Studies on the Biosynthesis of Guanidine-Containing Cyclic Lipopeptides from *Lysobacter* spp.

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Patricia Arlt

aus Schwäbisch Gmünd

Tübingen

2022

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	18.07.2022
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter/-in:	Prof. Dr. Harald Gross
2. Berichterstatter/-in:	Prof. Dr. Leonard Kaysser

Declaration

I hereby declare that this thesis entitled “Studies on the Biosynthesis of guanidine-containing cyclic lipopeptides from *Lysobacter* species” is an original report of my research, has been written by myself and has not been submitted for any previous degree. The experimental work is almost entirely my own work; the collaborative contributions have been indicated clearly and acknowledged. Due references have been provided on all supporting literature and resources.

Parts of this work will be published in:

Arlt, P., Hashizume, H., Igarashi, M., Gross, H., Genome Sequence of *Lysobacter* sp. Strain BMK333-48F3, the Producer Strain of Potent Lipopeptide Antibiotics of the Tripropeptin Family. Microbiol Resour Announc 2021, 10 (49), e0096921.

Arlt, P., Müll, M., Miess, H., Saad, H., Hashizume, H., Igarashi, M., Kries, H., Gust, B., Gross, H. Identification and analysis of the plusbacin gene cluster yields insights into the biosynthesis of one class of guanidine-containing cyclic lipopeptides [unpublished manuscript]

_____

Tübingen,

Acknowledgments

At this point, I would like to thank everybody who supported me during the realization of my doctorate.

In particular, I would like to thank Prof. Dr. Harald Groß for accepting me as a PhD student, providing me with an interesting topic, excellent working conditions, good supervision and scientific support.

I also owe gratitude to Prof. Dr. Leonard Kaysser for acting as a second referee, as well as his competent professional assessment.

My special appreciation goes to PD Dr. Bertolt Gust for his unfailing scientific support during my entire doctorate and for the many pieces of advice in the practical day-to-day laboratory work.

Moreover, I would like to thank him and PD Dr. Evi Stegmann for their willingness to act as examiners during the oral disputation.

I naturally thank Maximillian Müll and Dr. Hajo Kries for the productive cooperation and their help in the determination of A domain specificity.

Thanks go also to Dr. Dorothee Wistuba for the (HR-) LCMS measurements and Hamada Saad for his reliable and competent help with the analysis.

I would also like to express my appreciation to my laboratory colleagues of the S2 lab for the excellent working atmosphere and the helpful discussions and advice.

Moreover, a big thanks goes to Irina Helmle and Alicia Engelbrecht. It is impossible to put into words how much I have appreciated you and your friendship during the last years.

Furthermore, I’m very grateful to all my fellow PhD students and members of the Pharmaceutical Biology department for their friendship and ongoing support in and outside of the lab.

Last but not least, I want to pay a warm tribute to my family and friends for their care, patience, encouragement and love not only during the time of my doctorate but also throughout my life.

„Success consists of going from failure to failure without loss of enthusiasm. “

- *Winston Churchill*

Table of Contents

Zusammenfassung

Lipopeptid-Antibiotika bilden eine pharmazeutisch interessante Substanzklasse, die eine Reihe von klinisch relevanten Antibiotika umfasst. Guanidin-haltige zyklischen Lipopeptide, die durch das natürlich vorkommende Empedopeptin (EMP), die Tripropeptine (TPP) und die Plusbacine (PLUS) repräsentiert werden, zeigen eine starke antibakterielle Wirkung gegen eine Vielzahl von gram-positiven Krankheitserregern. Trotz der recht großen Unterschiede in der Wirkstärke der wichtigsten Vertreter dieser Gruppe ist die strukturelle Ähnlichkeit innerhalb der Guanidin-haltigen zyklischen Lipopeptide auffällig. EMP, TPP sowie PLUS weisen ein Octapeptid-Makrolacton als Teil ihrer Struktur auf, welches mit einem Lipidschwanz acyliert ist. Während die Aminosäurenzusammensetzung der südlichen Hemisphäre des Makrozyklus identisch ist, treten in der nördlichen Hemisphäre Modifikationen auf. Ebenso kann der Fettsäureschwanz entweder verzweigt oder unverzweigt auftreten, sowie in seiner Länge variieren.

Um den biosynthetischen Locus für die Produktion von Tripropeptin zu lokalisieren, wurde das gesamte Genom des Tripropeptin-Produzenten *Lysobacter* BMK333-48F3 sequenziert. Auf Grundlage dieser Daten wurde ein mutmaßliches biosynthetisches Gencluster (BGC) identifiziert, das aus nichtribosomalen Peptidsynthetase (NRPS) Strukturgenen und benachbarten akzessorischen Genen besteht, die für Dioxygenasen kodieren. Zusätzlich wurde das BGC von Plusbacin in *Lysobacter firmicutimachus* PB-6250^T mittels Knockouts verifiziert.

Um Einblicke in die Biosynthese der Guanidin-haltigen zyklischen Lipopeptide zu erhalten, wurden Knockout- sowie Überexpressionsstudien durchgeführt. Hierbei waren die Hydroxylierungsprozesse einzelner Aminosäuren und die Aufgabe der im Gencluster kodierten Dioxgenasen von besonderem Interesse.

In-frame Deletion von *plbD* und *plbE* sowie heterologe Überexpression von A Domänen, identifizierten je eine der beiden Hydroxyasparaginsäuren, die sich in der südlichen Hemisphäre des Makrozyklus befinden, als Target der Dioxygenasen. Die β-Hydroxylierung erfolgt nach der Aktivierung von Asparaginsäure durch die zugehörige A-domäne und daher, nachdem die Aminosäure kovalent an die T Domäne geladen wurde.

Summary

Lipopeptide antibiotics form a pharmaceutically interesting class of compounds that includes a number of clinically relevant antibiotics. Guanidine-containing cyclic lipopeptides represented by the naturally occurring empedopeptin (EMP), tripropeptins (TPP), and plusbacins (PLUS) exhibit potent antibacterial activity against a variety of Gram-positive pathogens. Despite the rather large differences in the potency of the major members of this group, the structural similarity within the guanidine-containing cyclic lipopeptides is striking. EMP, TPP as well as PLUS consist of octapeptide macrolactone core, which is acylated with a lipid tail. While the amino acid composition of the southern hemisphere macrocycle is identical, modifications occur in the northern hemisphere as well as in the fatty acid side chain, which can occur either branched or unbranched, as well as vary in length.

To localize the biosynthetic locus of tripropeptin, the genome of the tripropeptin producer *Lysobacter* BMK333-48F3 was sequenced. Based on these data, a putative biosynthetic gene cluster (BGC) consisting of nonribosomal peptide synthetase (NRPS) structural genes and adjacent accessory genes encoding dioxygenases was, identified. In addition, the BGC of plusbacin in *Lysobacter firmicutimachus* PB-6250^T was verified by knockouts.

To gain further insight into the biosynthesis of guanidine-containing cyclic lipopeptides, knockout as well as overexpression studies were performed. Here, the hydroxylation processes of individual amino acids and the role of the dioxgenases encoded in the gene cluster were of particular interest.

In-frame deletion of *plbD* and *plbE* as well as heterologous overexpression of A domains, identified one of the two hydroxyaspartic acids each, located in the southern hemisphere of the octapeptide macrolacton, as targets of the dioxygenases. β-hydroxylation occurs after activation of aspartic acid by the associated A domain and therefore after the amino acid is covalently loaded to the T domain.

Abbreviations

%	percent
×g	ground acceleration
Δ	standard error [ppm]
°C	degree Celsius
A (domain)	adenylation domain
aac(3)IV	apramycin resistance gene
ACN	acetonitrile
Ala	Alanine
antiSMASH	antibiotic and Secondary Metabolite Analysis Shell
Apra	apramycin
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
ATP	adenosine triphosphate
BGC	biosynthetic gene cluster
BLAST	Basic Local Alignment and Search Tool
bp	base pair
C (domain)	condensation domain
cm	centimeter
CoA	coenzyme A
Cre	cyclic recombination
Cy (domain)	cyclization domain
Cys	Cysteine
Da	Dalton
DAD	Diode Array Detector
ddH_2O	double distilled water
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside 5´-triphosphate
dsDNA	double stranded DNA
E (domain)	epimerization domain

E. coli	*Escherichia coli*
EDTA	ethylendiamine tetra-acetic acid
EMP	Empedopeptin
ESI	electrospray ionization
Fe(II)	ferrous iron
Fe(III)	ferric iron
fwd	forward
g	gram
gBlock	ds DNA fragment
gDNA	genomic DNA
Gen	Gentamicin
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
h	hour
HAMA	multiplexed hydroxamate assay
HCl	hydrochloric acid
HF	high fidelity
His	Histidine
His_6	hexahistidine
HPLC	High Performance Liquid Chromatography
HR	high resolution
Hya	Hydroxyaspartic acid
Hyp	Hydroxyproline
IC	inhibitory concentration
Ile	Isoleucine
IPTG	isopropyl s-D-1-thiogalactopyranoside
kb	kilo bases
L	liter
LC	Liquid Chromatography
LC-MS	Liquid Chromatography-Mass Spectrometry
Leu	Leucine

loxP	locus of X-over of P1 (Cre-recombinase recognition site)
Lys	Lysine
M	molar
m/z	mass-to-charge ratio
Mbp	mega/million base pairs
MCS	multiple cloning site
MeOH	methanol
MesG	7-methylthioguanosine
Met	Methionine
mg	milligram
MHz	Megahertz
MIC	minimum inhibitory concentration
min	minutes
mL	milliliter
MLP	MbtH-like protein
mM	millimolar
MRSA	methicillin-resistant *Staphylococcus aureus*
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MW	molecular weight
NaOH	sodium hydroxide
NEB	New England Biolab
NGS	next-generation sequencing
NRPS	Non-Ribosomal Peptide Synthetase
OD	optical density
OH	Hydroxyl-
ORF	Open Reading Frame
oriT	origin of transfer
PacBio	Pacific Bioscience
PAGE	polyacrylamide gel electrophoresis
PCP	Peptidyl Carrier Protein
PCR	Polymerase Chain Reaction

pH	potential of hydrogen
Phe	Phenylalanine
PKS	Polyketide Synthase
PLUS	Plusbacin
ppm	parts per million
Pro	Proline
PRSP	penicillin-resistant *Streptococcus pneumoniae*
rev	reverse
RP	reversed phase
rpm	rounds per minute
RT	room temperature
SDS	sodium dodecyl sulphate
sec	seconds
Ser	Serine
SPE	Solid Phase Extraction
T (domain)	thiolation domain
TE (domain)	thioesterase domain
Thr	Threonine
TIC	total ion chromatogram
TPP	Tripropeptin
t_R	retention time
Trp	Tryptophan
Tyr	Tyrosine
UV	ultraviolet
V	Volt
Val	Valine
VRE	vancomycin-resistant enterococci
XIC	extracted ion chromatogram
αKG	α-ketoglutarate
λ	wavelength
µ	micro
µg	microgram
µL	microliter

List of Tables

List of figures

I. Introduction

1. Cyclic Lipopeptides

The global rise of multi-resistant bacteria strains against most standard of care antibiotics due to mis- and overuse along with a rather dry pipeline has become one of the most concerning challenges in todays health care. This has urged not only doctors but also scientists to come up with improved therapeutic strategies and new antimicrobial agents.[1] Screening for latent antibiotic candidates the class of cyclic lipopeptides seems to be quite underutilized, hence showing tremendous therapeutic potential.[2, 3]

The common structural features among the group of cyclic lipopeptides include a fatty acid side chain linked to a short oligopeptide. The peptide core is usually built up by approximately 2-25 amino acids and forms a lactam or lactone ring (4-14 amino acids), either between two amino acids of the peptide backbone or between an amino acid and a hydroxyl- or amino-group of the lipid side chain.[2, 4]

A noteworthy diversity within this class derives from the peptide portion of the cyclic lipopeptide. The assembly of the peptide backbone is not limited to the 20 proteinogenic amino acids as it can also incorporate around 500 non-proteinogenic ones, including D- and L-configurated, aromatic and aliphatic amino acids as well as basic, acidic and cyclic amino acids, *α*- or *β*-type, *O*- or *N*-methylated and hydroxylated amino acids.[3, 5, 6] Striking is not only the number of non-proteinogenic amino acids but also the occurrence of unusually modified amino acids, which makes the compound less prone to proteolytic digestion by ubiquitous peptidases.[7]

A positive effect on the bioactivity of cyclic lipopeptides also lays within the composition of the lipid side chain. It varies not only in lengths (C_6-C_{18}) but also in its structural conformation resulting in the occurrence of either saturated or desaturated as well as unbranched or branched forms such as *β*-OH-groups, *iso*-, *anteiso*- methyl branched fatty acid side chains.[3]

1.1. Classes of Cyclic Lipopeptides

Natural occurring lipopeptides are secondary metabolites mostly produced by soil-borne or plant-associated bacterial species and can be subdivided based on

distinct structural features.[2] Hereinafter lipoglycopeptides, polymyxins, calcium-dependent lipopeptides and guanidine-containing cyclic lipopeptides will be further highlighted.

1.1.1. Lipoglycopeptides

Glycopeptide antibiotics were already introduced in the market in the 1950s with vancomycin being the first one.[8] Structural features of this group include a sugar moiety attached to a cyclic or polycyclic heptapeptide backbone. Teicoplanin, the second member of the glycopeptides is characterized by an additional acyl-aliphatic side chain.[9, 10]

Over the last decade three semi-synthetic derivatives, telavancin, oritavancin and dalbavancin (Figure 1), were registered and referred to as lipoglycopeptides due to their lipophilic side chain.[11-14]

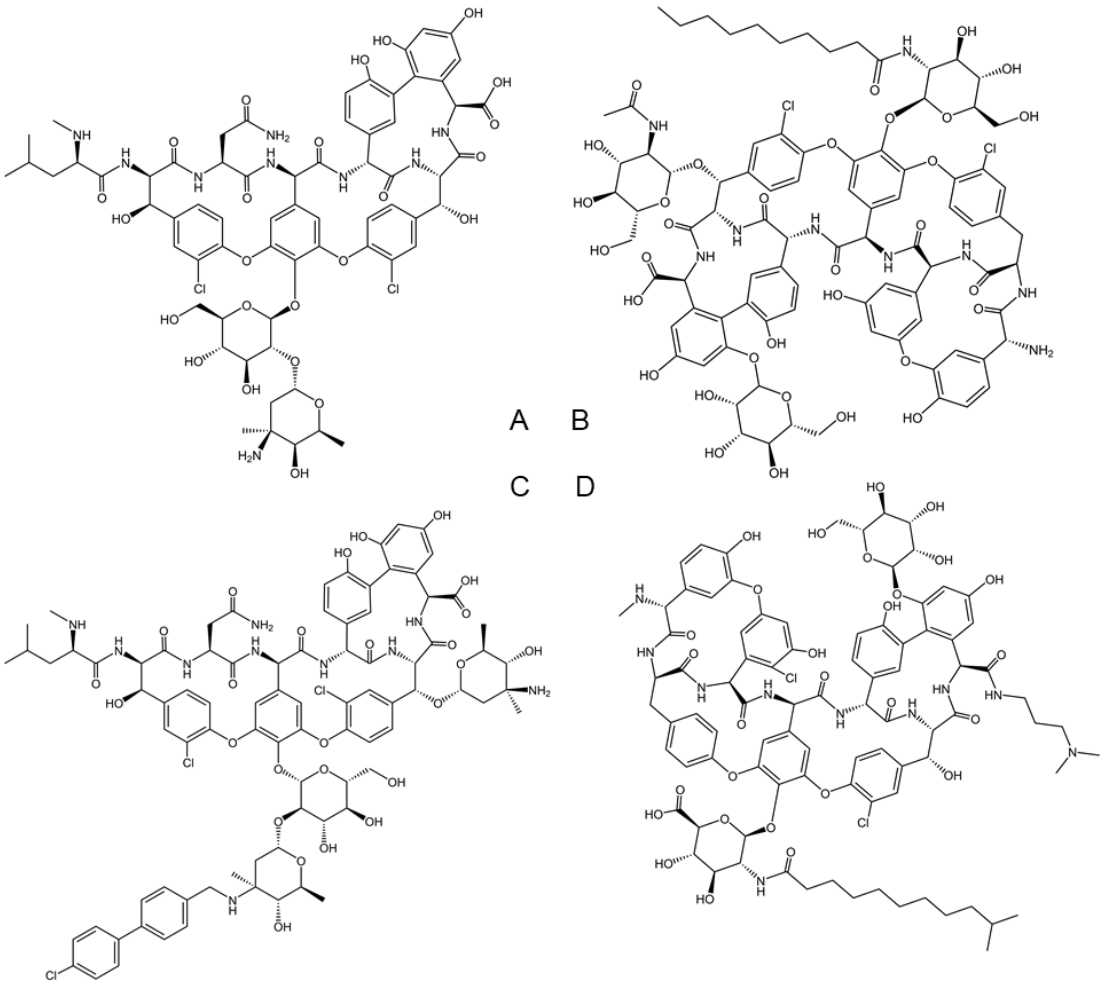


Figure 1: Representatives of the (lipo)glycopeptide antibiotic class: (A) vancomycin, (B) teicoplanin, (C) oritavancin and (D) dalbavancin

While dalbavancin shows a similar drug target as vancomycin by binding to the terminal D-Ala-D-Ala motif of the peptidoglycan precursor, thus inhibiting the late stages of cell wall synthesis, oritavancin and telavancin exhibit an additional mode of action by membrane anchoring via their lipophilic side chain linked to a disaccharide moiety.

The range of indication includes the treatment of skin and soft tissue infections caused by Gram-positive bacteria and hospital-acquired pneumonia. Due to their dual mode of action or stronger inhibition of the transpeptidase than the transglycosylase activity, all three drugs are potent against multi-resistant bugs including vancomycin-intermediate *S. aureus*.[11-14]

1.1.2. Polymyxins

The polymyxin lipopeptide antibiotic family comprise secondary metabolites produced by different species of the Gram-positive soil microbe *Paenibacillus*.[15, 16] In terms of their chemical structure, they share a basic decalipopeptide scaffold, which consists of a heptapeptide loop with a tripeptide side chain acylated by a fatty acid. A remarkable feature of this group is also the incorporation of up to five 2,4-diaminobutyric acid residues within the same molecule.[17, 18]

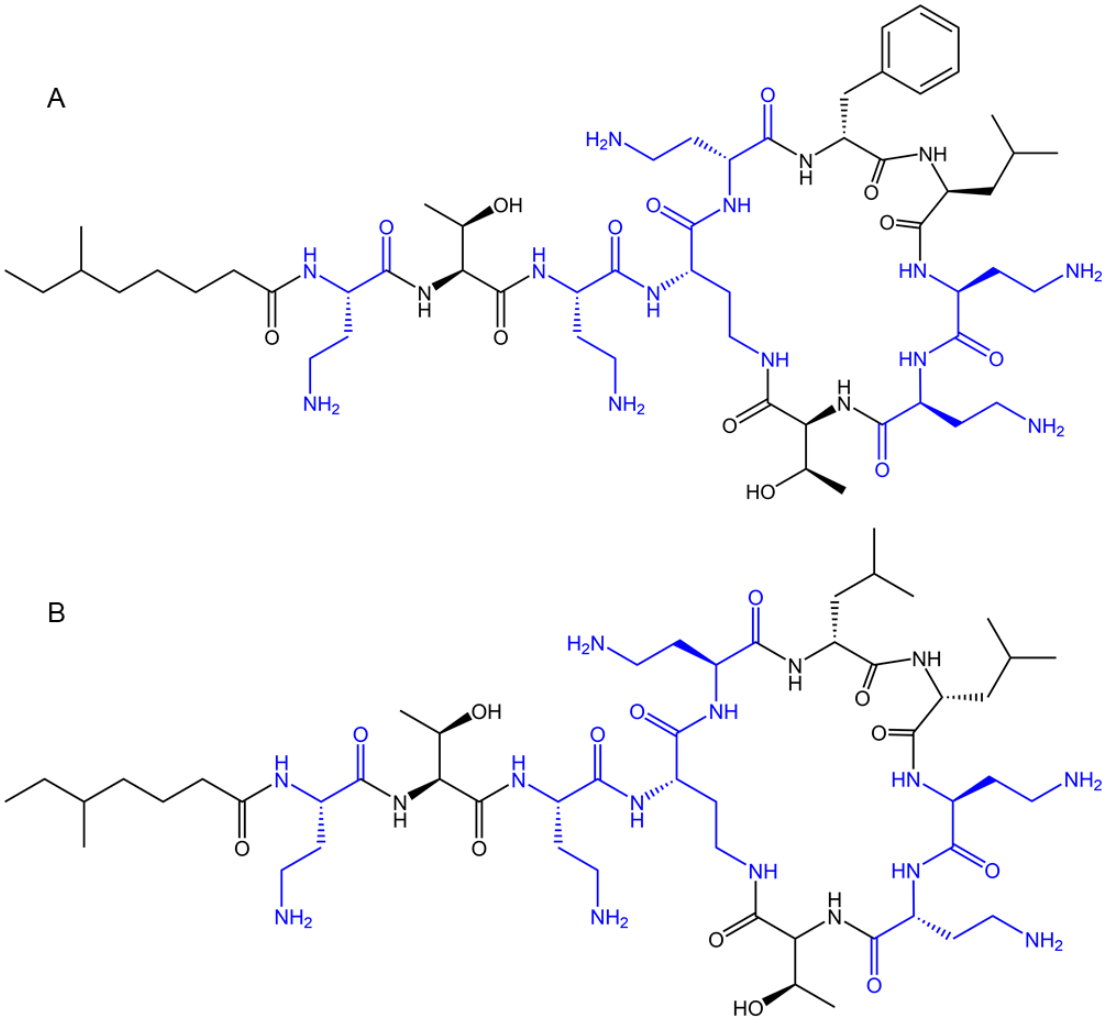


Figure 2: Representatives of polymyxins with their 2,4-diaminobutyric acid residues marked in blue: (A) polymyxin B and (B) colistin

Two representatives of the group, polymyxin B and colistin also known as polymyxin E (Figure 2), were already clinically approved in the 1960s. However, when this class was first introduced into clinical practice, side effects like neurotoxicity and nephrotoxicity were frequently observed leading to their gradually withdrawal from the market.[19] Current studies have proposed that polymyxins show considerably less toxic side effects after changing the dosing strategy, allowing them to slowly re-emerge as last-line treatment for severe infections.[20, 21]

Gram-negative bacteria are mainly targeted by the strong antibiotic activity of polymyxins although they show little effect on Gram-positive pathogens and anaerobic bacteria as well. This can be explained by their mode of action. It is indicated that polymyxin binds the lipid A component of lipopolysaccharide found in the outer membrane of Gram-negative bacteria. Polymyxins seem to be able to

move across the outer membrane by performing a self-promoted uptake.[22] Electrostatic interaction of polymyxin and lipid A cause displacing of calcium and magnesium ions which cross-link adjacent lipopolysaccharide molecules allowing the insertion of the hydrophobic motif and acyl tail of the polymyxin into the hydrophobic membrane. Therefore, this self-promoted uptake is leading to disruption of the cell membrane resulting in a leakage of ions and therefore in cell death.[23] More recently another target was suggested by pointing out that polymyxins are able to bind to ribosomal RNA and therefore interfering with protein translation *in vitro*.[24]

1.1.3. Calcium-dependent Lipopeptides

As the name implies, the representatives of this group share a requirement for calcium ions to display their full antibacterial activity.

One member of this group, daptomycin, forms oligomeric aggregates in the presence of calcium ions masking the overall anionic character of the drug. This allows the disruption of the negatively charged cytoplasmic membrane. Although the exact mode of action is still not completely understood, two hypotheses has been researched. Initially it was suggested that oligomeric aggregates of daptomycin form a pore-like complex resulting in membrane depolarization.[25] Since the depolarization effect is delayed, the insertion at specific lipid rafts seem a more likely target, thereby changing physicochemical properties of the bacterial membrane, affecting among others membrane-bound proteins, which are involved in cell wall biosynthesis and cell division processes.[26, 27]

As the mode of action already indicates, daptomycin is most potent against Gram-positive bacteria, including multi-resistant pathogens and is therefore used for treatment of skin infections, right-sided endocarditis and bacteremia.[28, 29]

The acidic cyclic depsipeptide produced by *Streptomyces roseosporus* consists of 13 amino acids arranged in a 10-membered cyclic lactone ring and three exocyclic amino acids, linked to a decanoyl fatty acid side chain.[30]

Structurally very closely related to daptomycin are friulimicin and amphomycin (Figure 3) as they all share common features including the macrocyclic decapeptide core and a lipid tail, interlinked by exocyclic amino acids.[31, 32]

Notable is also the presence of non-proteinogenic and D-configurated amino acids in all of them.

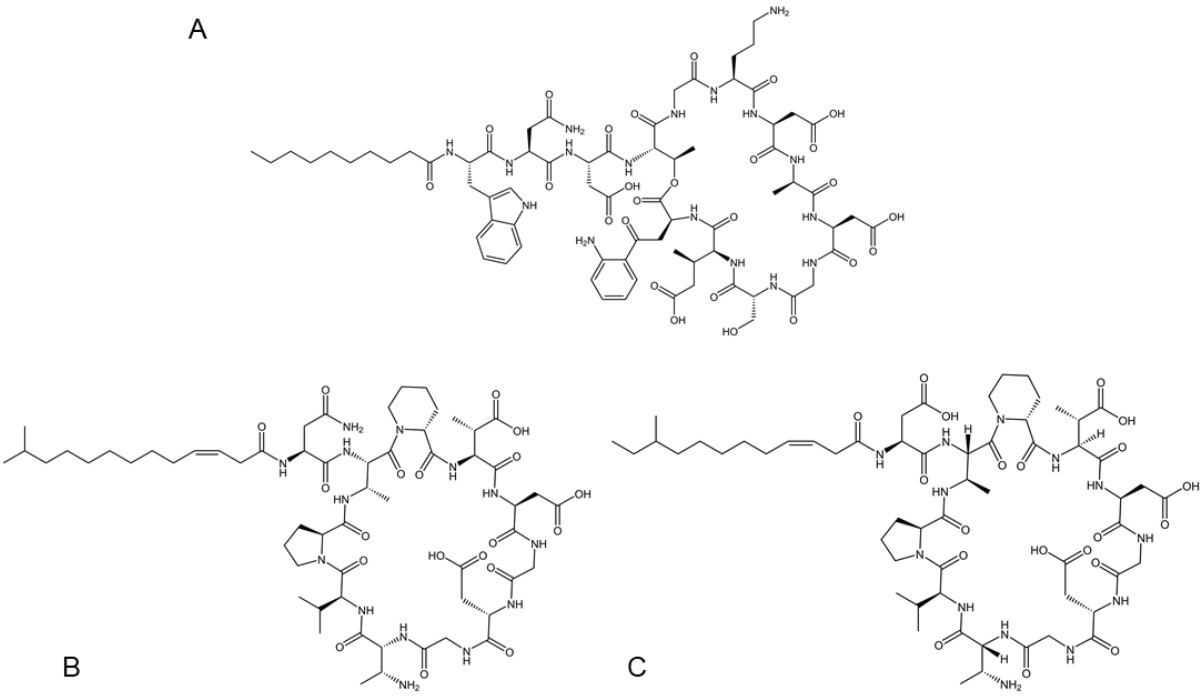


Figure 3: Representatives of calcium-dependent lipopeptides: (A) daptomycin, (B) friulimicin and (C) amphomycin

1.1.4. Guanidine-containing Cyclic Lipopeptides

Based on the position of the guanidine and other structural features this cyclic lipopeptide family can be subdivided into three groups (Figure 4).

The first sub-type of this cyclic lipopeptide family harbors its guanidine moiety in the lipid tail, whereas the other two are defined by the guanidine residue in the arginine structure. The second group additionally displays the structure of a chlorinated cyclic lipopeptide.

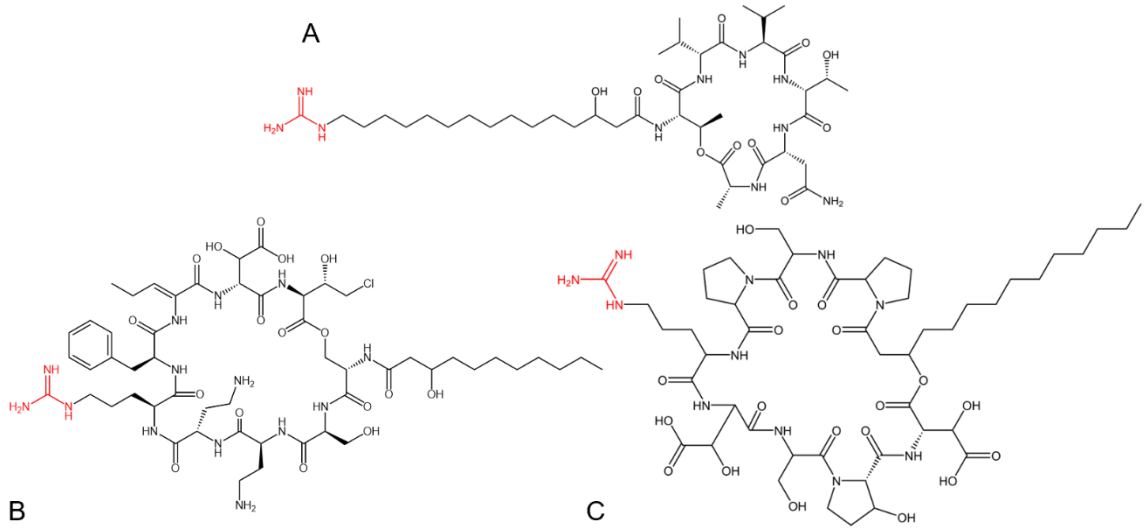


Figure 4: Representatives of guanidine-containing cyclic lipopeptides with the guanidine moiety marked in red: (A) fusaricidin, (B) syringomycin E and (C) empedopeptin

1.1.4.1. Subgroup 1 – Fusaricidin

Isolated from *Paenibacillus* sp., fusaricidins (Figure 4A) share a structural scaffold consisting of a cyclic depsipeptide containing a unique 15-guanidino-3-hydroxypentadecanoic acid chain attached via an amid bond. This lipid side chain is key for the antifungal and antibiotic activity against Gram-positive bacteria like methicillin-sensitive and -resistant *S. aureus* due to its interaction with phospholipid cell membranes.[33, 34]

1.1.4.2. Subgroup 2 - Syringomycins, Syringotoxins and Syringostatins

Syringomycins (Figure 4B), as well as syringotoxins and syringostatins belong to the cyclic lipodepsipeptides produced by *Pseudomonas* spp. as secondary metabolites. This group features the guanidine as part of an arginine and cyclization occurs through a serine residue.[35] These compounds are putative virulence factors involved in plant pathogenicity, but also bear antifungal potency, which is conferred by the C-terminal chlorinated threonine residue.[36, 37]

1.1.4.3. Subgroup 3 – Empedopeptin, Plubacins and Tripropetins

Although produced by different organisms, empedopeptin (Figure 4C), plusbacin and tripropeptin represent the last subgroup of the guanidine-containing cyclic lipopeptides and display the guanidine residue as part of the amino acid arginine.

In contrast to the group around syringomycin, the cyclization process does not take place exclusively via the amino acids but through the 3-OH group of the fatty acid side chain. All members of this family share a highly similar peptide backbone with five common amino acids, of which four belong to the non-proteinogenic type. Variations occur within three amino acids in the northern hemisphere of the cyclic core and the fatty acid side chain, which can differ in length and branching (Figure 5).

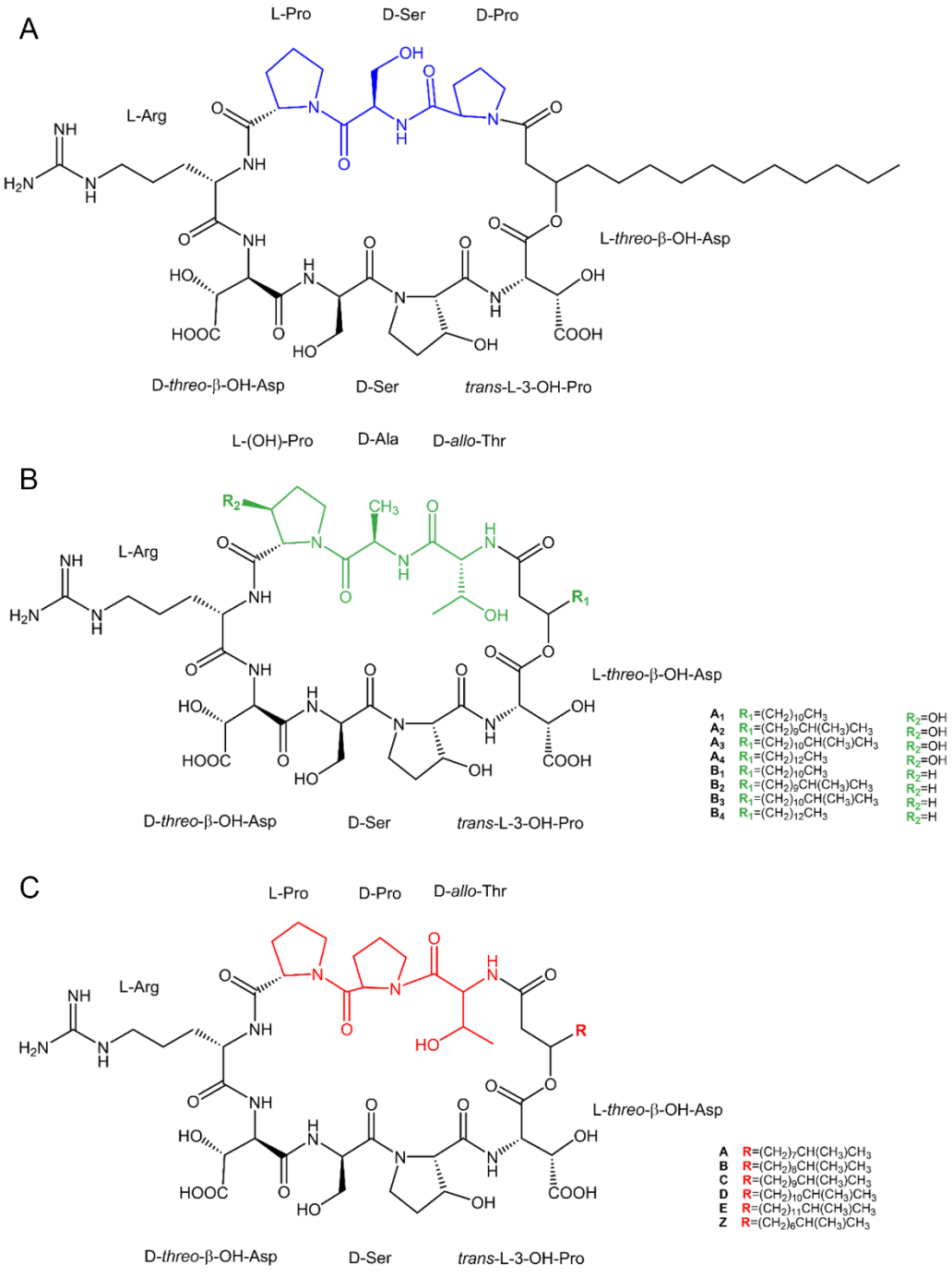


Figure 5: Structural characteristics of (A) empedopeptin, (B) plusbacin and (C) tripropeptin

1.1.4.3.1. Empedopeptin

Isolated in the 1980s from the Gram-negative bacterial strain *Empedobacter* sp. ATCC 31962, empedopeptin is an amphoteric lipodepsipeptide consisting of a cyclic octapeptide core and a 3-hydroxy-myristic acid tail. The macrolactone core is not only assembled by D- and L-configurated but also by modified amino acids like hydroxyaspartic acid and hydroxyproline (Figure 5).[38]

The richness of non-proteinogenic amino acids is remarkable, as well as the potent antibacterial activity of this compound. *In vivo* and *in vitro* experiments, empedopeptin was shown to successfully fight multi-drug-resistant Gram-positive aerobic and anaerobic pathogens like methicillin- and penicillin-resistant *S. aureus*, *Streptococcus pyogenes* and *Clostridium difficile* in combination with low cytotoxicity and good pharmacokinetics (Table 1, Table 2, Table 3).[38]

Table 1: Bioactive properties of empedopeptin - Antibacterial activity against Gram-positive clinical isolates[38]

Test organism	No. of strains	Geometric mean of MIC (µg/mL)	
		empedopeptin	vancomycin
S. aureus (methicillin-resistant)	24	0.9	0.5
S. epidermidis	18	3.1	1.9
S. epidermidis (methicillin-resistant)	3	4.0	2.0
S. agalactiae	7	5.1	0.6
S. pneumoniae	8	0.3	0.5
S. pyrogenes	7	0.3	0.5
S. viridans	5	2.5	0.7
Listeria monocytogenes	7	0.3	1.0
Clostridium difficile	10	4.0	1.2

Table 2: Bioactive properties of empedopeptin - Antibacterial activity dependent on Ca^{2+} [39]

Test organism	MIC (µg/mL)	
	Without Ca^{2+}	+1.25 mm Ca^{2+}
S. aureus ATCC 29213	32	4
S. aureus N315 (MRSA)	32	4
S. aureus SG 511	8	1
Staphylococcus simulans 22	16	1
S. pneumoniae DSM 11865 (PRSP)	1	0.25
Streptococcus pyogenes ATCC 10389	1	0.25
Enterococcus faecalis ATCC 29212	16	8
Enterococcus faecium BM4147 (VRE)	32	16
B. subtilis 168	8	1
M. luteus DSM 1790	0.125	<0.062
E. coli W3110	>64	>64

Table 3: Bioactive properties of empedopeptin - Mouse in vivo median protective dose (PD_{50}) against common Staphylococci and C. perfringens strains [38]

Test organism	PD_{50} (mg/kg)	
	empedopeptin	vancomycin
S. aureus Smith	3.3	1.3
S. aureus BX-1633 (methicillin-resistant)	3.6	2.5
S. aureus A15097 (penicillinase producer)	1.1	1.1
S. aureus A20609 (penicillinase producer)	2.4	0.8
S. pneumoniae A9584	0.94	0.74
S. pyrogenes A20201	0.82	0.82
Clostridium perfringens A9635	6.8	1.3

These tremendous therapeutic effects might be explained by the proposed dual mode of action. Studies have revealed that the cell wall building block lipid II is empedopeptin's most likely primary target. In the presence of calcium ions, the lipopeptide chelates with undecaprenyl phyrophosphate-containing cell wall precursors in a 2:1 molar ratio. Additionally, secondary targets like undecaprenyl-pyrophosphate carrier (C_{55}-PP) and the cell-wall teichoic acid precursors lipid III and lipid IV were pinpointed, proposing synergetic interactions combating Gram-positive pathogens by disrupting different mechanisms within the peptidoglycan biosynthesis.[39]

1.1.4.3.2. Plusbacin

Similar to empedopeptin and tripropeptin, plusbacins (Figure 6) display a cyclic core with eight amino acids cyclized through a lactone linkage between L-*threo*-β-hydroxyaspartic acid and a hydroxy fatty acid subunit. Within the plusbacins, there a two subclasses, A and B. Plusbacins A (A_{1-4}) exhibit L-hydroxyproline at the third position of the macrolactone core whereas in the plusbacin B (B_{1-4}) series this moiety is switched with proline. The variation of the subgroups of plusbacin can be traced back to the differences of the lipid side chain.[40]

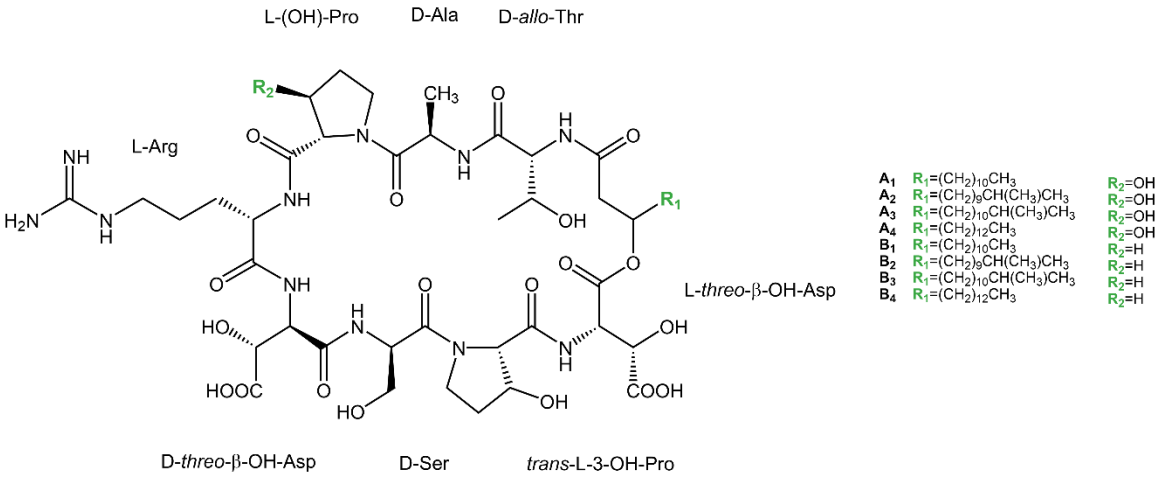


Figure 6: Chemical structure of plusbacins

The plusbacin series was discovered in the early 1990s and it was isolated from the Gram-negative bacterium *Lysobacter firmicutimachus*.[40, 41] Bioactivity studies showed that all plusbacins have antibiotic qualities against Gram-positive bacteria, including multi-resistant human pathogens, while Gram-negative bacteria are not harmed.[40, 42] Considering a rising vancomycin resistance, especially plusbacin A_3 seems to be a quite interesting drug candidate. Although also interfering with peptidoglycan biosynthesis, plusbacin seems to target a nascent step within this mechanism by hindering the formation of lipid intermediates while vancomycin is binding to the D-Ala-D-Ala residue of the lipid II pentapeptide. Due to their distinct modes of action, plusbacin A_3 still shows activity against vancomycin resistant bugs while overall displaying similar minimal inhibition concentration values as vancomycin (Table 4).[42]

Even though the exact mode of action is not completely understood, studies propose that the impact of the isotridecanyl side chain on the activity of plusbacin

A_3 cannot be ignored. By analyzing *Staphylococcus aureus,* it was demonstrated that an incorporation of the isotridecanyl side chain into the bacterial cell wall takes place (Table 5).[43]

Table 4: Bioactive properties of plusbacin derivatives - MIC values for all plusbacin derivatives as well as vancomycin against Gram-positive bacteria[40, 42]

Test organism	MIC (µg/mL)								
	plusbacins								vancomycin
	A_1	A_2	A_3	A_4	B_1	B_2	B_3	B_4	
S. aureus 209 JC-1	0.05	0.1	0.1	0.1	0.1	0.1	0.2	0.4	
S.aureus Smith	1.6	0.8	0.4	0.4	1.6	0.8	1.6	1.6	0.78
S. aureus SR5597	1.6	0.8	0.4	0.4	1.6	0.8	0.8	0.8	
S. aureus SR5598	1.6	0.8	0.8	0.4	1.6	0.8	1.6	0.8	
S. aureus SR5580	1.6	0.8	0.8	0.4	1.6	0.8	1.6	1.6	
S. aureus SR5584	1.6	1.6	0.8	0.4	3.1	1.6	1.6	1.6	
S. epidermidis A14990	0.8	0.2	0.2	0.4	0.8	0.4	0.4	0.8	
E. faecalis SR1004	6.3	3.1	1.6	1.6	6.3	1.6	1.6	3.1	1.56
E. faecalis SR4512	6.3	3.1	1.6	1.6	12.5	3.1	3.1	6.3	
E. faecalis SR7914			1.56						>50
E. faecalis SR7917			3.13						>50

Table 5:Bioactive properties of plusbacin derivatives - MIC values of plusbacin A3 and deslipo-plusbacin A3 against Gram-negative and Gram-positive bacteria[43]

Test organism	MIC (µg/mL)	
	plusbacin A_3	*deslipo*-plusbacin A
S. pyogenes M49 NZ131	0.2	>50
S. agalactiae A909	6.25	>50
S aureus ATCC33591	1.56	>50
E. faecalis ATCC51299	6.25	>50
E. coli ATCC25922	>50	>50
P. aeruginosa ATCC27853	>50	>50

1.1.4.3.3. Tripropeptin

Being the most potent out of the subgroup of cyclic guanidine-containing lipopeptides, tripropeptins A-E and Z share the same peptide nucleus but only differ in the lipid side chain (Figure 5). A comparison of the antibacterial activity within this compound complex shows that tripropeptin D seems to be the representative with the most therapeutical effects (Table 6). Noteworthy is the fact that tripropeptin D has the second longest acyl side chain, hinting that there might be a correlation between the length of the acyl chain and the molecules bioactivity.[44-48]

Table 6: Bioactive properties of tripropeptins - MIC values against Gram-positive pathogens for tripropeptin derivatives and vancomycin[44, 45]

Test organism	MIC (µg/mL)						
	tripropeptins						vancomycin
	A	B	C	D	E	Z	
Staphylococcus aureus FDA 209P	1.56	1.56	1.56	0.39	0.78	12.5	0.39
S. aureus Smith	0.78	1.56	1.56	0.39		6.25	0.39
S. aureus MS9610	6.25	3.13	1.56	0.78		25	0.39
S. aureus MS16460 (MRSA)	6.25	6.25	3.13	1.56		50	
S. aureus MS16497 (MRSA)	6.25	3.13	1.56	0.78		25	
S. aureus MS16526 (MRSA)	3.13	3.13	1.56	0.78	0.78	25	0.78
S. aureus TY-00933 (MRSA)	6.25	3.13	3.13	0.78		25	0.78
S. aureus TY-03454 (MRSA)	6.25	3.13	1.56	0.78		25	0.39
S. aureus TY-03456 (MRSA)	6.25	3.13	3.13	0.78		25	0.78
S. aureus TY-04282 (MRSA)	6.25	3.13	1.56	0.78	0.78	25	0.39
Enterococcus faecalis JCM5803	50	25	12.5	3.13		100	0.78
E. faecalis NCTC 12201 VCMR	50	12.5	6.25	3.13	1.56	100	>400
E. faecalis NCTC 12203 VCMR	50	50	25	6.25	3.13	>100	400
E. faecalis JCM5804	50	25	12.5	6.25		>100	0.78
E. faecalis NCTC 12202 VCMR	50	25	12.5	6.25	3.13	>100	>400
E. faecalis NCTC 12204 VCMR	50	25	12.5	6.25	3.13	>100	400

Like empedopeptin, tripropeptin C also seems to have a calcium-dependent mode of action. Forming an equimolar complex with undecaprenyl-pyrophosphate and calcium ions, tripropeptin C inhibits membrane-associated steps of peptidoglycan biosynthesis by hindering the transglycosylation step or flippase activity.[49]

2. Non-ribosomal Peptide Synthetases – Biosynthesis of Cyclic Lipopeptides

A myriad of naturally occurring bioactive peptides have been described in the past. Although the exact biosynthesis differs, they can be classified in two categories, ribosomally or ribosome-independent synthesized peptides. With a few exceptions[50], cyclic lipopeptides belong to the second group and are therefore assembled by modular multienzyme complexes, the so-called NRPSs. In contrast to the method of proteins being synthesized by ribosomes, NRPS-derived peptides are often drastically modified and contain a variety of non-proteinogenic amino acids.

The assembly line of NRPSs, which unlike the ribosome is independent of mRNA, is organized in modules subdivided into three essential subunits called adenylation (A) domain, thiolation (T) or peptidyl carrier protein (PCP) domain and condensation (C) domain. Depending on the structure of the secondary metabolite, further domains such as epimerases (E), cyclization (Cy) and methyltransferase (MT) domains must also be existent, and an additional thioesterase (TE) domain usually marks the termination module.

2.1. Elongation of Non-ribosomal Peptides

2.1.1. Adenylation (A) Domains

The first step of the elongation process of non-ribosomal peptides is controlled by an approximately 550 amino acid-long A domain. Being a so-called adenylate-forming enzyme, the A domain is responsible for selective substrate recognition and activation.[51]

The A domains of NRPSs share several conserved sequences, which are crucial for their proper function. Studies on crystal structures revealed that A domains consist of two domains - a small C-terminal and a large N-terminal domain- with the active site at the junction linking them together.[52-54]

Noteworthy seems also the association of some A domains with MbtH-like proteins (MLPs), which serve as chaperones and oftentimes are directly encoded in the gene cluster.[55, 56]

Representing the gatekeepers of the assembly line, A domains contain an 8-10 residue long motif necessary for the interaction with the amino acid. Changes within this region lead to altered substrate specificity.[57-59]

The A domain specifically selects any incorporated amino acid and activates it by adenylation using Mg^{2+}-ATP. The resulting aminoacyl-adenylate intermediate gets transferred onto the thiol group of the 4'-phosphopantetheinyl cofactor of the peptidyl carrier protein releasing AMP (Figure 7).[51]

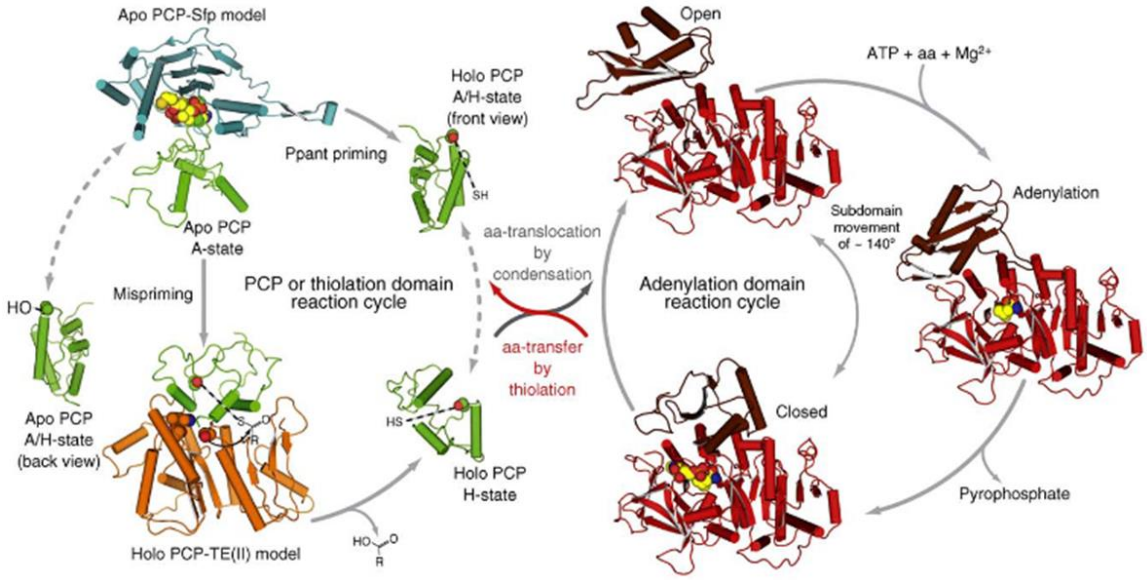


Figure 7: Peptide carrier protein with attached 4'-phospho-pantetheine and adenylation domain cycle adapted from Strieker et al.[60]

2.1.1.1. Adenylation Domain Specificity

Responsible for the recruitment of the amino acids, the order of the A domains along the NRPS assembly line usually controls the primary sequence of the final peptide chain. Therefore, it is desirable to predict the specificity of the A domains, to foresee the putative primary structure of the peptide on a genetic level.

2.1.1.1.1. Bioinformatic Prediction

Co-crystallization studies of A domains and amino acid monomers allowed further insights into the structure of the binding pocket. Based on the active site configuration and the residues which interact with the substrate, a specificity-conferring code of A domains, commonly referred to as Stachelhaus code, was proposed.[61] Further refinements do not only take the 10 interacting active site

residues into account but also use all amino acid residues within 8 Å of the A domain substrate to predict its specificity. Nowadays, the prediction process using just the genetic sequence is quite simplified as there are many bioinformatical tools built on this specificity-conferring code such as NRPSpredictor2.[62] It is worth mentioning that these bioinformatics tools are extremely useful, however they do not obviate the need for experimental adenylation enzyme characterization to verify the A domain specificity.

2.1.1.1.2. Experimental Adenylation Enzyme Characterization

To understand the principles of the different specificity assays, the underlaying biochemical mechanism of the A domain needs to be highlighted.

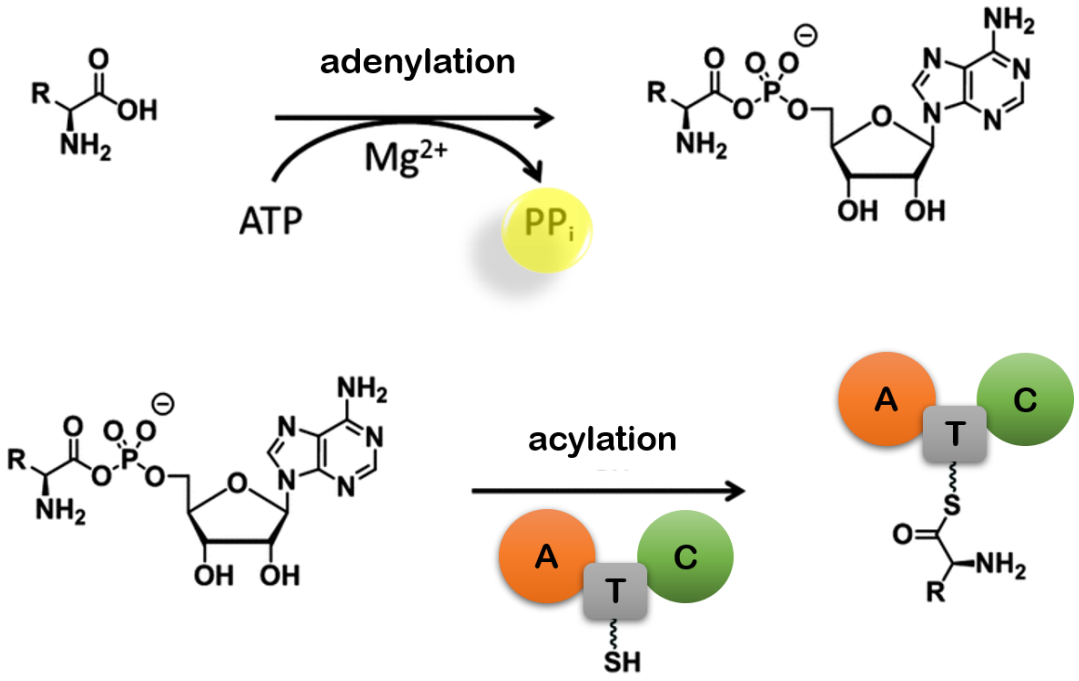


Figure 8: Enzymatic reaction catalyzed by the A-domain of the NRPS assembly line

The adenylation enzyme conducts a two half-step reaction comprising of an adenylation followed by an acylation (Figure 8). Binding the carboxylic acid substrate, which is usually an amino acid and ATP, a nucleophilic attack results in pyrophosphate alongside an acyl-adenylate, which is tightly bound to the active site. In the second half-reaction, the transfer of the acyl moiety of the activated amino acid onto the phosphopantetheinyl arm of the PCP domain yields the thioester tethered carboxylic acid building block.[63, 64]

A commonly applied assay to determine A domain activity is the pyrophosphate exchange radioassay (Figure 9). Detecting the incorporation of ^{32}P-PP$_i$ into ATP, this assay is quite sensitive although the handling of radioactivity is a major drawback.[65] This disadvantage is eliminated by using a mass spectrometric-based pyrophosphate exchange assay which employs γ-$^{18}O_4$-ATP.[66] It needs to be taken in consideration that both exchange assays are only able to assess the first half-step adenylation reaction of the A domain in a reverse direction. This can lead to a distortion of the kinetic parameters, since the absence of an acceptor molecule such as the T domain means that the acyl-adenylate intermediate is just slow getting out of the active site, which might lead to a slow turnover.[64]

Both aforementioned methodologies are discontinuous end point assays whereas release assays collect data about the liberation of PP$_i$ in the forward reaction in a continuous manner.[67, 68] Since the tightly bound aminoacyl-adenylates in this assay is quite detrimental, adding a reactive surrogate to take over the task of the PCP domain achieves a faster enzymatic turnover. Adding hydroxylamine as a quencher (Figure 9), the hydroxylamine-MesG coupled adenylation assay continuously monitors the formation of 7-methylthioguanine, derived from the conversion of 7-methylthioguanosine by purine nucleoside phosphorylase. This enzyme uses phosphate as a substrate which resulted from the enzymatic cleavage of pyrophosphate produced during the first-half reaction of the A domain.[63, 64]

The hallmark of the lately developed multiplexed hydroxamate assay (HAMA) is its ability to take substrate competition into account. Like the hydroxylamine-MesG adenylation assay, HAMA also uses hydroxamate quenching but differs in the detection of adenylation activity as hydroxamate products are evaluated by UPLC-MS/MS. This method allows parallel testing of dozens of competing amino acid substrates in one single chromatographic run.[69]

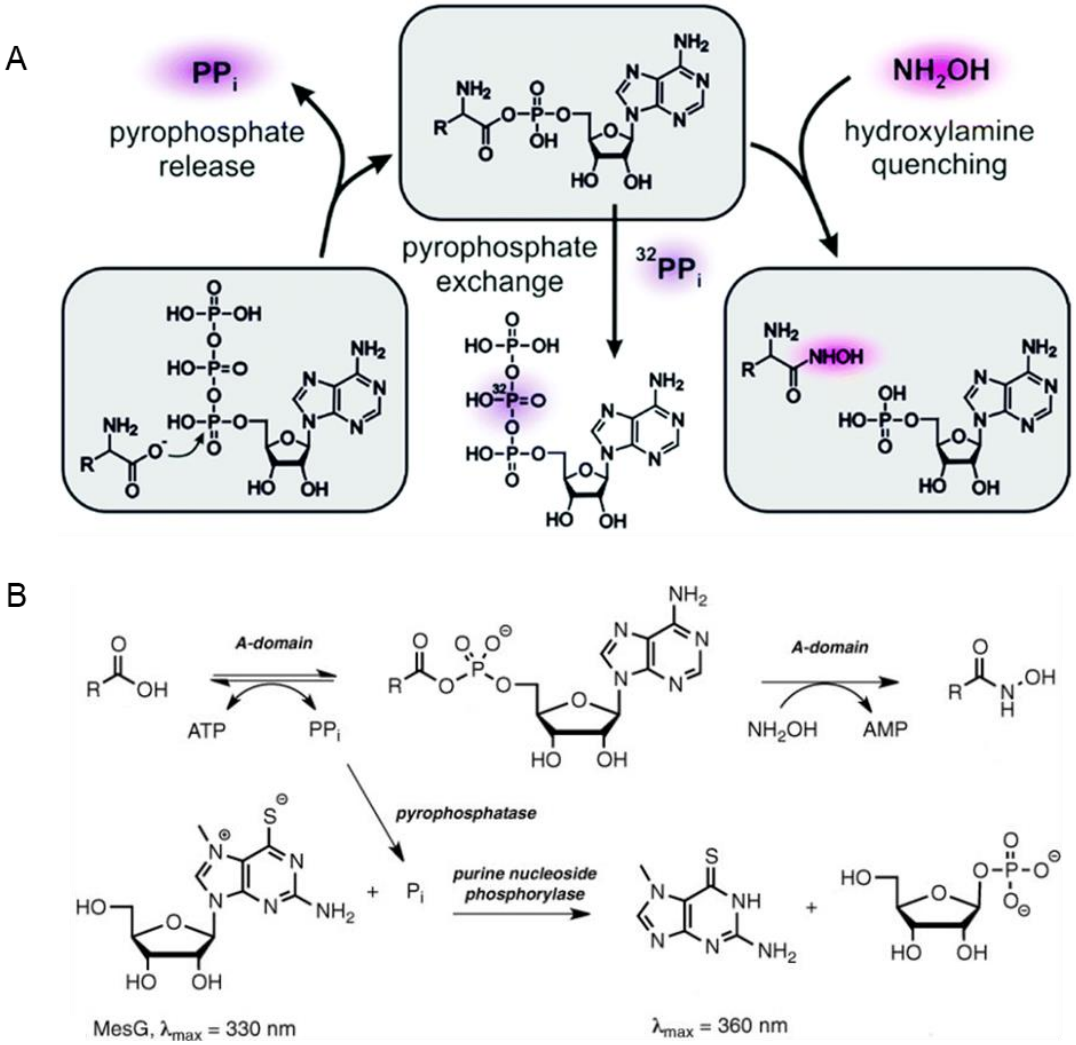


Figure 9: (A) Different assay types for the measurement of Nonribosomal Peptide Synthetase Adenylation Domain Activity adapted from Stanišić et al.[69] and (B) the principle of hydroxamate formation assay using the example of hydroxamate-MesG assay adapted from Duckworth et al.[63]

2.1.2. Thiolation (T) Domains/ Peptide Carrier Protein (PCP) Domains

The previously mentioned PCP domain, also known as T-domain, is the second essential unit of a NRPS assembly line. Immediately located downstream of the A domain, the T domain inheres just 80-100 amino acids making it the smallest catalytic element of the multi-enzyme complex.

Prior to taking part in the elongation of the non-ribosomal peptide, the PCP domain needs to undergo priming during which the inactive apoform is activated to its holoform (Figure 10). By forming a phosphodiester linkage between the 4'-phosphopantetheine of a coenzyme A and the conserved serine residue of PCP, phosphopantetheinyl transferases (PPTases) prime each T domain of the NRPS post- translationally.[70]

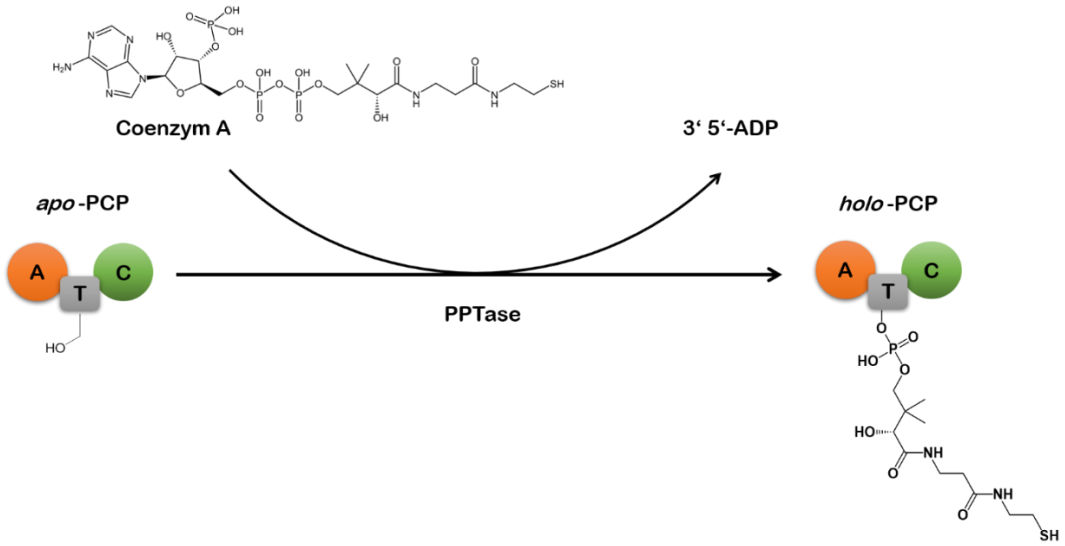


Figure 10: Activation of the NRPS T domain

The by ATP hydrolysis activated aminoacyl adenylate intermediated is relocated and fixed to the terminal thiol group of the prosthetic moiety 4'-phosphopantetheinyl cofactor of the PCP. The covalent bond during this step allows for further modifications of the peptide such as N-methylation or epimerization catalyzed by additional domains.

During assembly, the growing peptide intermediate is continuously passed on from one PCP domain to the following, thereby elongating the chain by one amino acid at a time.[55]

2.1.3. Condensation (C) Domain

The actual elongation, thus the formation of peptide bonds is catalyzed by the C domain linking the upstream and downstream building blocks. The C domain, which is about 450 amino acids long, mediates a nucleophilic attack of the amino group of the downstream amino acid on the activated upstream thioester group, resulting in the growing peptide chain being bound to the downstream PCP domain (Figure 11).[55, 71]

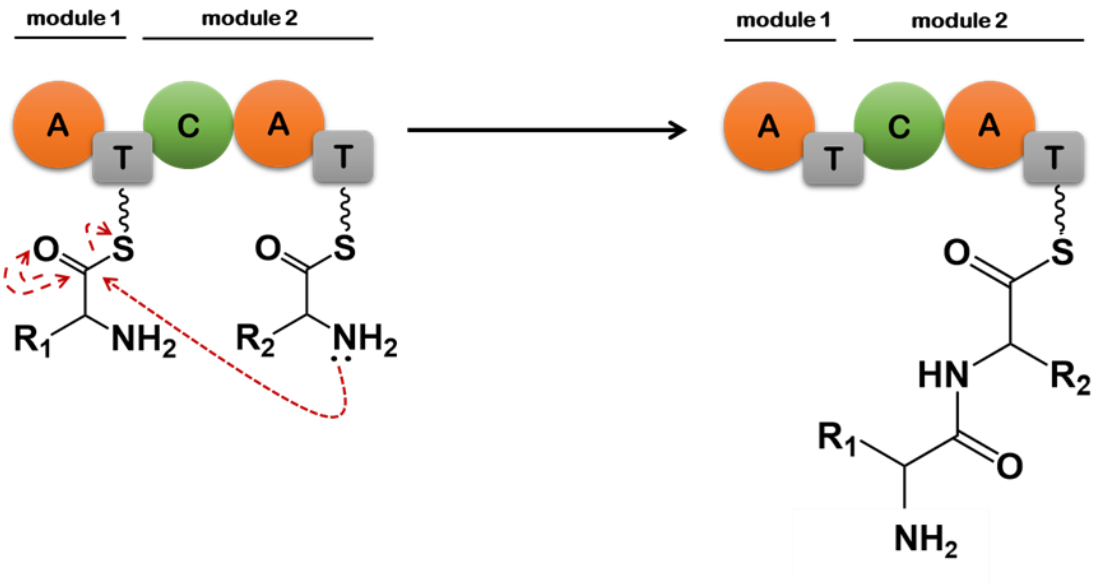


Figure 11: Peptide bond formation through the NRPS C domain

Structurally, the C domains can be described as a V shaped pseudo-dimer consisting of two lobes, the N-terminal lobe and the C-terminal lobe with an active site at their junction.[55]

The acceptor- and donor binding sites of the C domain display a selectivity for the electrophile and nucleophile, respectively which is not limited to the substrate identity as the binding sites are also exhibiting stereo selectivity.[57] Especially the acceptor site appears to be highly selective, proving that although the A domains primarily singles out the suitable substrate, C domains do not accept any substrates presented to them.[72]

Further insight into the structure of C domains revealed a conserved sequence required for proper elongation of the peptide chain via a nucleophilic attack. Especially the second histidine residue within the core motif HHxxxDG seems to have an important role for catalytic activity of the C domain.[71]

Although all C domains generally perform peptide bond formation, there are a few functionally distinct NRPS C domains known. The $C_{starter}$ domain does not catalyze the formation of a peptide bond between two amino acids but instead attaches a fatty acid to the first amino acid through an acylation reaction. The $^{L}C_{L}$ domain links two L-amino acids whereas the $^{D}C_{L}$ domain attaches an L-amino acid to the activated peptide chain ending with a D-amino acid. Similar to the described $^{D}C_{L}$ domain is the dual epimerisation/condensation (E/C) domain, but due to the

epimerase activity the chirality of the last in the growing peptide integrated L-residue is flipped into its D-form.[73]

2.2. Termination Stage

2.2.1. Thioesterase (TE) Domains

Following the elongation process, the mature peptide needs to be released from the NRPS assembly line. Positioned at the C-terminus of the final NRPS module, TE domains terminate the assembly line in a two half-step reaction (Figure 12). The roughly 250 amino acids long domain is equipped with a serine residue within a catalytic triad in its active site binding the peptidyl group from the PCP to form an O-acyl intermediate.[74] Subsequently, three possible reactions can be catalyzed. The intermolecular attack of a water molecule results in the hydrolysis and in the release of a linear mature peptide. Alternatively, a cyclic peptide can be produced. Either the cyclization can occur intramolecular due to an attacking nucleophilic group of the peptidyl moiety attached to the TE domain or the nucleophile being part of a newly synthesized peptidyl moiety attached to the PCP domain. The latter case leads oligomeric substrate as a mature peptide.[75-77]

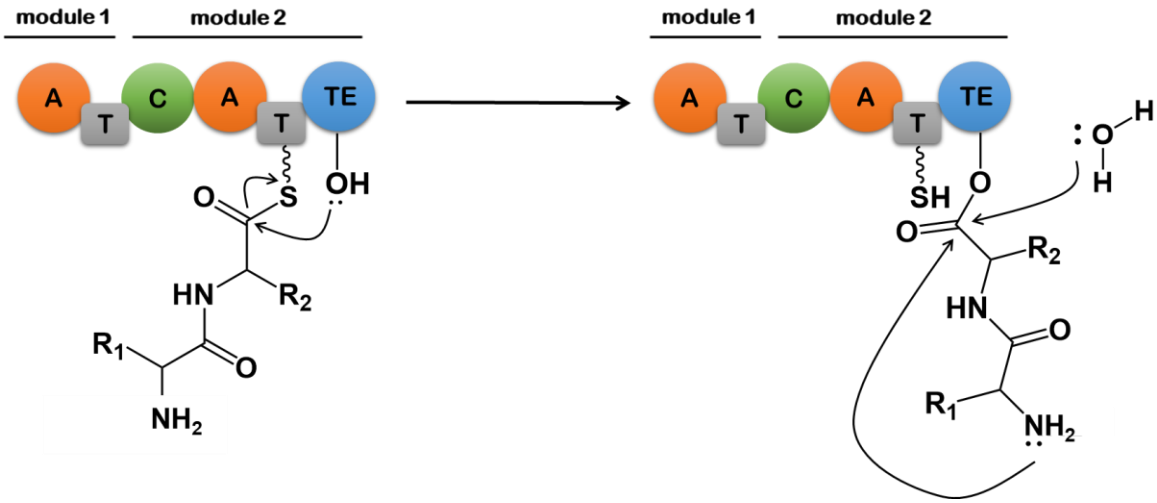


Figure 12: TE domain functionality as terminator of the NRPS assembly line

2.3. Tailoring Enzymes that Modify Non-ribosomal Peptides

Besides the previously described cascade of multi-modular assembly steps of non-ribosomal peptides by amino acid recruiting, amide-bond formation, translocation of the growing peptide chain and finally thioester lysis to release either a linear or

cyclic product, further enzymatic reactions can occur. In addition to the possible assembly of a vast variety of unusual amino acids, this leads to a high structural diversity within the secondary metabolites. Alleging an example some compounds like erythromycin, novobiocin or even vancomycin bear desoxysugars linked to their structure via glycosyltransferases.[78-80]

Along with the attachment of sugars, some common modifications coordinated by dedicated tailoring enzymes are methylation, epimerization, heterocyclization or hydroxylation.[81-83]

Depending on the genetic blueprint and the exact enzymatic mechanism, these specialized enzymes are either part of the NRPS assembly line at a module where unusual alterations occur or stand-alone enzymes. Noteworthy is also the timing of the catalyzed action as the modification may occur prior or after the release of the peptide. Tailoring reactions that do not take place at the full-length peptide can also be subdivided. Either the precursor gets enzymatically altered prior to entering the assembly line, or the tailoring enzyme acts during elongation, while the growing peptide chain is covalently bound to the T domain.[84]

The following chapter focuses on some tailoring enzymes that introduce hydroxy amino acids into a non-ribosomal peptide.

2.3.1. Oxygenases as Tailoring Enzymes

Prevalent in nature, oxygenases are a group of enzymes that oxidize a variety of substrates by incorporation of molecular oxygen. The enzymatic action is unlike those of oxidases, which catalyze dehydrogenation. However, both groups share most of their cofactors such as heme, nonheme iron and flavin.

Transferring either one or two oxygen atoms, this enzyme class can be subdivided into monooxygenases and dioxygenases, respectively.

2.3.1.1. Monooxygenases

The group of monooxygenases is catalyzing the integration of just one atom from molecular oxygen into its substrate. Needing a reducing agent to reduce one of the two molecular oxygen atoms into water, the monooxygenase reaction displays an

oxygenation as well as an oxidase reaction. Consequently, this subgroup is in some cases even known as a mixed function oxidase. Frequently also being referred to as hydroxylase, this term is misleading. Although monooxygenases are generally able to conduct hydroxylation reactions, dioxygenases are too.

Depending on the source of the electron donor, monooxygenases can be further subdivided. In the class of internal monooxygenases some parts of the substrate being hydroxylated serves as an electron donor itself.[85] An example thereof is lactate oxidative decarboxylase from *Mycobacterium phlei* or the flavoprotein monooxygenase EncM, involved in enterocin biosynthesis.[86, 87] As opposite to the substrate supplying the reducing agent, more common types of monooxygenases, also known as external monooxygenases, require various co-substrates as types of electron donors. To name just a few, a wide range of flavin-dependent monooxygenase-like bacterial luciferase or heme-containing monooxygenases such as cytochrome P450 monooxygenases categorize in this group.[85, 86, 88, 89]

2.3.1.2. Dioxygenases

The earmark of the catalyzed enzymatic reaction of dioxygenases is the incorporation of two atoms of molecular oxygen into either one substrate or two substrates. The first instance in which one substrate acts as oxygen acceptor for both atoms, the enzyme is referred to as intramolecular dioxygenase, whereas intermolecular dioxygenases contribute to an enzymatic reaction in which both substrates act as an acceptor for one oxygen atom each. As far as is known one substrate in the latter case invariably is α-ketoglutarate.[85]

Needing a cofactor for proper function, dioxygenases can contain copper as a prosthetic group, or iron in form of nonheme or in rare cases heme.[85, 90]

2.3.1.2.1. 2-Oxoglutarate-Dependent Oxygenases

Comprising a large family of enzymes, 2-oxoglutarate- also known as α-ketoglutarate-dependent oxygenases are widely spread within most forms of life. The ubiquity of this enzyme class goes along with a multitude of oxidative transformations that are crucial in many biochemical processes of bacteria and higher organisms. Notable is the involvement of those enzymes in the metabolism

of fatty acids, DNA and RNA repair, posttranslational modifications of collagen and in the biosynthesis of a variety of secondary metabolites including many antibiotics.[91-93]

Catalyzing oxidative transformations such as hydroxylation, halogenation, ring formation or expansion and desaturation, this enzyme class acts on a broad spectrum of substrates incorporating O_2 using mononuclear non-heme iron and the co-substrate α-ketoglutarate, which are converted into succinate, CO_2 and the desired product.[93]

Studies on several α-ketoglutarate-dependent oxygenases revealed common structural features. Binding their mononuclear iron at a triad of two histidine and one carboxylate of either glutamic or aspartic acid residues, this 2-His-1-carboxylate motif is located within the opening of a double-stranded β-helix fold.[94-96]

The well-studied model enzyme TauD is considered the archetype Fe(II)/2OG oxygenases, catalyzing substrate hydroxylation. With taurine being the substrate, hydroxylation results in an unstable intermediate which spontaneously degenerates to aminoacetaldehyde and sulfite.[97] A closer look into the enzymatic mechanism illustrated in Figure 13 postulates that the hydroxylation reaction starts with three water molecules coordinated to the active site Fe(II). Two of the metal-bound water molecules are replaced by 2-oxoglutarate as it occupies the metallocenter in a bidentate coordination, binding to the iron atom with its keto as well as with its carboxylate group. As the primary substrate links to the less conserved regions of the enzyme near the active site Fe(II), the third remaining water molecule gets displaced. The vacant coordination interacts with O_2 presumably resulting in a Fe(III)-superoxo species, which attacks the C atom of the keto group of the co-substrate. The formation of this peroxohemiketal bicyclic intermediate triggers the oxidative decarboxylation of α-ketoglutarate. Consequently, CO_2 is being released while an Fe(IV)-oxo intermediate with bound succinate stays in the active center. The ferryl intermediate shows oxidizing properties causing radical formation by hydrogen atom abstraction. To complete the process, a hydroxyl radical rebound finalizes the hydroxylation and the dissociation of the hydroxylated compound. Alternatively, the deprotonation of the

Fe(III)-OH intermediate, which was yielded during radical formation is hypothesized. This results in the dissociation of the products by an Fe(III)-oxo followed by an Fe(II)- alkoxo intermediate step.[98-100]

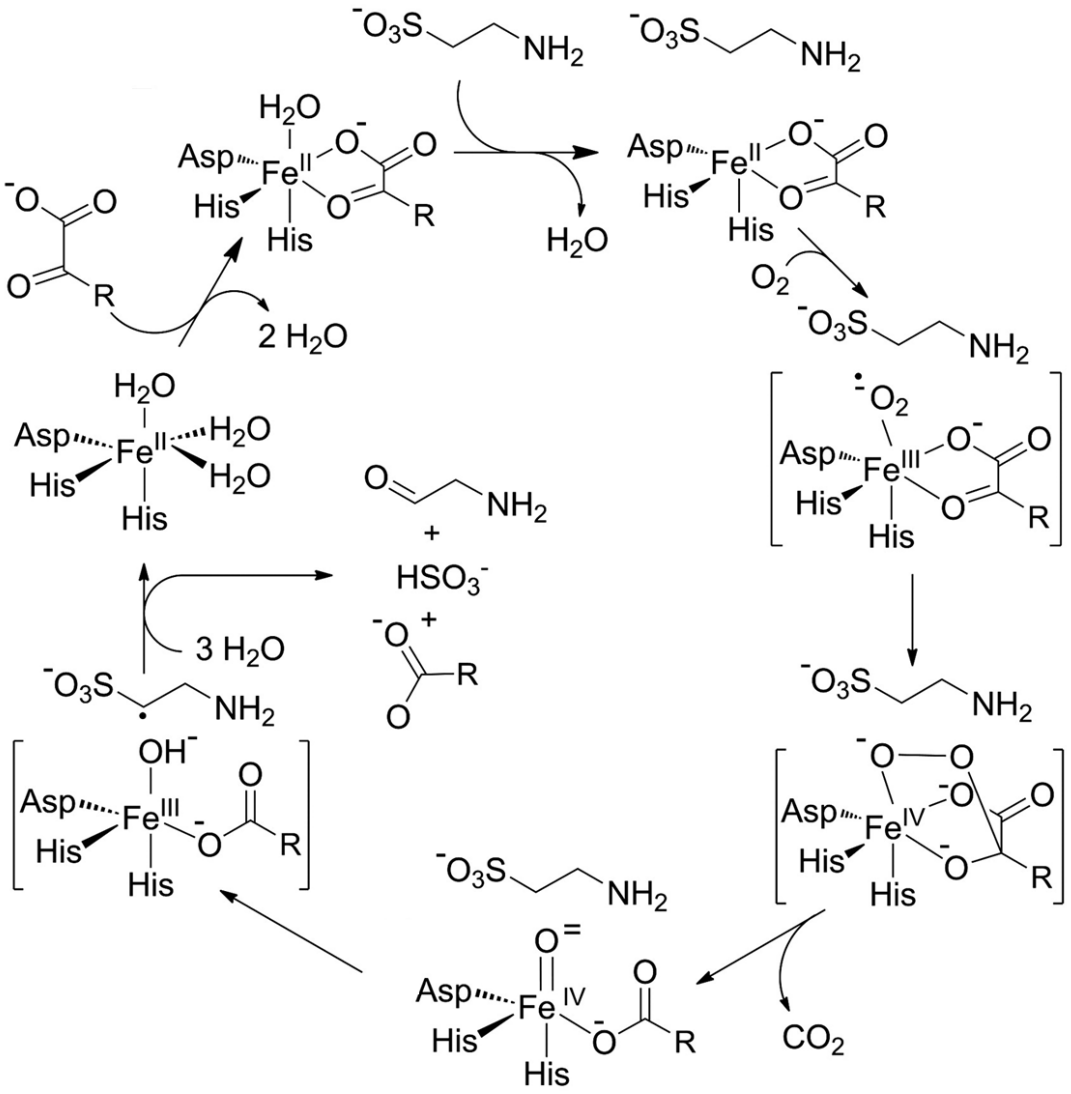


Figure 13: Postulated enzymatic mechanism catalyzed by TauD as a model enzyme for 2-oxoglutarate-dependent dioxygenases adapted from Grzyska et al.[100]

3. Aim of Research

A hallmark of nonribosomal peptides is the occurrence of nonproteinogenic amino acid residues within their structure. The group of guanidine-containing cyclic lipopeptides around empedopeptin, plusbacin, and tripropeptin share not only a common structural framework but also the presence of hydroxy-proline and hydroxy-aspartic acid in their macrolactone core. The exact hydroxylation process in the context of their NRPS biosynthesis is unclear, however, studies indicate that it is of critical importance for bioactivity.

Therefore, the aim was not only the bioinformatic investigation of the genome of *Lysobacter sp*. BMK333-48F3 to identify the gene cluster of tripropeptin as well as the experimental verification of the plusbacin gene cluster, but also to gain insight into the biosynthetic processes of guanidine-containing cyclic lipopeptides.

The identification of the exact targets of the two dioxygenases encoded in the gene cluster is of increased interest. Noteworthy is that the cluster comprises only two oxygenases, but the plusbacins have three or even four hydroxylated amino acids. To clarify this question, knockout studies of the two oxygenase genes *plbD* and *plbE* shall be performed.

The question of the timing of the hydroxylation also arose. Either the hydroxylation reaction occurs in-line or as a post-assembly modification of the lipopeptide backbone in a tailoring reaction catalyzed by the dioxygenases or the process is regulated by the corresponding A-domain, which selects, activates and integrates only hydroxylated amino acid precursors. For this purpose, the respective A domains have to be heterologously expressed and their specificity should be determined.

Moreover, a very low production of plusbacins by the producer strain was observed. Speculating that the hydroxylation process might be a bottleneck in the biosynthesis of plusbacin, an overexpression of *plbD* and *plbE* in *L. firmicutimachus* PB-6250^T is envisioned.

II. Materials and Methods

1. Materials

1.1. Devices and Equipment

Table 7: Instruments used in this study

Instrument	**Manufacturer**
µCuvette® G1.0	Eppendorf
Cryofreezer	ThermoScientific
Autoclave (Systec VX-150)	Systec
BioPhotometer D30	Eppendorf
Centrifuge (5424 R)	Eppendorf
Centrifuge (Heraeus Multifuge 4KR)	ThermoScientific
Centrifuge (Sorvall RC6 Plus)	Thermo Fisher Scientific
Clean bench (Safe 2020)	ThermoScientific
Electroporator (MicroPulser™)	Bio-Rad
Gel Doc Imaging System XR+	Bio-Rad
Gel electrophoresis apparatus	VWR
HPLC system	Waters
Incubation shaker Multitron Pro	Infors
LC (1100 Series; coupled to MS)	Agilent
Mass spectrometer (MS; QTRAP 3200)	AB Sciex
Microscope (Leica DM750)	Leica Microsystems
Mini-PROTEAN® Tetra Cell System	Bio-Rad
Multichannel pipette (12 ×10 µL)	Capp
Multichannel pipette (12 ×100 µL)	Capp
Multichannel pipette (8 × 100 µL)	Eppendorf
Incubating orbital shaker (falcon tubes)	VWR
pH meter (FiveEasy)	Mettler Toledo
Pipettes (Discovery Comfort: 1000 µL; 100 µL; 10 µL; 2 µL)	HTL Lab Solutions
Purelab™ Flex	Elga Veolia
Hei-VAP Precision rotary evaporator	Heidolph
Unichiller	Huber
Scale (B22002)	Oxford
Ultrasonics™ Sonifier™ SFX250	Branson
Special accuracy weighing scale (BP210D)	Sartorius
Thermo Cycler peqSTAR 96x Universal	VWR
Thermo Mixer	VWR
Vortex-Genie 2	Scientific Industries

Table 8: Consumables used in this study

Material	**Manufacturer**
Cuvettes (1 mL)	Sarstedt
Electroporation cuvettes (0.2 cm gap)	VWR
Eppendorf tubes (20 µL / 200 µL / 1.5 mL / 2.0 mL)	Sarstedt
Falcon tubes (15 mL / 50 mL)	Sarstedt
Filter paper discs	Sigma-Aldrich
Filtropur S 2 µm	Sarstedt
Multiwell plates	Brand
Vivaspin™ protein concentrator spin column	Satorius
Parafilm M	Brand
PCR tubes / PCR 96 well plates	VWR
Petri dishes (92 x 16 mm)	Sarstedt
Pipette tips (1000 µL; 100 µL; 10 µL)	Sarstedt
Strata™-XL 100 µm polymeric reversed phase 2 g/ 12 mL, Giga tubes,	Phenomenex
Syringe InjektTM (10 mL / 20 mL)	B. Braun Melsungen AG

1.2. Substrates and Premixed Solutions

Table 9: List of chemicals used in this study

Chemical	**Manufacturer**
Acetic acid	Sigma-Aldrich
Acetonitrile HPLC grade	VWR
Acrylamide	Bio-Rad
Agar	AppliChem
Agarose basic	AppliChem
Ammonium acetate	AppliChem
Ammonium molybdate tetrahydrate	Sigma-Aldrich
Anhydrotetracycline	Sigma-Aldrich
Apramycin sulfate salt	Sigma-Aldrich
Bacto tryptone	Difco
Bacto yeast extract	Difco
Bromophenol blue	Sigma
Calcium chloride dihydrate	Roth
Carbenicillin disodium salt	Roth
Casaminoacids	Difco
Chloramphenicol	Merck
Chloroform HPLC grade	Merck

Comassie Brilliant Blue G250	Merck
Comassie Brilliant Blue R250	Merck
Copper (II) chloride dihydrate	Sigma-Aldrich
D-(+)-glucose	Sigma-Aldrich
D-(+)-maltose monohydrate pure	AppliChem
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich
dNTP Mix 10 mM (2.5 mM of each dNTP)	Meridian Bioscience
Ethanol	Sigma-Aldrich
Ethylenediaminetetraacetic acid (EDTA)	Roth
Gentamicin	Sigma-Aldrich
Glucose	Sigma-Aldrich
Glycerol (98 %)	Roth
Hydrochloric acid (HCl)	Sigma-Aldrich
HyperLadder	Meridian Bioscience
Imidazole	Sigma-Aldrich
IPTG	Meridian Bioscience
Kanamycin sulfate	Roth
L-(+)-arabinose	AppliChem
L-Asp	Roth
L-Hya	Abcam
L-Hyp	ABCR
Loading dye (6x)	Thermo Scientific
L-Pro	Roth
L-proline	Sigma-Aldrich
Magnesium chloride hexahydrate	Sigma-Aldrich
Magnesium sulfate heptahydrate	AppliChem
Manganese (II) chloride tetrahydrate	Sigma-Aldrich
Mannitol	Sigma-Aldrich
Methanol HPLC grade	VWR
Methanol LC-MS grade	VWR
Ni-NTA resin	Qiagen
peqGREEN	VWR
peqGreen	Peqlab
Phenol:chloroform:isoamyl alcohol (25:24:1)	AppliChem
Phenylmethanesulfonyl fluoride (PMSF)	Sigma-Aldrich
Polypeptone	Roth
Potassium acetate	AppliChem
Potassium chloride	Sigma-Aldrich
Potassium dihydrogenphosphate	Sigma-Aldrich
Potassium sulfate	Roth
Propan-2-ol	Sigma-Aldrich
R2A agar	Fluka

Silica gel 60 H SiO2	Merck
Sodium acetate trihydrate	Merck
Sodium borate	Sigma-Aldrich
Sodium chloride	Sigma-Aldrich
Sodium dodecyl sulfate (SDS)	AppliChem
Sodium hydroxide	Sigma-Aldrich
Sucrose	Sigma-Aldrich
Tetracycline	Sigma-Aldrich
Tetramethyl ethylendiamine (TEMED)	Fluka
Thiostrepton	Sigma-Aldrich
Trifluoroacetic acid (TFA)	Sigma-Aldrich
Tris (Tris(hydroxymethyl)-aminomethan)	Roth
Tris(hydroxymethyl)-methylamino-ethanesulfonic acid (TES)	Roth
Trizma-Base	Sigma
Tryptic soy broth (TSB)	Difco
Zinc chloride	Sigma-Aldrich
β-Mercaptoethanol	Bio-Rad

1.3. Kits and Enzymes

Table 10: Kits and Enzymes used in this study

Name	**Supplier**
Cre Recombinase	New England BioLabs (NEB)
Genomic-tip 100/G	Qiagen
HiYield® Genomic DNA Mini Kit	Süd-Laborbedarf SLG
innuPREP DOUBLEpure Kit	Analytik Jena
Lysozyme	AppliChem
Monarch® DNA Gel Extraction Kit	New England BioLabs (NEB)
Pfu DNA Polymerase	New England Biolabs
Proteinase K	Meridian Bioscience
Q5 High-Fidelity Polymerase	New England BioLabs (NEB)
QIAGEN Genomic DNA Preparation Kit	Qiagen
QIAGEN Plasmid Plus Midi Kit	Qiagen
Restriction endonucleases	New England BioLabs (NEB)
RNase A 100 mg/ml	Qiagen
Taq DNA Polymerase	New England BioLabs (NEB)

1.4. Antibiotics

Antibiotic stock solutions were prepared according to supplier recommendations, filter sterilized and stored at -20°C. When required, antibiotics were added under a clean bench to their final concentration. Prior to adding the antibiotic, it was ensured that the medium was not warmer than 50°C.

Table 11: Antibiotics used in this study

Antibiotic	Stock Concentration	Working Concentration
Apramycin ($Apra^R$)	50 mg/mL	25 µg/mL
Chloramphenicol (Cm^R)	25 mg/mL	12.5 µg/mL
Gentamicin (Gm^R)	50 mg/mL	50 µg/mL
Kanamycin (Kan^R)	50 mg/mL	50 µg/mL
Tetracycline (Tet^R)	25 mg/mL	12.5-25 µg/mL

1.5. Bacterial Strains

Table 12: Bacterials strains used in this study

Name	**Description**	**Source or reference**
E. coli DH10B	*F- endA1 recA1 galE15 galK16 nupG rpsL ΔlacX74 Φ80lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) λ-*	Invitrogen
E. coli DH5α	*F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK–mK+), λ-*	Thermo Fisher Scientific
E. coli NEB5α	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	NEB
E. coli XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB laclqZΔM15* Tn*10*]; TcR	Stratagene

E. coli BW25113/pIJ790	K-12 Derivat: *ΔaraBAD, ΔrhaBAD* Kan^R recombination plasmid: λ/RED (*gam, bet, exo*), *araC, rep101ts* Cm^R	(Datsenko and Wanner 2000)[101] (Gust et al. 2003)[102]
E. coli BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS* *λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	NEB
B. subtilis 168	*trpC2*	Laboratory Stock
L. firmicutimachus PB-6250	*plb*	IPOD

1.6. Plasmids and Fosmids

Table 13: Plasmids and Fosmids used in this study

Plasmids		
Name	**Description**	**Source or reference**
pIJ774	pBS SK(+)-derivative, P1-loxP-oriT-*aac(3)IV*-loxP-P2	(Gust et al. 2003)[102]
pGEM®-T	Linearized cloning vector with T-overhangs for direct cloning of PCR- products with A-overhangs, *lacZ'*, ori, f1-origin, Carb^R	Promega
pJQ200SK	Suicide vector for gene replacement Gm^R, p15A, oriT, *sacB*	Addgene (Quand et al. 1993)[103]
pBBR1-MCS5	derivative of the broad-host-range (bhr) cloning vector pBBR1MCS	(Kovach et al. 1995)[104]

pET-28 a (+)	cloning and expression of recombinant proteins in *E. coli* 6XHis$_6$-Tag, *lac*I, Kan^R, thrombin site	Novagen
pJQ200SK/∆plbNRPS	PCR products from gDNA of PB-6250^T comprising regions up- and downstream of the disruption site of plbNRPS within pJQ200SK via Gibson Assembly, Gm^R	This study
pJQ200SK/∆*plbD*	PCR products from pCC1Fos/∆*plbD* comprising regions up- and downstream of ∆*plbD* within pJQ200SK via Gibson Assembly, Gm^R	This study
pJQ200SK/∆*plbE*	PCR products from pCC1Fos/∆plbE comprising regions up- and downstream of ∆plbE within pJQ200SK via Gibson Assembly, Gm^R	This study
pBBR1-MCS5-*plbD*	*plbD* gene flanked by *Eco*RI/*Hind*III in pBBR1-MCS5	This study
pBBR1-MCS5-*plbE*	*plbE* gene flanked by *Eco*RI/*Hind*III in pBBR1-MCS5	This study
pET-28 a (+)-A$_3$	plbA$_3$ flanked by *Eco*RI/*Hind*III in pET-28a(+)	This study
pET-28 a (+)-A$_5$	plbA$_3$ flanked by *Eco*RI/*Hind*III in pET-28a(+)	This study
pET-28 a (+)-A$_7$	plbA$_3$ flanked by *Eco*RI/*Hind*III in pET-28a(+)	This study

pET-28 a (+)-A_8	plbA_3 flanked by *Eco*RI/*Hind*III in pET-28a(+)	This study
pET-28 a (+)-AT_7	plbAT_7 flanked by *Eco*RI/*Hind*III in pET-28a(+)	This study
Fosmids		
pCC1Fos/plbOx	pCC1FOS based fosmid containing the partial plb gene cluster including *plbD* und *plbE*, Cm^R	Henrike Miess, PhD thesis
pCC1Fos/Δ*plbD*_apra	pCC1Fos/plbOx derivative, *plbD* replaced with P1-loxP-*aac(3)IV*-loxP-P2; $Apra^R$, Cm^R	This study
pCC1Fos/Δ*plbE*_apra	pCC1Fos/plbOx derivative, *plbE* replaced with P1-loxP-*aac(3)IV*-loxP-P2; $Apra^R$, Cm^R	This study
pCC1Fos/Δ*plbD*	pCC1Fos/plbOx derivative, Δ*plbD*, Cm^R	This study
pCC1Fos/Δ*plbE*	pCC1Fos/plbOx derivative, Δ*plbE*, Cm^R	This study

1.7. Oligonucleotides

Oligonucleotide primers were synthesized from either Eurofins Genomics or Integrated DNA Technologies IDT.

Table 14: Oligonucleotides used in this study

Oligonucleotides used for cloning	
Primer	Sequence
koGA plbNRPS3 up_fwd	CGAATTCCTGCAGCCCGGGGGTCGCGGTTTGCCTGGAAC
koGA plbNRPS3 up_rev	TCTTCAGCGCCTGGGTCACCAACAGCACGTGATC
koGA plbNRPS3 down_fwd	ACGTGCTGTTGGTGACCCAGGCGCTGAAGATCG

koGA plbNRPS3 down_rev	CCACCGCGGTGGCGGCCGCTGCCGACGGCAGCATGTATTC
plbD_fwd	GATCGTCGGCCGCCATTGGTATCTGGGACGATGAGCATGATT CCGGGGATCCGTCGACC
plbD_rev	CGCGTCGTTTGGAAAGTGATGGTCGGAGAATCGGGCTCATGT AGGCTGGAGCTGCTTCG
plbE_fwd	TCCGACCATCACTTTCCAAACGACGCGAACACGAACATGATTC CGGGGATCCGTCGACC
plbE_rev	AATTCTCGTGCCGTGAAGTGTGGTGGTTGGGGCGACTCATGT AGGCTGGAGCTGCTTCG
koGA plbD_fwd	CGAATTCCTGCAGCCCGGGGTCGCTTATGTGGTCGGCCTC
koGA plbD_rev	CCACCGCGGTGGCGGCCGCTTGACCTTGGACGGCCCGG
koGA plbE_fwd	CGAATTCCTGCAGCCCGGGGCCGCCCATCGCTACTACGCTC
koGA plbE_rev	CCACCGCGGTGGCGGCCGCTGCGGCTGGCCAAGTCGGC
plbA3 fwd	TGGGTCGCGGATCCGAATTCCAGGACATCTATCCGTTGGC
plbA3 rev	TCGAGTGCGGCCGCAAGCTTCACCGTGGCGGCCAGTTCCG
plbA5 fwd	TGGGTCGCGGATCCGAATTCCCGTTGCCGTTGTCGCATTC
plbA5 rev	TCGAGTGCGGCCGCAAGCTTCGTGGCGGCGGCGAGGTCGG
plbA7 fwd	TGGGTCGCGGATCCGAATTCCAGGACATCTATCCGCTGGC
plbA7 rev	TCGAGTGCGGCCGCAAGCTTCACGTCCCGGGCCAACTCGC
plbA8 fwd	TGGGTCGCGGATCCGAATTCTCCAAGCCGCTGCCGATGTC
plbA8 rev	TCGAGTGCGGCCGCAAGCTTGATCAGGGCCGCCAGCGCGC
Oligonucleotides used for sequencing	
Primer	Sequence
Screening plbNRPS fwd	ACGACGATCACGTGCTGTTG
Screening plbNRPS rev	ACCTGGAACACCGGGTGATG
Screening plbD fwd	GCGTCGTTTGGAAAGTGATG
Screening plbD rev	GAGTATCTTGCGCTGATCGG
Screening plbE fwd	GTAGGCGTTCTTGACCTTGC
Screening plbE rev	GCGTCGTTTGGAAAGTGATG

1.8. Media

Unless otherwise stated, all media were prepared with double distilled water and autoclaved at 120°C and 2 bar for 20 min. If required, sterile supplementary components were added at the time of use. Media were stored at room temperature or 4°C.

LB (Luria Bertani) Medium Liquid / Agar

Bacto tryptone	10 g
Bacto yeast extract	5 g
NaCl	10 g
ddH_2O	ad 1000 mL
For LB agar: Agar	20 g

SOB (Super Optimal Broth)

Bacto tryptone	20 g
Bacto yeast extract	5 g
NaCl	0.5 g
KCl	0.186 g
ddH_2O	ad 1000 mL

The pH was adjusted to pH 7 with 1M NaOH before sterilization

SOC (Super Optimal Broth with Catabolite Repression)

Bacto trypton	20 g
Bacto yeast extract	5 g
NaCl	0.5 g
KCl	0.186 g
ddH_2O	ad 980 mL
1 M Glucose (sterile)	20 mL

The pH was adjusted to pH 7 with 1M NaOH before sterilization. After autoclaving, 20 mL of sterile 1 M glucose solution was added.

2X YT

Bacto tryptone	16 g
Bacto yeast extract	10 g
NaCl	5 g
ddH_2O	ad 1000 mL
For 2XTY agar: Agar	20 g

R2A

Bacto yeast extract	0.5 g
Proteose peptone No. 3	0.5 g
Casamino acids	0.5 g
Dextrose	0.5 g
Soluble starch	0.5 g
Sodium pyruvate	0.3 g
KH_2PO_4	0.3 g
$MgSO_4$ x $7H_2O$	50 mg
ddH_2O	ad 1000 mL
For R2A agar: Agar	20 g

1.9. Buffers and Solutions

1M Glucose

Glucose powder	180.156 g
ddH_2O	ad 1000 mL

To sterilize, the solution was filtered using a 0.2 µm filter.

Tris-HCl

Tris base	121.1 g
HCl (9 %)	Adjust the pH with the necessary volume.
ddH_2O	ad 1000 mL

0.5 M EDTA

EDTA disodium salt, dihydrate	186.1 g
NaOH pellets/NaOH 1N	20g / 150 ml
ddH_2O	ad 1000 ml

The solution was stirred on a magnetic stirrer until dissolved completely and the pH was adjusted to 8.0.

1 M NaOH

NaOH	39,997 g
ddH_2O	ad 1000 mL

1M HCl

HCl 38 %	236.84 ml
ddH_2O	ad 1000 mL

5 M Potassium acetate

Potassium acetate (KOAc)	490.75 g
ddH_2O	ad 1000 mL

Blue bromophenol (Loading Dye)

Glycerol 99 %	4.29 mL
0.5 M EDTA (pH 7.5)	1 mL
Blue bromophenol	5 mg
ddH_2O	ad 9 mL

10 % Glycerol for Electrocompetent Cells

Glycerol	100 mL
ddH_2O	ad 1000 mL

100 mM $CaCl_2$

$CaCl_2$ x 2 H_2O	14.702 g
ddH_2O	ad 1000 mL

100 mM $CaCl_2$ + Glycerol 15 %

$CaCl_2$	1.1098 g
Glycerol (99.5 %)	15 mL
ddH_2O	ad 100 mL

1 M IPTG (Isopropyl-β-D-thiogalactopyranoside)

IPTG	2.3831 g
ddH_2O	ad 10 mL

To sterilize, the solution was filtered using a 0.2 µm filter and stored at -20°C.

50X TAE

Tris base	242 g
Acetic acid	57.1 mL
EDTA / 0.5 M EDTA (pH 8.0)	37.2 g / 100 mL
ddH_2O	ad 1000 mL

Buffer B1 (Bacterial Lysis Buffer)

EDTA	18.61 g
Tris-base	6.06 g
10 % Tween-20 solution	50 mL
10 % Triton X-100 solution	50 mL
ddH_2O	ad 1000 mL

Buffer B2 (Bacterial Lysis Buffer)

Guanidine HCl	286.59 g
100 % Tween-20	200 mL
ddH_2O	ad 1000 mL

Buffer QBT (Equilibration Buffer)

NaCl	43.83 g
MOPS (free acid)	10.46 g
Pure isopropanol	150 mL
10 % Triton X-100 solution	15 mL
ddH_2O	ad 1000 mL

Buffer QC (Washing Buffer)

NaCl	54.44 g
MOPS (free acid)	10.46 g
Pure isopropanol	150 mL
ddH_2O	ad 1000 mL

Buffer QF (Elution Buffer)

NaCl	73.05 g
Tris-base	6.06 g
Pure isopropanol	150 mL
ddH_2O MiliQ	ad 1000 mL

Sol I

Tris-base	6.057 g
EDTA	2.922 g
(alternatively 0.5 M EDTA solution	20 mL)
Glucose	9.008 g
ddH_2O	ad 1000 mL

After autoclaving, RNase A was added to a final concentration of 20 µg/mL.

Sol II

1 M NaOH	200 mL	20 mL
SDS 10 %	100 mL	10 mL
ddH_2O	ad 1000 mL	ad 100 mL

The solution was prepared directly prior to use.

Sol III

5 M Potassium acetate (KOAc)	600 mL
Acetic acid	115 mL
ddH_2O	ad 1000 mL

Lysis Buffer

Tris-HCl	50 mM
NaCl	500 mM
Glycerol	10 % (v/v)
β-Mercaptoethanol	10 mM
TWEEN 20	1 % (v/v)
Imidazole	20 mM

β-Mercaptoethanol was added directly prior to use.

Wash Buffer

Tris-HCl	50 mM
NaCl	500 mM
Glycerol	10 % (v/v)
β-Mercaptoethanol	10 mM
Imidazole	20 mM

β-Mercaptoethanol was added directly prior to use.

Elution Buffer

Tris-HCl	50 mM
NaCl	500 mM
Glycerol	10 % (v/v)
β-Mercaptoethanol	10 mM
Imidazole	250 mM

β-Mercaptoethanol was added directly prior to use.

Laemmli Buffer

Tris-HCl pH 6.8	250 mM
SDS	8 % (w/v)
Glycerin	40 % (v/v)
Blue bromophenol	0.004 % (w/v)
β-Mercaptoethanol	20 % (v/v)
ddH_2O	ad 10 mL

10X SDS Electrophoresis Buffer

Tris-HCl pH 8.3	250 mM
Glycin	1.5 M
SDS	1 %
ddH_2O	ad 1000 mL

Fixation Solution

Acetic acid	10 % (v/v)
Methanol	20 % (v/v)
ddH_2O	70 %

Dye Solution

Acetic acid	10 % (v/v)
Methanol	45 % (v/v)
Coomassie Brilliant Blue R-250	0.25 % (w/v)
ddH_2O	44.75 %

Destaining Solution

Acetic acid	10 % (v/v)
Methanol	45 % (v/v)
ddH_2O	45 %

2. Methods

2.1. *In silico* Analysis of DNA

For genome-wide identification, annotation and analysis of secondary metabolite biosynthesis gene clusters antiSMASH version 4.0, 5.0 and 6.0 was used. To perform alignment analysis and to determine the percentages of homology and identities between genes and proteins, Basic Local Alignment and Search Tool (BLAST)[105] as well as MultiGeneBlast and Artemis Comparison Tool was utilized. To plan cloning experiments, design primers and oligonucleotides, analyze and visualize DNA, software like SnapGene and Clone Manager was used.

2.2. Microbiology

2.2.1. Microbial Overnight Cultures

If not otherwise stated, bacterial overnight cultures of about 10 mL were inoculated with a single colony and usually grown overnight (16-20 h) at 30°C or 37°C in the appropriate medium with or without antibiotics. Cultivation in liquid medium was performed in an orbital shaker at 200 rpm.

2.2.2. Bacterial Glycerol Stocks

For long-term storage, glycerol stocks were prepared by thoroughly mixing 1 mL of overnight bacterial culture with 0.5 mL 80 % glycerol (v/v) and stored at -80°C.

To recover the bacteria, cells from the frozen glycerol stock were streaked on appropriate agar plates.

2.2.3. Cultivation of Bacteria

E. coli and *B. subtilis* cultures were usually grown over night (16-20 h) at 37°C in liquid or solid LB-medium. Cultivation in liquid medium was performed while shaking at 200 rpm. Suitable antibiotics were added to the medium.
For A domain production in *E. coli*, the incubation temperature was lowered to 20°C after induction with IPTG.
For recombineering experiments with *E. coli* BW25113, the incubation was performed at 30°C.

L. firmicutimachus was cultured for 1-2 days at 30°C in liquid or solid medium with the appropriate antibiotics or medium supplements. Depending on the experiment, LB medium or R2A was chosen.
For production of secondary metabolites, such as plusbacins or derivatives, a pre-culture of the *L. firmicutimachus* wild type or mutant strain was used to inoculate R2A. The cultivation process was performed at 28°C for 4 days at 200 rpm.

2.2.4. Monitoring Microbial Growth

To determine bacterial growth, optical density was measured against a sterile medium blank reference at the wavelength of 600 nm (OD_{600}) by using the BioPhotometer® D30 (Eppendorf).

2.2.5. Disc Diffusion Assay

Liquified LB agar was inoculated with the indicator strain *B. subtilis* 168 and poured into a petri dish. After solidifying, sterile filter discs were placed onto the surface of the agar plate and 10 µL of the test substance, or corresponding solvent as a negative control, was pipetted to the discs. The disc diffusion test plate was cultivated overnight at 37°C.

2.2.6. Antagonistic Assay

Overnight cultures of *Lysobacter* sp. were used to streak the bacteria on R2A agar plates in a line (thickness 0.5 – 1 cm). After 1-2 days of incubation at 30°C, *B. subtilis* 168 was transferred to the agar plate turned 90 degrees compared to

the first streaking. To evaluate the antibiotic activity of *Lysobacter* sp. the agar plate was cultured overnight at 30°C and checked for growth inhibition of *B. subtilis* 168.

2.3. Molecular Biology

2.3.1. DNA Sequencing

The nucleotide sequence of recombinant vectors and PCR products was determined using Sanger sequencing by Eurofins Genomics / GATC Services (Ebersberg). For all custom sequencing services, cycle sequencing technology (dideoxy chain termination / cycle sequencing) on ABI 3730XL sequencers was used.

2.3.2. Genomic DNA Isolation

2.3.2.1. Mini Preparation of Genomic DNA

For the isolation of small amounts of genomic DNA (gDNA) to use in PCR or other enzymatic reactions, the manufacturer's protocol of the Hi Yield® Genomic DNA Mini Kit for Bacteria was followed. To determine the concentration of the isolated gDNA, optical density measurements at a wavelength of 260 nm (OD_{260}) using a BioPhotometer® D30 (Eppendorf) was performed. gDNA was stored at – 20°C for later use.

2.3.2.2. Maxi Preparation of Genomic DNA

To isolate larger amounts of high quality gDNA, 10-20 mL bacterial overnight culture was processed using a QIAGEN® Genomic DNA Preparation Kit in combination with Genomic-tip 100/G, following the manufacturer's protocol. Minor changes to the protocol were made. In brief, after harvesting the cells, the bacterial pellet was resuspended in 7 mL buffer B1 containing 14 µL RNase A (100 mg/mL). After vortexing, 160 µL of lysozyme (100 mg/mL) and 150 µL of QIAGEN proteinase K were added. The suspension was incubated for 1 h at 37°C and mixed with 2.4 mL of buffer B2 by inverting for several times. The sample was incubated for 1 h at 50°C until achieving a clear lysate. After vortexing for 10 sec, the sample was loaded on a previously with 4 mL buffer QBT equilibrated Genomic-tip 100/G. The DNA was allowed to drip into the column resin by gravitational flow. The Genomic-tip 100/G was washed twice with 7.5 mL buffer QC followed by the elution of gDNA with 5 mL pre-warmed buffer QF. To precipitate the gDNA, 3.5 mL of

isopropanol (RT) was added to the eluate and inverted carefully about 20 times. The gDNA was recovered by spooling with a glass rod, then washed in 2 mL 70% ice-cold ethanol and dried. Subsequently the purified DNA was transferred to a screw-cap microcentrifuge tube containing a suitable amount of 10 mM Tris-HCl (pH 8.0) and dissolved completely by incubating at 37°C overnight or for 1-2 h at 55°C. To determine the concentration and quality of the isolated gDNA, optical density measurements at a wavelength of 260 nm (OD_{260}) using a BioPhotometer® D30 (Eppendorf) and agarose gel electrophoresis by running a 1.5 % agarose gel for 16 h at 25 V were performed. gDNA was stored at –20°C for later use.

2.3.3. Isolation of Plasmid or Fosmid DNA

To gain high quality and purity plasmid or fosmid DNA, a QIAGEN Plasmid Plus Midi Kit was applied according to the manufacturer's protocol. Standard isolation of plasmid or fosmid DNA was performed by alkaline lysis and DNA precipitation.[106]

2.3.3.1. Mini Preparation

1.5 mL of an overnight liquid culture was centrifuged (15000 rpm, 1 min, RT) and the resulting pellet was resuspended in 100 µL Sol I containing RNase A. To lyse the cells, 200 µL of freshly prepared Sol II was added. Proteins and genomic DNA were precipitated by addition of 150 µL Sol III and inverting the tube several times. After centrifugation (15000 rpm, 5 min, 4°C) the plasmid/fosmid DNA contained in the supernatant was purified by phenol-chloroform extraction and concentrated by alcoholic precipitation. For phenol-chloroform extraction, the supernatant was mixed with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) by vortexing and centrifuged (15000 rpm, 20 min, 4°C). The aqueous phase was then transferred to a new tube, mixed with the same volume of ice-cold 100 % isopropanol, incubated on ice for about 10 min and centrifuged (15000 rpm, 30 min, 4°C). The precipitated DNA was washed with 500 µL 70 % ethanol (15000 rpm, 10 min, 4°C), dried at 50°C, and finally dissolved in 30-50 µL 10 mM Tris-HCl buffer (pH8). For long-term use, the isolated DNA was stored at -20°C.

2.3.3.2. Maxi Preparation

50 mL of an overnight culture was harvested by centrifugation (4400 rpm, 15 min, RT). The supernatant was discarded while the pellet was thoroughly resuspended in 5 mL Sol I containing RNase A by vortexing, followed by an incubation step at RT for about 10 min. 10 mL of freshly prepared Sol II was added and mixed gently by inverting the tube several times. After a 10 min incubation on ice, 7.5 mL of ice-cold Sol III was gently mixed with the sample by inverting the tube for several times. Another 10 min incubation step on ice, was followed by centrifugation at the maximum speed for 10 min at 4°C. Equal volume isopropanol was added to the supernatant and incubated on ice for 10 min to precipitate the DNA and centrifuged (4400 rpm, 30 min, 4°C). Prior to dissolving the pellet in 700 µl 10 mM Tris-HCl (pH 8.0), the DNA pellet was washed with 1 mL of 70 % ethanol, air-dried for 5 min at RT. In order to get pure DNA, the sample was mixed with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) by vortexing and centrifuged (15000 rpm, 20 min, 4°C). The upper phase was transferred to a new tube and precipitated again by mixing with 70 µl of 3 M sodium acetate (pH 5.2) and 1.8 mL of 100 % ethanol by inverting several times. After centrifugation (4400 rpm, 30 min, 4°C), the precipitated DNA was washed with 500 µL 70 % ethanol (4400 rpm, 10 min, 4°C) and dried for several minutes at 50°C. Finally, the pellet was dissolved in 300 µL 10 mM Tris-HCl buffer (pH8) and stored at -20°C.

2.3.4. Purification of DNA Solutions

For purification and concentration of DNA, alcoholic precipitation and phenol-chloroform extraction were performed or an ion exchange column (innuPREP DOUBLEpure Kit) was used. Purification using the ion exchange columns was performed according to the manufacturer's instructions.

2.3.5. Determination of DNA Concentration

For concentration determination of DNA solutions, µCuvette® G1.0 (Eppendorf) for the BioPhotometer® D30 (Eppendorf) was used and absorbance was measured at 260 nm. Likewise, concentration determination by agarose gel electrophoresis was performed. Here, the fluorescence intensities of the DNA samples examined were compared with the DNA marker (HyperLadders, Meridian Bioscience).

2.3.6. Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) is an enzyme-dependent method to amplify specific gene sequences within an existing DNA molecule and was used for screening experiments or for the amplification of certain genes and DNA fragments. According to the type of the experiment, several varying programs and polymerases were applied. For general purposes like regular screening experiments and to determine the optimum annealing temperatures of primers, either a mixture of lab-made Taq:Pfu (1:1) or a commercial Taq polymerase (NEB) was used. To generate PCR fragments for cloning experiments, the DNA was amplified with Q5 High-fidelity polymerase (NEB). All primers used for the experiments (Table 14) were purchased from Eurofins Genomics or Integrated DNA Technologies (IDT). All PCR reactions were conducted by using a peqSTAR 96X Thermocycler (PeqLab, Darmstadt) and different PCR programs described below.

Following the PCR, the amplicons were analyzed by gel electrophoresis for the correct fragment size, and, if required for further experiments, purified.

2.3.6.1. General PCR Conditions and Gradient PCR

To determine the optimal annealing temperature for each primer pair, a gradient PCR was performed according to the following pipetting scheme (Table 15) and method specifications (Table 16). The so specified annealing temperature was used for later amplification and screening purposes. The elongation time was set to 60 sec for each 1 kb of product size.

Table 15: Standard PCR reaction master mix

Ingredient	25 µL reaction	50 µL reaction	Final Concentration
10X TP buffer	2.5 µL	5.0 µL	1X
10 mM dNTP mix	0.75 µL	1.5 µL	300 µM
Primer fwd (10 µM)	0.5 µL	1.0 µL	0.05-1 µM
Primer rev (10 µM)	0.5 µL	1.0 µL	0.05-1 µM
Template DNA	0.5 µL	1.0 µL	Less than 500 ng
DMSO	1.25 µL	2.5 µL	5 %
Taq:Pfu polymerase	0.5 µL	1.0 µL	-
ddH_2O	up to 25 µL	up to 50 µL	-

Table 16: Standard amplification conditions

Step	Temperature	Time	Cycles
Initial denaturation	95°C	5 min	1
Denaturation	95°C	15 sec	
Annealing	50-70°C	30 sec	30
Elongation	72°C	1 min	
Final elongation	72°C	5 min	1
Cooling	4°C	∞	

2.3.6.2. High-Fidelity PCR

High-fidelity PCR was performed according to the following pipetting scheme (Table 17) and method specifications (Table 18) for purposes requiring high accuracy during DNA amplification such as cloning or sequencing. The elongation time was set to 30 sec for each 1 kb of product size.

Table 17: PCR reaction master mix using Q5 polymerase

Ingredient	25 µL reaction	50 µL reaction	Final Concentration
5X Q5 Reaction Buffer	5 µL	10 µL	1X
10 mM dNTPs	0.5 µL	1 µL	200 µM
Primer fwd (10 µM)	1.25 µL	2.5 µL	0.5 µM
Primer rev (10 µM)	1.25 µL	2.5 µL	0.5 µM
Template DNA	variable	variable	< 1,000 ng
Q5 High-Fidelity DNA Polymerase	0.25 µL	0.5 µL	0.02 U/µl
5X Q5 High GC Enhancer (optional)	(5 µL)	(10 µL)	(1X)
Nuclease-Free Water	to 25 µL	to 50 µL	

Table 18: Amplification conditions using Q5 polymerase

Step	Temperature	Time	Cycles
Initial denaturation	98°C	30 sec	1
Denaturation	98°C	10 sec	
Annealing	50-72°C	30 sec	30
Elongation	72°C	30 sec	
Final elongation	72°C	2 min	1
Cooling	4°C	∞	

2.3.6.3. Colony PCR

Colony-PCR was used as a rapid, high throughput screening approach, carried out under the following conditions (Table 20) to determine the success of DNA transformation and crossover events based on homologous recombination without isolating DNA. This method allowed quick identification of successfully manipulated bacteria clones since it does not require DNA purification. However, it lacks accuracy and was therefore followed by a constant PCR to confirm the positive results.

For master mix preparation (Table 19), single colonies were picked as template using a sterile toothpick and transferred into 50 µL of autoclaved water.

Table 19: Colony PCR reaction master mix

Ingredient	25 µL reaction	50 µL reaction	Final Concentration
10X TP buffer	2.5 µL	5.0 µL	1X
10 mM dNTP mix	0.75 µL	1.5 µL	300 µM
Primer fwd (10 µM)	0.5 µL	1.0 µL	0.05-1 µM
Primer rev (10 µM)	0.5 µL	1.0 µL	0.05-1 µM
Bacteria suspension	1 µL	2 µL	Less than 500 ng
DMSO	1.25 µL	2.5 µL	5 %
Taq:Pfu polymerase	0.5 µL	1.0 µL	-
ddH_2O	up to 25 µL	up to 50 µL	-

Table 20: Colony PCR amplification conditions

Step	Temperature	Time	Cycles
Initial denaturation	95°C	10 min	1
Denaturation	95°C	15 sec	
Annealing	50-70°C	30 sec	30
Elongation	72°C	1 min	
Final elongation	72°C	5 min	1
Cooling	4°C	∞	

2.3.6.4. Amplification of Resistance Cassettes from pIJ774

To obtain the resistance cassettes from pIJ774 for PCR targeting, 10 µg of the aforementioned plasmid was digested with *Eco*RI and *Hind*III. The resistance cassette was then isolated and purified using preparative agarose gel electrophoresis. Extension of the resistance cassettes was performed under the following conditions (Table 22) using special overhanging primers to prepare the PCR master mix (Table 21).

Table 21: PCR reaction master mix for elongation of DNA fragments for recombineering

Ingredient	50 µL reaction	Final Concentration
5X Q5 Reaction Buffer	10.0 µL	1X
10 mM dNTP mix	4 µL	200 µM
Primer fwd (100 µM)	0.5 µL	50 µM
Primer rev (100 µM)	0.5 µL	50 µM
Template DNA	2 µL	100 ng
5X High GC Enhancer	10.0 µL	1X
Q5 polymerase	0.5 µL	1 Unit
ddH_2O	up to 50 µL	-

Table 22: Amplification conditions for elongation of DNA fragments for recombineering

Step	Temperature	Time	Cycles
Initial denaturation	94°C	30 sec	1
Denaturation	94°C	10 sec	
Annealing	50°C	30 sec	10
Elongation	72°C	60 sec	
Denaturation	94°C	10 sec	
Annealing	55°C	30 sec	20
Elongation	72°C	60 sec	
Final elongation	72°C	2 min	1
Cooling	4°C	∞	

2.3.7. DNA Restriction Digest

Restriction of DNA was performed using appropriate restriction endonucleases (New England Biolabs NEB) in accordance with the manufacturer's instructions. Regular control restriction reactions were incubated for 1 h at 37°C and analyzed by agarose gel electrophoresis. Genomic DNA, fosmid clones and vectors that were used in cloning experiments were incubated from 1-4 h or overnight at 37°C followed by a precipitation step. The concentration of the DNA solution after precipitation was quantified.

2.3.8. Gibson Assembly

The exonuclease-based method allows the assembly of multiple DNA fragments of various length in a desired order due to homologous regions. The reaction is performed in a one-tube reaction under isothermal conditions. For successful Gibson Assembly three enzymatic activities are needed: a 5'-exonuclease generates overhangs, a polymerase fills the gaps within the annealed single-stranded regions, and a DNA ligase seals the nick of filled gaps.[107] The generated recombinant DNA can be used for various applications.

Gibson Assembly was performed as described by Gibson et al[107] with minor changes. The suggested concentration of each DNA fragment was mixed with NEBuilder® HiFi DNA Assembly Master Mix (NEB). The mixture of a final volume of 20 µL was incubated at 45°C for 90 minutes and 2 µL of the resulting solution was used for subsequent transformation of *E. coli* cells.

2.3.9. Agarose Gel Electrophoresis

Agarose gel electrophoresis was performed either in analytical or preparative scale for isolated gDNA and plasmids, restriction digestions and all PCR products, with 0.5-2 % (w/v) agarose gels to separate DNA fragments (size range: 0.2 to 10 kbp) or for concentration determination.

The appropriate amount of agarose powder was completely dissolved in 1X TAE buffer by heating the solution. After cooling, peqGREEN dye was added directly to the hand-warm agarose solution according to the manufacturer's instructions. The agarose gel was prepared with a comb compatible with the number of samples being analyzed and placed in an electrophoresis chamber (PerfectBlue™, Peqlab). After solidification, the gel was overlayed with 1X TAE buffer.

The samples were prepared by mixing with 5X bromophenol blue loading dye and loaded on the gel. DNA separation was achieved by running the gel for 1 h at 120 V, for 2-3h at 90 V or overnight at 25 V (VWR® Power Source) and visualized under UV light (312 nm) using the GelDoc XR+ system (Bio-Rad). A suitable DNA ladder (HyperLadders, Meridian Bioscience) was used for sizing and quantification purposes.

2.3.10. DNA Extraction from Agarose Gel

For isolation and purification of a specific DNA fragment, a preparative agarose gel electrophoresis using a 1 % agarose gel was performed. The required DNA fragment was excised with a sterile scalpel and extracted from the gel either with an innuPREP DOUBLEpure or a Monarch® DNA Gel Extraction kit by following the manufacturer's instructions. The concentration of the purified DNA was determined.

2.3.11. Preparation of Competent Cells from different *E. coli* Strains

2.3.11.1. Electroporation

100 mL of LB medium was inoculated with 1 mL of an overnight *E. coli* culture and incubated at 37°C and 200 rpm to an OD_{600} of 0.5-0.7. The following steps were performed on ice. Cells were then harvested by centrifugation (4400 rpm, 10 min, 4°C) and washed twice with 40 mL of ice-cold 10 % glycerol solution (v/v). The cell pellet was resuspended in about 500 µL of ice-cold 10 % glycerol solution (v/v) and stored in 50 µL aliquots at -80°C or used instantly.

2.3.11.2. Chemically Competent Cells

100 mL of LB medium was inoculated with 1 mL of an overnight *E. coli* culture and incubated at 37°C and 200 rpm to an OD_{600} of 0.5-0.7. The following steps were performed on ice. Cells were harvested by centrifugation (4400 rpm, 10 min, 4°C) and washed with 20 mL of ice-cold 0.1 M $CaCl_2$ solution. The pellet was resuspended in 20 mL ice-cold 0.1 M $CaCl_2$ solution followed by a 20 min incubation step on ice. After centrifugation (4400 rpm, 20 min, 4°C), the cells were resuspended in 2 mL ice-cold 0.1 M $CaCl_2$ containing 15 % glycerol (v/v) and aliquoted to 100 µL in sterile screw-cap microcentrifuge tubes. The competent cells were either used directly or stored at -80°C.

2.3.12. Transformation into *E. coli*

2.3.12.1. Electroporation

50 µL of electrocompetent cells were mixed on ice with 1-5 µL (approximately 100 ng) of DNA and transferred to a sterile ice-cold electroporation cuvette (0.2 cm gap). Electroporation was performed at 2.5 kV/25 µF/200-400 Ω using an electroporation apparatus (MicroPulser™, Bio-Rad). The cuvette was immediately

placed on ice after electroporation, and 1 mL of chilled LB or SOC medium was added. The transformation mixture was transferred to a 1.5 mL reaction tube and incubated for 1 h at 37°C and 200 rpm. The transformed cells were centrifuged, resuspended in approximately 100 µL of medium, and spread on pre-warmed LB plates containing the appropriate selection antibiotic. The agar plates were incubated overnight at 37°C.

2.3.12.2. Heat Shock Transformation

100 µL of chemically competent cells were gently mixed with 1-5 µL (approximately 100 ng) DNA. After incubation on ice for 30 min, the heat shock was performed at 42°C from 30-60 sec depending on the transferred DNA fragment size. The cells/DNA mixture was incubated on ice for another 5 min. 900 µL LB or SOC medium was added to the transformed cells and incubated for 1h at 37°C and 200 rpm. The mixture was centrifuged and the cells were resuspended in about 100 µL medium before streaking on pre-warmed selection plates and incubation overnight at 37°C.

2.3.13. Preparation of Competent Cells from *L. firmicutimachus* PB-6250^T

100 mL of LB medium was inoculated with 2-3 mL of an overnight *L. firmicutimachus* PB-6250^T culture and incubated at 30°C and 200 rpm to an OD_{600} of 0.5-0.7. The following steps were performed on ice. Cells were then harvested by centrifugation (4400 rpm, 10 min, 4°C) and washed twice with 40 mL of ice-cold 10 % glycerol solution (v/v). The cell pellet was resuspended in about 500 µL of ice-cold 10 % glycerol solution (v/v) and stored in 50 µL aliquots at -80°C or used immediately.

2.3.14. Transformation into *L. firmicutimachus* PB-6250^T

50 µL of electrocompetent cells were mixed on ice with 1-5 µL (approximately 100 ng) of DNA and transferred to a sterile ice-cold electroporation cuvette (0.2 cm gap). Electroporation was performed at 2.5 kV/ 25 µF/ 200-400 Ω using an electroporation apparatus (MicroPulser™, Bio-Rad). The cuvette was immediately placed on ice after electroporation, and 1 mL of chilled LB or R2A medium was added. The transformation mixture was transferred to a 1.5 mL reaction tube and incubated for 3-4 h at 30°C and 200 rpm. The transformed cells were centrifuged,

resuspended in approximately 100 µL of medium, and spread on pre-warmed LB or R2A plates containing the appropriate antibiotic used for selection. The agar plates were incubated overnight at 30°C.

2.3.15. Red/ET-mediated Recombination in *E. coli*

Gene knockouts were achieved by PCR-targeted gene replacement.[102, 108] Red/ET-mediated recombination was conducted to insert linear DNA fragments into the fosmid. The required DNA fragments like resistance cassettes were generated by PCR using primers with 39-nt extensions homologous to the respective regions up- and downstream of the gene of interest.

To prepare for Red/ET-mediated recombination, *E. coli* BW25113/pIJ790 cells were transformed with pCC1Fos/plbOx. 10 mL SOB medium containing the required antibiotics was inoculated with the pre-culture of the resulting *E. coli* BW25113 strain. For induction of the λ-Red genes, 100 µL of a 1 M arabinose solution was pipetted to the culture and incubated at 30°C and 200 rpm. After reaching an OD_{600} of 0.5, cells were harvested (5 min, 3000 rpm, 4°C), washed two times with 10 mL of ice cold 10 % glycerol and resuspended in 100 µL ice cold 10 % glycerol. The prepared cells were then transformed with 100 ng of the appropriate DNA fragment. Positive clones were selected using suitable antibiotic selection plates and incubated at 37°C. For differentiation between wild-type and mutated DNA, isolated fosmid DNA was retransformed in *E. coli* XL1-Blue. The reisolated fosmids were screened for successful recombination by restriction analysis or colony PCR and sequencing. After verification, the floxed resistance cassette based on pIJ774 was excised via an *in vitro* reaction with a Cre-recombinase (NEB). The reaction was prepared as described in the manufacurer's protocol containing 1U Cre-recombinase, 1x reaction buffer and about 100 ng DNA in a final volume of 10 µL. After incubation (30 min, 37°C), the mixture was introduced into *E. coli* XL1-Blue cells by transformation. Resulting apramycin sensitive clones were screened for successful loss of the resistance cassette via colony PCR and verified via restriction analysis and sequencing.

2.3.16. Construction of the *plb* Disruption Cassette

To construct the deletion cassette for the NRPS region of the plusbacin gene cluster, homology arms of 2053 bp upstream and 1933 bp downstream were respectively amplified using Q5 polymerase and primers koGA plbNRPS3 up_fwd/rev and koGA plbNRPS3 down_fwd/rev. The knockout cassette was assembled using Gibson Assembly by cloning of homology arms into the *Bam*HI-*Xba*I-digested pJQ200SK[103] backbone. The final construct was verified using the primers Screening plbNRPS fwd and rev for sequencing.

2.3.17. Plasmid construction for In-frame Deletion of *plbD* and *plbE*

Gene knockouts were achieved by PCR-targeted gene replacement.[102] Red/ET-mediated recombination was applied to introduce a linear apramycin resistance cassette derived from pIJ774 into pCC1Fos/plbOx. The required cassettes were amplified by PCR using primers plbD_fwd/rev and plbE_fwd/rev with 39-nt extensions homologous to the respective regions up- and downstream of the gene of interest. Positive colonies containing pCC1Fos/Δ*plbD*_apra or pCC1Fos/Δ*plbE*_apra were selected on antibiotic selection plates and successful recombination was verified using colony PCR using the primer Screening_plbD_fwd/rev and Screening_plbE_fwd/rev. Excision of the resistance marker was performed in an *in vitro* reaction using a Cre-recombinase. Apramycin sensitive clones were screened via colony PCR using the same primer pairs and successful excision of the resistance cassette on the fosmid pCC1Fos/Δ*plbD* or pCC1Fos/Δ*plbE* was verified via sequencing. For the generation of the knockout vectors pJQ200SK/Δ*plbD* and pJQ200SK/Δ*plbE*, the fosmid pCC1Fos/Δ*plbD* and pCC1Fos/Δ*plbE* served as template DNA to amplify the knocked-out regions using the primers koGA plbD_fwd/rev and koGA plbE_fwd/rev, respectively. To assemble the fragment and the *Bam*HI-*Xba*I-digested pJQ200SK backbone, a Gibson Assembly was performed. The final construct was verified using sequencing.

2.3.18. Construction of Double-crossover Mutants

Electro-competent cells of *L. firmicutimachus* in 10 % ice-cold glycerol were prepared, transferred to a sterile 2 mm gapped cuvette and mixed with the

knockout vector (50 µL cells with 500 ng DNA). Following a single pulse 2.5 kV/25 µF/200-400 Ω (MicroPulser™, Bio-Rad), ice-cold LB medium was added to the cell suspension immediately and incubated for 3 h at 30°C without shear force. Cells were plated onto pre-warmed R2A selection plates containing the appropriate amount of antibiotics and cultivated 48 h at 30°C. Gentamicin-resistant single-crossover mutants were screened and counter-selected for double crossover events after growing individual single crossover clones in R2A containing sucrose. The site of double crossover recombination was genetically verified by sequencing.

2.3.19. Vector Design for Overexpression of Dioxygenases

Vectors containing the genes *plbD* and *plbE* were designed using suitable software and ordered from the Gene Synthesis of BioCat (Heidelberg).

2.3.20. Vector Design for Heterologous A Domain Expression

Vectors containing the genetic information for heterologous A domain expression in *E. coli* BL21 (DE3) were designed using suitable software and either ordered from the Gene Synthesis of BioCat (Heidelberg) or acquired by cloning.

For the generation of the expression vectors gDNA of *L. firmicutimachus* PB-6250^T served as template DNA to amplify the A domain in question. To assemble the fragment and the *Hind*III-*Eco*RI-digested pET-28 a (+) backbone a Gibson Assembly was performed. The final construct was verified using sequencing.

2.4. Biochemistry

2.4.1. Denaturing Polyacrylamid Gel Electrophoresis (SDS-PAGE)

SDS-PAGE was used to separate a protein mix in a size-dependent manner with a Mini-PROTEAN® Tetra Cell System (Bio-Rad). The method was performed as reported by Laemmli.[109] SDS gels were prepared directly prior of use as described below (Table 23).

Table 23: Recipe for SDS gel preparation

	Resolving gel			Stacking gel
Ingredient	10 %	12 %	15 %	5 %
ddH2O [mL]	7.9	6.6	4.6	3.4
30 % Acrylamide Mix [mL]	6.7	8.0	10	0.83
Tris (1.5M, pH 8.8) [mL]	5.0	5.0	5	0.63
10 % SDS [µL]	200	200	200	50
10 % APS [µL]	200	200	200	50
TEMED [µL]	8	8	8	5

First, the separating gel was poured by adding SDS, APS and TEMED one after each other in the correct order and immediately casting the mixture between 1.0 mm spacer and short plates locked in a casting holder. To level the surface, the separating gel was overlaid with isopropanol. The stacking gel was made after polymerization of the separating gel by pouring 5 % polyacrylamide solution on top and inserting a comb.

Before usage, the comb was removed and the gel pockets were rinsed with water to remove excess acrylamide. The gel chamber was set up by inserting the SDS gel and flooding with 1X SDS running buffer.

Before loading 10 µL of sample onto the gel, an equal volume of 2X Laemmli buffer was added and denatured for 10 min at 95°C. Additionally, 5 µL Precision Plus ProteinTM Dual Color Standard (Bio-Rad) was used as size control onto the gel. Gel electrophoresis was carried out with a working voltage of 160 V.

To visualize protein bands, the gel was rocked in fixing solution for 5 min, followed by 15 min staining in Coomassie dye (R-250) and bleaching for 2 h in destaining solution or water.

2.4.2. Heterologous Production of A Domains in *E. coli*

Heterologous overproduction and purification of 6×His-tagged proteins from *E. coli* were carried out using expression vector pET-28 a (+) (Novagene) and *E. coli* BL21 (DE3) (NEB) as the recombinant host strain.

The respective *E. coli* BL21 (DE3) strains containing the recombinant vectors for A domain expression were cultivated for small-scale production in 100 mL and for large-scale production in up to 2 L 2X YB broth supplemented with 50 µg/mL kanamycin at 37°C and 250 rpm. After reaching an OD_{600} between 0.6 and 1.2, the temperature was adjusted to 18°C and isopropylgalactoside (IPTG) was added for induction to a final concentration of 0.25 mM. After additional cultivation for 16-20 h the cells were harvested (4000 rpm, 20 min, 4°C) and immediately processed for protein purification.

2.4.3. Purification of Recombinant A Domains

2.4.3.1. Small Scale Approach

The whole purification process occurred on ice.

The cell pellet of 50 mL culture was resuspended in 4 mL lysis buffer containing 0.5 mg/mL lysozyme, 0.5 mM PMSF and 10 mM β-mercaptoethanol. Cells were disrupted by sonification using a Digital Sonifier® Cell Disruptor (Branson, Danbury, CT USA) and the conditions described below.

Amplitude	30 %
Time	10 min
Pulse	5 sec/ 5 sec pause

The lysate was cleared by centrifugation (18000 rpm, 30 min, 4°C) and gently mixed with 1 mL Ni-NTA resin/slurry at 4°C for 30 min. A column was loaded with the mixture and washed twice with 4 mL wash buffer. After elution with 2 mL elution buffer, the protein sample was aliquoted in 4 fractions of 0.5 mL and frozen at -80°C. The A domains were either directly analyzed or stored for later usage.

2.4.3.2. Large Scale Approach

Large scale production of the investigated A domains occurred prior to specificity testing in cooperation with Maximillian Müll and Hajo Kries (Independent Junior Research Group Biosynthetic Design of Natural Products, Leibniz Institute for Natural Product Research and Infection Biology e.V., Hans Knöll Institute (HKI Jena) Jena, Germany).

The protein expression was performed as described by Stanišić et al.[69] with some modifications. The cell pellet was resuspended in 30 mL lysis buffer (50 mM TRIS

[pH 7.4], 500 mM NaCl, 20 mM imidazole, 2 mM TCEP) and 100 µL protease inhibitor mix (Sigma, P8849) was added. Following sonication, the lysate was cleared by centrifugation at 19,000 rpm for 30 min at 4°C. The supernatant was loaded onto a column packed with 2 mL of Ni-IDA suspension (Rotigarose, Roth) after equilibration with lysis buffer. To get rid of impurities, the column was washed twice with 20 mL lysis buffer. The elution step of the target protein was performed using 4 x 0.75 mL elution buffer (50 mM TRIS [pH 7.4], 500 mM NaCl, 300 mM imidazole, 2 mM TCEP). After pooling the protein-containing fractions, the samples were concentrated to a volume of about 1 mL with a Vivaspin20 (Sartorius) centrifugal filter with 10 kDa MWCO, and then washed with two times 5 mL storage buffer (50 mM Tris, 200 mM NaCl, pH 7.4). Protein concentration was adjusted to 50 µM by diluting in storage buffer and glycerol was added to a final concentration of 10 %. Aliquots of 100 µL were prepared and directly used for A domain specificity testing or flash frozen in liquid nitrogen for storage at -80°C. To specify protein concentrations, the absorbance at 280 nm was measured in Take3 plates on an Epoch2 microplate reader (Biotek) using calculated extinction coefficients. SDS-PAGE was performed to monitor A domain purity.

2.4.4. A Domain Specificity Assays

The determination of the specify of the respective A domains were performed in cooperation with Maximillian Müll and Hajo Kries (Independent Junior Research Group Biosynthetic Design of Natural Products, Leibniz Institute for Natural Product Research and Infection Biology e.V., Hans Knöll Institute (HKI Jena) Jena, Germany).

2.4.4.1. HAMA

HAMA was performed as described by Stanišić et al.[69] The reaction was conducted at RT in 100 µL volume containing 50 mM TRIS (pH 7.6), 5 mM $MgCl_2$, 150 mM hydroxylamine (pH 7.5-8, adjusted with NaOH), 5 mM ATP, 1 mM TCEP and a mix of 5 mM proteinogenic amino acids in 100 mM TRIS (pH 8) to a final concentration of 1 mM. The reactions were started by adding 1 µM enzyme and incubated for 1 h. UPLC-MS analysis was performed on a Xevo TQ-S micro (Waters GmbH) using established parameters.[69]

Table 24: Parameters for hydroxamate detection via MRM

	Compound	Parent (*m/z*)	Daughter (*m/z*)	Dwell time (s)	Cone voltage (V)	Collision energy (V)
1	GlyHA	90.8200	29.9400	0.022	34	8
2	AlaHA	104.9000	43.9100	0.022	22	8
3	ProHA	131.0381	69.9546	0.022	24	12
4	D-ValHA	132.8666	71.9115	0.022	22	10
5	ThrHA	134.9104	73.9076	0.022	26	8
6	CysHA	136.8704	75.8737	0.022	28	12
7	L-Val-d8-HA	140.9200	79.9600	0.022	22	10
8	PipHA	144.9943	83.9160	0.022	32	14
9	IleHA	147.0143	85.9673	0.022	30	10
10	AspHA	148.9543	87.9160	0.022	32	10
11	L-Leu-d7-HA	154.0500	93.0100	0.022	30	10
12	LysHA	162.0281	83.9435	0.022	20	18
13	GluHA	163.0281	83.9542	0.022	24	18
14	MetHA	165.0281	103.8765	0.022	26	8
15	HisHA	171.0481	109.9154	0.022	22	10
16	D-PheHA	180.9943	119.9363	0.022	30	10
17	L-Phe-d5-HA	186.9843	124.9700	0.022	30	10
18	ArgHA	190.0243	69.9363	0.022	14	16
19	TyrHA	196.9843	135.9516	0.022	30	12
20	TrpHA	219.9404	166.9987	0.022	30	16

Chromatography was performed on a Waters ACQUITY H-class UPLC system (Waters) with an injection volume of 3 µL. Water with 0.1 % formic acid (A) and acetonitrile with 0.1 % formic acid (B) were used as strong and weak eluent, respectively. Amino acid hydroxamates were separated on an ACQUITY UPLC BEH Amide column (1.7 µm, 2.1 x 50 mm) with a linear gradient of 10-50 % A over 5 min (flow rate 0.4 mL/min) followed by 4 min re-equilibration. Water containing 0.1 % formic acid was used as a needle wash between samples. Data acquisition and quantitation were done using the MassLynx and TargetLynx software (version 4.1). Hydroxamate peaks were integrated manually and concentrations were determined in comparison with synthetic standards.[69]

2.4.4.2. MesG/hydroxylamine Assay

The MesG/hydroxylamine assay was performed similar to a protocol described by Wilson and Aldrich.[64] Stock solutions of L-Asp, L-Hya, L-Pro and L-Hyp were

prepared in substrate buffer (100 mM Tris, pH 8, adjusted with HCl). The highest substrate concentrations used in the assay were 5 mM for L-Asp, 2 mM for L-Hya, and 0.625 mM for L-Pro and L-Hyp. Dilution series of the substrates were prepared in 2-fold steps with substrate buffer. The reactions also contained 150 mM hydroxylamine (adjusted to pH 7.5 with NaOH), 100 µM 7-methyl-6-thioguanosine, 5 mM adenosine-5'-triphosphate (ATP), 1 mM *tris*-(2-carboxyethyl)phosphine (TCEP), 0.4 U/mL inorganic pyrophosphatase, 1 U/mL purine nucleoside phosphorylase (PNP), 50 mM Tris (pH 7.6), 5 mM $MgCl_2$, and 2.5 µM enzyme (PlbA3 or PlbA8). The background reaction was monitored using a reaction with substrate buffer without amino acid. The reactions were performed in flat-bottom 384-well plates (781620, Brand) and the absorbance at 355 nm was measured with a Synergy H1 (BioTek) microplate reader at 30°C. The background was subtracted from the other reads and the initial velocities $v_0/[E]_0$ were fitted to the Michaelis-Menten equation by nonlinear regression using R version 3.4.2. Errors on the catalytic constants k_{cat} and K_M were calculated by determining the error of fit with two biological replicates. Each biological replicate was measured as technical triplicate and averaged.

2.5. Chemical and Analytical Methods

2.5.1. Solvent Extraction of Bacterial Cultures

After cultivation, the production cultures were acidified to pH 3-4 with hydrochloric acid. The culture broth was overlaid with butanol in a 1:1 ratio and extracted twice for at least 2 h with orbital shaking at 60 rpm. The solvent was evaporated using rotary evaporation to yield the crude extract.

2.5.2. Column Chromatography

The crude extract was fractionated and purified via solid phase extraction (SPE). The procedure was executed with a commercial cartridge (Strata™-XL 100 µm polymeric reversed phase 2 g / 12 mL, Giga tubes, Phenomenex®).

For equilibration of the cartridge 10 mL 100 % ddH_2O was used. The extract of interest was dissolved in 12 mL 100 % ddH_2O and added to the cartridge. This step equaled the first fraction of elution of the following gradient displayed in the following table.

Table 25: SPE gradient conditions

Solvent	Percentage	Volume
H_2O	100	12 mL
H_2O:MeOH	85:15	12 mL
H_2O:MeOH	70:30	12 mL
H_2O:MeOH	50:50	12 mL
H_2O:MeOH	35:65	12 mL
H_2O:MeOH	20:80	12 mL
MeOH	100	12 mL
Acetone	100	12 mL

Subsequently, each fraction of the SPE gradient was dried by rotary evaporation.

2.5.3. High Performance Liquid Chromatography

Further purification and separation of the SPE fractions were achieved using High Performance Liquid Chromatography (HPLC). The samples were run on a HPLC system consisting of a Waters 1525 Binary Pump with a 717 plus Autosampler, Waters In-Line Degasser AF and Waters 996 Photodiode Array Detector. A Phenomenex Luna C18 column (250 x 10 mm, 5 µm) in combination with the displayed gradient was used.

Table 26: HPLC gradient conditions

Time [min]	Gradient		Flow [mL/min]
	Acetonitrile [%]	ddH_2O + 0.1 % TFA [%]	
00:00	10	90	1
03:00	10	90	1
10:00	30	70	1
20:00	35	65	1
32:00	50	50	1
42:00	100	0	1
50:00	100	0	1
57:00	10	90	1
60:00	10	90	1

2.5.4. Linearization of Plusbacin

To improve the output of subsequent mass spectrometric analysis, ring cleavage of plusbacin or its derivatives were conducted. 250 µL of the HPLC purified fractions containing plusbacin or derivatives were hydrolyzed with 400 µL 1 M

sodium hydroxide for 90 min at 45°C or overnight at RT. After neutralization with 200 µL 2 M hydrochloric acid the samples were evaporated to dryness using rotary evaporation. The linearized plusbacin was dissolved in LC-MS grade methanol for further analysis.

2.5.5. Mass Spectroscopy

2.5.5.1. Low Resolution HPLC-ESI-MS

Low resolution (LR) HPLC-ESI (electrospray ionization)-MS analysis was performed to detect specific masses in complex natural product extracts. Sample separation occurred on an Agilent 1100 Series HPLC coupled with an ABSciex 3200 QTRAP® mass spectrometer for further analysis.

Prior to LC-MS analysis, sample were dissolved in 100 % methanol (LC-MS grade) and centrifuged to remove particles. The system was run with a Phenomenex Luna C18(2) column (250 x 2.0 mm, 5 µm) at a flow rate of 0.2 mL/min and according to the described settings. Data sets were recorded in positive ionization mode.

Table 27: LR-HPLC-ESI-MS gradient conditions

Time [min]	Gradient		Flow [µL/min]
	Methanol [%]	ddH_2O + 0.1 % TFA [%]	
01:00	40	60	200
10:00	50	50	200
30:00	100	0	200
45:00	100	0	200
60:00	40	60	200
65:00	40	60	200

The detection of the mass to charge ratio (*m/z*) values and further analysis was conducted with Analyst® 1.6 software (ABSciex).

2.5.5.2. High Resolution (HR) -HPLC-ESI-MS

HR-HPLC-ESI-MS and MS/MS experiments were conducted by Dr. Dorothee Wistuba (MS department, Institute of Organic Chemistry, University of Tuebingen, Germany) on a Bruker ToF-MS maXis Impact ESI-HR-MS, using positive scan modes.

For data analysis, Bruker Compass DataAnalysis 4.4 and MetaboScape 3.0 were applied.

III. Results

1. Biosynthetic Capacity for Secondary Metabolite Production of *Lysobacter* sp. Strain BMK333-48F3

Lysobacter sp. strain BMK333-48F3 is known for its ability to produce tripropeptin antibiotics.[44-47] Therefore, the genome of the strain was sequenced at BaseClear (Leiden, The Netherlands) in cooperation with the Department of Microbiology, Institute of Microbial Chemistry, Tokyo, Japan.[110] A *de novo* hybrid assembly was created by using results from a 10-kb Pacific Biosciences (PacBio) genomic library sequenced on a PacBio RS II instrument employing one single-molecule real-time (SMRT) cell and a genomic Nextera XT paired-end library sequenced on an Illumina HiSeq 2500 platform. The following table summarizes the hybrid assembly results.

Table 28: Sequencing metrics for Lysobacter sp. strain BMK333-48F3[110]

Parameter	Findings of BMK333-48F3
PacBio Sequencing	
No. of reads	1,452,682
Mean read length (bp)	3,450
No. of mapped reads	1,061,956
Avg coverage (X)	790
Illumina sequencing	
Read length (nucleotides)	2X150
No. of reads	7,259,870
Yield (Mbp)	870
Avg quality score	37.95
Avg coverage (X)	162
Median insert size (bp)	320
De novo hybrid assembly	
Genome size (bp)	5,227,231
GC content (%)	69.7
No. of contigs	7
No. of scaffolds	3
N_{50} (bp)	5,224,492
No. of gaps	4
Avg gap size (bp)	352
Total number of genes	4,487
No. of coding genes	4,395
No. of predicted biosynthetic gene clusters	12

To investigate the complete biosynthetic capacity for secondary metabolism of *Lysobacter* sp. strain BMK333-48F3, the assembled genome was analyzed *in silico* using the annotation tool antiSMASH.[111] According to the antiSMASH analysis, twelve biosynthetic gene clusters were predicted. The characteristics and similarities to known gene clusters are described in Table 29.

Table 29: Secondary metabolite gene clusters identified in Lysobacter sp. strain BMK333-48F3 by genome mining

Cluster	Type	Most similar known cluster	Similarity
1	NRPS-T1PKS	Rhizomide	100 %
2	CDPS	Lysocin	19 %
3	NRPS	BE-43547A1	13 %
4	Arylpolyene	Xanthomonadin	57 %
5	Redox-Cofactor		
6	RiPP-like		
7	NRPS		
8	Lanthipeptide-Class-II		
9	NRPS/NRPS-like/Cyanobactin	Lysobacin	5 %
10	NRPS-T1PKS	HSAF	87 %
11	RiPP-like		
12	NRPS	WAP-8294A2	30 %

1.1. Analysis of the Biosynthetic Gene Cluster of Tripropeptin

To locate the biosynthetic locus of the tripropeptins within the genome of *Lysobacter* sp. strain BMK333-48F3, all predicted NRPS gene clusters were investigated. Only gene cluster 9 showed the required potential to encode for tripropeptin as seen in Figure 14.

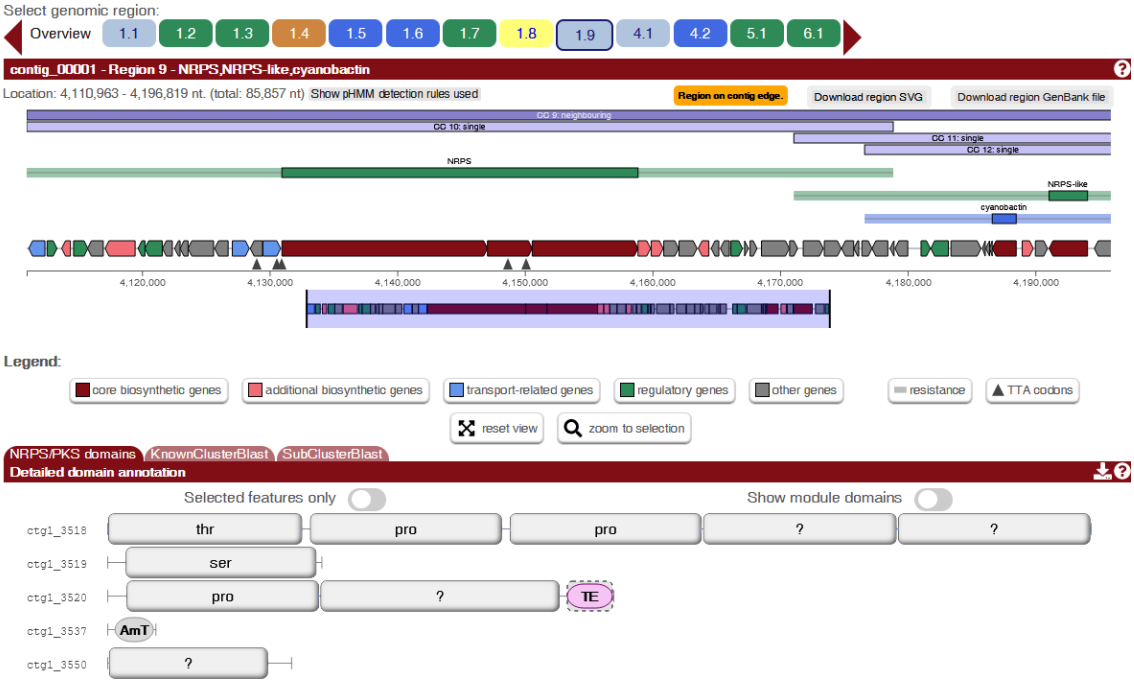


Figure 14: antiSMASH analysis of the predicted gene cluster 9 of Lysobacter sp. strain BMK333-48F3

A closer look into the genomic architecture of cluster 9, unveiled a NRPS assembly line consisting of eight modules stretched over three NRPS genes. *ctg1_3518*, *ctg1_3519* and *ctg1_3520* are referred to in the following as *tppA*, *tppB* and *tppC*, respectively. The NRPS genes are joined by two putative dioxygenase genes labeled as *tppD* and *tppE*. To predict the peptide sequence and to gain further insight into the biosynthesis encoded by the displayed gene cluster, the A as well as the C domains were analyzed *in silico*.

1.1.1. Bioinformatic Analysis of A Domains

The performed NRPSPredictor2 analysis[62] revealed the specificity of five A domains included in the gene cluster and therefore predicts a peptide sequence of Thr-Pro-Pro-X-X-Ser-Pro-X for the product. This serves as a decent starting point for further analysis since the known peptide sequence of tripropeptins is Thr-Pro-Pro-Arg- Hya-Ser-Hyp-Hya.[44-47] To figure out the characteristics of the other A domains a Stachelhaus prediction was also taken into consideration.[61]

Table 30: A domain specificities of the tripropeptin gene cluster of Lysobacter sp. strain BMK333-48F3 using Stachelhaus prediction

No.	Predicted Specificity	Stachelhaus sequence	Stachelhaus code match
A_1	Thr	DFWNIGMVHK	100 % (strong)
A_2	Pro	DVQFIAHVVK	100 % (strong)
A_3	Pro	DVQFIAHVVK	100 % (strong)
A_4	Arg	DVenVGAInK	70 % (weak)
A_5	Asp	DMaeLGMVDK	80 % (moderate)
A_6	Ser	DVWHVSLVDK	100 % (strong)
A_7	Pro	DVQFIAHVVK	100 % (strong)
A_8	Asp	DMveLGMVDK	80 % (moderate)

According to the A domain specificities displayed in Table 30, the NRPS product of the assigned gene cluster is Thr-Pro-Pro-Arg- Asp-Ser-Pro-Asp. This matches the peptide core of tripropeptin antibiotics perfectly, besides the hydroxylation of one proline at position 7 and both aspartic acids at position 5 and 8.

1.1.1.1. *In silico* Analysis of Proline-specific A Domains

Interestingly, the structural core of tripropetin contains unhydroxylated prolines as well as one hydroxyproline. It is unclear how the hydroxylation process is regulated. Since A domains are well known for their substrate specificity, there might be a chance that hydroxyproline instead of its unhydroxylated precursor is activated by the associated A domain. If the different hydroxylation pattern is due to the A domains, deviations in the protein sequences were expected. To clarify this question, protein sequences of the respective A domains were extracted *in silico* and compared not only to one another but also to those which are involved in the biosynthesis of the other closely related guanidine-containing cyclic lipopeptides. The protein sequence alignment containing all proline or hydroxyproline specific A domains of empedopeptin, plusbacin and tripropeptin from *Collimonas fungivorans* Ter331 as well as *Lysobacter* sp. strain BMK333-48F3 was visualized using the online multiple sequence alignment tool clustal omega[112] and is shown in Figure 15 below.

Overall, the protein sequences display large fragments of identical amino acids or of those with strongly similar properties, indicating highly conserved regions. After an amino acid count of about 120 amino acids a slightly bigger region of deviations

within the sequence alignment can be observed. Assuming this section might be part of the active site, the comparison between the protein sequences in this area was of special interest. As seen in Figure 15, most variations occur among the A domains of the different bacterial producer strains. Noteworthy is also that some variances of the A domains within the same species can be observed. Considering only the A domains of *Lysobacter* sp. strain BMK333-48F3, the alignment makes it obvious that A_2 and A_3 are practically identical, whereas A_7 deviates somewhat from their sequence.

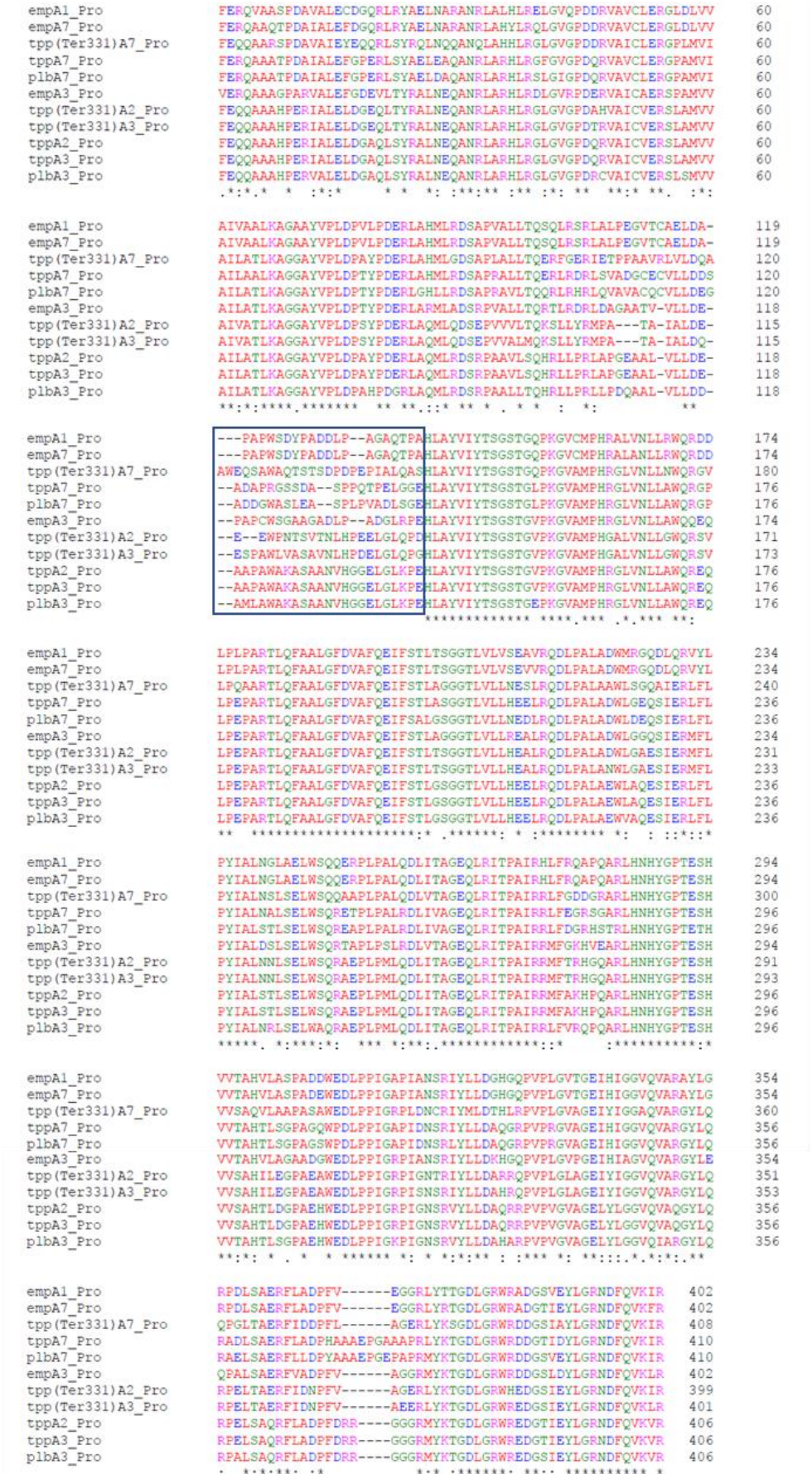


Figure 15: Sequence alignment of proline specific A domains of guanidine-containing cyclic lipopeptides

To gain a better understanding on how the differences in the peptide sequence of the proline specific A domains of *Lysobacter* sp. strain BMK333-48F3 impact the folding and therefore their function and specificity, a 3D model was simulated using SWISS-MODEL.[113-115] Since the A domain of module 2 and 3 display an identical peptide sequence, only one of them was used to create the 3D alignment shown in Figure 16 together with the A domain from module 7.

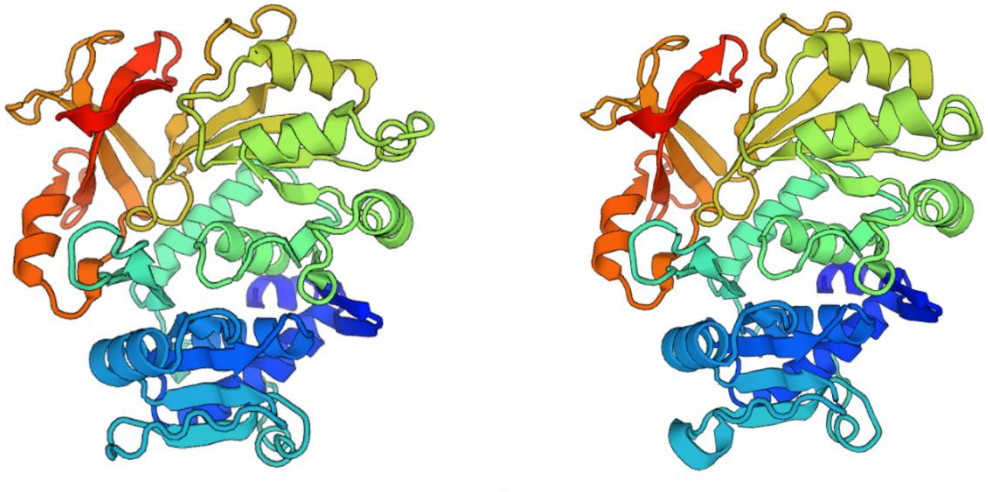


Figure 16: 3D models of the A domain of module 2 (left side) and 7 (right side) of the tripropeptin gene cluster of Lysobacter sp. strain BMK333-48F3 created using SWISS-MODEL

Similarities in protein folding was visualized by 3D alignment of both structures (Figure 17). The Local Distance Difference Test (lDDT) score was used by SWISS-MODEL to identify local deviations of the protein structures in red whereas the green areas indicate consistencies in the structure of the A domains.

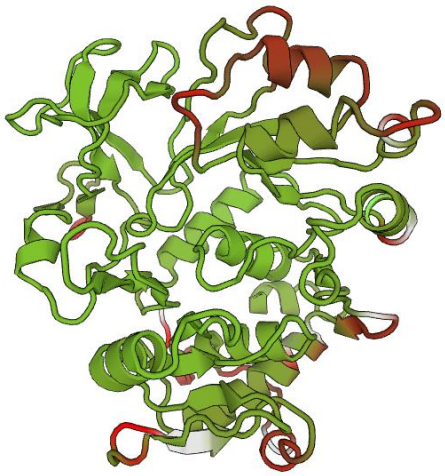


Figure 17: 3D alignment of the A domain of module 2 and 7 of the tripropeptin gene cluster of Lysobacter sp. strain BMK333-48F3 created using SWISS-MODEL

The 3D model shows minor discrepancies of the folding pattern of the investigated A domains. Those dissimilarities were not focused on just one specific area of the A domain like its binding site making it hard to predict their effect on the function.

In addition to the examinations which were limited to the A domains involved in the biosynthesis of the closely related empedopeptin, plusbacin and tripropeptin, their protein sequences were compared to other known A domains of NRPS assembly lines associated with the integration of a proline. Using clustal omega[112], a phylogenetic analysis was performed to visualize if the A domains of module 7 of the guanidine-containing cyclic lipopeptides show any striking peculiarities. To align the protein sequences the neighbour-joining method without distance corrections was selected. The resulting phylogenetic tree is displayed in Figure 18.

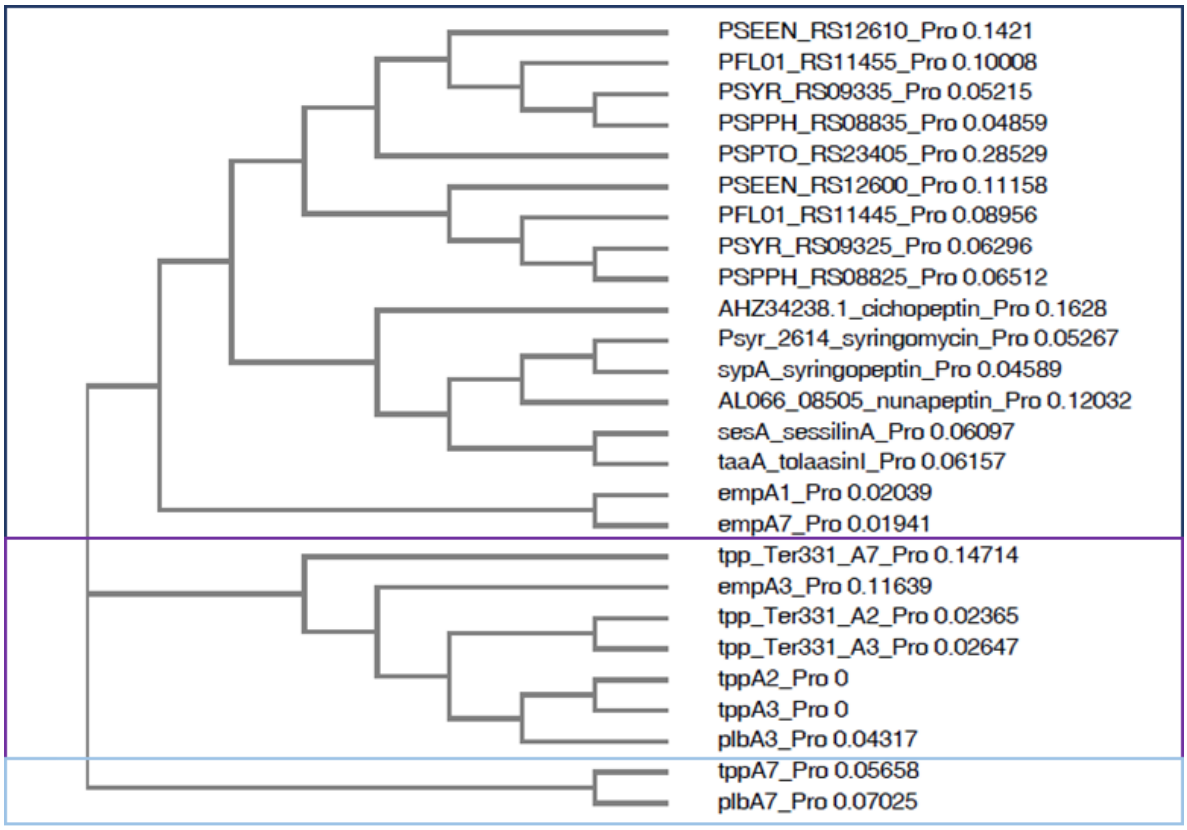


Figure 18: Phylogenetic analysis of proline specific A domains

Although most A domains encoded in the NRPS assembly line of empedopeptin, plusbacin or tripropeptin cluster together, the A domains of module 1 and 7 of the empedopeptin biosynthesis share a branch with mostly *Pseudomonas* spp. but also *Caldimona* sp. and *Actinoplanes* sp. whereas the A domains of module 7 involved in the biosynthesis of plusbacin and tripropeptin in *Lysobacter* spp. build

an independent clade. Since the A domains of module 7 encoded in the gene clusters of empedopeptin (empA7), plusbacin (plbA7) and tripropeptin (tppA7/ tpp_Ter331_A7) do not cluster in one group, the *in silico* analysis is indecisive. Therefore, just based on the bioinformatic studies no clear statement could be made as to whether the A domains display a specificity for proline or hydroxyproline.

1.1.1.2. *In silico* Analysis of Aspartic Acid-specific A Domains

The tripropeptin structure not only comprises hydroxyproline but also hydroxyaspartic acid, therefore similar *in silico* analysis as in previously described were conducted to indicate whether the A domains are responsible for the incorporation of the already modified amino acids. The A domain sequences of module 5 and 8 were extracted from the tripropeptin gene cluster and used for 3D modeling shown in Figure 19. Since both aspartic acids present themselves in a hydroxylated form, the A domain encoded in the NRPS assembly line of syringomycin, which is known to incorporate unhydroxylated aspartic acid[116], was used as a comparison.

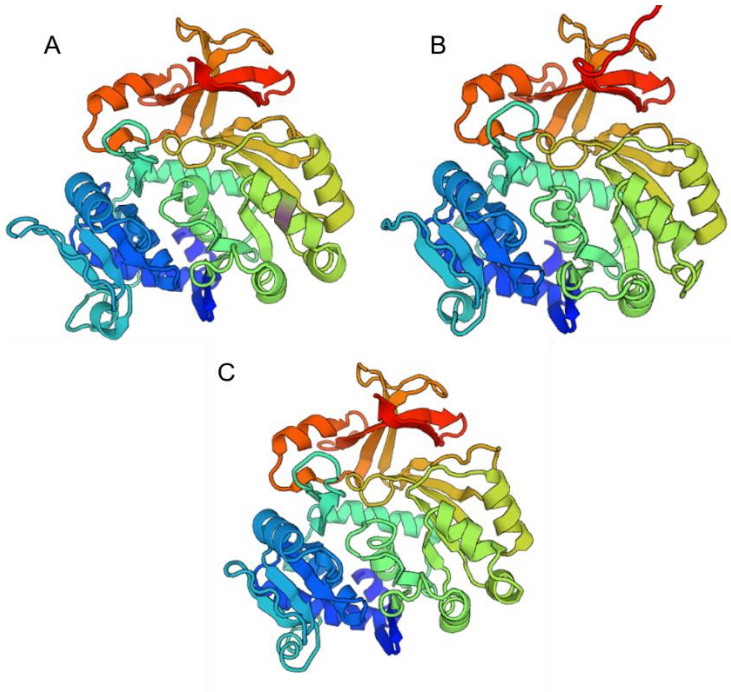


Figure 19: 3D models of the A domain of module 5 (A) and 8 (B) of the tripropeptin gene cluster of Lysobacter sp. strain BMK333-48F3 as well as Asp-specific A domain from the NRPS assembly line of syringomycin (C) created using SWISS-MODEL

Similarities in protein folding were visualized by 3D alignment of the structures displayed in Figure 20. The lDDT score was used by SWISS-MODEL to identify local deviations of the protein structures.

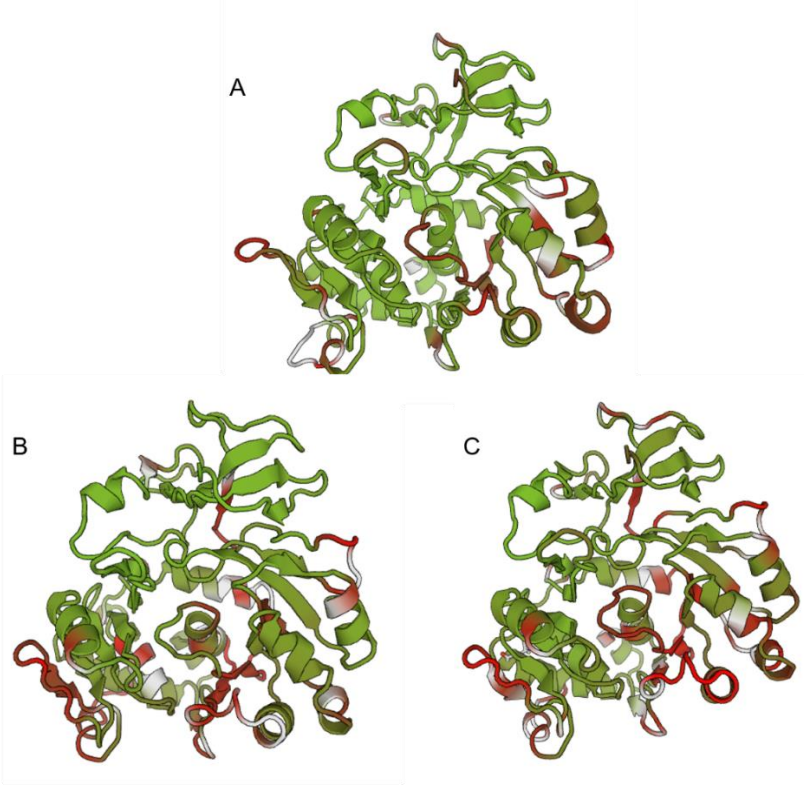


Figure 20: 3D alignment of Asp-specific A domains; (A) 3D alignment of both A domains encoded in the tripropeptin gene cluster of Lysobacter sp. strain BMK333-48F3, (B) 3D alignment of the A domain of modul 5 of the tripropeptin gene cluster and the Asp-specific A domain of syringomycin gene cluster, (C) 3D alignment of the A domain of modul 8 of the tripropeptin gene cluster and the Asp-specific A domain of syringomycin gene cluster

The deviations between both A domains activating aspartic acid or its derivative during the biosynthesis of tripropeptin seem minor. Seeing more prominent variations comparing the aspartic acid specific A domains of tripropeptin to the A domain activating aspartic acid during syringomycin biosynthesis, could indicate differences in their specificity. Already slight changes in the protein folding can excessively alter the A domains accessibility for its substrate and therefore allow binding of hydroxyaspartic acid. The 3D modeling is still indecisive since minor changes of protein folding are not only due to protein function.

To gain deeper insights in the characteristics of the A domains involved in the biosynthesis of guanidine containing cyclic lipopeptides, a phylogenic analysis was performed.

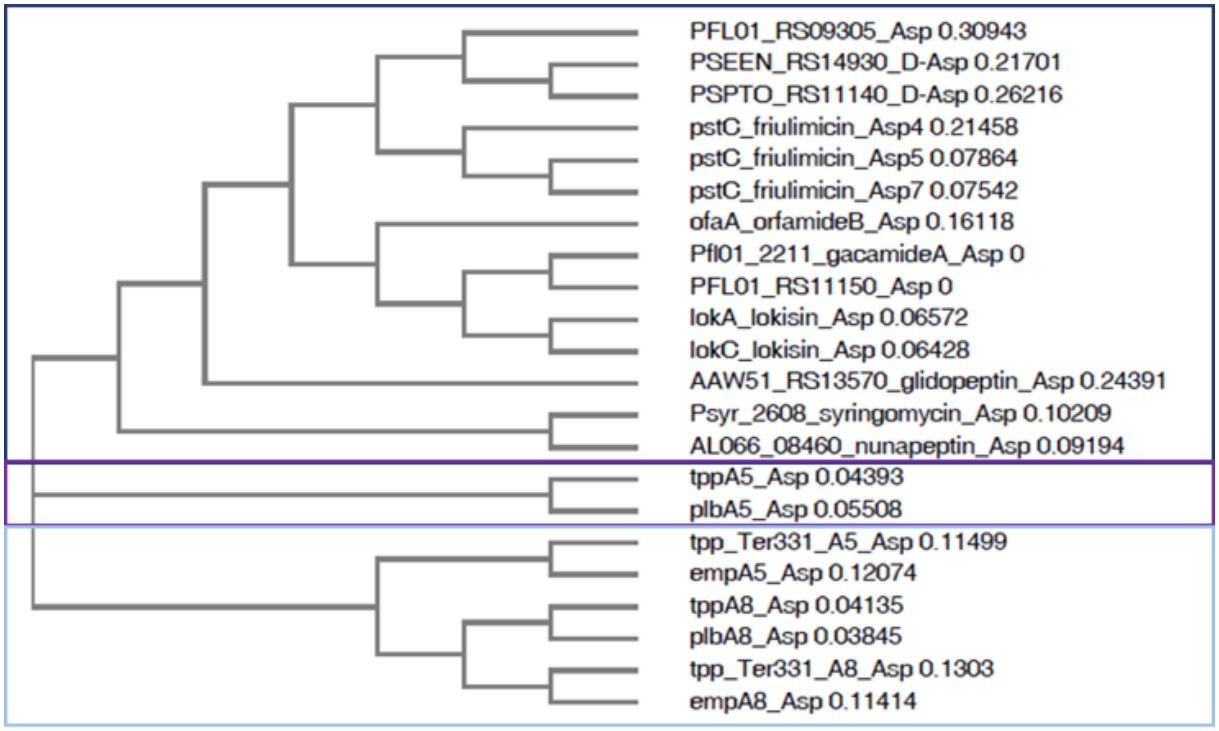


Figure 21: Phylogenetic analysis of aspartic acid specific A domains

As the phylogenetic tree in Figure 21 makes it immediately apparent that the investigated aspartic acid specific A domains mainly cluster in 3 groups. The first group contains A domains with a specificity for aspartic acid from *Pseudomonas* spp. but also *Caldimona* sp. and *Actinoplanes* sp. The other two phylogenetic cluster summarizes the A domains involved in the activation of aspartic acid in the biosynthesis of guanidine containing cyclic lipopeptides of the four known producer strains from the bacterial genera *Empedobacter*, *Lysobacter* and *Collimonas*. Noteworthy is also that the A domains involved in the biosynthesis of tripropeptin of both producer stains do no cluster together. However, the A domains of both *Lysobacter* producing guanidine containing cyclic lipopeptide, *Lysobacter* sp. strain BMK333-48F3 and PB-6250^T, are more similar. Interestingly, the A domains of module 5 of both *Lysobacter* producing tripropeptin and plusbacin, form an independent phylogenetic branch.

1.1.2. Bioinformatic Analysis of C Domains

To complete a fully functional module for elongation of the peptide, a condensation domain is required in addition to the A and T domain, since this domain catalyzes the peptide bond linkage and thus the actual elongation step. Hence, the C domain

provides a second center for substrate control in addition to the A domain. In contrast, the donor side of the C domain showed much lower substrate specificity.[117] However, C domains following an epimerization (E) domain select between D- and L-configured amino acids at a remarkably high degree, thus providing a filter for the successfully stereoinverted amino acid.[117-119] The macrolactone core of tripropeptin comprises a multitude of D-configured amino acids, which might be explained by the corresponding C domains. Bioinformatic examination revealed the C domain of the first module as a $C_{Starter}$ domain joined by four dual epimerization/condensation domains and three $^{L}C_{L}$ domains. Phylogenetic analysis using clustal omega illustrates that all C domains cluster in their respective group (Figure 22).

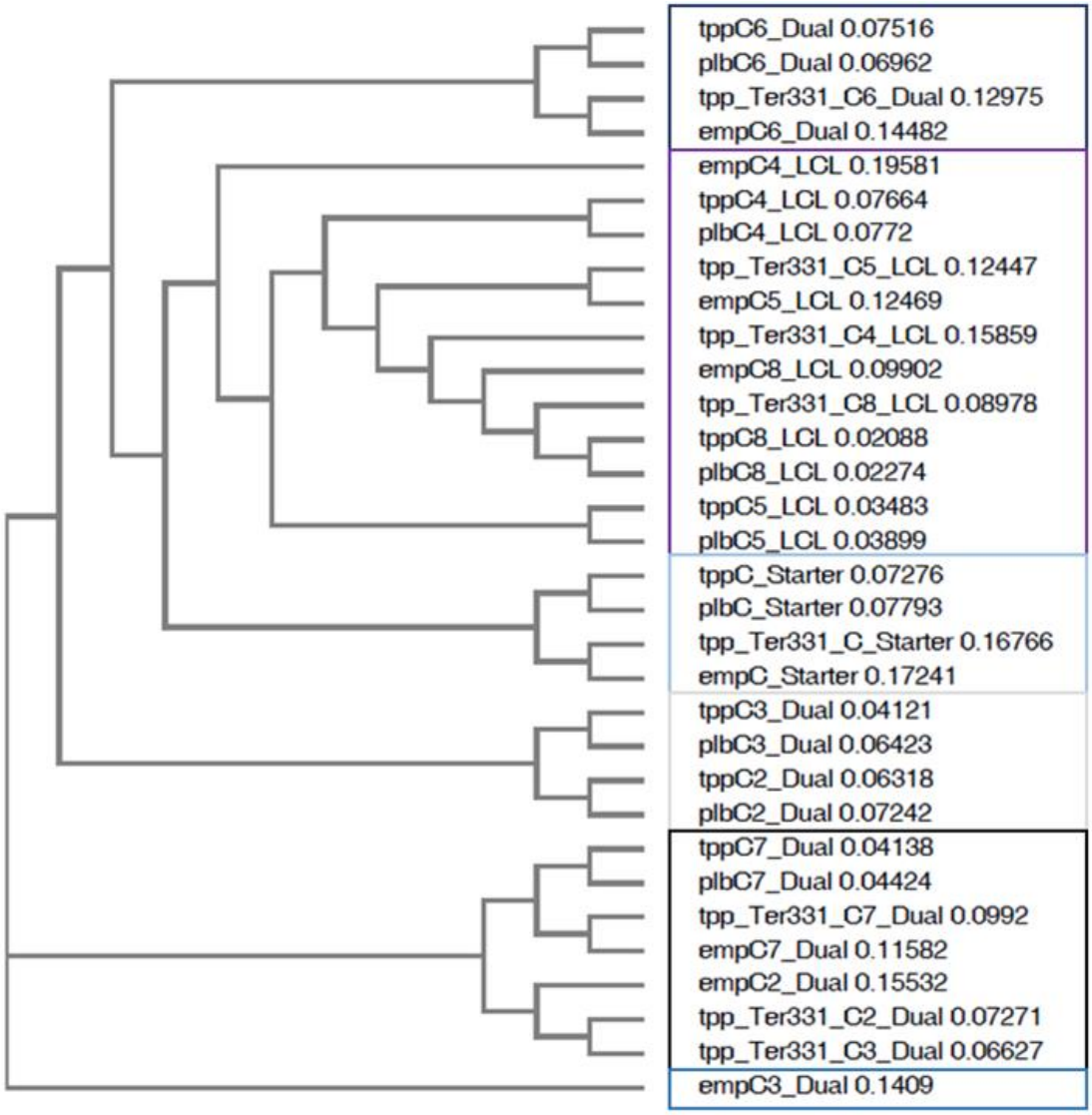


Figure 22: Phylogenetic tree of C domains involved in the biosynthesis of guanidine containing cyclic lipopeptides empedopeptin, plusbacin and tripropeptin

Based on the known tripropeptin structure and the order of the individual C domains in the NRPS assembly line, all dual E/C domains appear to be intact and fulfill their function during biosynthesis.

1.1.3. Biosynthesis of Tripropeptin

Based on the *in silico* analysis the following biosynthetic pathway for tripropeptin formation in *Lysobacter* sp. strain BMK333-48F3 was proposed.

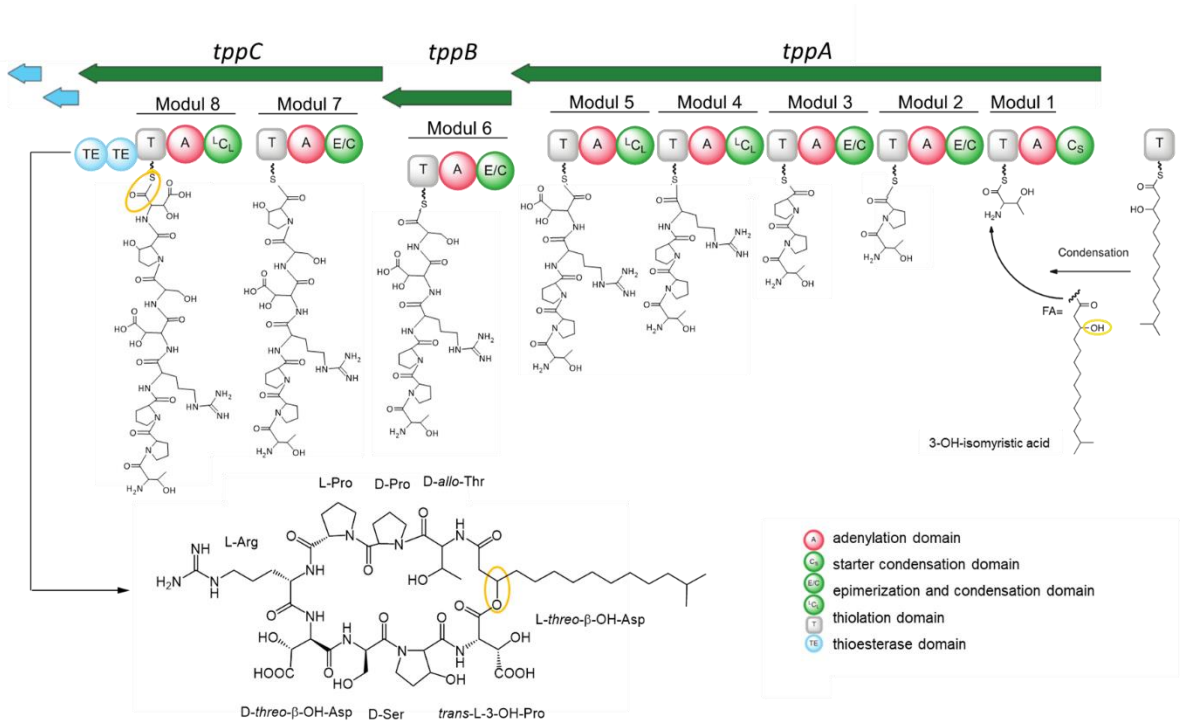


Figure 23: Putative biosynthetic pathway of tripropeptin production in Lysobacter sp. strain BMK333-48F3

The biosynthesis occurs on a NRPS assembly line (Figure 23). Three genes *tppA-C* encode the necessary eight NRPS modules. The C domain of module 1 was identified as a $C_{Starter}$ domain achieving the attachment of the lipid tail. L-Thr is selected and activated by the A domain of module 1 and transferred to the PCP domain, whereas the same happens on module 2 with L-Pro. Both amino acids are subsequently linked by a peptid bond catalyzed by the C domain. Since it is a E/C domain L-Thr gets stereoinverted to D-*allo*-Thr. After the A domain of module 3 selects and activates another L-Pro and transfers it to the corresponding thiolation domain, a peptide bond is formed linking both prolines. Consequently L-Pro from module 2 gets converted to D-Pro by the E/C domain. Since the next C domain is a $^{L}C_{L}$ domain, L-Pro recruited on module 3 preserves its stereochemistry even after

peptide bond formation with L-Arg. This modular assembly continues until all eight amino acids are incorporated into the peptide backbone of tripropeptin. Following the octapeptide formation, the peptide chain reaches the tandem TE domain. The termination process is catalyzed by the first TE domain through intermolecular cyclization and product release, whereas the second TE domain functions as a repair enzyme in case of a misprimed T domain.

The above-described biosynthetic mechanisms and the characteristics of the involved C domains explain the presence of the D-configurated amino acids. However, *in silico* analysis do not provide conclusive information on the hydroxylation processes of prolines and aspartic acid in the context of the NRPS biosynthesis. Experimental approaches are necessary to determine if the A domains control the incorporation of hydroxylated amino acids or instead activate the regular amino acid, which gets modified in a tailoring reaction by the cluster encoded dioxygenases *tppD* and *tppE*.

2. Studies on the Biosynthetic Gene Cluster of Plusbacin in *Lysobacter firmicutimachus* PB-6250^T

Seeing the tremendous similarities in the structures of empedopeptin, plusbacins and tripropeptins, a comparison between the gene clusters was performed and visualized using MultiGeneBlast[120] (Figure 24).

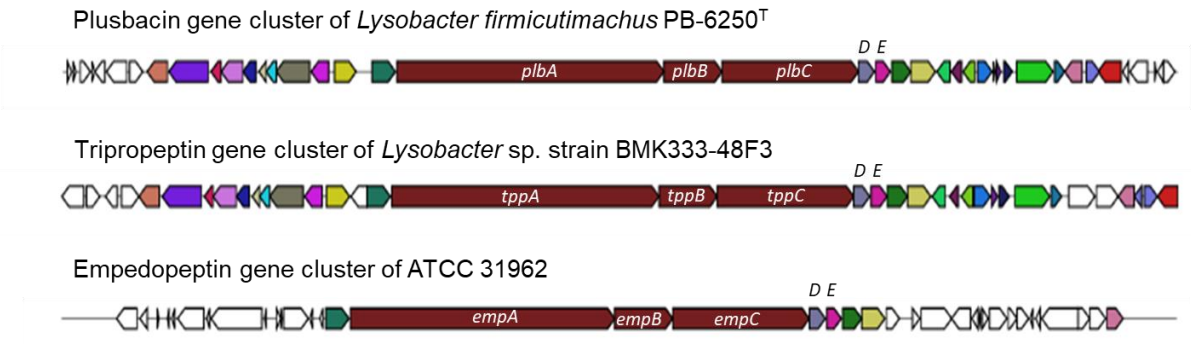


Figure 24: Gene loci of plusbacin, tripropeptin and empedopeptin in different producer strains

A BLAST analysis was used to visualize the actual similarities between the three gene loci of the guanidine-containing cyclic lipopeptides. The comparison of *plbA-E* and *tppA-E* displayed a coverage of 99 % and identity of 86.80 % whereas *empA-E* showed less coverage and identity in relation to *plbA-E* with values of 94 % and

75.29 %, correspondingly. A comparison of *tppA-E* to *empA-E* resulted in similar results showing a coverage of 95 % and identity of 76.05 %. A BLAST analysis of the dioxygenases also revealed high resemblances between the three representatives of this group with query and identity values over 80 % for *plbD/tppD/empD* and *plbE/tppE/empE*, respectively, suggesting the same function within the biosynthesis of the compounds.

For further experimental investigations the plusbacin producer strain *Lysobacter firmicutimachus* PB-6250^T was chosen.

2.1. Experimental Verification of the Plusbacin Gene Cluster

To confirm the plusbacin gene cluster, a cassette to knock out parts of the NRPS assembly line was designed. The general strategy is illustrated in Figure 25.

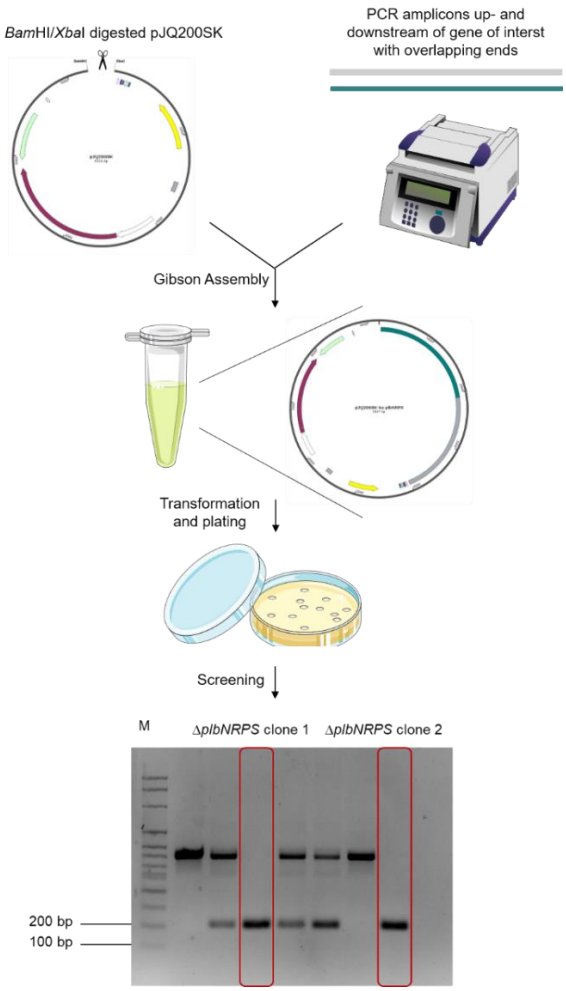


Figure 25: General workflow to generate ΔplbNRPS mutants

The knockout cassette for deletion of a region of the NRPS assembly line of plusbacin was assembled by cloning of homology arms into the *Bam*HI-*Xba*I-digested pJQ200SK backbone using Gibson Assembly. After verification of proper cloning via sequencing the resulting plasmid was transformed into *L. firmicutimachus* to disrupt the plusbacin gene cluster.

The secondary metabolite profile of the wild type producer strain PB-6250^T as well as the knockout mutants were elucidated. Therefore, the *Lysobacter* sp. strains were cultured and the crude extracts were obtained by acidification and extraction using butanol. For mass analysis, the sample extracts were subjected to LC-MS and additionally tested for antimicrobial activity against the indicator strain *B. subtilis* 168 (Figure 26). Moreover, an antagonistic assay was conducted.

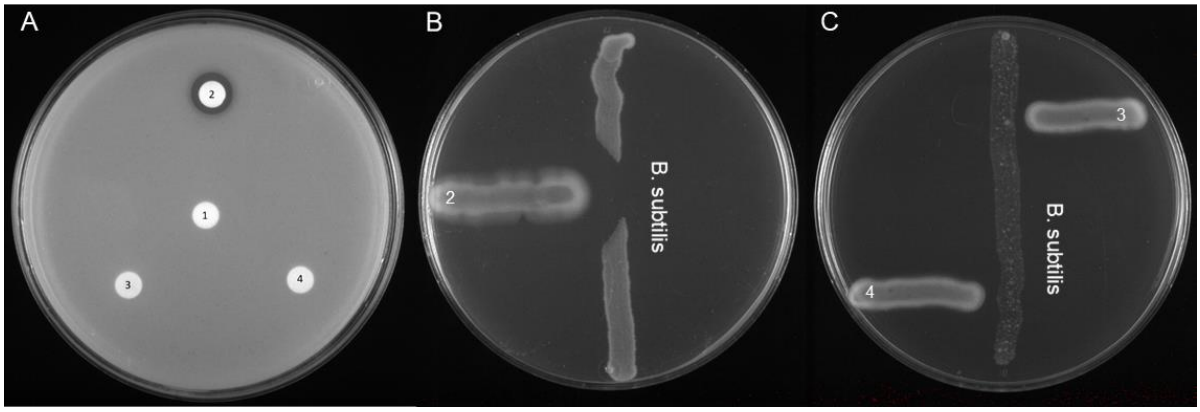


Figure 26: Analysis of ∆plbNRPS mutant strains: Antibiotic properties of the extract of PB6250^T (2) and ∆plbNRPS mutant strains (3,4) with a negative control (1) against B. subtilis 168 (left) and antagonistic assay against B. subtilis 168 (right)

The performed antagonistic assay as well as the disc diffusion bioassay indicated a total loss of antibiotic activity of the ∆*plbNRPS* mutants against *B. subtilis* as no zone of inhibition could be observed. To further investigate whether the loss of antibiotic activity is due to the elimination of plusbacin production, the extracted secondary metabolites of PB-6250^T and its mutant strains were submitted to LC-MS analysis. Plusbacin production in the wild type strain could be confirmed as seen in Figure 27. The masses of interest, shown in Table 31, were extracted from the total ion chromatogram (TIC). The resulting extracted ion chromatogram (XIC) shows peaks at retention times between 30 and 33 min corresponding to all plusbacin derivatives. A closer look revealed that plusbacins A_1 and B_1 elute first, followed by plusbacin A_2 and B_2 and plusbacin A_3, A_4, B_3 and B_4 last, represented by the peak at a retention time of 30,6 min, 31,4 min and 32,2 min respectively.

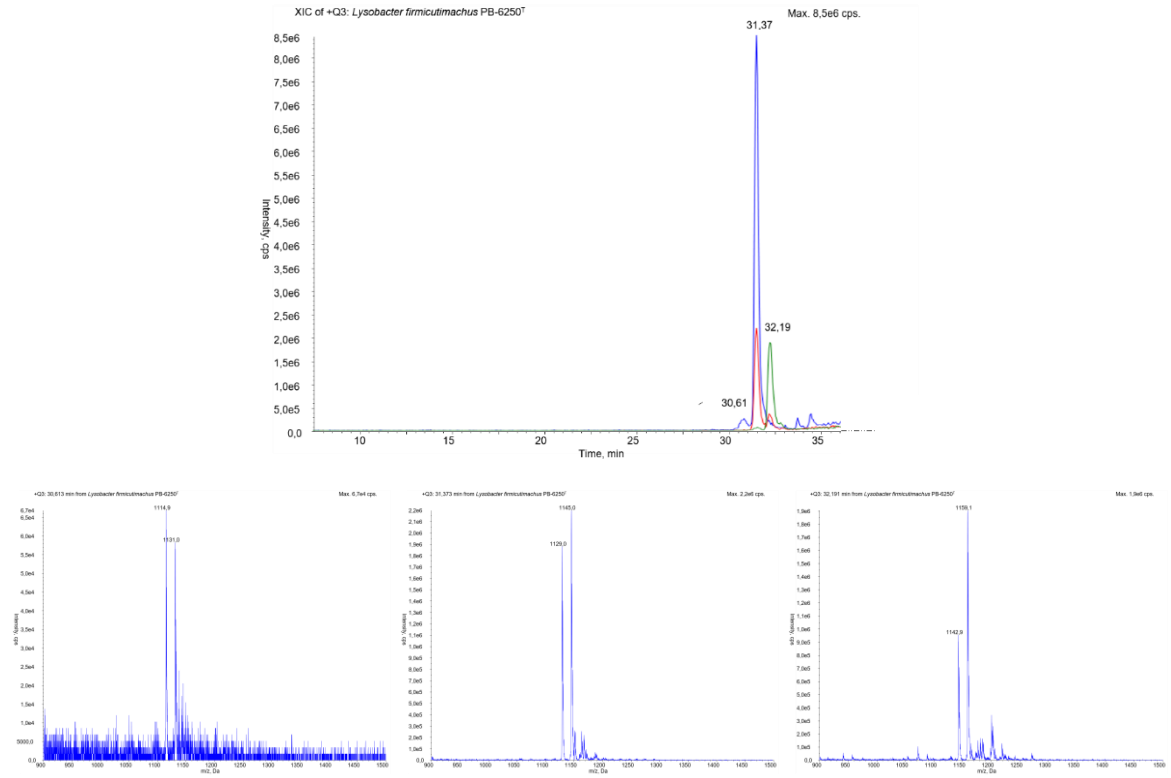


Figure 27: LC-MS screening of the secondary metabolite profile of PB-6250^T

Notable is the much lower peak intensity of plusbacin A_1 and B_1 indicating a decreased concentration of those derivatives compared to the plusbacins carrying a longer fatty acid side chain.

Table 31: Mass-to-charge ratios of plusbacins

Plusbacin *m/z* $[M+H]^+$			
A_1	1130.5569	B_1	1114.5613
A_2	1144.5735	B_2	1128.5796
A_3	1158.5883	B_3	1142.5945
A_4	1158.5893	B_4	1142.5934

Not only the extract of the wild type was examined but also those of the mutant strains harboring a deletion in the NRPS genes of the plusbacin gene cluster. Again, an extracted ion chromatogram was generated by extracting the masses of all plusbacins (Figure 28). In contrast to the wild type strain, no peaks were observed at the expected retention time of 30 min to 33 min. Therefore, it can be

postulated that the plusbacin production within the knockout strains is completely abolished whereas it is still intact within the wild type producer stain PB-6250^T.

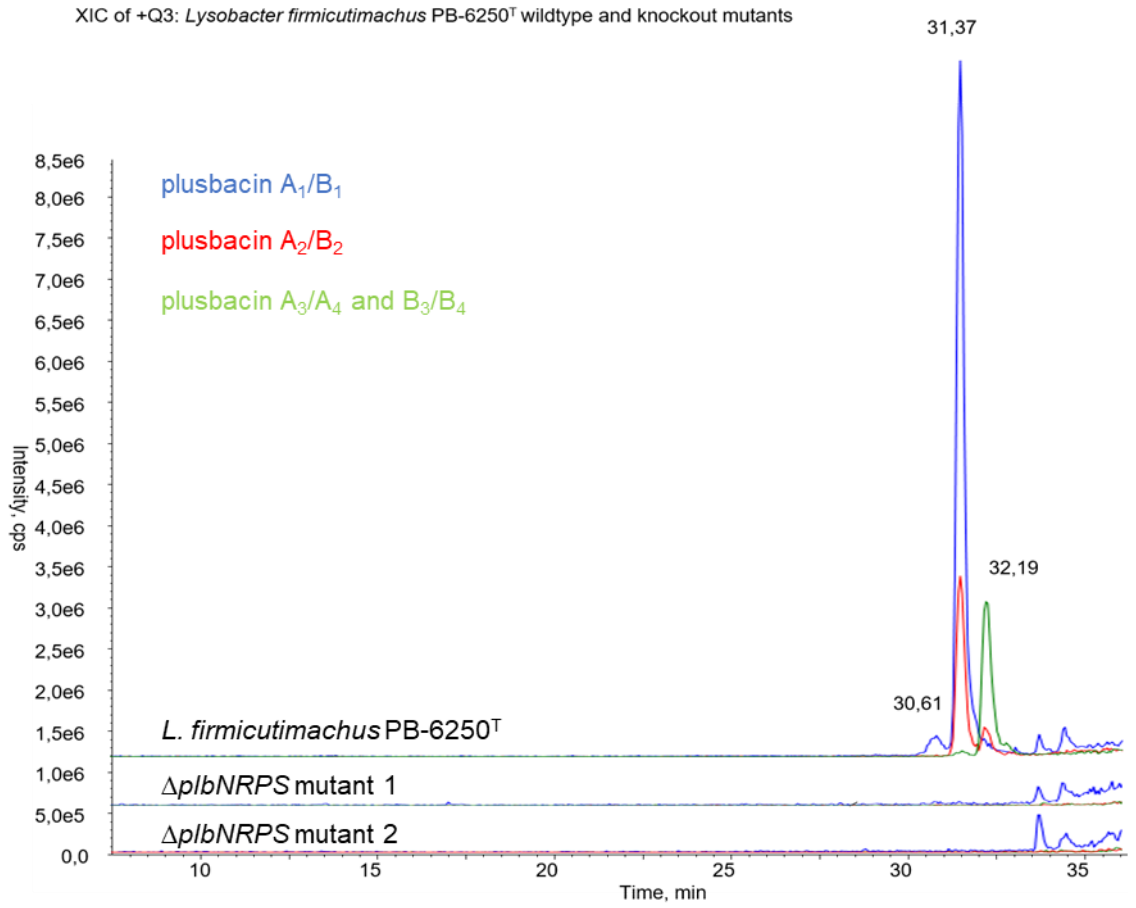


Figure 28: LC-MS analysis of the extracts of wild type plusbacin producer strain PB-6250^T and two ΔplbNRPS mutants

2.2. Studies on the Dioxygenases of the Plusbacin Gene Cluster

The plusbacin gene cluster revealed two putative dioxygenases *plbD* and *plbE*. The exact function and target remain to be clarified, since the macrolactone core of plusbacin comprises four hydroxylated amino acids for the A series and three for the B series. Several hypotheses as to how exactly the hydroxylation process takes place have been considered. Hydroxylation of the amino acids in question could occur prior to peptide assembly. Hereby, the precursors are hydroxylated by the dioxygenase and the modified amino acids are then selected and processed by respective A domains. To investigate this theory, the appropriate A domains were heterologously expressed and tested for their substrate selectivity, including not only the proteinogenic amino acid, but also the hydroxylated derivatives. Also

taken into consideration was the option that hydroxylation of amino acids could occur during or after the incorporation into the peptide backbone by *plbD* and *plbE*. Therefore, knockout studies were conducted as described in the methods section. To further clarify the impact of *plbD* and *plbE* to the plusbacin production, the dioxygenases were overexpressed in the producer strain *Lysobacter firmicutimachus* PB-6250^T.

2.2.1. Overexpression of the Dioxygenases

Overexpression studies on the dioxygenases *plbD* and *plbE* were conducted to investigate the effects on the production of plusbacin. The dioxygenase genes were chemically synthesized and individually cloned into pBBR1MCS-5. The corresponding replicative plasmids were introduced into the wild type plusbacin producer strain PB-6250^T to achieve non-integrative overexpression. Three biological individual mutants of a certain overexpression strain were cultivated on a 100 mL scale. For further analysis the crude extracts were obtained by acidifying and extracting the cultures twice with butanol.

Disc diffusion assays were conducted with the crude extracts of the *Lysobacter* sp. strains as shown in Figure 29 below.

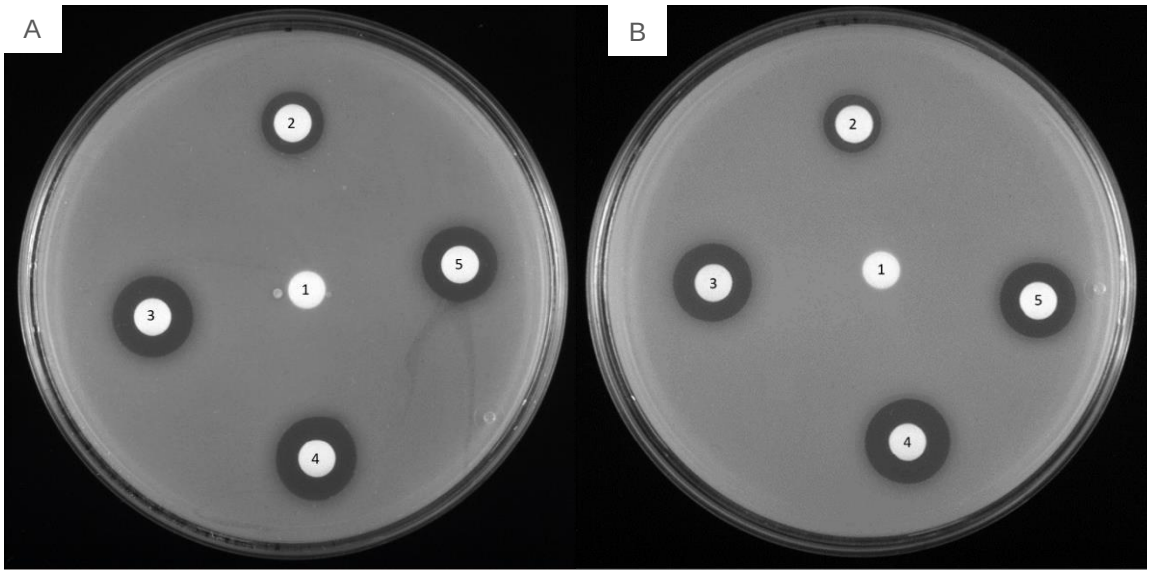


Figure 29: Analysis of dioxygenase overexpression mutant strains: Antibiotic properties of the extract of PB6250^T harboring the empty vector pBBR1MCS-5 (2) and overexpression strains (3,4,5) of plbD (A) and plbE (B) with a negative control (1) against B. subtilis 168

Zones of inhibition around the paper discs wetted with crude extract of the wild type stain are roughly half the size of the ones of the overexpression strains of *plbD* and *plbE*. Normally, the disc diffusion test is regarded more of a qualitative method to determine antimicrobial activity as antibiotic agent vary in their diffusion characteristics affecting the zone of inhibition size. However, in this experimental setting the assay is only performed with extracts of *L. firmicutimachus* strains containing plusbacin as antibiotic agent and therefore displaying the same diffusion characteristics, the disc diffusion test can therefore be considered semi-quantitative. Consequently, the diameter of the inhibition zone mainly depends on the concentration of plusbacin present in the crude extract the filter disc is soaked in. Hence, the bigger zones of inhibition of the mutant strains indicating a successful overproduction of plusbacin.

Additionally, the obtained crude extracts were subjected to LC-MS measured in triplicate. The results of the overexpression approach are depicted in Figure 30.

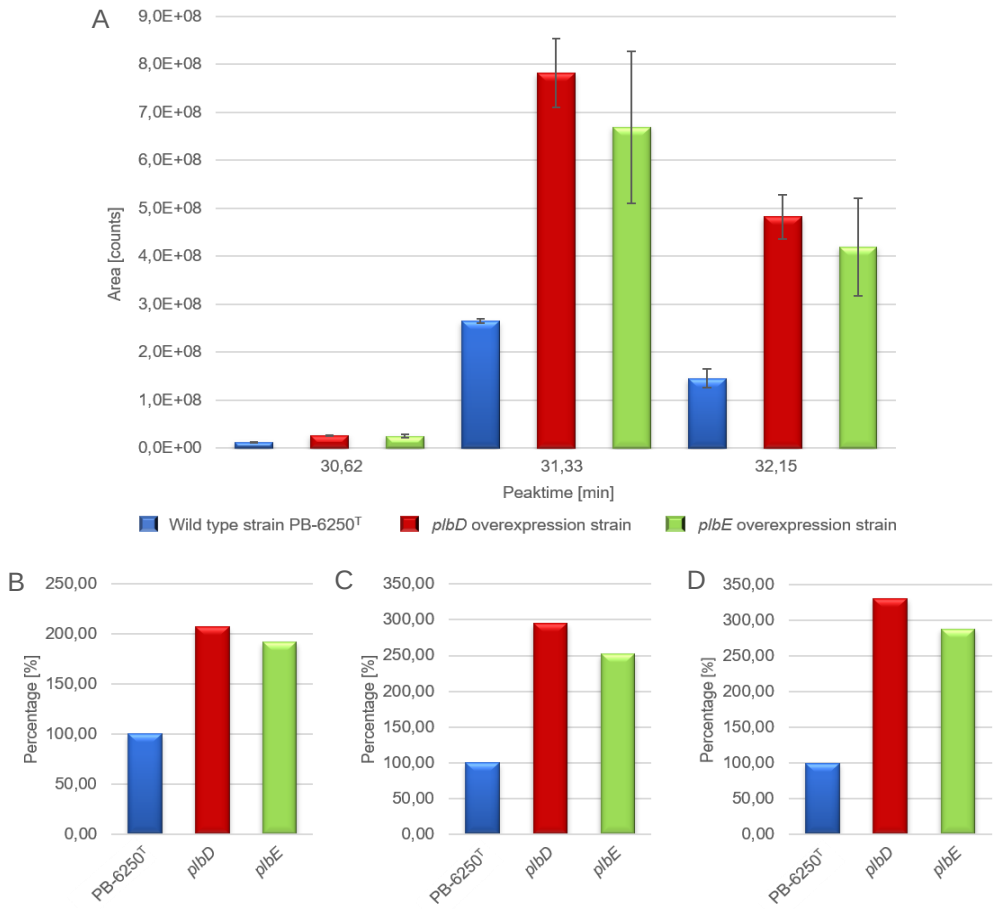


Figure 30: (A) Plusbacin production of PB-6250^T harboring the empty vector pBBR1MCS-5 and mutant strains overexpressing plbD and plbE and the percentual increase in production of plusbacin (B) A_1 and B_1, (C) A_2 and B_2 as well as (D) A_3/A_4 and B_3/B_4

As the peak area is directly proportional to the amount of plusbacin that is detected in the sample, this value was used for analysis (Figure 30A). Since there is no commercially available plusbacin standard to generate a proper calibration curve, a determination of the absolute plusbacin concentration was not possible. Therefore, a comparative approach was taken to correlate the production of the overexpression mutants to the amount of plusbacin produced in the wild type producer strain harboring the empty vector pBBR1-MCS5. On average, overexpression of each individual dioxygenase resulted in a 2- to 3-fold increase in plusbacin production (Figure 30B-D).

2.2.2. Target Identification of PlbD and PlbE

To gain further insight into the hydroxylation process catalyzed by the dioxygenases encoded by the plusbacin gene cluster, knockouts of *plbD* and *plbE* were performed. The secondary metabolite profile of the mutant strains was analyzed using mass spectrometry methods to pinpoint the exact target of each dioxygenase.

2.2.2.1. Generation of Double Crossover Knockout Mutants

Since a fosmid harboring parts of the plusbacin gene cluster including *plbD* and *plbE* was generated in a former study, the initial strategy was to generate knockouts by deleting the individual dioxygenase genes on pCC1Fos/plbOx via Red/ET-mediated recombination (Figure 31). For replacement of *plbD* and *plbE* respectively, the linearized and purified apramycin resistance cassette from pIJ774 was amplified. To achieve an appropriate PCR amplicon, elongated primers, containing homologous sequences to the adjacent up- and downstream regions of the dioxygenase genes were used. The resulting mutated foemids pCC1Fos/Δ*plbD*_apra and pCC1Fos/Δ*plbE*_apra were confirmed by restriction analysis. To create in-frame deletions, the resistance marker was eliminated in an *in vitro* reaction taking advantage of Cre-lox recombination. Screening of apramycin sensitive clones was performed and positive candidates were verified via restriction analysis, PCR and sequencing. The generated recombinant fosmids pCC1Fos/Δ*plbD* and pCC1Fos/Δ*plbE* were introduced in *L. firmicutimachus*

PB-6250^T to generate in-frame deletion mutants by double crossover recombination.

Even after screening hundreds of clones, no knockout mutant displaying a double crossover event could be identified.

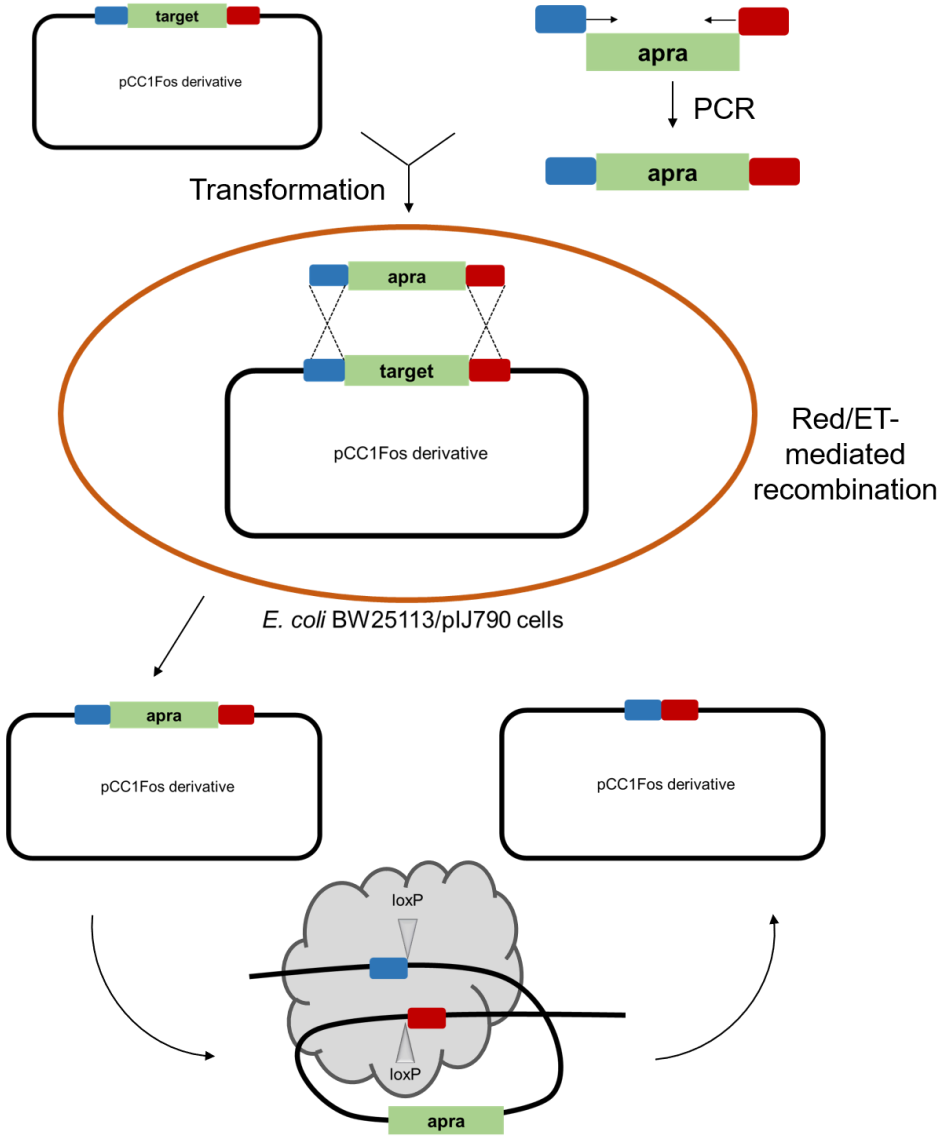


Figure 31: Cloning strategy for knockout vectors to generate in-frame deletions of plbD and plbE using Red/ET-mediated recombineering

Since pCC1Fos/Δ*plbD* and pCC1Fos/Δ*plbE* does not seem to be suitable to generate in-frame deletions in the genome of *L. firmicutimachus*, a new set of knockout plasmids were designed (Figure 32). pJQ200SK was chosen as a vector backbone because it acts as a suicide plasmid in PB-6250^T that is unable to

replicate in *Lysobacter* sp. strains. Furthermore, the vector harbors a *sacB* gene conferring sucrose sensitivity and therefore allowing counterselection.

For the generation of the knockout vectors pJQ200SK/Δ*plbD* and pJQ200SK/Δ*plbE*, the fosmid pCC1Fos/Δ*plbD* and pCC1Fos/Δ*plbE* served as template DNA to amplify the in-frame knocked-out using primers containing homologous arms. To assemble the fragment and the *Bam*HI-*Xba*I-digested pJQ200SK backbone, a Gibson Assembly was performed. The final constructs were verified using PCR and sequencing.

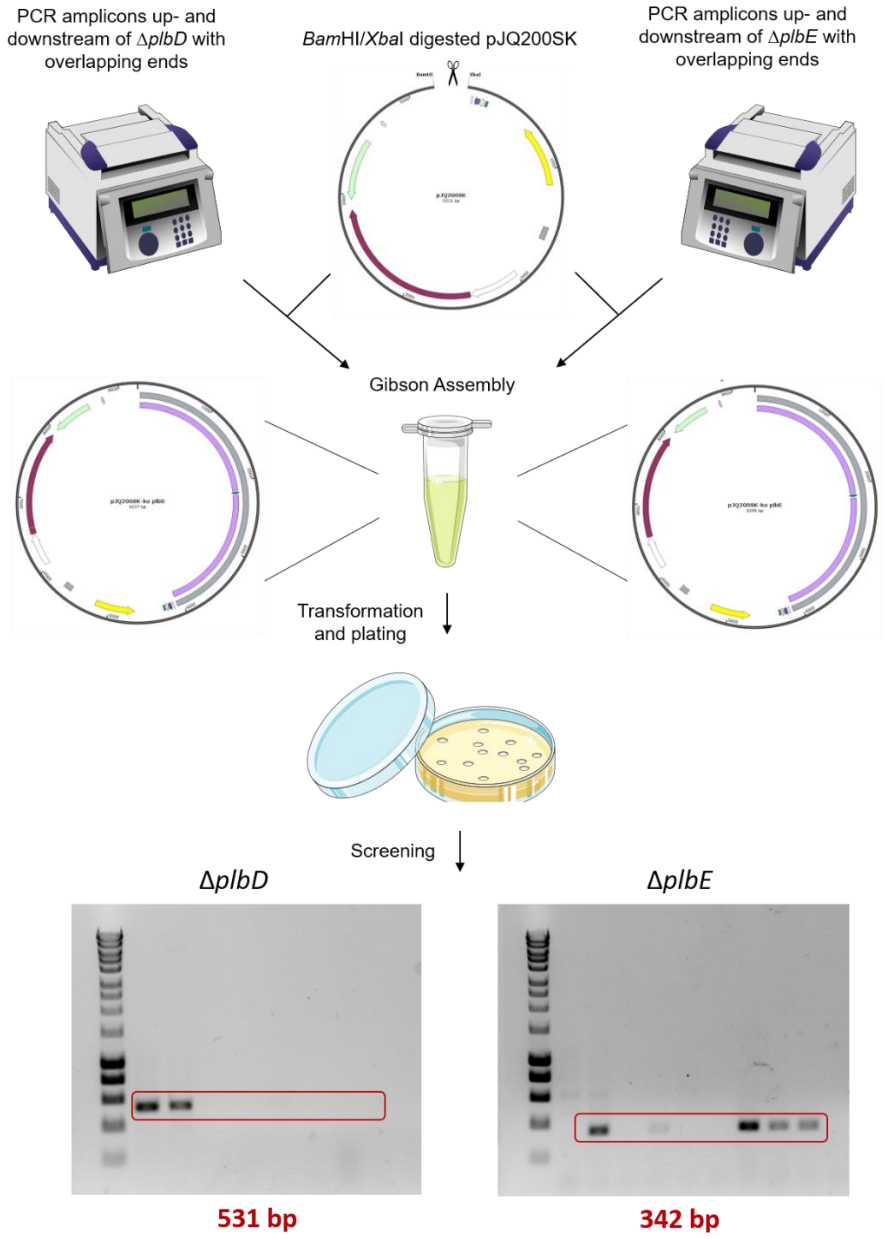


Figure 32: Cloning strategy for Gibson Assembly of pJQ200SK/ΔplbD and pJQ200SK/ΔplbE

The wild type producer strain was transformed using the assembled knockout constructs to enable homologous recombination and therefore in-frame deletion of *plbD* and *plbE* in the genome of PB-6250^T. The counterselectable marker *sacB* was instrumentalized to promote the death of the transformants containing it.

Hence, bacterial cells integrating the knockout construct based on the backbone of the suicide vector pJQ200SK containing *sacB* by a single crossover event retain a copy of the counterselectable marker in the genome and are therefore eliminated in the presence of sucrose. Although *sacB* counterselection can be associated with high false-positive rates, the screening method detailed in Figure 33 led to the generation and identification of double crossover mutants. The successful in-frame deletion was confirmed by PCR and sequencing.

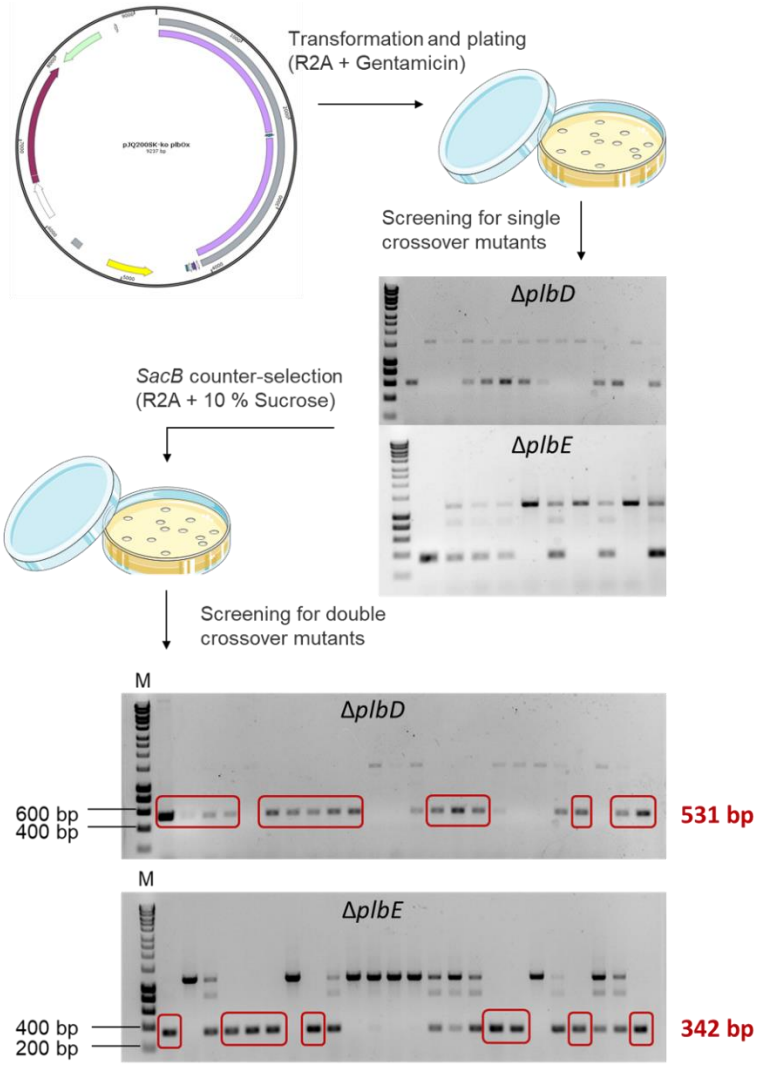


Figure 33: Workflow for the generation and identification of in-frame deletions of plbD and plbE using the counterselectable marker sacB

2.2.2.2. Analysis of Secondary Metabolites

After generation and verification of two biologically individual in-frame deletion mutants of each dioxygenase, their secondary metabolite profile was examined. Thus, the wild type producer strain as well as the knockout mutants were cultured first on a 100 mL scale and afterwards on a 9 L scale. The crude extract was obtained as described earlier after acidification and extraction with butanol twice. The generated samples were measured using LC-MS. To analyze the data, plusbacin masses were extracted from the TIC (Figure 34).

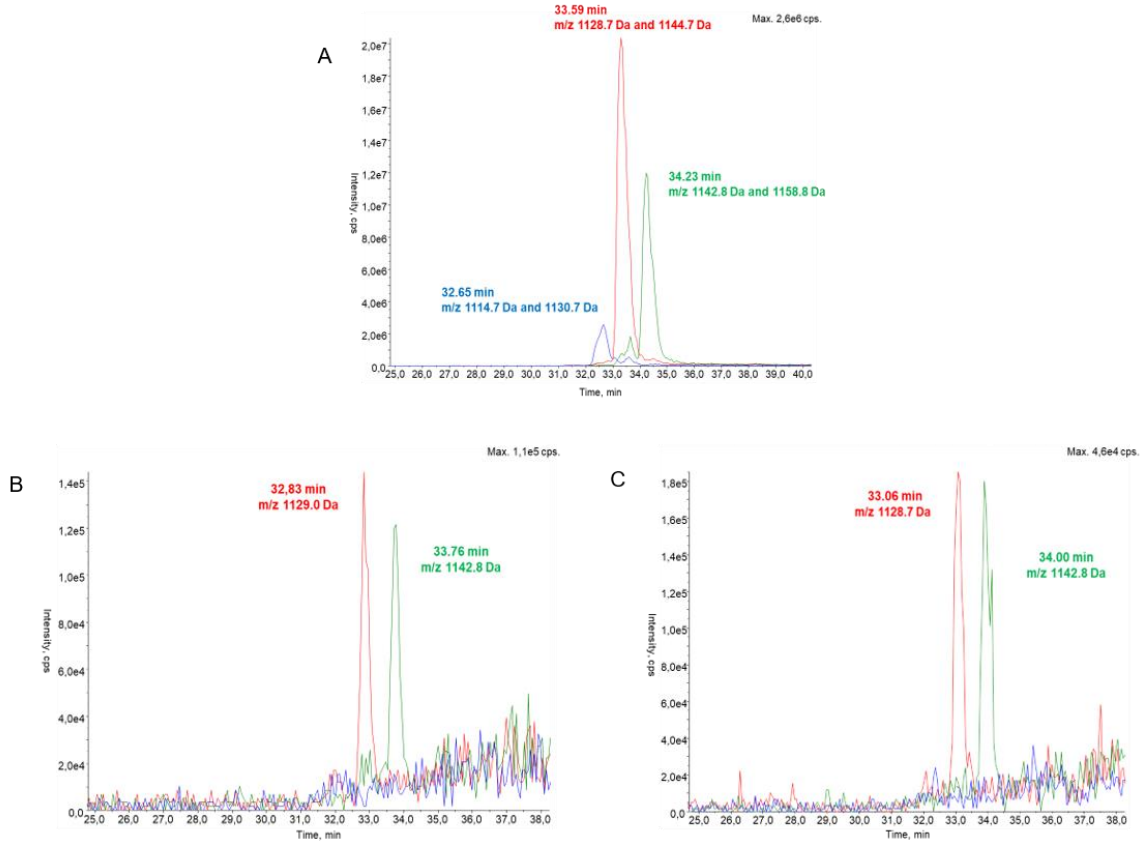


Figure 34: Extracted ion chromatogram of the extracted ions for plusbacin A1-A4 and plusbacin B1-B4 of (A) the wild type extract, (B) ΔplbD extract and (C) ΔplbE extract. The data was generated on a low-resolution LC-MS system (AB Sciex 3200 QT mass spectrometer coupled with an Agilent 1100 Series HPLC)

Since the XIC of the biologically individual mutants of the respective knockouts of the dioxygenases are comparable, the data of only one of each is shown in Figure 34. All plusbacin derivates are present in the wild type sample as all masses of interest (Table 31) could be identified. Comparing the data of the knockout extracts to those of the wild type producer strain, a lower concentration of the present plusbacin derivatives seem to be accumulated. This is indicated by a drastic decrease in peak intensity, which also makes the background noise more

apparent. Striking is also that the masses assigned to plusbacins of the A series and plusbacin B_1 are missing. Being unable to identify the masses of plusbacin A_1 and B_1 might be related to the already mentioned lower intensities of all peaks in general. As these peaks are already the least pronounced ones in the wild type, they may disappear in the background noise in the spectra of the knockout extracts. Another possibility would simple be that the knockout mutants lack the ability to produce those compounds due to the in-frame deletion in their genome. Not only the masses of plusbacin A_1 and B_1 are absent, but also those of all other plusbacin derivatives of the A series. Hence, the production of the whole plusbacin A series seems to be abolished. Mass analysis of both in-frame deletion mutants of the dioxygenases showed only two peaks at ~1129 Da and ~1143 Da, which correlate at first glance to plusbacin B_2 and B_3/B_4.

Comparing the masses of the plusbacin A series to the B series a mass difference of 16 Da is apparent. This variance is linked to the loss of an oxygen atom and can be traced back to the replacement of the hydroxylproline by proline. Since *plbD* and *plbE* encode for dioxygenases, is it hypothesized that an in-frame deletion would result in the loss of a hydroxyl group and therefore a mass shift of 16 Da. Thus, based on the LC-MS data, it can only be postulated that the knockout mutants assemble plusbacin derivatives with equal *m/z* values as those of the B series. The produced compounds therefore carry like plusbacin B only three hydroxylated amino acids instead of four. However, this analysis does not clarify if the compound produced by the mutant strains also lack the hydroxylation at the proline assembled by the third A domain in the NRPS assembly line or at another amino acid and position in the macrolactone core. Inconclusive is also whether the in-frame deletion of *plbD* has the same effect on plusbacin production as the deletion of *plbE* has.

To examine the antibiotic properties of the compounds produced by the knockout mutants, bioassays were performed. The results are displayed in Figure 35 showing no zone of inhibition around the discs which the crude extracts of the mutants. The absence of a zone of inhibition indicates either no or too low antibiotic activity to inhibit the growth of *B. subtilis*. Since 10 µL of the respective crude extracts were used for the bioassay, the concentration of the antibiotic agents in the sample varies and therefore also play a role in the formation of the zone of

inhibition. According to peak intensity of the LC-MS measurements, the compound concentration in the mutant extracts are significantly lower than in the wild type producer strain. The results of the disc diffusion assay are in agreement with those of the antagonistic assay. While the wild type shows clear inhibition of *B. subtilis*, the knockout mutants show only slight to no antibiotic activity.

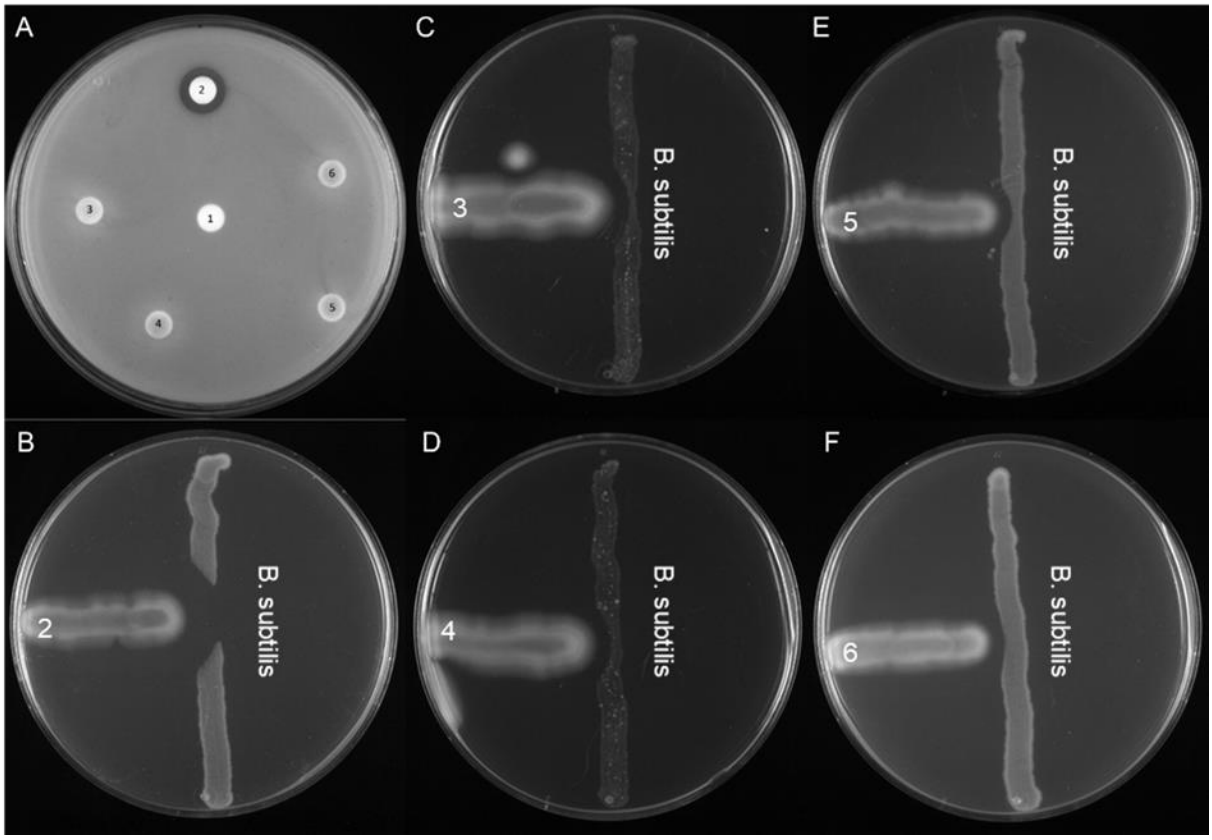


Figure 35: Antibiotic properties of plusbacin wild type producer strain PB-6250^T and the dioxygenase knockout mutants: (A) antibiotic properties of the extract of PB6250^T (2) and ∆plbD (3,4) as well as ∆plbE mutant strains (5,6) with a negative control (1) against B. subtilis 168, (B) antagonistic assay of PB-6250^T versus B. subtilis 168, (C/D) antagonistic assay of PB-6250^T/∆plbD versus B. subtilis 168, (E/F) antagonistic assay of PB-6250^T/∆plbE versus B. subtilis 168

2.2.2.3. Enrichment and Ring Cleavage

To determine the hydroxylation pattern of the compounds produced by the knockout mutants, HR-HPLC-ESI-MS and MS/MS experiments were executed. Since measurements with diluted crude extracts failed, further purification of the samples and enrichment of the substance in question needed to be carried out.

First, a solid phase extraction procedure was performed using Strata™-XL 100 µm polymeric reversed phase 2 g/12 mL giga tubes. All fractions were subjected to LC-MS for mass analysis. Plusbacin derivatives were identified in the 80 %

methanol fractionation as well as in 100 % methanol. Due to containing impurities and concentration of the compound in question, the 80 % methanol fraction was selected for further HPLC purification.

Additional separation was achieved by implementing five cut offs on a HPLC system using a Luna C18 column supplied from Phenomenex. Based on the LC-MS data, the fraction collected at a retention time between 25 min and 30 min was chosen for tandem mass spectrometry. Due to poor fragmentation, the initial results were inconclusive. To enable the peptide bonds in the macrolactone core to break easier, the ring was linearized chemically prior to MS/MS analysis.

Therefore, samples were treated for 90 min at 45°C or overnight at room temperature. LC-MS measurements provided information on which reaction conditions were more suitable for cutting open the peptide backbone of plusbacins.

Table 32: Theoretical masses of plusbacins and their linearized versions

A series	Theoretical mass [closed ring]	Theoretical mass [linearized]
A1	1129.5503	1147.5608
A2	1143.5659	1161.5765
A3	1157.5816	1175.5921
A4	1157.5816	1175.5921
B series	Theoretical mass [closed ring]	Theoretical mass [linearized]
B1	1113.5554	1131.5659
B2	1127.5711	1145.5816
B3	1141.5867	1159.5972
B4	1141.5867	1159.5972

The LC-MS data was screened for the presence of all the masses of interest displayed in Table 32. Only the *m/z* rations of all the linearized plusbacins were detected, indicating a successful ring cleavage under both conditions. Figure 36 shows the data for plusbacin A_2 and B_2 in an exemplary manner.

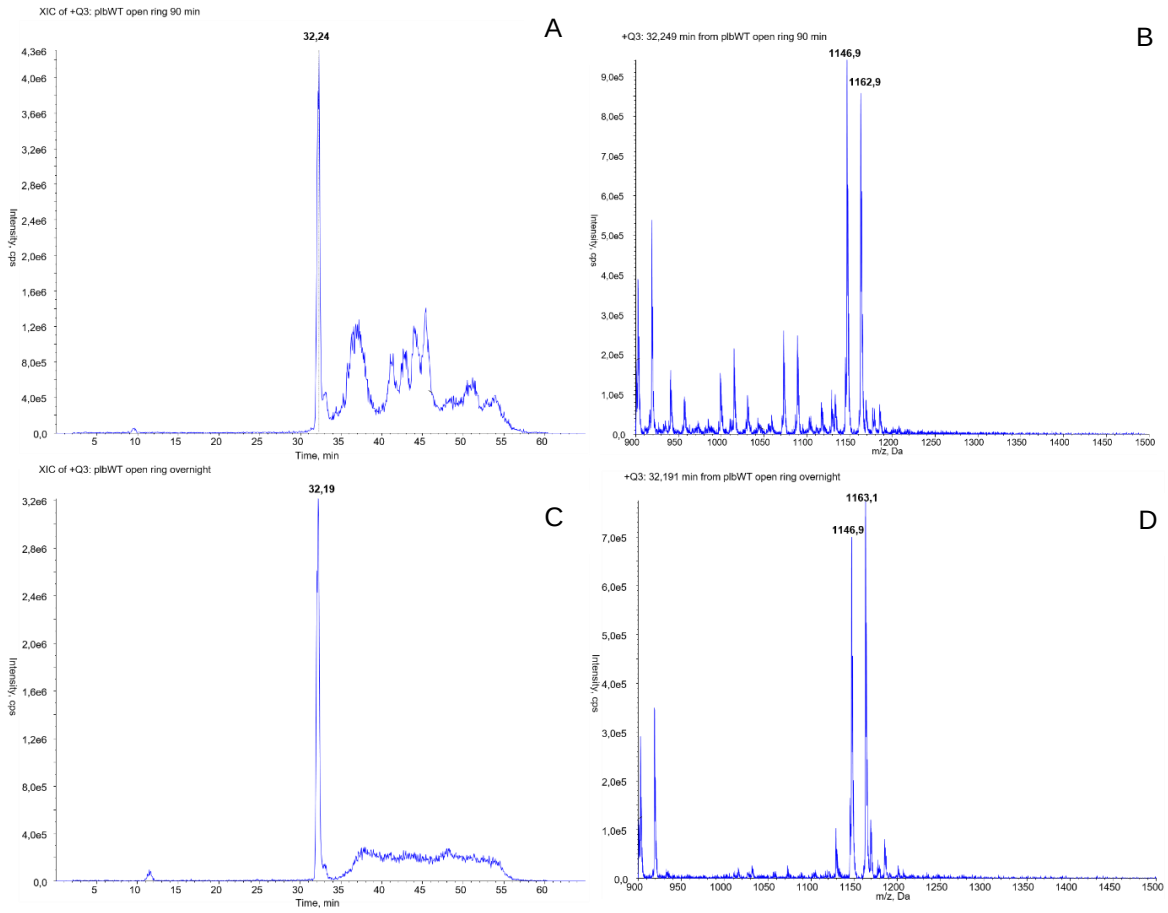


Figure 36: LC-MS screening of linearized plusbacin A_2 and B_2; (A) XIC of wild type producer strain linearized for 90 min at 45 ˚C, (B) LR-LC-MS Q3 positive mode scan of wild type producer strain linearized for 90 min at 45 ˚C at a retention time of ~32 min, (C) XIC of wild type producer strain linearized overnight at RT, (D) LR-LC-MS Q3 positive mode scan of wild type producer strain linearized overnight at RT at a retention time of ~32 min

Since both reaction conditions seem to be suitable to linearize plusbacin, all MS/MS samples were prepared overnight at room temperature owing to the amount of fewer unwanted by-products.

2.2.2.4. Tandem MS Spectrometry

To pinpoint the location of the missing hydroxylation in the in-frame deletion mutants of *plbD* and *plbE*, Liquid Chromatography/High-resolution Electron Spray Ionization Mass Spectrometry (LC/HRESI-MSMS) measurements were performed. Additionally, a sample of the wild type producer strain was also subjected for tandem mass spectrometry to allow a comparison of the MS^2 spectra of the plusbacin A- and B-series to the derivatives, produced by the mutant strains.

Exemplary, MS/MS data of plusbacin derivatives bearing a 3-hydroxy-13-methyltetradecanoic acid side chain are presented (Figure 37, Figure 38, Figure

39, Figure 40) and analyzed below (Table 33, Table 34, Table 35, Table 36, Figure 41, Figure 42, Figure 43, Figure 44).

The analysis of the MS^2 spectra reveals slightly different fragmentation patterns in the y- and b-ions for plusbacin A2 and B2 as well as the compounds produced by both knockout mutant. As the deviations are probably due to the hydroxylation pattern, a discrimination between Hya/Asp and Hyp/Pro, respectively, is necessary. Therefore, the investigation was focused on fragments that contain either of the amino acid in position 3, 5, 7 or 8 of the macrolactone peptide core. Comparing the mass of the y_1 ions revealed a doubly charged ion at a *m/z* value of 134 Da while all other spectra are displaying a *m/z* value of 150 Da. Since the y_1 ion correlates to the amino acid in eighth position of the peptide core, the hydroxylation of this aspartic acid is catalyzed by the dioxygenase encoded by *plbE*. The subsequent variations of *m/z* value are due to the Hya/Asp assembled as last amino acid in the NRPS assembly line in the respective compounds. Interestingly, the MS^2 spectra of the in-frame deletion strains Δ*plbD* and Δ*plbE* unveil the same *m/z* values for y_5 and the following ions. Unfortunately, the y_4 ion could not be identified in either mutant. To clarify if the hydroxylation occurs on Arg_4 or Asp_5, the b ions were considered. The masses of the b_4 and b_5 ions pinpoint the aspartic acid in position five of the macrolactone core as target of the dioxygenase PlbD.

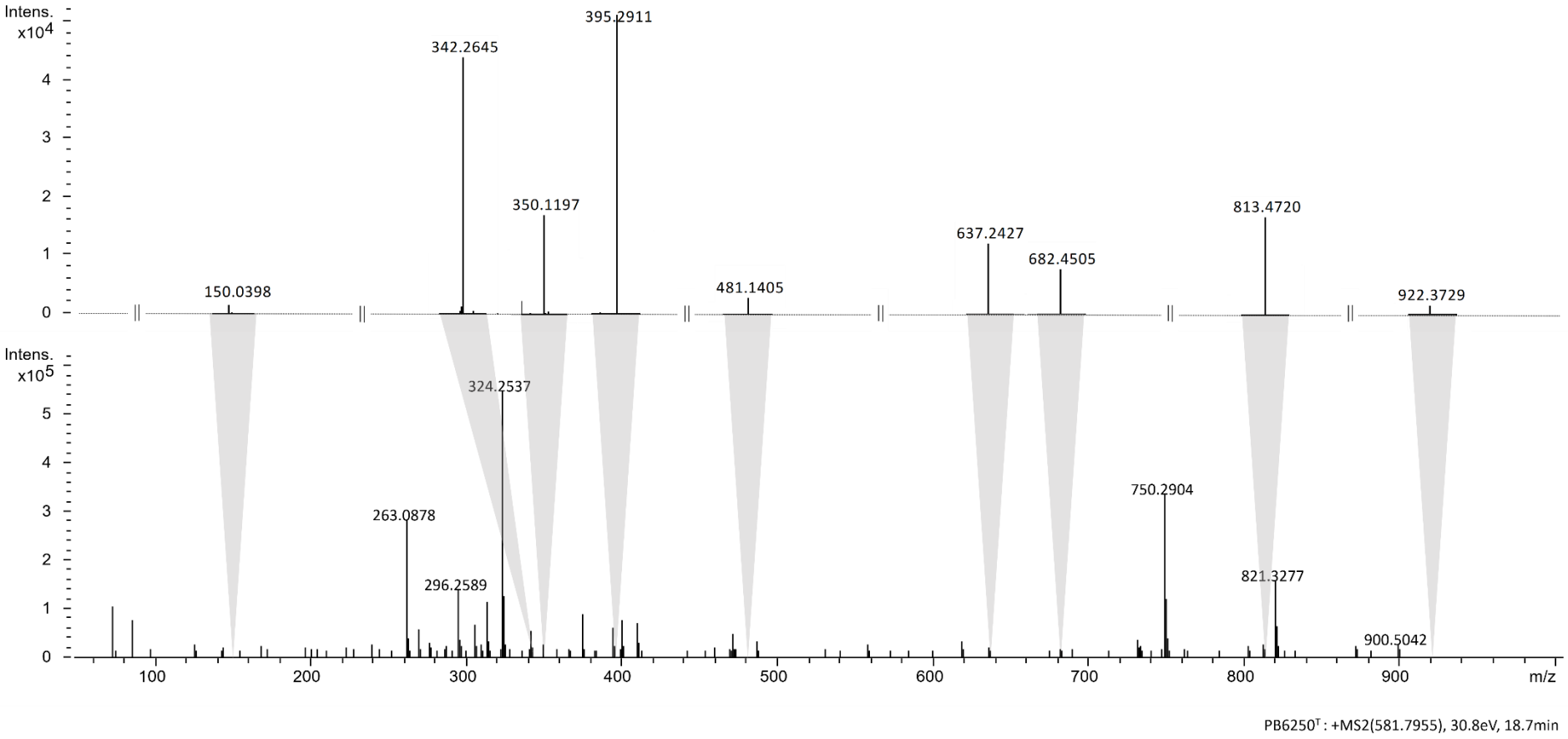


Figure 37: Annotated MS² spectrum of plusbacin A2

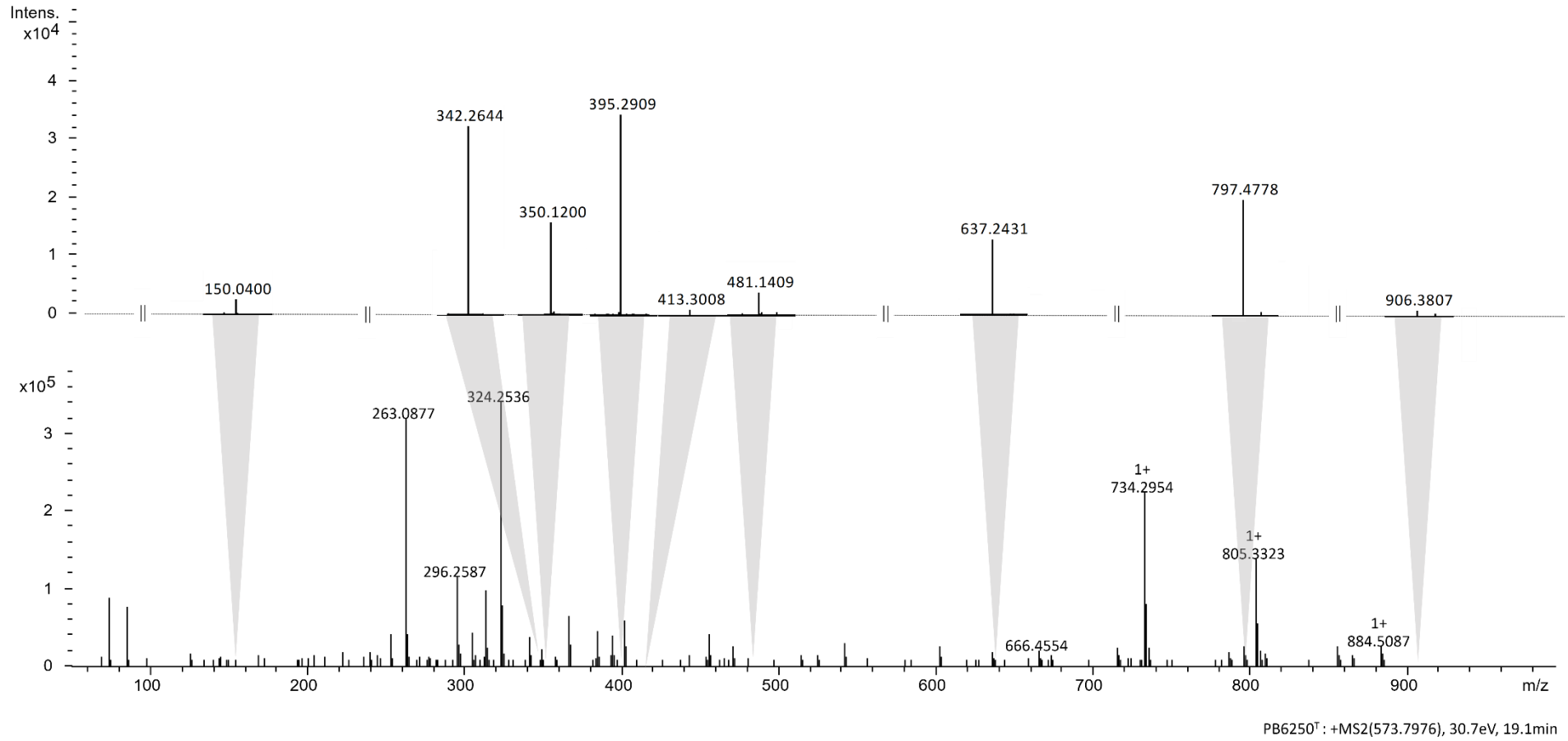


Figure 38: Annotated MS^2 spectrum of plusbacin B2

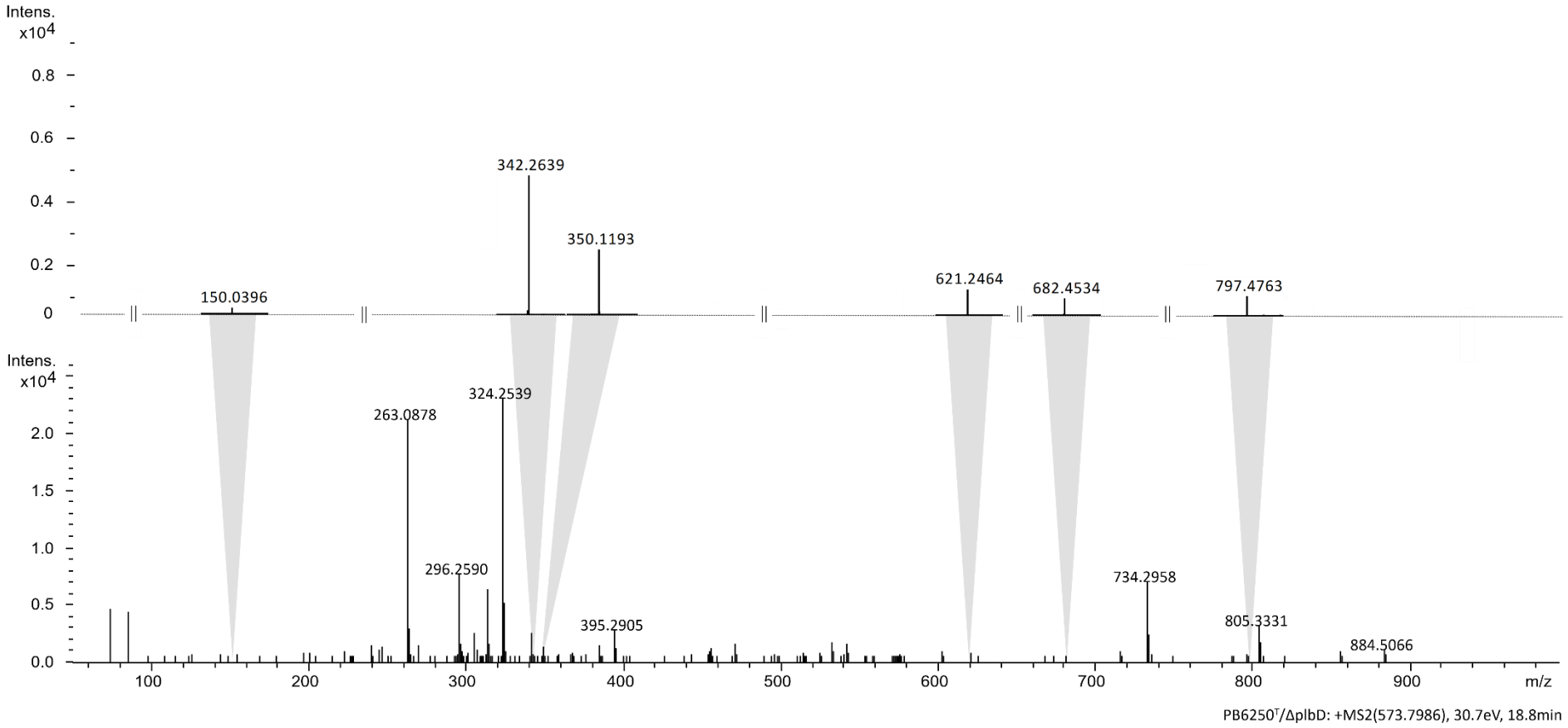


Figure 39: Annotated MS^2 spectrum of plusbacin derivative produced by ΔplbD strain

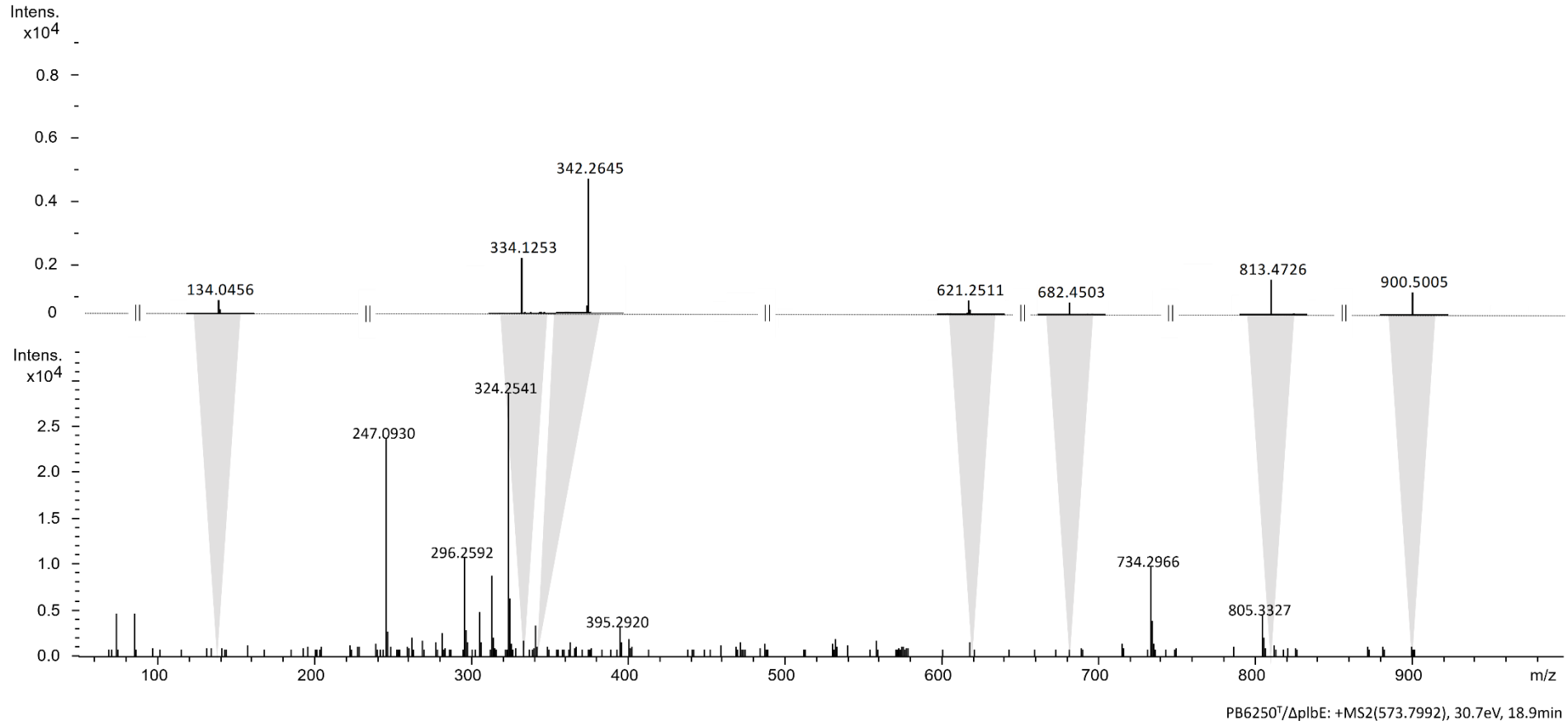


Figure 40: Annotated MS² spectrum of plusbacin derivative produced by ΔplbE strain

Table 33: Assigned fragments of plusbacin A2

Fragment	Ion Formula	Meas. *m/z*	Calc. *m/z*	Mass deviation
a_1 -H_2O	$C_{18}H_{35}NO_2$	296.2589	296.2590	Δ-0.337542 ppm
b_1	$C_{19}H_{36}NO_4$	342.2645	342.2644	Δ0.292172 ppm
b_1 -H_2O	$C_{19}H_{34}NO_3$	324.2537	324.2539	Δ-0.616801 ppm
b_2 -H_2O	$C_{22}H_{39}N_2O_4$	395.2911	395.2910	Δ0.252978 ppm
b_4	$C_{33}H_{60}N_7O_8$	682.4505	682.4503	Δ0.293062 ppm
b_5	$C_{37}H_{65}N_8O_{12}$	813.4720	813.4722	Δ-0.245860 ppm
b_6	$C_{40}H_{70}N_9O_{14}$	900.5042	900.5042	Δ0.000000 ppm
y_1	$C_4H_8NO_5$	150.0398	150.0402	Δ-2.665952 ppm
y_2	$C_9H_{15}N_2O_7$	263.0878	263.0879	Δ-0.380101 ppm
y_3	$C_{12}H_{20}N_3O_9$	350.1197	350.1200	Δ-0.856849 ppm
y_4	$C_{16}H_{25}N_4O_{13}$	481.1405	481.1418	Δ-2.701906 ppm
y_5	$C_{22}H_{37}N_8O_{14}$	637.2427	637.2429	Δ-0.313852 ppm
y_6	$C_{27}H_{44}N_9O_{16}$	750.2904	750.2906	Δ-1.770404 ppm
y_7	$C_{30}H_{49}N_{10}O_{17}$	821.3277	821.3280	Δ-0.365262 ppm
y_8	$C_{34}H_{56}N_{11}O_{19}$	922.3729	922.3754	Δ-2.710393 ppm

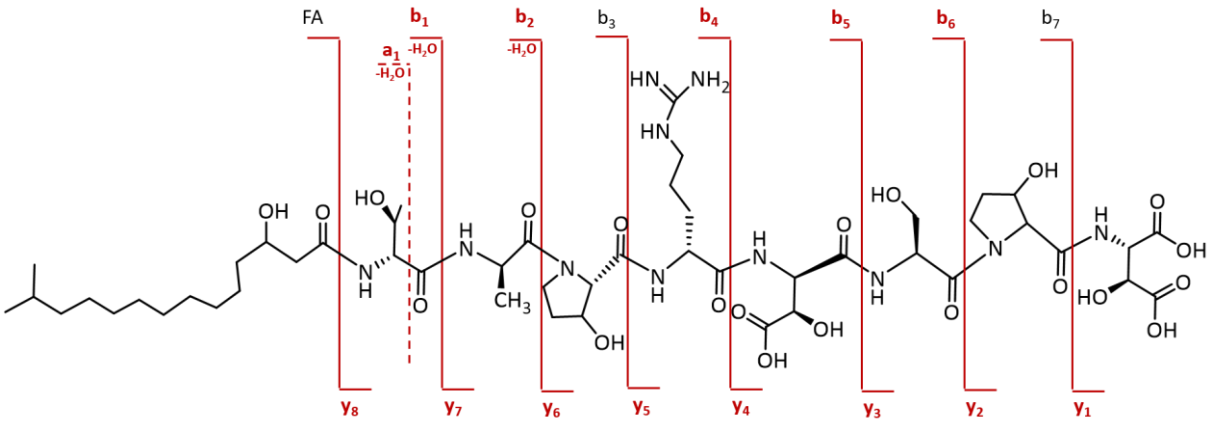


Figure 41: Schematic fragmentation of plusbacin A2; identified fragments are indicated in red

Table 34: Assigned fragments of plusbacin B2

Fragment	Ion Formula	Meas. *m/z*	Calc. *m/z*	Mass deviation
a_1 -H_2O	$C_{18}H_{35}NO_2$	296.2587	296.2590	Δ-1.012627 ppm
b_1	$C_{19}H_{36}NO_4$	342.2644	342.2644	Δ0.000000 ppm
b_1 -H_2O	$C_{19}H_{34}NO_3$	324.2536	324.2539	Δ-0.925201 ppm
b_2 -H_2O	$C_{22}H_{39}N_2O_4$	395.2909	395.2910	Δ-0.252978 ppm
b_2	$C_{22}H_{41}N_2O_5$	413.3008	413.3015	Δ-1.693679 ppm
b_4	$C_{33}H_{60}N_7O_7$	666.4554	666.4554	Δ0.000000 ppm
b_5	$C_{37}H_{65}N_8O_{11}$	797.4778	797.4773	Δ0.626977 ppm
b_6	$C_{40}H_{70}N_9O_{13}$	884.5087	884.5093	Δ-0.678342 ppm
y_1	$C_4H_8NO_5$	150.0400	150.0402	Δ-1.332976 ppm
y_2	$C_9H_{15}N_2O_7$	263.0877	263.0879	Δ-0.760202 ppm
y_3	$C_{12}H_{20}N_3O_9$	350.1200	350.1200	Δ0.000000 ppm
y_4	$C_{16}H_{25}N_4O_{13}$	481.1409	481.1418	Δ-1.870550 ppm
y_5	$C_{22}H_{37}N_8O_{14}$	637.2431	637.2429	Δ0.313852 ppm
y_6	$C_{27}H_{44}N_9O_{15}$	734.2954	734.2957	Δ-0.408555 ppm
y_7	$C_{30}H_{49}N_{10}O_{16}$	805.3323	805.3328	Δ-0.620861 ppm
y_8	$C_{34}H_{56}N_{11}O_{18}$	906.3807	906.3805	Δ0.220658 ppm

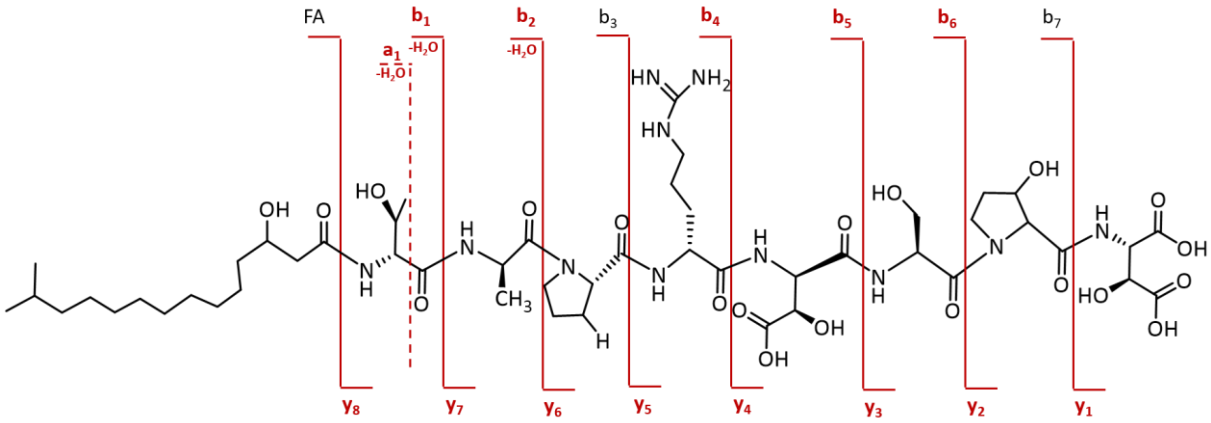


Figure 42: Schematic fragmentation of plusbacin B2; identified fragments are indicated in red

Table 35: Assigned fragments of a plusbacin derivative produced by ΔplbD strain

Fragment	Ion Formula	Meas. *m/z*	Calc. *m/z*	Mass deviation
a_1 -H_2O	$C_{18}H_{35}NO_2$	296.2590	296.2590	Δ0.000000 ppm
b_1	$C_{19}H_{36}NO_4$	342.2639	342.2644	Δ-1.460859 ppm
b_1 -H_2O	$C_{19}H_{34}NO_3$	324.2539	324.2539	Δ0.000000 ppm
b_2 -H_2O	$C_{22}H_{39}N_2O_4$	395.2905	395.2910	Δ-1.264891 ppm
b_4	$C_{33}H_{60}N_7O_8$	682.4534	682.4503	Δ4.542455 ppm
b_5	$C_{37}H_{65}N_8O_{11}$	797.4763	797.4773	Δ-1.253954 ppm
b_6	$C_{40}H_{70}N_9O_{13}$	884.5066	884.5093	Δ-3.052540 ppm
y_1	$C_4H_8NO_5$	150.0396	150.0402	Δ-3.998928 ppm
y_2	$C_9H_{15}N_2O_7$	263.0878	263.0879	Δ-0.380101 ppm
y_3	$C_{12}H_{20}N_3O_9$	350.1193	350.1200	Δ-1.999315 ppm
y_5	$C_{22}H_{37}N_8O_{13}$	621.2472	621.2480	Δ-1.287731 ppm
y_6	$C_{27}H_{44}N_9O_{15}$	734.2958	734.2957	Δ0.136185 ppm
y_7	$C_{30}H_{49}N_{10}O_{16}$	805.3331	805.3328	Δ0.372517 ppm

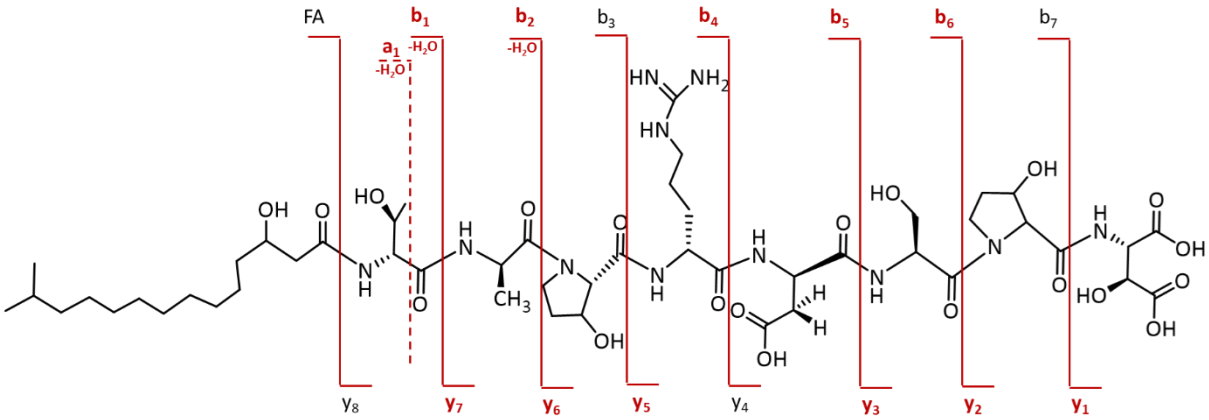


Figure 43: Schematic fragmentation of a plusbacin derivative produced by ΔplbD strain; identified fragments are indicated in red

Table 36: Assigned fragments of a plusbacin derivative produced by ΔplbE strain

Fragment	Ion Formula	Meas. *m/z*	Calc. m/z	Mass deviation
a_1 -H_2O	$C_{18}H_{35}NO_2$	296.2592	296.2590	Δ0.675085 ppm
b_1	$C_{19}H_{36}NO_4$	342.2645	342.2644	Δ0.292172 ppm
b_1 -H_2O	$C_{19}H_{34}NO_3$	324.2541	324.2539	Δ0.616801 ppm
b_2 -H_2O	$C_{22}H_{39}N_2O_4$	395.2920	395.2910	Δ2.529782 ppm
b_4	$C_{33}H_{60}N_7O_8$	682.4503	682.4503	Δ0.000000 ppm
b_5	$C_{37}H_{65}N_8O_{12}$	813.4726	813.4722	Δ0.491719 ppm
b_6	$C_{40}H_{70}N_9O_{14}$	900.5005	900.5042	Δ-4.108809 ppm
y_1	$C_4H_8NO_4$	134.0456	134.0453	Δ2.238049 ppm
y_2	$C_9H_{15}N_2O_6$	247.0930	247.0930	Δ0.000000 ppm
y_3	$C_{12}H_{20}N_3O_8$	334.1253	334.1250	Δ0.897868ppm
y_5	$C_{22}H_{37}N_8O_{13}$	621.2511	621.2480	Δ4.989956 ppm
y_6	$C_{27}H_{44}N_9O_{15}$	734.2966	734.2957	Δ1.225664 ppm
y_7	$C_{30}H_{49}N_{10}O_{16}$	805.3327	805.3328	Δ-0.124172 ppm

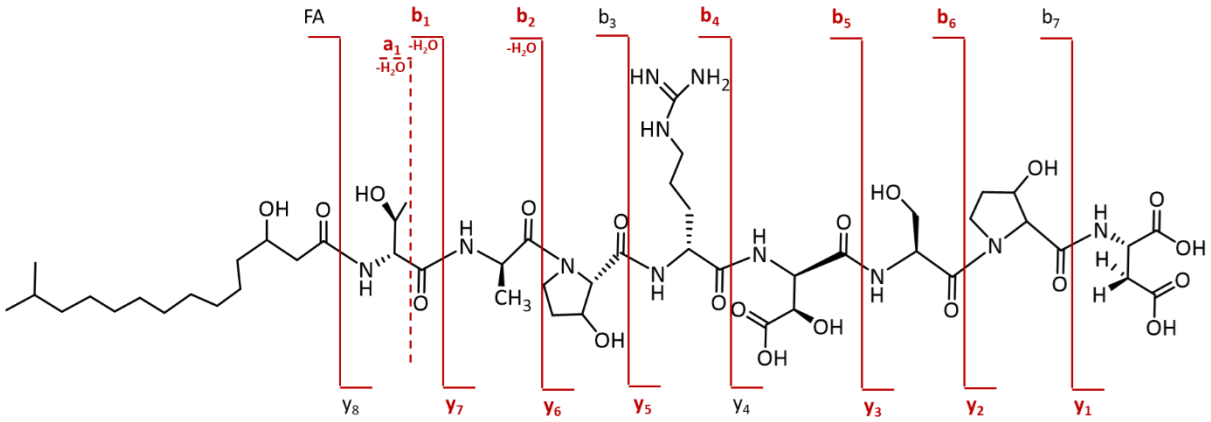


Figure 44: Schematic fragmentation of a plusbacin derivative produced by ΔplbE strain; identified fragments are indicated in red

2.3. Studies on A Domains of the Plusbacin Gene Cluster

Tandem mass spectrometry allowed the identification of aspartic acid as target of the dioxygenase gene products encoded in the plusbacin gene cluster. However, the timing of the hydroxylation process within the NRPS context remains unclear. In order to get to the bottom of this question, the respective A domains with a predicted specificity for either Pro/Hyp or Asp/Hya (Figure 45) were heterologously expressed and examined for their specificity in cooperation with the Hajo Kries' working group, in particular Maximillian Müll.

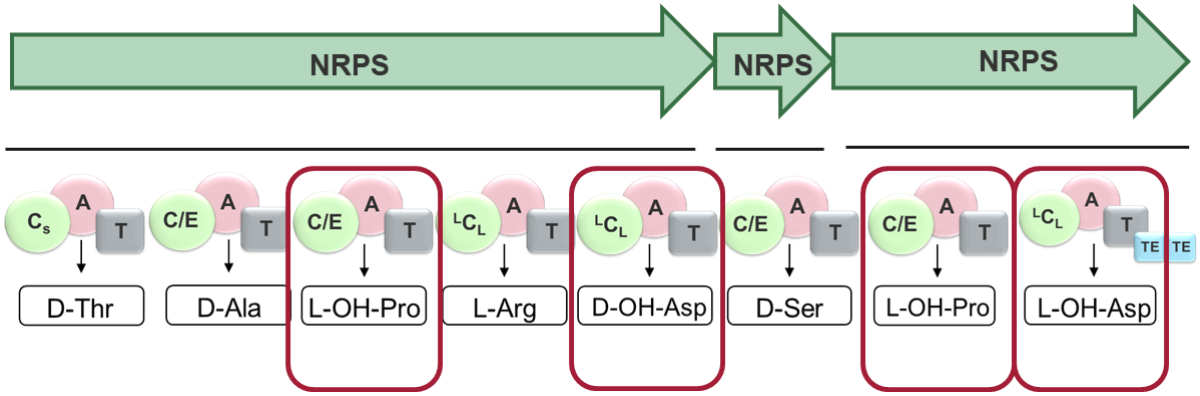


Figure 45: Investigated A domains and their location within the NRPS assembly line of plusbacin

Therefore, expression vectors were designed using pET-28a(+) and either chemically synthesized or generated by amplification via PCR followed by Gibson Assemby. Heterologous expression of the A domains was achieved under suitable conditions after transforming *E. coli* BL21 (DE3) with the expression plasmid. To test whether the method of purification was appropriate, the A domains were first heterologously expressed on a small scale. A representative SDS gel of PlbA8 is shown below.

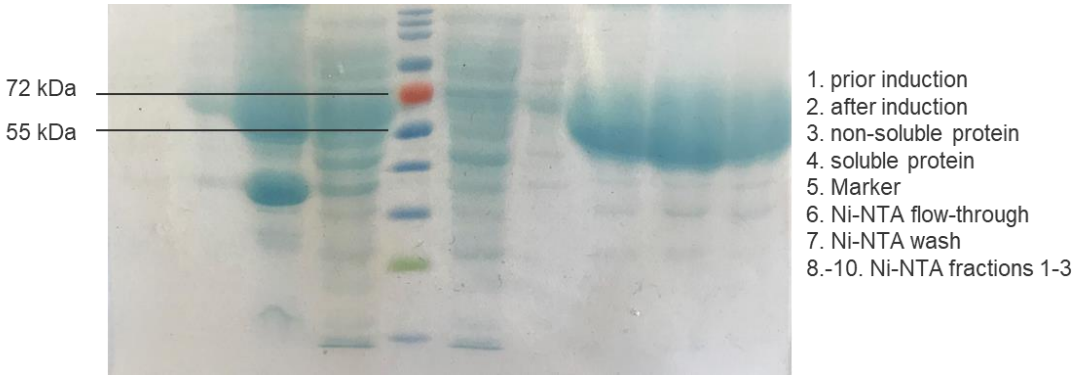


Figure 46: SDS page of PlbA8 (expected MW 65.1 kDa) after small scale heterologous expression

To determine the specificity, heterologous expression of the A domains was performed on a large scale prior to testing using MesG/hydroxylamine assay[63, 64] and hydroxamate assay (HAMA)[69].

One hallmark of the HAMA is the detection of adenylation activity under competitive conditions, since it allows parallel testing of the amino acids. Therefore, the results reflect the real circumstances during protein production on an NRPS assembly line in a cell best. Noteworthy is that synthetic hydroxamate standards are necessary for successful HAMA performance which were not available for all amino acids including Hyp and Hya. Nevertheless, HAMA was executed to test the adenylation activity of the A domains for the proteinogenic amino acids expect for serine and asparagine (Figure 47). Although only PlbA3 and PlbA8 resulted in a signal of the HAMA assay, the yielded results agree with the *in silico* prediction of the specificity of the respective A domains. Other proteinogenic amino acids were not activated.

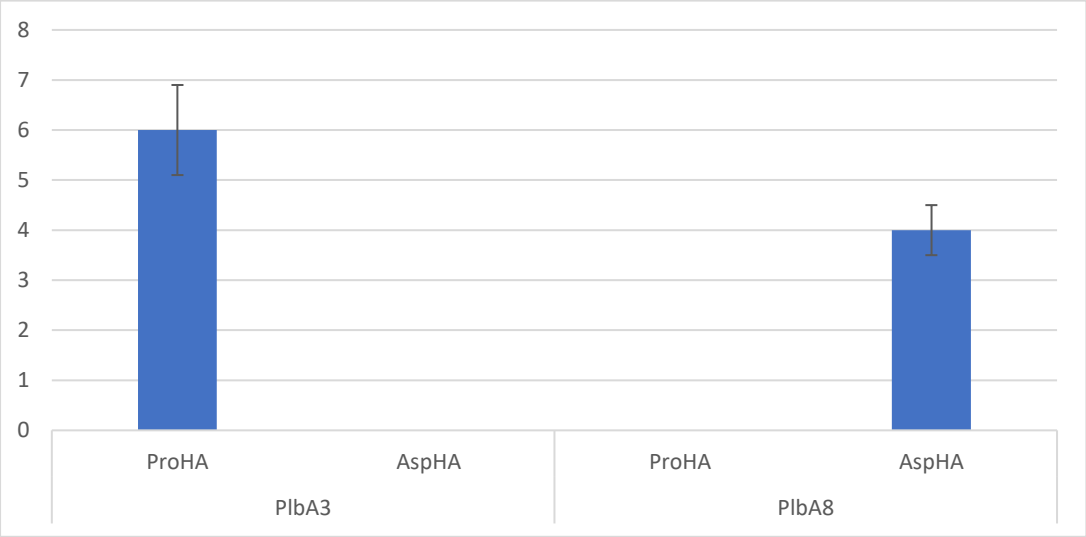


Figure 47: Detected hydroxamate concentration with the UPLC-tQ-MS after testing adenylation activation of A domains of interest using HAMA

To differentiate whether the A domains in question prefer the canonical amino acid or the hydroxylated form as its substrate, an MesG/hydroxylamine assay was performed to detect the saturation kinetics. For PlbA3, Michaelis-Menten kinetics were generated with Pro and Hyp whereas the ones for PlbA8 were created with Asp and Hya (Table 37, Figure 48). The specificity constants (k_{cat}/K_M) which were recorded with the canonical amino acid for PlbA3 and PlbA8 are within the typically observed range of catalytic parameters generated with this assay. An about 10-fold lower activity of PlbA8 was observed for Hya under the same assay conditions,

whereas no activity at all was recorded for PlbA3 with Hyp. Notable is the poor expression and difficulties while purification of PlbA3 (Figure 49) which might result in the underestimation of the catalytic parameters.

*Table 37: Kinetic parameters for PlbA3 and PlbA8 determined for the substrates Pro, Hyp, Asp and Hya**

	PlbA3		PlbA8	
	Pro	Hyp	Asp	Hya
k_{cat} [min^{-1}]	0.20 ± 0.02	n.d.	3.3 ± 0.4	n.d.[#]
K_M [mM]	0.09 ± 0.02	n.d.	4.1 ± 0.8	n.d.[#]
k_{cat}/K_M [min^{-1} mM^{-1}]	2.4 ± 0.7	n.d.	0.8 ± 0.5	0.08 ± 0.04

*n.d.: not detectable. Error margins indicate the error of fit. #Not determined because substrate saturation could not be reached.

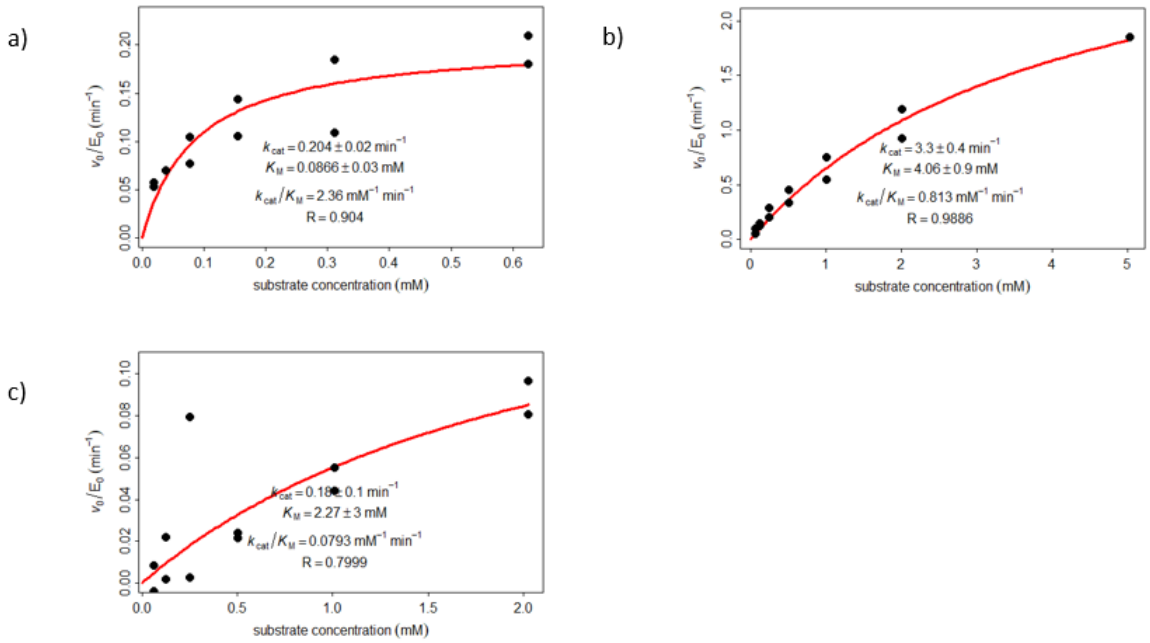


Figure 48: Michaelis-Menten kinetics with a) PlbA3 and Pro, b) PlbA8 and Asp, c) PlbA8 and Hyp, measured with the MesG/hydroxylamine assay in two biological replicates

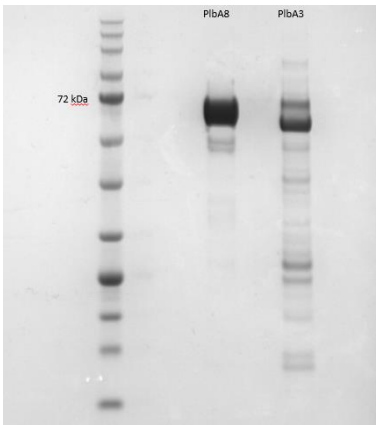


Figure 49: SDS Page of PlbA8 (left; expected MW 65.1 kDa) and PlbA3 (right; expected MW 67.1 kDa)

IV. Discussion

Infectious diseases, especially those caused by antibiotic-resistant bacteria, emensly burden global health.[121] Therefore, the constant discovery and development of antimicrobial drugs displaying novel modes of action is crucial. Since most clinically used antibiotics nowadays are natural products or derivatives thereof, originally isolated from actinomycetes and fungi,[122] new soures must be discovered.

Lately, the attention has been drawn to underexplored bacterial genera like *Lysobacter*.[123] Not only the guanidine-containing cyclic lipopeptides tripropeptin and plusbacin investigated in this study are produced by *Lysobacter* spp. as secondary metabolites, but also numerous other pharmaceutically and structurally interesting compounds like WAP-8294A2[124] and hypeptin.[125] To unveil the complete biosynthetic potential of this genus, genome sequencing is a powerful tool. Within a genome of about ~ 6 Mbp, *Lysobacter* spp. encode on average for 12 to 16 secondary metabolites. Among those are bacteriocins, siderophores, linaridins, arylpolyenes, terpenes, phenazines, and resorcinols, but the majority of the biosynthetic gene cluster seem to encode NRPS as well as NRPS hybrids and lantipeptides.[126] Therefore, the genus *Lysobacter* represents rather a peptide production specialist unlike actinomycetes or fungi, which exhibit a larger chemical diversity. This goes along with the characteristics of the tripropeptin producer *Lysobacter* sp. strain BMK333-48F3.[110] The genome size of 5.2 Mbp is on the smaller whereas the G+C content of 69.7 % is on the higher side but is still within the typical range of this genus, which normally displays G+C ratios of 61 % to 70 %. Harboring a total of 12 biosynthetic gene cluster, including many encoding NRPS or NRPS-like products, BMK333-48F3 displays a potential for secondary metabolite production comparable to other *Lysobacter* spp. like *L. antibioticus* strain 76, *L. capsici* strain 55 and *L. gummosus* strain 3.2.11.[126, 127]

One hallmark of *Lysobacter* sp. strain BMK333-48F3 is its ability to produce tripropeptin. Annotation of the assembled genome of BMK333-48F3 allowed the identification of the respective biosynthetic gene cluster. *In silico* analysis supported this finding. A domain prediction using NRPSPredictor2 and the Stachelhaus code align with the previously reported peptide core sequence of tripropeptin[44-48]. Moreover, the genetic architecture matches the previously reported tripropeptin gene cluster of *Collimonas*[128], as well as the ones of

empedopeptin[129] and plusbacin, which was experimentally verified in this study. The respective gene clusters comprise three NRPS genes followed by two dioxygenases. Subdivided into eight modules, the NRPS assembly line is flanked by a $C_{Starter}$ domain at the beginning and a tandem TE domain at the end.

Since this group of guanidine containing cyclic lipopeptides contain a remarkable amount of non-proteinogenic amino acids including several hydroxyamino acids, the presence of oxygenases within the gene cluster is not surprising. Nevertheless, empedopeptin, tripropeptin and plusbacin have two hydroxyaspartic acid and at least one hydroxylated proline embedded in their peptide backbone. The presence of two dioxygenases within the biosynthetic gene cluster does therefore not clearly explain the hydroxylation mechanism. To clarify the target of the dioxygenases two hypotheses were taken into consideration:

I. Are the dioxygenases promiscuous or do they display a clear specificity for either proline or aspartic acid?
II. If the dioxygenases are specific to one of the amino acids, are they able to substitute each other or only catalyze the hydroxylation reaction at a specific position of the peptide backbone?

Hydroxylation of proline as well as aspartic acid have previously been reported in the literature.

In mammalian cells, hydroxyproline is a key component of collage as it is crucial for its stability. The hydroxylation process is catalyzed posttranslationally by the enzyme prolyl hydroxylase. As a member of alpha-ketoglutarate-dependent hydroxylases, the enzyme acts in a similar manner as other dioxygenases and therefore requires alpha-ketoglutaric acid, Fe^{2+} and ascorbate.[130]

The hydroxylation of proline was also described in various natural products of fungi and bacteria. The fungal secondary metabolite pneumocandin B_0 (Figure 50A) is produced by *Glarea lozoyensis* and harbors hydroxyproline. *Trans*-4- and *trans*-3-hydroxyproline, as well as *trans*-3-hydroxy-4-methylproline are necessary for the biosynthesis of pneumocandin and are provided by only one enzyme (Figure 50C). The nonheme mononuclear iron oxygenase GLOXY2 (GloF) converts free proline to the different hydroxylated forms.[131-133]

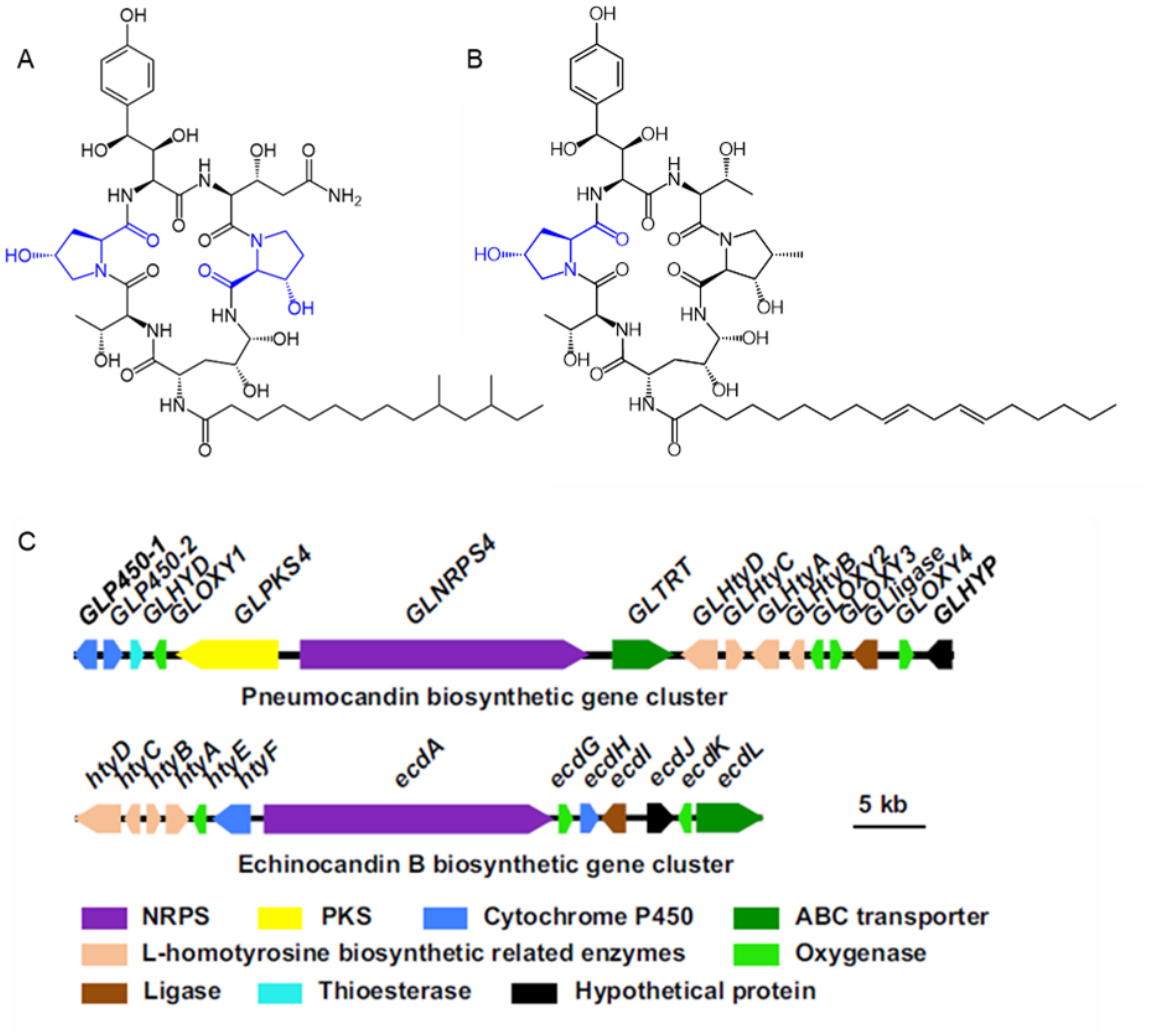


Figure 50: Chemical structures of pneumocandin B_0 (A) and echinocandin B (B) and organization of the respective biosynthetic gene clusters (C) adapted from Li et al.[133]

Another fungal secondary metabolite harboring hydroxyproline is echinocandin (Figure 50B). Genome sequencing of *Aspergillus pachycristatus* revealed the BGC (Figure 50C) including a gene coding for a 2-oxoglutarate-dependent proline hydroxylase. HtyE generates either 4-hydroxyproline or 3-hydroxy-methyl-proline using proline or methyl-proline as substrates, respectively. Therefore, HtyE seems to be the homologous protein of GloF in pneumocandin biosynthesis.[133, 134]

Besides producing echinocandin, fungi of the genus *Aspergillus* are also known to produce a wide variety of other bioactive metabolites, one of these are the burnettramic acids (Figure 51A). This unusual class of bolaamphiphilic pyrrolizidinediones is especially striking since they exhibit 4-hydroxyproline as part of a bicyclic pyrrolizidinedione in their structure. The hydroxylation of proline seems

to be catalyzed by BuaE encoded in the BGC of burnettramic acids in *A. burnettii* (Figure 51B).[135]

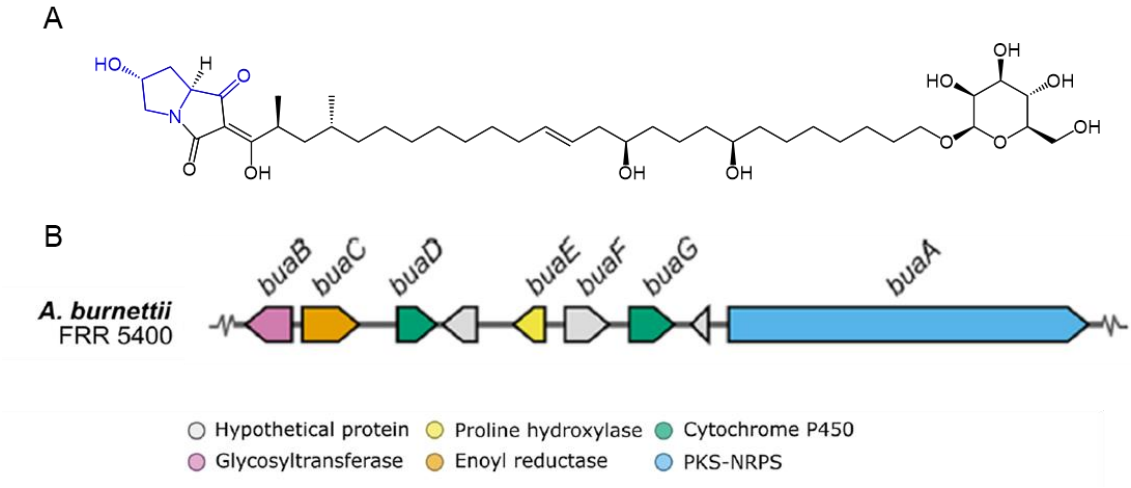


Figure 51: Chemical structure of burnettramic acid A (A) and organization of the respective biosynthetic gene cluster (B) adapted from Li et al.[135]

Bearing two hydroxyprolines, cicadapeptins (Figure 52) are secondary metabolites generated by *Cordyceps heteropoda* ARSEF1880. Although the gene cluster is not known, one could speculate that the hydroxation process of proline works in a similar manner as prior described. Thus, the gene cluster would harbor only one gene encoding a proline hydroxylase directly adjacent to the other genes needed for biosynthesis.

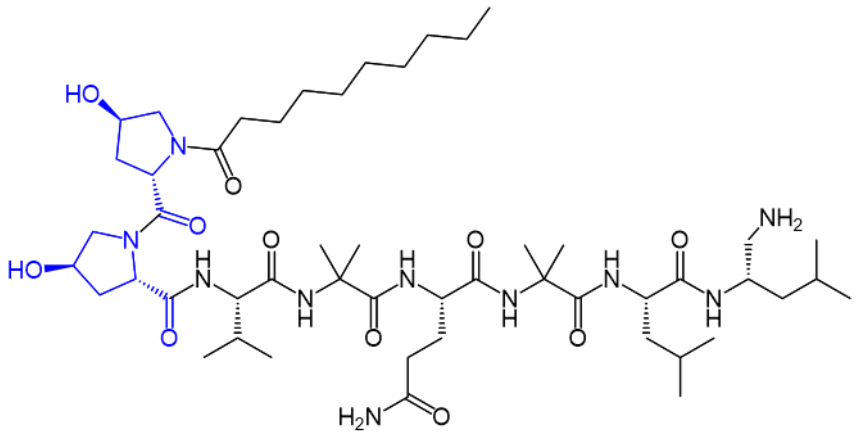


Figure 52: Chemical structure of cicadapeptin I

Due to the compounds from fungal origin, their biosynthesis may not be so easily applicable to those of prokaryotes, since a sequence analysis indicates that the eukaryotic hydroxylase is not associated to bacterial enzymes catalyzing the hydroxylation of proline.[131, 132]

Recently, the BGC (Figure 53B) coding for heinamides (Figure 53A) was discovered in *Nostoc* sp. UHCC 0702. Harboring non-proteinogenic amino acids, several enzymes supplying those were identified. Hydroxyproline is provided by LxaN, which catalyzes the hydroxylation of L-proline. Noteworthy is the locus of *lxaN* in the genome, since it is located approximately 389 kb apart from the *lxa* gene cluster. Unlike fungi, cyanobacteria are prokaryotes and could therefore serve as a better model, indicating that bacterial proline hydroxylases might not cluster within but rather up- or downstream of the BGC.[136]

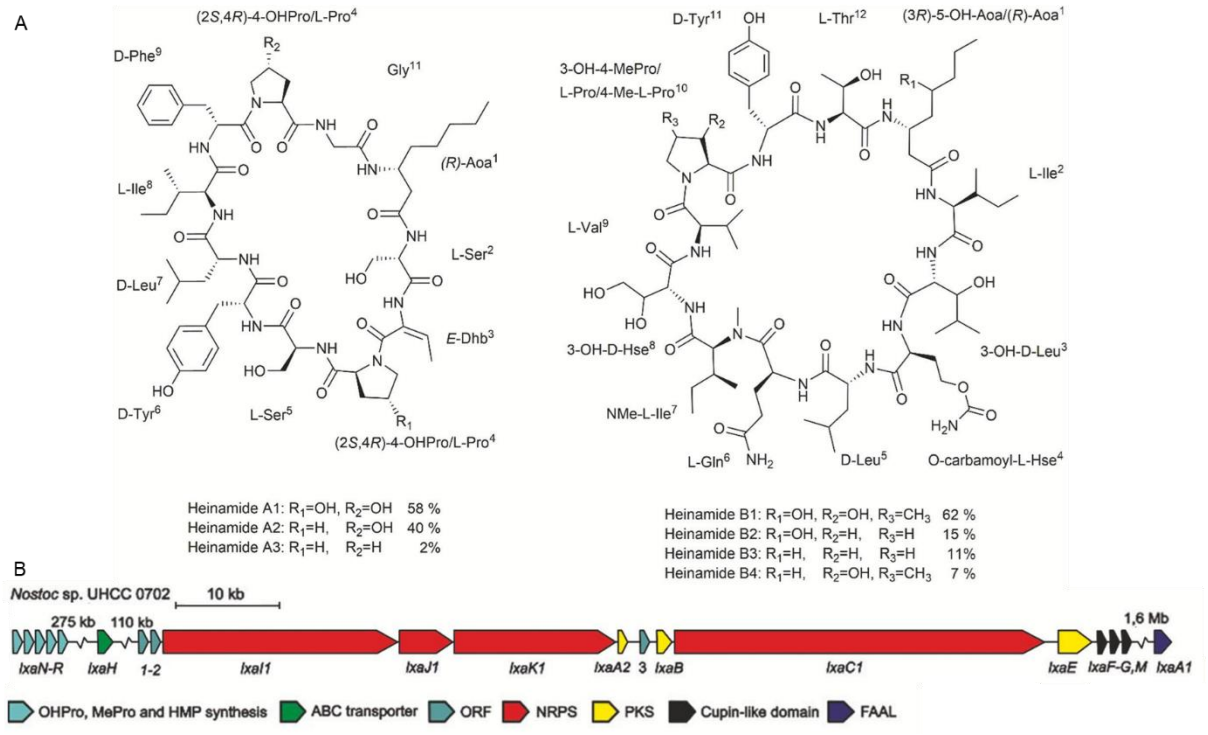


Figure 53: Chemical structure of heinamides (A) and organization of the respective biosynthetic gene clusters (B) adapted from Heinilä et al.[136]

An example for bacterial secondary metabolites containing hydroxyproline is etamycin (Figure 54).

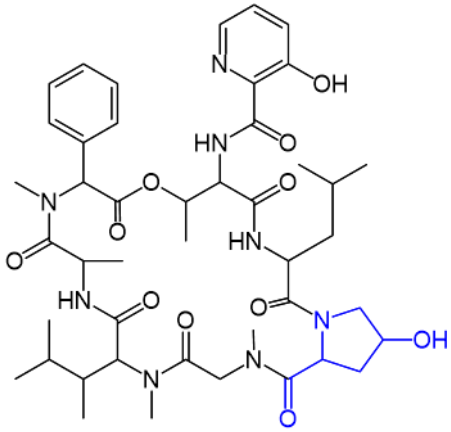


Figure 54: Chemical structure of etamycin

As an iron- and 2-oxoglutarate-dependent dioxygenase, proline 4-hydroxylase is responsible for the turnover of proline to hydroxyproline during the biosynthesis in *Streptomyces griseoviridus* P8648. The enzyme was purified from crude cell-free extract and characterized.[137, 138] Although its genetic locus is not known, a 4-hydroxylase enzyme does not seem to be encoded in the BGC of etamycin.[139]

According to the other secondary metabolites bearing hydroxyproline, it seems likely that only one gene and therefore one resultant enzyme is responsible for the hydroxylation of all prolines during biosynthesis of the guanidine containing cyclic lipopeptides. In principle, it would be possible that the hydroxylation process of the prolines during the biosynthesis of guanidine-containing cyclic lipopeptides is catalyzed by an enzyme located in the gene cluster, as is the case with the fungal secondary metabolites, but it seems much more likely that this is not the case. Although there is evidence in the literature that proline hydroxylases in bacteria might not be encoded in the gene cluster, it was also not possible to find a suitable enzyme not clustering with the other biosynthetic genes.

Not only hydroxylation of proline is a previously described feature of secondary metabolites, but also the occurrence of hydroxyaspartic acid. As previously described, serobactin (Figure 55) produced by *Herbaspirillum seropedicae*, for instance, carries two hydroxy-aspartyl residues. The BGC and proposed biosynthetic pathway suggest the presence of a TauD domain embedded in the NRPS assembly line directly succeeding the C domain as well as a stand-alone enzyme, which is an aspartyl β-hydroxylase not associated to a NRPS domain.[140]

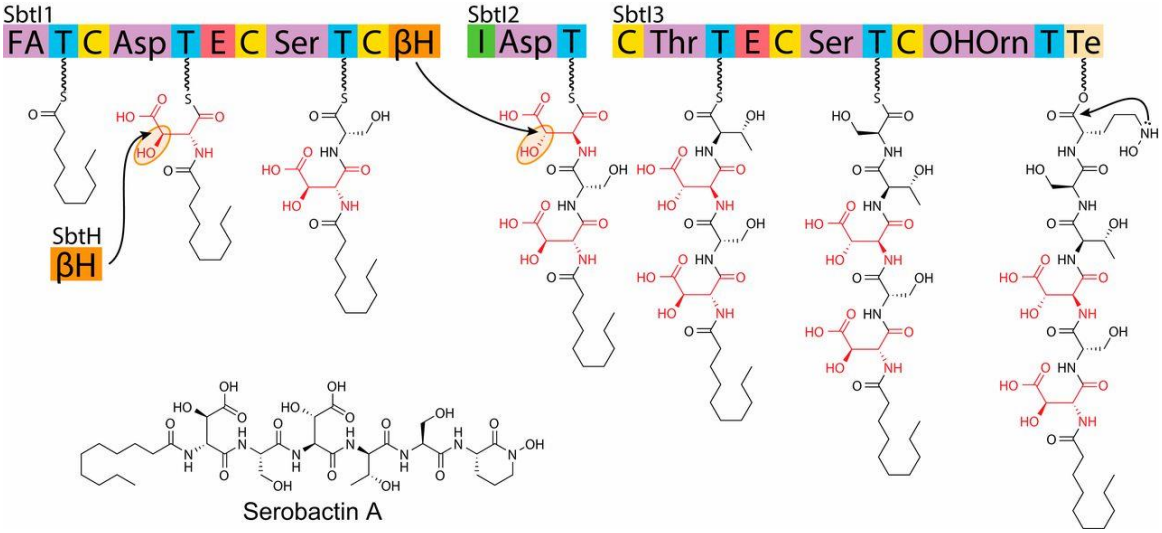


Figure 55: Proposed biosynthesis of serobactin adapted from Reitz et al.[141]

Another secondary metabolite displaying two hydroxyaspartic acids is cupriachelin (Figure 56A). Produced by *Cupriavidus necator*, the compound seems to be assembled in a similar manner as serobactin. During the biosynthesis of cupriachelin, the hydroxylation process of one aspartic acid is catalyzed by the C-terminal NRPS associated TauD domain of CucF whereas the other aspartyl residue gets hydroxylated by the dioxygenase CucE (Figure 56B).[142]

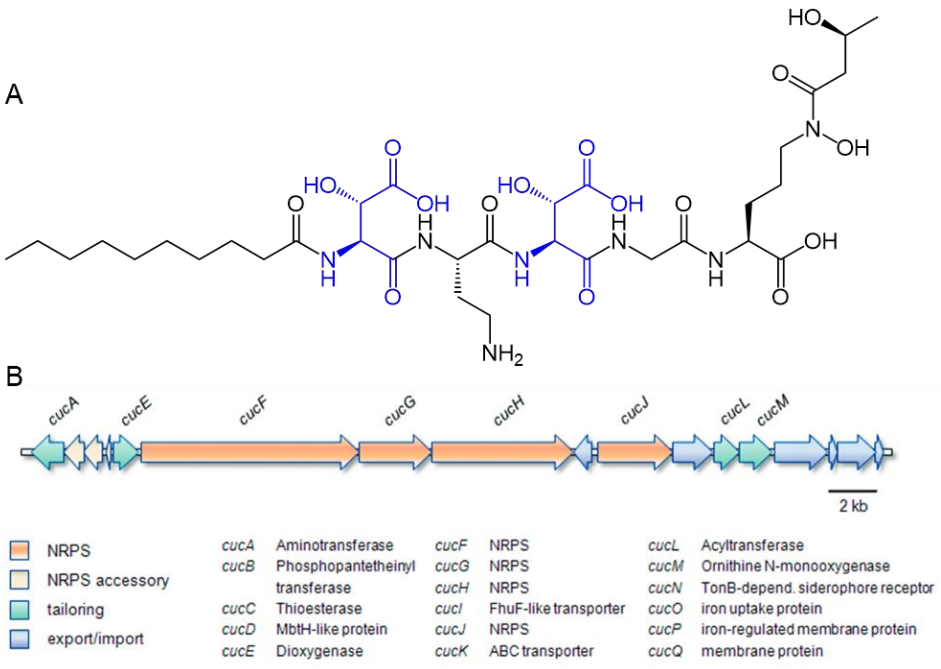


Figure 56: Chemical structure of cupriachelin (A) and organization of the respective biosynthetic gene cluster (B) adapted from Kreutzer et al.[142]

Another siderophore like serobactin and cupriachelin, carrying two hydroxy-aspartyl residues, is pacifibactin (Figure 57A). Containing the NRPS gene cluster coding for pacifibacin (Figure 57B), *Alcanivorax pacificus* harbors the stand-alone hydroxylase PfbF and a TauD domain embedded in the NRPS assembly in order to perform a hydroxylation of one aspartic acid each during synthesis.[143]

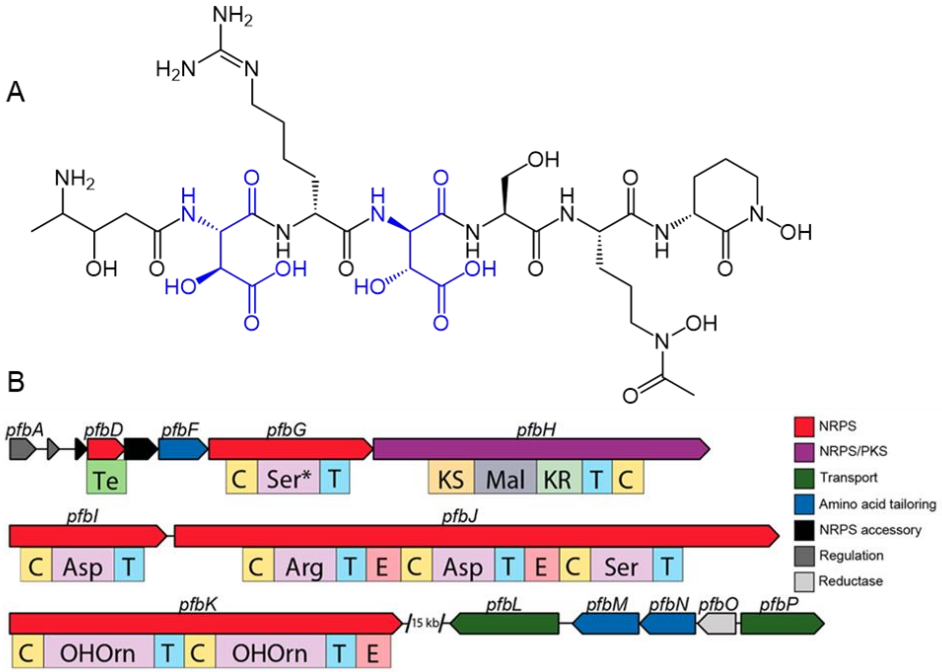


Figure 57: Chemical structure of pacifibacin (A) and organization of the respective biosynthetic gene clusters (B) adapted from Hardy et al.[143]

As not only NRPS-associated aspartyl β-hydroxylase domains are involved in the generation of secondary metabolites, some rely only on stand-alone enzymes for the hydroxylation of aspartic acid. During syringomycin production in *Pseudomonas syringae,* for instance, SyrP, a aspartyl hydroxylase, generates the L-*threo*-β-hydroxy-aspartic acid (Figure 58).[116]

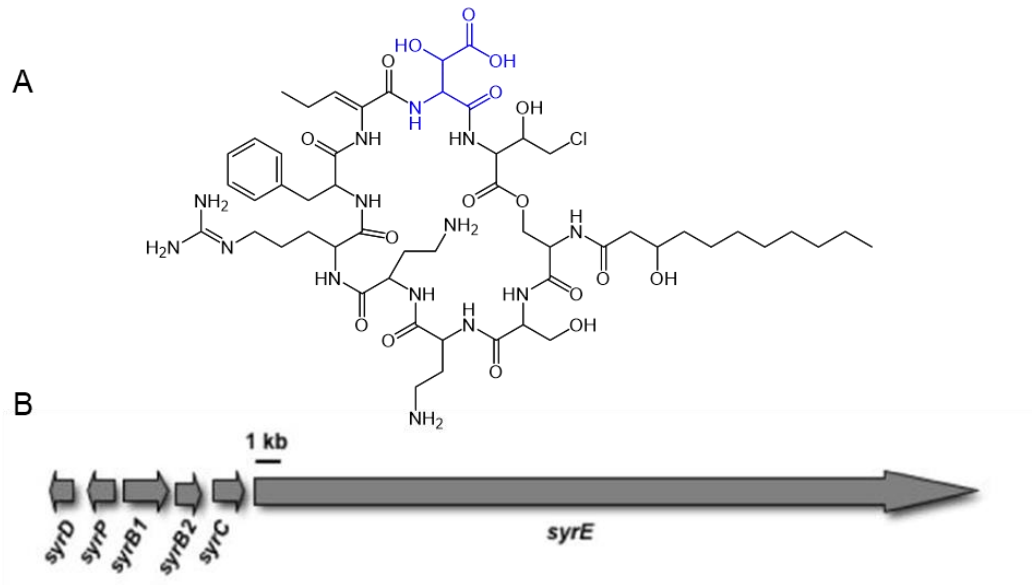


Figure 58: Chemical structure of syringomycin E (A) and organization of the respective biosynthetic gene cluster (B) adapted from Singh et al.[116]

Since syringomycin contains only one hydroxy aspartic acid, the BGC also contains only one aspartyl hydroxylase. However, during the biosynthesis of the aforementioned siderophores, it was shown that each hydroxylated aspartic acid has its own designated enzyme. However, these cases are not exactly comparable, as the BGCs of EMP, TPP and PLUS each contain two stand-alone enzymes, not an oxygenase and an NRPS-associated β-hydroxylase domain.

In silico analysis of PlbD and PlbE predicted them as putative dioxygenases belonging to the TauD/TfdA family. Since bioinformatics tools were not able to determine the target of the dioxygenases with certainty, experimental determination was crucial. Initially, each oxygenase gene was knocked out using the Red/ET system on a fosmid carrying the partial plusbacin gene cluster including *plbD* and *plbE*. The generated fosmid derivatives were supposed to create in-frame deletions in the wild type plusbacin producer strain *L. firmicutimachus* PB-6250^T via homologous recombination. Despite, screening the thousands of clones, no successful double crossover event could be identified. Since double crossover events seem to occur with a very low frequency in *L. firmicutimachus* PB-6250^T, the method was rendered impractical for the creation of unmarked gene deletions.

As genetic manipulation of *Lysobacter* species is still challenging due to a lack of a wide variety of knockout tools. A suicide vector containing a counter-selectable marker was chosen to design new in-frame deletion vectors. Suicide vectors known

to allow gene deletion in various *Lysobacter* spp. are pJQ200SK, pEX18GM and pMW91CM.[103, 124, 144, 145] To manipulate the plusbacin gene cluster pjQ200SK was chosen due to commercial availability, although all three suicide vectors contain the *sacB* gene as a counter-selectable marker to help generating clean knockouts derived from a double crossover event. Coding for levansucrase, the *sacB* gene from *Bacillus subtilis* promotes cell death due to the conversion of sucrose into levan, which has a lethal effect to gram-negative bacteria.[146] Thus, bearing *sacB* in the genome due to a single event of homologous or non-specific recombination *Lysobacter* transfomants display sucrose sensitivity. Therefore, the counter-selective marker was a powerful tool for positive selection of knock out mutants, which underwent successful double crossover.

Since LC-MS experiments suggested the loss of one hydroxy group associated with the deletion of each dioxygenase, the exact position still needed to be clarified using tandem mass spectrometry. To determine the amino acid sequence of a peptide, it must undergo fragmentation during MS/MS measurement. Within the peptide backbone, there are typically three different breaking points. The resulting a, b or c ions expand from the amino terminus, while the carboxy terminal fragments are called x, y and z ions (Figure 59).[147]

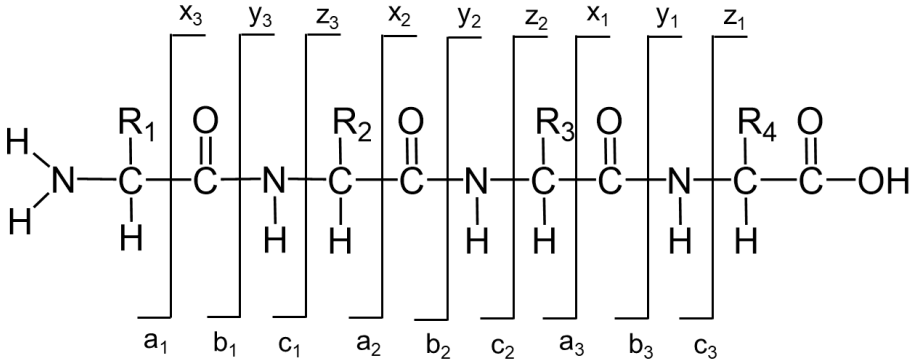


Figure 59: Nomenclature for fragmentation ions in tandem mass spectrometry of peptides

Most commonly, b and y ions are observed during peptide fragmentation allowing a sequence analysis starting at the amino or carboxy terminus. The MS^2 spectra resulting from tandem mass spectrometry of plusbacin or the derivatives produced by the knockout mutants display most of the b and y ions. The absence of some ions is easily explained since peptide fragmentation does not occur sequentially. The fragmentation process happens rather randomly meaning it does not start at

the amino or carboxy-terminus, respectively, proceeding one amino acid at the time along the peptide chain. Moreover, given a specific collision energy, some of the fragmentation events are preferred over others resulting in variations in the occurrence of observed peaks in the spectra or the absence of fragments altogether.[148]

Since most b and y ions were identified in the measured MS^2 spectra, pinpointing the missing hydroxylation caused by the in-frame deletions was possible. According to the MS/MS measurements, both dioxygenases show a specificity for aspartic acid. Although knocking out *plbD* results in a different metabolite profile than *plbE*, indicating different targets of the dioxygenases. Therefore, the dioxygenases are probably not able to stand in for each other but hydroxylate just one of the aspartic acids each in the structure of plusbacin as shown in Figure 60.

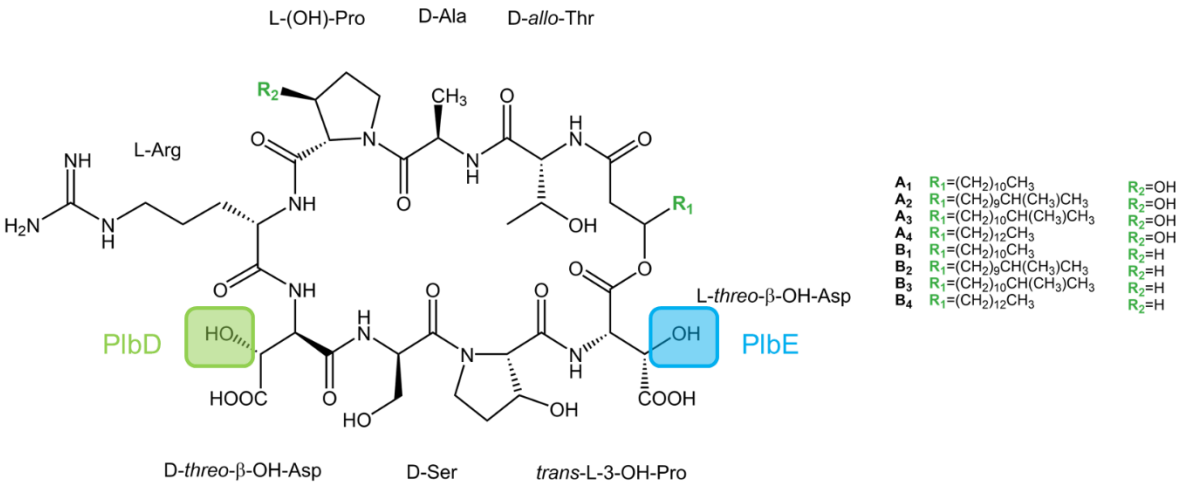


Figure 60: Target of PlbD and PlbE

As the dioxygenases each hydroxylate a different aspartic acid within the plusbacin core, the question arises as to how they differentiate between them. The aspartic acids are present in different configurations in the macrolactone core of plusbacin, so the hydroxylation process would likely occur as a post-translational modification making the dioxygenases stereoselective.

Hence, two hypotheses were considered concerning the timing of the hydroxylation process in the NRPS context.

I. Aspartic acid acts as a precursor and is hydroxylated by the dioxygenase prior to the incorporation of the amino acid into plusbacin. Therefore, the responsible A domain recognizes and activates hydroxy aspartic acid.

II. Aspartic acid is recognized and activated by the A domain in question during plusbacin production and the hydroxylation catalyzed by the dioxygenases occurs post-assembly.

To gain more insights into the timing of the hydroxylation of the aspartic acids, the respective A domains were expressed heterologously to perform adenylation activity assays. HAMA was preferred over other assays to determine the specificity of the A domains in question since it best reflects the situation prevailing in the cell. Other assays test one amino acid substrate at a time while HAMA allows parallel profiling representing competition conditions. Since synthetic hydroxamate standards are crucial for A domain specificity profiling using HAMA[69], the adenylation activity for Hya and Hyp could not be detected due to the lack of suitable hydroxamates. Hence, a MesG/hydroxylamine assay[63, 64] was performed additionally.

Only two of the tested A domains showed adenylation activity. PlbA3 activated proline, while PlbA8 displayed specificity for aspartic acid during HAMA testing. This agrees with the *in silico* prediction. To test whether the proteinogenic amino acid or the hydroxylated form is preferred by the A domains, saturation kinetics were recorded individually. Both A domains exhibited significantly lower or even no substrate activation under the same conditions for the respective hydroxy amino acid suggesting the canonical amino acid as the A domain substrate. Although the hydroxylated form may be activated by the A domain during *in vitro* testing, it is most likely not able to outperform the regular amino acid as a substrate within the cell.

The lack of activity in PlbA5 and PlbA7 might be due to obtaining the protein in an inactive form caused by poor border determination of the A domains. In order to identify functional A domains, sequence alignment with other functional adenylation domains assisted by suitable software tools is mandatory.

Another reason for poor production rates and the lack of adenylation activity of the heterologously expressed A domains might be the need to co-express an MbtH-

like protein. The superfamily of MbtH-like proteins is named after a small protein found in the gene cluster encoding for the biosynthesis of the siderophore mycobactin in *Mycobacterium tuberculosis*.[149] It turned out that *mbtH* homologs appear in various NRPS gene cluster. It is known that co-expression of MbtH-like proteins with NRPS derived domains can activate their enzymatic activity *in vitro*.[150, 151] It is not yet fully understood why some A domains require MbtH-like proteins and others do not, but the nature of some amino acid residues, which does not have direct contacts with the substrates, appear to have an influence on the interaction. It was demonstrated that the exchange of Leu383 for Met in CloH, which is a tyrosine-adenylating enzyme involved in biosynthesis of clorobiocin, resulted in a MbtH-like protein independence of the A domain. It is postulated that the interaction between an A domain and a MbtH-like protein causes conformational changes that lead to an increase in activity.[152] Although no *mbtH* gene could be identified in the biosynthetic gene cluster of plusbacin, an *mbtH* homolog was found elsewhere in the genome of PB-6250^T. Therefore, it cannot be ruled out that co-expression might help the production rate as well as trigger activity of the A domains.

Although not all A domains in question could be analyzed *in vitro*, it is still reasonable to postulate, that hydroxylation occurs post-assembly. Following the recruitment of the amino acid by the A domains, the dioxygenases PlbD and PlbE act on aspartic acid which is linked to the T carrier protein domain by a pantetheinyl thioester. Hence, the hydroxylation reaction of aspartic acids executed by PlbD and PlbE is similar to SyrP involved in the biosynthesis of pseudomonal syringomycin E.[116] 2-ketoglutarate dependent dioxygenase which catalyze a β-hydroxylation reaction on T domain bound amino acids are also known to be present in *Lysobacter*. Recently HynE was identified to hydroxylate asparagine during hypeptin biosynthesis acting on the amino acid covalently linked to the T domain.[125]

Plusbacin A_1 and B_1 in the wild type strain seem to be produced at the lowest concentration. The associated peak shows by far the lowest peak area, which relates to the concentration in the sample. LC-MS measurements of the mutant extracts lack a peak correlating to plusbacin derivatives harboring the shortest fatty acid side chain. The absence a mass of ~1115 Da might be due to a significant

drop in the production rate in the in-frame deletion mutants. This raises the question if the poor production rates of the compounds produced by the knockout mutants is just due to the genetic engineering or indicates that the hydroxylation process catalyzed by the dioxygenases might be a bottleneck within its biosynthesis. Thus, overexpression of PlbD and PlbE should provide more insights. Plusbacin production rates of the overexpression mutants were measured in a relative manner compared to the wild type producer strain. Overexpression of each of the dioxygenase genes, *plbD* and *plbE,* resulted in a 2- to 3-fold increase in plusbacin production. Seeing that the production of all plusbacin derivatives was boosted by about the same factor, the overexpression does not seem to shift the expression profile in the direction of one of the derivatives. For further evaluation a plusbacin standard would be desirable to allow quantitative concertation determination. In addition to the investigated strains, which only overexpress one of the dioxygenases, the simultaneous overexpression of both oxygenases would be of interest.

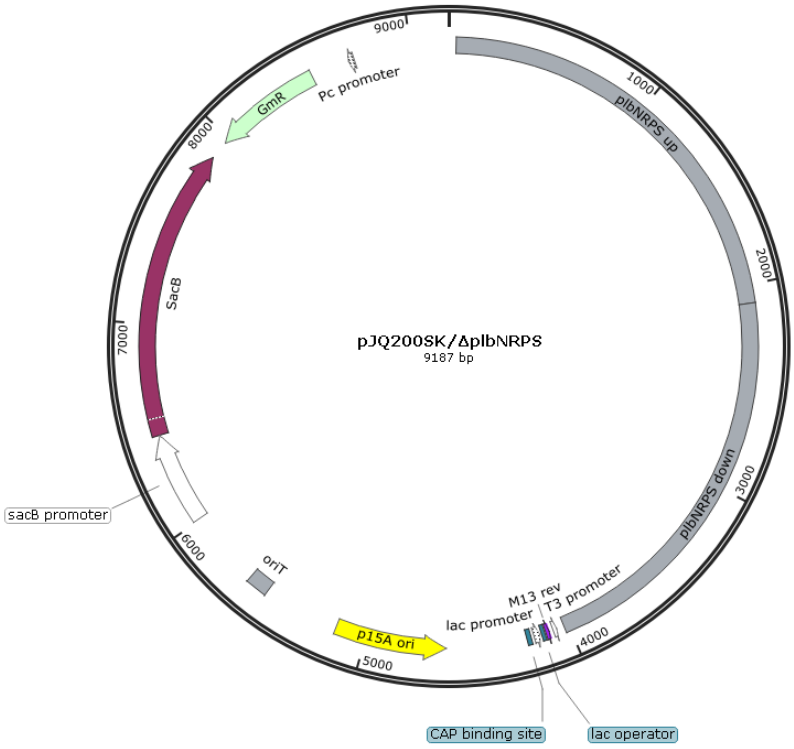
Besides playing an important role during biosynthesis, hydroxylation of aspartic acids catalyzed by the encoded aspartyl hydroxylases PlbD and PlbE are crucial for bioactivity of plusbacin in particular. Inhibiting bacterial cell-wall biosynthesis, plusbacin exhibits potent antibacterial effects against a wide variety of gram-positive bacteria.[153] Thus, *Bacillus subtilis* 168 served as indicator strain for bioactivity testing using disc diffusion assays as well as antagonistic bioassays. Both assays showed antibacterial activity against the indicator stain with extracts from the wild type producer strain whereas extracts from the knockout mutants displayed no effects regarding the growth of *Bacillus subtilis*. Since pinpointing the aspartic acids as targets of the dioxgenases, the absence of β-hydroxy-aspartyl residues in the cyclic depsipeptide core of the plusbacin structure results in the loss of bioactivity. These results are consistent with previously published results, stating that hydroxylation of the aspartic acids is required for proper antibacterial function of plusbacins.[154]

Based one genetic and structural similarities of the main representatives within this group of cyclic guanidine-containing lipopeptide, it can be hypothesized that the results are not unique to plusbacin biosynthesis. Although recently published data regarding the biosynthesis of empedopeptin might suggest otherwise. After

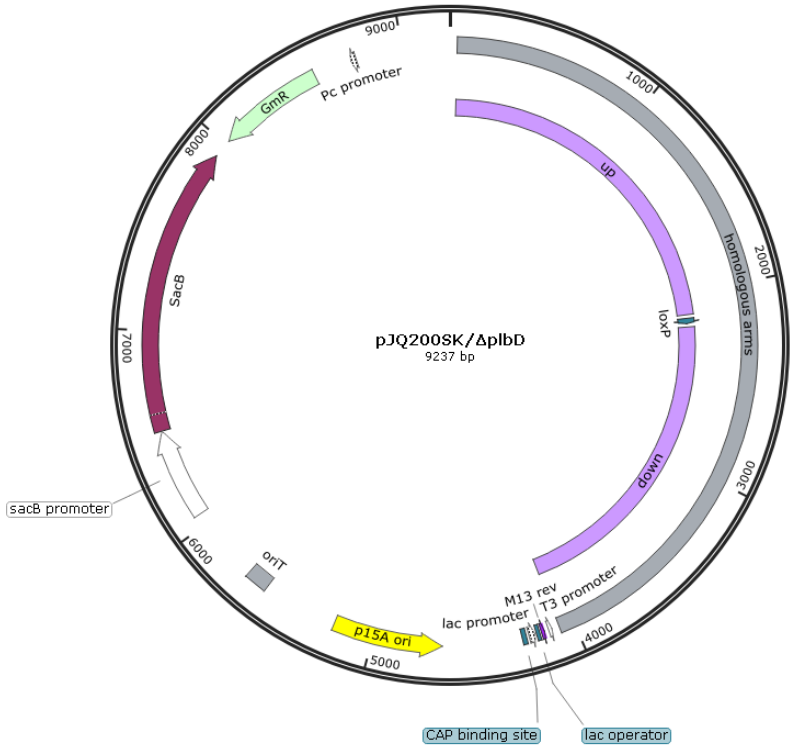
identifying empedopeptin derivatives bearing only one out of its three hydroxylated residues with Hya at position five in a knockout mutant of one dioxygenase and a compound lacking its hydroxy groups at Pro7 and Asp5 and one with no hydroxylation at all after deleting the other dioxygenase, it was speculated that hydroxylation pattern of EMP was achieved by the two dioxygenases encoded in the BGC functioning in a synergistic manner. However, pinpointing the selectivity of the dioxygenases might have been challenging since the structure elucidation was performed by comparative metabolomic analysis of the wild type empedopeptin producer and its in-frame deletion mutants.[129]

Hence, it might be assumed that the oxgenases of the empedopeptin and tripropeptin gene cluster are also responsible for the hydroxylation of the aspartic acid residues as described for plusbacin acting on the amino acid after recruitment of the respective A domain.

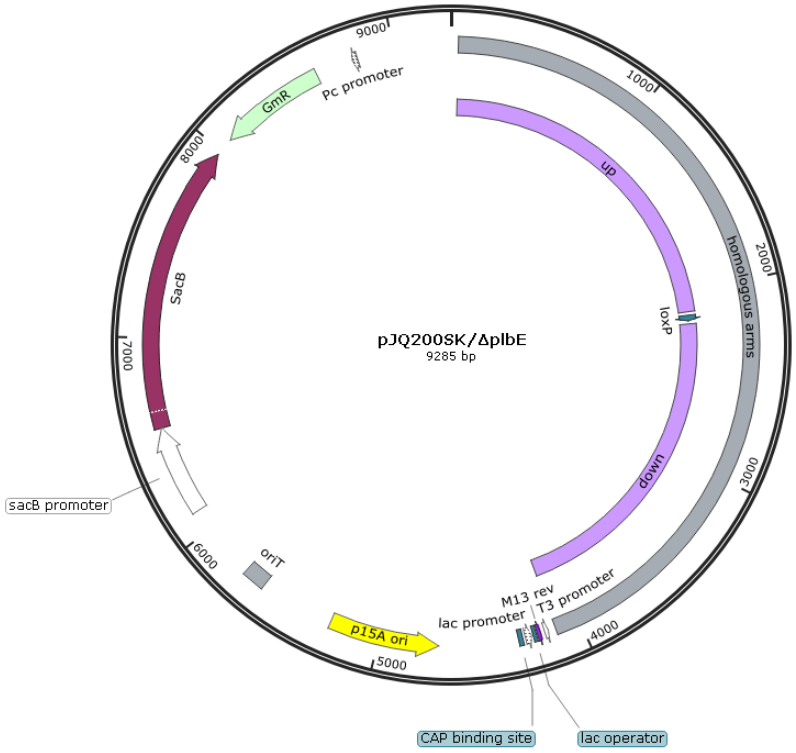
V. Appendix



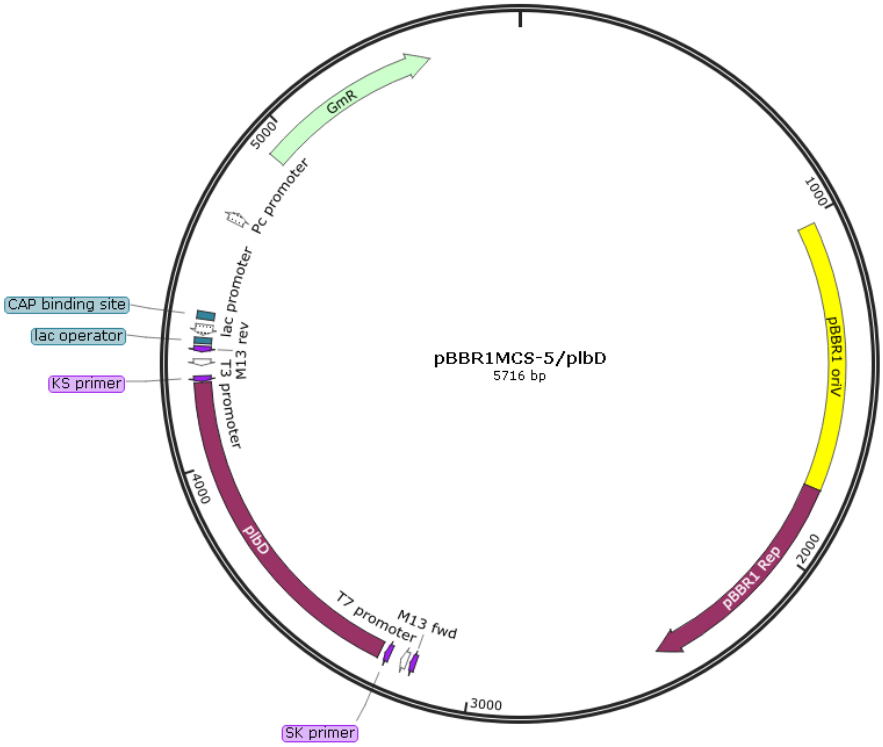
Appendix Figure 1: Vector map of pJQ200SK/ΔplbNRPS



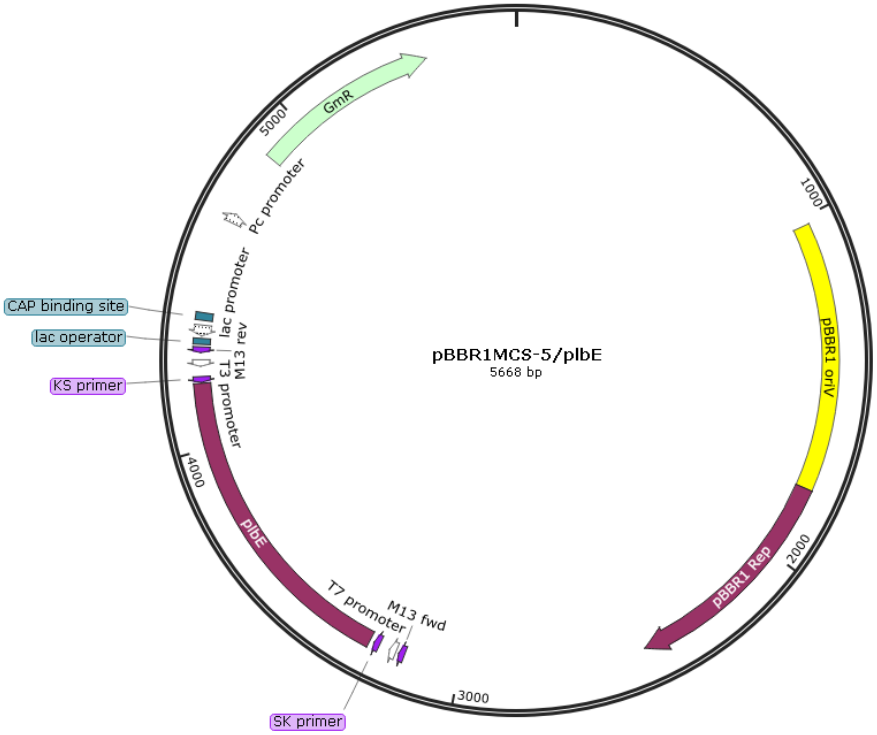
Appendix Figure 2: Vector map of pJQ200SK/ΔplbD



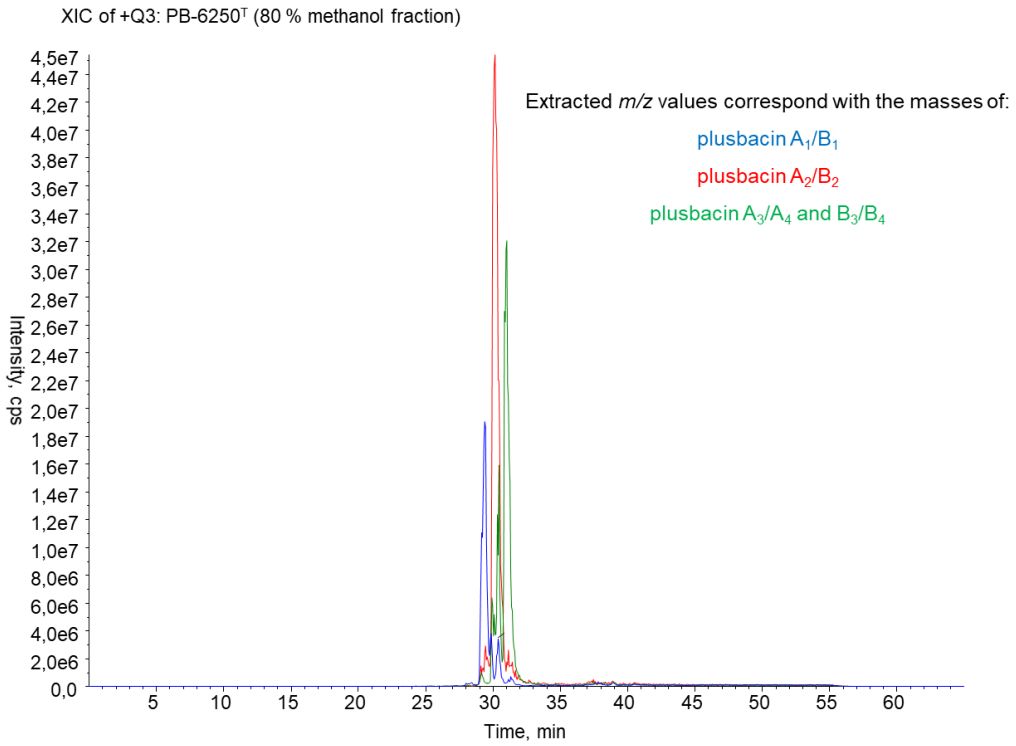
Appendix Figure 3: Vector map of pJQ200SK/ΔplbE



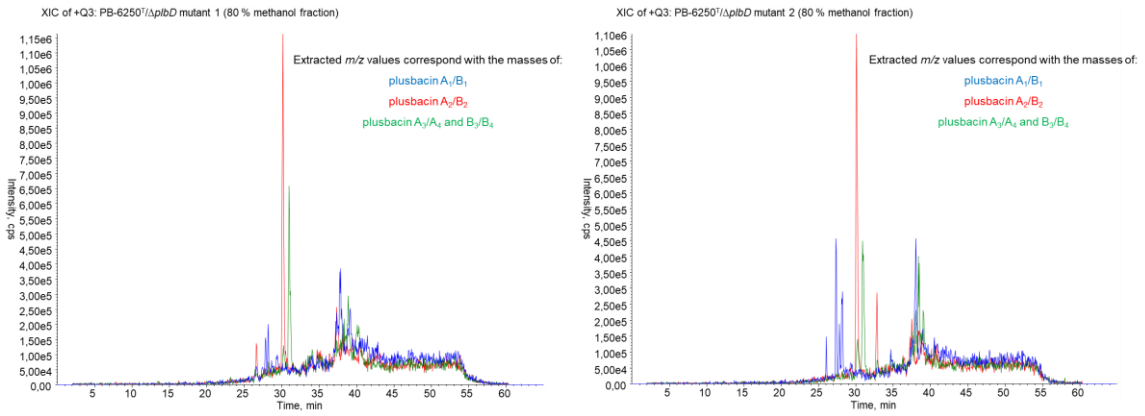
Appendix Figure 4: Vector map of pBBR1MCS-5/plbD



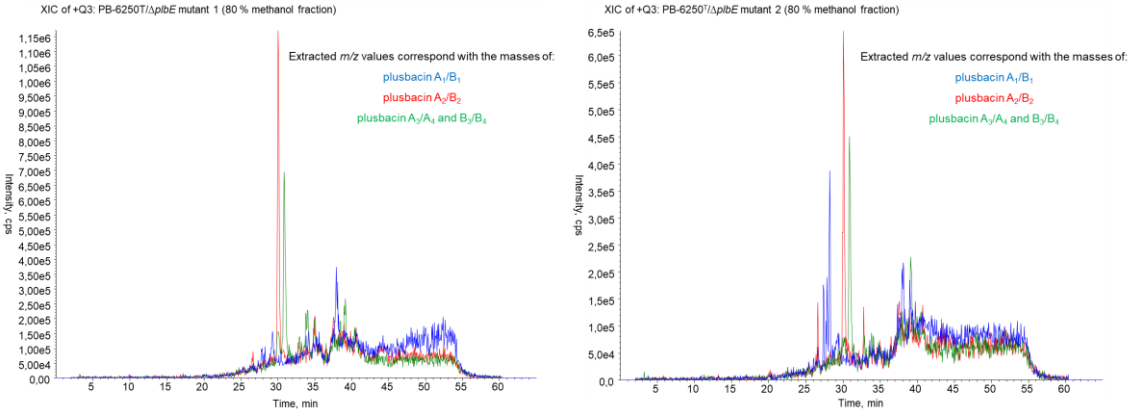
Appendix Figure 5: Vector map of pBBR1MCS-5/plbE



Appendix Figure 6: Extracted ion chromatogram of the extracted ions for plusbacin A1-A4 and plusbacin B1-B4 of the 80 % methanol fraction of the of the wild type plusbacin producer strain PB-6250^T extract after solid phase extraction using Strata™-XL 100 µm polymeric reversed phase 2 g/12 mL giga tubes prior to HPLC purification and enrichment



Appendix Figure 7: Extracted ion chromatogram of the extracted ions for plusbacin A1-A4 and plusbacin B1-B4 of the 80 % methanol fraction of the extract of ΔplbD mutants after solid phase extraction using Strata™-XL 100 µm polymeric reversed phase 2 g/12 mL giga tubes prior to HPLC purification and enrichment



Appendix Figure 8: Extracted ion chromatogram of the extracted ions for plusbacin A1-A4 and plusbacin B1-B4 of the 80 % methanol fraction of the extract of ΔplbE mutants after solid phase extraction using Strata™-XL 100 µm polymeric reversed phase 2 g/12 mL giga tubes prior to HPLC purification and enrichment

Appendix Table 1: A domain sequences

PlbA3
RRVLYDCNDTRRDYAAPALVHEAFEQQAAAHPERVALELDGAQLSYRALNEQANRL ARHLRGLGVGPDRCVAICVERSLSMVVAILATLKAGGAYVPLDPAHPDGRLAQMLRD SRPAALLTQHRLLPRLLPDQAALVLLDDAMLAWAKASAANVHGGELGLKPEHLAYVIY TSGSTGEPKGVAMPHRGLVNLLAWQREQLPEPARTLQFAALGFDVAFQEIFSTLGSG GTLVLLHEELRQDLPALAEWVAQESIERLFLPYIALNRLSELWAQRAEPLPMLQDLITA GEQLRITPAIRRLFVRQPQARLHNHYGPTESHVVTAHTLSGPAEHWEDLPPIGKPIGN SRVYLLDAHARPVPVGVAGELYLGGVQIARGYLQRPALSAQRFLADPFDRRGGGRM YKTGDLGRWREDGSIEYLGRNDFQVKVRGYRIELGEIEARLIGIDGVREAVVLARDER AGETQLVAYLIAEPAAAKPDPAQLRAQLSQGLPDYMLPTAYVTLDAWPLTPNGKLDR KALPAPDSEDYGRRAYVEPQGEWERALAAIWSQLLGVDRVGRDDDFFELGGHSLLA VQLISQVRERFDAELALSTLFVQPRLAELAATVAA
PlbA5
VVHGFNASGTTDNEDLLHRLFERQAAAQPATPALVYDGQALSYAELNARANRIAHHL RSLGLRPDDRVALCLERSLELGVAWWGTLKAGGACVPLDPVHPDERLAHMLADSAP IAVLTQSHLRSRLQVPAGCTVLSLDEAPESAPWAQASAQDPDPDEIGLAPSHLAYVIY TSGSTGLPKGVMVEHRNVLNFLLGMEQRIHGPAPDCRRIAWNSSFGFDMAAKAWG QLCFGRSVHLLSERTRLDAEALLDYIERHAIEAMECTPSHLRMLQAAGLLRGRGAGM RKLLLGGEALDLAAWKTLTEADAVVFHNMYGPTECSVDATCGPVAGAAPQIGRPMP GARVYVLDEHGEPVPIGVPGEIFIGGAGVARGYLHRPGLSAERFVRDPFARGADARM YRTGDLGRWRDDGTIEYLGRNDFQLKIRGYRIELGEIEARLAGLEGVREAVVVAHAD RPGDPRLIAYLLCEPGAQPDPAGLRERLGAQLPDYMLPAAYVVLDAWPLNANGKLD RKALPPPDDDGVAHRTYAAPANDMERRLAAIWAGLLGVERVGRDDNFFELGGHSAL AIQLIHLMSEQQLQVDVQMVFNAPTLADLAAAT
PlbA7
QVLHGFNHRRRDYAAHGELVHALFERQAAATPDAIALEFGPERLSYAELDAQANRLA RHLRSLGIGPDQRVAVCLERGPAMVIAILATLKAGGAYVPLDPTYPDERLGHLLRDSA PRAVLTQQRLRHRLQVAVACQCVLLDEGADDGWASLEASPLPVADLSGEHLAYVIYT SGSTGLPKGVAMPHRGLVNLLAWQRGPLPEPARTLQFAALGFDVAFQEIFSALGSG GTLVLLNEDLRQDLPALADWLDEQSIERLFLPYIALSTLSELWSQREAPLPALRDLIVA GEQLRITPAIRRLFDGRHSTRLHNHYGPTETHVVTAHTLSGPAGSWPDLPPIGAPIDN SRLYLLDAQGRPVPRGVAGEIHIGGVQVARGYLQRAELSAERFLLDPYAAAEPGEPA PRMYKTGDLGRWRDDGSVEYLGRNDFQVKIRGYRIELGEIEARLAEVEGVREAVVIA REDVPGDKRLVAYWVGEAGTALEPAELRARLGAALPDYMLPGAFVPLEALPLTPNGK LDRKALPAPDGLAFVHRAYEATIGEIETTLAQIWCELLGLERVGRQDNFFELGGHSLL AVRLISQLRERLGIELPLSALFTHPQLSELARDVAE
PlbA8
LVLQDFNESHVAAAEPALVHALIERQAATQPDATAIEYGGERLSYAELNARANRLAHH LRGLGVRPDDRIAVCAERSLELVVALLAALKAGAAYVPLDPVYPDERLAHMLEDSGA VALLTQRRLERRLHASPACATLLLDDPQPAWAQAPDANPDPAQIGLNPAHLAYVIYTS GSTGTPKGVMVEHRNVAYFLHAMEACIHGLEPDCRRVAWNSSFGFDMAVKAWGQL AHGRSVHLLPEAARLSAEDLLGFLETHAIEAMECTPSHLRLMQGAGLLQGRAPSLRK LLLGGEAIDSATWSALAAAEDRLFFNMYGPTECSVDASCGVIDGRRPHIGRVMPGAR IYLLDEAGQPVPVGVAGEIHIGGAGVARGYLHRPELTAERFLHDPFAADPQARMYKT GDLGRWREDGTVEYLGRNDFQVKLRGFRIELGEIEARLASQPGVREAAVIAREDSPG DKRLVAYVVGLAGAPTPDPAQLRAGLAPHLPEYMLPSAFVSLDALPLTANGKLDRKA LPAPDGQGLALSRYEPPQGEAELAIAALWSELLGVEKIGRHDNFFDLGGYSLMVFQVI EGLKQKGYEVALQDVLLAQQLSALAALIER

References

1. Ventola, C. L., The antibiotic resistance crisis: part 1: causes and threats. P t 2015, 40 (4), 277-83.
2. Hashizume H.; Nishimura Y., Cyclic Lipopeptides Antibiotics. Stud. Nat. Prod. Chem. 2008, 35, 693-751.
3. Schneider, T.; Müller, A.; Miess, H.; Gross, H., Cyclic lipopeptides as antibacterial agents - potent antibiotic activity mediated by intriguing mode of actions. Int J Med Microbiol 2014, 304 (1), 37-43.
4. Gross, H.; Loper, J. E., Genomics of secondary metabolite production by *Pseudomonas* spp. Nat Prod Rep 2009, 26 (11), 1408-46.
5. Walsh, C. T.; O'Brien, R. V.; Khosla, C., Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polyketide scaffolds. Angew Chem Int Ed Engl 2013, 52 (28), 7098-124.
6. Marahiel, M. A., A structural model for multimodular NRPS assembly lines. Nat Prod Rep 2016, 33 (2), 136-40.
7. Mootz, H. D.; Marahiel, M. A., Biosynthetic systems for nonribosomal peptide antibiotic assembly. Curr Opin Chem Biol 1997, 1 (4), 543-51.
8. Aminov, R., History of antimicrobial drug discovery: Major classes and health impact. Biochem Pharmacol 2017, 133, 4-19.
9. Phillips, G.; Golledge, C. L., Vancomycin and teicoplanin: something old, something new. Med J Aust 1992, 156 (1), 53-7.
10. Yushchuk, O.; Ostash, B.; Truman, A. W.; Marinelli, F.; Fedorenko, V., Teicoplanin biosynthesis: unraveling the interplay of structural, regulatory, and resistance genes. Appl Microbiol Biotechnol 2020, 104 (8), 3279-3291.
11. Belley, A.; McKay, G. A.; Arhin, F. F.; Sarmiento, I.; Beaulieu, S.; Fadhil, I.; Parr, T. R., Jr.; Moeck, G., Oritavancin disrupts membrane integrity of *Staphylococcus aureus* and vancomycin-resistant enterococci to effect rapid bacterial killing. Antimicrob Agents Chemother 2010, 54 (12), 5369-71.
12. Domenech, O.; Francius, G.; Tulkens, P. M.; Van Bambeke, F.; Dufrêne, Y.; Mingeot-Leclercq, M. P., Interactions of oritavancin, a new lipoglycopeptide derived from vancomycin, with phospholipid bilayers: Effect on membrane permeability and nanoscale lipid membrane organization. Biochim Biophys Acta 2009, 1788 (9), 1832-40.
13. Das, B.; Sarkar, C.; Biswas, R.; Pandey, S., Review: dalbavancin--a novel lipoglycopeptide antimicrobial for gram positive pathogens. Pak J Pharm Sci 2008, 21 (1), 78-87.
14. Lampejo, T., Dalbavancin and telavancin in the treatment of infective endocarditis: a literature review. Int J Antimicrob Agents 2020, 56 (3), 106072.
15. Yoshino, N.; Endo, M.; Kanno, H.; Matsukawa, N.; Tsutsumi, R.; Takeshita, R.; Sato, S., Polymyxins as novel and safe mucosal adjuvants to induce humoral immune responses in mice. PLoS One 2013, 8 (4), e61643.
16. DeCrescenzo Henriksen, E.; Phillips, D. R.; Peterson, J. B., Polymyxin E production by *P. amylolyticus*. Lett Appl Microbiol 2007, 45 (5), 491-6.
17. Velkov, T.; Roberts, K. D.; Nation, R. L.; Thompson, P. E.; Li, J., Pharmacology of polymyxins: new insights into an 'old' class of antibiotics. Future Microbiol 2013, 8 (6), 711-24.
18. Yu, Z.; Qin, W.; Lin, J.; Fang, S.; Qiu, J., Antibacterial mechanisms of polymyxin and bacterial resistance. Biomed Res Int 2015, 2015, 679109.

19. Falagas, M. E.; Kasiakou, S. K., Toxicity of polymyxins: a systematic review of the evidence from old and recent studies. Crit Care 2006, 10 (1), R27.
20. Landman, D.; Georgescu, C.; Martin, D. A.; Quale, J., Polymyxins revisited. Clin Microbiol Rev 2008, 21 (3), 449-65.
21. Gupta, S.; Govil, D.; Kakar, P. N.; Prakash, O.; Arora, D.; Das, S.; Govil, P.; Malhotra, A., Colistin and polymyxin B: a re-emergence. Indian J Crit Care Med 2009, 13 (2), 49-53.
22. Hancock, R. E., Alterations in outer membrane permeability. Annu Rev Microbiol 1984, 38, 237-64.
23. Hermsen, E. D.; Sullivan, C. J.; Rotschafer, J. C., Polymyxins: pharmacology, pharmacokinetics, pharmacodynamics, and clinical applications. Infect Dis Clin North Am 2003, 17 (3), 545-62.
24. McCoy, L. S.; Roberts, K. D.; Nation, R. L.; Thompson, P. E.; Velkov, T.; Li, J.; Tor, Y., Polymyxins and analogues bind to ribosomal RNA and interfere with eukaryotic translation in vitro. Chembiochem 2013, 14 (16), 2083-6.
25. Straus, S. K.; Hancock, R. E., Mode of action of the new antibiotic for Gram-positive pathogens daptomycin: comparison with cationic antimicrobial peptides and lipopeptides. Biochim Biophys Acta 2006, 1758 (9), 1215-23.
26. Pogliano, J.; Pogliano, N.; Silverman, J. A., Daptomycin-mediated reorganization of membrane architecture causes mislocalization of essential cell division proteins. J Bacteriol 2012, 194 (17), 4494-504.
27. Müller, A.; Wenzel, M.; Strahl, H.; Grein, F.; Saaki, T. N. V.; Kohl, B.; Siersma, T.; Bandow, J. E.; Sahl, H. G.; Schneider, T.; Hamoen, L. W., Daptomycin inhibits cell envelope synthesis by interfering with fluid membrane microdomains. Proc Natl Acad Sci U S A 2016, 113 (45), E7077-e7086.
28. Johnson, A., Daptomycin in the treatment of skin, soft-tissue and invasive infections due to Gram-positive bacteria. Future Microbiol 2006, 1 (3), 255-65.
29. Sauermann, R.; Rothenburger, M.; Graninger, W.; Joukhadar, C., Daptomycin: a review 4 years after first approval. Pharmacology 2008, 81 (2), 79-91.
30. Baltz, R. H.; Brian, P.; Miao, V.; Wrigley, S. K., Combinatorial biosynthesis of lipopeptide antibiotics in *Streptomyces roseosporus*. J Ind Microbiol Biotechnol 2006, 33 (2), 66-74.
31. Baltz, R. H.; Miao, V.; Wrigley, S. K., Natural products to drugs: daptomycin and related lipopeptide antibiotics. Nat Prod Rep 2005, 22 (6), 717-41.
32. Vértesy, L.; Ehlers, E.; Kogler, H.; Kurz, M.; Meiwes, J.; Seibert, G.; Vogel, M.; Hammann, P., Friulimicins: novel lipopeptide antibiotics with peptidoglycan synthesis inhibiting activity from *Actinoplanes friuliensis* sp. nov. II. Isolation and structural characterization. J Antibiot (Tokyo) 2000, 53 (8), 816-27.
33. Bionda, N.; Stawikowski, M.; Stawikowska, R.; Cudic, M.; López-Vallejo, F.; Treitl, D.; Medina-Franco, J.; Cudic, P., Effects of cyclic lipodepsipeptide structural modulation on stability, antibacterial activity, and human cell toxicity. ChemMedChem 2012, 7 (5), 871-82.
34. Kajimura, Y.; Kaneda, M., Fusaricidins B, C and D, new depsipeptide antibiotics produced by *Bacillus polymyxa* KT-8: isolation, structure elucidation and biological activity. J Antibiot (Tokyo) 1997, 50 (3), 220-8.
35. Sieber, S. A.; Tao, J.; Walsh, C. T.; Marahiel, M. A., Peptidyl thiophenols as substrates for nonribosomal peptide cyclases. Angew Chem Int Ed Engl 2004, 43 (4), 493-8.

36. Quigley, N. B.; Gross, D. C., Syringomycin production among strains of *Pseudomonas syringae* pv. *syringae*: conservation of the *syrB* and *syrD* genes and activation of phytotoxin production by plant signal molecules. Mol Plant Microbe Interact 1994, 7 (1), 78-90.
37. Dalla Serra, M.; Fagiuoli, G.; Nordera, P.; Bernhart, I.; Della Volpe, C.; Di Giorgio, D.; Ballio, A.; Menestrina, G., The interaction of lipodepsipeptide toxins from *Pseudomonas syringae* pv. *syringae* with biological and model membranes: a comparison of syringotoxin, syringomycin, and two syringopeptins. Mol Plant Microbe Interact 1999, 12 (5), 391-400.
38. Konishi, M.; Sugawara, K.; Hanada, M.; Tomita, K.; Tomatsu, K.; Miyaki, T.; Kawaguchi, H.; Buck, R. E.; More, C.; Rossomano, V. Z., Empedopeptin (BMY-28117), a new depsipeptide antibiotic. I. Production, isolation and properties. J Antibiot (Tokyo) 1984, 37 (9), 949-57.
39. Müller, A.; Münch, D.; Schmidt, Y.; Reder-Christ, K.; Schiffer, G.; Bendas, G.; Gross, H.; Sahl, H. G.; Schneider, T.; Brötz-Oesterhelt, H., Lipodepsipeptide empedopeptin inhibits cell wall biosynthesis through Ca^{2+}-dependent complex formation with peptidoglycan precursors. J Biol Chem 2012, 287 (24), 20270-80.
40. Shoji, J.; Hinoo, H.; Katayama, T.; Matsumoto, K.; Tanimoto, T.; Hattori, T.; Higashiyama, I.; Miwa, H.; Motokawa, K.; Yoshida, T., Isolation and characterization of new peptide antibiotics, plusbacins A1-A4 and B1-B4. J Antibiot (Tokyo) 1992, 45 (6), 817-23.
41. Miess, H.; van Trappen, S.; Cleenwerck, I.; De Vos, P.; Gross, H., Reclassification of *Pseudomonas* sp. PB-6250T as *Lysobacter firmicutimachus* sp. nov. Int J Syst Evol Microbiol 2016, 66 (10), 4162-4166.
42. Maki, H.; Miura, K.; Yamano, Y., Katanosin B and plusbacin A(3), inhibitors of peptidoglycan synthesis in methicillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 2001, 45 (6), 1823-7.
43. Kim, S. J.; Singh, M.; Wohlrab, A.; Yu, T. Y.; Patti, G. J.; O'Connor, R. D.; VanNieuwenhze, M.; Schaefer, J., The isotridecanyl side chain of plusbacin-A3 is essential for the transglycosylase inhibition of peptidoglycan biosynthesis. Biochemistry 2013, 52 (11), 1973-9.
44. Hashizume, H.; Hattori, S.; Igarashi, M.; Akamatsu, Y., Tripropeptin E, a new tripropeptin group antibiotic produced by *Lysobacter* sp. BMK333-48F3. J Antibiot (Tokyo) 2004, 57 (6), 394-9.
45. Hashizume, H.; Igarashi, M.; Hattori, S.; Hori, M.; Hamada, M.; Takeuchi, T., Tripropeptins, novel antimicrobial agents produced by *Lysobacter* sp. I. Taxonomy, isolation and biological activities. J Antibiot (Tokyo) 2001, 54 (12), 1054-9.
46. Hashizume, H.; Hirosawa, S.; Sawa, R.; Muraoka, Y.; Ikeda, D.; Naganawa, H.; Igarashi, M., Tripropeptins, novel antimicrobial agents produced by *Lysobacter* sp. J Antibiot (Tokyo) 2004, 57 (1), 52-8.
47. Hashizume, H.; Hirosawa, S.; Sawa, R.; Muraoka, Y.; Ikeda, D.; Naganawa, H.; Igarashi, M., Corrigendum: Tripropeptins, novel antimicrobial agents produced by *Lysobacter* sp. II. structure elucidation. J Antibiot (Tokyo) 2016, 69 (12), 889-891.
48. Hashizume, H.; Igarashi, M.; Sawa, R.; Adachi, H.; Nishimura, Y.; Akamatsu, Y., Corrigendum: A new type of tripropeptin with anteiso-branched chain fatty acid from *Lysobacter* sp. BMK333-48F3. J Antibiot (Tokyo) 2016, 69 (12), 892-893.

49. Hashizume, H.; Sawa, R.; Harada, S.; Igarashi, M.; Adachi, H.; Nishimura, Y.; Nomoto, A., Tripropeptin C blocks the lipid cycle of cell wall biosynthesis by complex formation with undecaprenyl pyrophosphate. Antimicrob Agents Chemother 2011, 55 (8), 3821-8.
50. Wiebach, V.; Mainz, A.; Siegert, M. J.; Jungmann, N. A.; Lesquame, G.; Tirat, S.; Dreux-Zigha, A.; Aszodi, J.; Le Beller, D.; Süssmuth, R. D., The anti-staphylococcal lipolanthines are ribosomally synthesized lipopeptides. Nat Chem Biol 2018, 14 (7), 652-654.
51. Marahiel, M. A.; Stachelhaus, T.; Mootz, H. D., Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. Chem Rev 1997, 97 (7), 2651-2674.
52. Conti, E.; Stachelhaus, T.; Marahiel, M. A.; Brick, P., Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. Embo j 1997, 16 (14), 4174-83.
53. May, J. J.; Kessler, N.; Marahiel, M. A.; Stubbs, M. T., Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. Proc Natl Acad Sci U S A 2002, 99 (19), 12120-5.
54. Gulick, A. M.; Starai, V. J.; Horswill, A. R.; Homick, K. M.; Escalante-Semerena, J. C., The 1.75 A crystal structure of acetyl-CoA synthetase bound to adenosine-5'-propylphosphate and coenzyme A. Biochemistry 2003, 42 (10), 2866-73.
55. Hur, G. H.; Vickery, C. R.; Burkart, M. D., Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. Nat Prod Rep 2012, 29 (10), 1074-98.
56. Baltz, R. H., Function of MbtH homologs in nonribosomal peptide biosynthesis and applications in secondary metabolite discovery. J Ind Microbiol Biotechnol 2011, 38 (11), 1747-60.
57. Lautru, S.; Challis, G. L., Substrate recognition by nonribosomal peptide synthetase multi-enzymes. Microbiology (Reading) 2004, 150 (Pt 6), 1629-1636.
58. Challis, G. L.; Ravel, J.; Townsend, C. A., Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem Biol 2000, 7 (3), 211-24.
59. Stachelhaus, T.; Marahiel, M. A., Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. J Biol Chem 1995, 270 (11), 6163-9.
60. Strieker, M.; Tanović, A.; Marahiel, M. A., Nonribosomal peptide synthetases: structures and dynamics. Curr Opin Struct Biol 2010, 20 (2), 234-40.
61. Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A., The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem Biol 1999, 6 (8), 493-505.
62. Röttig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O., NRPSpredictor2--a web server for predicting NRPS adenylation domain specificity. Nucleic Acids Res 2011, 39 (Web Server issue), W362-7.
63. Duckworth, B. P.; Wilson, D. J.; Aldrich, C. C., Measurement of Nonribosomal Peptide Synthetase Adenylation Domain Activity Using a Continuous Hydroxylamine Release Assay. Methods Mol Biol 2016, 1401, 53-61.
64. Wilson, D. J.; Aldrich, C. C., A continuous kinetic assay for adenylation enzyme activity and inhibition. Anal Biochem 2010, 404 (1), 56-63.
65. Otten, L. G.; Schaffer, M. L.; Villiers, B. R.; Stachelhaus, T.; Hollfelder, F., An optimized ATP/PP(i)-exchange assay in 96-well format for screening of

adenylation domains for applications in combinatorial biosynthesis. Biotechnol J 2007, 2 (2), 232-40.
66. Phelan, V. V.; Du, Y.; McLean, J. A.; Bachmann, B. O., Adenylation enzyme characterization using gamma -(18)O(4)-ATP pyrophosphate exchange. Chem Biol 2009, 16 (5), 473-8.
67. Kittilä, T.; Schoppet, M.; Cryle, M. J., Online Pyrophosphate Assay for Analyzing Adenylation Domains of Nonribosomal Peptide Synthetases. Chembiochem 2016, 17 (7), 576-84.
68. Katano, H.; Watanabe, H.; Takakuwa, M.; Maruyama, C.; Hamano, Y., Colorimetric determination of pyrophosphate anion and its application to adenylation enzyme assay. Anal Sci 2013, 29 (11), 1095-8.
69. Stanišić, A.; Hüsken, A.; Kries, H., HAMA: a multiplexed LC-MS/MS assay for specificity profiling of adenylate-forming enzymes. Chem Sci 2019, 10 (44), 10395-10399.
70. Drake, E. J.; Miller, B. R.; Shi, C.; Tarrasch, J. T.; Sundlov, J. A.; Allen, C. L.; Skiniotis, G.; Aldrich, C. C.; Gulick, A. M., Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. Nature 2016, 529 (7585), 235-8.
71. Condurso, H. L.; Bruner, S. D., Structure and noncanonical chemistry of nonribosomal peptide biosynthetic machinery. Nat Prod Rep 2012, 29 (10), 1099-110.
72. Bloudoff, K.; Schmeing, T. M., Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: discovery, dissection and diversity. Biochim Biophys Acta Proteins Proteom 2017, 1865 (11 Pt B), 1587-1604.
73. Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H., Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. BMC Evol Biol 2007, 7, 78.
74. Marahiel, M. A., Working outside the protein-synthesis rules: insights into non-ribosomal peptide synthesis. J Pept Sci 2009, 15 (12), 799-807.
75. Hoyer, K. M.; Mahlert, C.; Marahiel, M. A., The iterative gramicidin s thioesterase catalyzes peptide ligation and cyclization. Chem Biol 2007, 14 (1), 13-22.
76. Alonzo, D. A.; Magarvey, N. A.; Schmeing, T. M., Characterization of cereulide synthetase, a toxin-producing macromolecular machine. PLoS One 2015, 10 (6), e0128569.
77. Keating, T. A.; Ehmann, D. E.; Kohli, R. M.; Marshall, C. G.; Trauger, J. W.; Walsh, C. T., Chain termination steps in nonribosomal peptide synthetase assembly lines: directed acyl-S-enzyme breakdown in antibiotic and siderophore biosynthesis. Chembiochem 2001, 2 (2), 99-107.
78. Staunton, J.; Wilkinson, B., Biosynthesis of Erythromycin and Rapamycin. Chem Rev 1997, 97 (7), 2611-2630.
79. Steffensky, M.; Mühlenweg, A.; Wang, Z. X.; Li, S. M.; Heide, L., Identification of the novobiocin biosynthetic gene cluster of *Streptomyces spheroides* NCIB 11891. Antimicrob Agents Chemother 2000, 44 (5), 1214-22.
80. van Wageningen, A. M.; Kirkpatrick, P. N.; Williams, D. H.; Harris, B. R.; Kershaw, J. K.; Lennard, N. J.; Jones, M.; Jones, S. J.; Solenberg, P. J., Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. Chem Biol 1998, 5 (3), 155-62.

81. Kennedy, J.; Auclair, K.; Kendrew, S. G.; Park, C.; Vederas, J. C.; Hutchinson, C. R., Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. Science 1999, 284 (5418), 1368-72.
82. Lauer, B.; Russwurm, R.; Bormann, C., Molecular characterization of two genes from *Streptomyces tendae* Tü901 required for the formation of the 4-formyl-4-imidazolin-2-one-containing nucleoside moiety of the peptidyl nucleoside antibiotic nikkomycin. Eur J Biochem 2000, 267 (6), 1698-706.
83. Konz, D.; Klens, A.; Schörgendorfer, K.; Marahiel, M. A., The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases. Chem Biol 1997, 4 (12), 927-37.
84. Wilkinson, B.; Micklefield, J., Chapter 14. Biosynthesis of nonribosomal peptide precursors. Methods Enzymol 2009, 458, 353-78.
85. Hayaishi, O., Oxygenases. In Encyclopedia of Biological Chemistry, 2013; pp 371-374.
86. Toplak, M.; Matthews, A.; Teufel, R., The devil is in the details: The chemical basis and mechanistic versatility of flavoprotein monooxygenases. Arch Biochem Biophys 2021, 698, 108732.
87. Piel, J.; Hertweck, C.; Shipley, P. R.; Hunt, D. M.; Newman, M. S.; Moore, B. S., Cloning, sequencing and analysis of the enterocin biosynthesis gene cluster from the marine isolate '*Streptomyces maritimus*': evidence for the derailment of an aromatic polyketide synthase. Chem Biol 2000, 7 (12), 943-55.
88. Ceccoli, R. D.; Bianchi, D. A.; Rial, D. V., Flavoprotein monooxygenases for oxidative biocatalysis: recombinant expression in microbial hosts and applications. Front Microbiol 2014, 5, 25.
89. Urlacher, V. B.; Girhard, M., Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. Trends Biotechnol 2012, 30 (1), 26-36.
90. Efimov, I.; Basran, J.; Thackray, S. J.; Handa, S.; Mowat, C. G.; Raven, E. L., Structure and reaction mechanism in the heme dioxygenases. Biochemistry 2011, 50 (14), 2717-24.
91. Fletcher, S. C.; Coleman, M. L., Human 2-oxoglutarate-dependent oxygenases: nutrient sensors, stress responders, and disease mediators. Biochem Soc Trans 2020, 48 (5), 1843-1858.
92. Ryle, M. J.; Hausinger, R. P., Non-heme iron oxygenases. Curr Opin Chem Biol 2002, 6 (2), 193-201.
93. Herr, C. Q.; Hausinger, R. P., Amazing Diversity in Biochemical Roles of Fe(II)/2-Oxoglutarate Oxygenases. Trends Biochem Sci 2018, 43 (7), 517-532.
94. Hegg, E. L.; Que, L., Jr., The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. Eur J Biochem 1997, 250 (3), 625-9.
95. Clifton, I. J.; McDonough, M. A.; Ehrismann, D.; Kershaw, N. J.; Granatino, N.; Schofield, C. J., Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins. J Inorg Biochem 2006, 100 (4), 644-69.
96. Aik, W.; McDonough, M. A.; Thalhammer, A.; Chowdhury, R.; Schofield, C. J., Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. Curr Opin Struct Biol 2012, 22 (6), 691-700.

97. Eichhorn, E.; van der Ploeg, J. R.; Kertesz, M. A.; Leisinger, T., Characterization of alpha-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. J Biol Chem 1997, 272 (37), 23031-6.
98. Proshlyakov, D. A.; McCracken, J.; Hausinger, R. P., Spectroscopic analyses of 2-oxoglutarate-dependent oxygenases: TauD as a case study. J Biol Inorg Chem 2017, 22 (2-3), 367-379.
99. Price, J. C.; Barr, E. W.; Tirupati, B.; Bollinger, J. M., Jr.; Krebs, C., The first direct characterization of a high-valent iron intermediate in the reaction of an alpha-ketoglutarate-dependent dioxygenase: a high-spin FeIV complex in taurine/alpha-ketoglutarate dioxygenase (TauD) from *Escherichia coli*. Biochemistry 2003, 42 (24), 7497-508.
100. Grzyska, P. K.; Appelman, E. H.; Hausinger, R. P.; Proshlyakov, D. A., Insight into the mechanism of an iron dioxygenase by resolution of steps following the FeIV=HO species. Proc Natl Acad Sci U S A 2010, 107 (9), 3982-7.
101. Datsenko, K. A.; Wanner, B. L., One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 2000, 97 (12), 6640-5.
102. Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F., PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc Natl Acad Sci U S A 2003, 100 (4), 1541-6.
103. Quandt, J.; Hynes, M. F., Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. Gene 1993, 127 (1), 15-21.
104. Kovach, M. E.; Elzer, P. H.; Hill, D. S.; Robertson, G. T.; Farris, M. A.; Roop, R. M., 2nd; Peterson, K. M., Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene 1995, 166 (1), 175-6.
105. Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997, 25 (17), 3389-402.
106. Birnboim, H. C., A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol 1983, 100, 243-55.
107. Gibson, D. G.; Young, L.; Chuang, R. Y.; Venter, J. C.; Hutchison, C. A., 3rd; Smith, H. O., Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 2009, 6 (5), 343-5.
108. Gust, B.; Chandra, G.; Jakimowicz, D.; Yuqing, T.; Bruton, C. J.; Chater, K. F., Lambda red-mediated genetic manipulation of antibiotic-producing *Streptomyces*. Adv Appl Microbiol 2004, 54, 107-28.
109. Laemmli, U. K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970, 227 (5259), 680-5.
110. Arlt, P.; Hashizume, H.; Igarashi, M.; Gross, H., Genome Sequence of *Lysobacter* sp. Strain BMK333-48F3, the Producer Strain of Potent Lipopeptide Antibiotics of the Tripropeptin Family. Microbiol Resour Announc 2021, 10 (49), e0096921.
111. Blin, K.; Shaw, S.; Kloosterman, A. M.; Charlop-Powers, Z.; van Wezel, G. P.; Medema, M. H.; Weber, T., antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 2021, 49 (W1), W29-w35.
112. Madeira, F.; Park, Y. M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A. R. N.; Potter, S. C.; Finn, R. D.; Lopez, R., The

EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 2019, 47 (W1), W636-w641.
113. Guex, N.; Peitsch, M. C.; Schwede, T., Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis 2009, 30 Suppl 1, S162-73.
114. Bienert, S.; Waterhouse, A.; de Beer, T. A.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T., The SWISS-MODEL Repository-new features and functionality. Nucleic Acids Res 2017, 45 (D1), D313-d319.
115. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T., SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 2018, 46 (W1), W296-w303.
116. Singh, G. M.; Fortin, P. D.; Koglin, A.; Walsh, C. T., beta-Hydroxylation of the aspartyl residue in the phytotoxin syringomycin E: characterization of two candidate hydroxylases AspH and SyrP in *Pseudomonas syringae*. Biochemistry 2008, 47 (43), 11310-20.
117. Belshaw, P. J.; Walsh, C. T.; Stachelhaus, T., Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. Science 1999, 284 (5413), 486-9.
118. Clugston, S. L.; Sieber, S. A.; Marahiel, M. A.; Walsh, C. T., Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidine synthetase is a (D)C(L) catalyst. Biochemistry 2003, 42 (41), 12095-104.
119. Ehmann, D. E.; Trauger, J. W.; Stachelhaus, T.; Walsh, C. T., Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases. Chem Biol 2000, 7 (10), 765-72.
120. Medema, M. H.; Takano, E.; Breitling, R., Detecting sequence homology at the gene cluster level with MultiGeneBlast. Mol Biol Evol 2013, 30 (5), 1218-23.
121. Cassini, A.; Högberg, L. D.; Plachouras, D.; Quattrocchi, A.; Hoxha, A.; Simonsen, G. S.; Colomb-Cotinat, M.; Kretzschmar, M. E.; Devleesschauwer, B.; Cecchini, M.; Ouakrim, D. A.; Oliveira, T. C.; Struelens, M. J.; Suetens, C.; Monnet, D. L., Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. Lancet Infect Dis 2019, 19 (1), 56-66.
122. Bérdy, J., Bioactive microbial metabolites. J Antibiot (Tokyo) 2005, 58 (1), 1-26.
123. Xie, Y.; Wright, S.; Shen, Y.; Du, L., Bioactive natural products from *Lysobacter*. Nat Prod Rep 2012, 29 (11), 1277-87.
124. Zhang, W.; Li, Y.; Qian, G.; Wang, Y.; Chen, H.; Li, Y. Z.; Liu, F.; Shen, Y.; Du, L., Identification and characterization of the anti-methicillin-resistant *Staphylococcus aureus* WAP-8294A2 biosynthetic gene cluster from *Lysobacter enzymogenes* OH11. Antimicrob Agents Chemother 2011, 55 (12), 5581-9.
125. Wirtz, D. A.; Ludwig, K. C.; Arts, M.; Marx, C. E.; Krannich, S.; Barac, P.; Kehraus, S.; Josten, M.; Henrichfreise, B.; Müller, A.; König, G. M.; Peoples, A. J.; Nitti, A.; Spoering, A. L.; Ling, L. L.; Lewis, K.; Crüsemann, M.; Schneider, T., Biosynthesis and Mechanism of Action of the Cell Wall Targeting Antibiotic Hypeptin. Angew Chem Int Ed Engl 2021, 60 (24), 13579-13586.

126. Panthee, S.; Hamamoto, H.; Paudel, A.; Sekimizu, K., *Lysobacter* species: a potential source of novel antibiotics. Arch Microbiol 2016, 198 (9), 839-45.
127. de Bruijn, I.; Cheng, X.; de Jager, V.; Expósito, R. G.; Watrous, J.; Patel, N.; Postma, J.; Dorrestein, P. C.; Kobayashi, D.; Raaijmakers, J. M., Comparative genomics and metabolic profiling of the genus *Lysobacter*. BMC Genomics 2015, 16, 991.
128. Song, C.; Schmidt, R.; de Jager, V.; Krzyzanowska, D.; Jongedijk, E.; Cankar, K.; Beekwilder, J.; van Veen, A.; de Boer, W.; van Veen, J. A.; Garbeva, P., Exploring the genomic traits of fungus-feeding bacterial genus *Collimonas*. BMC Genomics 2015, 16, 1103.
129. Ho, S. T.; Ho, Y. N.; Lin, C.; Hsu, W. C.; Lee, H. J.; Peng, C. C.; Cheng, H. T.; Yang, Y. L., Integrated Omics Strategy Reveals Cyclic Lipopeptides Empedopeptins from *Massilia* sp. YMA4 and Their Biosynthetic Pathway. Mar Drugs 2021, 19 (4).
130. Hutton, J. J., Jr.; Trappel, A. L.; Udenfriend, S., Requirements for alpha-ketoglutarate, ferrous ion and ascorbate by collagen proline hydroxylase. Biochem Biophys Res Commun 1966, 24 (2), 179-84.
131. Petersen, L.; Olewinski, R.; Salmon, P.; Connors, N., Novel proline hydroxylase activities in the pneumocandin-producing fungus *Glarea lozoyensis* responsible for the formation of trans 3- and trans 4-hydroxyproline. Appl Microbiol Biotechnol 2003, 62 (2-3), 263-7.
132. Houwaart, S.; Youssar, L.; Hüttel, W., Pneumocandin biosynthesis: involvement of a trans-selective proline hydroxylase. Chembiochem 2014, 15 (16), 2365-9.
133. Li, Y.; Lan, N.; Xu, L.; Yue, Q., Biosynthesis of pneumocandin lipopeptides and perspectives for its production and related echinocandins. Appl Microbiol Biotechnol 2018, 102 (23), 9881-9891.
134. Mattay, J.; Houwaart, S.; Hüttel, W., Cryptic Production of trans-3-Hydroxyproline in Echinocandin B Biosynthesis. Appl Environ Microbiol 2018, 84 (7).
135. Li, H.; Gilchrist, C. L. M.; Lacey, H. J.; Crombie, A.; Vuong, D.; Pitt, J. I.; Lacey, E.; Chooi, Y. H.; Piggott, A. M., Discovery and Heterologous Biosynthesis of the Burnettramic Acids: Rare PKS-NRPS-Derived Bolaamphiphilic Pyrrolizidinediones from an Australian Fungus, *Aspergillus burnettii*. Org Lett 2019, 21 (5), 1287-1291.
136. Heinilä, L. M. P.; Fewer, D. P.; Jokela, J. K.; Wahlsten, M.; Ouyang, X.; Permi, P.; Jortikka, A.; Sivonen, K., The structure and biosynthesis of heinamides A1-A3 and B1-B5, antifungal members of the laxaphycin lipopeptide family. Org Biomol Chem 2021, 19 (25), 5577-5588.
137. Onishi, M.; Okumura, Y.; Okamoto, R.; Ishikura, T., Proline hydroxylation by cell free extract of a streptomycete. Biochem Biophys Res Commun 1984, 120 (1), 45-51.
138. Lawrence, C. C.; Sobey, W. J.; Field, R. A.; Baldwin, J. E.; Schofield, C. J., Purification and initial characterization of proline 4-hydroxylase from *Streptomyces griseoviridus* P8648: a 2-oxoacid, ferrous-dependent dioxygenase involved in etamycin biosynthesis. Biochem J 1996, 313 (Pt 1) (Pt 1), 185-91.
139. Xie, Y.; Wang, B.; Liu, J.; Zhou, J.; Ma, J.; Huang, H.; Ju, J., Identification of the biosynthetic gene cluster and regulatory cascade for the

synergistic antibacterial antibiotics griseoviridin and viridogrisein in *Streptomyces griseoviridis*. Chembiochem 2012, 13 (18), 2745-57.
140. Rosconi, F.; Davyt, D.; Martínez, V.; Martínez, M.; Abin-Carriquiry, J. A.; Zane, H.; Butler, A.; de Souza, E. M.; Fabiano, E., Identification and structural characterization of serobactins, a suite of lipopeptide siderophores produced by the grass endophyte *Herbaspirillum seropedicae*. Environ Microbiol 2013, 15 (3), 916-27.
141. Reitz, Z. L.; Hardy, C. D.; Suk, J.; Bouvet, J.; Butler, A., Genomic analysis of siderophore β-hydroxylases reveals divergent stereocontrol and expands the condensation domain family. Proc Natl Acad Sci U S A 2019, 116 (40), 19805-19814.
142. Kreutzer, M. F.; Kage, H.; Nett, M., Structure and biosynthetic assembly of cupriachelin, a photoreactive siderophore from the bioplastic producer *Cupriavidus necator* H16. J Am Chem Soc 2012, 134 (11), 5415-22.
143. Hardy, C. D.; Butler, A., Ambiguity of NRPS Structure Predictions: Four Bidentate Chelating Groups in the Siderophore Pacifibactin. J Nat Prod 2019, 82 (4), 990-997.
144. Qian, G.; Wang, Y.; Qian, D.; Fan, J.; Hu, B.; Liu, F., Selection of available suicide vectors for gene mutagenesis using *chiA* (a chitinase encoding gene) as a new reporter and primary functional analysis of *chiA* in *Lysobacter enzymogenes* strain OH11. World J Microbiol Biotechnol 2012, 28 (2), 549-57.
145. Lin, D.; McBride, M. J., Development of techniques for the genetic manipulation of the gliding bacteria *Lysobacter enzymogenes* and *Lysobacter brunescens*. Can J Microbiol 1996, 42 (9), 896-902.
146. Gay, P.; Le Coq, D.; Steinmetz, M.; Berkelman, T.; Kado, C. I., Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. J Bacteriol 1985, 164 (2), 918-21.
147. Roepstorff, P.; Fohlman, J., Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed Mass Spectrom 1984, 11 (11), 601.
148. Biemann, K., Contributions of mass spectrometry to peptide and protein structure. Biomed Environ Mass Spectrom 1988, 16 (1-12), 99-111.
149. Quadri, L. E.; Sello, J.; Keating, T. A.; Weinreb, P. H.; Walsh, C. T., Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. Chem Biol 1998, 5 (11), 631-45.
150. Felnagle, E. A.; Barkei, J. J.; Park, H.; Podevels, A. M.; McMahon, M. D.; Drott, D. W.; Thomas, M. G., MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases. Biochemistry 2010, 49 (41), 8815-7.
151. Zhang, W.; Heemstra, J. R., Jr.; Walsh, C. T.; Imker, H. J., Activation of the pacidamycin PacL adenylation domain by MbtH-like proteins. Biochemistry 2010, 49 (46), 9946-7.
152. Boll, B.; Taubitz, T.; Heide, L., Role of MbtH-like proteins in the adenylation of tyrosine during aminocoumarin and vancomycin biosynthesis. J Biol Chem 2011, 286 (42), 36281-90.
153. O'Connor, R. D.; Singh, M.; Chang, J.; Kim, S. J.; VanNieuwenhze, M.; Schaefer, J., Dual Mode of Action for Plusbacin A(3) in *Staphylococcus aureus*. J Phys Chem B 2017, 121 (7), 1499-1505.

154. Katsuyama, A.; Paudel, A.; Panthee, S.; Hamamoto, H.; Kawakami, T.; Hojo, H.; Yakushiji, F.; Ichikawa, S., Total Synthesis and Antibacterial Investigation of Plusbacin A(3). Org Lett 2017, 19 (14), 3771-3774.