yeast extract	4 g/L
D-glucose	4 g/L
(Agar agar	20 g/L)
p.H.	7.5

2xYT

Tryptone	16 g/L
Yeast extract	10 g/L
NaCl	5 g/L
p.H.	7.0

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

Tryptic soy broth (TSB)

Tryptone	17 g/L
Peptone from soy	3 g/L
D-glucose	2.5 g/L
NaCl	5 g/L
(Agar agar	20 g/L)
K ₂ HPO ₄	2.5 g/L
p.H.	7.2

Glucose and K₂HPO₄ were autoclaved separately and added after cooling the medium to room temperature

Production medium (PM)

D-glucose	10 g/L
Dextrin from corn	24 g/L
Peptone from soy	8 g/L
Yeast extract	5 g/L
NaCl	1 g/L
p.H.	7.2

Stock solutions

Kanamycin (Kan)	50 g/L
Apramycin (Apra)	50 g/L
Carbenicillin (Carb)	50 g/L
Chloramphenicol in EtOH (Ca)	30 g/L
Nalidixic acid (Nal)	30 g/L
Anhydrotetracycline in DMSO	10 g/L
L-arabinose (Ara)	1 M
MgCl ₂	2 M

Media were prepared using deionised H₂O and autoclaved at 121 °C for 20 min and 2 bar vapor pressure. For sterilisation of ISP2 and PM, D-glucose and Dextrin from corn were autoclaved separately from the other components in smaller volume to avoid Maillard reactions. Solid media were allowed to cool down to ~50 °C before adding selective antibiotics dilution or other additives, poured into sterile petri dishes and left open to solidify and dry. Liquid and solid media were stored at room temperature, stock solutions were sterile filtered (aside from Chloramphenicol) and stored at -20 °C.

III.2.3. Cultivation of bacteria

E. coli for cloning purposes were grown in LB broth or agar supplemented with appropriate antibiotics in 1:1000 dilution from the stock solutions at 37° with exception of *E. coli* BW25113 + pKD46 which was grown in SOB-medium at 30 °C. Cell growth was monitored by measuring the OD₆₀₀ using an Eppendorf BioSpectrometer in single use cuvettes with 10 mm light path.

E. coli precultures were grown in 5 mL LB in reaction tubes supplemented with appropriate antibiotics. Larger volume cultivations took place in 100 mL or 300 mL Erlenmeyer flasks. *E. coli* cryocultures were prepared by mixing 1 mL of an over-night culture of *E. coli* with sterile glycerol and storing them in -80 °C

Streptomycetal strains were cultivated and selected on MS agar plates containing appropriate antibiotics. For isolation of genomic DNA and as preculture, *Streptomyces* were grown in ISP2 for 2-3 days. *Streptomyces* were stored as spore suspension in 20% glycerol in -80. Streptomycetal spores were collected from MS agar plates as described in Practical *Streptomyces* Engineering (John Innes Foundation, 2000).

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

III.2.4. Primers used in this study

Table 2: Primers used in this study

Name	Sequence (5' -> 3')
plJ773cass_f	ATTCCGGGGATCCGTCGACC
ermEp1	CCTCCCACCCGCTGGATCCTACCAACCGGCACGATTGTCCAGCCCACAACAGCATCGCGGTGCCACGT GTGGACCGCGTCGGTCAGATCCTCCCCGCA
ermEp2	TGCTGTTGTGGGCACAATCGTGCCGGTTGGTAGGATCCAGCGgtaggagg
tcp830	CCTCCCAGATCTCTATCACTGATAGGGATCCTACCACTATCAATGATAGAGTAGCCAACAGCTGTAGG CTGGAGCTGCTTC
pumB_f	CATGGTATAAATAGTGGCGTGGATAGATACCTACGAGC
pumB_r	GTGGTCACCTCCTCATAACCGCTCCGTCTCCA
pumD_f	CGGTATGAGGAGGGTGACCACGTGACGGGCACC
pumD_r	CGATCATGCAGCCGCCTCCTCAGGTCCATTGACTGAGAG
pumG_f	GGAGGCGGCCTGCATGATCGCGGCATGTCGCT
pumG_r	GATGCCTCCACGTTTATCAGTCACAGGTCCGCAAGAGCCT
pumG+F_r	TGTAAAGGCCTCCGTCACAGGTCCGCAAGAGCCT
pumE_f	CTGATAAACGTGGAGGCATCATGTCGATTTCCCTGTCGTT
pumE_r	CAGAACTCCCTCCTCACTCGGATCCGTCCCGGC
pumEdH_r	GACGCATCGGTACTCCCTCCTCACTCGGATCCGTCCCGGC
pumF_f	CCTGTGACGGAGGCCTTTACAGTGTGGAACGTC
pumF_r	GATGCCTCCACGTTTATCAGTCATCGACCGCTCCGGGATC
pumI_f	GGAGGGAGTACCGATGCGTCAGGGCTTCGATGA
pumH_f	CCGAGTGAGGAGGGAGTTCTGGTGATCATTGAGGGC
pumJ_r	GGACAGCCTCCCGGGGACGCTCAGCGGCGGAGGACCAACT
pumK_f	GCGTCCCCGGGAGGCTGTCCATGGCGTTGCTGCTGCTCAA
pumN_r	TATGTAGCTTTTCGACATATTAGGCCAACGGCCGGTAAC
pCAP03_1f	TGTCGAAAGCTACATATAAG
pCAP03_1r	AACTGTTCCGCCAGGCTCAAG
pCAP03_2f	CTTGAGCCTGGCGAACAGTT
pCAP03_2r	GCCACTATTTATACCATGGG
Rec_uni_f	CCCTGTCGCCCTCCTATTGGCTTCCGATTATCTTCTTGGCAGCTCACGGTAACTGATG
PUM_ermE*_rec	GCCAGTCGGCCGGCTCGTAGGTATCTATCCACGCCACTATCCTCCTACCCGCTGGATCCT
PUM_tcp830_rec	GCCAGTCGGCCGGCTCGTAGGTATCTATCCACGCCACTATCCTCCCAGATCTCTATCACT
Pum_screen	GCATGGGCGTTCATGCTGAC
ermE*_rectest	ATCTTGACGGCTGGCGAGAG
tcp830_rectest	CAGCTGTTGGCTACTCTATC

Primers for isothermal assembly of the genes were created by outfitting the respective binding region with an artificial GGAGG ribosomal binding site (RBS) 8 random nucleotides upstream the start codon. Additionally, primers were outfitted with ~20 bp (melting temperature of the

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

overlapping region >50 °C) overlap to the respective adjacent fragment. Primers were purchased from Eurofins Genomics (Ebersberg, Germany).

III.2.5. Molecular biology techniques

Plasmid DNA was isolated using the innuPREP plasmid mini kit 2.0 (AnalytikJena, Jena, Germany) according to protocol. Genomic DNA was extracted using the innuPREP bacteriaDNA kit (AnalytikJena, Jena, Germany). PCR amplification for cloning purposes was performed using Q5 DNA polymerase (NEB Biolabs, New Brunswick, USA) using the following setup:

Table 3: Q5-Polymerase PCR setup. GC-Enhancer was added if necessary, the annealing temperature was set to 50°, 60° or 70° and the Elongation time was set calculated to match the fragment size using 45 s for amplification of 1000 nt.

Ingredient	Volume [μl]	Step	Temperature [°C]	Time [s]	
Q5-Buffer	5	1	init. Denaturation	98	60
GC-Enhancer	(5)	2	Denaturation	98	10
dNTPs	0.5	3	Annealing	var	20
primer_f (20 pM)	0.625	4	Elongation	72	var
primer_r (20 pM)	0.625	5	final Elongation	72	600
template DNA	1				
Q5-Polymerase	0.15				
H2O	ad 25				

Test-PCRs were performed using GoTAQ (Promega, Madison, USA) using the following setup:

Table 4: GoTAQ polymerase PCR setup. For colony test PCRs the reaction was filled with H₂O and a small portion of the tested colony was dispersed in the PCR mixture. Elongation time was calculated with 60 s for amplification of 1000 nt.

Ingredient	Volume [μl]	Step	Temperature [°C]	Time [s]	
Green GoTAQ buffer	4	1	init. Denaturation	98	60
MgCl ₂	1	2	Denaturation	98	10
DMSO	1	3	Annealing	var	20
dNTPs	0.33	4	Elongation	72	var
primer_f (20 pM)	0.33	5	final Elongation	72	600
primer_r (20 pM)	0.33				
template DNA	0.5				
GoTAQ	0.1				
H2O	ad 20				

Restriction was performed using standard techniques and NEB enzymes, DNA fragments were analysed on 1% or 2% TAE-Agarose gels using MidoriGreen as loading dye and stain and a uv-free

blue light-transilluminator. GeneRuler 1kb Plus was used as marker. DNA for cloning purposes was excised from the gel and purified using the Zymoclean large fragment DNA recovery kit according to the manufacturers instruction. DNA concentrations were determined with an Eppendorf BioSpectrometer using a 1 mm light path.

E. coli were transformed with DNA using standard electroporation methodology; In brief: *E. coli* cells were grown in 50 mL LB at 37 °C or 50 mL SOB at 30 °C for several hours to an OD₆₀₀ of ~0.5. Cells were harvested by centrifugation at 10000 rcf for 3 min and washed twice with ice cold 10% glycerol. After the final washing step, the resulting pellet was resuspended in the return flow and aliquoted to 50 µL. DNA was mixed with 50 µL of competent cells, added to ice cold 2 mm electroporation cuvettes and an electric pulse of 2.5 kV was applied for ~ 5 s using a BioRad MicroPulser. The electroporation mixture was taken up in 800µL LB medium and incubated at 37 °C for 1h. Subsequently, transformants were plated on LB plates containing appropriate antibiotics.

III.2.6. Construction of pGEM-teasy_Apra-ermE* and pGEM-teasy_Apra-tcp830

Promotor sequences of ermE* and tcp830 were obtained from Siegl *et al.*, 2013 and Dangel *et al.*, 2010 and ordered as primers with overlaps to the Apramycin antibiotic resistance cassette of pIJ773 and an artificial RBS behind the promotor region. pIJ773 was isolated and the promoters ermE* and tcp830 were fused to the Apra antibiotic resistance cassette by PCR using Q5 as polymerase, pIJ773 as template, pIJ773cass_f as a forward primer and tcp830, ermEp1 and ermEp2 as reverse primers. Due to the size of the ermE* promotor, amplification had to take place in two rounds amplification, fusing the first segment of the promotor by using ermEp1 first and subsequently using the resulting product as template for the second amplification with ermEp2 as reverse primer. Final constructs were outfitted with an A overhang by incubating the Q5 PCR fragments with GoTaq polymerase using the standard reaction mixture without primers at 72° for 1 h. Final products were gel purified and ligated into pGEM-t easy using pGEM-t easy vector systems (Promega) according to the manufacturers instructions and introduced to *E. coli* XL1 blue by electroporation. Correct fusion of the promoters was corroborated by sequencing (Seqlab,).

III.2.7. Construction of the integrative *Streptomyces* PUM expression plasmids pCAP03-PUM Δ F Δ H_ermE*/tcp830, pCAP03-PUM Δ F_ermE*/tcp830 and pCAP03-PUM_ermE*/tcp830

Streptomyces sp. DSM26212 was grown in 30 ml ISP2 medium for 3 days and genomic DNA was extracted, *E. coli* Top10 + pCAP03 was grown in 5ml LB_{kan} over-night and the plasmid was isolated. The PUM BGC was amplified in several parts with primer pairs pumB_f/r, pumD_f/r, pumG_f/r, pumE_f/pumEdH_r, pumI_f/pumJ_r and pumK_f/pumN_r for pCAP03-PUM Δ F Δ H, pumB_f/r, pumD_f/r, pumG_f/r, pumE_f/r, pumH_f/pumJ_r and pumK_f/pumN_r for pCAP03-PUM Δ F and pumB_f/r, pumD_f/r, pumG_f/pumG+F_r, pumE_f/pumE_r, pumF_f/r, pumH_f/pumJ_r and pumK_f/pumN_r for pCAP03-PUM respectively, using isolated genomic DNA as a template. pCAP03 was amplified in two parts using the primer pairs pCAP03_1f/r and pCAP03_2f/r and linearised (*Nde*I, *Xho*I) pCAP03 as a template. DNA Fragments were gel purified and fused by isothermal assembly (Gibson *et al.*, 2009) using self-made isothermal assembly master mix. In brief, 15 μ L of isothermal assembly master mix were thawed on ice and equimolar concentration of all fragments were mixed in 5 μ L and added to the master mix. The reaction mixture was incubated at 50 °C for 1 h and subsequently dialysed by drop dialysis using MF-Millipore VSWP membranes. Subsequently, *E. coli* Top 10 electrocompetent cells were transformed with 2 μ L of dialysed reaction mixture and correct assembly of pCAP03-PUM Δ F Δ H, pCAP03-PUM Δ F and pCAP03-PUM was corroborated by restriction analysis using double digests with *Hind*III/*Not*I and *Hind*III/*Nco*I respectively.

E. coli BW25113 + pKD46 was transformed with pCAP03-PUM Δ F Δ H, pCAP03-PUM Δ F and pCAP03-PUM and selected on LB_{Carb/Kan} agar. Apramycin resistance cassette/promotor fusions were amplified with the primers Rec_uni_f as forward primer and PUM_ermE*_rec and PUM_tcp830_rec as reverse primer respectively, adding 40 nt homologous overhangs for introduction of promoters into the construct. The PCR reaction was directly purified from the mixture using the ZymoResearch large fragment DNA recovery kit and eluted in 10 μ L of H₂O. *E. coli* BW25113 + pKD46 + pCAP03-PUM was grown in 20 mL SOB supplemented with Carb and Kan for maintenance of the plasmids and 0.1 % Arabinose for induction of the λ RED genes at 30 °C to an OD₆₀₀ of ~ 0.5. Cells were harvested, washed twice with ice cold 10 % glycerol and resuspended in the return flow. Subsequently, cells were transformed with 4 μ L of the previously recovered PCR fragment, taken up in 800 μ L of SOB and left to recover at 37 °C. After recovery, cells were plated on LB_{Apra} and incubated over night at 37 °C to induce loss of the plasmid pKD46. The correct integration was corroborated by colony PCR using GoTaq and the primers Pum_screen, binding to *pumB* and ermE*_retest and tcp830_retest respectively, binding to the integrated promotor region one side and to *pumB* on the other.

III.2.8. Conjugation of PUM expression plasmids *S. coelicolor* M1146, *S. lividans* TK24 and *S. albidoflavus* J1074

Conjugation to *S. coelicolor* M1146, *S. lividans* TK24 and *S. albidoflavus* J1074 was carried out by triparental conjugation as described by Gust *et al.*, 2003. In brief, methylation deficient *E. coli* ET12567 was transformed with pCAP03, pCAP03-PUM Δ F Δ H_ermE*/tcp830, pCAP03-PUM Δ F_ermE*/tcp830 and pCAP03-PUM_ermE*/tcp830 respectively and selected with Apramycin and Chloramphenicol. 20ml cultures of the freshly transformed *E. coli* ET12567 and *E. coli* ET12567 carrying the conjugation helper plasmid pUB307 were grown from over-night precultures cultures in LB_{Ca/Apra} and LB_{Ca/Kan} respectively to an OD₆₀₀ of ~ 0.6. All four cultures were washed twice with LB medium and resuspended in the return flow. 20 μ L of Streptomyces spores were introduced to 200 μ L 2xYT medium, heat-shocked for 10 min at 50° C and left to cool down. Subsequently, 100 μ L of the washed *E. coli* ET12567 + pUB307 and *E. coli* ET12567 carrying pCAP03, pCAP03-PUM Δ F Δ H_ermE*/tcp830, pCAP03-PUM Δ F_ermE*/tcp830 and pCAP03-PUM_ermE*/tcp830 respectively were added to the Streptomyces spores and the mixtures were plated on MS plates supplemented with 10 mM MgCl₂. The plates were incubated at 30° C over night, then overlaid with 0.5 mL H₂O containing 1.2 mg Apramycin and 0.5 mg Nalidixic acid and incubated at 30° C for ~ 3 days more. Candidate transconjugands were transferred to new MS agar plates containing apramycin, kanamycin and nalidixic acid and the successful integration of the PUM expression plasmid was corroborated by PCR using GoTaq and the primers Pum_screen, binding to pumB and ermE*_rectest and tcp830_rectest respectively. Correct integration of the empty pCAP03 was inferred from the antibiotic resistance of transconjugands.

III.2.9. Enrichment of PUM from *Streptomyces* sp. DSM26212 culture broth

Streptomyces sp. DSM26212 grown in 2 L PM for 5 days was lyophilised and redissolved in 50 % of the culture volume in H₂O. The solution was filtered through paper filters and the solution was applied to a strong cation exchange column (Dowex W50x2) previously washed with 5 % HCl and equilibrated with H₂O. The column was washed with 5 – 10 bed volumes 20 mM NaAc p.H. 6 and p.H. 7. Subsequently, semi pure PUM was eluted with 100 mM NH₄Ac p.H. 9. Fractions were lyophilised and analysed on LC-MS/MS and MALDI-MS.

III.2.10. Heterologous expression of the PUM BGC

Precultures of transgenic *Streptomyces* carrying PUM expression plasmids pCAP03-PUM Δ FAH_ermE*/tcp830, pCAP03-PUM Δ F_ermE*/tcp830 and pCAP03-PUM_ermE*/tcp830 respectively as well as empty vector negative control and WT producer *Streptomyces* sp. DSM26212 were grown in Erlenmeyer flasks containing 20 mL ISP2 medium in the absence of antibiotics for two days at 200rpm and 30 °C. For analysis of PUM production by MALDI-MS measurements, MS, ISP2 and TSB plates were inoculated from *S. coelicolor* M1146 + pCAP03, *S. coelicolor* M1146 + pCAP03-PUM Δ FAH_ermE* and *Streptomyces* sp. DSM26212 precultures. *S. coelicolor* M1146 + pCAP03-PUM Δ FAH_tcp830 was not inoculated due to the inability to induce the expression in the grown culture on plate after a given time. For analysis of PUM production in liquid medium, 100 mL cultures in ISP2 and TSB were inoculated with 1 % v/v from the precultures and grown for 5 days in Erlenmeyer flasks at 30 °C and 200 rpm. For the induction of the tcp830 promotor, anhydrotetracycline was added to 2 mg/L after one days. The cultures were grown for 3-5 days and PUM production was analysed by LC-MS/MS directly from the culture. For *S. coelicolor* M1146 + pCAP03-PUM_ermE*, cultivation was extended to 20 days.

III.2.11. LC/MS analysis of PUM

For the analysis of PUM, a microTOFq II (Bruker) ESI-qTOF-HRMS mass spectrometer (Bruker Daltonics, Bremen, Germany) coupled to an Agilent 1290 UPLC system with an Acquity UPLC BEH C18 1.7 μ m (2.1x100 mm) and Acquity UPLC BEH C18 1.7 μ m VanGuard Pre-Column (2.1x5 mm) column setup was used. The UPLC system was run using a gradient (A:H₂O, 0.1% FA; B: MeCN, 0.1% FA; Flow: 600 μ L/min): 0 min: 95%A; 0.30 min: 95%A; 18.00 min: 4.75%A; 18.10 min: 0%A; 22.50 min: 0%A; 22.60 min: 95%A; 25.00 min: 95%A, the column oven was set to 45°C and injection volume was set to 5 μ L. Data were analysed using the DataAnalysis software package from Bruker. Samples were prepared by clearing the supernatant by centrifugation at 20000 rcf and directly applying the supernatant to the LCMS.

III.2.12. Matrix assisted laser desorption/ionisation mass spectrometry (MALDI-MS)

MALDI-MS measurements of *S. coelicolor* M1146 + pCAP03-PUM Δ H_ermE* was performed by growing the strain with the WT producer and empty vector negative control on MS, ISP2 and TSB agar plates. Measurements were conducted as described in Kompauer *et al.*, 2017. In brief, samples were prepared by transferring grown cell material to a stainless steel sample carrier and overlaying

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

said material with 30 mg/mL 2,5-dihydroxybenzoic acid as matrix. Subsequently material was ionised with an “AP-SMALDI5 AF” ion source (TransMIT GmbH, Giessen, Germany) coupled to a Fourier transform orbital trapping mass spectrometer (Q Exactive HF, Thermo Scientific GmbH, Bremen, Germany). Data were converted to .mzXML using the Xcalibur MS software package (Thermo Scientific GmbH, Bremen, Germany) and analysed with mzMINE.

III.3. Results

III.3.1. Construction of pGEM-teasy_Apra-ermE* and pGEM-teasy_Apra-tcp830

In order to integrate promoters to the final construct by λ RED mediated recombination, first step was the creation of *ready-to-use* promotor/resistance-cassette fusions. These are outfitted with ~40 nt homology arms by PCR to determine the insertion site and subsequently can be used for recombination.

After successful fusion of the synthesised promoters ermE* and tcp830 to the pIJ773 *aac(3)* – gene (Apramycin resistance) and subsequent ligation into pGEMT easy, the vectors were digested with *EcoRI* and analysed on agarose gels. All test restrictions showed the expected bands with the vector backbone being ~ 3 kB in size and the resistance cassette-promoter fusions being ~ 1.5 kB (Fig.: 8). Sequencing of the plasmids using standard T7 and SP6 sequencing primers corroborated the successful creation of the promoter-resistance cassette fusion.

III.3.2. Cloning of the PUM BGC

To rearrange the individual genes of the BGC into the orders given in Fig.: 7, the individual fragments of the PUM BGC were amplified together with the two fragments of the pCAP03 vector and subsequently fused by isothermal assembly. The resulting fragments all matched the expected sizes: For pCAP03-PUM Δ F Δ H, pumB_f/r 1222 nt, pumD_f/r 644 nt, pumG_f/r 1356 nt, pumE_f/pumEdH_r 1297 nt, pumI_f/pumJ_r 2666 nt, pumK_f/pumN_r 5101 nt, pCAP03_1f/r 5574 nt and pCAP03_2f/r 4955 nt, for pCAP03-PUM Δ F pumB_f/r 1222 nt, pumD_f/r 644 nt, pumG_f/r 1356 nt, pumE_f/r 1290 nt, pumH_f/pumJ_r 3356 nt, pumK_f/pumN_r 5101 nt, pCAP03_1f/r 5574 nt and pCAP03_2f/r 4955 nt and for pCAP03-PUM pumB_f/r 1222 nt, pumD_f/r 644 nt, pumG_f/pumG+F_r, pumE+F_f/pumE_r, pumF_f/r, pumH_f/pumJ_r 3356 nt and pumK_f/pumN_r 5101 nt, pCAP03_1f/r 5574 nt and pCAP03_2f/r 4955 nt. After isothermal assembly of all fragments, colonies were grown and respective plasmids were tested by restriction with *HindIII/NotI* and *HindIII/NcoI*. For pCAP03-PUM Δ F Δ H, 2 out of 10 colonies, for pCAP03-PUM Δ F 6 out of 9 colonies and for pCAP03-PUM Δ F 1 out of 4 colonies showed the expected restriction pattern for both test restrictions based on *in-silico* simulation. The cloning process is exemplified for pCAP03-PUM Δ F and its promotor derivatives in Fig.: 8. Plasmid maps for pCAP03-PUM Δ H and pCAP03-PUM and test restriction for pCAP03-PUM Δ H are given in Suppl.Fig.: 1 - 3.

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

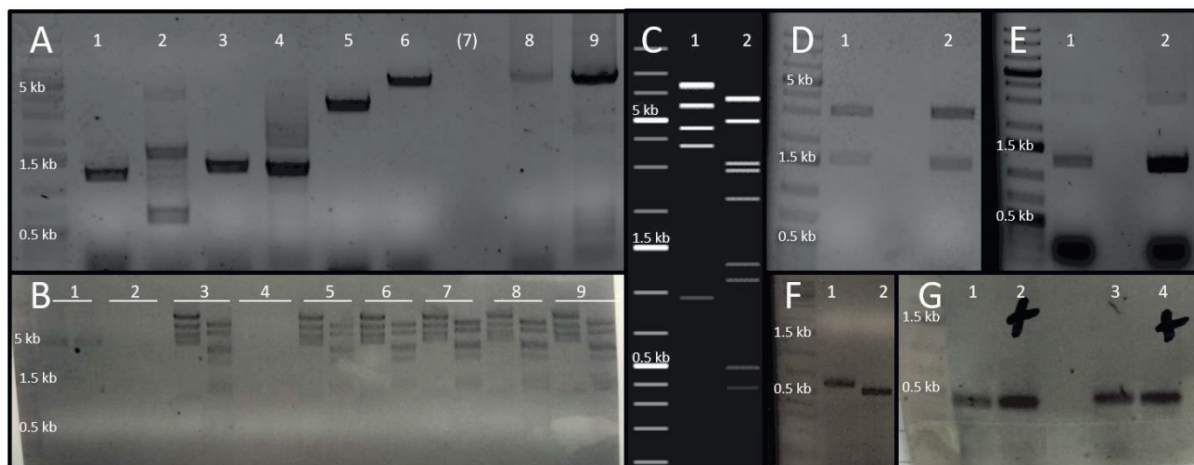


Figure 8: Cloning procedure for construction of pCAP03-PUM Δ F Δ H, pCAP03-PUM Δ F and pCAP03-PUM promoter derivatives and subsequent conjugation to *Streptomyces* heterologous hosts exemplified by pCAP03-PUM Δ F: A) PCRs for *pumB* (1), *pumD* (2), *pumG* (3), *pumE* (4), *pumH* - *pumI* (5), *pumK* - *pumL* (6), pCAP03_part1 (8) and pCAP03_part1 (9); B) Test restrictions of 9 colonies of assembled pCAP03-PUM with *HindIII/NotI* (left) and *HindIII/NcoI* (right); C) *in-silico* simulation of the restriction pattern of pCAP03-PUM with *HindIII/NotI* (1) and *HindIII/NcoI* (2); D) Restriction of pGEM-teasy_Apra-ermE* (1) and pGEM-teasy_Apra-tcp830 (2) with *EcoRI*; E) PCRs for the Apramycin resistance cassette – promoter fusions Apra-ermE* (1) and Apra-tcp830 (2); F) Test PCRs for corroboration of correct integration of Apra-ermE* (1) and Apra-tcp830 (2) into pCAP03-PUM; G) Test PCRs to verify the successful conjugation of pCAP03-PUM-tcp830 (1) and pCAP03-PUM-ermE* (3) and with respective positive controls (2,4) exemplified for *S. coelicolor* M1146.

Outfitting the promoter-resistance cassette fusion with overhangs matching pCAP03-PUM Δ F Δ H, pCAP03-PUM Δ F and pCAP03-PUM by PCR resulted in well visible bands of correct size (ermE* 1186 nt, tcp830 1118 nt) respectively. After λ RED mediated introduction of the promoters, 5 colonies each growing on LB_{Apra}, for selection and enrichment of modified pCAP03-PUM Δ F Δ H, pCAP03-PUM Δ F and pCAP03-PUM vectors, were randomly picked and screened for correct integration of the promoters into pCAP03-PUM Δ H and pCAP03-PUM by colony PCR. All screened colonies showed the expected bands of ~ 500 nt (exemplified in Fig. 8).

While pCAP03-PUM Δ F Δ H and its promoter derivatives were only conjugated to *S. coelicolor* M1146 in a first experiment, pCAP03-PUM Δ F and pCAP03-PUM and its promoter derivatives were conjugated to the three commonly used Streptomyces heterologous hosts *S. coelicolor* M1146, *S. lividans* TK24 and *S. albidoflavus* J1074. Conjugation of pCAP03-PUM Δ F_tcp830 to *S. albidoflavus* J1074 failed by not yielding a single colony on selective antibiotic containing MS agar and conjugation of pCAP03-PUM was only successful for *S. coelicolor* M1146. Colonies were picked over to new plates. Integration of pCAP03-PUM Δ F Δ H, pCAP03-PUM Δ F and pCAP03-PUM derivatives into the respective surrogate hosts genome was inferred from growth on selective antibiotics Apramycin and Kanamycin. Precultures were inoculated from plates, DNA from 0.5 mL of grown preculture was isolated and correct integration of the expression plasmids was corroborated by PCR using isolated plasmid from the previous step as positive control (exemplified in Fig.: 8). Conjugation of pCAP03-

PUMΔF_tcp830 to *S. albidoflavus* J1074 and pCAP03-PUM to *S. lividans* TK24 and *S. albidoflavus* J1074 was not repeated.

III.3.3. PUM production in *Streptomyces* sp. DSM26212

The nucleoside-analog inhibitor PUM represents a highly hydrophilic compound. First, the experimental setup for detection of PUM was tested. Therefore, the reported wild type producer strain *Streptomyces* sp. DSM26212 was cultivated in production medium (PM) medium for 5 days. The fermentation broth was cleared by centrifugation and filtration and PUM was enriched by strong cation exchange chromatography as previously described (Maffioli *et al.*, 2017). Subsequent high resolution LCMS analysis revealed the presence of PUM as $[M+H]^+$ ion in the enriched fraction (Fig.: 9). The identity of this ion could be corroborated by MS/MS fragmentation pattern analysis. A plausible decay mechanism yielding the observed fragments was proposed and fragments observed in the experiment could be matched.

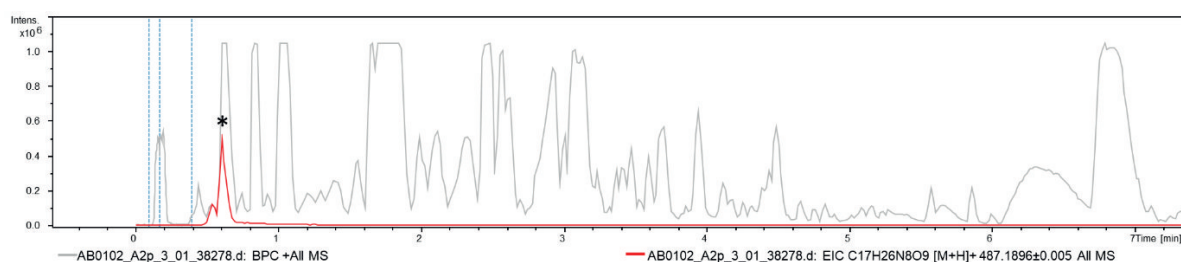


Figure 9: High resolution LCMS measurement of a PUM-enriched fraction from *Streptomyces* sp. DSM26212. Grey: Base peak chromatogram (BPC), Red: Extracted ion chromatogram (EIC) of $C_{17}H_{26}N_8O_9$ $[M+H]^+$ ± 0.0005 . The PUM peak in the EIC is marked with an asterisk (*) (Image created with help of Dr. M. Patras).

Identity of the molecule was corroborated by MS/MS analysis of the parent ion showing 14 prominent fragments with 388.1463 (OH-Gln-APU), 371.1197 (OH-Gln-APU -NH₃), 370.1357 (OH-Gln-APU -H₂O), 353.1089, 335.0986, 257.1008, 244.0927 (APU), 226.0821 (APU -H₂O), 217.0817, 209.0555, 208.0715, 191.0450, 165.0771 and 155.0447 m/z. Eleven fragments could be matched with plausible MS/MS decay products of PUM (Fig.: 10). In contrast, PUM could not be detected from semi-purified source material by MALDI-MS.

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

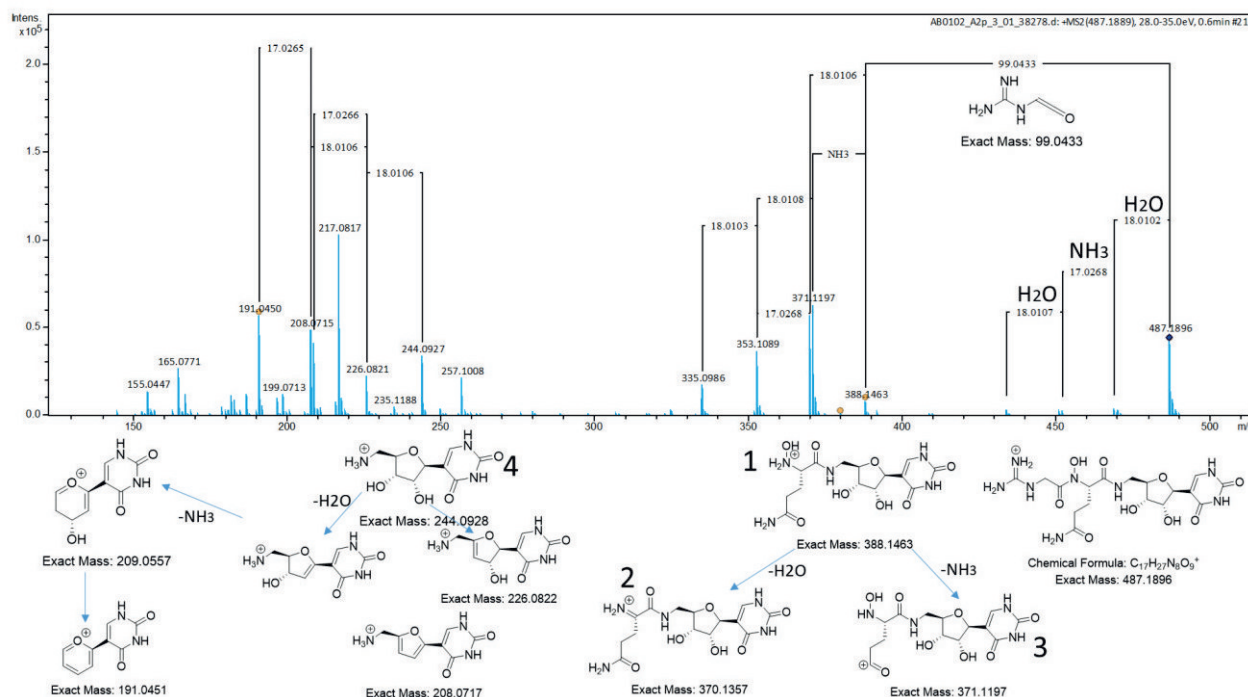


Figure 10: Fragmentation pattern analysis of the parent ion m/z 487.1896, corresponding to PUM. Top: Detected MS/MS fragmentation pattern. Bottom: Plausible MS/MS decay products, matching the recorded fragmentation pattern. (Image created with help of Dr. M. Patras).

III.3.4. Heterologous expression of the PUM Δ F Δ H BGC in *S. coelicolor* M1146 on agar plate using MALDI-MS

100 mL cultures of *Streptomyces* sp. DSM26212, *S. coelicolor* M1146 + pCAP03, *S. coelicolor* M1146 + pCAP03-PUM Δ F Δ H_ermE*, *S. coelicolor* M1146 + pCAP03-PUM Δ F Δ H_tcp830 were grown in ISP2, TSB and PM media were analysed for PUM production after 3 and 5 days, however no PUM production could be observed in LCMS measurements in the heterologous host strains. Since PUM could not be detected in in this first experiment, it was attempted to detect PUM by MALDI-MS measurements in cultures directly grown on agar plates. Due to inability to induce cultures on plates after a given time, only *S. coelicolor* M1146 + pCAP03-PUM Δ H_ermE* was tested. MALDI-MS measurements revealed presence of an ion with 487.1852 m/z in the WT producer strain and an ion with 487.1863 m/z in *S. coelicolor* M1146 + pCAP03-PUM Δ H_ermE* grown on ISP2 agar plates however with poor intensity. Both ions correspond well to the exact mass of PUM C₁₇H₂₆N₈O₉ [M+H]⁺ = 487.1896 Da with 9 ppm error and 7 ppm error respectively. Ions with matching m/z could not be detected in the negative control.

III.3.5. Heterologous expression of the PUM BGCs in liquid media

Fermentation of transconjugands was done in complex medium, i.e. in ISP2 and TSB. Thereby, the culture carrying the *tcp830* promoter were induced by addition of anhydrotetracycline (2 mg/L) to the cultures after one day of growth. PUM production was analysed after 3 and 5 days of fermentation by LCMS using the cleared culture broth. *S. coelicolor* M1146 – pCAP03 (empty vector) served as negative control. However, heterologous expression using construct pCAP03-PUM Δ F Δ H (see Fig.: 7, version I) did not result in detectable PUM production at all. Therefrom, it was hypothesized that a lack of correctly phosphorylated uridine precursors prevents detectable PUM production. To verify this assumption, the adenylate kinase encoding *pumH*, which catalyses the phosphorylation of a uridine-based substrate, was added to the construct, resulting in pCAP03-PUM Δ F (Fig.: 7, version II). Heterologous expression using the latter construct enabled detection of minor traces of an ion corresponding to the calculated PUM sum formula. Such an ion was not present in the negative control. This result pointed towards the fact that this could be regarded as the minimal gene set necessary for production of PUM. However, it was questionable if this very low expression yield justifies this as the minimal BGC. Even though it seems that construct pCAP03-PUM Δ F already contains all necessary genes for manufacture of PUM, a satisfying production is not possible, although a closely related strain was used as heterologous host. It was expected that the weak production could be achieved despite the lack of an important gene. To enable a valid PUM biosynthesis, it was tested if a further gene, i.e. *pumF*, is necessary.

It was envisaged to test if a further gene is essential to enable a valid PUM biosynthesis. A promising candidate was *pumF*, which is located right in the middle of the BGC and could be annotated as a putative serine/threonine kinase. The protein shows homology to the gene *mur33* (77 % query coverage; 43 % identity) in the BGC coding for Muraymycin C1, the latter another nucleoside natural product. BLASTp search of *pumF* and *mur33* showed homology (48 % and 50 %) to predicted serine/threonine protein kinases of Actinomycetes. Based on these observations, the *pumF* gene was incorporated into the expression construct pCAP03-PUM (Fig.: 7). Fermentation of a heterologous host carrying the expression construct pCAP03-PUM_ermE*, in which the PUM genes are under control of the ermE* promotor, in ISP2 medium led to detection of an ion with 487.1856 *m/z* after 5 days incubation time. This matched the PUM sum formula of C₁₇H₂₆N₈O₉ [M+H]⁺ (8 ppm error) and retention time; thereby, being absent in the empty vector control. Prolonged incubation (20 days) of the strain led to accumulation of this ion and allowed more accurate detection with 487.1896 *m/z* (0 ppm error) (Fig.: 11). Furthermore, targeted MS/MS fragmentation of this ion revealed the presence of the characteristic signature ion of 244.0920 *m/z*, which is indicative for the

III. HETEROLOGOUS EXPRESSION OF THE MINIMAL PSEUDOURIDIMYCIN BIOSYNTHETIC GENE CLUSTER IN STREPTOMYCETAL HETEROLOGOUS HOSTS

APU fragment (3 ppm error). Thus, the identity of the molecule as PUM was verified and the (indirect) involvement of PumF in PUM biosynthesis was shown.

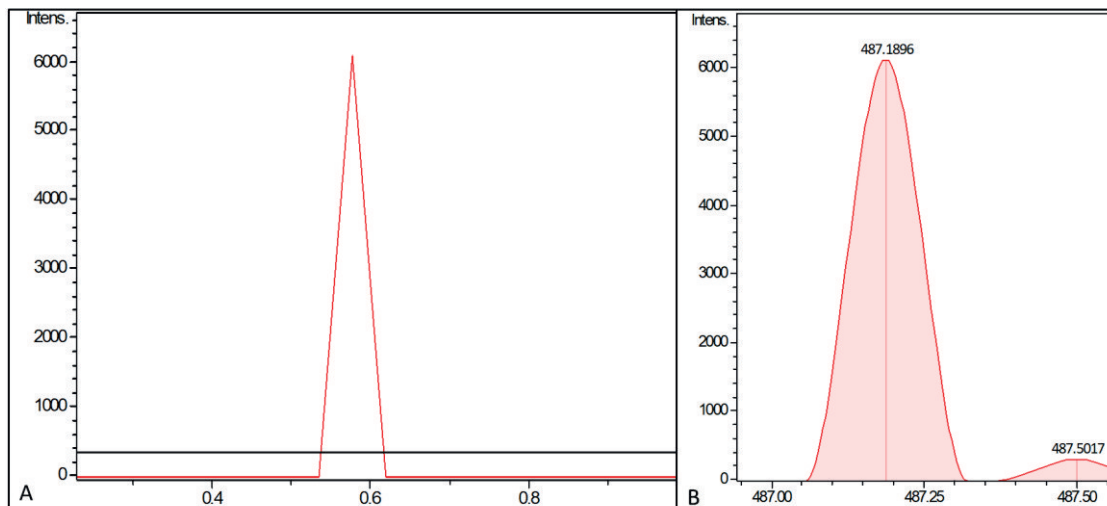


Figure 11: LCMS analysis of *Streptomyces coelicolor* M1146 fermentation broth. Displayed are A) the extracted ion chromatogram (EIC) for $C_{17}H_{26}N_8O_9$ $[M+H]^+ \pm 0.0005$ corresponding to PUM. The strain extract of *S. coelicolor* M1146 + pCAP03 (empty vector control in black) and *S. coelicolor* M1146 + pCAP03-PUM_ermE* (carrying the expression construct version III in red) are shown. Fermentation was performed in ISP2 medium. B) Measured m/z of the peak, fitting exactly to the calculated value for PUM ions (i.e., 487.1896 m/z).

For pCAP03-PUM Δ F constructs that were expressed in *S. albidoflavus* J1074 and *S. lividans* TK24, the picture is similar to the expression in *S. coelicolor* M1146 showing ions matching the PUM with only weak accuracy and intensity. No ions corresponding to PUM could be detected in cultures grown in TSB medium and in strains that were harbouring the expression constructs in which the streamlined BGC was under the control of the tcp830 promoter.

III.4. Discussion

III.4.1. Cloning the PUM BGC

In order to heterologously express the PUM BGC with and without the *pumH* gene, the whole gene cluster had to be cloned into a suitable *E. coli*/*Streptomyces* shuttle/expression vector. pCAP03 was chosen due to its well documented record for the heterologous expression of *Actinomycetal* NP BGCs (Tang *et al.*, 2015, Bauman *et al.*, 2019) and its relation to SuperCos1, a cosmid with the ability to carry large DNA fragments reliably in *E. coli*. Due to the fragmented nature of the BGC in the original producer *Streptomyces* sp. DSM26212, severe rearrangement of the individual genes was necessary. Therefore, the employed cloning strategy was based on the amplification of individual genes and larger stretches containing 3 and 4 genes already in the desired order and subsequent isothermal assembly into the final molecule. In contrast to other BGC cloning methods like TAR cloning, this allows higher flexibility for design of the final construct as well as higher tolerances for the DNA template quality. This is preferable for the work with *Actinomycetal*, *Cyanobacterial* and other “tricky to work with” DNA, making it a method with rapidly increasing popularity in the NP research community (D’Agostino and Gulder, 2018).

During the cloning process, no major obstacles were encountered. All PCRs worked well under standard conditions and isothermal assembly yielded a high amount of correctly assembled constructs. Subsequent modification of the constructs with artificial promoters and subsequent conjugation of the final constructs to *Streptomyces* heterologous hosts as well described by Gust *et al.*, 2003 worked flawlessly. This enabled construction of two separate versions of the BGC with *pumH* present and absent respectively under control of the strong promoters *ermE** and *tcp830*. Only failure in the course of this experiments was the inability to conjugate pCAP03-PUM-*tcp830* to *S. albidoflavus* J1074 in the first try, however this was accepted since the conjugation of the *ermE** promoter construct worked. Underlying idea for the use of inducible and constitutive promoters was to bypass possible autotoxicity issues when expressing the BGC constitutively.

III.4.2. Analysis of PUM production in *Streptomyces* sp. DSM26212

One major problem encountered in earlier stages of this study was the inability to detect PUM, even from the WT producer with HPLC-MS equipment used at that time. In order to remove background noise and to enrich PUM for better analysis, the first purification step (strong cation ion exchange) was performed with cleared supernatant from a 2 L fermentation of *Streptomyces* sp. DSM26212 in PM, essentially recreating the first step of the published isolation procedure. Subsequent

measurement on UPLC-MS clearly revealed the PUM peak with its extremely short retention time of 0.6 min while it could not be detected from enriched sample material by MALDI-MS. In turn, measurements directly from cleared *Streptomyces* DSM26212 supernatant on this UPLC-MS enabled rapid detection of PUM and therefore were method of choice for further investigation.

This inability to detect PUM on HPLC-MS equipment used at that time could be narrowed down to the employed gradient. In the PUM initial publication the authors employed a 100% H₂O tolerant RP C18 column with a gradient of 100% H₂O to 10:90 MeCN:H₂O in 20 min, reporting a retention time for PUM of 12 min corresponding to an elution at 6 % MeCN. This is well congruent with the inability to see PUM in the first few experiments since the HPLC-MS gradient started at 10 % MeOH, which elutes PUM in the very first solvent peak. Again, this is also well fitting to the very short retention time in the subsequently employed UPLC-MS instrument, since the gradient starts at 5 % MeCN. Unfortunately, the column used does not tolerate < 5 % organic solvent, so changing the gradient on the existing setup was no option. Furthermore, the applied UPLC-MS system was optimised for high-throughput measurements, so that no machining time could be expended for optimisation. This in conclusion shows that the very hydrophilic PUM is not effectively retained by C18 material, impeding its analysis. This also hints to the fact that the hydrophilic nature of PUM, making it insoluble in organic solvents commonly used for maceration such as ethyl acetate or butanol, is one of the reasons for its only very recent discovery. The original publication reports a yield of PUM from *Streptomyces* sp. DSM26212 cultivation of 98 mg/L pure compound. Despite questions arising about the purity of this reported material due to medium pressure chromatography instead of HPLC purification, this is in terms of bacterial natural products a tremendous yield for a non-optimised producer strain. Compared to other well investigated *Actinomycetal* and non-*Actinomycetal* natural products like Erythromycin with ~ 45 mg/L (Production patent US2653899A, 1952), Darobactin A with 2 mg/L (Imai *et al.*, 2019) and Coralopyronin A 0.66 mg/L (Erol *et al.*, 2010), PUM is produced very well by and can be isolated in substantial amount from the original producer strain. This contrasts with the relatively weak intensity of the PUM ion measured from enriched material when compared to the BPC. PUM was by no means one of the major peaks even in the semi purified material (Fig.: 9). This might be explained by *per se* bad ionisation of the molecule under the given conditions, a high ion suppression by other ions at the same retention time or a combination of both. Investigation revealed however, that the predominant ionisation of PUM was $[M + H]^+$ and other ionisations with common adducts or losses (+ Na, + NH₄, - H₂O, -NH₄) could not be detected.

In contrast to the difficulties detecting PUM in the first place, corroboration of the identity of PUM was effortless due to good MS/MS fragmentation of the molecule. It was possible to identify all

fragment masses reported in the original publication, albeit with significantly higher accuracy. This enabled the identification of all major plausible MS/MS decay products.

III.4.3. Heterologous expression of three streamlined versions of the PUM BGC

Due to its recent discovery, interesting activities and a novel mode of action, PUM is a compound of interest for further investigation and possible pharmaceutical development. In order to obtain better knowledge of this compound, goals of this study were i) confirmation of the minimal BGC necessary for the production of PUM and ii) generate a heterologous producer strain of PUM in order to isolate the compound for further investigation. If successful, this heterologous expression system could in turn serve as a platform to produce PUM biosynthetic precursors by selecting the respective set of genes necessary or even to derivatise the structure by incorporating suited genes for subsequent modifications at the PUM backbone. This would be facilitated tremendously by using a refactored BGC that can easily be cloned and modified in *E. coli* and that does not provoke regulatory issues when modifying the very fragmented BGC in the WT producer strain. The general feasibility of the PUM BGC for the envisioned tinkering was already demonstrated by Sosio *et al.*, 2018 while characterising the BGC in detail. Nine catalytic steps were proposed for the formation of PUM and matched with plausible enzymatic functions within the BGC, proposing the minimal PUM BGC (Sosio *et al.*, 2018). Here one by one deletion of the biosynthetic genes from the WT producer strain resulted in accumulation of the respective precursors. While the involvement of *pumH* in PUM biosynthesis was corroborated by the KO experiments, it was also shown that deletion of *pumH* from the WT producer did not lead to abolition of PUM biosynthesis, hinting towards a rather regulatory function of *pumH*. It was reasoned that PumH as adenylate kinase alters the phosphorylation pattern of the precursor uridine, which in turn could be processed into pseudouridine. In this case, correctly phosphorylated uridine could be sourced in small amount from tRNA turnover in primary metabolism (Sosio *et al.*, 2018). This gives evidence that PumH has a gatekeeper function in funneling uridine-based nucleotides (UMP or UDP) as precursor molecules towards the PUM pathway. This is important for the producing organism, to ensure that the uridine pool designated to transcription and translation is not drained for PUM production. In this way, PumH would decouple PUM biosynthesis from the primary metabolism uridine pool. However, heterologous expression of the identified genes (in- and excluding *pumH*) did not lead to clearly detectable production of PUM (i.e., only constructs including *pumH* resulted in traces of PUM at the detection limit). This provoked the question whether the proposed biosynthetic route covers all steps and hence, if the proposed minimal PUM BGC is correct or if further important functions are encoded in the BGC.

The *pumF* gene is located right in the middle of the BGC and codes for a putative serine/threonine kinase with a homolog gene, i.e. *mur33*, in the BGC coding for the nucleoside natural product Muraymycin C1. Enzyme phosphorylation is one of the most common posttranslational modifications, often involved in regulation of enzyme activity, i.e. activating or inactivating a given enzyme (Cousin et al., 2013). Hence, it seemed reasonable to assume that this gene is indirectly involved in the PUM biosynthesis, by regulating the activity of one or more enzymes participating in assembly of the molecule. Heterologous expression of the previously identified minimal BGC incorporating *pumF* led to detectable PUM formation. Thereby, involvement of PumF in the PUM biosynthesis was corroborated.

It has to be emphasized that the presented results do not contradict the initially proposed biosynthetic route (Sosio *et al.*, 2018), instead heterologous expression is a further proof of the PUM BGC. However, the here collected data indicate that the genes *pumH* and *pumF* should be considered as well as components of the minimal BGC necessary for PUM biosynthesis. Based on the obtained results, we propose that the genes *pumD*, *E*, *G*, *I*, *J*, *K*, *M*, *N* code for enzymes directly participating in de novo formation of PUM, while *pumH* serves as decoupling mechanism of PUM biosynthesis from primary metabolism as demonstrated before by KO experiments (Sosio *et al.*, 2018). The *pumF* gene encodes a regulatory protein controlling the enzymatic activity of one (or more) of the participating enzymes by post-translational phosphorylation. Due to presence of minor traces of an ion consistent with PUM, it remains possible that the target enzyme of PumF can retain (poor) activity even in the wrong phosphorylation state or that minor amounts of protein undergo phosphorylation from other serine/threonine kinases from the surrogate hosts genetic background with certain promiscuity. Despite the proof that presence of *pumF* is important for PUM biosynthesis, the precise function of PumF remains to be elucidated in future studies, since currently the kinase activity is not proven and a protein target is unknown. Furthermore, heterologous production of PUM was achieved; however, the yields were too low to use this strain as a plug and play production platform for PUM and its biosynthetic precursors. Therefore, it would be a long way to optimize production in heterologous strains and future work should continue to focus on WT producer strains, which are also genetically accessible and could be optimized in untargeted and target ways

IV. DAROBACTIN

IV.1. Introduction

Recently a novel antibiotic that selectively kills Gram- pathogens called Darobactin (Fig.: 12) was discovered. It is produced by several *Photorhabdus* strains and it shows good *in-vitro* activity against Gram- pathogens, e.g. *E. coli* (MIC 4 µg/ml), *Klebsiella pneumoniae* (MIC 2 µg/ml) and *Acinetobacter baumannii* (MIC 8 µg/ml) (Imai *et al.*, 2019). In contrast, Gram+ organisms and certain members of the bacterial gut microflora from the evolutionary distinct *Bacteroidetes* phylum are not affected by this compound. Additionally, testing the efficacy in a mouse septicemia model against *E. coli* and *P. aeruginosa* revealed good *in-vivo* efficacy in the tested scenarios. While untreated infected mice died within 24 hours, mice treated with Darobactin dose dependently showed higher survival rates up to 100 % survival at 2.5 mg/kg in both infection models (Imai *et al.*, 2019).

Investigating the antibiotic target of Darobactin surprisingly did not yield any hit in the screens for well-investigated targets like lipid A, teichoic acids or RNA polymerase (RNAP). Transcriptome analysis revealed a general activation of envelope stress pathways, which together with the selective Gram- activity hinted towards a new, cryptic mode of action against the outer membrane. Few other antibiotics also interfere with the outer membrane, e.g. the polymyxins and other bacteriocidal compounds subsumed under the term antimicrobial peptides (AMPs). These however do not have distinct modes of action but a more general destabilising effect to lipid bilayer membranes. Usually these rather non-specific modes of action result in severe unwanted side effects like general cytotoxicity, nephrotoxicity or neurotoxicity when pushed into pharmaceutical application. These preoccupations increase the need to find anti Gram- compounds with distinct modes of action, ideally limited exclusively to bacterial targets. Finally, breeding of resistant *E. coli* MG1655 clones and subsequent whole genome analysis revealed that all resistant clones accumulate several mutations in one specific gene, i.e. *bamA*. This gene encodes BamA, the central component of the β -barrel-assembly-machinery (Bam) in the outer membrane (OM) of Gram- bacteria. The Bam complex consists of the proteins BamA, BamB, BamC, BamD and BamE. Thereby, all but BamA are located in the periplasm and BamA as the central component protruding the outer membrane. The Bam complex plays a vital role in the assembly and insertion of outer membrane proteins (OMP) into the OM and its function was completely elucidated just recently. BamA consists of 16 antiparallel β -strands, which integrate into the outer membrane as a non-covalently closed β -barrel structure. Further polypeptide-translocation associated (POTRA) domains are connected to the periplasmatic face of BamA acting as a “lid” to the β -barrel structure. Nascent OMPs get secreted into the periplasm and assembled into the β -barrel structure of BamA (Fig.: 12). The POTRA “lid”-like

structure restricts further access to the β -barrel. Then, the non-covalently linked lateral gate of BamA opens to release the protein to the outer membrane. To assist this assembly process, the thickness and order of the lipid bilayer surrounding BamA is lowered. Interference with this complex system will have significant effect on the formation and maintenance of the OM and thus, general fitness of the cell. BamA as target of Darobactin is well congruent with the observed activity spectrum. Gram+ bacteria, lacking an outer membrane obviously also lack a Bam complex of any kind, while the Gram- *Bacteroidetes* phylum is evolutionary distinct from the tested Gram-pathogens belonging to the rather closely related *Enterobacteriales* (*E. coli*, *Salmonella typhimurium*, *Klebsiella pneumoniae*, *Shigella sonnei*, *Proteus mirabilis* and *Enterobacter cloacae*) and *Pseudomonadales* (*Pseudomonas aeruginosa*, *Acinetobacter baumannii* and *Moraxella catarrhalis*) orders in the γ -*Proteobacteria* class. Comparison of *bamA* genes of *E. coli* and *Bacteroides fragilis* reveals vast differences in overall length of the genes as well as low similarity on protein level. BamA so far was not identified or validated as feasible antibiotic target, despite efforts to find compounds interfering with its function. BamA's exposed location outside the cell make it an attractive target since the outer membrane usually restricts access of drugs to intracellular targets. Until very recently, only monoclonal antibodies and nanobodies were discovered that specifically target and inactivate BamA (Storek *et al.*, 2018). This changed in late 2019, when nearly at once, three publications individually reported small molecules with activity specifically against BamA:

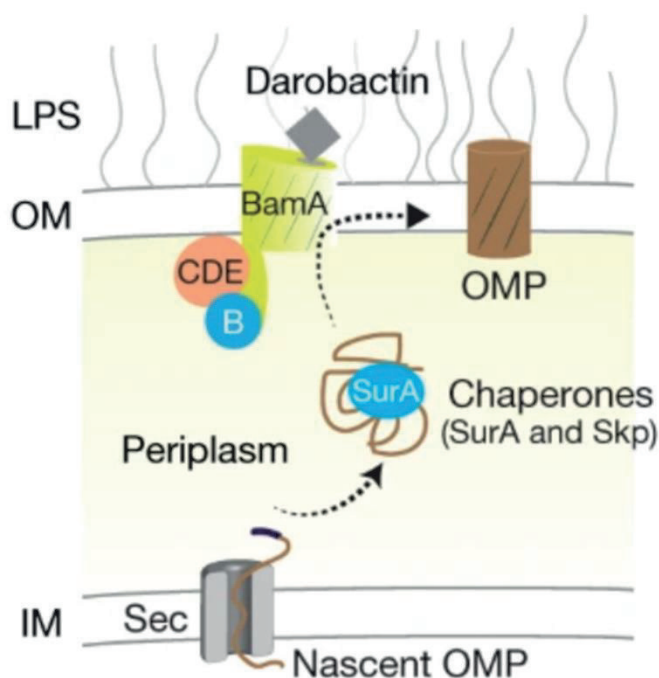


Figure 12: Schematic illustration of insertion of OMPs into the OM in *E. coli*: Nascent OMPs are secreted into the periplasm where they are bound by chaperones and transported to BamA. Subsequently, BamA introduces the protein into the OM. Darobactin was identified to bind to BamA and arrest it in closed state, preventing introduction of OMPs to the OM (Figure from Imai *et al.*, 2019).

MRL-494 (Fig.: 12), a cationic amphiphilic compound with an MIC of $\sim 16 \mu\text{g/mL}$ ($25 \mu\text{M}$) against *E. coli* and $\sim 8 \mu\text{g/mL}$ ($12.5 \mu\text{M}$) against MRSA was shown to inhibit the biosynthesis of OM proteins in *E. coli*. It was discovered by screening a synthetic compound library for activity specifically against the OM of *E. coli*. When applied in sub-lethal concentration together with Rifampicin, the compound drastically reduced the MIC of Rifampicin by permeabilization of the OM. Further investigation showed that MRL-494 binds near to or at BamA directly, which is responsible for the inhibition of OMP formation. However, the even better activity against MRSA shows an additional antibiotic mechanism, hence it was speculated that the cationic amphiphilic nature of the compound also tends to interfere with lipid bilayers in general (Hart *et al.*, 2019).

Furthermore, a series of chimeric peptidomimetic compounds (Fig.: 12) was reported by Luther *et al.*, 2019. The reported peptidomimetic antibiotic compounds exhibited strong bactericidal effect against Gram- pathogens ($\text{MIC} \leq 2 \mu\text{g/mL}$). Proceeding from the identification of two separate pharmacophores in Murepavidin, chimeric peptidomimetics were created. Here, a mimetic of the Murepavidin pharmacophore was fused to a mimetic of the cyclic Polymyxin B pharmacophore. The resulting compounds have excellent activities against Gram- pathogens, while no activity against Gram+ pathogens and eukaryotic lipid bilayers was observed. Consequential mouse *in-vivo* testing revealed lower propensity for renal damage compared to the lipopeptide antibiotic Colistin. (Luther *et al.*, 2019).

The third publication reporting a small molecule inhibitor of BamA was the previously mentioned finding of Darobactin. It is a heptapeptidic natural product with the amino acid sequence $\text{W}^1\text{-N}^2\text{-W}^3\text{-S}^4\text{-K}^5\text{-S}^6\text{-F}^7$, containing only proteinogenic amino acids and two distinct unusual ring closures between the first, third and fifth amino acid as key features: One aromatic-aliphatic ether (C-O-C) bridge from the aromatic residue of W^1 and the β -carbon atom of W^3 and one aromatic-aliphatic covalent (C-C) bond between the aromatic residue of W^3 and the β -carbon of K^5 . Darobactin was discovered from several members of the *Photorhabdus* genus in a bioprospecting approach. Members of the *Photorhabdus* genus are commonly associated with higher organisms (Clarke 2008). Especially well investigated is their symbiosis with *Heterorhabditis* nematodes, where *Photorhabdi* facilitate the infection of insect larvae by the *Heterorhabditidis* nematode host. Underlying idea was the assumption that compounds isolated from bacteria living in symbiosis with higher organisms (such as nematodes) are by evolution designed to exhibit lower toxicity towards their eukaryotic host and hence probably towards mammals as well. This would drastically facilitate the compounds development into a drug, since toxicity and resulting severe side effects are a major reason for failure of novel drugs in (pre-)clinical trials.

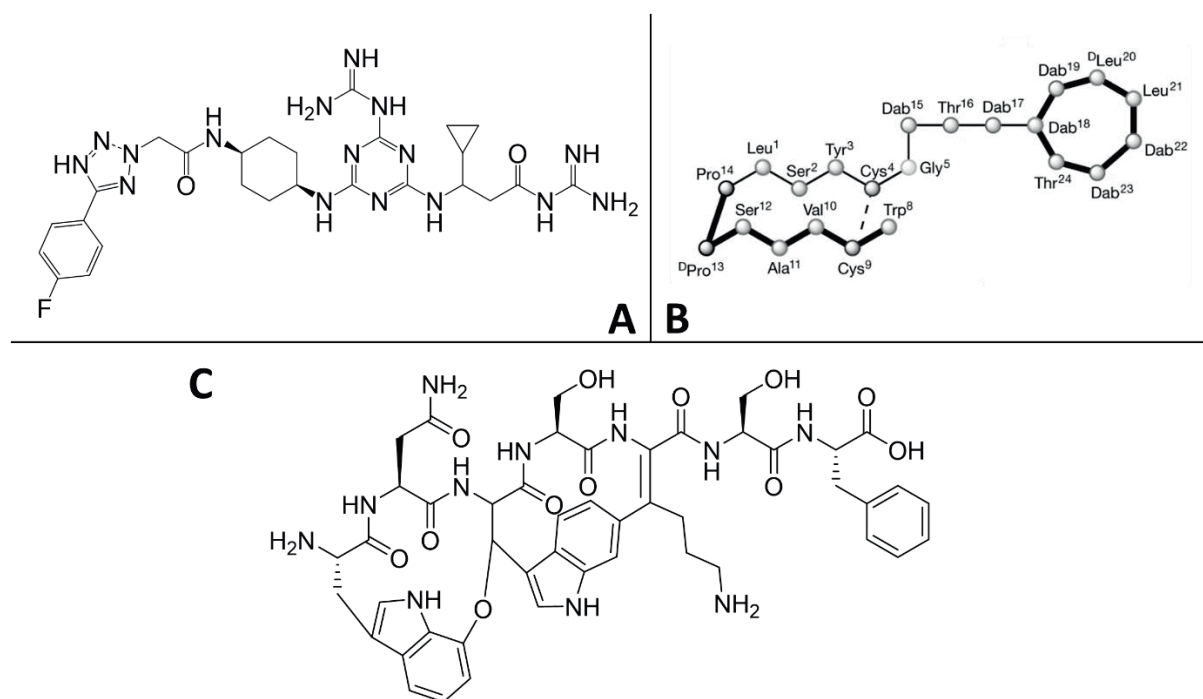


Figure 13: Recently discovered small molecule BamA inhibitors: A) MRL-494, B) Example for a series of peptidomimetic antibiotic compounds (Luther *et al.*, 2019, modified), C) Darobactin A.

After finding the compound and initial testing of its antibacterial properties, attention shifted towards its biosynthesis. Most intriguingly with its dual ring system, the compound has two rare features: The (eastern) aromatic-aliphatic covalent (C–C) bond and the (western) aromatic-aliphatic ether (C–O–C) bond without any precedent. The C–C bond bears strong similarities to Streptide (Schramma *et al.*, 2015), a ribosomally synthesized post-translationally modified peptide (RiPP), which also features a covalent (C–C) bond between the aromatic residue of a tryptophan and the α -carbon of a lysin. While Streptide was identified as streptococcal bacterial pheromone and hence does not share any activity with Darobactin, the occurrence of a similar C–C bond in both molecules indicates a similarity in biosynthesis in both compounds.

IV.1.1. Ribosomally synthesised post-translationally modified peptides (RiPPs)

Although members of this NP class, e.g. Nisin are known since the 1930s, the term RiPP was coined just recently and universal nomenclature was put in place in 2013 (Arnison *et al.*, 2013). RiPPs represent a class of peptidic natural products, with a plethora of bioactivities however, essentially the same initial biosynthetic steps. Among others, RiPPs with antibacterial (Nisin), antifungal (Pinensins), antiviral (Labyrinthopeptin), cytotoxic (Amanitins) and protease inhibiting (Microviridin) activities have been reported (Dang and Süssmuth, 2017, Mohr *et al.*, 2015).

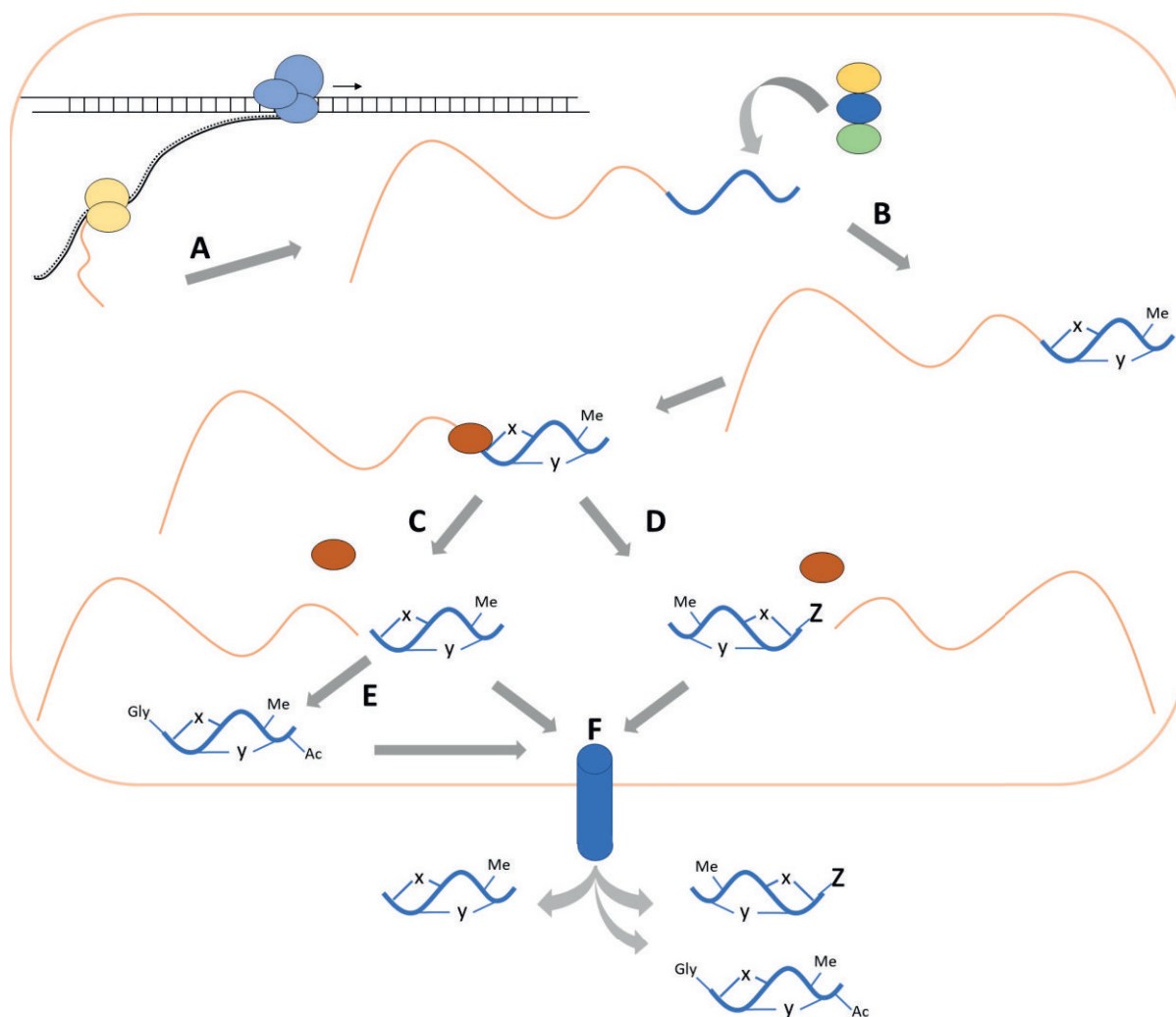


Figure 14: Schematic overview of general RiPP biosynthesis: A) Transcription of the propeptide gene yields the linear, unmodified propeptide with leader (yellow) and core (blue) AAs. B) The core peptide portion of the propeptide is modified by posttranslational modification. Modifications include intramolecular ring closures of various kind (formation of lanthionine bridges, intramolecular esters, intramolecular peptide bonds and more), N-, O-, and C-methylations and others. C) Peptidase mediated cleavage of leader and modified core peptide. D) Peptidase mediated cleavage of leader and modified core peptide and simultaneous modification of the cleavage site by the peptidase. E) Further augmentation of the product such as glycosylation or acylation by enzyme acting on the leaderless modified product. F) Export of the mature product to the environment.

The often-used term secondary metabolites is not quite precise for RiPPs since the biosynthesis of RiPPs involves the standard protein biosynthesis and certain RiPPs were identified as essential for the producing organism in the form of cofactors. RiPPs can be further subdivided into multiple subclasses such as lasso peptides, lanthipeptides, thiopeptides and more (Hudson and Mitchel, 2018). Initial step in the biosynthesis of every RiPP is the translation of a (usually small) peptide called propeptide. This propeptide is subdivided into leader- and core-peptide regions. While the backbone of the final product is composed of the core AAs, the leader peptide governs the posttranslational modifications and can be found at N- or C-terminal position of the core AAs. The leader peptide interacts with specific recognition sites of the modifying enzymes and by that

interaction, seats the core AAs in the active site cavity of the enzyme. Common posttranslational modifications that can take place on an L-amino acid peptide chain are dehydratations of hydroxyl containing amino acids like serine and condensation of sulphur containing amino acid side chains to form intramolecular lanthionine bridges, e.g. in Lanthipeptide biosynthesis. Furthermore, epimerisations and O-, N- and C- methylations, e.g. in Proteusins. Subsequently, the leader is cleaved off and the mature product is exported to the environment (Fig.: 14). The resulting products can vary from being close to the transcribed AA template such as the Lasso peptides, where major modification is a change in topology (linear vs knotted) to a degree of modification that can completely obscure the ribosomal origin such as the Proteusins. Especially the latter are modified extensively, incorporating D-AAs and rare AAs that do not resemble anything derived from primary metabolism such as tert-butyl AAs that are generated by RaS mediated C- methylation of valine (Bhushan *et al.*, 2019). These observed epimerisations and C– methylations are result of radical SAM (RaS) enzymes, which are commonly participating in maturation of RiPPs. RaS enzymes act by breaking down S-adenosyl-L-methionine to the 5'-deoxyadenosyl radical (Yokoyama and Lilla, 2018). This energy rich radical is in turn able to abstract a hydrogen atom; hence, functionalise inactivated carbon bonds. These secondary radicals can subsequently react with other atoms from the molecule or other molecules presented by the enzyme. This can result in formation of (in terms of organic chemistry) highly unusual moieties such as the mentioned C–C covalent bond in Streptide.

IV.1.2. Scope of this study

In the following, the contribution to the elucidation of the Darobactin biosynthesis i.e. the discovery and proof of its BGC will be elaborated thoroughly. This contribution directly resulted in co-authorship of the publication “A new antibiotic selectively kills Gram- pathogens” (Imai *et al.*, 2019). Furthermore, the consequential (i) discovery of analogues and (ii) mutasynthetic creation thereof, as well as (iii) discovery of three new naturally occurring analogues from the marine bacteria *Pseudoalteromonas luteoviolacea* H33 and H33S and initial characterisation of these analogues by tandem mass spectrometry (MS/MS) and activity assays will be covered in detail.

IV.2. Material and Methods

IV.2.1. Bacterial strains used in this study

Table 5: List of bacterial strains used in this study

Strain	Genotype	Reference
<i>P. khanii</i> HGB1456	WT	Tailliez et al. 2010
<i>P. khanii</i> DSM3369	WT	Tailliez et al. 2010
<i>E. coli</i> S17 + pSAMbt	TpR SmR recA, thi, pro, hsdR-M+RP4: 2-Tc:Mu: Km Tn7 λpir + pSAMbt	Goodman et al. 2009
<i>E. coli</i> WM3064	hsdS lacZΔM15 RP4-1360 Δ(araBAD)567 ΔdapA1341::[erm pir(wt)]	William Metcalf (unpublished)
<i>E. coli</i> WM3064 + pNB01	hsdS lacZΔM15 RP4-1360 Δ(araBAD)567 ΔdapA1341::[erm pir(wt)] + pNB01	this work
<i>P. khanii</i> DSM3369 + pNB01	WT + chrom. Integrated pNB01	this work
<i>E. coli</i> WM3064 + pNPTS12	hsdS lacZΔM15 RP4-1360 Δ(araBAD)567 ΔdapA1341::[erm pir(wt)] + pNPTS12	Lassack et al.
<i>E. coli</i> BW25113 + pKD46	lacI+rrnBT14 ΔlacZΔM15 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pKD46	Datsenko & Wanner, 2000
<i>E. coli</i> WM3064 + pNB02	hsdS lacZΔM15 RP4-1360 Δ(araBAD)567 ΔdapA1341::[erm pir(wt)] + pNB02	this work
<i>P. khanii</i> DSM3369 ΔdarABCDE	WT ΔdarABCDE	this work
<i>E. coli</i> XL1blue + pIJ773	endA1 gyrA96(nalR) thi-1 recA1 relA1 lac glnV44 F' ::Tn10 proAB+ lacIq Δ(lacZ)M15 hsdR17(rk- mK+) + pIJ773	Gust et al. 2003
<i>E. coli</i> DH5α + pCAP03	DH5α + pCAP03	Tang et al.2015
<i>E. coli</i> Top10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-	commercially available (invitrogen)
<i>E. coli</i> Top10 + pNB03	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ- + pNB03	this work
<i>E. coli</i> Top10 + pNB03-darABCDE	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ- + pNB03-darABCDE	this work
<i>E. coli</i> Top10 + pNB03Daro	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ- + pNB03Daro	this work
<i>E. coli</i> ET12567 + pUB307	dam-13::Tn9, dcm-6, hsdM + pUB307	Flett et al., 1997
<i>P. khanii</i> DSM3369 ΔdarABCDE + pNB03	WT ΔdarABCDE + pNB03	this work
<i>P. khanii</i> DSM3369 ΔdarABCDE + pNB03-darABCDE	WT ΔdarABCDE + pNB03-darABCDE	this work
<i>P. khanii</i> DSM3369 ΔdarABCDE + pNB03Daro	WT ΔdarABCDE + pNB03Daro	this work
<i>E. coli</i> BW25113 + pNB03-darABCDE	lacI+rrnBT14 ΔlacZΔM15 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNB03-darABCDE	this work
<i>E. coli</i> BW25113 + pNB03Daro	lacI+rrnBT14 ΔlacZΔM15 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNB03Daro	this work
<i>E. coli</i> Top10 + pNB03Daromod	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ- + pNB03Daromod	this work
<i>E. coli</i> BW25113 + pNB03Daromod DarobactinB + fol	lacI+rrnBT14 ΔlacZΔM15 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNB03Daromod DarobactinB + fol	this work
<i>E. coli</i> BW25113 + pNB03Daromod DarobactinC + fol	lacI+rrnBT14 ΔlacZΔM15 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNB03Daromod DarobactinC + fol	this work
<i>E. coli</i> BW25113 + pNB03Daromod DarobactinD + fol	lacI+rrnBT14 ΔlacZΔM15 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNB03Daromod DarobactinD + fol	this work
<i>E. coli</i> BW25113 + pNB03Daromod DarobactinE + fol	lacI+rrnBT14 ΔlacZΔM15 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNB03Daromod DarobactinE + fol	this work
<i>E. coli</i> BW25113 + pNB03Daromod DarobactinA - fol	lacI+rrnBT14 ΔlacZΔM15 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNB03Daromod DarobactinA - fol	this work

Strain	Genotype	Reference
<i>E. coli</i> BW25113 + pNBDaromod DarobactinB - fol	lacI+rrnBT14 ΔlacZWJ16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNBDaromod DarobactinB - fol	this work
<i>E. coli</i> BW25113 + pNBDaromod DarobactinC - fol	lacI+rrnBT14 ΔlacZWJ16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNBDaromod DarobactinC - fol	this work
<i>E. coli</i> BW25113 + pNBDaromod DarobactinD - fol	lacI+rrnBT14 ΔlacZWJ16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNBDaromod DarobactinD - fol	this work
<i>E. coli</i> BW25113 + pNBDaromod DarobactinE - fol	lacI+rrnBT14 ΔlacZWJ16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNBDaromod DarobactinE - fol	this work
<i>E. coli</i> BW25113 + pNBDaromod DarobactinF - fol	lacI+rrnBT14 ΔlacZWJ16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNBDaromod DarobactinF - fol	this work
<i>E. coli</i> BW25113 + pNBDaromod DarobactinA - Phe	lacI+rrnBT14 ΔlacZWJ16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNBDaromod DarobactinA - Phe	this work
<i>E. coli</i> BW25113 + pNBDaromod DarobactinA - SerPhe	lacI+rrnBT14 ΔlacZWJ16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 rph-1 Δ(araB-D)567 Δ(rhaD-B)568 ΔlacZ4787(::rrnB-3) hsdR514 rph-1 + pNBDaromod DarobactinA - SerPhe	this work
<i>E. coli</i> DH5α + pRSFduett	DH5α + pRSFduett	commercially available (Novagen)
<i>E. coli</i> BAP1mut	BL21(DE3) ΔprpRBCD::T7prom-sfp,T7prom-prpE bamA(1300 A>G, 1334 A>C, 2113 G>A)	created by Zerlina G Wuisan (Darobactin resistant, unpublished)
<i>E. coli</i> BAP1mut + pRSFduett - Darobactin A	BL21(DE3) ΔprpRBCD::T7prom-sfp,T7prom-prpE bamA(1300 A>G, 1334 A>C, 2113 G>A) + pRSFduett - Darobactin A	created by Zerlina G Wuisan (unpublished)
<i>E. coli</i> BAP1mut + pRSFduett - Darobactin B	BL21(DE3) ΔprpRBCD::T7prom-sfp,T7prom-prpE bamA(1300 A>G, 1334 A>C, 2113 G>A) + pRSFduett - Darobactin B	this work
<i>E. coli</i> BAP1mut + pRSFduett - Darobactin C	BL21(DE3) ΔprpRBCD::T7prom-sfp,T7prom-prpE bamA(1300 A>G, 1334 A>C, 2113 G>A) + pRSFduett - Darobactin C	this work
<i>E. coli</i> BAP1mut + pRSFduett - Darobactin D	BL21(DE3) ΔprpRBCD::T7prom-sfp,T7prom-prpE bamA(1300 A>G, 1334 A>C, 2113 G>A) + pRSFduett - Darobactin D	this work
<i>E. coli</i> BAP1mut + pRSFduett - Darobactin E	BL21(DE3) ΔprpRBCD::T7prom-sfp,T7prom-prpE bamA(1300 A>G, 1334 A>C, 2113 G>A) + pRSFduett - Darobactin E	this work
<i>E. coli</i> BAP1mut + pRSFduett - Darobactin F	BL21(DE3) ΔprpRBCD::T7prom-sfp,T7prom-prpE bamA(1300 A>G, 1334 A>C, 2113 G>A) + pRSFduett - Darobactin F	this work
<i>Pseudoalteromonas luteoviolacea</i> H33	WT	Vynne et al., 2013
<i>Pseudoalteromonas luteoviolacea</i> H33S	WT	Vynne et al., 2014
<i>E. coli</i> MG1655	WT + aceA-mScarlet	ADC NeU Boston collection
<i>E. coli</i> MG1655 <i>bamA6</i>	<i>bamA6</i> [duplication of Q217 and K218, Hagan & Kahne, 2018] + aceA-mVenus	ADC NeU Boston collection
<i>E. coli</i> ATCC35218	WT	DSM5923
<i>E. coli</i> NRZ14408 KPC-2	KPC-2 (clinical isolate; Gentamycin, Rifampicin and Tetracycline resistant)	Chakraborty (JLU Giessen)
<i>E. coli</i> K0416 VIM-1	VIM-1 (clinical isolate; Tetracycline resistant)	Chakraborty (JLU Giessen)
<i>E. coli</i> Survcare 052 NDM-5	NDM-5 (clinical isolate)	Chakraborty (JLU Giessen)
<i>E. coli</i> MMG11 OXA-48	OXA-48 (clinical isolate; Gentamycin and Tetracycline resistant)	Chakraborty (JLU Giessen)
<i>Pseudomonas aeruginosa</i> PAO 1	WT	DSM22644
<i>Pseudomonas aeruginosa</i> PAO 750	efflux null	IME-BR Fraunhofer collection
<i>Klebsiella pneumoniae</i> DSM30104	WT	DSM30104
<i>Acinetobacter baumannii</i> ATCC19606	WT	DSM30007
<i>Salmonella enteritidis</i> ATCC13076	WT	IME-BR Fraunhofer collection

IV.2.2. Media and additives used in this study**Lysogenic broth (LB)**

Peptone from casein	10 g/L
Yeast extract	5 g/L
NaCl	5 g/L
(Agar-agar	20 g/L)
p.H.	7.0

Marine broth (MB)

Marine Broth (Difco)	40 g/L
(Agar-agar	20 g/L)
p.H.	7.0

M1

ASW 200 %	500 mL/L
MOPS	40 mM
Glycerol	3 g/L
Casitone	5 g/L
p.H.	8.2

M2

ASW 200 %	500 mL/L
MOPS	40 mM
Glucose	3 g/L
Casitone	5 g/L
p.H.	8.2

M3

ASW 200 %	500 mL/L
MOPS	40 mM
Glucose	3 g/L
Casitone	5 g/L
Urea	1 g/L
p.H.	8.2

M4

ASW 200 %	500 mL/L
MOPS	40 mM
Glucose	3 g/L
Casitone	5 g/L
NH ₄ -Ac	1 g/L
p.H.	8.2

M5

ASW 200 %	500 mL/L
MOPS	40 mM
Glucose	3 g/L
Casitone	5 g/L

IV – DAROBACTIN

NH ₄ Cl	1 g/L
p.H.	8.2

M6

ASW 200 %	500 mL/L
MOPS	40 mM
Glucose	3 g/L
Casitone	5 g/L
Na-Glutamate	1 g/L
p.H.	8.2

M7

ASW 200 %	500 mL/L
MOPS	40 mM
Starch	3 g/L
Peptone	5 g/L
p.H.	8.2

M8

ASW 200 %	500 mL/L
MOPS	40 mM
Starch	3 g/L
Malt extract	5 g/L
p.H.	8.2

M9

ASW 200 %	500 mL/L
MOPS	40 mM
Mannitol	3 g/L
Peptone	2.5 g/L
p.H.	8.2

M10

ASW 200 %	500 mL/L
MOPS	40 mM
Soybean flour	2.5 g/L
Peptone	2.5 g/L
p.H.	8.2

Mueller Hinton broth II (cation adjusted)

Beef Extract	3 g/L
Acid Hydrolysate of Casein	17.g g/L
Starch	1.5 g/L
Ca ²⁺	0.025 g/L
Mg ²⁺	0.0125 g/L

Stock solutions

Kanamycin (Kan)	50 g/L
Apramycin (Apra)	50 g/L
Carbenicillin (Carb)	50 g/L
Rifampicin	12.8 g/L
Gentamycin	12.8 g/LI
Tetracycline	12.8 g/LI
Colistin	12.8 g/LI
L-arabinose (Ara)	1 M
2,6-Diaminopimelic acid (DAP)	60 mM
Isopropyl β -D-1-thiogalactopyranoside (IPTG)	1 M

Media were prepared using deionised H₂O and autoclaved at 121 °C for 20 min and 2 bar vapor pressure. Solid media were allowed to cool down to ~50 °C before adding selective antibiotics dilution or other additives, poured into sterile petri dishes and left open to solidify and dry. Liquid and solid media were stored at room temperature, stock solutions were sterile filtered and stored at -20 °C. Media components and chemicals were obtained by standard laboratory suppliers e.g. CarlRoth or Sigma.

IV.2.3. Primers used in this study

Table 6: List of oligonucleotide primers used in this study

Name	Sequence 5' - 3'
pNB01-ins_f	GAAGCAGGGTTATGCAGCGGAAAGCCACCAGGCTTTGCAATTTAAGATG
pNB01-ins_r	ATACCTCTTTGTTGGCTGCGATCACCCGGCTGGCATCGGTTATCTGGGAC
pNB01-R6K-Ori_f	CGTTGTATCGCTCTTGAAGG
pNB01-R6K-Ori_r	CCGCTGCATAACCCTGCTTC
pNB01-Kan_f	CGCAGCCAACAAAGAGGTAT
pNB01-Kan_r	CCTTCAAGAGCGATACAACGCGCGGAACCCCTATTTGTTT
pNB02-upstr_f	TTTGACGTTGGAGTCCACGTGTTATGGACGTGGCAAACGCGGTTCTTGAC
pNB02-upstr_r	TTGAAATATCAGGATAGCATTGCGCTCGCTCACCCGGTCACATAGTTCG
pNB02-dwnstr_f	ATGCTATCCTGATATTTCAAATGCAAGTAAATGTTTCATCATAATAACC
pNB02-dwnstr_r	TTCTTGACGAGTTCTTCTGAGATGGGTTGATATCCACTGATATAAATCTC
pNB02-pNPTS_f	TCGAGCTCTAAGGAGGTTATAAAAAATGAACATCAAAAAGTTTGCAAAACAAGCA
pNB02-pNPTS_r	ACGTGGACTCCAACGTCAAA
pNB02-blapr_f	ACTCTTCCTTTTCAATATTATTGAAGCAT
pNB02-blapr_r	TGCATTTTTTATAACCTCCTTAGAGCTCGAATTCC
pNB02-aph_f	TCAGAAGAAGTCAAGCAAGGCGA
pNB02-aph_r	TCAATAATATTGAAAAAGGAAGAGTATGATTGAACAAGATGGATTGCACG
DCOscr_f	ATCTCCATCAAAGCGCTACC
DCOscr_r	CCGCGCTGCAACTCGAAATC
pNB03-ori_f	CTCTAAGGAGGTTATAAAAAGCGGCCGCATCCCTTAACGTGAGTTTTC

pNB03-ori_r	GGTCGACGGATCCCCGGAATAGCGGAAATGGCTTACGAAC
pNB03-Aprar_f	ATTCCGGGGATCCGTCGACC
pNB03-Aprar_r	TGTAGGCTGGAGCTGCTT
pNB03-araC_f	AAGCAGCTCCAGCCTACATCAGAAGAACTCGTCAAGAAGGCGA
pNB03-araC_r	TTTTTATAACCTCCTTAGAGCTCGAATTCC
pNB03_f	TCCCTTAACGTGAGTTTTCG
pNB03_r	TTTTTATAACCTCCTTAGAGCTCGAA
pNB03-darA_f	GCTCTAAGGAGGTTATAAAAAATGCATAATACCTTAAATGAAACCGTTAAA
pNB03-darA_r	AATAGCATTCATTTATGGCTCTCCTTTTAAATTTCTGGAAGCTTT
pNB03-darB_f	AAAGCTTCCAGGAAATTTAAAAGGAGAGCCATAAATGAATGCTATT
pNB03-darE_r	CGAAAACCTCACGTTAAGGGATTACGCCGCGATGGTTTGTGTTTTATT
pNB03-clidiss-int_r	TTATGGCTACCGTTCCCTTACTTAAA
pNB03-clidiss-darB2_f	GTAAGGAACGGTAGCCATAATTGCCTGATGAAATTTCTATTGGTT
pNB03-clidiss-darC_f	GTAAGGAACGGTAGCCATAAATGGATATAGAAATCAGAAAAAAGGCATA
pNB03-clidiss-darD_f	GTAAGGAACGGTAGCCATAAATGATTAGCATGATGAATGTCTGTAAATCA
pNB03-clidiss-darE_f	GTAAGGAACGGTAGCCATAAATGGACACAATAATCCCCATAAAAT
pNBDaroMod-darA_r	TAGGTTTATTGCTTAATTCGTTTAGTGCTT
pNBDaroMod-lacsp_f	CGAATTAAGCAATAAACCTAAAGTCTTCTCAGCCGCTACA
pNBDaroMod-lacsp_r	ACCTGATGGGATAAGCTTTAATGTCTTCACCGGTGGAAAG
pNBDaroMod-int_f	TAAAGCTTATCCCATCAGGTTATTT
pRSF-Daro_f*	GTATAAGAAGGAGATATACAATGCATAATACCTTAAATGA
pRSF-Daro_r*	TGCTCAGCGGTGGCAGCAGCTTACGCCGCGATGGTTTGT
DaroB_f	CCTAAGATCCCTGAGATCACGGCCTGGAACTGGACAAAAAGATTCCAGGAAATT
DaroB_r	TTTAAATTTCTGGAATCTTTTGTCCAGTTCCAGGCCGTGATCTCAGGGATCT
DaroC_f	CCTAAGATCCCTGAGATCACGGCCTGGTCATGGTCAAGATCATTCAGGAAATT
DaroC_r	TTTAAATTTCTGGAATGATCTTGACCATGACCAGGCCGTGATCTCAGGGATCT
DaroD_f	CCTAAGATCCCTGAGATCACGGCCTGGAACTGGTCAAGAAGCTTCCAGGAAATT
DaroD_r	TTTAAATTTCTGGAAGCTTCTTGACCAGTTCCAGGCCGTGATCTCAGGGATCT
DaroE_f	CCTAAGATCCCTGAGATCACGGCCTGGTCATGGTCAAAGAGCTTCCAGGAAATT
DaroE_r	TTTAAATTTCTGGAAGCTCTTTGACCATGACCAGGCCGTGATCTCAGGGATCT
DaroA-fol_f	CCTAAGATCCCTGAGATCACGGCCTGGAACTGGTCAAAAAGCTTC
DaroA-fol_r	TTTAGAAGCTTTTGTACCAGTTCCAGGCCGTGATCTCAGGGATCT
DaroB-fol_f	CCTAAGATCCCTGAGATCACGGCCTGGAACTGGACAAAAAGATTC
DaroB-fol_r	TTTAGAATCTTTTGTCCAGTTCCAGGCCGTGATCTCAGGGATCT
DaroC-fol_f	CCTAAGATCCCTGAGATCACGGCCTGGTCATGGTCAAGATCATTC
DaroC-fol_r	TTTAGAATGATCTTGACCATGACCAGGCCGTGATCTCAGGGATCT
DaroD-fol_f	CCTAAGATCCCTGAGATCACGGCCTGGAACTGGTCAAGAAGCTTC
DaroD-fol_r	TTTAGAAGCTTCTTGACCAGTTCCAGGCCGTGATCTCAGGGATCT
DaroE-fol_f	CCTAAGATCCCTGAGATCACGGCCTGGTCATGGTCAAAGAGCTTC
DaroE-fol_r	TTTAGAAGCTCTTTGACCATGACCAGGCCGTGATCTCAGGGATCT
DaroF-fol_f	CCTAAGATCCCTGAGATCACGGCCTGGAAAGTGGTCAAAGAATCTT
DaroF-fol_r	TTTAAAGATTCTTTGACCACTTCCAGGCCGTGATCTCAGGGATCT
DaroA-F_f	CCTAAGATCCCTGAGATCACGGCCTGGAACTGGTCAAAAAGC
DaroA-F_r	TTTAGCTTTTGTACCAGTTCCAGGCCGTGATCTCAGGGATCT
DaroA-SF_f	CCTAAGATCCCTGAGATCACGGCCTGGAACTGGTCAAAA
DaroA-SF_r	TTTATTTTGACCAGTTCCAGGCCGTGATCTCAGGGATCT

Primers for isothermal assembly of the genes were created by outfitting the respective binding region with an artificial GGAGG ribosomal binding site (RBS) 8 random nucleotides upstream the start codon. Additionally, primers were outfitting with ~20 bp (melting temperature of the overlapping region >50 °C) overlap to the respective adjacent fragment. Primers were purchased from Eurofins Genomics (Ebersberg, Germany).

IV.2.4. Cultivation of bacteria

E. coli for cloning purposes were grown in LB broth or agar supplemented with appropriate antibiotics in 1:1000 dilution from the stock solutions at 37°. Cell growth was monitored by measuring the OD₆₀₀ using an Eppendorf BioSpectrometer (Eppendorf AG, Hamburg, Germany) in single use cuvettes with 10 mm light path.

E. coli precultures were grown in 5 mL LB in reaction tubes supplemented with appropriate antibiotics. Larger volume cultivations took place in 100 mL or 300 mL Erlenmeyer flasks. *E. coli* cryocultures were prepared by mixing 1 mL of an over night culture of *E. coli* with sterile glycerol and storing them in -80 °C.

Photorhabdus strains were treated identically to *E. coli*, however incubation temperature was 30 °C.

Pseudoalteromonas was maintained MB or MB agar at 30 °C. Precultures were grown in 20 mL MB and other media were inoculated from the preculture with 0.5 % – 1 % v/v.

IV.2.5. Bioinformatic analysis of Darobactin BGCs

Darobactin BGCs were detected by performing protein BLAST (BLASTp), either using (partial) protein sequences as query entry. Genetic loci were determined by performing tBLASTn, aligning the protein sequence as query to the respective genome. Sequences of ~ 15 kbp up and downstream the genetic loci with (putative) Darobactin BGCs in FASTA format were downloaded from the NCBI database and reannotated using PROKKA, run on a GALAXY server hosted by the JLU Giessen. Reannotated BGCs were analysed for the genetic setup of the BGCs in different bacterial species. DarA and DarE orthologs were aligned using MEGA7. Phylogenetic trees were constructed by aligning the respective protein sequence in Mega7 using the ClustalW algorithm and subsequent calculation of a Maximum Likelihood tree.

IV.2.6. General molecular biology techniques

Plasmid DNA was isolated using the innuPREP plasmid mini kit 2.0 (AnalytikJena, Jena, Germany) according to protocol. Genomic DNA was extracted using the innuPREP bacteriaDNA kit (AnalytikJena, Jena, Germany). PCR amplification for cloning purposes was performed using Q5 DNA polymerase (NEB Biolabs, New Brunswick, USA) using the following setup:

Table 7: Q5-Polymerase PCR setup. GC-Enhancer was added if necessary, the annealing temperature was set to 50°, 60° or 70° and the Elongation time was set calculated to match the fragment size using 45 s for amplification of 1000 nt.

Ingredient	Volume [μl]	Step	Temperature [°C]	Time [s]	
Q5-Buffer	5	1	init. Denaturation	98	60
GC-Enhancer	(5)	2	Denaturation	98	10
dNTPs	0.5	3	Annealing	var	20
primer_f (20 pM)	0.625	4	Elongation	72	var
primer_r (20 pM)	0.625	5	final Elongation	72	600
template DNA	1				
Q5-Polymerase	0.15				
H2O	ad 25				

Test-PCRs were performed using GoTAQ (Promega, Madison, USA) using the following setup:

Table 8: GoTAQ polymerase PCR setup. For colony test PCRs the reaction was filled with H2O and a small portion of the tested colony was dispersed in the PCR mixture. Elongation time was calculated with 60 s for amplification of 1000 nt.

Ingredient	Volume [μl]	Step	Temperature [°C]	Time [s]	
Green GoTAQ buffer	4	1	init. Denaturation	98	60
MgCl ₂	1	2	Denaturation	98	10
DMSO	1	3	Annealing	var	20
dNTPs	0.33	4	Elongation	72	var
primer_f (20 pM)	0.33	5	final Elongation	72	600
primer_r (20 pM)	0.33				
template DNA	0.5				
GoTAQ	0.1				
H2O	ad 20				

Restriction was performed using standard techniques and NEB enzymes (NEB Biolabs, New Brunswick, USA), DNA fragments were analysed on 1% or 2% TAE-Agarose gels using MidoriGreen as loading dye and stain and a uv-free blue light-transilluminator. GeneRuler 1kb Plus (ThermoFisher, Waltham, USA) was used as marker. DNA for cloning purposes was excised from the gel and purified using the Zymoclean large fragment DNA recovery kit according to manufacturers instruction. DNA concentrations were determined with an Eppendorf BioSpectrometer (Eppendorf AG, Hamburg,

Germany) using a 1 mm light path cuvette. DNA Fragments to be fused by isothermal assembly were gel purified and fused using self-made isothermal assembly master mix (Gibson et al., 2009) using NEB enzymes (NEB Biolabs, New Brunswick, USA). In brief, 15 µL of isothermal assembly master mix were thawed on ice and equimolar concentration of all fragments were mixed in 5µL and added to the master mix. Reaction mixtures were incubated at 50 °C for 1 h. Prior to transformation of *E. coli* with assembled DNA, the DNA was desalted by drop dialysis using MF-Millipore VSWP membranes.

E. coli were transformed with DNA using standard electroporation methodology; In brief: *E. coli* cells were grown in 50 mL LB at 37 °C at 30 °C for several hours to an OD₆₀₀ of ~0.5. Cells were harvested by centrifugation at 10000 rcf for 3 min and washed twice with ice cold 10% glycerol. After the final washing step, the resulting pellet was resuspended in the return flow and aliquoted to 50 µL. DNA was mixed with 50 µL of competent cells, added to ice cold 2 mm BioRad electroporation cuvettes and an electric pulse of 2.5 kV was applied for ~ 5 s using a BioRad MicroPulser (BioRad, Hercules, USA). The electroporation mixture was taken up in 800µL LB medium and incubated at 37 °C for 1h. Subsequently, transformants were plated on LB plates containing appropriate antibiotics.

IV.2.7. LCMS analysis of Darobactin and Darobactin analogue containing samples

LCMS based detection of Darobactin A and its derivatives was performed on various LCMS systems, depending on the quality needed. Lower (relative) accuracy measurements were performed on a Dionex Ultimate3000 (Thermo Scientific, Darmstadt, Germany) using an EC10/2 Nucleoshell C18 2.7µm column (Macherey-Nagel, Düren, Germany) HPLC coupled to a microTOFq II (Bruker) ESI-qTOF-HRMS. The LC system was run in the following gradient (A:H₂O; B: MeOH, ; Flow: 200 µL/min): 0-5 min 90% A, 5-35min 90%-0% A, 35-50min 100% B, 50-60min 90% A.

Higher (relative) accuracy measurements and targeted fragmentation analysis experiments were performed on an Agilent 1290 UPLC system with an Acquity UPLC BEH C18 1.7 µm (2.1x100 mm) and Acquity UPLC BEH C18 1.7 µm VanGuard Pre-Column (2.1x5 mm) column setup coupled to DAD and ELSD detectors and a maXis II (Bruker) ESI-qTOF-UHRMS. The LC part was run in a gradient (A:H₂O, 0.1% FA; B: MeCN, 0.1% FA; Flow: 600 µL/min): 0 min: 95%A; 0.30 min: 95%A; 18.00 min: 4.75%A; 18.10 min: 0%A; 22.50 min: 0%A; 22.60 min: 95%A; 25.00 min: 95%A and the column oven was set to 45°C.

During the experiments, the LC system of the microTOFq II was exchanged with an Agilent 1290 UPLC system, mimicking the LC part of the maXis II UHRMS system.

LCMS samples from culture supernatant if not stated otherwise, were desalted and concentrated ~ 6 times using C18 stage tips before LCMS analysis. Stage tips were washed with 200 µL 100 % MeOH and equilibrated using 200 µL 95:5 H₂O/MeCN + 0.1 % FA and centrifugation at 3000 rcf. Subsequently, 400 - 500 µL of cleared supernatant were applied to the stage tips and the C18 matrix was washed with 30 µL 95:5 H₂O/MeCN + 0.1 % FA to remove salts. Finally, the samples were eluted using two times 30 µL 20:80 H₂O/MeCN + 0.1 % FA and 5 µL of eluate were injected

For LCMS analysis of the pellet, cells were harvested by centrifugation and the pellet was resuspended in 10 % of the harvested volume 20:80 H₂O/MeCN + 0.1 % FA. The pellet was kept in an ultrasound bath for 10 min, cell debris was removed by centrifugation (max speed 10 min) and 5 µL of extract were injected into the LCMS.

IV.2.8. Interspecific conjugation between *E. coli* WM3064 and *Photorhabdus khanii* DSM3369

E. coli WM3064 cells were transformed with the vector by electroporation and correct assembly was corroborated by PCR. Conjugation between *E. coli* WM3064 and *P. khanii* DSM3369 was performed by growing both strains to an OD₆₀₀ of ~ 0.6. After washing twice with LB, the cells were mixed in 1:3 ratio of *E. coli*/*P. khanii*, plated out on LB agar supplemented with diaminopimelic acid (0.3 mM) and incubated at 37° C for three hours, followed by overnight incubation at 30° C. The bacterial lawn was resuspended in LB and plated on LB agar with kanamycin (50 µg/mL) in serial dilution. Transconjugants were plated on fresh LB plates containing appropriate selective antibiotics.

Triparental conjugation between *P. khanii* DSM3369 $\Delta darABCDE$, *E. coli* Top10 carrying expression plasmids and *E. coli* ET12567 + pUB307 was performed in the same way, however the recipient strain and both *E. coli* strains were mixed in 3:1:1 ratio.

IV.2.9. Genetic inactivation of DarE in *Photorhabdus khanii* DSM3369

In a first attempt to generate a precise $\Delta darABCDE$ mutant, we tried to electroporate linear DNA into *P. khanii* DSM3369 hoping to directly select double crossover mutants. In brief, the upstream and downstream regions of the cluster were amplified and fused upstream and downstream the *aph* gene conferring kanamycin resistance from pCAP03 (Tang *et al.*, 2015) combined with the beta lactamase (*bla*) promoter region from pSAM_Bt (Goodman *et al.*, 2009). This construct was PCR amplified and electroporated into *P. khanii* DSM3369 using various settings, however we were unable to generate any kanamycin resistant colonies using this strategy.

Instead, we employed conjugational transfer for genetic modifications in *P. khanii* DSM3369. Chromosomal DNA of *P. khanii* DSM3369 was isolated and *darE* was partially amplified using the primers pNB01-ins_f and pNB01-ins_r. The R6K replicon and oriT region from pSAM_Bt was amplified by pNB01-R6K-Ori_f and pNB01-R6K-Ori_r, the *aph* gene together with the *bla* promoter from the previously described linear construct was amplified with pNB01-Kan_f and pNB01-Kan_r. The R6K replicon does not replicate in the absence of λ pir, ensuring no extrachromosomal replication can take place in *P. khanii* DSM3369. All fragments were amplified using Q5 polymerase and gel purified. The fragments were fused by isothermal assembly, creating the vector pNB01.

After assembly, *E. coli* WM3064 was transformed with the vector and the vector was conjugated to *P. khanii* DSM3369 and transconjugants with disrupted *darE* ORF were selected on LB_{Kan}. Subsequently, *darE* disrupted mutants as well as *P. khanii* DSM3369 WT were grown in LB_{Kan} and LB respectively for 3 days and analysed by LCMS.

IV.2.10. Markerless deletion of *darABCDE* in *Photorhabdus khanii* DSM3369

Chromosomal DNA of *P. khanii* DSM3369 was isolated and roughly one kb upstream and downstream of the BGC were amplified using the primer pairs pNB02-upstr_f/r and pNB02-dwnstr_f/r. The R6K origin of replication (ori), the origin of transfer (oriT) and the levansucrase gene *sacB* lacking the promoter region from *Bacillus subtilis* was amplified in one piece from the vector pNPTS138 (Lassak *et al.*, 2010) using the primer pair pNB02-pNPTS_f/r. The arabinose inducible expression system of pKD46 (Datsenko & Wanner, 2000) with the adjacent (*bla*) promoter was amplified by the primer pair pNB02-blapr_f/r and the *aph* gene from pCAP03 (Tang *et al.*, 2015), conferring resistance to kanamycin was amplified using the primer pair pNB02-aph_f/r. All fragments were amplified with Q5 DNA polymerase and gel purified. Subsequently all fragments were fused by isothermal assembly, generating the plasmid pNB02.

In the vector pNB02, the *aph* gene was placed under control of the *bla* promoter and the *sacB* gene under the control of the inducible *araB* promoter, ensuring the functionality of the selection systems in *P. khanii* DSM3369. The *sacB* gene is frequently used as a suicide selection marker, intracellularly producing the polysaccharide levan from sucrose. The oriT enabled conjugational transfer of pNB02 and the R6K replicon does not replicate in the absence of λ pir, ensuring no extrachromosomal replication can take place in *P. khanii* DSM3369.

After assembly, *E. coli* WM3064 cells were transformed with the vector by electroporation and correct assembly was corroborated by PCR and restriction following standard procedures and pNB02 was transferred to *P. khanii* DSM3369 by interspecific conjugation. Kanamycin resistant single cross

over transconjugants were grown in LB to an OD₆₀₀ of ~ 0.6, the expression of SacB was induced by adding arabinose to a concentration of 0.2% (w/v) and incubated for two hours. Subsequently the culture was plated out on LB agar supplemented with 0.2% (w/v) arabinose and 10% sucrose and incubated at 30° C for 48 hours. Single colonies were selectively picked on LB_{Kan} and LB_{Ara/Suc} agar. Sensitivity to kanamycin indicated the loss of the plasmid and therefore a successful double crossover event. Double crossover mutants were picked to fresh LB agar and screened for the loss of the gene cluster by PCR using the primer pair DCOscr_f/r.

IV.2.11. Construction of darA-E expression plasmids pNB03 – darA-E and pNBDaro

In a further experiment the Darobactin BGC was (re)introduced into *P. khanii* DSM3369 $\Delta darABCDE$ and *E. coli* BW25113 on the *Photorhabdus/E. coli* expression plasmid pNB03. To avoid issues with the regulation system between the propeptide and the modifying enzymes, two constructs were created : i) one where all intergenic regions were removed and the genes *darA-darE* expressed streamlined under the control of the arabinose inducible araB promoter and ii) one with the intergenic region intact, essentially mimicking the natural layout of the BGC with darA under the control of the inducible araB promoter.

pNB03 was created by amplification of the p15A ori from pACYC177 using the primers pNB03-ori_f/r, the arabinose expression system and kanamycin resistance of pNB02 using the primers pNB03-araC_f/r and the oriT and the *aac(3)* gene conferring resistance to apramycin from pIJ773 (Gust *et al.*, 2003) using the primer pair pNB03-Aprar_f/r. Subsequently all fragments were gel purified and assembled as described previously. After assembly, *E. coli* TOP10 cells were transformed with the vector and correct assembly was corroborated by test restriction.

To introduce the Darobactin BGC without intergenic region between darA and darB to the plasmid pNB03-darA-E, pNB03 was linearised using the primers pNB03_f/r, *darA* was amplified using pNB03-darA_f/r, *darB – darE* were amplified using pNB03-darB_f/pNB03-darE_r. For introduction of the native BGC on the plasmid pNBDaro, the vector was linearised and the entire cluster was amplified using pNB03-darA_f and pNB03-darE_r. All fragments were gel purified and assembled as described previously, both vectors were transferred to *E. coli* TOP10 cells and correct assembly was corroborated by test restriction.

IV.2.12. Homologous and heterologous overexpression of the Darobactin BGC in *P. khanii* DSM3369 Δ darABCDE and *E. coli* BW25113

The empty pNB03 as well as pNB03-darABCDE and pNBDaro were transferred to *P. khanii* DSM3369 Δ darABCDE by triparental conjugation. Since *P. khanii* DSM3369 is naturally resistant to carbenicillin and the kanamycin resistance of pUB307 lacks the bla promoter, final selection took place on LB agar supplemented with kanamycin and carbenicillin. Kanamycin resistant transconjugants were grown in LB_{Kan}, the plasmid was isolated and the identity of the transferred plasmid was verified by PCR using pNB01-Kan_f/r. Additionally, *E. coli* BW25113 was transformed with the plasmids pNBDaro and pNB03 – darA-E and transformants were selected on LB_{Kan}.

Subsequently, *P. khanii* DSM3369 WT, *P. khanii* DSM3369 Δ darABCDE + pNB03 (empty vector) and *P. khanii* DSM3369 Δ darABCDE + pNB03-darA-E, *P. khanii* DSM3369 Δ darABCDE + pNBDaro, *E. coli* BW25113 + pNB03-darA-E and *E. coli* BW25113 pNBDaro were grown in LB and LB_{Kan} supplemented with 0.2 % (w/v) arabinose respectively for 3 days and analysed by LCMS.

IV.2.13. Identification of the minimal Darobactin BGC

For the identification of the minimal Darobactin BGC, the expression plasmid pNB03 – DaroBGC was shortened, leaving out the ORFs of the transporter genes darB,C,D one after another resulting in four new expression plasmids, each one lacking one more transporter gene. The pNB03 backbone was linearised using the primers pNB03_f/r, darA together with the intergenic region was amplified using darA_f and pNB03-cliss-int_r. darB2 to darE, darC to darE, darD to darE and darE alone were amplified by the primers pNB03-cliss-darB2_f, pNB03-cliss-darC_f, pNB03-cliss-darD_f and pNB03-cliss-darE_f, always using pNB03-darE_r as reverse primer. The backbone pNB03, darA and the adjacent intergenic region and the shortened residual BGC parts respectively were fused by isothermal assembly, creating the vectors pNBDaro- Δ darB1, pNBDaro- Δ darB, pNBDaro- Δ darBC and pNBDaro- Δ darBCD. Subsequently, *E. coli* BW25113 was transformed with the constructs and correct assembly was corroborated by restriction. Shortened versions of the BGC were expressed as described before. Darobactin A concentration in the supernatant and in the pellet was analysed by LCMS.

IV.2.14. Mutasynthetic production of Darobactin analogues B,C,D,E,F by modification of WT Darobactin A BGC

The goldengate plug and play vector pNBDaroMod for easy modification of the propeptide was assembled essentially using the same primers for the creation of pNBDaro, however the last 50 nt of the propeptide before the stop codon were replaced with a lacZ spacer flanked by two BbsI restriction sites from pCRISPOMYCES-2 (Cobb *et al.*, 2015) allowing scarless incorporation of annealed oligonucleotides into the expression system. Therefore pNB03 was linearised with the primers pNB03_f/r, the truncated darA ORF was amplified by with pNB03-darA_f/pNBDaroMod-darA_r, the lacZ spacer was amplified from pCRISPOMYCES-2 with pNBDaroMod-lacsp_f/r and the rest of the BGC was amplified using pNBDaroMod-int_f and pNB03-darE_r. The fragments were fused by isothermal assembly and transferred to *E. coli* Top10 cells. Subsequently, the identity of the plasmid was corroborated by positive blue white selection and test restriction.

Darobactin analogues were created by modification of the Darobactin A propeptide to different core amino acids by in silico exchanging respective codons to the desired amino acid. Sequences were ordered as complementary primers with 4 nt sticky ends to the BbsI restriction sites to pNB03-ADCmod. To replace the core amino acids by GoldenGate cloning, we followed the instructions outlined in Cobb *et al.*, 2015. In brief, oligonucleotides were annealed by mixing 5 µL 100 µM complementary oligos each in 90 µL 30 mM HEPES buffer p.H. 7.8, heating the mixture to 95 °C for 5 min and ramping the temperature down to 4 °C at 0.1 °C/s. The annealed oligos were subsequently assembled into pNBDaroMod using the following reaction mixture and thermocycler setup.

Table 9: Reaction mixture and thermocycler setup for GoldenGate assembly of annealed oligonucleotides into pNBDaromod

Ingredient	Volume µL	Step	Temp [°C]	Time [s]	
Backbone	X (100ng)	1	37	10	9x to 1
Insert	0.3	2	16	10	
T4 ligase buffer	2	3	50	5	
T4 ligase	1	4	65	20	
T4 polynucleotide kinase	1	5	4	store	
BbsI	1				
H2O	ad 20				

Subsequently, *E. coli* BW25113 was transformed with 2 µL dialysed the reaction mixture. Correct assembly was corroborated by sequencing or chemical profiling by LCMS.

The following primers were used to mutate the Darobactin A score amino acid sequences to the respective analogue sequences (Tab.: 10):

Table 10: List of oligonucleotide pairs used for GoldenGate assembly into pNBDaromod. + or – QEI indicates presence or absence of the QEI follower AAs present in the WT DarA protein sequence. -F7 and -S6F7 indicates lack of the Darobactin A core AAs in position 6 and 7 respectively.

Darobactin analogue sequence	Oligo pair
B + QEI follower	DaroB_f/r
C + QEI follower	DaroC_f/r
D + QEI follower	DaroD_f/r
E + QEI follower	DaroE_f/r
A - QEI follower	DaroA-fol_f/r
B - QEI follower	DaroB-fol_f/r
C - QEI follower	DaroC-fol_f/r
D - QEI follower	DaroD-fol_f/r
E - QEI follower	DaroE-fol_f/r
F - QEI follower	DaroF-fol_f/r
A - F7	DaroA-F_f/r
A - S6F7	DaroA-SF_f/r

Darobactin BGCs containing mutated propeptides were expressed using the same conditions used for the heterologous expression of Darobactin A in *E. coli* BW25113 and production of Darobactin analogues was monitored by LCMS after three days.

To increase the amounts of analogues produced for analysis of the MS/MS fragmentation pattern, the mutated clusters were recloned into the high copy number expression plasmid pRSF-duett under control of the strong T7 promoter. pRSF-duett was digested with NdeI and AvrII, respective clusters were amplified using the primers pRSF-Daro_f/r and fragments were fused by isothermal assembly. Subsequently, Darobactin resistant *E. coli* BAP1mut was transformed with the resulting constructs and correct assembly was corroborated by chemical profiling by LCMS.

IV.2.15.MS/MS based characterisation of Darobactin analogues

E. coli Bap1mut carrying the Darobactin BGC as well as the respective analogue BGCs were grown in 20 mL LB_{Kan}/IPTG for 3 days. Cells were removed by centrifugation. A solid phase extraction column was packed with 0.5 g of C18 derivatised silica, washed with 10 CV 100% MeOH and equilibrated with 10 CV H₂O + 0.1 % formic acid. Subsequently, the cleared supernatant was applied to the column and the matrix was washed with 2 CV H₂O + 0.1 % formic acid. Material bound to the C18 matrix was eluted using 10 CV 80:20 MeCN/H₂O + 0.1 % formic acid and the eluate was dried in a GeneVac SpeedVac system (SPScientific, Ipswich, UK) over night. Dried concentrate was dissolved in 200 µL H₂O + 0.1 % formic acid analysed on LCMS. For analysis, 5 µL of the concentrate were injected and the respective mass was specifically targeted for MS/MS fragmentation.

IV.2.16. Isolation of Darobactin and its derivatives

Darobactin analogue producing *E. coli* strains were grown in LB containing appropriate antibiotics and IPTG in 1 L per flask in 5 L or 2 L flasks. Cells were separated from the medium and extracted with 80% MeCN. Crude extract was dried in vacuo and separated by FLASH medium pressure liquid chromatography using a C18 F0120 column with a H₂O:MeCN gradient (0 min – 28 min 5% MeCN, 28 min – 37 min 5 % MeCN – 15 % MeCN, 37 min – 50 min 15 % MeCN, 50 min – 60 min 15 % MeCN – 30 % MeCN, 60 min – 80 min 30 % - 100 % MeCN). Fractions were collected along the UV chromatogram, analysed by LCMS and Darobactin (analogue) containing fractions were further fractionated on HPLC run in a gradient (0 min - 5 min 15 % MeCN, 5 min - 25 min 15 % - 25 % MeCN, 25 min - 30 min 25 % - 60 % MeCN, 30 min – 39 min 60 % - 100 % MeCN). Darobactin (analogues) were collected as pure compounds in a last HPLC run using a gradient of 0 min – 5 min 10 % MeCN, 5 min - 31 min 10 % - 23.5 % MeCN, 31 min – 38 min 100% MeCN and weight out in glass vials.

Darobactin B and Marinodarobactin A and B producing strains were isolated from the respective medium. Therefore, Darobactin B producing *E. coli* and *Pse. luteoviolacea* H33 were cultivated in LB_{Kan/IPTG} and Medium 2 respectively. Cleared supernatant was passaged over a 4L bed volume column packed with XAD-16N and subsequently washed with H₂O. Compounds were eluted with 80 % MeOH + 0.1 % FA and the MeOH was evaporated in a rotary evaporator. The MeOH free elution fraction was loaded to a SP Sepharose XL strong ion exchange column (220 mL bed volume, GE healthcare, Chalfont St Giles, UK) and subsequently eluted with 10 CV 50 mM NH₄Ac with pH. 3, 5, 7, 9 and 11. Darobactin analogue containing fractions (LCMS) were pooled, lyophilised and purified on HPLC, using the program mentioned above.

IV.2.17. Detection of natural Darobactin analogues Marinodarobactin A and B and

Didehydrodarobactin A from *Pseudoalteromonas luteoviolacea* H33 and H33s

In order to detect novel Darobactin analogues from *Pseudoalteromonas* strains a quasi OSMAC approach was performed. Eleven media (MB + M1 – M10) were inoculated from overnight cultures in MB and grown for 3 days. Subsequently, the cells were removed by centrifugation and the supernatant was concentrated as described previously, using a 0.5 g C18 SPE column and subsequent evaporation in a GeneVac SpeedVac system (SPScientific, Ipswich, UK). The concentrate was dissolved in 200 µL H₂O + 0.1 % formic acid and antimicrobial activity was tested with *E. coli* MG1655 and *E. coli* MG1655 BamA6 with 20 µL, 10 µL and 5 µL of each generated crude extract. This was repeated with the 4 extracts with the highest activity, this time running an extract of uninoculated medium as negative control. Relative inhibitory activity was calculated as inh. % = 100 * [1 - (AU end - AU start) / (AU neg. ctrl - AU start)].

The crude extracts were subsequently subjected to LCMS for GNPS molecular networking analysis. LCMS chromatograms were converted to .mzXML format using Compass DataAnalysis and uploaded to the GNPS server using FileZilla. GNPS molecular networking was run, setting empty media as one group, Semi purified Darobactin A from a previous experiment as another. *Pseudoalteromonas* crude extracts were then distributed to the other groups based on media they were grown in. Molecular networking was run under standard conditions and subsequently rerun under relaxed conditions by setting the number of matching fragments to 2 (Standard: 6). GNPS molecular networks were analysed using CytoScape (ISB, Seattle, USA).

IV.2.18.Determination of MIC values of Darobactin A and Darobactin analogues

Test panel for the determination of MIC values for Darobactin derivatives consisted of *E. coli* ATCC35218, *E.coli* NRZ14408 KPC-2, *E.coli* K0416 VIM-1, *E. coli* Survcare 052 NDM-5, *E. coli* MMGI1 OXA-48, *Pseudomonas aeruginosa* PAO 1, *Pseudomonas aeruginosa* PAO 750, *Klebsiella pneumoniae* DSM30104, *Acinetobacter baumannii* ATCC19606 and *Salmonella enterica* ATCC13076. For determination of MIC values, test strains were grown over night and fresh cation adjusted Mueller Hinton broth II was inoculated to 5×10^5 cells. Pure compound was dissolved in H₂O to 6.4 mg/ml and growth inhibition was tested in 100 µL medium in 96 well microtiter plates at 37° C and 180 rpm measuring turbidity (OD₆₀₀) at the start of the experiment and after ~ 18 h using a LUMIstar®Omega BMG Labtech plate reader. Darobactin derivatives were dissolved to a concentration of 12.8 g/L or 6.4 g/L respectively and tested in triplicate with 12 dilution steps and concentrations from 64 µg/mL – 0.031 µg/mL. Uninoculated medium and antibiotic free inoculated medium were used to define the lower and upper boundary of the turbidity range, Rifampicin, Tetracycline, Colistin and Gentamycin were used as control antibiotics. Relative inhibition was calculated as $\text{rel. inh. \%} = 100 * [1 - (\text{AU sample} - \text{AU Low}) / (\text{AU High} - \text{AU Low})]$.

IV.3. Results

IV.3.1. Identification of the Darobactin BGC

After discovery of Darobactin and its characterisation as an antimicrobial compound, further investigation of its biosynthetic origin i.e. identification of the corresponding BGC was necessary. Since all AA positions were proteinogenic, a ribosomal origin and subsequent modification was inferred. First tBLASTn search of the AA sequence WNWSKSF against the genome of the first identified Darobactin producer *P. khanii* HGB1456 did not yield any hit. However, NCBI Blastp search of the Darobactin A core amino acids against genes exclusively from the *Photorhabdus* species revealed a small hypothetical protein of 68 amino acids with the core AAs WNWSKSF at the C - terminus, followed by the AAs Q, E and I in a number of deposited sequences (Fig.: 15). Closer investigation of the genomic context of this gene in the reported hits revealed a small unidirectional BGC of ~6500 nt, consisting of 5 individual ORFs with a little stretch between the ORF of the small first hypothetical protein and the following ORF (Fig.: 15). This (putative) RiPP BGC is bordered upstream by a pyruvate kinase (*pyrK*) gene and downstream by a lauroyl-Kdo(2)-lipid IV(A) myristoyltransferase (*msbB*) gene facing the opposite direction. While the first gene could not be assigned any function based on sequence homology, the function of the other genes could be readily identified as coding for ABC transenvelope transport proteins (ORF 2-4) and for a member of the radical SAM (RaS or rSAM) superfamily of enzymes (ORF 5). These 5 ORFs were provisionally termed *darA*, *darB*, *darC*, *darD* and *darE*. Subsequent resequencing and reinvestigation of the genomes of the identified producer strain *P. khanii* HGB1456 verified it carrying this BGC. Investigating the genetic loci of the adjacent genes in members of the genus *Photorhabdus* and closely related *Xenorhabdus* strains revealed that only a subset of *Photorhabdus* strains carry the BGC while other *Photorhabdus* strains and all *Xenorhabdus* strain do not. Comparison between *Photorhabdus/Xenorhabdus* strains with and without the putative BGC, showed that where the BGC is missing, *pyrK* and *msbB* genes are in direct neighbourhood, indicating that the putative Darobactin A BGC is limited to the five identified genes *darA* – *darE* (Fig.: 15). This is backed by the observation that the overall GC content of the hypothetical BGC is relatively low with 33 % and differs vastly from the adjacent *pyrK* and *msbB* genes with ~ 50 %. Surprisingly, while the BGC in the Darobactin producing strain *P. khanii* DSM3369 is nearly identical in nucleotide sequence to the one from *P. khanii* HGB1456, in *P. khanii* DSM3369 the *darB* ORF is split into two separate ORFs. Closer investigation revealed a point mutation, introducing a TAA stop codon 763 nt into the *darB* ORF closely followed by an in-frame methionin (start) codon, effectively splitting the *darB* ORF into two separate ORFs with 100 % identity to their respective parts of the intact *darB* from *P. khanii* HGB1456 (Fig.: 15). Reannotation of the putative BGCs from all *Photorhabdus* genomes deposited in

GenBank additionally revealed a second small ORF upstream *darA* in the four *Phototrhhabdus* strains *Phototrhhabdus temperata* ssp. *thracensis* DSM 15199, *Phototrhhabdus heterorhabditis* VMG, *Phototrhhabdus asymbiotica* ATCC43949 and *Phototrhhabdus australis* PB68.1 (Fig.:) with strong (~ 70 %) overall similarity to the respective *darA* but with WNWTKRF (3 sequences, provisionally termed Darobactin B) and WKWSKNL (1 sequence, *P. temperata* DSM 15199) as core amino acid sequence followed by P and I as a follower, hinting towards existence of natural analogues of Darobactin, coded in a separate propeptides.

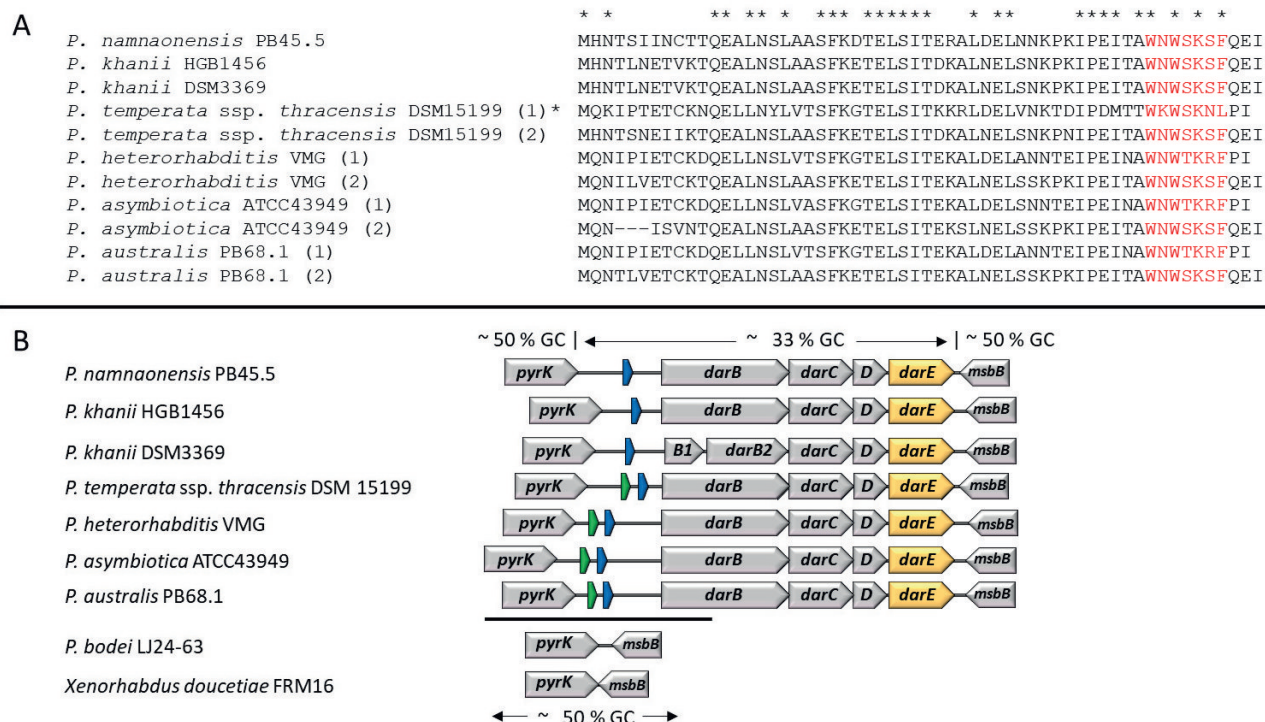


Figure 15: A) Alignment of putative translated Darobactin propeptide AA sequences from various *Phototrhhabdus* strains (core AA sequences highlighted in red). Some *Phototrhhabdus* strains carry two propeptides with different core AA sequence and follower AAs (indicated by nr. 1). The propeptide marked * was not incorporated into the alignment due to severe aberration in the core AAs as well as in overall sequence, possibly hinting towards a degenerated version of the propeptide. B) Putative Darobactin BGCs from various *Phototrhhabdus* strains in their genomic context compared to the genomic context of *P. bodei* LJ24-63 and *X. doucetiae* FRM16, indicating the borders of the putative BGC. Putative propeptides are coloured green and blue, *darE* is coloured yellow. Additionally, the difference in GC-content between the putative BGC and the adjacent genes is highlighted.

BLASTp on the translated amino acid sequence of *darE* revealed additionally to the already identified *darE* genes in *Phototrhhabdus* > 100 ortholog genes with high coverage (> 90 %) and moderate to high identity (> 50 %). In the *Vibrio* (3 hits), *Pseudoalteromonas* (9 hits) and *Yersinia* (> 100 hits) genera, all belonging to the Gammaproteobacteriacea, orthologs to *darE* and subsequently also orthologs to all other genes of the putative Darobactin BGC were found. Constructing a phylogenetic tree on genus level, using all identified *darE* and ortholog AA sequences unsurprisingly shows that individual sequences cluster by species. In the grand scheme however, it revealed that *darE* orthologs from

Vibrio and *Pseudoalteromonas* species form a distinct clade with *darE* from *Phototrhhabdus* forming a direct sister clade and *darE* orthologs from *Yersinia* species branching off more remote (Fig.: 16). This however is not reflected in the overall architecture of the BGCs. While putative Darobactin A BGCs in *Phototrhhabdus* and *Yersinia* strains show sequence dissimilarities in the respective enzymes, they maintain the exact architecture with orthologs to all five genes from the first identified BGCs (Fig.: 16) alongside a similarly low GC content of ~ 33 %. ORFs split by point mutations similar to *P. khanii* DSM3369 can also be found in *Yersinia* strains. In contrast, BGCs in *Vibrio* and *Pseudoalteromonas* species have a higher average GC content of ~ 40 % and slightly differ in genetic architecture. While the putative BGCs in *Vibrio* species essentially follow the same architecture as the ones in *Phototrhhabdus* and *Yersinia*, the BGCs are extended by one more ~ 1700 nt gene for an ABC type transenvelope transporter downstream of *darE* that does not have a match in the aforementioned species BGCs (Fig.: 16). The BGCs in *Pseudoalteromonas* species essentially share the same architecture as the ones found in *Vibrio*, however one additional gene of 1065 nt is located upstream of *darB* in between *darA* and *darB*. BLASTp on the translated amino acid sequence of this gene identified a HEXXH motif located on the C – terminal side of the protein, usually indicating the presence of a RaS SPASM domain closer to the N - terminus. In this case however, neither a full SPASM domain nor the CXXXCXXC motif, indicative of a RaS enzyme can be found. Additionally, a small subset of the *Pseudoalteromonas* population carrying the BGC carries another gene downstream the *Vibrio/Pseudoalteromonas* exclusive ABC transporter gene. This gene was annotated as FAD dependent oxidoreductase based on sequence homology (Fig.: 16). Regardless of species no peptidases for the final trim of the molecule were found within the BGCs.

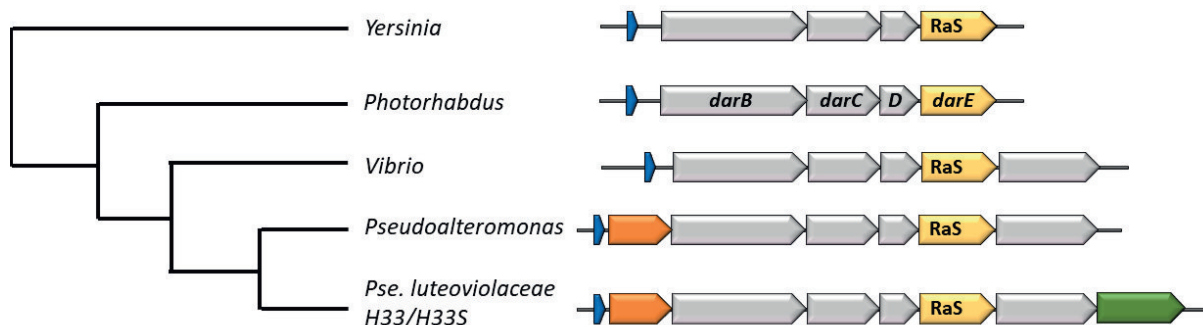


Figure 16: Topological maximum likelihood phylogenetic tree of *DarE* and its orthologs from *Yersinia*, *Vibrio* and *Pseudoalteromonas* and the respective BGC architecture. Blue: propeptides; Grey: transport related genes; Yellow: *darE* orthologs; Orange: gene of unknown function (ann. hypothetical protein); Green: gene of unknown function (ann. FAD dependent oxidoreductase).

Closer investigation of *darA* and its orthologs from all genera showed that only certain *Phototrhhabdus* BGCs carry a separate propeptide with altered AA sequence. In all other genera, only one propeptide ORF per BGC could be found. Aligning all identified *darA* orthologs from all genera revealed that

propeptide AA sequences essentially differ by genus, however maintaining a high homology within the genus boundaries for the genera *Photobacterium*, *Vibrio* and *Pseudoalteromonas*. Comparison of *Vibrio* and *Pseudoalteromonas* propeptides additionally reveals stunningly high homology between all sequences from both genera and no aberration from the Darobactin A core AA sequence at all. (Tab.: 11).

Table 11: Alignment of translated propeptide AA sequences from different genera, indicating the type of propeptide classified among core AA sequence (highlighted in red). Darobactin F propeptide was not incorporated into the alignment due to severe aberration of core AA and general AA sequence to other *Photobacterium* propeptides, hinting towards a degenerated propeptide.

Strain	Darobactin type	Sequence
		* ** * ** ***** *** ***** ** * ***** * *
<i>P. khanii</i> DSM3369	A	MHNTLNETVKTQEALNSLAASFKEITELSDKALNELSNKPKIPEITAWNWSKSFQEI
<i>P. heterorhabditis</i> VMG (1)	B	MQNIPETCKDQELLNSLVTSFKGTELSITEKALDELANNTEIPEINAWNWTGRFPI
<i>P. heterorhabditis</i> VMG (2)	A	MQNILVETCKTQEALNSLAASFKEITELSDKALNELSSKPKIPEITAWNWSKSFQEI
<i>P. temperata</i> DSM15199 (1)	(F)	MQKIPTETCKNQELNLYLVTSFKGTELSITKKRLDELVNKTDIPDMTTWWSKSNLPI
<i>P. temperata</i> DSM15199 (2)	A	MHNTSNEIIKTQEALNSLAASFKEITELSDKALNELSNKPKIPEITAWNWSKSFQEI
		***** ***** * ** * *****
<i>V. vulnificus</i> CG100	A	MIIVEKEKVSISEKLDALMSSFSEMNLELRKFDQEKVNSINIAPPITAWNWSKSF
<i>Pse. luteoviolacea</i> H33	A	MIVEAPKEKVSISEKLDALKSSFSNQLNINAVDQARVDSISVAPPITAWNWSKSFQEI
		* * * * * * * * *****
<i>Y. enterocolitica</i>	A	MFTSNQSNERNINTHLMALKTKLESLEQSFKNLFSINDHEIENLRRSNSNNQITAWNWSKSFQQ
<i>Y. enterocolitica</i>	D	MYTSHQSDLNTHNGKLIKTKLEALDESFNNSLHISYDEIEKIKNNLSKSKITAWNWSRSFAEE
<i>Y. bercovieri</i>	A	MYTSHHTDRKTSNSNLMALKAKLESLEQSFKNLFSINDHEIENLKNNNFNNEITAWNWSKSFQQ
<i>Y. bercovieri</i>	E	MYTSHHTDRKTSNSNLMALKAKLESLEQSFKNLFSINDHEIENLKNNNFNNEITAWWSKSFQQ
<i>Y. frederiksenii</i>	A	MYTSHQPEKTSNTNLIALRTKLESLEESFKNSGLSIDAQEIENLKNSESENKITAWNWSKSFQQ
<i>Y. rhodei</i>	A	MYTSRNPGEIIPSNIMALLTKGLSLEESFKNNALTNNEIENLKNSEVNNKITAWNWSKSFQQ
<i>Y. aldovae</i>	A	MYNSTSHQKSHSVNNATALRSKLLSLQESFKSIPIHININKIEDLINSKSNKITAWNWSKSFQD
<i>Y. aldovae</i>	D	MFTSNQSNERNINNAHLMALKAKLESLEESFKNNLTHISDNEIEKIKSNTLRSKITAWNWSRSFAEE
<i>Y. massiliensis</i>	A	MSISFQHQRKNNQNLIALKSKLQSLGESFSHSLYISNSELDKIRNSLAKTKITAWNWSKSFQEN
<i>Y. pseudotuberculosis</i>	D	MNPSSQSTVEKNNVNLKLSKLSLEESFKNNPLYITSNEIDEVKNNTLHTKITAWNWSRSFAED
<i>Y. pseudotuberculosis</i>	C	MNPSSQSVVEKSNVNLKLSKLSLEESFKNNPLYITSNEIDEIKNNLTHSKITAWWSRSFAED

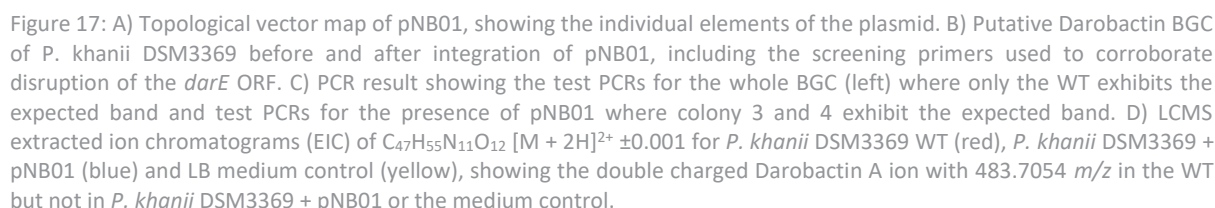
This contrasts with propeptides from *Yersinia* strains. While certain propeptide varieties from closely related strains may have high identity in AA sequence to each other, the pool of *Yersinia* propeptide sequences has only little overall homology (Tab.: 11).

It was also observed that the vast majority of propeptides maintains the Darobactin A core AA sequence (WNWSKSF) with only few exemptions. WWSRSF (provisionally termed Darobactin C) is carried by a small set of *Yersinia pseudotuberculosis* (6) and *Yersinia pestis* (4) genomic sequences, WWSRSF (provisionally termed Darobactin D) can be found in 28 sequences from various *Yersinia* species and WWSKSF (provisionally termed Darobactin E) can be found in two sequences from

Yersinia bercovieri (Fig.: 16). Constructing a tree using aligned *Yersinia* propeptide sequences shows that most propeptides with Darobactin A (WNWSKSF) and Darobactin E (WSWSKSF) core AA sequence cluster by species (although species denomination is difficult in the *Yersinia* genus). Darobactin C and D propeptide sequences however do not cluster by species and form a distinct clade from the Darobactin A and E type propeptides, including different follower AAs compared to the majority of Darobactin A propeptide sequences.

IV.3.2. Genetic inactivation of darE and knock out of darABCDE in *Photorhabdus khanii* DSM3369

In order to confirm the involvement of the previously identified putative BGC in Darobactin A biosynthesis, the darE ORF coding for the presumably essential biosynthetic enzyme was disrupted by single crossover disruption using the vector pNB01. pNB01 was created by PCR amplification (Suppl. Fig.: 4) and isothermal assembly of the necessary parts and transformation of *E. coli* WM3064. After transformation, plenty of Kanamycin resistant colonies could be observed, indicating correct assembly of the plasmid. *E. coli* WM3064 was used to shuttle pNB01 to *P. khanii* DSM3369, effectively disrupting the darE ORF as proven by PCR (Fig.: 17). Using the described methodology, *P. khanii* DSM3369 was easy to conjugate, resulting in multiple conjugation plates with >100 colonies per plate. In contrast, several attempts to conjugate *P. khanii* HGB1456 using the same methodology did not result in any Kanamycin resistant colony. Integration of pNB01 in *P. khanii* DSM3369 was tested in 5 colonies by PCR, with two colonies showing the expected bands (Fig.: 17). After 3 days cultivation and LCMS analysis of WT *P. khanii* DSM3369 and *P. khanii* DSM3369-pNB01 in LB and LB_{Kan} respectively, Darobactin A production was analysed. Darobactin A was identified in the WT producer strain *P. khanii* DSM3369 due to presence of an ion with 483.7054 *m/z* and a retention time of ~ 21 min, corresponding well to C₄₇H₅₅N₁₁O₁₂ [M+H]²⁺ (theor. 483.7089 *m/z*, 7 ppm error). In *P. khanii* DSM3369-pNB01, no corresponding ion could be detected (Fig.: 17).



93

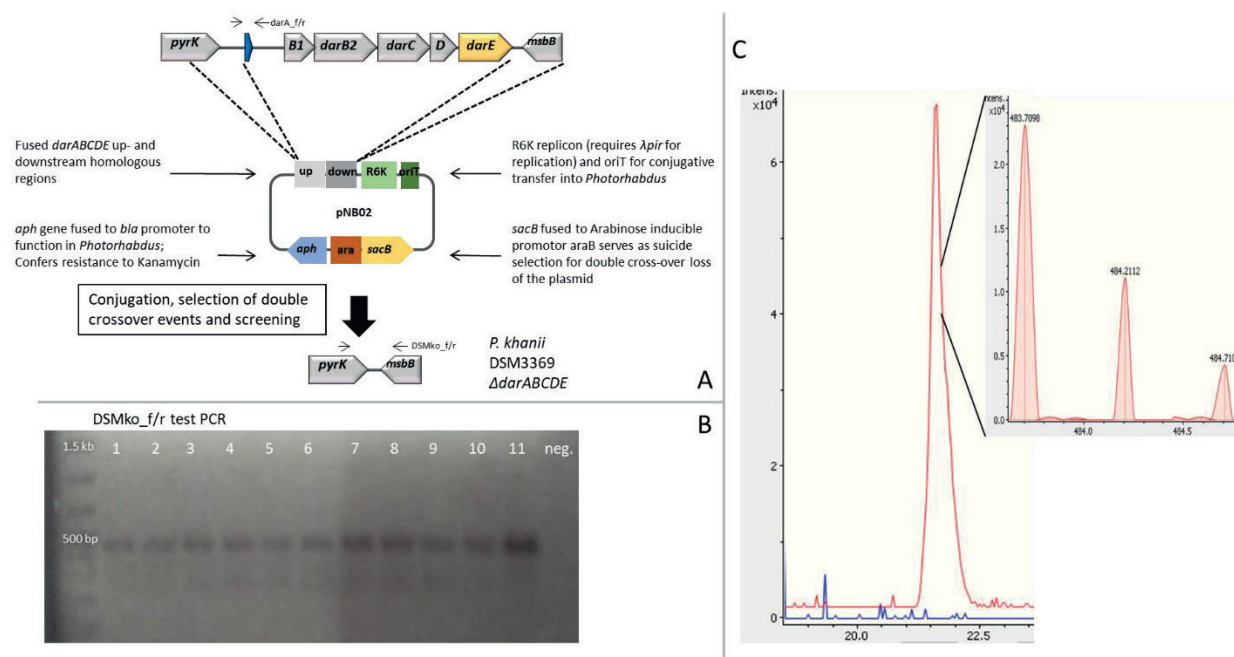


Figure 18: A) Topological vector map of pNB02, showing the individual elements of the plasmid together with the targeted genomic region in *P. khanii* DSM3369 and the genomic region after successful double crossover recombination KO. Primers used for the screening of the KO are annotated at the respective position. B) Screening PCR for the loss of the Darobactin A BGC using the primers DSMko_f/r, indicating the successful deletion of the Darobactin BGC in 11 clones and a negative control containing no template DNA. C) LCMS EICs of $C_{47}H_{55}N_{11}O_{12}$ $[M + 2H]^{2+} \pm 0.01$ for *P. khanii* DSM3369 WT (red) and *P. khanii* DSM3369 Δ *darABCDE* (blue), showing the double charged Darobactin A ion with 483.7098 m/z in the WT but not in *P. khanii* DSM3369 Δ *darABCDE*.

IV.3.3. Homologous and heterologous overexpression of the Darobactin A BGC in *Photorhabdus khanii* DSM3369 Δ *darABCDE* and *E. coli* BW25113

For homologous and heterologous overexpression of the Darobactin A BGC, the *E. coli*/*Photorhabdus* expression vector pNB03 was created by amplification of the individual parts and subsequent fusion by isothermal assembly. Test restriction with *SacI* and *SalI* revealed correct assembly in 5 out of 5 screened clones. Subsequently, the BGC was introduced to pNB03 with (pNB03-*darA-E*) intergenic region between *darA* and *darB*. Test restriction with *SalI*/*SacI* showed that 1 out of 3 and 4 out of 5 tested clones respectively carried the correctly assembled plasmids (Suppl. Fig.: 6). Constructs were transferred to *P. khanii* DSM3369 Δ *darABCDE* and *E. coli* BW25113 by conjugation and transformation, both yielding plenty of Kanamycin resistant clones. Successful conjugation was corroborated by colony PCR for the presence of *darA* and the *aph* gene on the pNB03 backbone (Suppl. Fig.: 6). Successful transformation of *E. coli* BW25113 was inferred by resistance to Kanamycin and Apramycin.

Cultivation of *P. khanii* DSM3369 Δ *darABCDE* + pNB03 (empty vector), *P. khanii* DSM3369 Δ *darABCDE* + pNB03-*darA-E*, *P. khanii* DSM3369 Δ *darABCDE* + pNB03-*darA-E*, *E. coli* BW25113 + pNB03-*darA-E*, *E. coli* BW25113 + pNB03-*darA-E*, *E. coli* BW25113 + pNB03-*darA-E* and WT *P. khanii* DSM3369 for three days in LB containing

appropriate antibiotics and L-arabinose as inducer unsurprisingly showed no production of Darobactin A in the negative control and good production on the WT *P. khanii* DSM3369. Surprisingly however, no production could be observed in *P. khanii* DSM3369 $\Delta darABCDE$ + pNB03-*darA-E* and only little production in *P. khanii* DSM3369 $\Delta darABCDE$ + pNBDaro. In contrast, the culture with *E. coli* BW25113 + pNB03-*darA-E* showed relatively high intensity of the double charged Darobactin A ion, resulting in a well visible peak in the EIC, roughly half the size the of WT *P. khanii* DSM3369. Comparison between *E. coli* BW25113 + pNB03-*darA-E* and *E. coli* BW25113 + pNBDaro showed that Darobactin A intensity in EIC was ~ 1.7 fold elevated in *E. coli* BW25113 + pNBDaro and thereby close to the intensity of the Darobactin A peak in the WT positive control strain (Fig.: 19).

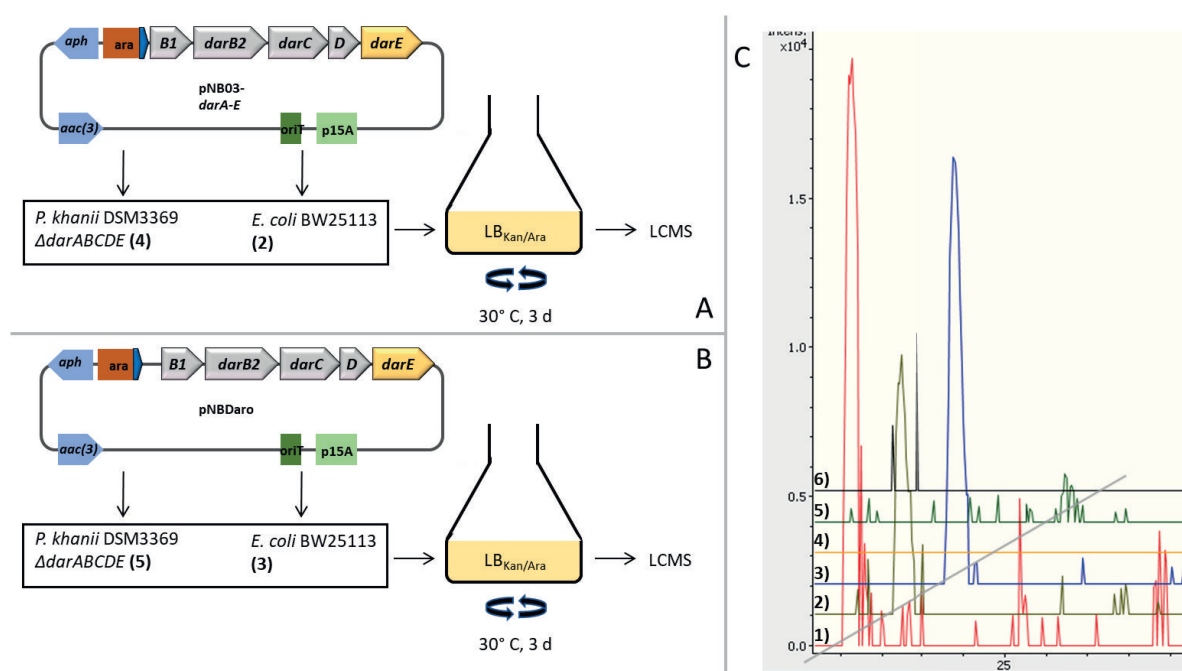


Figure 19: A) Topological vector map of pNB03-*darA-E* (intergenic region deleted between *darA* and *darB*) and subsequent homologous and heterologous expression workflow. B) Topological vector map of pNBDaro (with intergenic region between *darA* and *darB*) and subsequent homologous and heterologous expression workflow. C) LCMS EICs of $C_{47}H_{55}N_{11}O_{12} [M + 2H]^{2+} \pm 0.01$ for *P. khanii* DSM3369 WT (1), *E. coli* BW25113 + pNB03-*darA-E* (2), *E. coli* BW25113 + pNBDaro (3), *P. khanii* DSM3369 $\Delta darABCDE$ + pNB03-*darA-E* (4), *P. khanii* DSM3369 $\Delta darABCDE$ + pNBDaro (5) and *P. khanii* DSM3369 $\Delta darABCDE$ + pNB03 (6). The grey line illustrates the offset of the individual chromatogram traces.

IV.3.4. Identification of the Darobactin minimal BGC

For identification of the minimal Darobactin BGC, the identified transporter genes were removed one by another from the plasmid pNBDaro in order to test i) the possible involvement of one of the identified transporter genes in crosslinking or final trim of Darobactin A and ii) the impact of the transporter genes on the localisation of Darobactin A in the producing culture (medium vs. cell precipitate). Plasmids were assembled using standard methodology (Suppl. Fig.: 7) and all constructs

were expressed as described. Subsequent analysis of Darobactin A production by LCMS and comparison of Darobactin A concentration in supernatant and pellet revealed that the enzymes in question are not involved in the biosynthesis of Darobactin A at all. Darobactin A was produced by all constructs, even with the one only containing *darA* and *darE* showing that these two genes are sufficient to synthesize Darobactin A in *E. coli* BW25113. Comparison of Darobactin A concentration between the individual constructs in supernatant and pellet also showed that Darobactin A is always present outside the cell, regardless of the presence of the transporter genes encoded in the BGC (Fig.: 20). Concentrations in the supernatant were similar in all investigated constructs but pNBDaro- $\Delta darB1$, where only little amounts could be detected in both supernatant and cell pellet. The concentration of Darobactin A in the pellet was relatively high in pNBDaro- $\Delta darB$ and pNBDaro- $\Delta darBC$, while in pNBDaro- $\Delta darB1$ hardly any and in pNBDaro- $\Delta darBCD$ only little amounts Darobactin A could be detected (Fig.: 20). These differences might also be well explained by differences in growth or external factors since previous experiments show certain fluctuation among different cultures of the same heterologous Darobactin A producer.

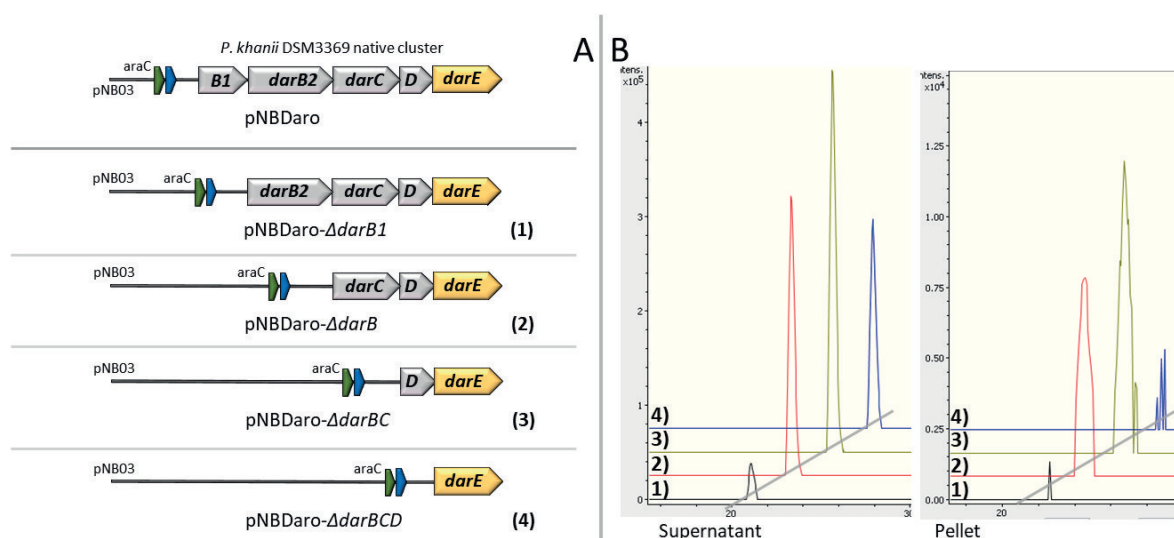


Figure 20: A) Genetic setup of the native cluster on pNBDaro compared to the setup of respective deletion plasmids expressed in *E. coli* BW25113 with numbers indicating the position in the EICs in B). B) EICs of $C_{47}H_{55}N_{11}O_{12} [M + 2H]^{2+} \pm 0.01$ in culture supernatant (left) and cell pellet (right). 1) *E. coli* BW25113 + pNBDaro- $\Delta darB1$, 2) *E. coli* BW25113 + pNBDaro- $\Delta darB$, 3) *E. coli* BW25113 + pNBDaro- $\Delta darBC$ and 4) *E. coli* BW25113 + pNBDaro- $\Delta darBCD$, showing the presence of Darobactin A in every tested construct in supernatant and cell pellet. The grey line illustrates the offset of the individual chromatogram traces.

IV.3.5. Mutasynthetic production of Darobactin analogues B-F

Since RiPP biosynthetic pathway rely on ribosomal assembly of the precursor and subsequent modification of key amino acids, they are very susceptible to mutasynthetic derivation of the core structure by exchange of certain amino acids on codon level. In order to test said versatility for the investigated Darobactin BGC, the vector pNBDaromod was created (Suppl. Fig.: 8), which allowed the quick modification of the core peptide by GoldenGate cloning of synthesised DNA fragments while the rest of the BGC remains untouched. In this way, derivations from the core AA structure can be designed *in-silico* and directly incorporated in an easy 2 step reaction. Correct assembly of pNBDaromod was inferred from the antibiotic resistance of *E. coli* BW25113 transformants and a positive blue-white reaction on X-Gal containing medium. This plasmid allows scarless incorporation of altered core AA sequences into the Darobactin A propeptide gene *darA* (Fig.: 21). In a first experiment, the AA sequences of previously identified hypothetical Darobactin analogs B-E containing the Darobactin A follower sequence QEI were assembled into pNBDaromod (Fig.: 21). Transformation of the GoldenGate assembly mix and growth on IPTG an X-Gal containing agar predominantly resulted in growth of white colonies, with blue colonies being less than 1 % for all reactions. White transformants were picked and propeptides of three clones respectively were sequenced, revealing that all clones obtained the desired modification of the *darA*. Since all sequenced clones showed the respective desired modification, sequencing was omitted in further experiments and correct assembly was inferred from negative blue-white selection. Expression of the strains carrying individual propeptides in LB_{Kan/Ara} at 30° C however did not yield any hit for the expected predicted masses, except from the Darobactin E construct which showed the expected mass of the double charged ion in very low amounts with similar retention time to Darobactin A (Suppl. Fig.: 9).

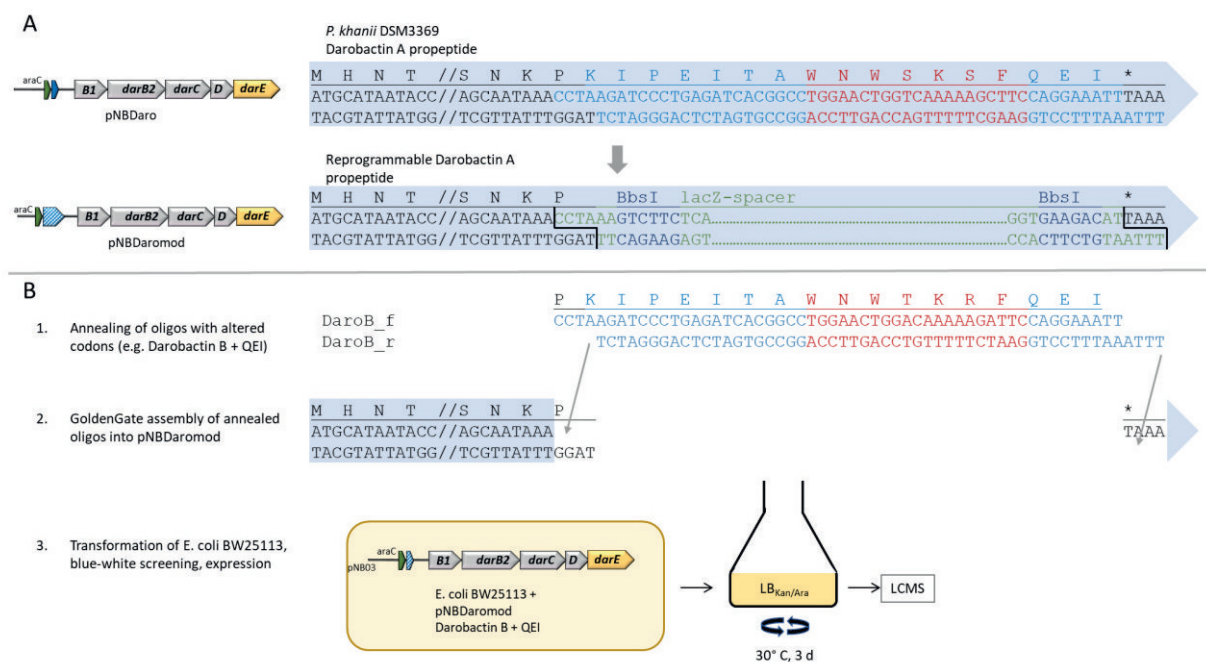


Figure 21: A) Layout of pNBDaro compared to the layout of pNBDaromod with lacZ spacer and BbsI cutting sites for GoldenGate mediated exchange of DarA core AAs. B) Workflow of Darobactin analogue generation exemplified with Darobactin B core AA sequence + QEI follower. B1) Annealing of synthesised oligonucleotides that match the desired core AA sequence. B2) GoldenGate assembly of annealed oligos into pNBDaromod. B3) Transformation of *E. coli* BW25113 with the altered plasmid, selection, expression and LCMS analysis.

To investigate whether the QEI follower of the Darobactin A propeptide is elemental to the maturation of Darobactin A, the QEI follower was removed. Expression and subsequent LCMS analysis revealed good Darobactin A production, comparable to heterologous expression of the unaltered BGC, showing that the follower AAs in the different propeptide sequences are expendable for the production of Darobactin (Fig.: 21). Hence, sequences were altered to the respective Darobactin AA sequence, this time omitting the follower AAs. Expression and subsequent LCMS analysis now revealed the presence of the expected masses of all putative natural derivatives with only little aberration however in changing intensities (Fig.: 21). Double charged ions of Darobactin B were detected with 525.2533 m/z (4 ppm error), C with 484.2110 m/z (9 ppm error) D with 497.7143 m/z (5 ppm error), E with 470.2064 m/z (6 ppm error) and F with 487.2496 m/z (3 ppm error). Masses of truncated Darobactin A analogues, i.e. lacking AA positions -F⁷ and -S⁶F⁷ completely, could not be observed. Retention times all were similar to the 20.5 min of Darobactin A with Darobactin C, D and E eluting nearly identically. Darobactin B and F however eluted slightly earlier at ~ 19 min.

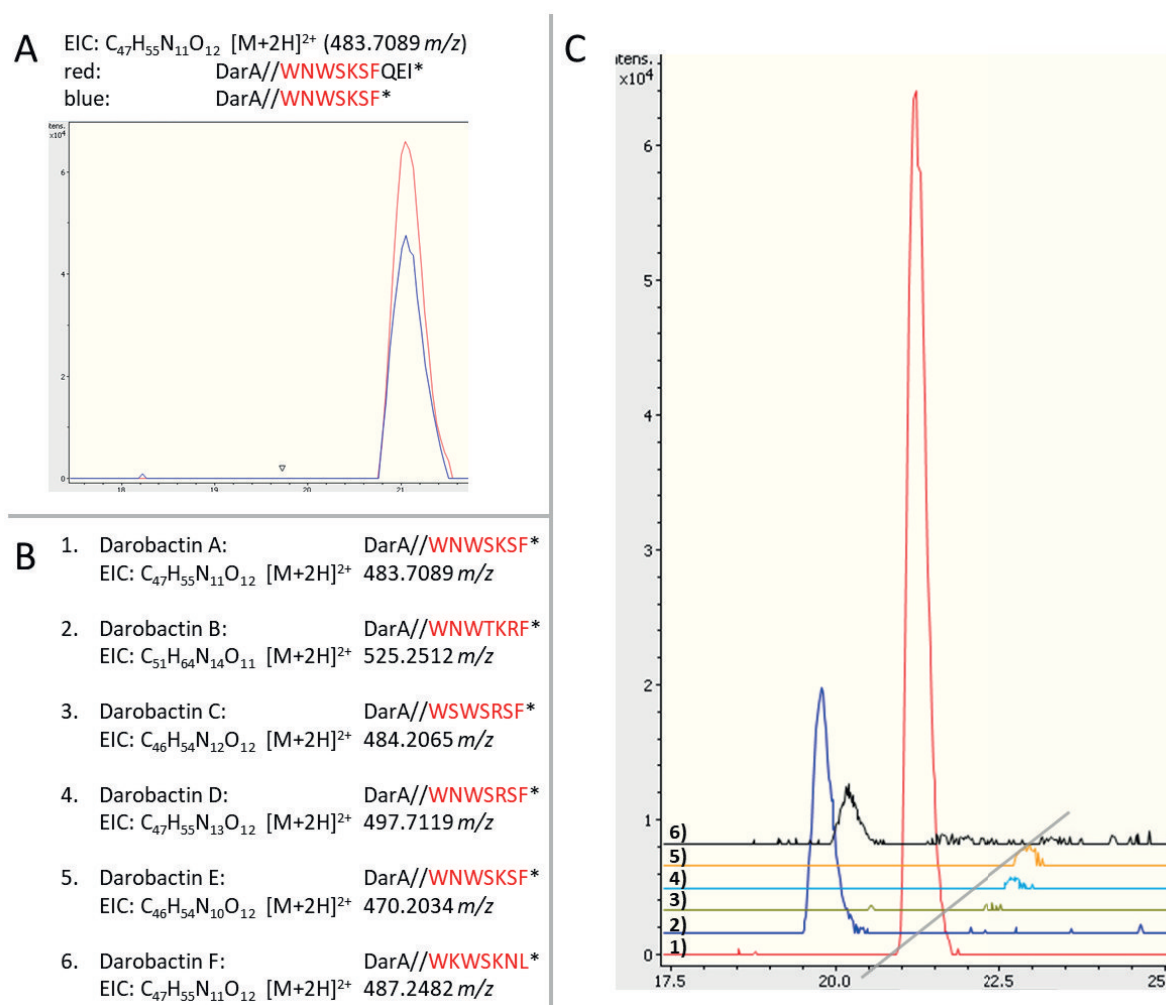


Figure 22: A) Comparison of Darobactin A production with QEI and without QEI follower AAs in the DarA propeptide. B) DarA with different core AA and their respective expected product sum formula and calculated double charged ion m/z . Numbers indicate respective position in the EICs in C). C) EICs of Darobactin analogues B (2), C (3), D (4), E (5) and F (6) with Darobactin A (1) as control, showing the creation of all natural Darobactin A derivatives by mutasynthesis. The grey line illustrates the offset of the individual chromatogram traces.

Comparison of the EIC for Darobactin derivatives to the LCMS chromatogram of the Darobactin A construct which served as negative control also showed that all detected masses for the derivatives were absent. Darobactin A was detected with by far the highest intensity, followed by Darobactin B, F and E, the latter showing < 10 % of the intensity of the Darobactin A peak. In contrast, Darobactin derivatives C and D with modification of AA 5 (K → R) were only detected in very minor amounts, showing even lower intensity than Darobactin E (Fig.: 22) with Darobactin C having the lowest intensity of all.

To obtain higher production yield of Darobactin derivatives and confirm the identity of the observed molecules, modified clusters were recloned into the high copy number *E. coli* expression vector pRSF-duett under control of the strong T7 promoter and transferred to *E. coli* Bap1mut (Darobactin resistant). Darobactin analogue production was performed in $LB_{Kan/IPTG}$ as described and culture

supernatant was concentrated 100 fold using C18 solid phase extraction and subjected to LCMS/MS analysis. Using targeted fragmentation for the respective masses however, only revealed the fragmentation patterns of Darobactin A, B, D and E. Darobactin C could neither be fragmented by auto MS/MS nor targeted fragmentation due to the low intensity. Darobactin F was identified a certain time after and it was decided to omit fragmentation analysis of this molecule.

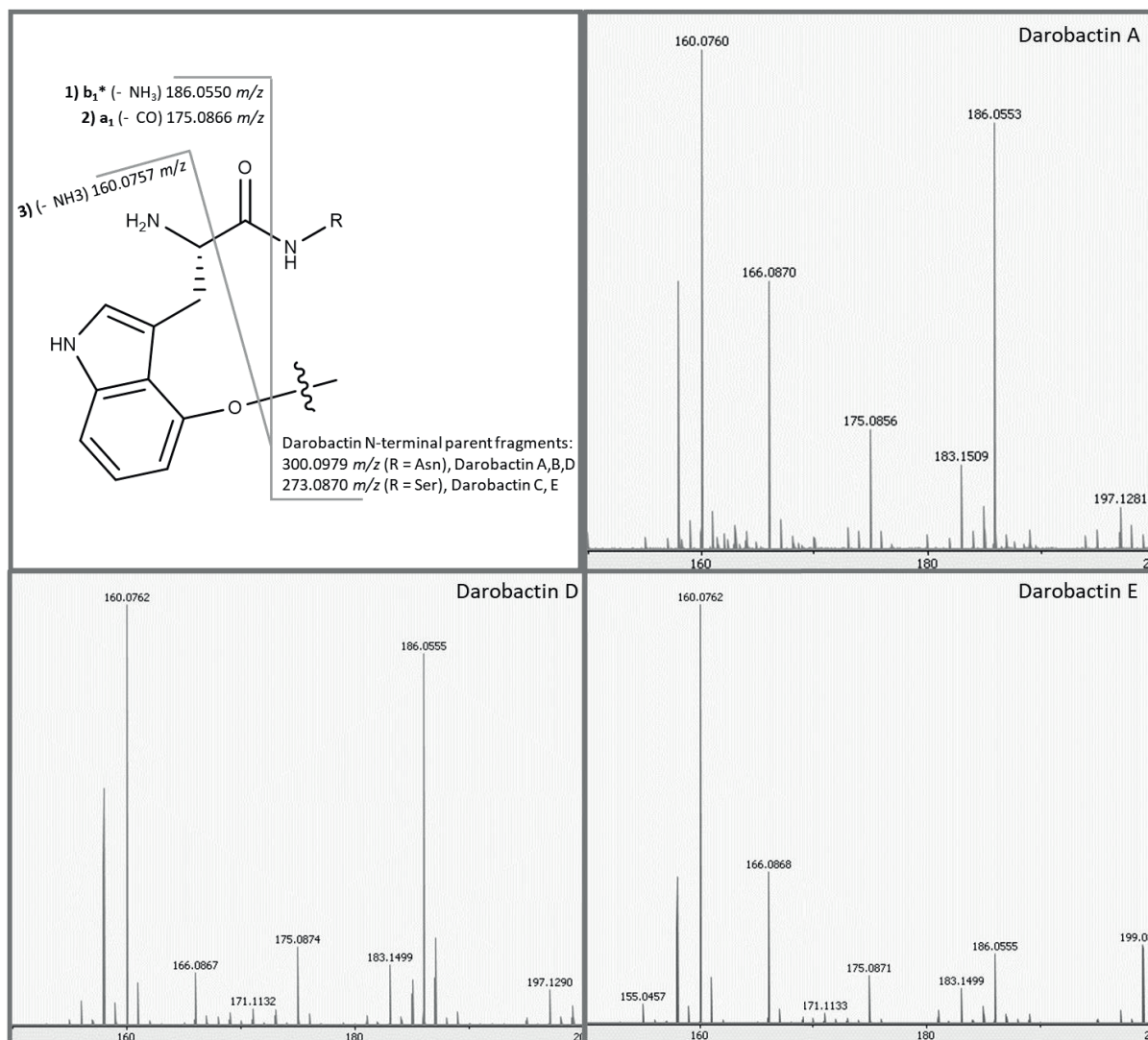


Figure 23: Structure of the N-terminal W of Darobactin Analogues with plausible MS/MS fragmentation pattern and MS/MS spectra of Darobactin A, D and E. Respective measured fragment ions are annotated. Darobactin B was measured on a different instrument that did not measure < 200 m/z .

Fragmentation of Darobactin A, D and E resulted in clear fragmentation patterns and fragment m/z values could be readily matched with plausible decay products of the respective parent ions. Comparison of the generated fragmentation spectra enabled the identification of characteristic fragments: The N-terminal W^1 cleaved among the ether bridge to W^3 , creating b_1^* (* = loss of NH_3) (186.0550 m/z) and a_1 (175.0866 m/z) fragments as well as the hydroxyindolylethanylium ion

(160.0757 m/z) residue of hydroxylated tryptophan (Fig.: 23) which were observed in the spectra of Darobactin A,D and E. Darobactin B fragmentation was measured on the lower resolution mass spectrometer with a detection range of 200 m/z to 2000 m/z , hence these small fragments could not be detected.

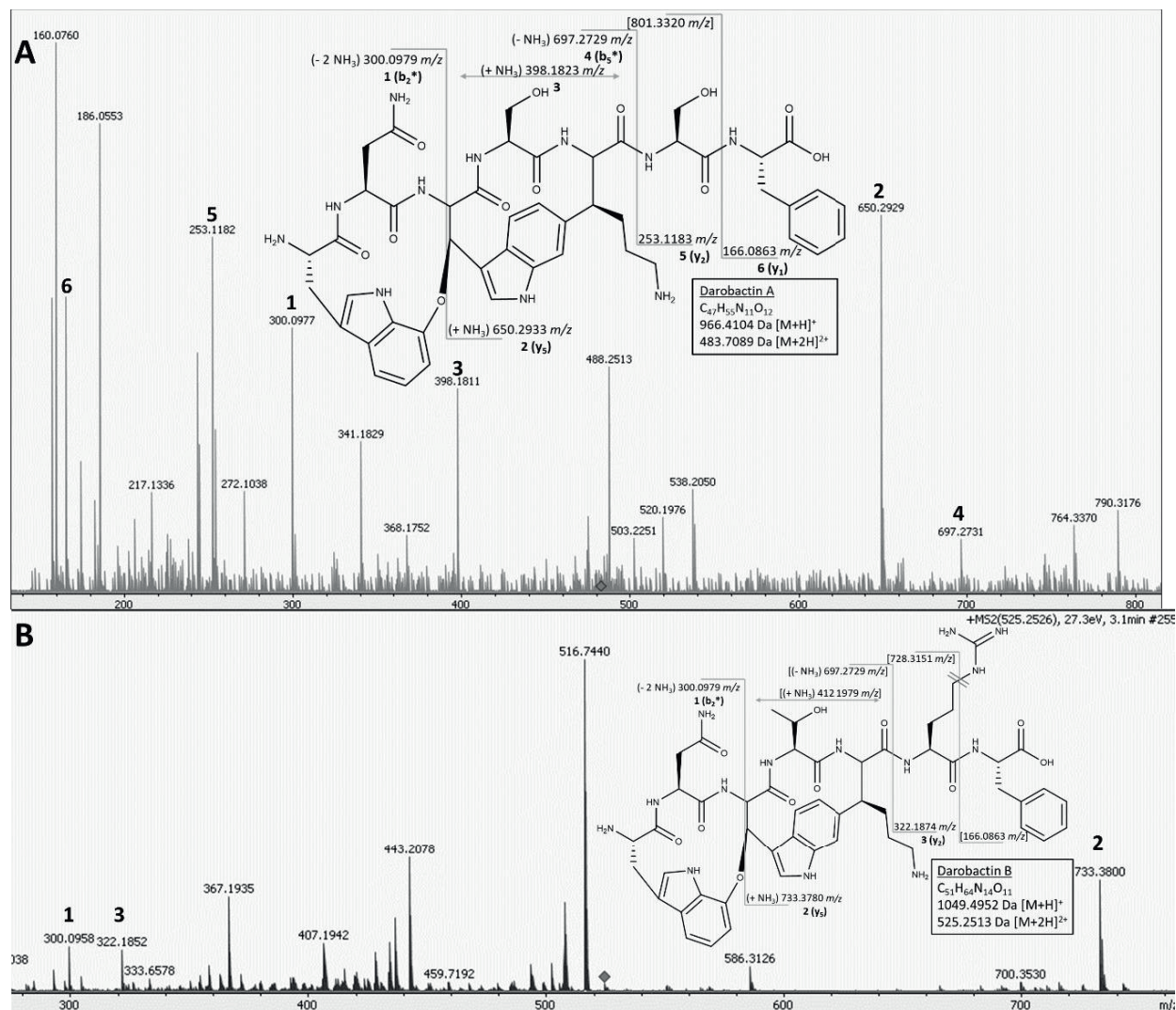


Figure 24: Structure, sum formula and theoretical exact mass of A) Darobactin A and B) Darobactin B with plausible MS/MS fragmentation pattern and MS/MS spectrum. The respective measured fragment ions that correspond to the shown fragmentation pattern are annotated.

b_2^* fragments of Darobactin A, B, D ($W^1\text{-N}^2$) and E ($W^1\text{-S}^2$), also cleaved among the aromatic-aliphatic ether bridge between $W^1\text{-W}^3$ with 300.0979 m/z (Fig.: 24) and 273.0870 m/z (Fig.: 25) respectively. From the C-terminus of the molecules, characteristic fragments for all molecules were identified as y_1 (F^7) with 166.0863 m/z and y_2 ($S^6\text{-F}^7$) 253.1183 m/z for Darobactin A, D and E. For Darobactin B, y_2 ($R^6\text{-F}^7$) has an m/z of 322.1874. y_5 ions with cleavage of the aromatic-aliphatic ether bridge ($W^1\text{-W}^3$) could be observed for Darobactin A and E (650.2933 m/z) (Fig.: 24, 25), Darobactin B (733.3780 m/z) (Fig.: 24) and Darobactin D (678.2994 m/z) (Fig.: 25). b_5^* ions could be identified for Darobactin A (697.2729 m/z) and Darobactin E (670.2620 m/z) Internal cleavage ions $N^2\text{-W}^3\text{-S}^4\text{-K}^5$ of Darobactin A

(538.2045 m/z) and $W^3-S^4-K^5$ of Darobactin A and E were also observed (398.1823 m/z) (Fig.: 24, 25). m/z values of all fragments were recorded with high accuracy (< 5 ppm error) aside from the Darobactin B fragments which were measured on a lower accuracy instrument (< 7 ppm error). This in depth analysis of the fragmentation patterns confirmed the identity of Darobactin A, D and E and additionally corroborated the successful creation of Darobactin A analogues by mutation of the Darobactin propeptide gene *darA* in this heterologous expression system (Fig.: 21)

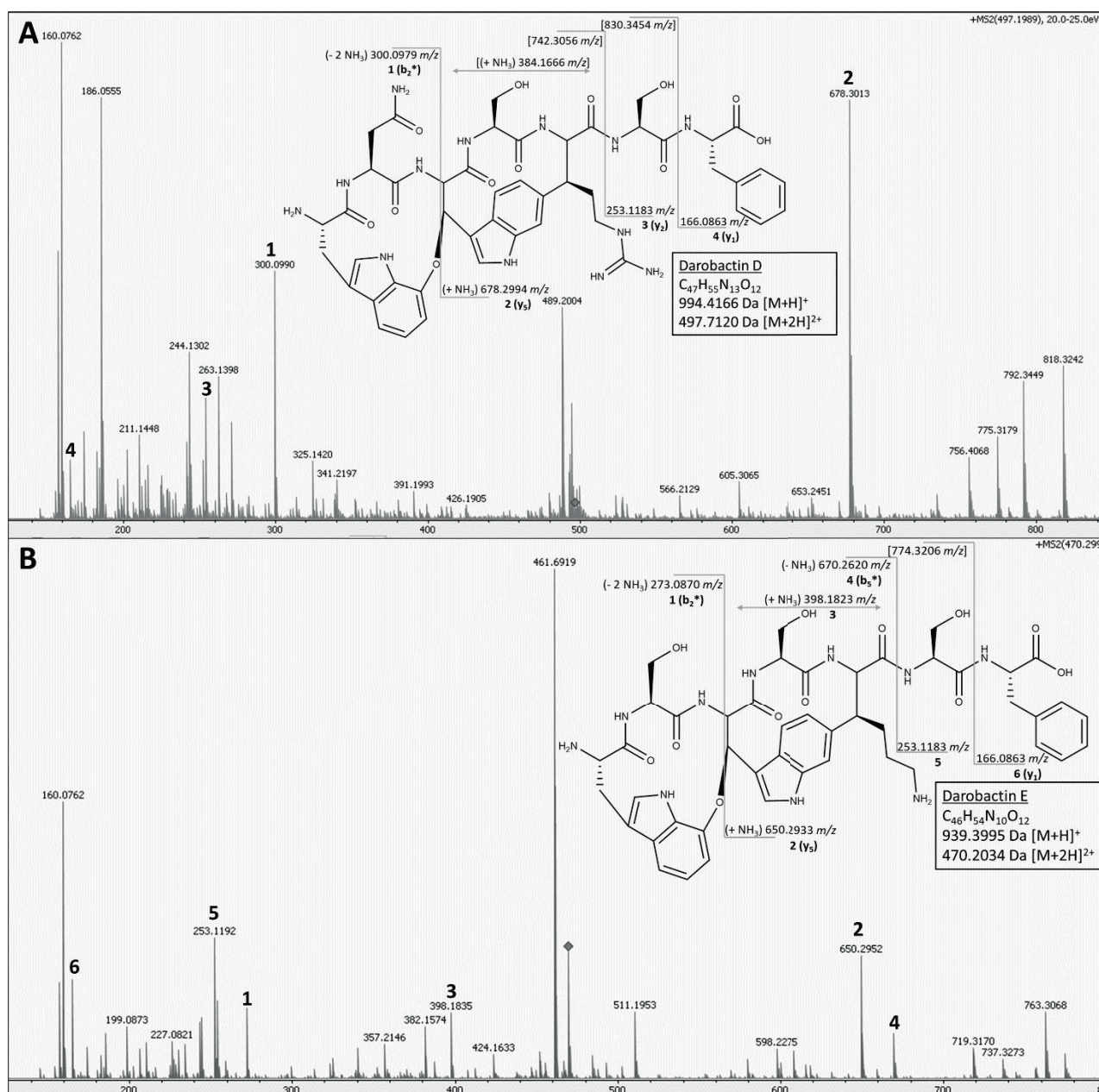


Figure 25: Structure, sum formula and theoretical exact mass of A) Darobactin D and B) Darobactin E with plausible MS/MS fragmentation pattern and MS/MS spectrum. The respective measured fragment ions that correspond to the shown fragmentation pattern are annotated.

Isolation of Darobactin B, D and E from large scale cultivation yielded 12 mg for Darobactin B, 0.6 mg for D and 1.3 mg for E. Large volumes of culture were necessary to obtain sufficient amounts for

activity tests. While Darobactin B could be isolated from 35 L of culture, ~ 100 L of culture in total were necessary to isolate Darobactin D and E. Attempts to purify Darobactin C even in low mg quantity failed.

IV.3.6. Discovery of Dehydrodarobactin A and Marinodarobactin A and B from *Pseudoalteromonas luteoviolaceae* H33 and H33S

Pse. luteoviolaceae H33 and H33S were grown in ten different media and SPE extracts were prepared as described. Antibiotic activities were screened in 96 well microtiter plates against *E. coli* MG1655 and MG1655 *bamA6* in three subsequent serial dilution steps. Testing extracts of cultures in 10 different media showed clear inhibition of *E. coli* MG1655 and MG1655 *bamA6* in all media tested (data not shown). Extracts of cultures grown in M4, M5, M6 and M8 showed dose independently the highest activity and were selected for further screening. Tests of the extracts of MB, M4, M5, M6 and M8 were repeated using extracts of uninoculated medium as a negative control, showing that the uninoculated medium also has certain antimicrobial activity especially at high concentration. Wells containing diluted extracts however could be read out and growth inhibition could be calculated, revealing that extracts from *Pse. luteoviolaceae* H33 and H33S grown in M8 clearly showed highest dose independent inhibition of WT *E. coli* MG1655 as well as the more sensitive (to purified Darobactin A) MG1655 *bamA6* test strain. Surprisingly, the M8 extracts showed higher inhibition of the WT than the more sensitive test strain. Similar results could be observed for cultures in all other media, also showing vast differences in activity between H33 and H33S extracts.

Table 12: Growth inhibitory effect of *Pseudoalteromonas luteoviolaceae* H33 and H33S extracts from different media against *E. coli* MG1655 and MG1655 *bamA6*. Low % values indicate good growth inhibition, high % values indicates low inhibition.

<i>bamA6</i>	MB	0.5	0.25	M4	0.5	0.25	M5	0.5	0.25	M6	0.5	0.25	M8	0.5	0.25
H33	nd	44,46	24,65	nd	35,42	29,10	nd	nd	35,98	M4	nd	35,55	nd	35,27	28,91
H33S	nd	70,71	39,86	nd	13,07	52,35	nd	nd	81,56	M5	nd	32,93	nd	39,19	27,36
WT <i>bamA</i>	MB	0.5	0.25	M4	0.5	0.25	M5	0.5	0.25	M6	0.5	0.25	M8	0.5	0.25
H33	nd	12,67	8,30	nd	60,66	81,27	nd	nd	80,52	nd	nd	77,45	nd	14,79	11,18
H33S	nd	98,73	90,43	nd	26,23	83,97	nd	nd	89,98	nd	nd	48,84	nd	10,78	8,52

Subsequently, generated extracts were subjected to LCMS analysis and investigated by GNPS molecular networking. Calculating a network using standard settings showed that Darobactin A was produced by both *Pse. luteoviolaceae* strains in all four media, however Darobactin A did not cluster with any other masses but formed a single node. Recalculation of the network, adjusting the number of matching fragments to 2 (Standard: 6) revealed a small three membered cluster with Darobactin

A, a mass of 523.663 m/z (+39.954 m/z) and a mass of 450.129 m/z (-73.534 m/z). Subsequent inspection of the individual LCMS runs confirmed the presence of Darobactin A in all culture extracts, with highest intensity in the M5 culture extract. The EIC of the double charged Darobactin A ion however additionally showed a second peak, eluting 0.8 min after Darobactin A. Inspection of this peak revealed that the EIC picked up the third peak of the isotope pattern of a double charged ion with 482.7014 m/z (Fig.: 26), being exactly 2 Da lighter than Darobactin A, perfectly fitting to the sum formula $C_{47}H_{53}N_{11}O_{12}$ (0.6 ppm aberration).



Figure 26: A) EICs of i) $C_{47}H_{55}N_{11}O_{12} [M+2H]^{2+}$ and ii) $C_{47}H_{54}N_{11}O_{12}Br [M+2H]^{2+}$, numbers representing Darobactin A (1), Dehydrodarobactin A (2), Bromodarobactin A (3) and Dehydrobromodarobactin A (4). The straight line indicates the offset for better visualisation. B) Sum formula, theoretical double charged m/z and measured isotopic pattern of Darobactin A, C) Sum formula, theoretical double charged m/z and measured isotopic pattern of Dehydrodarobactin A, D) Sum formula, theoretical double charged m/z and measured isotopic pattern of Bromodarobactin A and E) Sum formula, theoretical double charged m/z and measured isotopic pattern of Dehydrobromodarobactin A.

By comparison of the fragmentation pattern of the 450.129 m/z ion to Darobactin A, it could be ruled out that the 450.129 m/z ion was related to Darobactin A due to poor accordance and low number of matching fragments. The EIC of 523.663 m/z showed the ion with a shift in retention time of + 0.8 min compared to Darobactin A and it also showed the same dual peak that was observed for the EIC of Darobactin A (Fig.: 26). Investigation of the mass spectrum of the culture grown in M5 showed that 523.663 m/z is not the mass of the monoisotopic ion but the third peak of an unusual isotope pattern of a monoisotopic double charged ion with 522.6647 m/z (Fig.: 26). Calculating the difference of the exact mass of Darobactin A to the mass of the ion and taking into account the double charge showed a difference of 79.91 m/z , very close to the mass of the lighter stable bromine isotope ^{79}Br , indicating the presence of brominated Darobactin A derivatives. Simulating the isotope pattern for $\text{C}_{47}\text{H}_{54}\text{N}_{11}\text{O}_{12}\text{Br} [\text{M}+2\text{H}]^{2+}$ in Compass DataAnalysis showed a perfect fit to the measured ion with only 1.15 ppm aberration for the monoisotopic m/z signal (Fig.: 27).

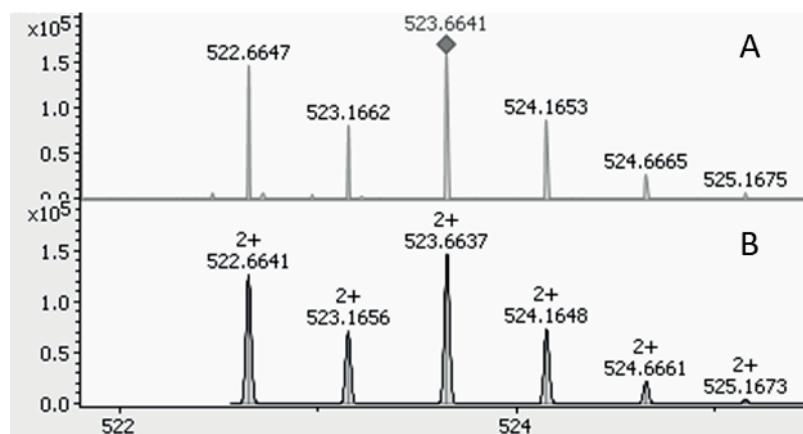


Figure 27: Measured isotopic pattern of Bromodarobactin A (A) compared to calculated (Bruker Compass DataAnalysis) isotopic pattern for the sum formula $\text{C}_{47}\text{H}_{54}\text{N}_{11}\text{O}_{12}\text{Br} [\text{M}+2\text{H}]^{2+}$ (B) confirming the deduced sum formula of the detected Darobactin analogue.

Looking at the second peak in the EIC of $\text{C}_{47}\text{H}_{54}\text{N}_{11}\text{O}_{12}\text{Br} [\text{M}+2\text{H}]^{2+}$, the ion showed the same – 2 Da mass shift with the characteristic bromination isotope pattern and + 0.8 min shift in retention time compared to the monobrominated Darobactin A as already observed for Darobactin A. Again, this is well congruent with a double charged ion with the sum formula $\text{C}_{47}\text{H}_{52}\text{N}_{11}\text{O}_{12}\text{Br}$ with only 0.4 ppm aberration. Hence, this – 2 Da mass shift is indicative of the presence of a double bond (-2H^+) that is not present in the previously identified molecules Darobactin A and monobrominated Darobactin A. Comparison of the fragmentation patterns of the ions with 482.7014 m/z , 522.6647 m/z and 521.6565 m/z corroborated the relation to Darobactin A and revealed that major parts of the molecules are in fact identical (Fig.: 28).

Fragmentation of (-2H) -Darobactin A yields the same N-terminal b_1^* (186.0550 m/z), a_1 (175.0866 m/z) and hydroxyindolyethanylium ion (160.0757 m/z) from W^1 as well as C-terminal y_1 (F^7) with

166.0863 m/z and y_2 (S^6-F^7) 253.1183 m/z fragment ions characteristic for Darobactin A. Compared to Darobactin A, y_5 (650.2933 m/z) and the internal cleavage ion $N^2-W^3-S^4-K^5$ of Darobactin A (538.2045 m/z) were shifted to 648.2776 m/z and 536.1888 m/z respectively and neither the characteristic 398.1823 m/z nor a -2H downshifted fragment could be detected. This indicates that the double bond has to be located on the $W^3-S^4-K^5$ backbone section, however the exact location remains unclear (Fig.: 28). Fragmentation of Br-Darobactin A yields the same C-terminal y_1 (F^7 , 166.0863 m/z), y_2 (S^6-F^7 , 253.1183 m/z) and y_5 (650.2933 m/z) characteristic for Darobactin A as well as the characteristic internal cleavage ions with 538.2045 m/z and 398.1823 m/z respectively, however completely lacking characteristic fragments of the N-terminus of Darobactin A. Inspection of the fragmentation spectrum revealed the presence of fragments congruent with $b_1^* + {}^{79}\text{Br}$ (263.9655 m/z), a ${}^{79}\text{Br}$ -bromohydroxyindolyethanylium ion (237.9862 m/z) and a corresponding $b_5^* + \text{Br}$ ion with 775.1834 m/z , all showing a distinctive second peak of similar intensity with a shift of +2 Da corresponding to the respective ${}^{81}\text{Br}$ fragment (Fig.: 29). This indicates that the bromination is located at the residue of W^1 , likely at the aromatic part of the indole moiety. Fragmentation of (-2H)-Br-Darobactin A essentially shows a combination of the brominated N-terminal section of Br-Darobactin A and the - 2 Da downshifted core fragments of (- 2H)-Darobactin A, indicating that the bromination is at the exact same location as in Br-Darobactin A while the double bond is in the same location as in (- 2H)-Darobactin A. Again, all fragment m/z values were recorded with high accuracy (< 5 ppm error).

Newly identified Darobactin analogues were provisionally termed Dehydrodarobactin A ($\text{C}_{47}\text{H}_{53}\text{N}_{11}\text{O}_{12}$), Marinodarobactin A ($\text{C}_{47}\text{H}_{54}\text{N}_{11}\text{O}_{12}\text{Br}$) and Marinodehydrodarobactin A ($\text{C}_{47}\text{H}_{52}\text{N}_{11}\text{O}_{12}\text{Br}$). These novel analogs could subsequently be found in all generated extracts from *Pse. luteoviolacea* H33 and H33S, however in lower intensity. Comparing the intensities of all four Darobactins from *Pse. luteoviolacea* H33 grown in M5 shows that corresponding dehydro and non-dehydro derivatives have roughly the same intensity, while brominated derivatives have slightly lower intensity. Further inspection of the recorded mass spectrum of *Pse. luteoviolacea* H33 grown in M5 shows masses fitting to respective double charged iodinated Darobactin A derivatives with respective sum formulae $\text{C}_{47}\text{H}_{54}\text{N}_{11}\text{O}_{12}\text{I}$ and $\text{C}_{47}\text{H}_{52}\text{N}_{11}\text{O}_{12}\text{I}$ with good accuracy but minor intensity (data not shown). In contrast, derivatives containing halogens lighter than Br (F or Cl) or multiple halogenation of the Darobactin backbone could not be detected at all.

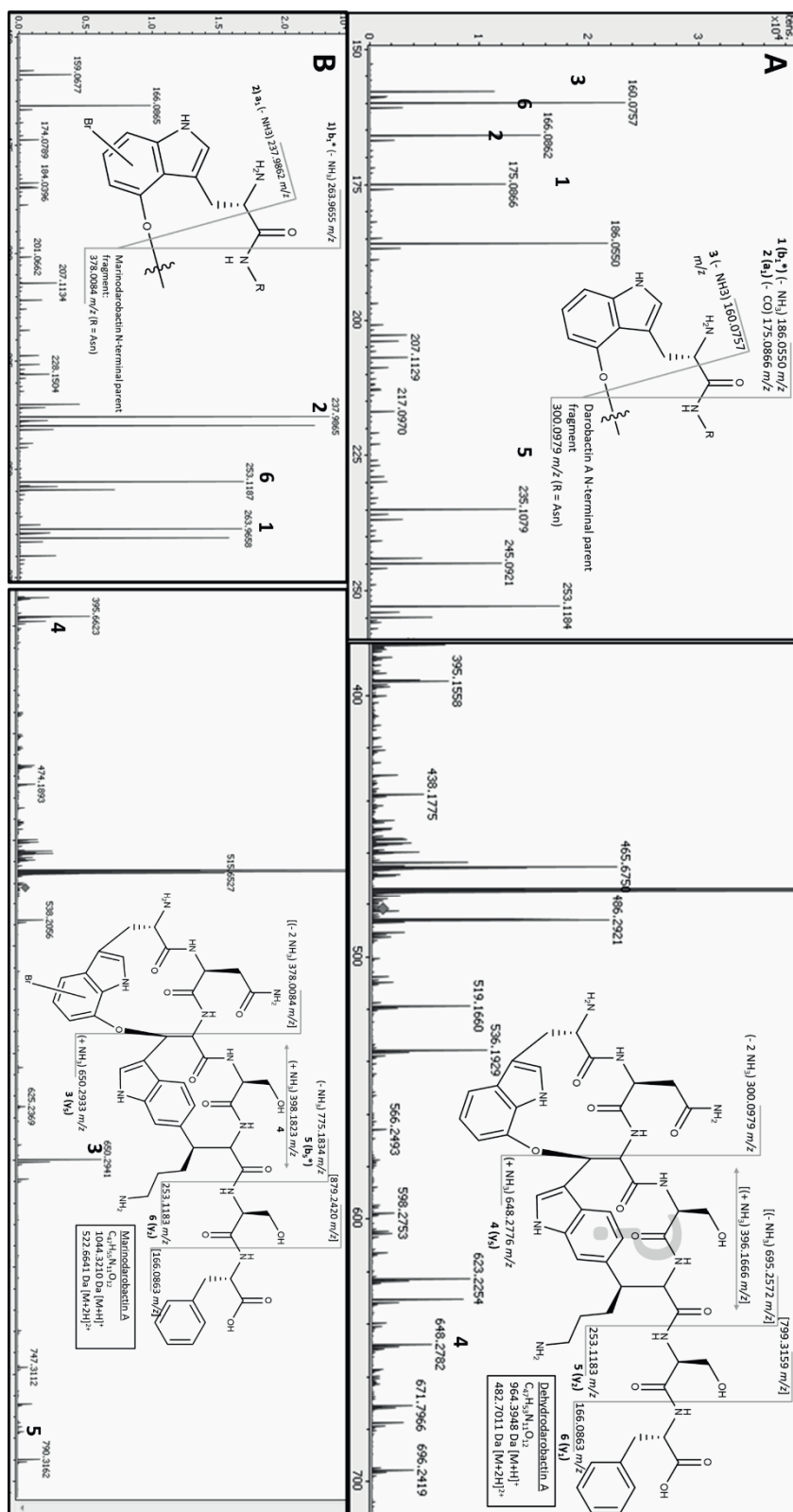


Figure 28: A) Structure, sum formula and theoretical exact mass of Dehydrdarobactin A and respective N-terminus in detail with plausible MS/MS fragmentation pattern and MS/MS spectrum with annotated respective measured fragment ions. Location of the double bond can be narrowed down to a specific fragment, however exact locations remains elusive (indicated by the greyed out ?) B) Structure, sum formula and theoretical exact mass of Marinodarobactin A and respective N-terminus with plausible MS/MS fragmentation pattern and MS/MS spectrum with annotated respective measured fragment ions. Location of the bromination can be narrowed down to the W1 residue, most likely the aromatic ring.

IV.3.7. MIC values of Darobactin analogues

MIC values of Darobactin analogs B, D, E and Marinodarobactin A and Marinodehydrodarobactin A were determined against a panel of clinically relevant Gram- pathogens. Darobactin A (supplied as pure compound from Prof. Kim Lewis) showed MIC values of 4 µg/mL or slightly above for all tested strains but *A. baumannii* (> 64 µg/mL) and *S. enterica* ATCC13076 (8 µg/mL). Darobactin B showed MICs of 1 µg/mL or lower but for *P. aeruginosa* (> 64 µg/mL), *K. pneumoniae* (8 µg/mL) and *A. baumannii* (32 µg/mL). Darobactin D in general showed less activity with MIC 64 µg/mL or higher for all tested strains. Darobactin E showed MIC values of 32 µg/mL for all strains but *A. baumannii* and *E. coli* Survcare052 NDM-5 (> 64 µg/mL) and *E. coli* ATCC35218 and *S. enterica* (64 µg/mL – 32 µg/mL). Marinodarobactin A showed MICs of 2 µg/mL or lower except for *P. aeruginosa* (> 64 µg/mL), *A. baumannii* (32 µg/mL) and *S. enterica* (8 µg/mL - 4 µg/mL). MICs of Marinodehydrodarobactin A were less uniform with all being between 16 µg/mL and 1 µg/mL but for *P. aeruginosa* with > 64 µg/mL (Tab.: 13). MIC values for the Rifampicin control were 16 µg/mL or lower except for *E. coli* NRZ14408 KPC-2 (> 64 µg/mL) and for the Gentamycin control were 4 or lower except for *A. baumannii* (16 µg/mL), *E. coli* NRZ14408 KPC-2 and *E. coli* MMGI1 OXA-48 (both 64 µg/mL or higher). A Tetracycline control showed MICs of 4 µg/mL or lower except for *S. enterica* ATCC13076 (16 µg/mL - 8 µg/mL) and *Pseudomonas aeruginosa* (> 64 µg/mL), however all clinical *E. coli* isolates were Tetracyclin resistant, so that Colistin was used as control for those, showing MICs all < 2 µg/mL (Tab.: 13).

Table 13: MIC values of Darobactin A and selected Darobactin analogues against clinically relevant gram – pathogens. Activities > 64 mg/mL were considered not active, n.d. indicates that this strain was not tested against the respective compound.

cpd tested (n=3)	<i>E. coli</i>	<i>E. coli</i>	<i>E. coli</i>	<i>E. coli</i>	<i>E. coli</i>	<i>P. aeru</i>	<i>P. aeru</i>	<i>K. pneu</i>	<i>A. bau</i>	<i>S. ent</i>
	ATCC	NRZ	K046	Survcare	MMGI1	PAO 1	PAO 750	DSM	ATCC	ATCC
	35218	14408		052				30104	19606	13076
	WT	KPC-2	VIM-1	NDM-5	OXA-48	WT	Efflux -	WT	WT	WT
Darobactin A	8 - 4	4	4	8 - 4	4	4	4	4	> 64	8
Darobactin B	1	1 - 0.5	1 - 0.5	1	1	> 64	n.d.	2 - 1	32	1
Darobactin D	64	64	32	64	64	32	64	64	> 64	64
Darobactin E	64 - 32	32	32	64	32	32	32	32	> 64	64 - 32
Marinodarobactin A	2	1	2 - 1	2 - 1	2	> 64	n.d.	1	32	8 - 4
Marinodehydrodarobactin A	8 - 4	1	4	16	4	> 64	n.d.	4 - 2	16	8 - 4
Rifampicin	4 - 2	> 64	8 - 4	16 - 8	16 - 8	16	16	8	2	16
Tetracycline/ Colistin*	0.5	2*	0.125*	0.125 - 0.063*	0.031	> 64	0.5	2 - 1	4 - 2	16 - 8
Gentamycin	0.25	> 64	4	0.25	64	0.063	0.125	0.125 - 0.063	16	0.125
	MIC [µg/mL]									

IV.4. Discussion

IV.4.1. Identification and corroboration of the Darobactin minimal BGC

When first presented with the Darobactin (A) structure by Prof. Kim Lewis, due to the unusual ring-forming modifications, curiosity about the biosynthetic origin i.e. the corresponding BGC, was sparked immediately. Initial speculation about non-ribosomal peptide (NRPS) biosynthesis was dismissed fast, due to the presence of only proteinogenic (L) AAs and no convincing match in an *in-silico* analysis of the producer genome. An antiSMASH analysis of the *Photorhabdus khanii* HGB1456 producer strains genome for NRPS BGCs. Additionally, the previously identified streptococcal pheromone Streptide (Schramma *et al.*, 2015), which also carries a very similar covalent bond between the aromatic residue of tryptophan and the α -carbon of lysine, was previously identified as RiPP, again hinting towards the ribosomal origin of the Darobactins. Subsequently, tBLASTn searches using the AA sequence WNWSKSF were performed against the producer strain's genome, however, giving no hits and further investigation and tBLASTn search of public databases, run without any restrictions did not yield any convincing result either. However, restricting the query to only *Photorhabdus* genomes quickly reported the candidate BGC in multiple strains of the *Photorhabdus* genus. The candidate BGC consisted of a short propeptide (*darA*), three transporter genes (*darBCD*) and one enzyme of the radical SAM (RaS) superfamily of enzymes (*darE*). Sizewise, this small cluster is comparable to the Streptide BGC, which also comprises 5 genes, only two of are involved in Streptide biosynthesis (Schramma *et al.*, 2015). In contrast to the Streptide BGC, three individual transporter genes could be identified coding for separate parts of the ABC type transenvelope transporter while no regulatory genes could be identified. While the overall size of the identified cluster is small compared to well investigated RiPPs such as the Thiopeptides (Liao *et al.*, 2009) and Lanthipeptides (Cheigh and Pyun, 2005), this is not surprising since the size of RiPP BGCs usually reflects the degree of functionalisation of the respective propeptide (Number of chemically different functionalisations $\approx$ number of genes). However surprisingly, the identified BGC cluster lacks any type of regulatory genes as well as dedicated peptidases, usually required for maturation of the molecule (Arnison *et al.*, 2013). Finding of this RaS enzyme instantly hardened the confidence that this BGC is responsible for Darobactin production, since the previously identified Streptide also requires a single RaS enzyme to manufacture the covalent C–C bond. With this information, all genes of the candidate BGC were searched in the producer strains genome sequence, only to find that the RaS encoding sequence was partially present at the edge of a contig. Resequencing of *Photorhabdus khanii* HGB1456 finally revealed the presence of the full candidate BGC in this strain, as well as in the second confirmed producer strain *Photorhabdus khanii* DSM3369. While at first it was not obvious which genes in the general vicinity of *darA* were part of the cluster, the BGC was easily narrowed

down to the final five genes. First indication that the BGC is limited to these five genes was that all they had significantly lower GC% content compared to the neighbouring genes, which is a hint to a common evolutionary origin that is different from the neighbouring genes with higher GC% (Fig.: 15). This also showed that this BGC was inserted into its genomic locus in one block. Second observation that backed this up was that some *Photorhabdus* and closely related *Xenorhabdus* strains do not carry the BGC. Here the locus of the neighbouring genes looks identical, just lacking the five genes of the Darobactin BGC in between them (Fig.: 15). Apparently, this represents a different state in evolution, where the Darobactin BGC was not yet integrated into its locus or was already lost; however, making it clear that the identified BGC is limited to these five genes. The layout of the locus and especially the low GC% content hint towards the acquisition of this BGC by *Photorhabdus* producer strains by horizontal gene transfer. The fact that possession of the BGC is rather the normal state than exception shows that derives an ecological benefit from keeping the cluster as is expected for the BGC of a potent antibacterial compound; thus supporting the assumption that the BGC codes for Darobactin.

To prove that this candidate BGC is responsible for the production of Darobactin, knock out (KO) experiments as well as heterologous expression experiments were designed and performed. While heterologous expression (and subsequent detection of the product) remains gold standard to confirm the responsibility of genes for a certain reaction or product, KO experiments are a reliable way to get corroboration of the involvement of specific genes. While successful heterologous expression of NP BGCs from the *Photorhabdus* genus in *E. coli* is well documented for non-ribosomal peptides and polyketides (Cai *et al.*, 2019, Zhou *et al.*, 2019) this is no guaranteed success. Especially, heterologous expression of the RaS enzyme with the required correct loading of iron-sulphur clusters (FeS) was seen as potential bottleneck due to the fact that heterologous production of iron-sulfur cluster containing proteins in *E. coli* can be troublesome. Co-overexpression of the *isc* operon seems to work for some proteins in *E. coli*, however it remains unclear which factors contribute to successful FeS loading of heterologously produced protein in *E. coli* (Jaganaman *et al.*, 2006).

For quick confirmation, the first experiment was to disrupt the *darE* ORF, which led to abolishment of Darobactin production. Subsequently, to test whether reintroduction of the genes restores production, the whole BGC was deleted markerlessly, again abolishing production of the compound. Surprisingly, only the second identified producer strain *Photorhabdus kanii* DSM3369 was susceptible to the genetic modification while *Photorhabdus kanii* HGB1456 could not be conjugated with the same constructs. This may be explained by the presence of a restriction system, degrading incoming DNA with non-matching methylation pattern to prevent unwanted horizontal gene transfer. This is only mildly surprising, since similar behaviour can be observed even in closely

related bacterial species such as *Streptomyces coelicolor* A3(2) and its derivatives compared to *S. lividans* TK24 (Practical *Streptomyces* Genetics, John Innes Foundation).

Subsequent plasmid borne reintroduction of the BGC under control of a stronger promotor resulted in only slightly visible production in *Phototrhaddus kanii* DSM3369. In contrast, inducing expression in *E. coli* BW25113 carrying the same plasmid, led to well visible production of Darobactin. This result was surprising at first, since naturally one would assume that the original producer would be the best host for overexpression, however this is not necessarily the case. Earlier studies in *Streptomyces* show that the amount of NP produced is not clearly linked to the strength of expression, i.e. the amount of mRNA produced, but rather is finely tuned among the seamless interaction of the participating enzymes (Myronovskyi and Luzhetskyy, 2016). Literature suggests that a lot of NP biosynthetic pathways undergo internal regulation mechanisms that integrate with the primary metabolism in one way or the other (Craney *et al.*, 2012). Changing expression levels of one of the enzymes might have detrimental downstream effects to the whole production due to feedback regulation within the BGC. In contrast, taking genes out of this regulatory context by rearrangement of the BGC, introduction of artificial promoters and subsequent heterologous expression sometimes can boost production as was demonstrated by overexpression of myxochromide (Perlova *et al.*, 2009).

Having proven the connection between the identified BGC and Darobactin production and KO, the further question at hand was which of the five identified enzymes are directly involved in the manufacture of Darobactin. While the role of DarA as propeptide was clear, it was unclear which enzymes are responsible for the introduction of the aromatic aliphatic ether bond (C–O–C) and the aromatic aliphatic covalent bond (C–C). Since Streptide has the same aromatic aliphatic covalent bond, it was reasonable to assume that this is mediated by the RaS DarE; however, the origin of the aromatic aliphatic ether bond remained unclear. Different hypotheses were i) that DarE installs the C–C covalent bond and the C–O–C bond is installed by one of the genes annotated as transporters or *in-trans* by a gene present in *Phototrhaddus* and *E. coli*, or ii) that DarE installs both rings into the propeptide, which were so far unprecedented to the best knowledge. Heterologous expression of the BGC, deleting *darB*, *darC* and *darD* one after the other, showed that simultaneous expression of *darA* and *darE* is sufficient to manufacture Darobactin, hence hypothesis ii) was most likely to be correct. Final proof of this hypothesis would be *in-vitro* reconstruction of the reaction which will be part of future studies. This experiment also tested another important factor in the Darobactin biosynthesis: The identified BGC completely lacks any peptidases associated with the final trim of the molecule from the modified propeptide and the propeptide lacks any peptidase recognition sites such as the double glycine motif, utilised in a multitude of RiPP maturation processes such as the

Lichenicidin pathway (Letzel *et al.*, 2014). Here again, different scenarios can be envisioned: i) final trim is performed by DarE during the modification steps, possibly being relevant for the reaction and hence assigning yet another function to DarE; ii) final trim is performed by the transporter proteins possibly during export of the cell as is the case for class II lanthipeptide maturation (Severi and Thomas, 2019) or iii) final trim is performed *in trans* by unspecific or specific ubiquitous peptidases from the producer strain and the heterologous hosts genetic background. By deletion of the transporter genes, it was shown that these are not involved in the final trim; however, it still is unclear which of the two remaining scenarios is correct. While it might be a far stretch to assume that Darobactin is trimmed *in-trans* by peptidases that are encoded in both, the original producers and the heterologous hosts genome, this might as well be very likely due to the evolutionary relationship as both are members of the order *Enterobacterales*. The alternative scenario i) that DarE is mediating the introduction of two rings with distinct different chemistries as well as cleavage of peptide bonds in two separate positions (C and N-terminal of the final Darobactin) appears too far of a stretch for an enzyme with very narrow and distinct mode of action. To summarise, despite collecting evidence that DarE is introducing the two ring systems into the DarA propeptide, the further maturation i.e. cleaving process of the modified propeptide into final Darobactin remains unclear.

With the experiments at hand, the minimal Darobactin BGC was identified and confirmed, i.e. it was shown that only the two proteins DarA and DarE are sufficient to produce Darobactin in *E. coli*. This however only opens door for more thorough investigation. While the biological formation of C-C covalent bonds is common, the formation of covalent bonds including non-activated carbon atoms is predominantly mediated by two chemical reactions. The first being Claisen condensation which is the defining reaction in biosynthesis of fatty acids and polyketides and the second being the formation of members of the radical SAM (RaS) superfamily of enzymes as found in the biosynthesis of Streptide, hence, most probably also Darobactin. However, while the mechanism for forming the Darobactin C-C bond can be expected to be the same as in Streptide, the mechanism for forming the C-O-C bond must be completely different. In Streptide (hence most probably in Darobactin) biosynthesis, S-adenosylmethionine gets broken down to a 5'-deoxyadenosyl radical which abstracts a hydrogen from the respective α -carbon forming a secondary radical in the K⁵ α -carbon position. This in turn is attacked by the aromatic residue of tryptophan forming the C-C bond (Schramma *et al.*, 2015). However, neither the energy of the 5'-deoxyadenosyl radical nor the secondary radical is sufficient to attack oxygen from abundant sources, such as H₂O, hinting towards a new, different reaction mechanism. While ether bond formation by RaS enzymes is reported, it is limited to thioether formation, such as the sulphur-to- α -carbon links, eponymous for the Sactipeptide

compound class (Flühe and Marahiel, 2013). The alternative of an *in trans* acting enzyme also might be true, possibly involving intermediate steps such as hydroxylation in one or both positions involved in the ring closure. However, no corresponding peptide modification at this aromatic position of tryptophan was previously reported in *E. coli* and no plausible intermediary product could be detected in the course of these experiments. In consequence, while on the one hand providing a good understanding of the Darobactin biosynthesis, these experiments on the other hand pave the way for further experiments to elucidate the exact reaction mechanisms involved in introduction of the C – O – C bond as well as the final trim of the molecule.

IV.4.2. Mutasynthetic production of Darobactin analogues

While investigating the Darobactin BGC, BLASTp searches for all five genes were performed not only against genomes from the *Photorhabdus* genus but also against all sequences deposited in GeneBank. This revealed a plethora of homologous or orthologous clusters from different bacterial genera such as *Photorhabdus*, *Yersinia*, *Vibrio* and *Pseudoalteromonas* with the majority found in *Yersinia*. Whether this overrepresentation of this cluster is due to the higher number of sequenced *Yersinia* genomes cannot be concluded. Peculiarly, some of the *Photorhabdus* strains do also carry a second propeptide with altered core amino acid sequence and follower AAs, already hinting towards more molecules of the Darobactin class to be discovered. Furthermore, some of the propeptides in orthologous clusters in *Yersinia* also showed altered core AA sequences, despite being the only propeptide in the respective cluster. The novel putative Darobactin analogues were termed Darobactin B for the second propeptide sequence in some *Photorhabdus* strains and CDE in *Yersinia* strains in no specific order. Closer inspection of the nucleotide sequences of *Yersinia* propeptides with altered core AA sequence showed that all these modifications to the core AA sequence can be explained with only minor point mutations to one original Darobactin A propeptide. Here it also has to be noted that, while the propeptide sequences in *Photorhabdus*, *Vibrio* and *Pseudoalteromonas* are very consistent, without big differences between individual species of the same genus, this is not true for the propeptide sequences from *Yersinia*, where the propeptides have much higher sequence variation. Whether this is due to the mere higher abundance of *Yersinia* propeptide sequences or due to higher evolution rate of *Yersinia* based on higher genetic instability cannot be answered.

This however immediately raised question, whether the observed sequences translate into i) manufacturable products with the given biosynthetic machinery and if so, ii) whether the products show similar activity compared to Darobactin A. To test this, the previous heterologous expression system was modified to incorporate the altered AA sequences into the Darobactin A propeptide from *P. khanii* DSM3369. After successful cloning, the resulting constructs were expressed and

analysed by LCMS, assuming the ring closures remain the same. However, no masses corresponding to the respective products except for Darobactin E with low intensity were detected. Subsequently, after verifying that the three follower AAs from the Darobactin A propeptide are expendable, the constructs were re-cloned, leaving out the follower AAs and all products could promptly be found by LCMS investigation. Here, vast differences in production level could be observed. While Darobactin B, the only sequence from a distinct additional propeptide was produced well, the other compounds were produced only in minor amounts. Worst production was observed for Darobactins C and D, the compounds with $K^5 \rightarrow R^5$, where the AA exchange is in one of the AAs undergoing posttranslational modification. Here, a clear order could be observed: In comparison to Darobactin A production Darobactin E ($N^2 \rightarrow S^2$) was produced in a roughly tenfold lower concentration and Darobactin D ($K^5 \rightarrow R^5$) even less. Darobactin C, being the combination of both ($N^2 \rightarrow S^2$; $K^5 \rightarrow R^5$) was produced lowest of all. The Darobactin F ($W^1-K^2-W^3-S^4-K^5-N^6-L^7$) sequence was discovered the latest and seems to be a degenerated version of a Darobactin B propeptide. Here, observed production was in the general range of Darobactin B, showing that even multiple modifications do not necessarily impair production drastically. This shows that DarE from *Photorhabdus* will accept and process modified propeptides. However, based on the result that production level dropped significantly when modifying the K^5 position, it seems to not tolerate any modification readily. Rather, the acceptance of modifications in the Darobactin core AA sequence needs to be tested for each position and AA possible. Furthermore, while omitting the follower AAs of the propeptide did not have any impact on the synthesis of the Darobactins, shortening the heptameric core AA sequence was not tolerated at all, hinting towards the reaction requiring the full seven AAs to fill the active site in order to manufacture the intramolecular rings as observed in Darobactin A. This shows that while the RaS reaction mechanism obviously is vital to the formation, the layout of the active site also plays a crucial role in the formation of the two Darobactin-defining ring systems. By defining the proximity of the respective reaction partners, they are restricted to using what is presented rather than what is thermodynamically preferable. This makes the interior design of the active site cavity responsible that the respective covalent bonds are only (or predominantly) formed in the observed positions. This also sparks the need to elucidate the 3D-structure of the DarE by crystallisation, to better exploit or possibly even modify the DarE active site to better fit altered core AA combinations.

While these experiments were only performed with DarE from *Photorhabdus kanii* DSM3369 and therefore cannot be directly transferred to the *Yersinia* DarE ortholog, these results hint that from the propeptides with altered AA sequences only the second propeptide from *Photorhabdus* coding for Darobactin B can be produced in the same order of magnitude as Darobactin A. It is reasonable to assume this being valid also for the DarE orthologs from other genera than *Photorhabdus*, due to

the vast majority of Darobactin A propeptides compared to limited numbers of other Darobactin core AA sequences. This indicates that only Darobactin A and B may represent functional compounds, while the Darobactin derivatives from *Yersinia* with altered AA composition might be considered artifacts of sequencing or artifacts of evolution as first step on the way to loss of the cluster.

To isolate the compounds identified by LCMS for detailed characterisation, the respective clusters needed to be re-cloned to a higher copy number vector with a strong promotor, e.g. the T7 promotor system, which was shown to increase production drastically (Zerlina G. Wuisan, unpublished). However, even with the better producing constructs, isolation was not an easy task. The amount recovered from heterologous production was in the same order as intensity of the compounds in the initial LCMS analysis. Hence, the highest isolated amount was Darobactin B followed by Darobactin E and D. Darobactin C could not be isolated at all, due to the low production yield and high loss during attempted purification.

Subsequent testing of the MICs against common Gram- pathogens as well as clinical isolates from Giessen revealed that Darobactin A and B are similarly active against the whole panel tested. Even clinical *E. coli* isolates that were completely resistant to Rifampicin, Tetracyclin and Gentamycin were killed by Darobactin A and B with low MIC values. In contrast, Darobactin D and E had significantly worse activity against the whole panel with 8 – 16 fold elevated MIC compared to Darobactin A. This shows that the altered core AAs found in *Yersinia* propeptides perform worse, which again raises question whether these are actually “real” compounds rather than artifacts from an evolutionary degeneration process (not taking into account possible sequencing errors). Evolution nearly without exception maximises instead of decreases the activity of NPs (Fewer and Metsä-Ketelä, 2020). This follows the simple idea that a more potent compound is simply better at killing competitors or that a more active compound does not require the same high production as a weaker ancestor, making it possible to save on resources for manufacturing. This leads to the conclusion that of the here studied compounds, only Darobactin A and B represent “real” NPs, while the modified core AA sequences in *Yersinia* strains represent either degenerated versions of the Darobactin A cluster or possibly (however unlikely) sequencing errors. In contrast to the presented mutasynthetic approach, investigation of organisms with the second propeptide (in case of *Photorhabdus*) or altered core AA sequence in the only propeptide (in case of *Yersinia*) for the respective Darobactin analogues was not successful at all. None of the here reported compounds could be found by activity or mass from the respective WT strain (personal communication with Robert Green, ADC, Northeastern University, Boston, USA). One final key issue that needs to be addressed is the difference in MIC of Darobactin A observed in these experiments compared to Darobactin B and to the published results. This can

easily explained by the different readout techniques (microtiter plate reader vs. read out by eye (Imai *et al.*, 2019).) and that MIC values of different compounds were recorded on different days as well as the huge variety of available pure compound to be tested (12 mg vs. 1.3 mg). Here, seemingly negligible factors can contribute to the observed difference such as precise dilution of the compounds or fitness of the precultures. Since the general trend remains the same i.e. the activity profile against the panel tested matches the published results, both results can be considered consistent.

Central advantage of the engineering of RiPP pathways is the ease of introducing modifications by simply changing codons in the respective propeptide. In contrast to large multidomain NRPS and PKS pathways that require seamless integration of all enzymes and usually allow only little tinkering, simply exchanging codons on DNA level is done fast and does require less preoccupation with identifying interlocking domains and general layout of the enzymatic interaction. The intriguing flexibility of RiPP pathways was recognised before and made use of for other NPs biosynthetic pathways. To date, a multitude of RiPPs pathways from different classes such as the Microcins (lasso peptide), the Prochlorosins (lanthipeptide) or the Thiostreptone (thiopeptide) underwent studies that tested and exploited the promiscuous nature of their respective pathways (Hudson & Mitchell, 2018). Most intriguing efforts in were recently reported: Mutasynthetic creation of a 10^6 membered library combined with simultaneous *in-strain* protein-protein interaction screening allowed identification of an inhibitor of the HIV p6 protein, crucial for maturation of viral particles in infected cells (Yang *et al.*, 2018). Furthermore, coupling mutasynthetic derivation of RiPPs with microfluidics facilitates labour intensive screening of clone libraries, as was demonstrated by identification of 126 novel antimicrobial lanthipeptides by microfluidics based screening (Schmitt *et al.*, 2019).

The presented experiments show that there is at least one natural Darobactin A analogue, which could be produced by engineering the biosynthetic machinery for Darobactin A, alongside more analogues that might be artefacts of evolution. This also shows that the essential biosynthetic enzyme DarE has a high tolerance to accept different AA combinations in the core peptide, however not all combinations will be produced well or probably even be produced at all. Especially modifications of AAs participating in the intramolecular ring formation will most probably not be tolerated easily as was demonstrated for position 5 by changing K⁵ -> R⁵. Having shown the general possibility of mutasynthetic derivation of the Darobactin core structure, this again opens door for further experiments. In the future, experiments will be performed to elucidate the tolerance of DarE for all AAs in all positions as well as combinations thereof. In turn, this will create a vast library for testing against pathogens in the hope to find even more active analogues than Darobactin A or analogues with an altered activity profile, e.g. better activity towards other relevant Gram-

pathogens. This will not only be limited to mutasynthesis but also semi-synthetic derivation of the core structure. This is why better production of Darobactin A and its analogues and better understanding thereof is vital for the further development of this class of compounds possibly into a marketable drug. Furthermore, this will also be a good tool to enable study of the structure activity relationship. This will also be necessary since the mode of action against BamA is still not fully understood and more data will help elucidating this in detail. Better understanding of the anti-BamA mode of action in consequence will also contribute to identifying more compounds with the same mode of action and thus help filling the antibiotic pipeline again.

IV.4.3. Identification of natural Darobactin A analogues from *Pseudoalteromonas luteoviolacea* H33 and H33S

Similar to the identification of Darobactin A analogues with altered core AA sequence, the following findings root in the bioinformatic investigation of the Darobactin homolog and ortholog BGCs from genomes deposited in GeneBank. However, what raised attention was not an alteration of the core AA sequence but the finding that some ortholog BGCs carry additional genes. While BGCs in *Vibrio* strains harbour an additional transporter gene that seems of no interest, certain *Pseudoalteromonas* species harbour two additional genes compared to the *Vibrio* BGC. One gene coding for a hypothetical protein and one gene that was annotated to code for an FAD-dependent oxidoreductase. This again immediately sparked hope that these enzymes modify the Darobactin core structure beyond exchange of certain AAs and fuelled the desire to identify the associated compounds.

In order to identify these associated compounds, an OSMAC approach was performed with *Pse. luteoviolacea* H33 and H33S and MS/MS based investigation revealed production of three Darobactin A derivatives in said organisms. The first observed modification was the introduction of a double bond (Dehydrodarobactin A), altering the chemical properties of Darobactin A to be less hydrophilic. The second modification found was the introduction of a bromine atom into the aromatic system of W¹ (Marinodarobactin A), also shifting the chemical properties towards lower hydrophilicity. Finally, it could be observed that both modifications at once (Marinodehydrodarobactin A) significantly shift retention time towards higher organic solvent content in the used UPLC-MS system, indicating an additive or even synergistic effect of both modifications on the hydrophilic nature of Darobactin A. Linking the respective modifications to their respective enzyme however will remain difficult due to usually poor heterologous expression of *Pseudoalteromonas* genes in *E. coli* and relative unease to introduce genetic modifications to *Pseudoalteromonas*. Halogenation of aromatic systems is usually performed via electrophilic

halogenation where initially the participating halogen is oxidised and the resulting product is subsequently attacked by the aromatic electron system (Busch *et al.*, 2019). This is well in tune with the reaction mechanism of an oxidoreductase and therefore a good hint towards the function of the FAD-dependent oxidoreductase in Darobactin modification. This also is well in tune with literature reports of halogenations of NPs and specifically the biosynthesis of brominated pyrrolic-phenolic compounds such as Pentabromopseudilin, the first marine antibiotic, found in a variety of marine bacteria strains (Burkholder *et al.*, 1966, Busch *et al.*, 2019). The dehydration however cannot be readily attributed to any of the enzymes coded in the BGC. Several dehydration reactions in α -carbon position are reported for non-ribosomal and ribosomal peptides, most commonly from bacteria and usually associated with antibacterial or antifungal activity (Siodlak, 2015). While introduction of dehydro AAs in natural products is thoroughly investigated, e.g. in the lanthipeptide biosynthesis (Yu *et al.*, 2013) and dehydration in general, e.g. in form of desaturation in fatty acid metabolism is well understood, it is not clear which kind of reaction takes place here. One possibility would be the introduction of a hydroxyl group at the desired position with subsequent dehydration, leaving behind the observed double bond. However, no plausible intermediate product could be identified by mass, which either indicates a one-step process or that a hypothetical hydroxylated intermediate product is not released in any detectable quantity. Other reaction mechanisms include the direct abstraction of a hydrogen atom by an oxidoreductase or a radical mechanism. Making an educated guess is further impeded by the fact that the location of the double bond is not yet clear. Due to the low recovery of the compounds carrying the double bond, its location could not be identified by NMR studies yet. In the light of these uncertainties, it still seems reasonable to assume that the halogenation is mediated by the FAD-dependent oxidoreductase (essentially making it a halogenase), since the proposed reaction mechanism is fitting well to the general function and only specifically monobrominated Darobactin was identified, ruling out *in-trans* promiscuous halogenation of the final molecule or its free AA precursors. For dehydration, since neither the exact location nor the reaction mechanism is clear, it is more difficult to speculate about the biosynthetic origin. Since the only two remaining genes of the cluster annotate as transporter and hypothetical protein of unknown function, the best guess would probably be the second gene in the BGC coding for the hypothetical protein with a domain that is commonly found in RaS enzymes however not directly linked to the RaS functionality. In conclusion, this modification cannot be readily explained by one of the genes contained in the BGC and it cannot be ruled out that it is performed *in trans* by another (promiscuous) enzyme from the genetic background of this strain. From the observed intensities in the LCMS, it is also reasonable to assume that Marinodehydrodarobactin A (- H₂, + Br) is the main product of this BGC while Dehydrodarobactin A (-

H₂) and Marinodarobactin A (+ Br) are side products of its biosynthesis. However, in which exact order these modifications take place cannot be answered yet.

Despite low recovery of the compounds, activities of Marinodarobactin A and B could be tested against the panel of clinically relevant pathogens. Here it could be observed that both modifications have only weak impact on the activity of Darobactin A. While Marinodarobactin A had similar activity profile compared to Darobactin A and B, Marinodehydrodarobactin A was only slightly weaker. It therefore can be assumed that both modifications do not or only slightly change the interaction of the molecules with their BamA target. Since only human pathogens were tested in this study, it might be that the dehydration optimises the compound for better activity against competitors of *Pse. luteoviolaceae* in its natural habitat such as *Vibrio* strains. To gain a holistic understanding, further investigation is needed. Therefore, the compounds need to be isolated in larger amounts to clarify structural uncertainties and to test activities against a broader panel, including marine bacteria. Simultaneously, progress in the understanding of the structure – activity relationship of the Darobactins is needed to elucidate how the molecules behave towards their target and what impact respective modifications have on this interaction.

In contrast to their influence on the activities, both modifications have strong influence on the chemical properties of the molecule, again leading to speculation about the role in its producer's habitat. From the experiments performed, it became obvious that production of Darobactin A and its marine derivatives in *Pse. luteoviolacea* H33 and H33S are linked to the formation of biofilms that settle on the surface of the medium, since Darobactin (analogue) production could only be observed upon biofilm formation. This biofilm formation also caused colour changes in the culture from yellow to dark purple. The colour change was inferred to be caused by violacein, a common bacterial indole pigment with medium antibacterial properties. Violacein and other natural product production has previously been linked to biofilm formation in *Pseudoalteromonas* strains, hence it is not surprising for Darobactin production to undergo a similar regulation (Ayé *et al.*, 2015). This biofilm formation in *Pseudoalteromonads* is regulated by an N-acetyl-homoserinelactone messenger systems (Ayé *et al.*, 2015) and plays an important role in the natural habitat. *Pseudoalteromonads* are commonly associated with higher organisms, where biofilm formation is one way to adhere to the eukaryotic host. In this environment, production of more lipophilic antibiotic compounds might be beneficial due to better penetration of biofilms and host tissue. This would make the novel identified Darobactin A analogues better suited for the environment of their producer. This however might also be beneficial for the development of these compounds into a drug, since enhanced tissue penetration is usually desirable for any drug supposed for systemic administration.

These experiments show that the initially identified extended Darobactin BGC in *Pse. luteoviolacea* H33 and H33S indeed is responsible for production of additional derivatives of Darobactin A by halogenation and dehydration of the core structure. Hereby, the producer strain is the first WT strain outside the genus of *Photorhabdus* in which Darobactin production could be observed. Again, this raises further questions that need to be addressed in future studies. First, the exact location of the bromination on the aromatic W¹ residue and the location of the introduced double bond needs to be elucidated by NMR studies and their biosynthetic origin, i.e. the responsible genes need to be identified and corroborated. One particularly interesting question in this regard is whether the responsible enzymes act on free Darobactin or on the propeptide. In order to further investigate the compounds properties, production has to be improved drastically, ideally by finding a suited heterologous expression system that is easy to upscale. Once production is ramped up, these novel identified compounds will undergo extensive *in-vitro* and *in-vivo* antimicrobial testing to determine whether they have any advantage in application over non-modified Darobactin A, especially regarding tissue penetration. Furthermore, upon clarification of the exact biosynthetic route, future studies will determine whether these modifications are biosynthetically limited to Darobactin A, or if they can be introduced to other mutasynthetically created Darobactin analogues. This would again drastically increase the number of possible structures to undergo closer investigation in order to find more active Darobactin analogues, e.g. analogues with altered activity profile or analogues in general better suited for further drug development.

V. SUMMARY AND OUTLOOK

The discovery and development of antibiotics in the early 20th century sparked a medicinal revolution, enabling humankind to effectively combat the threat of bacterial infectious diseases. Despite great success, the discovery and development of new antibiotics declined over the years, while simultaneously resistances to all clinically relevant antibiotics arose. This is of great concern, since this situation drastically impedes the ability to treat bacterial infections successfully and shows that reinvigoration of antibiotic research is desperately needed. In the presented thesis, three separate projects in the field of antibiotic natural product research are presented:

Due to the seeming depletion of terrestrial bioresources, especially soil borne *Actinomycetes* for discovery of novel antimicrobial natural products (NPs), bioprospecting other ecological niches, such as the marine environment, moved into focus in the recent years. Chapter I – “Antimicrobial Potential of Bacteria Associated with Marine Sea Slugs from North Sulawesi, Indonesia” describes the efforts undertaken to investigate nudibranch-associated bacteria for their potential to produce antimicrobial NPs. Nudibranch samples and to minor degree sponge material were retrieved by diving from the reef biodiversity hotspot around Bunaken Island, Indonesia and subsequently, bacteria were isolated. In total, 49 bacterial strains were retrieved and initially characterised for antibacterial activity. Among the isolated strains were mostly common marine bacterial genera such as *Vibrio* and *Pseudoalteromonas*; however, also rare bacterial species such as *Pelagibaca bermudensis* could be retrieved. Of 49 axenic isolates, 71 % exhibited at least weak antibacterial activities; however, three strains (2 *Pseudoalteromonas* and 1 *Vibrio* strain) exhibited anti-MRSA activity and one isolate (*Marinomonas communis*) exhibited anti-EHEC activity. This shows high potential for antimicrobial NP production within this habitat and can be considered a first dip into this untapped habitat, which will remain worth investigating for new bioactive metabolites in the future. While no novel compound could be isolated from this bacterial collection so far, reinvestigation will be worth it, especially with the knowledge gained in the projects described in Chapter II and III.

In Chapter II, the biosynthesis of the Streptomyces antibacterial NP Pseudouridimycin (PUM) was investigated. The very fragmented PUM biosynthetic gene cluster was previously identified and characterised by Sosio *et al.*, 2018 by generation of knock-out (KO) mutant versions of the WT producer strain. Based on these published results, three different refactored versions were designed and cloned to identify the minimal set of genes required for PUM biosynthesis. These refactored BGCs were expressed under control of strong promoters. Underlying idea was the generation of a plug and play platform for heterologous production of PUM and its biosynthetic precursors that can

easily be assembled in *E. coli* and shuttled to the respective Streptomyces surrogate host. During the course of these experiments, it became clear, that not all genes contributing to the biosynthesis of PUM were identified in the first characterisation. Expressing only genes to which a clear biosynthetic step could be assigned by Sosio *et al.*, 2018 did not result in formation of detectable amounts of PUM. Subsequent reinvestigation of the BGC allowed identification of the putative serine/threonine kinase encoding gene *pumF* as candidate gene to be involved in PUM biosynthesis. Incorporation of this gene into the streamlined construct led to detectable production of PUM and hence, clarification on the minimal set of genes required for PUM production. While PumF does not participate directly in the biosynthesis, by posttranslational phosphorylation it serves as a positive regulator of PUM biosynthesis on protein level. However, the ultimate goal of creating a plug and play platform for selective production of PUM and its precursors for further semi-synthetic derivation and subsequent study of this antibacterial NP class for possible further development could not be reached due to weak production of PUM in this heterologous system. Since regulation has to be considered to improve the yields, future efforts will be directed towards better understanding of the regulation within this BGC in order to generate a better suited heterologous expression system. Additionally, optimisation of the WT producer strain will remain a last resort option.

The final chapter describes the identification and initial characterisation of the Darobactin biosynthesis. Darobactin A is a novel antibacterial NP that inhibits BamA, an essential chaperone for outer membrane proteins, thus selectively killing only Gram- bacteria (MIC 2 µg/mL against *E. coli*). This is particularly interesting, since antibiotic resistant Gram- bacteria are considered the most imminent bacterial threat to human health today. Here, it was possible to identify and corroborate the BGC responsible for Darobactin A formation by generation of KO mutants of the WT producer strain *Photobacterium khanii* DSM3369 and heterologous expression in *E. coli*. Furthermore, identification of the Darobactin (minimal) BGC allowed subsequent identification of Darobactin A analogues in the genera *Photobacterium*, *Yersinia* and *Pseudoalteromonas*, all members of the class of γ -proteobacteria. These *in-silico* identified analogues could be readily identified by LCMS either from their WT producers *Pseudoalteromonas luteoviolacea* or after reprogramming the Darobactin A biosynthetic machinery from *P. khanii* DSM3369 to produce derivatives mutasynthetically in *E. coli*. Subsequently, the identified Darobactin analogues were purified and their antibacterial activities were tested against a panel of clinically relevant Gram- pathogens. Activity tests revealed that the analogues Darobactin B, Marinodarobactin A and Marinodehydrodarobactin A had similarly potent antibacterial properties compared to Darobactin A, while the activities of Darobactin D and Darobactin E were substantially lower. While none of the identified analogues exhibited drastically

better activity than Darobactin A, especially the derivatives from the marine environment exhibit significant changed chemical properties, *e.g.* their solubility in water.

Due to its good activities against Gram- bacteria as well as their novel mode of action, the Darobactins represent a promising class for further drug development. The presented results expand this class with novel analogues as well as with insights into their biosynthesis that will also allow creation of non-natural Darobactin analogues. Derivation of the Darobactin core structure will be important for thorough investigation of the structure activity relationship. Also, the identified Darobactin derivatives with significantly altered chemical properties might be especially interesting for further development, due to better or different pharmacological properties. It is envisaged to further develop this compound class and proceed to pre-clinical and hopefully at some point clinical trials to combat the threat of multi resistant Gram- bacterial infections. Furthermore, thorough investigation of the Darobactin mode of action will give insight into this fascinating new target and allow detection of more compounds (NPs and synthetic ones) with the same mode of action, thus contributing to filling the antibiotic pipeline again.

VI. REFERENCES

- Akita, H., Nakashima, N. and Hoshino, T. (2015). Bacterial production of isobutanol without expensive reagents. *Appl Microbiol Biotechnol.* 99(2), 991-999
- Amiri Moghaddam, J., Crüsemann, M., Alanjary, M., Harms, H., Dávila-Céspedes, A., Blom, J., Poehlein, A., Ziemert, N., König, G. M., & Schäberle, T. F. (2018). Analysis of the Genome and Metabolome of Marine Myxobacteria Reveals High Potential for Biosynthesis of Novel Specialized Metabolites. *Scientific reports*, 8(1), 16600
- Amos, G., Awakawa, T., Tuttle, R. N., Letzel, A. C., Kim, M. C., Kudo, Y., Fenical, W., Moore, B. S., & Jensen, P. R. (2017). Comparative transcriptomics as a guide to natural product discovery and biosynthetic gene cluster functionality. *PNAS*, 114(52), E111121–E111130
- Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K. D., Fischbach, M. A., Garavelli, J. S., Göransson, U., ... van der Donk, W. A. (2013). Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Natural product reports*, 30(1), 108–160
- Astrin, J., Stüben, P. (2008). Phylogeny in cryptic weevils: molecules, morphology and new genera of western Palaearctic Cryptorhynchinae (Coleoptera: Curculionidae). *Invertebrate Systematics, CSIRO*, 22 (5), 503–522
- Awakawa, T., Fujioka, T., Zhang, L., Hoshino, S., Hu, Z., Hashimoto, J., Kozono, I., Ikeda, H., Shin-Ya, K., Liu, W., & Abe, I. (2018). Reprogramming of the antimycin NRPS-PKS assembly lines inspired by gene evolution. *Nature communications*, 9(1), 3534
- Bauman, K. D., Li, J., Murata, K., Mantovani, S. M., Dahesh, S., Nizet, V., Luhavaya, H., & Moore, B. S. (2019). Refactoring the Cryptic Streptophenazine Biosynthetic Gene Cluster Unites Phenazine, Polyketide, and Nonribosomal Peptide Biochemistry. *Cell chemical biology*, 26(5), 724–736.e7
- Benedictow, O.J. (2006). The Black Death, 1346-1353: The Complete History
- Bhushan, A., Egli, P. J., Peters, E. E., Freeman, M. F., & Piel, J. (2019). Genome mining- and synthetic biology-enabled production of hypermodified peptides. *Nature chemistry*, 11(10), 931–939
- Bogdanov, A., Kehraus, S., Bleidissel, S., Preisfeld, G., Schillo, D., Piel, J. et al. (2014). Defense in the Aeolidioidean Genus Phyllodesmium (Gastropoda). *J Chem Ecol*, 40(9), 1013-1024
- Bogdanov, A., Hertzner, C., Kehraus, S., Nietzer, S., Rohde, S., Schupp, P.J. et al. (2016). Defensive Diterpene from the Aeolidioidean *Phyllodesmium longicirrum*. *J Nat Prod*, 79(3), 611-615
- Bowling, J. J., Kochanowska, A. J., Kasanah, N., & Hamann, M. T. (2007). Nature's bounty - drug discovery from the sea. *Expert opinion on drug discovery*, 2(11), 1505–1522
- Brodie, G.D., Wägele, H., Klusmann-Kolb, A. (1999). Histological investigations on *Dendrodoris nigra* (Stimpson, 1855) (Gastropoda, Nudibranchia, Dendrodorididae). *Molluscan Research*, 20(1), 79-94
- Busch, J., Agarwal, V., Schorn, M., Machado, H., Moore, B. S., Rouse, G. W., Gram, L., & Jensen, P. R. (2019). Diversity and distribution of the bmp gene cluster and its Polybrominated products in the genus *Pseudoalteromonas*. *Environmental microbiology*, 21(5), 1575–1585
- Burkholder, P. R., Pfister, R. M., & Leitz, F. H. (1966). Production of a pyrrole antibiotic by a marine bacterium. *Applied microbiology*, 14(4), 649–653.
- Bunch, R.L., McGuire, J.M. (1952). US2653899A - Erythromycin, its salts, and method of preparation
- Cai, X., Zhao, L., & Bode, H. B. (2019). Reprogramming Promiscuous Nonribosomal Peptide Synthetases for Production of Specific Peptides. *Organic letters*, 21(7), 2116–2120
- Cheigh, C. I., & Pyun, Y. R. (2005). Nisin biosynthesis and its properties. *Biotechnology letters*, 27(21), 1641–1648

- Choudhury, J.D., Pramanik, A., Webster, N.S., Lewellyn, L.E., Gachhui, R., Mukherjee, J. (2015). The Pathogen of the Great Barrier Reef Sponge *Rhopaloeides odorabile* Is a New Strain of *Pseudoalteromonas agarivorans* Containing Abundant and Diverse Virulence-Related Genes. *Mar Biotechnol*, 17(4), 463-478
- Cimino, G., Gavagnin, M. (2006). *Molluscs From Chemo-ecological Study to Biotechnological Application*. Heidelberg: Springer-Verlag GMBH, Germany
- Clarke D. J. (2008). Phototaxidus: a model for the analysis of pathogenicity and mutualism. *Cellular microbiology*, 10(11), 2159–2167
- Cobb, R. E., Wang, Y., & Zhao, H. (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS synthetic biology*, 4(6), 723–728
- Cousin, C., Derouiche, A., Shi, L., Pagot, Y., Poncet, S., & Mijakovic, I. (2013). Protein-serine/threonine/tyrosine kinases in bacterial signaling and regulation. *FEMS microbiology letters*, 346(1)
- Craney, A., Ozimok, C., Pimentel-Elardo, S. M., Capretta, A., & Nodwell, J. R. (2012). Chemical perturbation of secondary metabolism demonstrates important links to primary metabolism. *Chemistry & biology*, 19(8), 1020–1027
- D'Agostino, P. M., & Gulder, T. (2018). Direct Pathway Cloning Combined with Sequence- and Ligation-Independent Cloning for Fast Biosynthetic Gene Cluster Refactoring and Heterologous Expression. *ACS synthetic biology*, 7(7), 1702–1708
- Dang, T., & Süßmuth, R. D. (2017). Bioactive Peptide Natural Products as Lead Structures for Medicinal Use. *Accounts of chemical research*, 50(7), 1566–1576
- Dangel, V., Westrich, L., Smith, M. C., Heide, L., & Gust, B. (2010). Use of an inducible promoter for antibiotic production in a heterologous host. *Applied microbiology and biotechnology*, 87(1), 261–269
- Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12), 6640–6645
- Donia, M. Hamann, M.T. (2003). Marine natural products and their potential applications as anti-infective agents. *The Lancet*, 3(6), 338-348
- D'Onofrio, A., Crawford, J. M., Stewart, E. J., Witt, K., Gavrish, E., Epstein, S., Clardy, J., & Lewis, K. (2010). Siderophores from neighboring organisms promote the growth of uncultured bacteria. *Chemistry & biology*, 17(3), 254–264
- Erol, O., Schäberle, T. F., Schmitz, A., Rachid, S., Gurgui, C., El Omari, M., Lohr, F., Kehraus, S., Piel, J., Müller, R., & König, G. M. (2010). Biosynthesis of the myxobacterial antibiotic coralopyronin A. *Chembiochem : a European journal of chemical biology*, 11(9), 1253–1265
- Fewer, D. P., & Metsä-Ketelä, M. (2020). A pharmaceutical model for the molecular evolution of microbial natural products. *The FEBS journal*, 287(7), 1429–1449
- Finking, R., & Marahiel, M. A. (2004). Biosynthesis of nonribosomal peptides1. *Annual review of microbiology*, 58, 453–488
- Flett, F., Mersinias, V., & Smith, C. P. (1997). High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS microbiology letters*, 155(2), 223–229
- Flühe, L., & Marahiel, M. A. (2013). Radical S-adenosylmethionine enzyme catalyzed thioether bond formation in sactipeptide biosynthesis. *Current opinion in chemical biology*, 17(4), 605–612
- Forsythe, S.J. (2020). *The Microbiology of Safe Food*, 3rd Edition
- Gerwick, W.H. and Fenner, A.M. (2013). Drug Discovery from Marine Microbes. *Microbial Ecology*, 65(4), 800-806
- Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*, 6(5), 343–345

VI - REFERENCES

- Gomez-Escribano, J. P., & Bibb, M. J. (2011). Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microbial biotechnology*, 4(2), 207–215
- Goodman, A. L., McNulty, N. P., Zhao, Y., Leip, D., Mitra, R. D., Lozupone, C. A., Knight, R., & Gordon, J. I. (2009). Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell host & microbe*, 6(3), 279–289
- Gosliner, T. M., Behrens, D. W., Valdés, A. (2008). *Indo-Pacific Nudibranchs and Sea Slugs: A Field Guide to the World's Most Diverse Fauna*. San Francisco, CA: Sea Challengers, Gig Harbor, WA & California Academy of Sciences.
- Gosliner, T.M., Valdés, A., Behrens, D.W. (2015). *Nudibranch and Sea Slug Identification - Indo-Pacific*. Jacksonville, Florida: New World Publications.
- Grein, F., Müller, A., Scherer, K. M., Liu, X., Ludwig, K. C., Klöckner, A., Strach, M., Sahl, H. G., Kubitscheck, U., & Schneider, T. (2020). Ca²⁺-Daptomycin targets cell wall biosynthesis by forming a tripartite complex with undecaprenyl-coupled intermediates and membrane lipids. *Nature communications*, 11(1), 1455
- Gust, B., Challis, G. L., Fowler, K., Kieser, T., & Chater, K. F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proceedings of the National Academy of Sciences of the United States of America*, 100(4), 1541–1546
- Hart, E. M., Mitchell, A. M., Kononova, A., Grabowicz, M., Sheng, J., Han, X., Rodriguez-Rivera, F. P., Schwaid, A. G., Malinverni, J. C., Balibar, C. J., Bodea, S., Si, Q., Wang, H., Homsher, M. F., Painter, R. E., Ogawa, A. K., Sutterlin, H., Roemer, T., Black, T. A., Rothman, D. M., ... Silhavy, T. J. (2019). A small-molecule inhibitor of BamA impervious to efflux and the outer membrane permeability barrier. *PNAS*, 116(43), 21748–21757.
- Hashimoto, T., Hashimoto, J., Kozono, I., Amagai, K., Kawahara, T., Takahashi, S., Ikeda, H., & Shin-Ya, K. (2018). Biosynthesis of Quinolidomycin, the Largest Known Macrolide of Terrestrial Origin: Identification and Heterologous Expression of a Biosynthetic Gene Cluster over 200 kb. *Organic letters*, 20(24), 7996–7999
- Hertweck C. (2009). The biosynthetic logic of polyketide diversity. *Angewandte Chemie (International ed. in English)*, 48(26), 4688–4716
- Hoffmann, T., Krug, D., Bozkurt, N., Duddela, S., Jansen, R., Garcia, R., Gerth, K., Steinmetz, H., & Müller, R. (2018). Correlating chemical diversity with taxonomic distance for discovery of natural products in myxobacteria. *Nature communications*, 9(1), 803
- Hong, H., Samborsky, M., Zhou, Y., & Leadlay, P. F. (2019). C-Nucleoside Formation in the Biosynthesis of the Antifungal Malayamycin A. *Cell chemical biology*, 26(4), 493–501.e5
- Hopwood, D.A., Kieser, T., Bibb, M., Buttner, M., Chater, K. (2000). *Practical Streptomyces Genetics*. John Innes Foundation, Norwich, UK
- Hudson, G. A., & Mitchell, D. A. (2018). RiPP antibiotics: biosynthesis and engineering potential. *Current opinion in microbiology*, 45, 61–69
- Imai, Y., Meyer, K. J., Iinishi, A., Favre-Godal, Q., Green, R., Manuse, S., Caboni, M., Mori, M., Niles, S., Ghiglieri, M., Honrao, C., Ma, X., Guo, J. J., Makriyannis, A., Linares-Otoya, L., Böhringer, N., Wuisan, Z. G., Kaur, H., Wu, R., Mateus, A., ... Lewis, K. (2019). A new antibiotic selectively kills Gram-negative pathogens. *Nature*, 576(7787), 459–464
- Jaganaman, S., Pinto, A., Tarasev, M., & Ballou, D. P. (2007). High levels of expression of the iron-sulfur proteins phthalate dioxygenase and phthalate dioxygenase reductase in *Escherichia coli*. *Protein expression and purification*, 52(2), 273–279
- Jenke-Kodama, H., & Dittmann, E. (2009). Bioinformatic perspectives on NRPS/PKS megasynthases: advances and challenges. *Natural product reports*, 26(7), 874–883
- Katoh, K., Misawa, K., Kuma, K., Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucl Acids Res*, 30(14), 3059–3066

- Kaya, Y., & Ofengand, J. (2003). A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea, and eukarya. *RNA (New York, N.Y.)*, 9(6), 711–721
- Kling, A., Lukat, P., Almeida, D.V., Bauer, A., Fontaine, E., Sordello, S. et al. (2015). Antibiotics. Targeting DnaN for tuberculosis therapy using novel griselimycins. *Science*. 348(6239), 1106–1112
- Kloosterman, A. M., Shelton, K. E., van Wezel, G. P., Medema, M. H., & Mitchell, D. A. (2020). RRE-Finder: a Genome-Mining Tool for Class-Independent RiPP Discovery. *mSystems*, 5(5), e00267–20
- Letzel, A. C., Pidot, S. J., & Hertweck, C. (2014). Genome mining for ribosomally synthesized and post-translationally modified peptides (RiPPs) in anaerobic bacteria. *BMC genomics*, 15(1), 983
- Lewis K. (2020). The Science of Antibiotic Discovery. *Cell*, 181(1), 29–45
- Liao, R., Duan, L., Lei, C., Pan, H., Ding, Y., Zhang, Q., Chen, D., Shen, B., Yu, Y., & Liu, W. (2009). Thiopeptide biosynthesis featuring ribosomally synthesized precursor peptides and conserved posttranslational modifications. *Chemistry & biology*, 16(2), 141–147
- Lindsay, M. I., Jr, Herrmann, E. C., Jr, Morrow, G. W., Jr, & Brown, A. L., Jr (1970). Hong Kong influenza: clinical, microbiologic, and pathologic features in 127 cases. *JAMA*, 214(10), 1825–1832
- Ling, L. L., Schneider, T., Peoples, A. J., Spoering, A. L., Engels, I., Conlon, B. P., Mueller, A., Schäberle, 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. (2015). A new antibiotic kills pathogens without detectable resistance. *Nature*, 517(7535), 455–459
- Lucas-Elio, P., Hernandez, P., Sanchez-Amat, A., Solano, F. (2005). Purification and partial characterization of marinocine, a new broad-spectrum antibacterial protein produced by *Marinomonas mediterranea*. *Biochim Biophys Acta*, 1721(1-3), 193–203
- Luesch, H., Moore, R.E., Paul, V.J., Mooberry, S.L., Corbett, T.H. (2001). Isolation of Dolastatin 10 from the Marine Cyanobacterium *Symploca* Species VP642 and Total Stereochemistry and Biological Evaluation of Its Analogue Symplostatin 1. *J Nat Prod*, 64(7), 907–910
- Luther, A., Urfer, M., Zahn, M., Müller, M., Wang, S. Y., Mondal, M., Vitale, A., Hartmann, J. B., Sharpe, T., Monte, F. L., Kocherla, H., Cline, E., Pessi, G., Rath, P., Modaresi, S. M., Chiquet, P., Stiegeler, S., Verbree, C., Remus, T., Schmitt, M., ... Obrecht, D. (2019). Chimeric peptidomimetic antibiotics against Gram-negative bacteria. *Nature*, 576(7787), 452–458
- Maffioli, S. I., Zhang, Y., Degen, D., Carzaniga, T., Del Gatto, G., Serina, S., Monciardini, P., Mazzetti, C., Gugliera, P., Candiani, G., Chiriac, A. I., Facchetti, G., Kaltofen, P., Sahl, H. G., Dehò, G., Donadio, S., & Ebright, R. H. (2017). Antibacterial Nucleoside-Analog Inhibitor of Bacterial RNA Polymerase. *Cell*, 169(7), 1240–1248.e23
- Mansson, M., Gram, L., Larsen, T.O. (2011). Production of Bioactive Secondary Metabolites by Marine Vibrionaceae. *Mar Drugs*, 9(9), 1440–1468
- Matsunaga, S., Fusetani, N., Hashimoto, K., Koseki, K., Noma, M. (1986). Bioactive marine metabolites. Part 13. Kabiramide C, a novel antifungal macrolide from nudibranch egg masses. *J Am Chem Soc*, 108(4), 847–849
- McCarthy, P. J., Troke, P. F., & Gull, K. (1985). Mechanism of action of nikkomycin and the peptide transport system of *Candida albicans*. *Journal of general microbiology*, 131(4), 775–780
- Medema, M. H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M. A., Weber, T., Takano, E., & Breitling, R. (2011). antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic acids research*, 39(Web Server issue), W339–W346
- Mehbub, M.F., Lei, J., Franco, C., Zhang, W. (2014). Marine Sponge Derived Natural Products between 2001 and 2010: Trends and Opportunities for Discovery of Bioactives. *Mar Drugs*, 12(8), 4539–4577

VI - REFERENCES

- Mireille Ayé, A., Bonnin-Jusserand, M., Brian-Jaisson, F., Ortalo-Magné, A., Culioli, G., Koffi Nevry, R., Rabah, N., Blache, Y., & Molmeret, M. (2015). Modulation of violacein production and phenotypes associated with biofilm by exogenous quorum sensing N-acylhomoserine lactones in the marine bacterium *Pseudoalteromonas ulvae* TC14. *Microbiology (Reading, England)*, 161(10), 2039–2051
- Mohr, K. I., Volz, C., Jansen, R., Wray, V., Hoffmann, J., Bernecker, S., Wink, J., Gerth, K., Stadler, M., & Müller, R. (2015). Pinensins: the first antifungal lantibiotics. *Angewandte Chemie (International ed. in English)*, 54(38), 11254–11258
- Morens, D. M., Taubenberger, J. K., & Fauci, A. S. (2008). Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: implications for pandemic influenza preparedness. *The Journal of infectious diseases*, 198(7), 962–970
- Morris, D. E., Cleary, D. W., & Clarke, S. C. (2017). Secondary Bacterial Infections Associated with Influenza Pandemics. *Frontiers in microbiology*, 8, 1041
- Myronovskyi, M., & Luzhetskyy, A. (2016). Native and engineered promoters in natural product discovery. *Natural product reports*, 33(8), 1006–1019
- Newman, D.J. and Cragg, G.M. (2012). Natural Products as Sources of New Drugs over the 30 Years from 1981 to 2010. *J Nat Prod*. 75(3), 311-335
- Newman, D. J., & Cragg, G. M. (2016). Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of natural products*, 79(3), 629–661
- Offret, C., Desriac, F., Le Chevalier, P., Mounier, J., Jegou, C., Fleurie, Y. (2016). Spotlight on Antimicrobial Metabolites from the Marine Bacteria *Pseudoalteromonas*: Chemodiversity and Ecological Significance. *Mar Drugs*, 14(7), 129
- Palumbi, S.R. (1996). Nucleic acids II: the polymerase chain reaction. In: *Molecular Systematics*, ed. D.H. Hillis, C. Moritz, B.K. Mable (Sinauer & Associates Inc) 205–247
- Pasquina, L. W., Santa Maria, J. P., & Walker, S. (2013). Teichoic acid biosynthesis as an antibiotic target. *Current opinion in microbiology*, 16(5), 531–537
- Perlova, O., Gerth, K., Kuhlmann, S., Zhang, Y., & Müller, R. (2009). Novel expression hosts for complex secondary metabolite megasynthetases: Production of myxochromide in the thermophilic isolate *Corallococcus macrosporus* GT-2. *Microbial cell factories*, 8, 1
- Pesic, A., Steinhaus, B., Kemper, S., Nachtigall, J., Kutzner, H. J., Höfle, G., & Süssmuth, R. D. (2014). Isolation and structure elucidation of the nucleoside antibiotic strepturidin from *Streptomyces albus* DSM 40763. *The Journal of antibiotics*, 67(6), 471–477
- Peters, C., Collins, G.M. Benkendorff, K. (2011). Characterisation of the physical and chemical properties influencing bacterial epibiont communities on benthic gelatinous egg masses of the pulmonate *Siphonaria diemenensis*. *J Ex Mar Biol Ecol*, 432-433, 138-147
- Pidot, S., Ishida, K., Cyrulies, M., Hertweck, C. (2014). Discovery of Clostrubin, an Exceptional Polyphenolic Polyketide Antibiotic from a Strictly Anaerobic Bacterium. *Angew Chem Int Ed Engl*. 53(30), 7856-7859
- Rajalaxmi, M., Beema Shafreen, R., Iyer, P.M., Sahaya Vino, R., Balamurugan, K., Pandian, S.K. (2016). An in silico, in vitro and in vivo investigation of indole-3-carboxaldehyde identified from the seawater bacterium *Marinomonas* sp. as an anti-biofilm agent against *Vibrio cholerae* O1. *Biofouling*, 32(4), 1-12
- Reher, R., Kim, H. W., Zhang, C., Mao, H. H., Wang, M., Nothias, L. F., Caraballo-Rodriguez, A. M., Glukhov, E., Teke, B., Leao, T., Alexander, K. L., Duggan, B. M., Van Everbroeck, E. L., Dorrestein, P. C., Cottrell, G. W., & Gerwick, W. H. (2020). A Convolutional Neural Network-Based Approach for the Rapid Annotation of Molecularly Diverse Natural Products. *Journal of the American Chemical Society*, 142(9), 4114–4120

- Richter, S. G., Elli, D., Kim, H. K., Hendrickx, A. P., Sorg, J. A., Schneewind, O., & Missiakas, D. (2013). Small molecule inhibitor of lipoteichoic acid synthesis is an antibiotic for Gram-positive bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 110(9), 3531–3536
- Rosenqvist, P., Palmu, K., Prajapati, R. K., Yamada, K., Niemi, J., Belogurov, G. A., Metsä-Ketelä, M., & Virta, P. (2019). Characterization of C-nucleoside Antimicrobials from *Streptomyces albus* DSM 40763: Strepturidin is Pseudouridimycin. *Scientific reports*, 9(1), 8935
- Schäberle, T.F. and Hack, I.M. (2014). Overcoming the current deadlock in antibiotic research. *Trends Microbiol.* 22, 165-167.
- Schmitt, S., Montalbán-López, M., Peterhoff, D., Deng, J., Wagner, R., Held, M., Kuipers, O. P., & Panke, S. (2019). Analysis of modular bioengineered antimicrobial lanthipeptides at nanoliter scale. *Nature chemical biology*, 15(5), 437–443
- Schramma, K. R., Bushin, L. B., & Seyedsayamdost, M. R. (2015). Structure and biosynthesis of a macrocyclic peptide containing an unprecedented lysine-to-tryptophan crosslink. *Nature chemistry*, 7(5), 431–437
- Severi, E., & Thomas, G. H. (2019). Antibiotic export: transporters involved in the final step of natural product production. *Microbiology (Reading, England)*, 165(8), 805–818
- Shen B. (2003). Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current opinion in chemical biology*, 7(2), 285–295
- Siegl, T., Tokovenko, B., Myronovskyi, M., & Luzhetskyy, A. (2013). Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes. *Metabolic engineering*, 19, 98–106
- Sirirak, T., Brecker, L., Plubrukarn, A. (2013). Kabiramide L, a new antiplasmodial trisoxazole macrolide from the sponge *Pachastrissa nux*. *Nat Prod Res*, 27(13), 1213-1219
- Siodlak D. (2015). α,β -Dehydroamino acids in naturally occurring peptides. *Amino acids*, 47(1), 1–17
- Sosio, M., Gaspari, E., Iorio, M., Pessina, S., Medema, M. H., Bernasconi, A., Simone, M., Maffioli, S. I., Ebright, R. H., & Donadio, S. (2018). Analysis of the Pseudouridimycin Biosynthetic Pathway Provides Insights into the Formation of C-nucleoside Antibiotics. *Cell chemical biology*, 25(5), 540–549.e4
- Srinivas, N., Jetter, P., Ueberbacher, B. J., Werneburg, M., Zerbe, K., Steinmann, J., Van der Meijden, B., Bernardini, F., Lederer, A., Dias, R. L., Misson, P. E., Henze, H., Zumbunn, J., Gombert, F. O., Obrecht, D., Hunziker, P., Schauer, S., Ziegler, U., Käch, A., Eberl, L., ... Robinson, J. A. (2010). Peptidomimetic antibiotics target outer-membrane biogenesis in *Pseudomonas aeruginosa*. *Science (New York, N.Y.)*, 327(5968), 1010–1013
- Stamatakis, A. (2014). RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics*, doi: 10.1093/bioinformatics/btu033
- Stein T. (2005). *Bacillus subtilis* antibiotics: structures, syntheses and specific functions. *Molecular microbiology*, 56(4), 845–857
- Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., MacNair, C. R., French, S., Carfrae, L. A., Bloom-Ackermann, Z., Tran, V. M., Chiappino-Pepe, A., Badran, A. H., Andrews, I. W., Chory, E. J., Church, G. M., Brown, E. D., Jaakkola, T. S., Barzilay, R., & Collins, J. J. (2020). A Deep Learning Approach to Antibiotic Discovery. *Cell*, 180(4), 688–702.e13
- Storek, K. M., Auerbach, M. R., Shi, H., Garcia, N. K., Sun, D., Nickerson, N. N., Vij, R., Lin, Z., Chiang, N., Schneider, K., Weckslar, A. T., Skippington, E., Nakamura, G., Seshasayee, D., Koerber, J. T., Payandeh, J., Smith, P. A., & Rutherford, S. T. (2018). Monoclonal antibody targeting the β -barrel assembly machine of *Escherichia coli* is bactericidal. *PNAS*, 115(14), 3692–3697
- Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., Pulcini, C., Kahlmeter, G., Kluytmans, J., Carmeli, Y., Ouellette, M., Outtersen, K., Patel, J., Cavaleri, M., Cox, E. M., Houchens, C. R., Grayson, M. L., Hansen, P.,

- Singh, N., Theuretzbacher, U., ... WHO Pathogens Priority List Working Group (2018). Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *The Lancet. Infectious diseases*, 18(3), 318–327
- Tailliez, P., Laroui, C., Ginibre, N., Paule, A., Pagès, S., & Boemare, N. (2010). Phylogeny of *Photorhabdus* and *Xenorhabdus* based on universally conserved protein-coding sequences and implications for the taxonomy of these two genera. Proposal of new taxa: *X. vietnamensis* sp. nov., *P. luminescens* subsp. *caribbeanensis* subsp. nov., *P. luminescens* subsp. *hainanensis* subsp. nov., *P. temperata* subsp. *khanii* subsp. nov., *P. temperata* subsp. *tasmaniensis* subsp. nov., and the reclassification of *P. luminescens* subsp. *thracensis* as *P. temperata* subsp. *thracensis* comb. nov. *International journal of systematic and evolutionary microbiology*, 60(Pt 8), 1921–1937
- Tang, X., Li, J., Millán-Aguiñaga, N., Zhang, J. J., O'Neill, E. C., Ugalde, J. A., Jensen, P. R., Mantovani, S. M., & Moore, B. S. (2015). Identification of Thiotetronic Acid Antibiotic Biosynthetic Pathways by Target-directed Genome Mining. *ACS chemical biology*, 10(12), 2841–2849
- Tipper, D. J., & Strominger, J. L. (1965). Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proceedings of the National Academy of Sciences of the United States of America*, 54(4), 1133–1141
- Velkov, T., Deris, Z. Z., Huang, J. X., Azad, M. A., Butler, M., Sivanesan, S., Kaminskas, L. M., Dong, Y. D., Boyd, B., Baker, M. A., Cooper, M. A., Nation, R. L., & Li, J. (2014). Surface changes and polymyxin interactions with a resistant strain of *Klebsiella pneumoniae*. *Innate immunity*, 20(4), 350–363
- Vynne, N. G., Mansson, M. and Gram, L. (2013). Comparative Genomic and Metabolomic Analysis of Twelve Strains of *Pseudoalteromonas luteoviolacea*. Unpublished (NCBI Genebank submission NZ_AUXZ01000151.1)
- Watanakunakorn C. (1984). Mode of action and in-vitro activity of vancomycin. *The Journal of antimicrobial chemotherapy*, 14 Suppl D, 7–18
- Waters, A. L., Peraud, O., Kasanah, N., Sims, J. W., Kothalawala, N., Anderson, M. A. et al. (2014). An analysis of the sponge *Acanthostrongylophora igens*' microbiome yields an actinomycete that produces the natural product manzamine A. *Front Mar Sci*, <https://doi.org/10.3389/fmars.2014.00054>
- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Brucoleri, et al. (2015). antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res*, 43(W1), W237–W243
- Wegman, M. A., Janssen M. H. A., van Rantwijk, F., Sheldon, R. A. (2001). Towards Biocatalytic Synthesis of β -Lactam Antibiotics. *Adv. Synth. Catal*, 343
- Weisburg, W. G., Barns, S. M., Pelletier, D. A., Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol*, 173(2), 697–703
- Weist, S., & Süssmuth, R. D. (2005). Mutational biosynthesis—a tool for the generation of structural diversity in the biosynthesis of antibiotics. *Applied microbiology and biotechnology*, 68(2), 141–150
- Wilson, M. C., Mori, T., Rückert, C., Uria, A. R., Helf, M. J., Takada, K. et al. (2014). An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature*, 506, 58–62
- WHO Library Cataloguing-in-Publication Data. (2014). Antimicrobial resistance: global report on surveillance. Geneva: WHO Press, World Health Organisation
- Xu, F., Wu, Y., Zhang, C., Davis, K. M., Moon, K., Bushin, L. B., & Seyedsayamdost, M. R. (2019). A genetics-free method for high-throughput discovery of cryptic microbial metabolites. *Nature chemical biology*, 15(2), 161–168
- Yang, X., Lennard, K. R., He, C., Walker, M. C., Ball, A. T., Doigneaux, C., Tavassoli, A., & van der Donk, W. A. (2018). A lanthipeptide library used to identify a protein-protein interaction inhibitor. *Nature chemical biology*, 14(4), 375–380

VI - REFERENCES

- Yocum, R. R., Rasmussen, J. R., & Strominger, J. L. (1980). The mechanism of action of penicillin. Penicillin acylates the active site of *Bacillus stearothermophilus* D-alanine carboxypeptidase. *The Journal of biological chemistry*, 255(9), 3977–3986
- Yu, Y., Zhang, Q., & van der Donk, W. A. (2013). Insights into the evolution of lanthipeptide biosynthesis. *Protein science : a publication of the Protein Society*, 22(11), 1478–1489
- Zhou, Q., Bräuer, A., Adihou, H., Schmalhofer, M., Saura, P., Grammbitter, G., Kaila, V., Groll, M., & Bode, H. B. (2019). Molecular mechanism of polyketide shortening in anthraquinone biosynthesis of *Photorhabdus luminescens*. *Chemical science*, 10(25), 6341–6349

VII. SUPPLEMENTARY MATERIAL

Suppl.Table 1: List of primers used in Chapter II

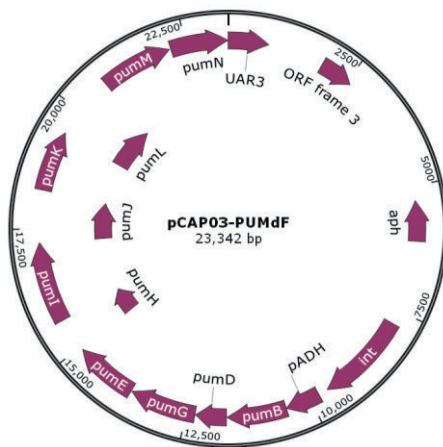
Primer name	Sequence 5' -> 3'	specificity	Literature
JS002=A7Rallg	SASRTCNCNGTNCGRRTASA	A-domain A7	this study, based on A7R of Ayuso-Sacido, A. and Genilloud, O. (2005)
A3F	GCSTACSYSATSTACACSTCSGG	A-domain A3, GC rich	Ayuso-Sacido, A. and Genilloud, O. (2005)
A7R	SASGTCVCCSGTSCGGTAS	A-domain A7, GC rich	Ayuso-Sacido, A. and Genilloud, O. (2005)
JS007	AARDSNGGNGSNGSNTAYBNCC	A-domain, A2	based on degNRPS-1F.i Schirmer et al. 2005
JS008	CKRWRNCCNCKNANYTTNACYTG	A-domain, A8	based on degNRPS-4R.i Schirmer et al. 2005
NRPSA-F	GGWCDACHGGHMANCCHAARGG	A-domain A3	Wu et al. 2011
NRPSA-R	GGCAKCCATYTYGCCARGTCNCCKGT	A-domain A7	Wu et al. 2011
AnSerin	CAYTTYGTNCCNWSNATGYT	A-domain, serine	this study
KF0001=KSDPQQFi	MGi GAR GCi HWi SMi ATG GAY Cci CAR CAi MG	general KS-specific primer	based on Fisch et al. 2009
KF0002=KSHGTGRi	GGR TCi Cci ARi SWi GTi Cci GTi CCR TG	general KS-specific primer	based on Fisch et al. 2009
KF0003=KSDPQQF	MGN GAR GCN NWN SMN ATG GAY CCN CAR CAN MG	general KS-specific primer	Fisch et al. 2009
KF0004=KSHGTGR	GGR TCN CCN ARN SWN GTN CCN GTN CCR TG	general KS-specific primer	Fisch et al. 2009

Suppl.Table 2: List of bacterial strains and their respective source organism retrieved; Chapter II

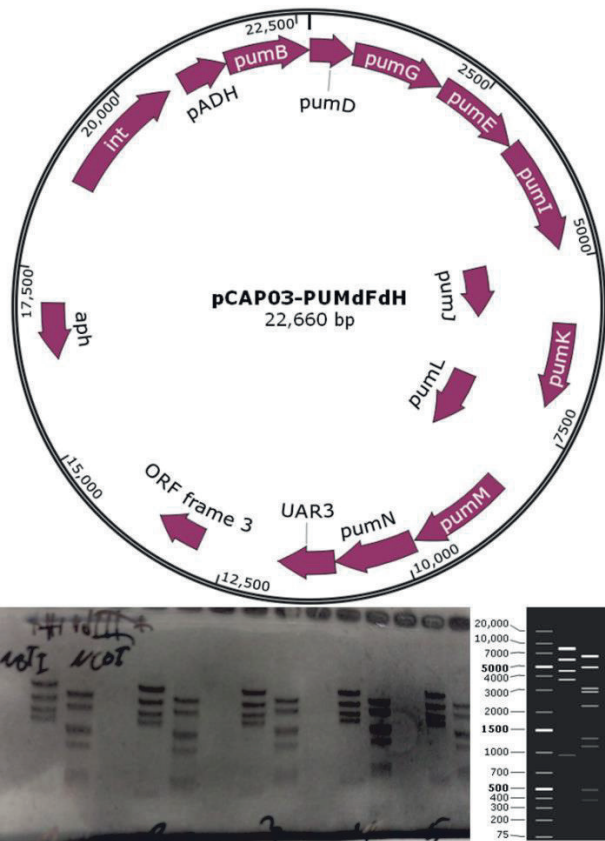
Source	Closest relative strain	Int. Strain nr.
<i>Chromodoris annae</i> (Chan15Bu-2)	<i>Vibrio coralliilyticus</i> strain OCN014	Bu15_01
	<i>Vibrio harveyi</i> strain ATCC 33843	Bu15_02
	<i>Vibrio harveyi</i> strain ATCC 33843	Bu15_03
	<i>Vibrio coralliilyticus</i> strain OCN014	Bu15_04
<i>Chromodoris diana</i> (Chdi15Bu-3 + Chdi15Bu-38 + Chdi15Bu-55)	<i>Bacillus subtilis</i> strain EDR4	Bu15_05
	<i>Serratia marcescens</i> strain SW2-9-3	Bu15_06
	<i>Pseudovibrio</i> sp. FO-BEG1 strain FO-BEG1	Bu15_07
	<i>Pseudoalteromonas</i> sp. CF6-1 (+uncultured organism clone)	Bu15_08
	<i>Pseudoalteromonas</i> sp. AS-43	Bu15_09
	<i>Vibrio harveyi</i> strain ATCC 3384	Bu15_10

	Pseudovibrio sp. FO-BEG1 strain FO-BEG1	Bu15_11
<i>Chromodoris sp. 30</i> (Chsp3015Bu-4)	Vibrio harveyi strain ATCC 33843	Bu15_12
	Pseudovibrio sp. FO-BEG1	Bu15_13
	Vibrio harveyi strain ATCC 33843	Bu15_14
<i>Chromodoris williani</i> (Chwi15Bu-2)	Vibrio tubiashii ATCC 19109	Bu15_15
<i>Chromodoris annae</i> eggmass (Chan15Bu-11E)	Bacillus aryabhattai isolate PSB57	Bu15_16
	Bacillus thuringiensis serovar indiana strain HD521	Bu15_17
	Pseudoalteromonas sp. BJ9	Bu15_18
	Pseudoalteromonas sp. BJ9	Bu15_19
	Marinomonas communis strain NBRC 102224 (+uncultured organism clone)	Bu15_20
	Pseudovibrio sp. FO-BEG1	Bu15_21
<i>Doriprismatic stellata</i> (Glst15Bu-1)	Providencia vermicola strain FFA6	Bu15_38
<i>Doriprismatic stellata</i> eggmass (Glst15Bu-1E)	Vibrio alginolyticus NBRC 15630	Bu15_32
	Vibrio alginolyticus NBRC 15630	Bu15_33
	Vibrio alginolyticus NBRC 15630	Bu15_34
	Pseudoalteromonas sp. AS-43	Bu15_35
	Pseudoalteromonas sp. BJ9 (+uncultured organism clone)	Bu15_36
<i>Doriprismatic stellata</i> sponge (Glst15Bu-1P)	Vibrio alginolyticus strain RE98	Bu15_37
<i>Hexabranhus sanguineus</i> eggmass (Hesa15Bu-1)	Longispora albida strain K97-0003	Bu15_22
	Marinobacter sp. L21-PYE-C23	Bu15_23
	Pseudoalteromonas sp. AS-43	Bu15_24
	Pseudoalteromonas sp. AS-43	Bu15_25
	Pseudoalteromonas sp. AS-43	Bu15_26
	Vibrio alginolyticus NBRC 15630	Bu15_27
	Vibrio sp. Ex25	Bu15_28
	Bacillus thuringiensis serovar indiana strain HD521	Bu15_29
	Vibrio alginolyticus NBRC 15630	Bu15_30
	Pseudoalteromonas rubra	Bu15_31
<i>Hexabranhus sanguineus</i> eggmass (Hesa15Bu-2)	Pseudomonas sp. PETBA	Bu15_40
	Pseudomonas sp. MBEA06	Bu15_47
	Microbacterium testaceum	Bu15_48
	Pelagibaca bermudensis	Bu15_49
	Microbacterium sp. SKJH-23	Bu15_41
	Microbacterium sp. LHR-08	Bu15_42
	Pseudomonas pseudoalcaligenes	Bu15_43
	Gordonia sp. VCM12	Bu15_44
	Gordonia terrae	Bu15_45
	Pseudomonas pachastrellae	Bu15_46
<i>Phyllidiella cf pustulosa</i> (Phpu15Bu-1)	Vibrio sp. BWDY-57 (+uncultured organism clone)	Bu15_39

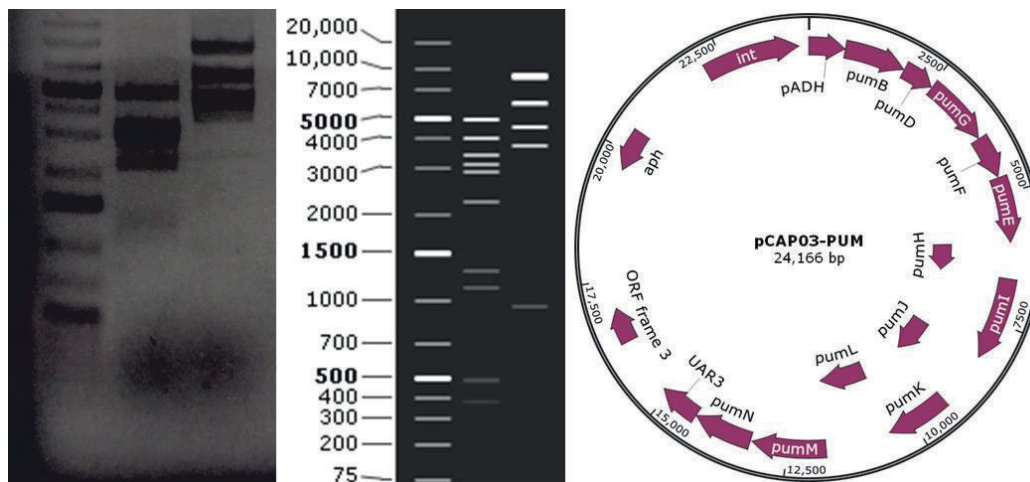
Suppl. Fig 1: Vector Map of pCAP03-PUMΔF



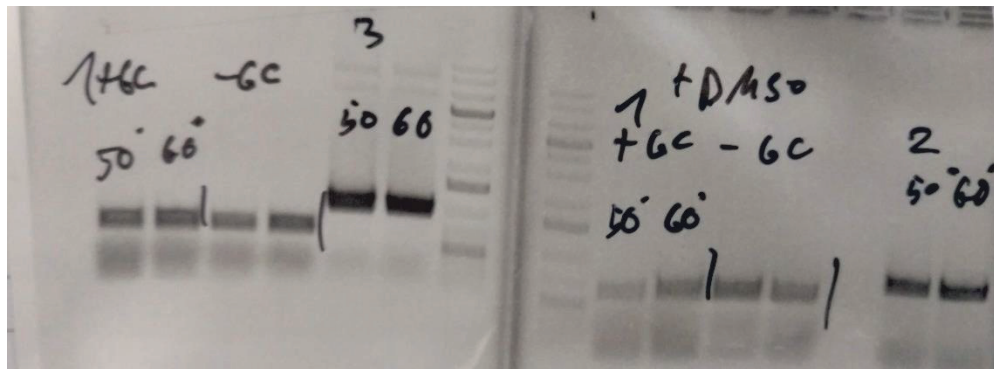
Suppl. Fig 2: Vector Map and restriction (*HindIII*/*NcoI* and *HindIII*/*NotI*) analysis of pCAP03-PUMΔFΔH



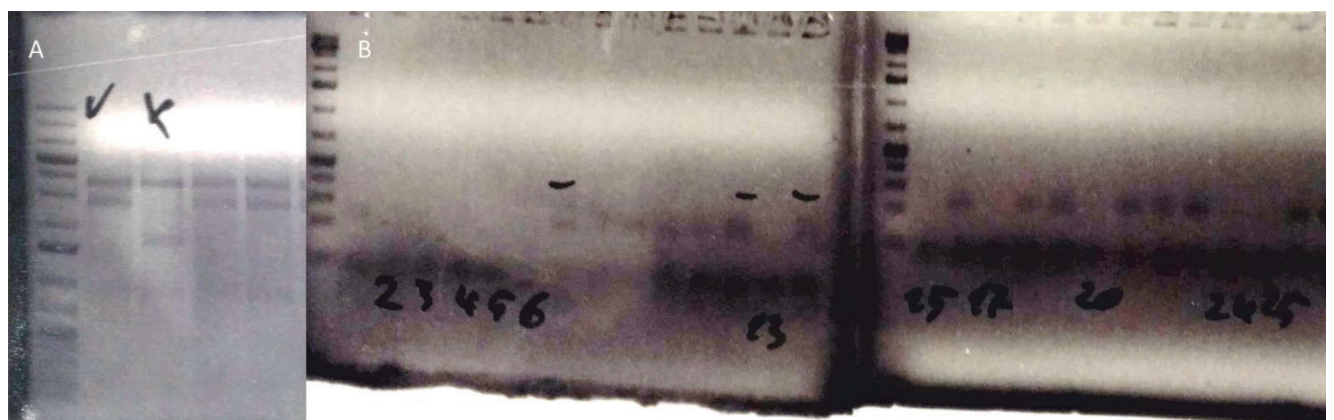
Suppl. Fig 3: Vector Map and restriction analysis (*HindIII/NcoI* and *HindIII/NotI*) of pCAP03-PUM Δ F Δ H



Suppl. Fig 4: PCRs for generation of pNB01



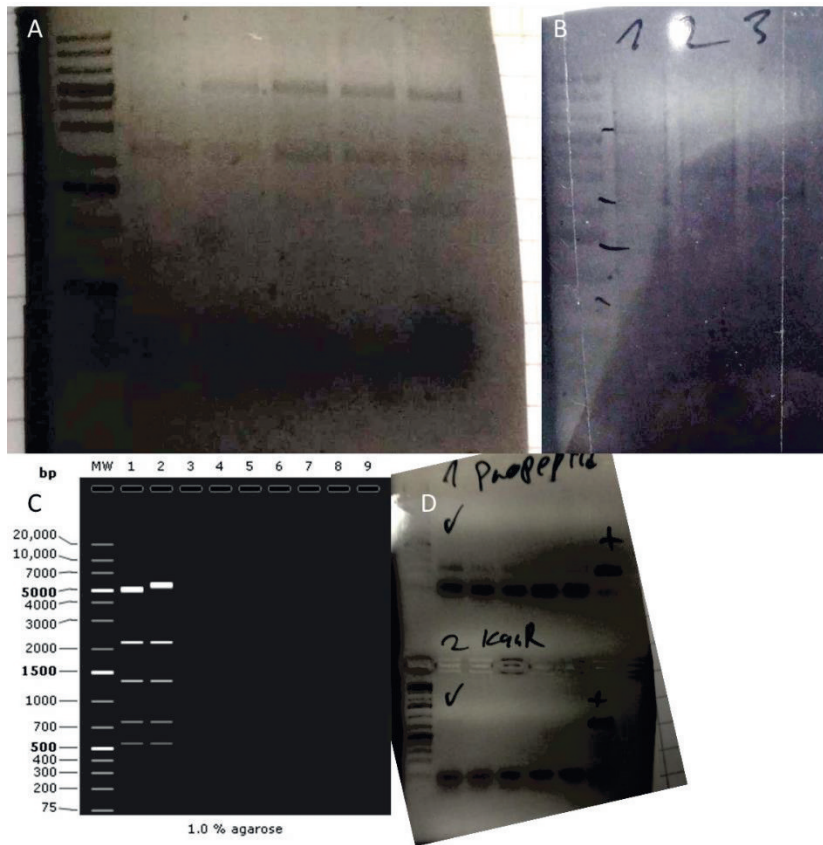
Suppl. Fig 5: A) restriction analysis of pNB02 (*HindIII/NcoI*). B) PCR for *darA*: presence of *darA* shows double crossover mutants of *P. khani* DSM3369 with pNB02, reverting the WT, absence of *darA* indicates successful KO. 11 colonies were tested negative for *darA*



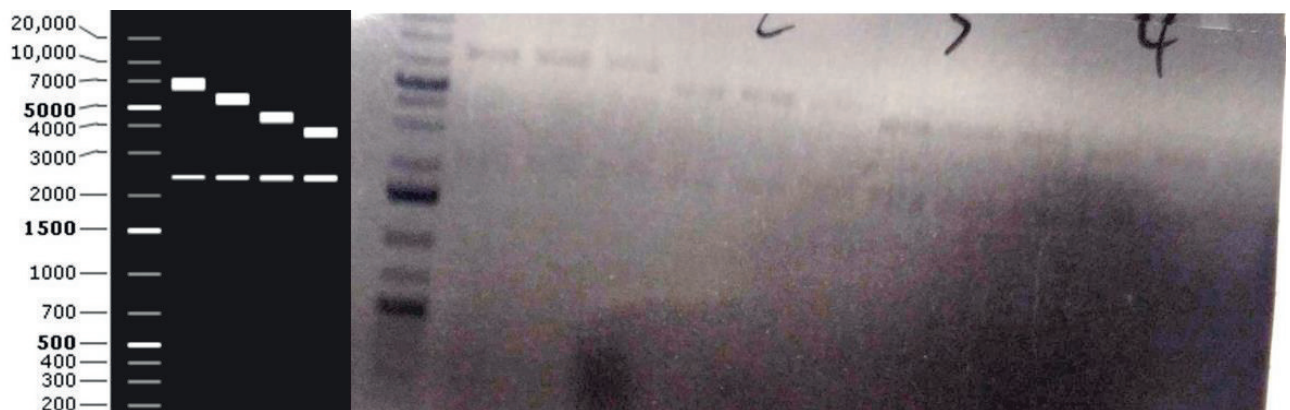
Suppl. Fig 6: A) restriction analysis of pNB03-darA-E (*Sall/SacI*). B) restriction analysis of pNB03-darA-E (*Sall/SacI*). C) in silico prediction of restriction analysis of pNB03-darA-E (left) and pNB03-darA-E (right) (*Sall/SacI*). D) Conjugation of pNB03-darA-E to

VII – SUPPLEMENTARY MATERIAL

P. khani DSM3369 Δ darABCDE; successful amplification of the darA (top) and the kan resistance cassette (bottom) indicates successful introduction of the vector (exemplified for pNBDaro).

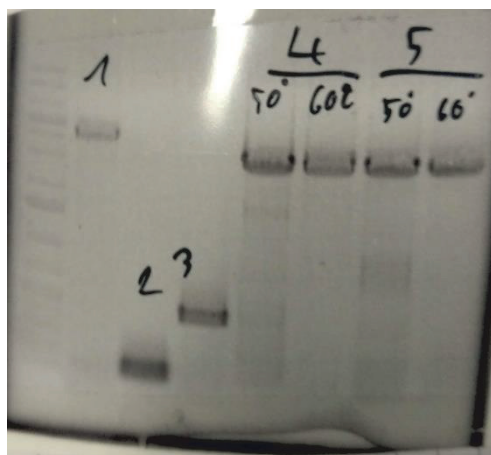


Suppl. Fig 7: Test restriction (*PvuI/NcoI*) for (from left to right) pNBDaro- Δ darB1, pNBDaro- Δ darB, pNBDaro- Δ darBC, pNBDaro- Δ darBCD

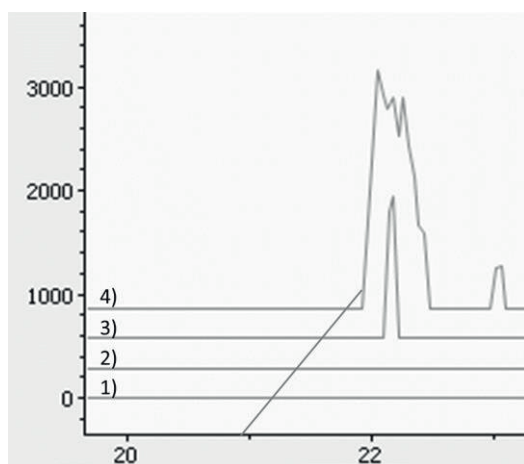


VII – SUPPLEMENTARY MATERIAL

Suppl. Fig 8: PCRs used for assembly of pNBDaromod: 1) pNB03, 2) darA fragment, 3) lacZ-spacer, 4) darB-E_part1, 5) darB-E_part2



Suppl. Fig 9: EICs of 1) $C_{51}H_{64}N_{14}O_{11} [M+2H]^{2+} \pm 0.01$ (Darobactin B), 2) $C_{46}H_{54}N_{12}O_{12} [M+2H]^{2+} \pm 0.01$ (Darobactin C), 3) $C_{47}H_{55}N_{13}O_{12} [M+2H]^{2+} \pm 0.01$ (Darobactin D) and 4) $C_{46}H_{54}N_{10}O_{12} [M+2H]^{2+} \pm 0.01$ (Darobactin E) measured from pNBDaromod derived constructs including the follower AAs QEI. Only traces of Darobactin E could be detected. The peak in 3) does not correspond to traces of Darobactin D but another closely mass



VIII. Publications

2020

Hertzer C, Kehraus S, Böhringer N, Kaligis F, Bara R, Erpenbeck D, Wörheide G, Schäberle TF, Wägele H, König GM, (2020). Antibacterial scalarane from *Doriprismatica stellata* nudibranchs (Gastropoda, Nudibranchia), egg ribbons, and their dietary sponge *Spongia* cf. *agaricina* (Demospongiae, Dictyoceratida). *Beilstein journal of organic chemistry*.

2019

Imai Y, Meyer KJ, Iinishi A, Favre-Godal Q, Green R, Manuse S, Caboni M, Mori M, Niles S, Ghiglieri M, Honrao C, Ma X, Guo JJ, Makriyannis A, Linares-Otaya L, Böhringer N, Wuisan ZG, Kaur H, Wu R, Mateus A, Typas A, Savitski MM, Espinoza JL, O'Rourke A, Nelson KE, Hiller S, Noinaj N, Schäberle TF, D'Onofrio A, Lewis K. (2019). A new antibiotic selectively kills Gram-negative pathogens. *Nature*.

2018

Kaligis F, Eisenbarth JH, Schillo D, Dialao J, Schäberle TF, Böhringer N, Bara R, Reumschüssel S, König GM, Wägele H. (2018). Second survey of heterobranch sea slugs (Mollusca, Gastropoda, Heterobranchia) from Bunaken National Park, North Sulawesi, Indonesia - how much do we know after 12 years? *Marine Biodiversity Records*.

2017

Fisch KM, Hertzer C, Böhringer N, Wuisan ZG, Schillo D, Bara R, Kaligis F, Wägele H, König GM, Schäberle TF. (2017). Potential of Indonesian Heterobranchs Found around Bunaken Island for the Production of Bioactive Compounds. *Mar Drugs*.

Böhringer N, Fisch KM, Schillo D, Bara R, Hertzer C, Grein F, Eisenbarth JH, Kaligis F, Schneider T, Wägele H, König GM, Schäberle TF. (2017). Antimicrobial Potential of Bacteria Associated with Marine Sea Slugs from North Sulawesi, Indonesia. *Front Microbiol*.

Böhringer N, Gütschow M, König GM, Schäberle TF. (2017). Phileucin - A Cyclic Dipeptide Similar to Phevalin (Aureusimine B) from *Streptomyces coelicolor* M1146. *Nat Prod Commun*.

2016

Böhringer N, Moghaddam JA, Burdziak A, Kunte HJ, Galinski EA, Schäberle TF. (2016). Different strategies of osmoadaptation in the closely related marine myxobacteria *Enhygromyxa salina* SWB007 and *Plesiocystis pacifica* SIR-1. *Microbiology*.