

Biosynthesis of selected natural products from entomopathogenic bacteria

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich 15
der Johann Wolfgang Goethe Universität
in Frankfurt am Main

von
Margaretha Anna Westphalen
aus Speyer

Frankfurt am Main 2021

(D30)

vom Fachbereich 15 der
Johann Wolfgang Goethe Universität als Dissertation angenommen.

Dekan: Prof. Dr. Sven Klimpel

Gutachter:

Prof. Dr. Helge B. Bode

Prof. Dr. Martin Grininger

Datum der Disputation:

Table of Contents

List of Abbreviations	11
Zusammenfassung	14
Summary	19
1. Introduction	22
1.1. <i>Xenorhabdus</i>	22
1.2 Specialised metabolites from <i>Xenorhabdus</i>	23
1.2.1 Gene regulation	26
1.2.2 Thiotemplated assembly line in NRPS and PKS	27
1.3 Non-ribosomal peptide synthetases	28
1.3.1 Adenylation domain	30
1.3.2 Peptidyl Carrier Protein (PCP) domain or thiolation domain	31
1.3.3 Condensation domain	32
1.3.4 Peptide release mechanisms	34
1.3.5 Additional domains.....	35
1.3.6 Tailoring enzymes and trans-acting enzymes	36
1.3.7 NRPS-like enzymes	36
1.4 PKS	37
1.4.1 Acyltransferase	40
1.4.2 Acyl Carrier Protein	40
1.4.3 Ketosynthase	40
1.4.4 Modifying domains: KR, DH, ER	41
1.4.5 Polyketide release mechanisms.....	43
1.5 Docking domains and hybrids of NRPS and PKS enzymes	44
1.5.1 Docking Domains.....	44
1.5.2 NRPS PKS and PKS-NRPS hybrids	45
1.6 NP identification and NP diversification (engineering).....	46
1.6.1 NP identification traditional and targeted approaches.....	46
1.6.2 Generating “new” natural products by NRPS PKS engineering approaches.....	48

1.7 Preliminary work on the BGCs involved in this work	51
1.7.1 Xenofuranone biosynthesis encoded by the <i>xfS</i> BGC.....	51
1.7.2 Glyoxypeptide biosynthesis encoded by the <i>xgs</i> BGC	52
1.7.3 Xildivaline biosynthesis encoded by the <i>xis</i> BGC	53
1.8 Aim of this work	54
2 Material and Methods	55
2.1 General Methods	55
2.1.1 Genomic DNA	55
2.1.2 Plasmid DNA.....	55
2.1.3 Polymerase Chain Reaction (PCR).....	55
2.1.4 Purification from agarose gel and/or PCR	56
2.1.5 Hot Fusion cloning.....	56
2.1.6 <i>E. coli</i> transformation.....	56
2.2 Xenofuranone project	57
2.2.1 Cultivation of strains	57
2.2.2 Microorganisms.....	57
2.2.3 Plasmids	58
2.2.4 Oligonucleotides.....	59
2.2.5 Gene deletions via homologous recombination	60
2.2.6 Heat shock transformation of <i>Xenorhabdus</i>	61
2.2.7 Feeding experiments with d ₃ -methionine or p-N ₃ -F.....	62
2.2.8 Sample preparation for HPLC-MS analysis	62
2.2.9 HPLC-MS analysis	62
2.2.10 Proteomics analysis	63
2.2.11 Bioinformatic analysis.....	63
2.3 Pyrrolizwilline project.....	63
2.3.1 Cultivation of strains	63
2.3.2 Microorganisms.....	64
2.3.3 Plasmids	66

2.3.4	Oligonucleotides.....	67
2.3.5	Promoter exchange.....	71
2.3.6	C _{starter} domain exchange and promoter exchange.....	71
2.3.7	Gene deletions.....	72
2.3.8	Heterologous expression in <i>E. coli</i> DH10B:: <i>mtaa</i>	72
2.3.9	Heat shock transformation of <i>Xenorhabdus</i>	72
2.3.10	Sample preparation for LC-MS analysis.....	72
2.3.11	LC-MS analysis.....	72
2.3.12	Pyrrrolizwilline purification.....	72
2.3.13	Agar diffusion assay.....	73
2.3.14	Bioinformatic and phylogenetic analysis.....	73
2.4	Glyoxpeptide project.....	74
2.4.1	Cultivation of strains.....	74
2.4.2	Microorganisms.....	75
2.4.3	Plasmids.....	76
2.4.4	Oligonucleotides.....	77
2.4.5	Deletion of <i>xgsD</i> and XisC_MonoOX domain encoding sequence.....	78
2.4.6	Sample preparation for LC-MS analysis.....	79
2.4.7	LC-MS analysis.....	79
2.4.8	MonoOx domain protein expression and purification.....	79
2.4.9	Sodiumdodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis....	80
2.4.10	Isothermal titration calorimetry.....	80
2.4.11	Cofactor analysis from denatured purified protein.....	80
2.4.12	Insertion of MonoOx domain in phototemtide synthetase B (PttB).....	81
2.4.13	Growth experiments.....	81
2.4.14	Bioinformatic and phylogenetic analysis.....	81
2.5	Xildivaline project.....	82
2.5.1	Cultivation of strains.....	82
2.5.2	Microorganisms.....	82

2.5.3	Plasmids	83
2.5.4	Oligonucleotides.....	84
2.5.5	Promoter exchange.....	85
2.5.6	Deletions of <i>xisE</i> and <i>xisD</i>	85
2.5.7	T domain point mutations	86
2.5.8	LC-MS analysis	86
2.5.9	4-bromothiophenol (BTP) addition.....	86
2.5.10	Growth experiments	87
2.5.11	Bioinformatic and phylogenetic analysis.....	87
3	Results	88
3.1	Xenofuranone	88
3.1.1	<i>xfkB</i> deletion.....	88
3.1.2	Promoter exchange based activation of the <i>xfB</i> BGC.....	89
3.1.3	Xenofuranone biosynthesis	92
3.1.4	4-Hydroxyphenylpyruvic acid incorporating xenofuranone derivatives	93
3.2	Pyrrrolizwilline	95
3.2.1	Pyrrrolizwilline identification	95
3.2.2	Pyrrrolizidine compounds in <i>Xenorhabdus</i> : Pyrrrolizwilline and pyrrrolizixenamide..	95
3.2.3	XhpAB	98
3.2.3.1	Heterologous expression in <i>E. coli</i>	98
3.2.4	P_{BAD_xhpA} activation in Δhfq background	101
3.2.5	SDR1 XhpC and SDR2 XhpD	101
3.2.6	PPTase XhpE	101
3.2.7	Condensation domain XhpF.....	102
3.2.8	α/β -Hydrolase XhpG	103
3.2.9	Overview of observed intermediates and postulated pyrrrolizwilline biosynthesis	104
3.2.10	Phylogenetic analysis of pyrrrolizwilline-like BGCs.....	107
3.2.11	TetR transcriptional regulator	109
3.2.12	Bioactivity testing	109

3.3 Glyoxpeptide	110
3.3.1 Glycerol adduct identification	110
3.3.2 MonoOx domain characterisation.....	112
3.3.3 Transporter XgsD.....	117
3.3.4 Distribution of XgsC_A2 domain with integrated MonoOx domain.....	118
3.4 Xildivaline	120
3.4.2 Methyltransferase XisE deletion.....	121
3.4.3 P _{BAD_xisA} activation in <i>X. hominickii</i> WT and <i>X. hominickii</i> Δ hfq	123
3.4.4 Peptide deformylase encoding gene <i>xisD</i> deleted in <i>X. hominickii</i> WT.....	125
3.4.5 XisB exhibits two adjacent thiolation domains	127
3.4.6 <i>In silico</i> analysis of XisB_KR.....	128
3.4.7 Two possible biosynthetic routes	130
4 Discussion	133
4.1 Xenofuranone biosynthesis	133
4.1.1 Xenofuranone decarboxylase XfsB.....	133
4.1.2 Methylcarboxylated compound 6	135
4.1.3 Xenofuranone methyltransferase XfsC	135
4.1.4 4-Hydroxyphenylpyruvic acid derivatives.....	137
4.2 Pyrrolizwilline	138
4.2.1 Two pyrrolizidine core units and the roles of XhpA and PxaA.....	139
4.2.2 Gene deletions and further biosynthesis steps.....	142
4.2.3 Phylogenetic analysis and bioactivity	146
4.3 Glyoxpeptide	147
4.3.1 Glycerol derivative	148
4.3.2 MonoOx domain characterisation.....	148
4.3.3 Phylogenetic distribution	152
4.3.4 Transporter XgsD.....	152
4.4 Xildivaline	152
4.4.1 Three main xildivalines.....	152

4.4.2 Additional derivatives detected upon P _{BAD_xisA} activation in <i>X. hominickii</i> WT ..	154
4.4.3 Methyltransferase XisE	155
4.4.4. Peptide deformylase XisD and possible function of xildivaline.....	155
4.4.5 Tandem T domains	156
4.4.6 Proposed biosynthesis	157
4.4.7 Phylogeny	158
4.5 Concluding remarks	159
5 References	160
6 Supplementary Information.....	171
7 Conference participation.....	188
8 Acknowledgement	189
9 Declaration	191
10 Author's declaration.....	191
11 Curriculum vitae	193

List of Abbreviations

A	Adenylation
aa	amino acid
A _{sub}	Small C-terminal subdomain of the adenylation domain
A _{core}	Large N-terminal core domain of the adenylation domain
ACN	Acetonitrile
ACP	Acyl carrier protein
AMPs	Antimicrobial peptides
AmT	Aminotransferase
ANL	Acyl-CoA synthetases, NRPS adenylation domains and Luciferase enzymes
antiSMASH	Antibiotics and secondary metabolite analysis shell
AT	Acyltransferase
ATP	Adenosine triphosphate
BGC	Biosynthetic gene cluster
BlastP	Basic Local Alignment Search Tool (for Protein sequences)
bp	base pairs
BTP	4-bromothiophenol
cAMP	cyclic adenosinmonophospat
C	Condensation
^d C	donor site at the N-terminus of the C domain
C _a	acceptor site at the C-terminus of the C domain
C _{term}	Terminal condensation domain
C _{starter}	Starter condensation domain
CAP	catabolite activator protein
CoA	Coenzyme A
COM	Communication-mediating
CRISPR	Clustered regularly interspaced short palindromic repeats
Cy	Cyclisation
D ₃ met	d ₃ -methionine
DD	Docking domain
DEBS	6-deoxyerythronolide synthase
DFP	Diisopropyl fluorophosphate
DH	Dehydratase
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
E	Epimerisation
easyPACId	Easy Promoter Activated Compound Identification
EDTA	Ethylenediaminetetraacetic acid
EIC	Extracted Ion Chromatogram
ER	Enoylreductase
ESI	Electrospray-ionisation
EtOH	Ethanol
F	Formylation
FAD	Flavine-adenine dinucleotide
FMN	Flavinmononucleotide
h	hour
Hfq	host factor bacteriophage q

LIST OF ABBREVIATIONS

HMM	Hidden Markov Model
4hpa	4-hydroxyphenylpyruvic acid
HPLC	High performance liquid chromatography
HR	High resolution
IJ	infective juvenile
IT	ion trap
ITC	Isothermal titration calorimetry
kDa	kilodalton
KR	Ketoreductase
KS	Ketosynthase
L	litre
LB	Lysogeny broth
LC	Liquid chromatography
LDD	Loading didomain
LrpO	Leucine-responsive regulatory protein O
μ L	microlitre
mL	millilitre
M	Methylation
MDR	Medium chain dehydrogenase/reductase
MeOH	Methanol
min	minute
MMDB	Molecular Modelling Database
MonoOx	Monoxygenase
mRNA	messenger RNA
MS	Mass spectrometry
MTHF	Methyltetrahydrofolate
NAD(P)H	Nicotinamide adenine dinucleotide (phosphate)
NaPDoS	Natural product domain seeker
NCBI	National Center for Biotechnology Information
NMR	Nuclear magnetic resonance
NP	Natural product
NRP	Non-ribosomal peptide
NRPS	Non-ribosomal peptide synthetase
OD	Optical density
OSMAC	One Strain Many Compounds
Ox	Oxidase
P	Pathogenic
PAGE	Polyacrylamide Gel Electrophoresis
PAX	Peptide antimicrobial from <i>Xenorhabdus</i>
PAM	Peptidylglycine- α -amidating monooxygenase
PCP	Peptidyl carrier protein
PCR	Polymerase Chain Reaction
PD	Peptide deformylase
PDB	Protein data bank
PFBHA	O-(2,3,4,5,6-Pentafluorobenzyl)hydroxylamine
PK	Polyketide
PKS	Polyketide synthase

LIST OF ABBREVIATIONS

PLA2	Phospholipase A2
p-N ₃ -F	Para-azido-phenylalanine
PO	Phenol oxidase
Ppa	Phenylpyruvic acid
PPAnt	4'phosphopantetheinyl
Ppm	Parts per million
PPTase	Phosphopantetheinyltransferase
PRISM	Prediction Informatics for Secondary Metabolomes
PUFA	Polyunsaturated fatty acid
R	Reductase
RNA	Ribonucleic acid
rRNA	ribosomal RNA
RNase	Ribonuclease
rpm	round per minute
RXP	Rhabdopeptide/xenortide-like peptides
SAM	S-Adenosylmethionine
SDR	Short chain reductase/ dehydrogenase
SDS	Sodiumdodecylsulfate
SM	Secondary metabolite
SMP30	Senescence marker protein 30
SMURF	Secondary Metabolite Unknown Regions Finder
SNAC	N-acetylcysteamine thioester
T	Thiolation
TE	Thioesterase
TFR	TetR family transcriptional regulator
TR	Terminal reductase
t _R	retention time
tRNA	transfer RNA
UV	Ultraviolet
V	Volt
WT	Wild type
XPP	XenoPhotoProduction
XU	exchange unit in NRPS engineering; comprises A-T-C domains
XUC	exchange unit in NRPS engineering, comprises C _a -A-T- ^d C domains

ZUSAMMENFASSUNG

Zusammenfassung

Diese Arbeit umfasst die Untersuchung vier unterschiedlicher Biosynthesegencluster (BGC) aus *Xenorhabdus*. *Xenorhabdus* sind Bakterien, die in mutualistischer Symbiose mit Nematoden der Gattung *Steinernema* leben und gemeinsam Insektenlarven befallen und töten. *Xenorhabdus* produziert eine Vielzahl an Sekundärmetaboliten, von denen einigen eine Rolle in diesem entomopathogenen Lebenszyklus zugewiesen wird. Dies kann beispielsweise die Unterdrückung der Immunantwort des Insekts oder die Abwehr von Fressfeinden sein.

Häufig werden diese Moleküle von Nicht-Ribosomalen Peptidsynthetasen (NRPS) oder NRPS-Polyketidsynthase-(PKS)-Hybriden produziert. NRPS sind modular aufgebaut und verknüpfen einzelne Aminosäurebausteine zu einem Peptid. Ein klassisches NRPS-Modul verfügt über eine Adenylierungsdomäne (A-Domäne), eine Thiolierungsdomäne (T-Domäne) und eine Kondensationsdomäne (C-Domäne). Die A-Domäne aktiviert einen spezifischen Aminosäurebaustein und lädt diesen auf den Phosphopantetheinylarm der T-Domäne. Die Kondensationsdomäne bildet eine Peptidbindung zwischen zwei Aminosäuren aus, die jeweils stromaufwärts und stromabwärts der C-Domäne an eine T-Domäne gebunden vorliegen. Ein klassisches PKS Modul besitzt eine Acyltransferasedomäne (AT-Domäne), eine T-Domäne und eine Ketosynthasedomäne (KS-Domäne). Die AT-Domäne überträgt Acyl-Einheiten auf die T-Domänen und die KS-Domäne verknüpft einzelne Acyl-Einheiten miteinander. Zusätzliche Domänen innerhalb der NRPS oder PKS sowie nachgeschaltete Enzyme können das entstehende bzw. vom Enzym freigesetzte Produkt zusätzlich modifizieren.

Häufig sind die für die Biosynthese verantwortlichen Gene in sogenannten Biosynthesegenclustern (BGCs) organisiert. Der Austausch des nativen Promotors vor dem ersten Gen durch einen eingebrachten, induzierbaren Promotor, z.B. P_{BAD} , ermöglicht die gezielte Aktivierung des BGCs und die Charakterisierung des Biosyntheseprodukts mittels Flüssigkeitschromatographie-Massenspektrometrie (LC-MS). Neben der Aufreinigung eines identifizierten Moleküls und anschließender Strukturaufklärung mittels Kernspinresonanz (NMR)-Spektroskopie, gibt es verschiedene LC-MS basierte Möglichkeiten, die Struktur eines Metabolits und seine Biosynthese zu analysieren.

Das im Rahmen dieser Arbeit zuerst vorgestellte Teilprojekt befasst sich mit der näheren Charakterisierung der Biosynthese von Xenofuranon. In vorangegangenen Arbeiten waren die zwei Hauptprodukte Xenofuranon A und B aus dem Bakterium *X. szentirmaii* isoliert und strukturell aufgeklärt worden. Im Rahmen meiner Masterarbeit wurde durch Deletion das verantwortliche BGC bestätigt. Das BGC umfasst zwei Gene, die für das NRPS-ähnliche XfsA und eine vorhergesagte Gluconolactonase XfsB kodieren. Obwohl NRPS-ähnliche Enzyme keine C-Domäne aufweisen, können sie zwei aromatische α -Ketosäuren unter Ausbildung

einer C-C-Einfachbindung verknüpfen. Eine im Rahmen dieser Arbeit angefertigte $\Delta xfsB$ Deletionsmutante zeigte die Produktion einer carboxylierten Xenofuranon-Vorstufe. Dies deutet darauf hin, dass XfsA die Bildung der carboxylierten Xenofuranon-Vorstufe aus zwei Phenylpyruvatmolekülen katalysiert. Diese Vorstufe wird dann von dem als Gluconolactonase annotierten Enzym XfsB decarboxyliert, sodass Xenofuranon B als Produkt vorliegt. Im nächsten Schritt der Xenofuranon Biosynthese entsteht Xenofuranon A durch S-Adenosylmethionin-abhängige O-Methylierung von Xenofuranon B. Das *xfs* BGC weist keine eigene Methyltransferase auf, weshalb die benötigte Methyltransferase zu Beginn dieser Arbeit nicht bekannt war. Die verantwortliche Methyltransferase wurde in dieser Arbeit durch eine vergleichende Proteom-Analyse identifiziert und mittels Gendeletion bestätigt. Die Proteom-Analyse basierte auf dem Vergleich zwischen *X. szentirmaii* WT und *X. szentirmaii* Δhfq , da sich die Produktion von Xenofuranon A unterscheidet, wenn das *xfs* BGC mittels Promotoraustausch in diesen beiden Stämmen aktiviert wird. Der induzierte *X. szentirmaii* P_{BAD_xfsA} produziert Xenofuranon A und B während der induzierte *X. szentirmaii* Δhfq P_{BAD_xfsA} vorwiegend Xenofuranon B und nahezu kein Xenofuranon A produziert. Es ist anzunehmen, dass das globale Regulatorprotein Hfq einen negativen regulatorischen Effekt auf die Expression der Methyltransferase hat. Andere Arbeiten haben gezeigt, dass die Hfq-Deletionsmutante von *X. szentirmaii* im Vergleich zum Wildtypstamm eine geringere Zahl an Naturstoffen produziert.

Die in Folge des Promotoraustauschs starke Aktivierung des *xfs* BGCs führte außerdem zur Detektion von zwei weiteren Xenofuranon-Derivaten. Diese beiden Derivate werden jeweils von einem Phenylpyruvat und einem *p*-Hydroxyphenylpyruvat gebildet und unterscheiden sich darin, welcher Baustein zuerst von XfsA eingebaut wird.

Das **nächste im Rahmen dieser Arbeit untersuchte BGC** wurde mittels antiSMASH-Analyse in Genomdaten von *X. hominickii* entdeckt. Daraufhin wurde das BGC mittels Promotoraustausch und nachfolgender Induktion aktiviert. Das Produkt wurde isoliert und die Molekülstruktur wurde mittels NMR-Spektroskopie bestimmt. Die Struktur wurde als Pyrrolizwillin bezeichnet, da sie zwei Pyrrolizidin Einheiten aufweist. Das aktivierte BGC weist sieben Gene auf und die Biosynthese von Pyrrolizwillin wurde mittels LC-MS-Analyse von erzeugten Deletionsmutanten, Promotoraustauschstämmen und einem Reprogrammierungsansatz untersucht. Die ersten zwei vom BGC kodierten Enzyme XhpA und XhpB weisen eine große Sequenzähnlichkeit zu PxA und PxB auf, welche für die Biosynthese von Pyrrolizinenamiden in *Xenorhabdus* verantwortlich sind. In der Pyrrolizinenamid-Biosynthese verwendet die NRPS PxA einen Acylrest, ein Serin und ein Prolin zur Bildung eines 5,6-bizyklischen Intermediates, das im Folgenden von der

ZUSAMMENFASSUNG

Monooxygenase PxaB oxidiert wird, sodass nach Hydrolyse und verkürzendem Ringschluss Pyrrolizinenamid entsteht. Die Ergebnisse der heterologen Expression von XhpA und XhpB in *E. coli* bestätigen, dass diese beiden Enzyme analog zu PxaA und PxaB Pyrrolizinenamid-ähnliche Bausteine bilden. Die initiale C_{starter}-Domäne wurde zwischen PxaA aus *X. szentirmaii* und XhpA aus *X. hominickii* getauscht. Der Tausch resultierte in der Produktion von Pyrrolizinenamiden, wobei die C_{starter}-Domäne von XhpA im Vergleich zur C_{starter}-Domäne von PxaA kürzere Acyl-Reste einbaut. Aufgrund der Struktur von Pyrrolizwillin wird angenommen, dass zwei Pyrrolizinenamid-ähnliche Bausteine für die Biosynthese benötigt werden. Der Pyrrolizwillinproduzent *X. hominickii* besitzt neben dem aktivierten BGC ein weiteres BGC, welches PxaA und PxaB kodiert. Es wird angenommen, dass XhpA und XhpB ausreichend sind für die Biosynthese der beiden benötigten Pyrrolizidineinheiten, da in einer *pxaAB* Deletionsmutante die Produktion von Pyrrolizwillin beobachtet wurde.

Nach der Biosynthese des Pyrrolizinenamid-Bausteins entfernt die α,β -Hydrolase XhpG den Acylrest. In einer $\Delta xhpG$ Deletionsmutante wird ein Pyrrolizinenamid mit C₂-Acylrest als einziges Intermediat entdeckt. Dies wird dahingehend interpretiert, dass die Hydrolyse des Acylrests vor weiteren Reaktionen stattfinden muss. Anschließend können die beiden Reduktasen XhpC und XhpD die Ketofunktion des Pyrrolizidinkerns in zwei aufeinanderfolgenden Zwei-Elektronen-Reduktionen zur ungesättigten C-C-Bindung reduzieren. Der derzeitige Vorschlag für die Biosynthese ist, dass einer der beiden benötigten reduzierten Pyrrolizidinbausteine mit Glyoxalsäure verlängert wird, bevor er mit dem zweiten reduzierten Pyrrolizidinbaustein dimerisiert. Im letzten Schritt erfolgt der Einbau von L-Valin, welcher möglicherweise durch die freistehende C-Domäne XhpF katalysiert wird. Ein essentieller Schritt in der Biosynthese nicht-ribosomaler Peptide (NRPs) ist die initiale posttranslationale Modifikation der T-Domänen von NRPS-Enzymen durch eine Phosphopantetheinyltransferase (PPTase). Trotz Deletion des PPTase kodierenden Gens *xhpE*, war die Bildung von Pyrrolizwillin möglich. Die Deletionsmutante der PPTase NgrA hingegen zeigte keine Produktion von Pyrrolizwillin, was impliziert, dass NgrA in die Biosynthese von Pyrrolizwillin involviert ist. Neben der Aktivierung von T-Domänen in XhpA wird auch die Möglichkeit diskutiert, dass L-Valin durch eine weitere NRPS, die eine posttranslationale Modifikation erfahren muss, bereitgestellt werden könnte. Vermutlich werden weitere Enzyme, wie beispielsweise eine CoA-Ligase, für die Biosynthese von Pyrrolizwillin benötigt, die nicht in dieser Arbeit identifiziert wurden. Möglicherweise stehen weitere benötigte Enzyme unter dem Einfluss von dem Regulatorprotein Hfq, da die Bildung von Pyrrolizwillin bei einem Promotoraustausch in *X. hominickii* Δhfq nicht beobachtet werden konnte. Weder für die Deletionsmutante noch die Plasmid-basierte Expression des Transkriptionsfaktor-Gens, das in räumlicher Nähe zum Pyrrolizwillin-BGC liegt, konnte eine Produktion von Pyrrolizwillin festgestellt werden. Zukünftige Untersuchungen könnten

Aufschluss über weitere Details der Biosynthese und eine mögliche Funktion im Lebenszyklus des Bakteriums geben.

Das dritte untersuchte BGC ist verantwortlich für die Biosynthese eines Tripeptids. In vorausgegangenen Arbeiten wurde die Biosynthese wie folgt postuliert: Die NRPS XgsA aktiviert L-Aspartat, epimerisiert dieses und setzt durch Reduktion 3-Amino-4-oxo-Butansäure frei. Diese wird von der Short-Chain-Dehydrogenase/Reduktase zu β -D-Homoserin reduziert, welches von der ersten A-Domäne der zweiten im BGC kodierten NRPS XgsC als Substrat erkannt wird. XgsC verlängert β -D-Homoserin mit α -Hydroxyglycin und L-Valin zu einem linearen Tripeptid, das in dieser Arbeit als Glyoxpeptid bezeichnet wird.

Neben Glyoxpeptid wurde in vorausgegangenen Arbeiten ein weiteres Derivat beobachtet. Dieses wurde im Rahmen dieser Arbeit als Glycerin-Derivat von Glyoxpeptid identifiziert, das auf das im Medium enthaltene Glycerin zurückzuführen ist. Bei Verzicht auf Glycerin im verwendeten Produktionsmedium wird nur Glyoxpeptid als Produkt beobachtet.

Eine in die Glycin-aktivierende A-Domäne inserierte Monooxygenase-Domäne (MonoOx-Domäne) ist verantwortlich für die α -Hydroxylierung von Glycin. Im Zuge dieser Arbeit konnte mittels Deletion des Abschnitts, der die MonoOx-Domäne kodiert, eine veränderte, funktionale NRPS generiert werden, die das nicht-hydroxylierte Glyoxpeptid produziert. Die MonoOx-Domäne ist in die Sequenz der Glycin-aktivierenden Adenylierungsdomäne zwischen den Sequenzmotiven A8 und A9 integriert. Unter anderem die Position der Insertion unterscheidet die XgsC_MonoOx-Domäne von den MonoOx-Domänen aus der Melithiazol und Myxothiazol Biosynthese, welche ebenfalls die α -Hydroxylierung von Glycin katalysieren, jedoch zwischen den Sequenzmotiven A4 und A5 in die Adenylierungsdomäne eingegliedert sind. Eine BlastP-Analyse zeigte, dass ähnliche, Glycin-aktivierende A-Domänen mit einer zwischen A8 und A9 integrierten MonoOx-Domäne auch in weiteren BGCs anderer Proteobakterien vorkommen. Untersuchungen der aufgereinigten MonoOx-Domäne mittels isothermaler Titrationskalorimetrie zeigten keine Bindung für die getesteten Kofaktoren FMN oder FAD. Außerdem konnte für die Insertion der XgsC_MonoOx-Domäne in die Glycin-aktivierende Adenylierungsdomäne der Phototemtidsynthetase keine Produktbildung beobachtet werden. Es bleibt zu untersuchen, inwiefern die MonoOx-Domäne eine Anwendung in der Reprogrammierung von NRPS finden kann. Es wird beschrieben, dass die α -Hydroxyfunktion durch Protonabstraktion von einer Base zur Spaltung des Peptids führen kann. Die Spaltprodukte sind demnach die N-terminale Peptidhälfte mit C-terminaler Amidfunktion und die C-terminale Peptidhälfte, die ein Glyoxal am N-Terminus trägt.

ZUSAMMENFASSUNG

Das vierte im Rahmen dieser Arbeit untersuchte BGC wurde im Vorfeld durch einen Promotortausch vor *xisA* in *X. hominickii* Δ *hfq* aktiviert. Dieses BGC kodiert für ein NRPS-PKS-NRPS-Hybridsystem sowie eine Peptiddeformylase und eine Methyltransferase. Da die Aufreinigung des Produkts für eine NMR-Spektroskopie basierte Strukturaufklärung im Rahmen der Masterarbeit von Janine Chekaiban nicht erfolgreich war, wurde die Molekülstruktur auf Grundlage verschiedener LC-MS-Experimente postuliert. Die beobachteten Produkte werden im Rahmen dieser Arbeit als Xildivaline benannt. Diese Arbeit schlägt basierend auf der MS² Fragmentierung eine alternative Position der *N*-Methylierung für Xildivalin vor. Außerdem werden zwei zusätzliche Derivate identifiziert, die nur nach Aktivierung des Promotors im Wildtypstamm und nicht bei Promotoraktivierung in der *Hfq*-Deletionsmutante zu beobachten sind. Die beiden zusätzlichen Derivate sind vermutlich auf eine Variation im PKS-Baustein des Moleküls zurückzuführen und könnten durch weitere Enzyme katalysiert werden, welche nur im Wildtypstamm und nicht in der *Hfq*-Deletionsmutante exprimiert werden.

Die Anwesenheit eines Michael-Akzeptors wird durch Addition des gefütterten Elektrophils 4-Bromothiophenol bestätigt. Die PKS *XisB* weist zwei benachbarte T-Domänen, sogenannte Tandem-T-Domänen, auf. Diese wurden jeweils durch Punktmutation des im aktiven Zentrum befindlichen Serins inaktiviert. Die jeweilige Inaktivierung der ersten oder zweiten T-Domäne zeigte eine verminderte Produktion der verschiedenen Xildivalin-Derivate im Vergleich zum Wildtyp-Stamm. Die Deletion von *xisD*, das eine Peptiddeformylase kodiert, zeigt keinen Einfluss auf das Produktspektrum. Die Deletion von *xisE*, das eine Methyltransferase kodiert, führte zur Produktion von nichtmethylierten Xildivalinen. Es bleibt zu klären, ob die *N*-Methylierung während der Produktbildung an der NRPS oder am freigesetzten Produkt stattfindet. Des Weiteren diskutiert diese Arbeit zwei verschiedene Biosynthesewege, die sich in der Reihenfolge der stattfindenden Reaktionen unterscheiden. Im Fokus steht dabei die C-terminale C-Domäne von *XisC*, welche als stromabwärts präsentiertes Substrat entweder das freigesetzte NRPS-PKS-Hybrid-Produkt oder das von *XisA* gebundene Isoleucin akzeptieren könnte.

Zusammenfassend hat diese Arbeit Teilschritte in den jeweiligen Biosynthesewegen von vier verschiedenen Naturstoffen aus *Xenorhabdus* charakterisiert. Das Beispiel der Xenofuranon-Methyltransferase zeigt eine zielgerichtete Methode für die Identifizierung von modifizierenden Enzymen, die nicht in unmittelbarer Nähe des BGCs kodiert sind. Pyrrolizillin ist ein neuartiger Naturstoff mit spannender Struktur und Biosynthese. Die MonoOx-Domäne ist möglicherweise eine Ergänzung für den zum NRPS-Engineering zur Verfügung stehenden Baukasten. Außerdem wird kurz die Rolle von *Hfq* als Vorteil und auch als Hindernis bei der Entdeckung von Naturstoffen und der Charakterisierung von Biosyntheseschritten thematisiert.

Summary

This work addresses the investigation of four different biosynthesis gene clusters from *Xenorhabdus*. *Xenorhabdus* is an entomopathogenic bacterium that lives in mutualistic symbiosis with its *Steinernema* nematode host and together they infect and kill insect larvae. *Xenorhabdus* is well known for the production of so-called specialised metabolites and some of them were predicted to function in the defense against the insect's immune system or against food predators. Many of these compounds are synthesised by non-ribosomal peptide synthetases (NRPSs) or NRPS-polyketide synthase (PKS)-hybrids. These enzymes are organised in a modular manner wherein NRPSs assemble amino acid building blocks into linear or cyclic peptides while PKSs generate polyketides from acyl-CoA derived extender units. NRPS and PKS enzymes respectively produce structurally very diverse molecules, often with the help of modifying domains and tailoring enzymes. In general, the genes involved in the biosynthesis are organised in so-called biosynthetic gene clusters (BGCs) in the genome of the producing strain. Exchanging the native promoter with an inducible promoter, e.g. P_{BAD} , allows the targeted activation of the BGC and in turn the analysis of the biosynthesis product via LC-MS analysis.

The first BGC investigated in this work is responsible for the biosynthesis of xenofuranones. The phenylpyruvate dimers xenofuranone A and B and the corresponding *xf*s BGC were identified from *X. szentirmaii* and a biosynthetic route had been proposed in previous work. The BGC encodes an NRPS-like enzyme XfsA and a predicted gluconolactonase XfsB. Based on gene deletions, this work shows that the NRPS-like enzyme XfsA produces a carboxylated furanone intermediate which is subsequently decarboxylated by XfsB to yield xenofuranone B. The next step in xenofuranone biosynthesis is the O-methylation of xenofuranone B to yield xenofuranone A. The BGC does not encode a methyltransferase and no candidate was identified in close proximity. A comparative proteomics approach allowed identification of three methyltransferase candidates and subsequent gene deletions confirmed one of the candidates to be responsible for methylation of xenofuranone B. The proteome analysis was based on the comparison of *X. szentirmaii* WT and *X. szentirmaii* Δhfq because distinct levels of the methylated xenofuranone A were observed when the *xf*s BGC was activated in either WT or Δhfq strain. Hfq is a global transcriptional regulator whose deletion is associated with the down regulation of natural product biosynthesis in *Xenorhabdus*. The strong P_{BAD} activation of the *xf*s BGC also allowed the detection of two novel xenofuranone derivatives. They are derived from the incorporation of 4-hydroxyphenylpyruvic acid as first or second building block replacing one of the two phenylpyruvic acid building blocks required for xenofuranone biosynthesis, respectively.

SUMMARY

The second BGC addressed in this work was identified from *X. hominickii* genome sequence in an antiSMASH analysis and had not been characterised before as to my knowledge. Activation of P_{BAD} introduced in front of the first gene, *xhpA*, resulted in the detection of a novel metabolite. Compound purification allowed NMR-based structure elucidation. The molecule exhibits two pyrrolizidine moieties and was named pyrrolizwilline (pyrrolizidine + twin (German: "Zwilling")). The BGC comprises seven genes and single gene deletions as well as heterologous expression in *E. coli* and NRPS engineering were conducted to investigate the biosynthesis. The first two genes *xhpA* and *xhpB* encode a bimodular NRPS and a monooxygenase which synthesise a pyrrolizixenamide-like structure, similar to PxA and PxB in pyrrolizixenamide biosynthesis. The pyrrolizwilline producer *X. hominickii* exhibits another BGC with similarity to the *pxa* BGC. Deletion of this second putative *pxa* BGC still allows pyrrolizwilline production, though in reduced amounts. The in this work presented biosynthesis proposes that XhpA and XhpB synthesise both pyrrolizixenamide-like molecules required for pyrrolizwilline formation. The acyl side chain incorporated by XhpA is removed by the α,β -hydrolase XhpG. The keto function is then reduced by two subsequent two electron reductions catalysed by XhpC and XhpD. One of these two reduced pyrrolizidine units most likely is extended with glyoxalate prior to dimerisation with the remaining pyrrolizidine moiety. To yield pyrrolizwilline, L-valine is attached to the precursor and the free-standing condensation (C) domain XhpF could be catalysing this reaction step. It is shown that the BGC's own PPTase is not essential for pyrrolizwilline biosynthesis as deletion of this gene only decreases but does not abolish compound production. PPTases catalyse the posttranslational modification of T domains, which are required for amino acid substrate binding and proper NRPS function. It is suggested that the PPTase NgrA catalyses this essential reaction in pyrrolizwilline biosynthesis since deletion of the corresponding gene abolished compound production.

Conducted growth inhibition assays did not reveal a bioactivity against the test strains and neither deletion nor plasmid based expression of the LysR-type transcriptional regulator encoded in proximity to the BGC resulted in pyrrolizwilline production. Future investigations might shed light on a possible function of this highly interesting pyrrolizwilline structure.

The third BGC investigated in this work had already been addressed in previous work. Therein, activation of the exchanged promoter in front of *xgsA* in *X. hominickii* had led to the identification of a novel tripeptide composed of β -D-homoserine, α -hydroxyglycine and L-valine and is referred to as glyoxpeptide in this work. In addition, a derivative was detected in previous work. This work demonstrates that this glyoxpeptide derivative is derived from glycerol present in the culture medium. Furthermore, this work aimed at a better characterisation of the monooxygenase (MonoOx) domain present in XgsC. This MonoOx domain is responsible for the α -hydroxylation of glycine and is embedded in between core motifs A8 and A9 within the glycine activating A domain. The position of insertion within the A domain is different to the

position of the previously described glycine hydroxylating MonoOx domains from melithiazol and myxothiazol biosynthesis which are found in between core motifs A4 and A5. Deletion of the MonoOx domain encoding sequence resulted in a functional NRPS that produces the non-hydroxylated tripeptide. ITC experiments with the purified MonoOx domain protein did not show binding of the analysed cofactors FMN or FAD. The attempt to apply this MonoOx domain in the glycine incorporating A domain of the phototemtide synthetase did not show compound production. It is suggested that the α -hydroxylation of glycine renders the tripeptide prone to hydrolysis via hemiacetal formation. Hence, the XgsC_MonoOx domain might be an interesting candidate for further NRPS engineering.

The fourth BGC addressed in this work had been activated by promoter exchange activation in *X. hominickii* Δhfq in previous work and based on LC-MS experiments, a structure and biosynthesis have been proposed. The BGC encodes an isoleucine-activating NRPS initiation module followed by a malonyl-extending PKS and a bimodular NRPS with predicted specificity for incorporation of two valines. In addition, the BGC exhibits a peptide deformylase and a methyltransferase. The detected products are referred to as xildivalines within this work and an alternative position for the *N*-methylation is proposed in contrast to the previous work. Furthermore, this work describes two additional derivatives which are detected when the promoter is exchanged and activated in the *X. hominickii* WT strain but are not detected when the promoter is exchanged and activated in the *X. hominickii* Δhfq strain. Both derivatives are proposed to differ in the PKS-derived moiety and could be installed by tailoring enzymes that are not expressed in the Δhfq strain. The existence of the previously predicted Michael acceptor is supported in this work by a 4-bromothiophenol feeding approach. Deletion of the methyltransferase encoding gene *xisE* results in the production of non-methylated xildivalines. It remains to be determined when the *N*-methylation of L-valine takes place. It is discussed that the methyltransferase could act on the NRPS released product but also during the assembly. The peptide deformylase is not involved in the proposed biosynthesis as xildivaline production is detected in a $\Delta xisD$ strain. The PKS XisB features two adjacent, so called tandem T domains. The inactivation of the first or the second T domain by point mutation causes decreased production titres of detected xildivalines in the respective mutant strain when compared to the wild type.

In conclusion, this work characterises different aspects of natural product biosynthesis in the above presented four different biosynthetic pathways and demonstrates the versatility of NRPSs. This work shows an efficient strategy to identify tailoring enzymes by the example of the xenofuranone methyltransferase. The regulatory effect of the global regulator protein Hfq on modifying enzymes is discussed as both helpful and challenging in the discovery of natural products and their biosynthesis. Furthermore, the XgsC_MonoOx domain is proposed as an additional tool for further NRPS engineering.

INTRODUCTION

1. Introduction

1.1. *Xenorhabdus*

Xenorhabdus species are Gram-negative, facultatively anaerob, rod-shaped bacteria and belong to the family of *Morganellaceae* within the phylum of γ -Proteobacteria.^[1] They are insect pathogenic and live in symbiosis with their nematode host from the genus of *Steinernema*.^[2] The bacteria have so far not been found as free-living forms in soil but were always identified from soil-dwelling nematodes or infected insects.^[3] Hence, *X. szentirmai* was isolated from the nematode *S. rarum* (Argentina) while *X. hominickii* species were isolated from *S. karii* (Kenya) or *S. monticolum* (Korea).^[4] ^[5] *X. hominickii* DSM 17903 addressed in this work was isolated from *S. karii*.^[5] The bacterium-nematode symbiosis is described as an obligate mutualism.^[6] The bacterium requires the nematode to be carried into the insect and in turn the bacteria are responsible for killing of the insect which leads to provision of nutrients for nematode reproduction and growth (Fig. 1).^[6]

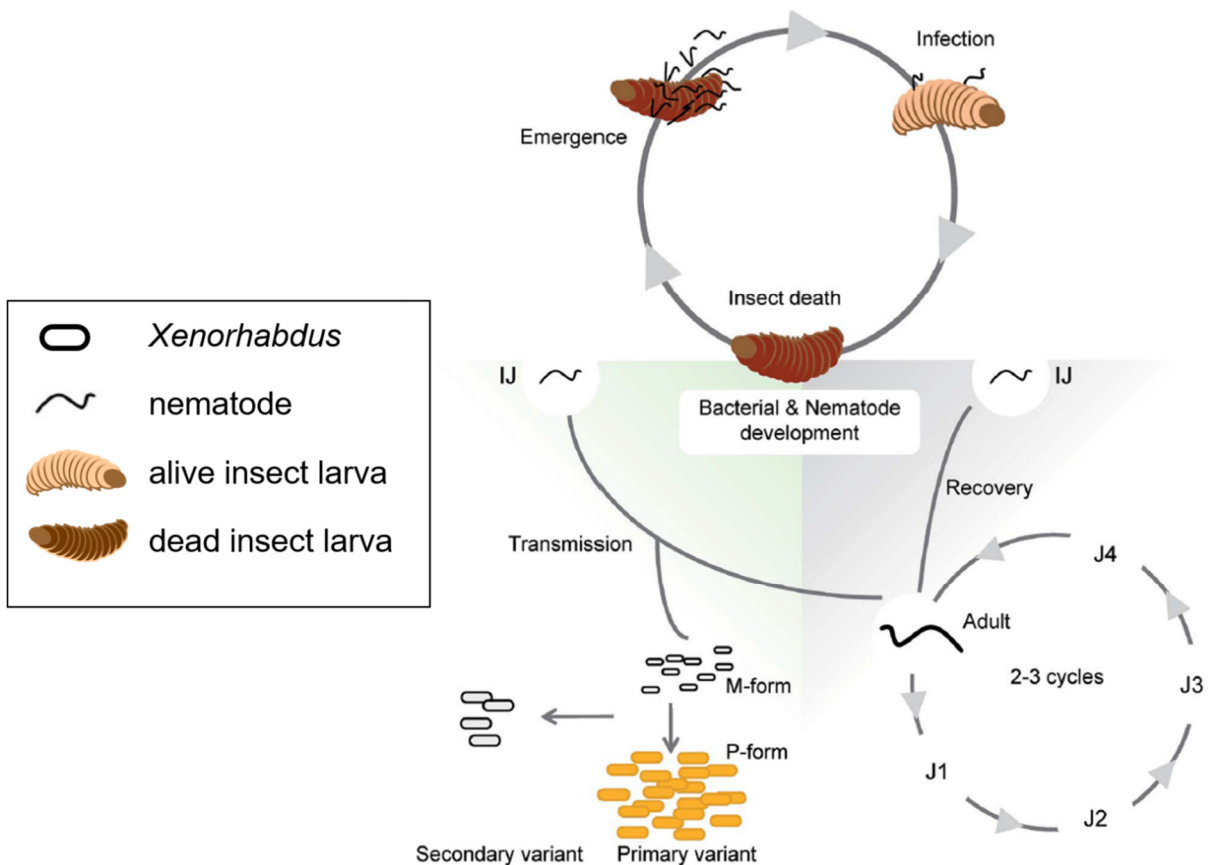


Figure 1. Life cycle of the entomopathogenic *Xenorhabdus/Steinernema*-complex. The bacterium colonises the infective juvenile nematode and therein enters the insect. Following infection and insect death, a new generation of infective juveniles leave the dead insect and takes up bacteria. The bacteria display phenotypic variation and show mutualistic (M) and pathogenic (P) forms. Taken from ^[7] and used in analogy to the *Photorhabdus/Heterorhabditis* complex.

The life cycle of the *Xenorhabdus-Steinernema* pair can be divided in three parts (Fig.1).^[8] At first, the bacteria colonise the receptacle, a vesicle found between the pharynx and the

intestine of the nematode.^[9] It is assumed that a small number of bacterial cells is sufficient to start higher cell number colonisation in the nematode.^[10] This stage of the nematode is termed infective juvenile (IJ) and is living in the soil and non-feeding.^[11] The IJ searches for an insect prey and enters the insect via natural openings or the cuticle and then releases the bacteria into the haemolymph.^[12] ^[13] There, the bacteria multiply and start to produce protein toxins to kill the insect.^[14] *Xenorhabdus* was shown to be able to kill the insect without contribution from the nematode.^[2] Following the insect's death, the nematode feeds on the insect cadaver and the bacterial biomass.^[11] The nematodes pursue two to three life cycles from larval stage to adult before a new generation of IJs reassociates with bacteria and leaves in search of the next insect prey.^[8] In general, the bacteria-nematode complex can sense the insect host via different parameters including olfactory cues and is able to infect a variety of different insect species.^[15,16] ^[16]

Phenotypic variation is a characteristic trait of *Xenorhabdus* and different phenotypic variants have been described.^[17] *Xenorhabdus* is usually isolated as the primary variant from nature.^[17] This primary variant is often pigmented and is observed to transform into a secondary variant under laboratory conditions.^[17] One phenotypic switch is described for *X. nematophila* primary variant to take place within the soil-dwelling IJ: The IJ takes up the mutualistic form of *Xenorhabdus* but as the IJ ages prior to insect infection, *Xenorhabdus* switches to a virulent form.^[18]

Xenorhabdus encodes a relatively high number of biosynthetic gene clusters (BGCs) in its genome and therefore is regarded as rich producer of specialised metabolites.^[19] In general, these metabolites are often derived from non-ribosomal peptide synthetases (NRPSs) or polyketide synthases (PKSs)-NRPS hybrids. The mechanisms of non-ribosomal peptide (NRP) and polyketide (PK) biosynthesis are described in sections 1.3 and 1.4. Prior to the introduction to NRPS and PKS, an overview of selected specialised metabolites identified from *Xenorhabdus* and their proposed function are presented in section 1.2.

1.2 Specialised metabolites from *Xenorhabdus*

Within its complex life style, *Xenorhabdus* produces a variety of specialised metabolites which can function to support nematode growth, kill the insect, evade the insect's defensive immune response, defeat predators from the insect cadaver or mediate cell-cell-communication.^[7] A variety of these secondary metabolites are synthesised by NRPS enzymes.^[7] Both NRPSs and PKSs follow a thiotemplated assembly line enzymology and rely on the action of a phosphopantetheinyltransferase (PPTase).^[20] In *X. szentirmaii* the deletion of the PPTase encoding gene resulted in reduced virulence and is likely caused by the depletion of the specialised metabolites.^[21] In the following, a selection of characterised specialised

INTRODUCTION

metabolites from *Xenorhabdus* is presented for which a function within the life cycle was assigned (Fig. 2).^[22]

Next to killing the insect and supporting nematode development, *Xenorhabdus*'s specialised metabolites are engaged in repression of the insect's immune response upon infection.^[22] The insect recognises the microbial infection via pattern recognition proteins and activates immune responses such as the phenol oxidase (PO) response.^[23] Phenol oxidase is involved in melanisation to repel invading pathogens.^[24] The precursor pro-PO is transported to the haemolymph and finally activated as the last step of an enzymatic cascade which includes serine proteases.^[23] *X. nematophila* produces the isonitrile and aminoglycoside containing tyrosine derivative rhabduscin that inhibits the PO in nanomolar concentrations (Fig. 2).^[25] In addition, it was suggested that within this PO cascade, serine proteases 1 and 2 could be targeted by rhabdopeptide/xenortide like peptides (RXP) due to the structural similarity of RXP with known protease inhibitors.^[22] Generally, RXPs comprise between 2-8 amino acids, N-methylation and a terminal amine, often tryptamine or phenylethylamine.^[26] RXPs show strong activity against protozoa with an increased activity for longer and/or N-methylated peptides.^[26] The bacterial infection also triggers activation of phospholipase A2 (PLA2) which is required for the biosynthesis of eicosanoids which in turn are mediators in the insect's immune response.^[27] This initial enzyme, PLA2, is inhibited by the *Xenorhabdus* metabolites benzylideneacetone or oxindole, respectively.^[22,28]

As a part of the humoral immune response upon microbial infection, insects produce antimicrobial peptides (AMPs) and release them into the haemolymph.^[29] It was suggested that the hydroxamate siderophore putrebactin could inhibit the histone deacetylase which results in an impaired antimicrobial peptide production.^[22] The proteasome is an additional player in the activation of the insect immune response as degradation of repressor proteins can stimulate certain immune response pathways.^[30] The insect proteasome was proposed to be targeted by the dithiolopyrrolone xenorhabdin.^[22] Moreover, xenorhabdin exhibits strong activity against parasites such as *Trypanosoma* or *Plasmodium* and could also function as repellent of food predators.^[31]

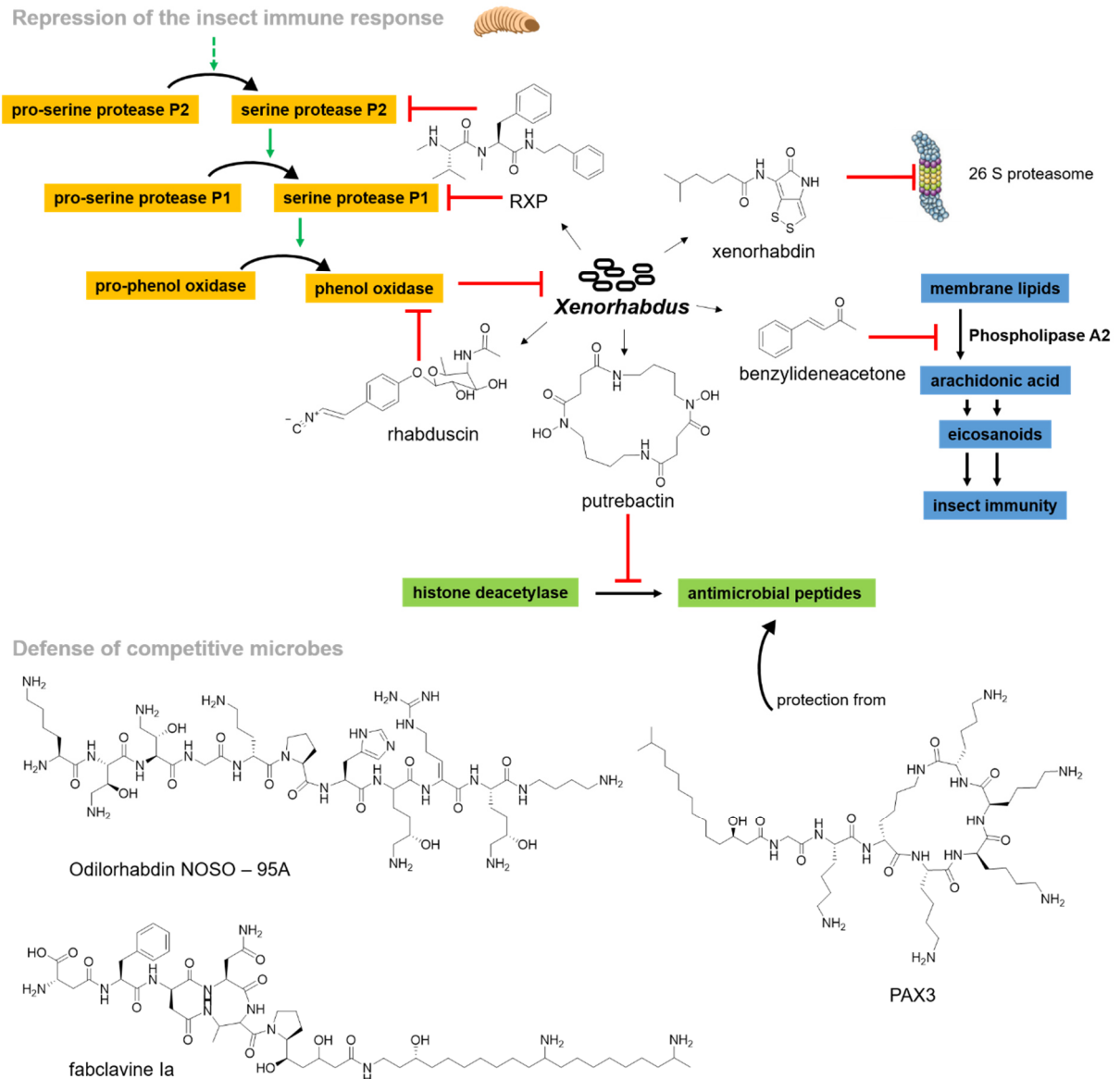


Figure 2. Selected secondary metabolites from *Xenorhabdus*. The metabolites are described to function in the defense of the insect's host immune response (top) or the defense of competitive microbes (bottom). Modified from Tobias *et al.* 2018.^[22]

Predator defense

The decapeptides odilorhabdins exhibit activity against Gram-positive and Gram-negative bacteria.^[32] Odilorhabdins were shown to bind to the small ribosomal subunit in a region that is not targeted by previously known antibiotics.^[32] Antifungal activity was observed for cyclic lipopeptides named PAX (Peptide Antimicrobial from *Xenorhabdus*) peptides.^[33] Recently, an additional function was assumed for PAX peptides that localise on the bacterial cell surface. It is suggested that the presentation of PAX peptides on the bacterial cell surface protects the bacterium from cationic antimicrobial peptides as part of the insect's immune response.^[34]

INTRODUCTION

Because of its broad range activity against bacteria, fungi and protozoa, the NRPS-PKS-hybrid product fabclavine is also proposed to function in the defense of food predators.^[35]

The presented specialised metabolites show a high structural diversity and the majority of them is biosynthesised with the help of NRPS enzymes. In the following chapters, these enzymes and their mechanisms are explained in more detail.

1.2.1 Gene regulation

The complex life style of *Xenorhabdus* requires a controlled gene expression and/or regulation in adaptation to the changing conditions during the life cycle. In bacteria, transcription and translation are linked in spatial and temporal aspects.^[36] An additional characteristic of prokaryotic gene expression is the organisation of functionally related genes under the control of one initial promoter in a so called operon.^[37] Various possibilities of gene regulation have been described and transcriptional regulators, two component systems and alternative sigma factors are common strategies among them.^[38] The sigma factor is required for the interaction between RNA polymerase and a promoter and often alternative sigma factors are engaged to coordinate stress responses.^[39] In two-component systems environmental conditions are sensed by a sensor kinase which transfers the phosphoryl group on to a response regulator which is often a DNA-binding transcription factor.^[40] Transcriptional regulators interfere at the level of transcription initiation and there can act as repressors or activators.^[41] Transcription factors are referred to as global or local based on the number of genes they regulate.^[42]

In *Xenorhabdus*, some transcriptional regulators were linked to the bacterial ecology and production of specialised metabolites. The global transcriptional regulator leucine-responsive regulatory protein O (LrpO) is associated with virulence modulation in *X. nematophila* switching from mutualistic to virulent behaviour.^[18] In addition, LrpO is described as activating regulator of the secondary metabolites xenocoumacin 1, xenortide A and rhabdopeptide 1 in *X. nematophila*.^[43] For the transcriptional regulator LeuO both activation and repression of different compounds were observed in *X. nematophila*.^[43] LeuO is assumed to activate production of xenotetrapeptide and xenematide A while it represses xenocoumacin and nematophin production.^[43] The LysR-type regulator HexA functions as repressor for the majority of secondary metabolites in *Photorhabdus*, closely related to *Xenorhabdus*.^[43,44] The RNA binding protein Hfq (host factor of the RNA bacteriophage Q β) is proposed to stabilise the mRNA of HexA.^[44] When *hfq* is deleted in *Photorhabdus*, this strain displays a significant reduction of secondary metaphosphobolite production and is not able to support mutualistic symbiosis with its nematode host from the genus *Heterorhabditis*.^[44] The small RNA ArcZ was identified to bind HexA mRNA and thus mediates the HexA control by Hfq.^[45] The overall reduced SM production also applies to *hfq* deletion in *Xenorhabdus* and this “null producer”

strain facilitates the search for novel metabolites upon induction of a certain BGC.^[46] This topic will be addressed in section 1.6.1 in more detail.

1.2.2 Thiotemplated assembly line in NRPS and PKS

NRPS and PKS enzymes belong to the family of megasynthases which show a modular organisation.^[47] Type A NRPS and type I PKS enzymes operate in a thiotemplated assembly line enzymology (Fig. 3B).^[47] In general, one module of the enzyme is responsible for the incorporation of one building block during the biosynthesis and the number of modules usually is collinear with the number of incorporated building blocks (type A NRPS and type I PKS).^[47]

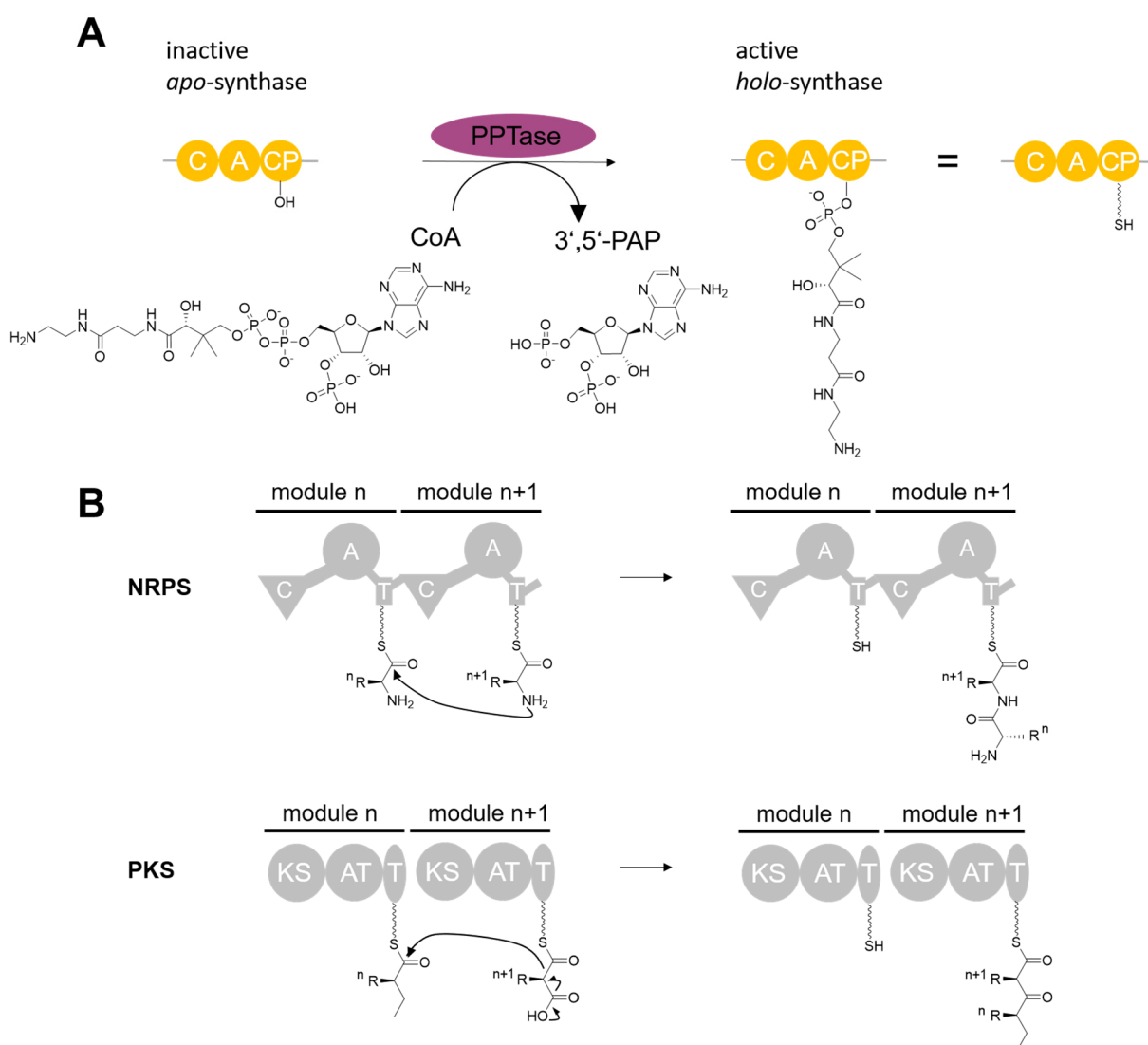


Figure 3. Phosphopantetheinylation in NRPS and PKS enzymes **A** PPTase transfers the phosphopantetheinyl arm from CoA to the (acyl-/peptidyl-) carrier protein (CP), also referred to as T domain. This modification allows the thiotemplated assembly line. Taken from Beld *et al.* 2014.^[48] **B** Collinear thiotemplated assembly-line enzymology for type A NRPS and type I PKS enzymes. Taken and modified from Du *et al.* 2001.^[49]

INTRODUCTION

The NRPS enzyme links two amino acid building blocks via peptide bond formation with the help of condensation domains.^[50] Similarly, a PKS condenses acyl moieties to generate polyketides using ketosynthase KS domains.^[47] In both enzymes the building blocks are thioester bound to carrier proteins (acyl- and peptidyl carrier domains), often referred to as thiolation (T) domains.^[47] The enzyme PPTase post-translationally modifies the inactive *apo*-synthase with the 4'phosphopantetheinyl (Ppant) arm from coenzyme A to yield the active *holo*-synthase (Fig. 3A).^[48] This PPTase step is essential for NRPS and PKS biosynthesis.^[48] The details of NRPS and PKS biosynthesis are introduced in the following chapters.

1.3 Non-ribosomal peptide synthetases

Non-ribosomal peptide synthetases are megasynthetases which show a modular organisation and operate in an assembly line enzymology to produce peptides.^[47] The classic NRPS module comprises an amino acid activating adenylation domain (A), a thiolation (T) domain to which the activated aminoacyladenylate is bound and a condensation (C) domain.^[51] At first, the adenylation domain activates the amino acid building block using ATP and then tethers the amino acid to the Ppant arm of the T domain.^[47] The condensation domain from the elongation module condenses the amino acid from upstream (donor site) with the downstream amino acid building block (acceptor site) (Fig. 4).^[51] The termination module releases the assembled product as a linear or cyclic peptide. This reaction is catalysed by the terminal domain, which is usually a thioesterase (TE) or terminal reductase (TR).^[51] It has to be noted that the domains undergo conformational changes during the catalytic cycle.^[52] The installed Ppant linkers additionally contribute to the flexible orientation of the substrates.^[47]

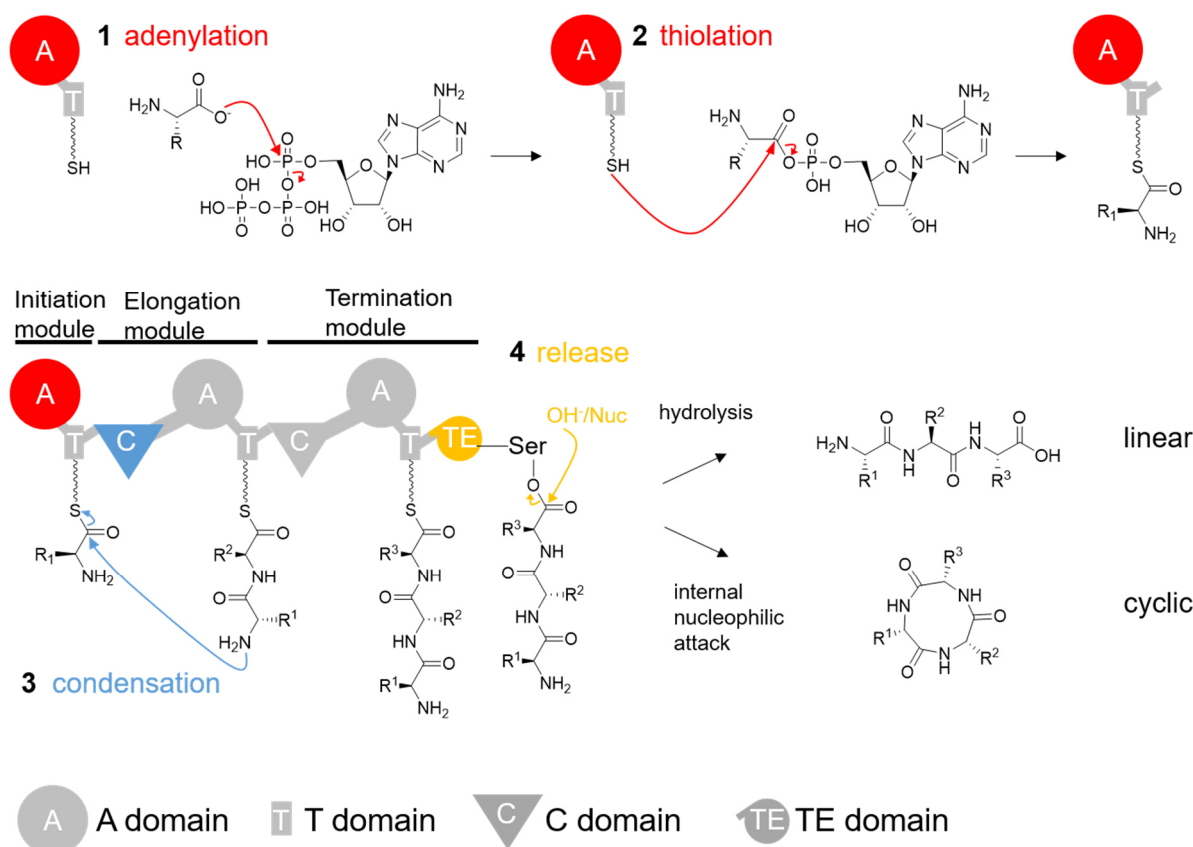


Figure 4. The typical NRPS cycle: adenylation (1), thiolation (2), condensation (3) and release (4) of the assembled NRP product. The A domain catalyses the adenylation and thiolation reaction (red). The C domain condenses two amino acids, each bound to the T domain (blue). The thioesterase releases the NRP as linear or cyclic product (yellow). Modified from Süßmuth & Mainz 2017.^[51]

In addition to the presented collinear type A NRPS further NRPS types have been described.^[53] Type B NRPSs show an iterative use of single modules which results in products in which the number of incorporated building blocks exceeds the number of modules in the synthetase.^[53] Type C NRPSs are non-linear and deviate from the standard C-A-T domain organisation.^[53] It is worth mentioning that also stand-alone NRPS enzymes exist such as the free-standing C domain VibH.^[51] The number of modules in an NRPS is highly variable as monomodular NRPSs are found as well as for example eighteen modules in the peptaibol synthetase, encoded in one single gene.^[54] An even larger number of building blocks is incorporated by the twenty-five module comprising syringopeptin synthetase which assembles from several proteins.^[55] The NRPS system can be encoded by one single gene or by multiple genes that encode separate subunits which assemble to form the megasynthetase as shown for tyrocidin synthetase.^[56] Interpolypeptide linkers can mediate the communication between the separate polypeptides and are presented in section 1.5.^[57] NRPSs are able to incorporate not only proteinogenic amino acids but also fatty acids, acyl moieties and non-natural amino acids. The

INTRODUCTION

details of this flexibility and the single domains found in NRPSs are presented in more detail in the following.

1.3.1 Adenylation domain

The adenylation domain belongs to the ANL (AcyI-CoA synthetases, NRPS adenylation domains and Luciferase enzymes) superfamily of adenyating enzymes.^[58] These are ATP and Mg^{2+} dependent enzymes and catalyse first the adenylation reaction to form the aminoacyladenylate and then transfer the amino acid to the Ppant arm bound to the thiolation domain.^[58] Crystal structures of the phenylalanine activating A domain (PheA) from gramicidin synthetase A (GrsA) allowed first insights into the mechanisms of the A domains.^[59,60] The A domain exhibits a larger subunit A_{core} and a more flexible smaller subunit A_{sub} and overall ten core motifs A1-A10 within the A domain were defined (Fig. 5A).^[59] During the catalytic cycle of the A domain, the orientation of the A_{sub} domain changes. Following binding of ATP, Mg^{2+} and the amino acid, the A_{sub} domain rotates from an open position to the thiolation conformation to allow the adenylation reaction.^[61] After pyrophosphate release the A_{sub} domain orients another approx. 140 °C prior to the thiolation reaction.^[61]

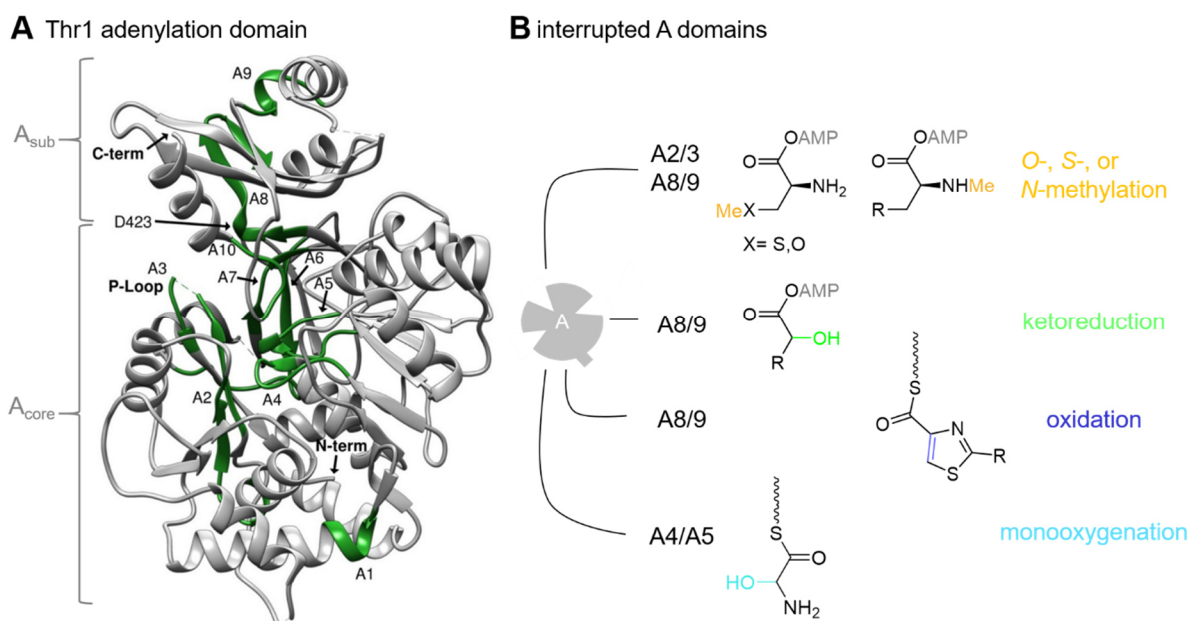


Figure 5. Structure of an adenylation domain and possible domain insertions. **A** Structure of free-standing Thr1 adenylation domain involved in 4-chlorothreonine biosynthesis in ribbon representation. The core motifs A1-A10 are presented in green colour. A_{core} and A_{sub} regions are indicated. Taken from Scaglione *et al.* 2017.^[62] **B** Overview of interrupted A domains. Methyltransferase, ketoreductase, oxidase and monooxygenase domains can be inserted in different positions in the A domain. The position of domain insertion is given. Taken and modified from Labby *et al.* 2015.^[63]

For selection of the amino acid ten residues were identified to play a role in the recognition of the amino acid building block.^[64] They are referred to as the Stachelhaus non-ribosomal code

and the identification of these residues allows the prediction of the activated substrate based on sequence information.^[65] Additionally, various biochemical assays were designed to experimentally determine A domain substrate specificity.^[66] In general, A domains show some substrate promiscuity for similar building blocks.^[67]

A domains contribute to the NRP diversity by being able to incorporate various building blocks, other than proteinogenic amino acids.^[68] Often these building blocks are provided from other pathways in the cell.^[68] But the modification can also take place after A domain activation and during the product assembly by functionalised domains of the NRPS.^[51] In some cases modifying domains are embedded within the adenylation domain.^[63] The domain insertions are found in different positions within the A domain, namely between core motifs A2/A3, A4/A5 and between A8/A9.^[63] Examples are methylation (M) domains, most commonly inserted in between core motifs A8-A9.^[63] In addition, M domain insertion occurs in between motifs A2 and A3 in the thiocoraline synthetase TioN.^[63] M domains have been shown to install methylations of N, S and O atoms.^[63] In yersiniabactin, C-methylation is installed by an M domain which is not integrated within an A domain but in between cyclisation and T domain.^[51] Additionally, modifying domain insertions between core motifs A4/A5 have been described for monooxygenase (MonoOx) domains catalysing α -hydroxylation of glycine in myxothiazol and melithiazol biosynthesis.^[63,69] Also observed in myxothiazol and melithiazol biosynthesis are the oxidase (Ox) domains which are integrated in between A8 and A9 core motifs and catalyse the oxidation of thiazolines to thiazoles and can be inserted in the A domain between A8 and A9 core motifs.^[69,70] Moreover, domains can be integrated into the NRPS following the T domain as in the bleomycin synthetase BImIII.^[70]

Adenylation domains can be accompanied by small MbtH-like proteins which are described to fulfil various functions in respect to NRP synthesis, among them supporting the adenylation reaction.^[71]

1.3.2 Peptidyl Carrier Protein (PCP) domain or thiolation domain

The peptidyl carrier protein (PCP) domain is also referred to as the thiolation domain and represents the smallest NRPS domain.^[72] The T domain forms a four- α -helix-bundle and is essential to NRP synthesis, although it is a non-catalytic domain (Fig. 6).^[51] The conserved serine residue in the second α -helix becomes post-translationally modified by the PPTase which catalyses the transfer of the PPant group from coenzyme A (Fig. 3).^[48] The transfer of the Ppant arm converts the *apo* T domain into its *holo* form.^[48] The A domain then catalyses covalent binding of the amino acid to the Ppant arm.^[73] The T domain is regarded to present the substrates to the relevant NRPS domains and therefore exhibits a certain flexibility.^[51] The T domain is shown to interact with other domains mainly via the α -helices 2 and 3.^[74]

INTRODUCTION

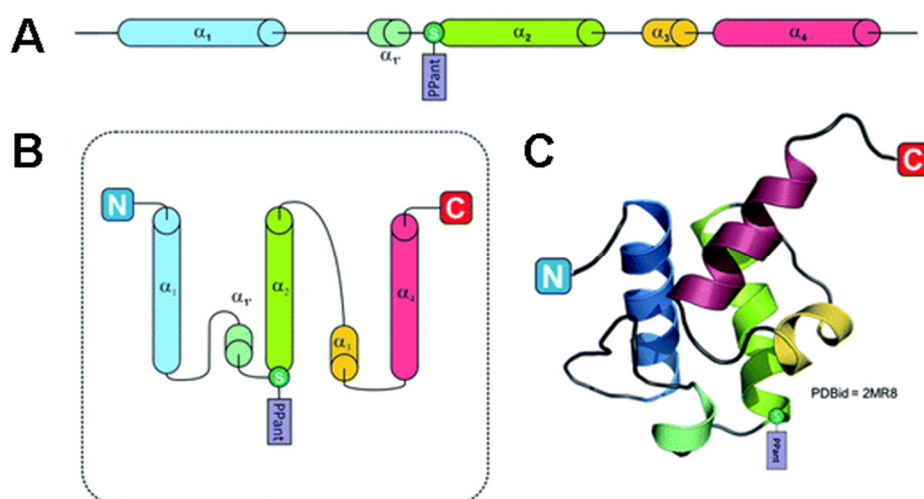


Figure 6. Schematic representation of the PCP domain. **A** Linear representation of typical PCP secondary structure. **B** Topology diagram of a canonical PCP. **C** Crystal structure of a PCP from module 7 of the teicoplanin-producing NRPS machinery, coloured according to panels A and B. The circled “S” represents the position of the conserved serine residue that bears the phosphopantetheine group added after protein synthesis. Figure and caption taken from Izoré & Cryle 2018.^[74]

1.3.3 Condensation domain

The condensation domain catalyses the condensation reaction between two PCP-domain tethered amino acid building blocks to yield an amide bond.^[50] The first C domain crystal structure was obtained from the free-standing C domain VibH involved in vibriobactin biosynthesis.^[75] The C domain forms a V-shaped pseudomer of two beta sheet lobes, the N-terminal and C-terminal lobes (Fig. 7).^[52] The two lobes are assumed to have arisen from gene duplication event.^[52] The C domain exhibits the conserved HHxxxDG motif in its active site for amide bond formation between the upstream donor and downstream acceptor substrates.^[50] The second histidine is held responsible for correct substrate orientation to allow peptide bond formation.^[52]

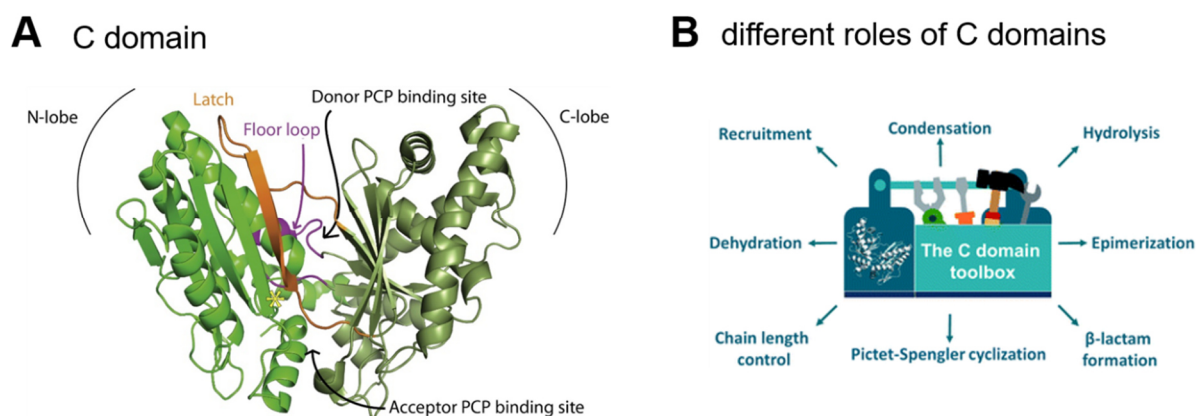


Figure 7. Structure and possible functions of a condensation domain. **A** Structure of the free-standing C domain VibH from vibriobactin biosynthesis. VibH shows a V-shaped pseudodimeric structure built from an N-terminal and C-terminal lobe. Entrance from T domain bound donor and acceptor substrates is indicated. The yellow asterisk marks the active site. Taken from Bloudoff & Schmeing 2017.^[52] **B** Overview of different functions fulfilled by C domains. Taken from Dekimpe & Masschelein 2021.^[76]

Condensation domains contribute to the high diversity of NRPs as they show a variety of different functions. Next to amide bond formation they are described to function in different ways including the formation of ester bonds, isopeptide bonds, β -lactams as well as the catalysis of Pictet-Spengler cyclisation, dehydration and Diels-Alder-cyclisation (Fig. 7).^[76] $L\text{-}C_L$ domains condense two L-amino acids while $D\text{-}C_L$ domains link a D-amino acid with a downstream L-amino acid.^[52] Epimerisation domains are phylogenetically related to C domains and catalyse the epimerisation of L- to D-amino acids.^[77] In addition to adjacent C and E domains, dual C/E domains are observed.^[51] It is possible to distinguish and predict these four types based on sequence information.^[77] C_{starter} domains as first domain of the NRPS condense an acyl moiety with the first amino acid presented by the downstream T domain (Fig. 8).^[77] The donor acyl moiety can be CoA bound or T domain bound and key residues within the C_{starter} domain were identified to be selective for the accepted chain length of the acyl moiety.^[78]

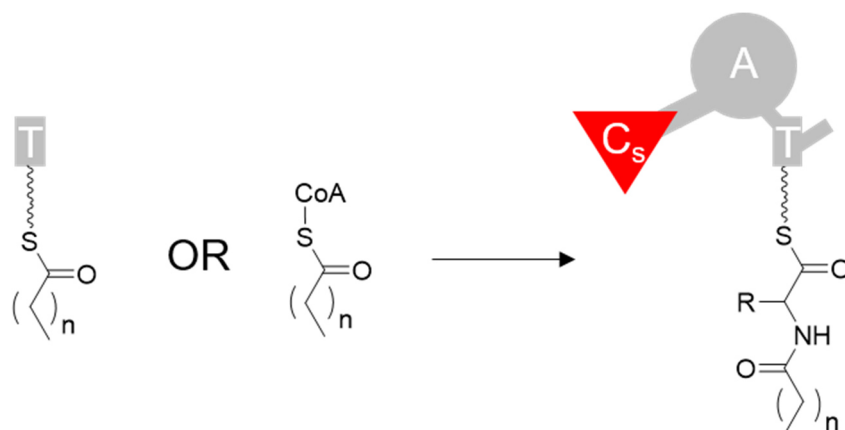


Figure 8. Standard reaction catalysed by C_{starter} domains. C_{starter} domain condenses an acyl rest with the downstream acceptor amino acid. The acyl chain can be T domain bound or CoA bound. Taken from Zhong *et al.* 2021.^[78]

In fungi, terminal C domains (C_{term}) were shown to function as last domain and to be responsible for the peptide release by catalysis of macrocyclisation (Fig. 9).^[79] The structure of the fungal C_{term} domain TqaA showed that the acceptor site is blocked by α -helix 2 as result of the replacement of an N-terminal loop of typical C domains with α -helix 1.^[80] The subtype hybrid C domain applies to domains that connect polyketide building blocks with downstream amino acid building blocks and thus represent a link between PKS and NRPS assemblies.^[81] Another example is the dehydration reaction of serine to dehydroalanine which is employed by the C domain subtype “modified AA”.^[81] Although their catalysed reactions are very diverse, C domains appear to show a so-called gate keeping function as each C domain is specific in its catalytic role and the substrates it accepts.^[52,82] To conclude, C domains are versatile in their function and can be located at different positions within the NRPS enzyme. They even occur as free-standing domains.^[52]

INTRODUCTION

1.3.4 Peptide release mechanisms

For release of the assembled peptide from the NRPS different strategies are observed and different domains, such as TE, terminal reductase or terminal C domain, can be involved (Fig. 9).^[51] TE domains belong to the family of α/β -hydrolases and can release the substrate in different ways.^[51] TE domains show a conserved catalytic triad of serine, histidine and aspartate which coordinates the transfer of the NRP product from the T domain to the TE domain.^[83] The then TE domain bound NRP can undergo nucleophilic attack, e.g. from water and be released as linear peptide.^[83] In addition, an intramolecular nucleophile, e.g. amine or hydroxyl group from the peptide side chain can result in macrocyclisation to form a macrolactone or macrolactame.^[83] While TE domains bind the substrate prior to release, terminal reductase domains act on the substrate which is bound to the T domain.^[51] Terminal reductase domains share similarities with short chain dehydrogenases (SDRs), usually depend on NAD(P)H and can catalyse two- or four-electron reductions.^[84] The obtained aldehyde from a two-electron reduction can also be processed by the R domain to form a macrocyclic product.^[51] In addition, some R domains, termed R* domains, conduct Dieckmann-type cyclisation yielding a tetramate unit.^[83] The terminal C domain which is prominent in fungi was already shortly introduced in chapter 1.3.3 and shows a release mechanism without covalent binding of the substrate that leads to intramolecular cyclisation.^[51] C domains positioned as the last domain in an NRPS have also been shown to condense the upstream building block with free/cytosolic amines resulting in an NRPS-released product.^[76] The melithiazol biosynthesis is an example where the monooxygenase domain integrated in the A domain of the last module of melithiazol synthetase catalyses α -hydroxylation of glycine.^[69] This modification results in the release of the C-terminal amide melithiazol precursor while the glyoxalate remains T domain bound.^[69] In conclusion, various mechanisms are described to promote product release and they all contribute to the structural diversity of NRPs.^[83]

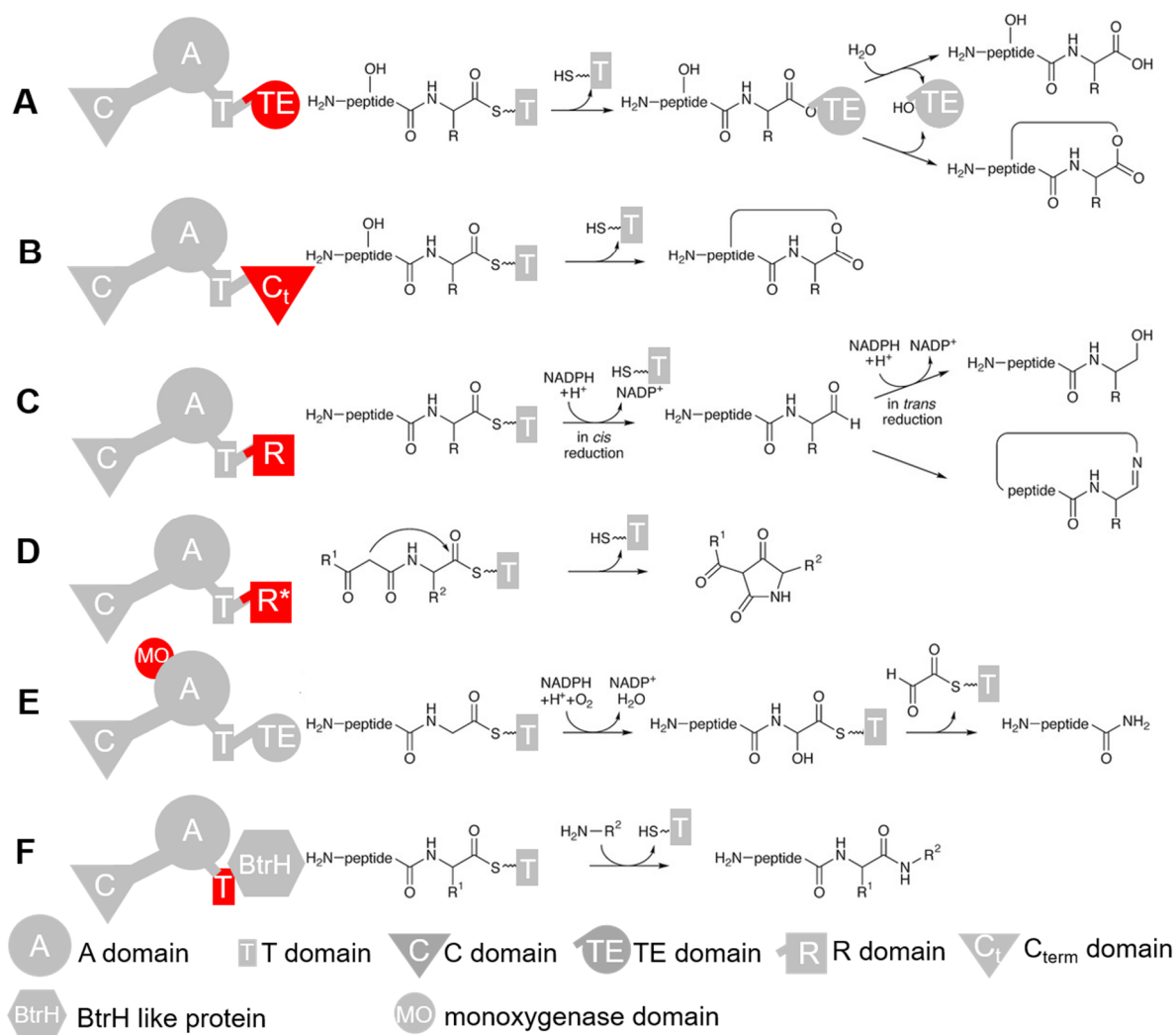


Figure 9. Overview of release mechanisms by NRPS enzymes. **A** TE domain binds the substrate as oxoester and promotes hydrolytic linear peptide release or intramolecular cyclisation. **B** Fungal C_{term} domains catalyse intramolecular cyclisation and thereby release the product. **C** R domains reductively release the product generating an aldehyde which can be further reduced or undergo cyclisation. **D** R^* domains conduct Dieckmann-type cyclisations of PK-NRP hybrids generating tetramates. **E** α -hydroxylation of glycine installed by the monooxygenase domain in the last module of melithiazol synthetase results in release of the amide product. **F** terminal T domains can recruit other enzymes for NRP release, e.g. BtrH like protein. Taken from Süßmuth & Mainz 2017.^[51]

1.3.5 Additional domains

In addition to the above presented domains, many more additional domains are described that contribute to the high structural diversity of NRPs. Among them are for example formylation (F) domains, M domains, epimerisation (E), cyclisation (Cy) or X domains.^[85] Formylation domains use formyltetrahydrofolate as donor to transfer formyl groups to the first activated building block.^[85] They are often positioned as initial domain, as shown in the case of linear gramicidin or anabaenapeptilide.^[86] E domains as well as Cy and X domains belong to the C domain superfamily.^[52] E domains catalyse the epimerisation of L- into D-amino acid and do not replace the C domains but are found adjacent to the T domain as tailoring domain.^[52] Cy

INTRODUCTION

domains can replace C domains as they first condense the donor with the acceptor substrate which is usually cysteine, serine or threonine.^[52] In a next step they reduce the side chain of the cysteine, serine or threonine and cyclise it with the carbonyl carbon of the before generated peptide bond.^[52] Thereby Cy domains introduce heterocycles to the assembled product.^[52] X domains are found in the terminal module of glycopeptide producing NRPSs and are assumed to orchestrate oxygenases to the peptide assembly to enable further modifications.^[87]

1.3.6 Tailoring enzymes and trans-acting enzymes

In addition to modifying domains that are encoded within the NRPS gene, also exogenously encoded enzymes can contribute to NRP biosynthesis. These can be tailoring enzymes that act on the released product or they can be trans-acting enzymes that are actively recruited to the NRPS assembly and when not properly presented stall the NRP synthesis.^[51] The glycopeptide UK-68 597 exhibits several post-NRPS modifications including chlorination, sulfation and glycosylation performed by tailoring enzymes encoded in proximity to the BGC.^[88] Furthermore, additional enzymes can be required for the provision of modified substrates or precursors that are incorporated by the NRPS.^[68]

1.3.7 NRPS-like enzymes

The termini minimal NRPS and NRPS-like enzyme refer to NRPS enzymes which exhibit an A and a T domain but lack a C domain.^[47,89] NRPS-like enzymes exhibit A-T-TE domain organisation and were shown to produce quinones such as atromentin or asterrequinone as well as furanone exhibiting compounds like microperfurane, ralfuranone or butyrolactone.^[90,91-93] The exact mechanisms for ring formation remain unknown but the thioesterase is suggested to play a key role in the formation of the backbone.^[94,95] In general, the adenylation domain of NRPS-like enzymes accepts aromatic α -keto acids and for the condensation reaction to be possible the first activated building block is oxoestered to the TE domain while the second building block is thioester-bound to the T domain.^[91,95] Thus the two building blocks are simultaneously bound by the NRPS-like enzyme.^[91] Both building blocks are activated by the same adenylation domain. Since NRPS-like enzymes usually activate aromatic α -keto acids, a specific code for α -keto acid activation was proposed.^[96] In ralfuranone biosynthesis the aminotransferase RaID was identified to convert L-phenylalanine into the corresponding α -keto acid.^[93]

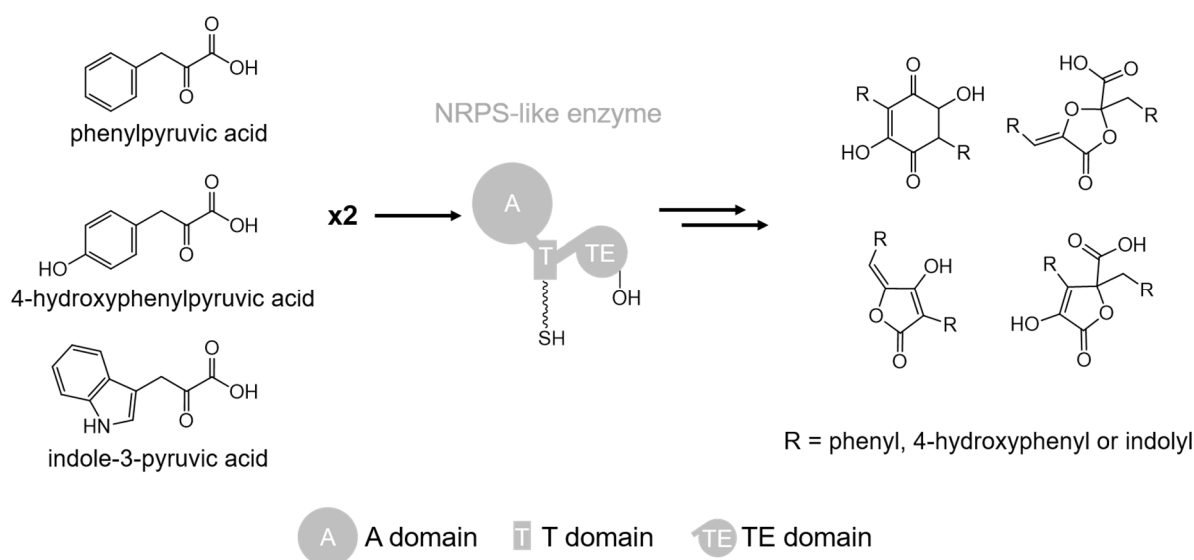


Figure 10. NRPS-like enzymes incorporate α -keto acids to yield furanone or quinone core structures. During biosynthesis, two identical building blocks are simultaneously bound to the NRPS, one oxoestered to the TE domain and one thioestered to the T domain. Modified after Wackler *et al.* 2012 and Hühner *et al.* 2019.^[94,96]

1.4 PKS

Polyketide biosynthesis resembles fatty acid biosynthesis and comprises activation of acyl units by an acyltransferase followed by decarboxylative Claisen condensation reactions to yield linear or cyclic polyketides.^[97] PKS enzymes are divided into three subtypes.^[97] Type I PKSs are multidomain enzyme complexes while type II PKSs are usually discrete enzymes that assemble to form the megasynthase.^[97] Type III PKSs act iteratively and do not rely on ACPs as they extend the substrate while it is CoA bound.^[97,98] This chapter mainly addresses type I PKS systems.

Type I PKSs form homodimers and can be further divided into iteratively or non-iteratively operating enzymes.^[99] In non-iterative PKSs, the ketosynthase conducts only one elongation cycle while the ketosynthase is repeatedly employed in iterative systems.^[97] Non-iterative type I PKS are distinguished in *cis*-AT PKS and *trans*-AT PKS systems depending on whether the AT domain is part of the enzyme or a discrete protein.^[97] The 6-deoxyerythronolide synthase (DEBS) involved in erythromycin biosynthesis represents a well studied example of a *cis*-AT type I PKS.^[97] A truncated version of this enzyme is depicted in Fig. 11 for simplified presentation of a PKS assembly.^[100] In standard *cis*-AT PKSs, the minimal PKS module comprises a ketosynthase domain, an acyltransferase domain and an acyl-carrier protein.^[97] The loading didomain (LDD) provides the first acyl unit for the condensation reaction with the downstream acyl unit bound to the T domain in module 1.^[100] The decarboxylative Claisen condensation is carried out by the ketosynthase domain in module 1.^[100] In this example, two

INTRODUCTION

more condensation cycles take place and the terminal domain, the thioesterase domain, releases the cyclic polyketide.^[100]

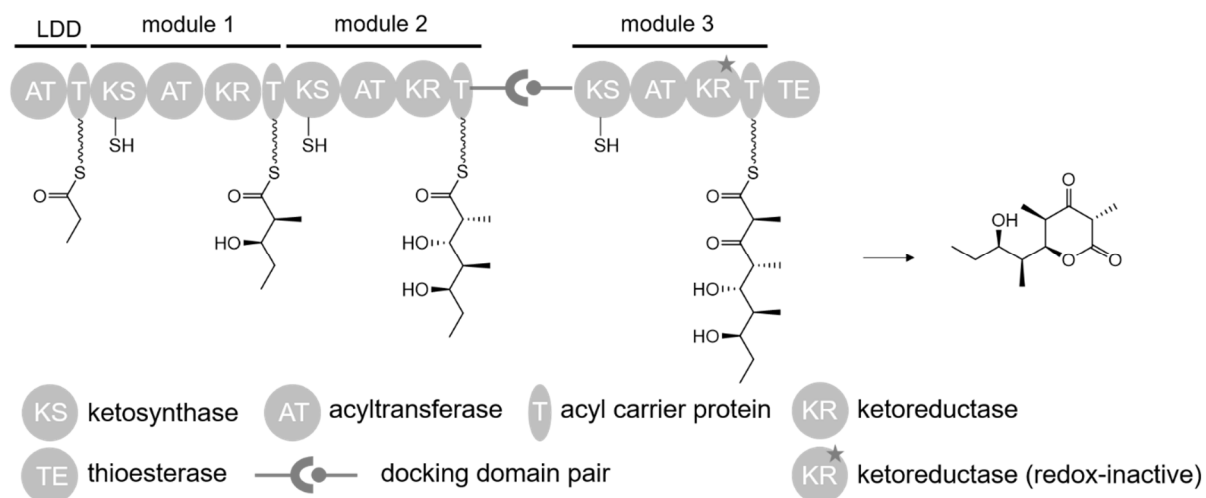


Figure 11. Exemplary type I *cis*-AT PKS assembly of a truncated DEBS enzyme. The loading didomain (LDD) incorporates the propionyl unit. Modules 1, 2 and 3 each extend a methylmalonyl unit, but the KR domain in module 3 is redox-inactive. The terminal thioesterase releases the cyclic tetraketide. Module 2 and 3 are encoded by distinct proteins but communicate via a docking domain pair. Taken from Robbins *et al.* 2016.^[100]

The LDD, module 1 and module 2 are part of one polypeptide chain and thus covalently linked. Module 3 is encoded by a separate gene but assembles as part of the megasynthase via so called docking domains (Fig. 11).^[100] Docking domains occur in pairs and link interpolypeptide chains. The donor domain is located at the C-terminus of DEBS1 and the acceptor domain is found at the N-terminus of DEBS2.^[101]

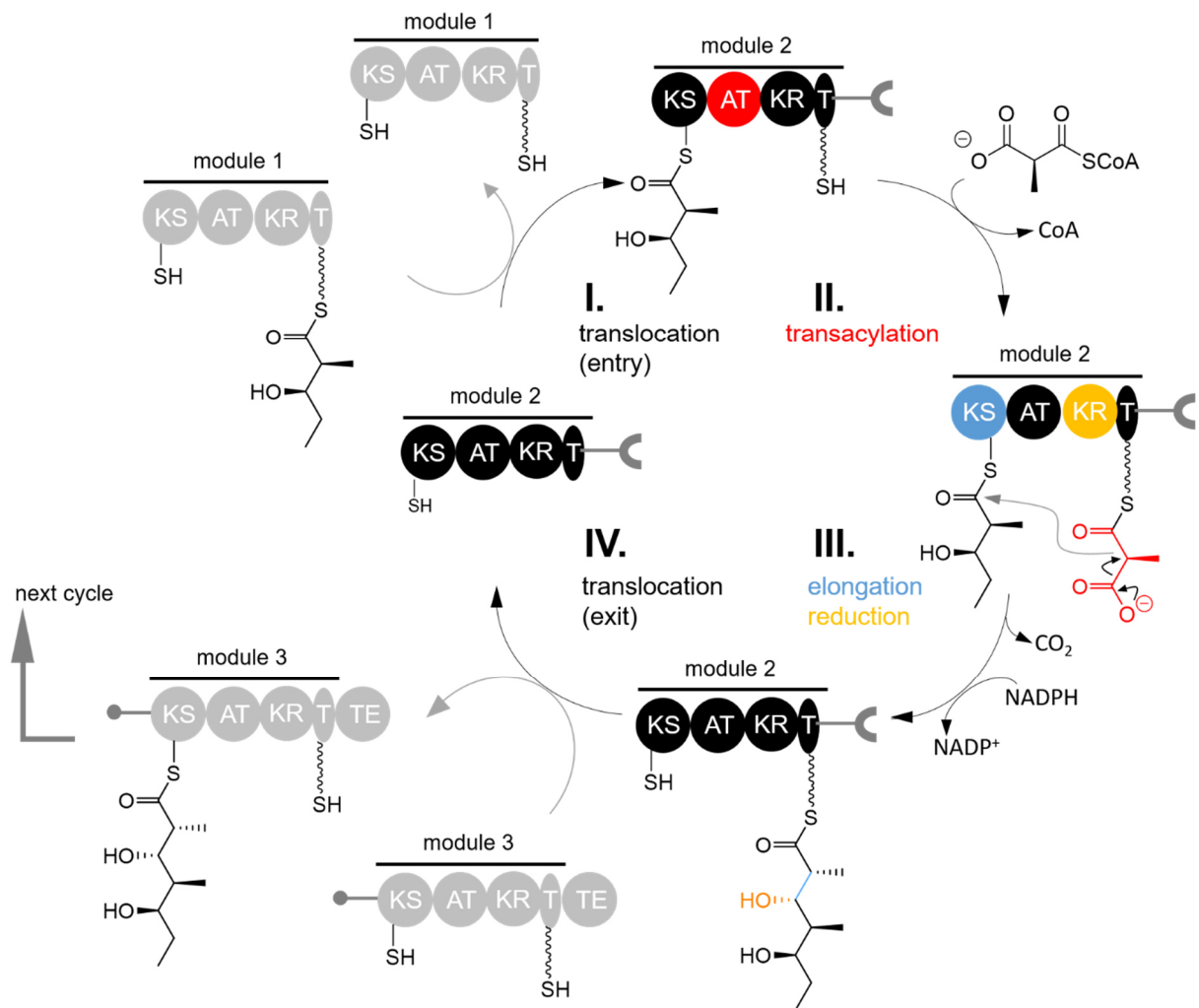


Figure 12. Chain elongation in PKS. Step I. Translocation of the diketide from the T domain of module 1 to the KS active site cysteine of module 2. Step II. Transacylation. AT catalyses the transfer of methylmalonyl extender unit from CoA onto the Ppant arm of the T domain in module 2. Step III. The ketosynthase catalyses a decarboxylative Claisen condensation. The ketoreductase reduces the β-keto group to the alcohol. Step IV. The extended triketide is translocated from module 2 T domain to module 3 KS, preparing this module for the next elongation. Taken from Robbins *et al.* 2016.^[100]

It is necessary for PK biosynthesis that the ACPs are post-translationally modified by PPTases with a Ppant arm as described in section 1.3.^[48] They will be equally referred to as T or ACP domains in this work. The acyltransferase loads the acyl substrate from CoA onto the Ppant arm of the T domain.^[100] Polyketide biosynthesis usually involves a few cycles of chain elongation and one general cycle is sketched in Fig. 12.^[100] At first, the initial acyl unit bound to the LDD T domain is transferred from the ACP to the KS domain where it is bound by a conserved cysteine residue (translocation).^[100] Next, the acyltransferase catalyses the transfer of the methylmalonyl unit from the coenzyme to the T domain (transacylation).^[100] The ketosynthase then catalyses the decarboxylative Claisen condensation between the KS bound diketide and the methylmalonate bound to the T domain.^[100] The ketoreductase then catalyses reduction of the keto group in β-position to the alcohol under NADPH consumption.^[100] The

INTRODUCTION

translocation of the formed triketide is transferred to the ketosynthase of the downstream module 3.^[100]

1.4.1 Acyltransferase

Acyltransferases load the acyl extender unit onto the Ppant arm of the ACP domain and this process requires two steps (Fig. 13). At first, the AT domain recognises the acyl coenzyme and binds the acyl unit.^[102] Next, the acyl unit is translocated to the Ppant arm of the T domain.^[103] AT domains specifically select their substrate and thus are regarded as gatekeepers for the incorporated acyl unit.^[104] Most AT domains incorporate malonyl or methylmalonyl extender units but also different extender units, e.g. chloroethylmalonyl, have been described.^[105]

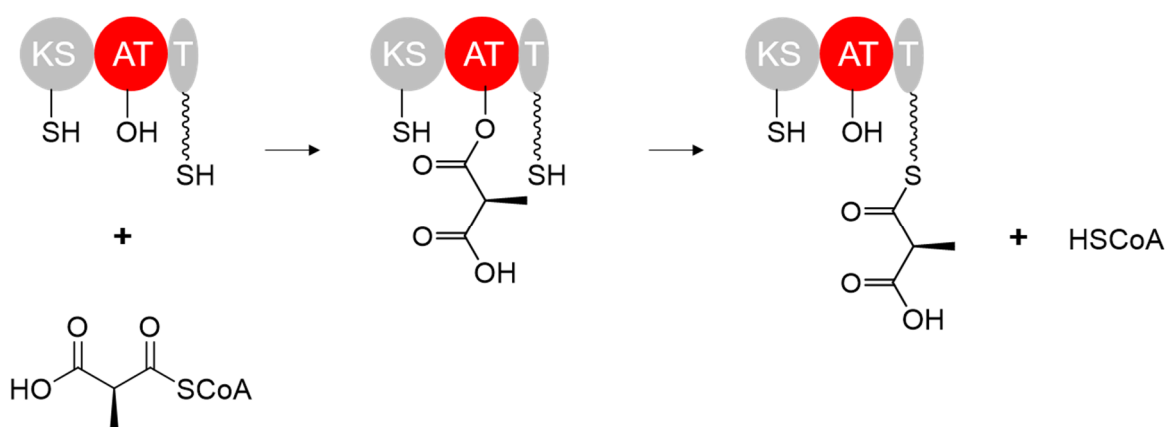


Figure 13. Schematic mechanism of the acyltransferase (AT) domain. AT domain catalyses transfer of the methylmalonyl extender unit from coenzyme A to the Ppant arm of the T domain via self-acylation by binding methylmalonate. Modified from Dunn *et al.* 2013.^[102]

1.4.2 Acyl Carrier Protein

ACPs comprise approximately 80 aa residues and form a four α -helix bundle.^[99] The PPTase posttranslationally modifies the serine in the (D/E)xGDxSL motif of the acyl carrier protein to turn the *apo* protein into the *holo* protein which then carries the Ppant arm. The ACP domain is involved in binding the acyl substrates prior to and following chain elongation and orchestrates the substrates between the domains.^[99]

1.4.3 Ketosynthase

Ketosynthases are β -ketoacyl ACP synthases that install the carbon-carbon bond following the decarboxylation (decarboxylative Claisen condensation).^[99] The KS domain exhibits conserved catalytic residues which include two conserved histidines and the cysteine residue which binds the upstream substrate.^[99] Mutation of the cysteine abolishes chain elongation but

still allows the KS to catalyse decarboxylation.^[106] A serine in proximity to the active site is proposed to link decarboxylation and chain elongation.^[106]

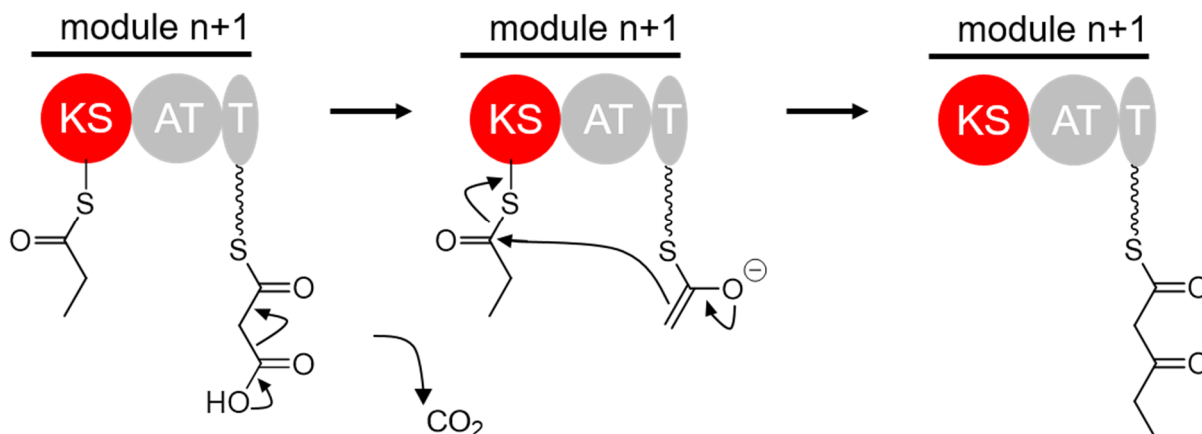


Figure 14. The ketosynthase (KS) domain and the decarboxylative Claisen condensation reaction. The ketosynthase domain binds the upstream propionyl unit via a conserved serine. Decarboxylation of the downstream T domain bound substrate allows nucleophilic attack of the enolate on the thioester. Taken from Chan *et al.* 2009.^[105]

New insights into the evolutionary development suggest that KS domains and upstream domains move together.^[107] These findings led to a refined definition of the PKS module now comprising the KS as terminal domain and the upstream processing domains.^[103,108] This is also shortly addressed in chapter 1.6.2.

1.4.4 Modifying domains: KR, DH, ER

Next to the KS, T and AT domain, a PKS module can comprise additional modifying domains.^[109] These include for example ketoreductase (KR), dehydratase (DH), enoylreductase (ER) or methyltransferase (M) domains.^[99] KR, DH and ER domains sequentially can reduce the β -carbonyl group, similarly to fatty acid biosynthesis.^[110] With respect to the generated product, these PKS are referred to as non-reducing, partially reducing and highly reducing PKS.^[97] An overview is given in Fig. 15.

INTRODUCTION

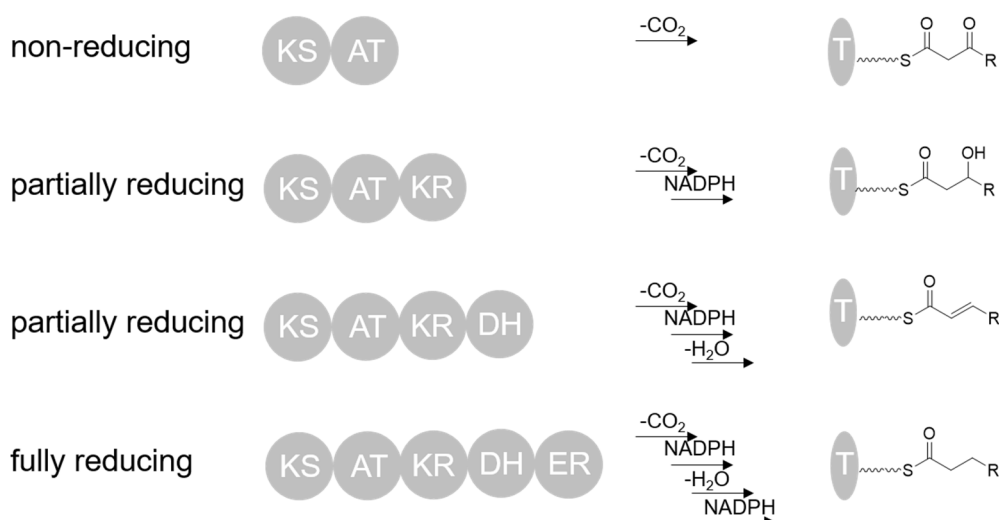
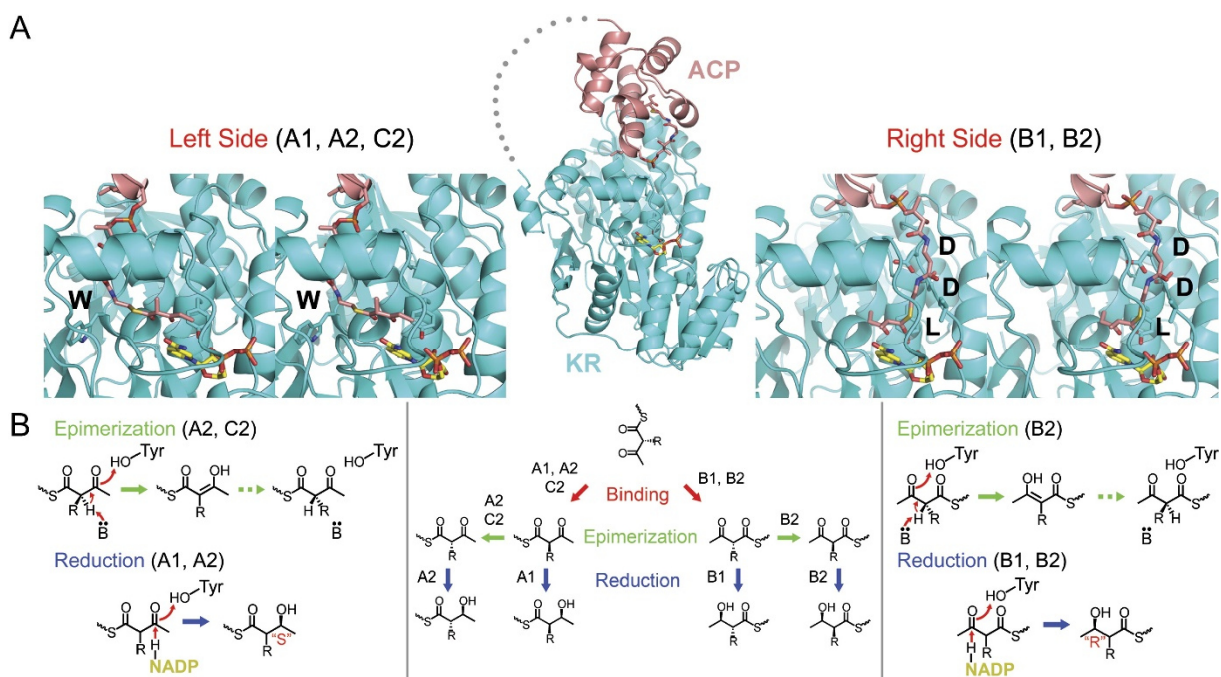


Figure 15. Non-reducing, partially and fully reducing PKS modules. KR, ER, DH domains process the β -carbonyl. Modified after Grninger 2020.^[111]

Ketoreductase domains belong to the SDR family and use NADPH for the reduction of the β -keto group to the alcohol function.^[112] KR domains are composed of a structural subdomain and a catalytic domain which harbours the catalytic residues tyrosine and serine.^[113] Dependent on their stereo outcome KR domains can be distinguished in different subtypes (Fig. 16).^[114] The catalysis of type A and type B KR is highly similar but the substrate orientation determines the stereo configuration of the ketoreduction.^[114] Type A KR domains generate S hydroxyl groups as their substrate enters the active site from the left side due to contact with a conserved tryptophan residue.^[114] The LDD motif in type B KR causes the substrate to enter the active site from the right side which results in R hydroxyl groups.^[114] Type C KR domains are not able to catalyse ketoreduction but C2 types may epimerise the α substituent.^[114]



The installed alcohol group can be further reduced by the DH domain resulting in an α,β -unsaturated bond.^[99] So far the stereospecific outcome cannot be predicted from the DH sequence but trans and cis configurations were observed.^[99,115] The ER domain is a member of medium chain dehydrogenase/reductase (MDR) superfamily and requires NADPH to reduce the enoyl.^[99]

1.4.5 Polyketide release mechanisms

Similar to NRPSs, PKS release can be mediated by C-terminal thioesterase (TE) and reductase (R) domains.^[83] TE domains can release the linear polyketide or catalyse the formation of macrocyclic products via the attack of an intramolecular nucleophile.^[83]

R-domains are NAD(P)H dependent short-chain dehydrogenase/reductases (SDR) that show a Rossmann fold and reductively release the product from the megasynthase.^[83] A first round of two electron reduction under consumption of NADPH yields the released aldehyde product via formation of a thio-hemiacetal intermediate of the T domain bound substrate.^[84] Following the aldehyde release, four electron reductase domains can perform a second two electron reduction generating an alcohol function, as observed in myxochelin biosynthesis.^[84] In contrast to the primary alcohol group resulting from the two subsequent two electron reductions, the aldehyde function of the single two electron reduction is often further

INTRODUCTION

processed.^[84] Often, linear peptides with a terminal aldehyde undergo macrocyclisation via imine formation as for example in nostocyclopeptide M1.^[84] The aldehyde can also be maintained after the first reduction and in this case often found in proteasome inhibitors.^[84] Barajas *et al.* 2015 solved the terminal R domain structure from the myxalamid synthetase in the NADPH bound state proposing that the C-terminal subdomain is responsible for cofactor binding and substrate specificity.^[116] Kinetic analysis led them to propose that the first reduction to the aldehyde is the slower reaction and that product release prior to the next reduction might be the rate-limiting step.^[116] Terminal R domains are also shown to catalyse Dieckmann cyclisations as mentioned before in chapter 1.3.4.^[83] In addition to TE and R domains, additional enzymes can be recruited to the assembly for product release, too.^[83] These include acyltransferases, lactamases or Baeyer-Villiger oxidases.^[83]

Following PKS release, discrete and/or tailoring enzymes can modify the released polyketide.^[117] These can include glycosyltransferases, oxidoreductases, methyltransferases, halogenases or acyltransferases.^[117]

1.5 Docking domains and hybrids of NRPS and PKS enzymes

NRPSs and PKSs can assemble from different protein subunits encoded by separate genes. Interpolypeptide linkers, so called docking domains, mediate the correct orientation and order of the different polypeptide chains in NRPSs, PKSs and hybrids thereof.^[118] The next section gives a short overview of identified docking domain pairs.

1.5.1 Docking Domains

Docking domains were first described for PKS and shortly thereafter for NRPSs in the tyrocidine synthetase.^[56,101] Similarly, linker regions were reported for NRPS-PKS hybrids from epothilone biosynthesis.^[119] An increasing number of different docking domain types is identified with ongoing research.^[120] Many docking domains exhibit α -helices but overall domain lengths and structures vary between the different observed types.^[120] *Cis* AT PKS docking domains usually form α -helical bundles and are categorised in two groups.^[120] Class 1 docking domains exhibit four α -helix bundles and are often found in actinobacteria.^[120] The second group show eight α -helix bundles and occur in PKS as well as PKS-NRPS hybrids.^[120] In NRPS, the originally termed communication-mediating (COM) domains are often unstructured and depend on their adjacent domain.^[120] In addition to α -helical elements, β -hairpin elements are found in NRPS docking domains, as for example in the N terminal docking domains from rhabdopeptide synthetases.^[120,121] Recently an additional type, composed of a three α -helical bundle was reported for the docking domain pair in the PAX producing NRPS from *Xenorhabdus*.^[122]

1.5.2 NRPS PKS and PKS-NRPS hybrids

Following the different NRPS linearity types there are also natural products derived from hybrids of NRPSs and PKSs.^[123] These products can be distinguished in two groups based on functionalisation of the two enzyme classes during the assembly.^[49] “Simple” PKS-NRPS hybrids that do not rely on direct functionalisation are for example cyclosporin A, microcystin, surfactin or coronatine.^[49] For coronatine formation a discrete ligase is responsible for fusion of NRPS and PKS product moieties.^[49,124] In other cases, the assembly line requires direct contact of the megasynthases.^[49] A prominent example is the biosynthesis of epothilones which features a PKS-NRPS-PKS enzyme organisation.^[119] C and KS domains can be regarded as interface between PKS and NRPS enzymes by condensing substrates of ACPs and PCPs.^[81] Ziemert and colleagues introduced the term hybrid C domains and described them to condense amino-acids with acyl extender units in PKS-NRPS hybrids.^[81] In conclusion hybridisation between NRPS and PKS contributes to the structural diversity of products generated by these megasynt(et)ases.^[123]

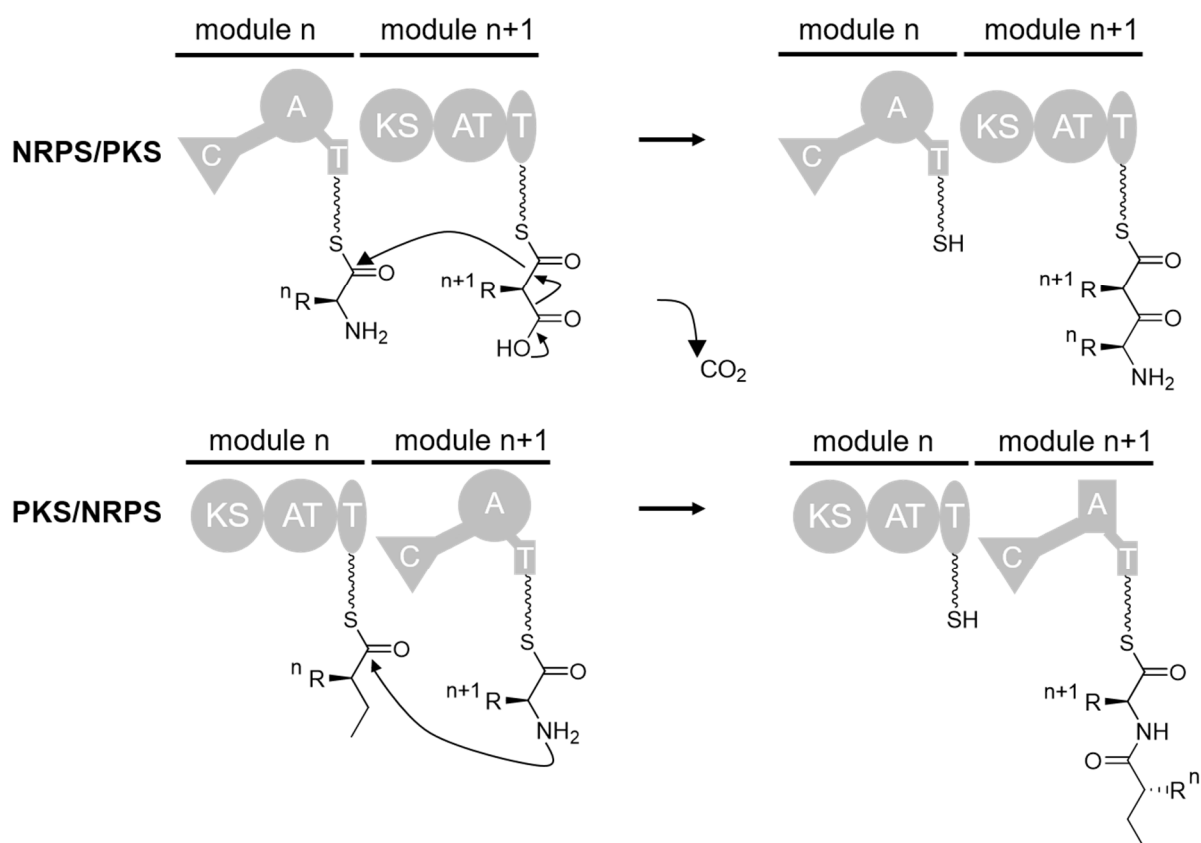


Figure 17. Hybrids of NRPS and PKS modules. The KS domain in a NRPS-PKS hybrid catalyses condensation between upstream amino acid and downstream acyl moiety. The hybrid C domain in a PKS-NRPS hybrid condenses the downstream amino acid with an upstream acyl moiety. Taken from Du *et al.* 2001.^[49]

1.6 NP identification and NP diversification (engineering)

1.6.1 NP identification traditional and targeted approaches

Traditional approaches for secondary metabolite identification are based on strain cultivation and subsequent compound purification, often combined with screening for biological activity.^[125] Altering the cultivation parameters, e.g. medium ingredients, cultivation flasks, oxygen rate or even the use of additives such as enzyme inhibitors can result in the production of additional or modified compounds and is often referred to as the One Strain MAny Compounds (OSMAC) approach.^{[126][127]} Additionally, cocultivation of two or more strains can lead to the discovery of novel compounds.^[128] Compound production is assumed to be only conducted when the products are required and might be triggered by environmental conditions such as interspecies communication in microbial competition or stress, e.g. nutrient depletion.^[129] The above described approaches for NP identification are often bioactivity guided and do not require genomic information.

The identification of the first biosynthetic genes as origin of natural product biosynthesis including NRPs and polyketides and the findings that the genes occur in cluster organisation were seen as the beginning of a new era of natural product research.^[130] As stated by Medema *et al.* 2015, a BGC comprises a set of two more genes which are found in physical proximity in the same genome and encode enzymes that are responsible for the biosynthesis of a specialised metabolite.^[131] BGCs that are not expressed under the investigated laboratory conditions are often referred to as “silent” but can now be predicted from bioinformatics and in turn be specifically activated for compound identification.^[132] Genome mining tools like ANTibiotics and Secondary Metabolite Analysis Shell (antiSMASH), Prediction Informatics for Secondary Metabolomes (PRISM) and Secondary Metabolite Unknown Regions Finder (SMURF) allow the identification of BGCs based on sequence information.^[133] In general, BGCs are detected based on similarity to conserved domains of core biosynthetic enzyme domains by Hidden Markov Model (HMM) profiles.^[133] When a candidate BGC is identified, various methods for the activation can be employed. They include heterologous expression of this BGC as well as deletion or overexpression of a specific transcription factor, if present.^[132] In a less targeted approach, global transcription factors are used.^[132] An example for such an untargeted approach is the expression of *X. nematophila* global regulator LeuO in *X. szentirmaii* which allowed the detection of two additional compounds.^[43] A highly efficient strategy for targeted BGC activation is the replacement of the natural promoter with a controlled inducible promoter.^[134] In *Xenorhabdus* the exchange of the natural promoter with the arabinose inducible P_{BAD} promoter is well established.^[31] This is achieved by conjugation of an integrative plasmid from an *E. coli* donor strain to the *Xenorhabdus* recipient. The plasmid carries the P_{BAD} in front of the first 300-600 bp of the target gene and therefore is integrated via homologous recombination in the genome. Hence, the P_{BAD} promoter is placed in front of

the target gene and can be activated with L-arabinose. Finally, LC-MS-based analysis allows compound identification for the investigated BGC (Fig. 18).^[31] In a more advanced manner, this targeted BGC activation can be applied in *Xenorhabdus* Δhfq strains and is then referred to as “easy Promoter Activated Compound Identification” (easyPACId).^[46] In this combined approach, the overall reduced compound production in Δhfq strains facilitates compound identification and bypassing purification prior to bioactivity testing. Ideally, culture extracts of the induced *Xenorhabdus* Δhfq P_{BAD} strain only show signals for the product(s) from the specifically activated BGC (Fig. 18).

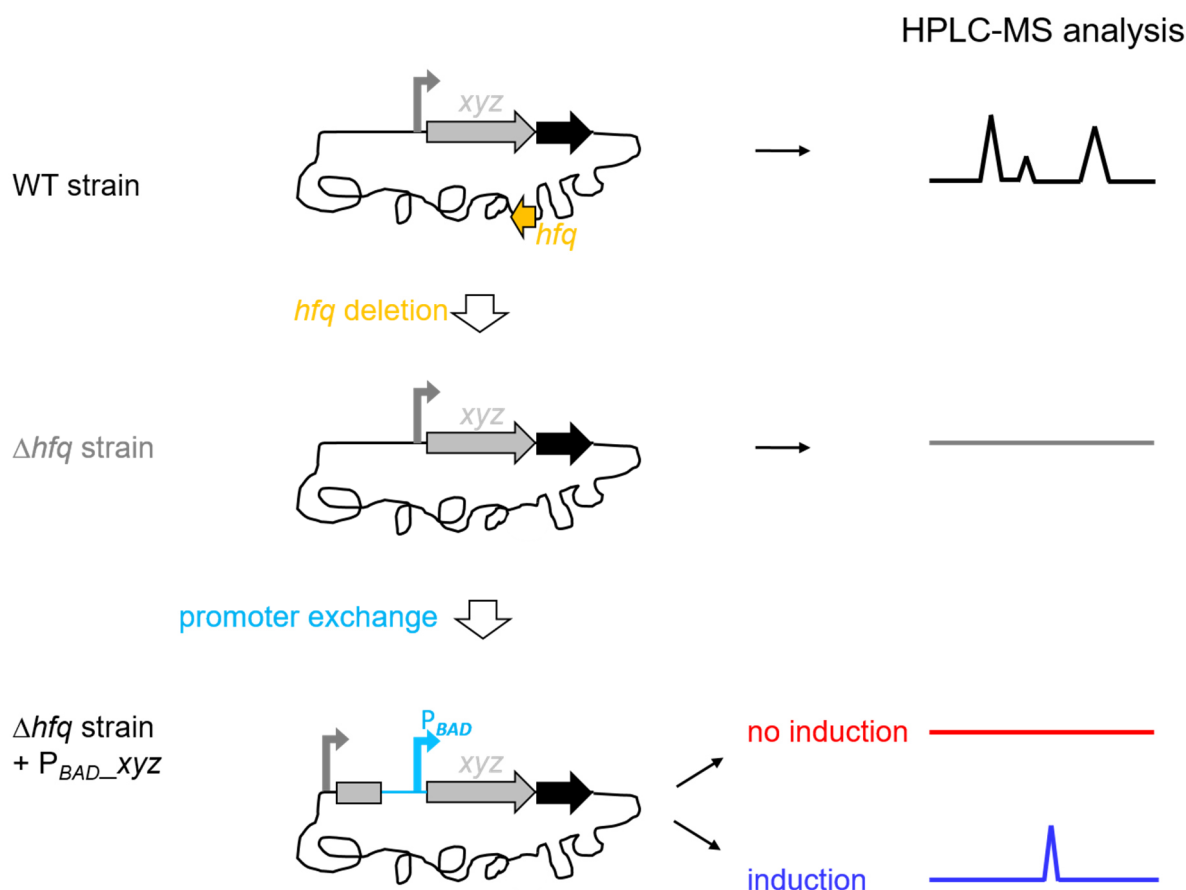


Figure 18. Facilitated compound identification by promoter exchange in Δhfq strains. Deletion of *hfq* in *Xenorhabdus* results in reduced secondary metabolite production. Promoter exchange based activation of target BGCs in Δhfq strains facilitates compound identification (easyPACId). Simplified from Bode et al. 2019.^[46]

NP and NP biosynthesis identification using LC-MS-based strategies

Following compound identification and purification, nuclear magnetic resonance (NMR) techniques are applied for structural elucidation which ultimately tells whether the studied molecule had previously been known.^[135] Low compound titres can complicate purification and subsequently limit NMR analysis. Different LC-MS-based strategies have been developed to gain structural information without previous purification and a few of them are presented in the

INTRODUCTION

following. First of all, high resolution mass spectrometry (MS) allows sum formula predictions for the investigated compound.^[136] Growth of the producer strain in ¹³C and ¹⁵N labelled medium results in the production of the compound with an increased molecular weight. This can be used to confirm the predicted sum formula from HR-MS data.^[137] In an “inversed feeding” experiment, non-labelled building blocks, e.g. amino acids, can be fed to the production cultures grown in ¹³C labelled medium. The incorporation of the “lighter” building block shifts the compound’s signal to a respective lower mass.^[137] In addition, comparison of retention time and MS² fragmentation in an LC-MS analysis between identified compound and a synthetic standard can be used in structure determination.^[138]

LC-MS analysis is also suitable to investigate how the compounds are produced. Required experiments can include heterologous expression of the BGC, specific gene deletions, precursor complementation, offloading with chain terminators or megasynthase engineering.^[139] ^[78] ^[140] Engineering approaches from NRPS and PKS enzymes are presented in more detail in the next chapter.

1.6.2 Generating “new” natural products by NRPS PKS engineering approaches

The assembly line enzymology and its modular architecture make NRPS and PKS interesting tools for applications in synthetic biology.^[141] A plethora of (re-)engineering strategies have been discovered and this chapter will provide a brief summary of selected strategies, reviewed by Beck, Garzon & Weber. Engineering approaches can be applied to enhance the biological activity, to introduce chemical labels which can be used for localisation and enzyme characterisation studies as well as to study details of the biosynthesis.^[142]

Classical architecture of type I PKS and NRPS systems

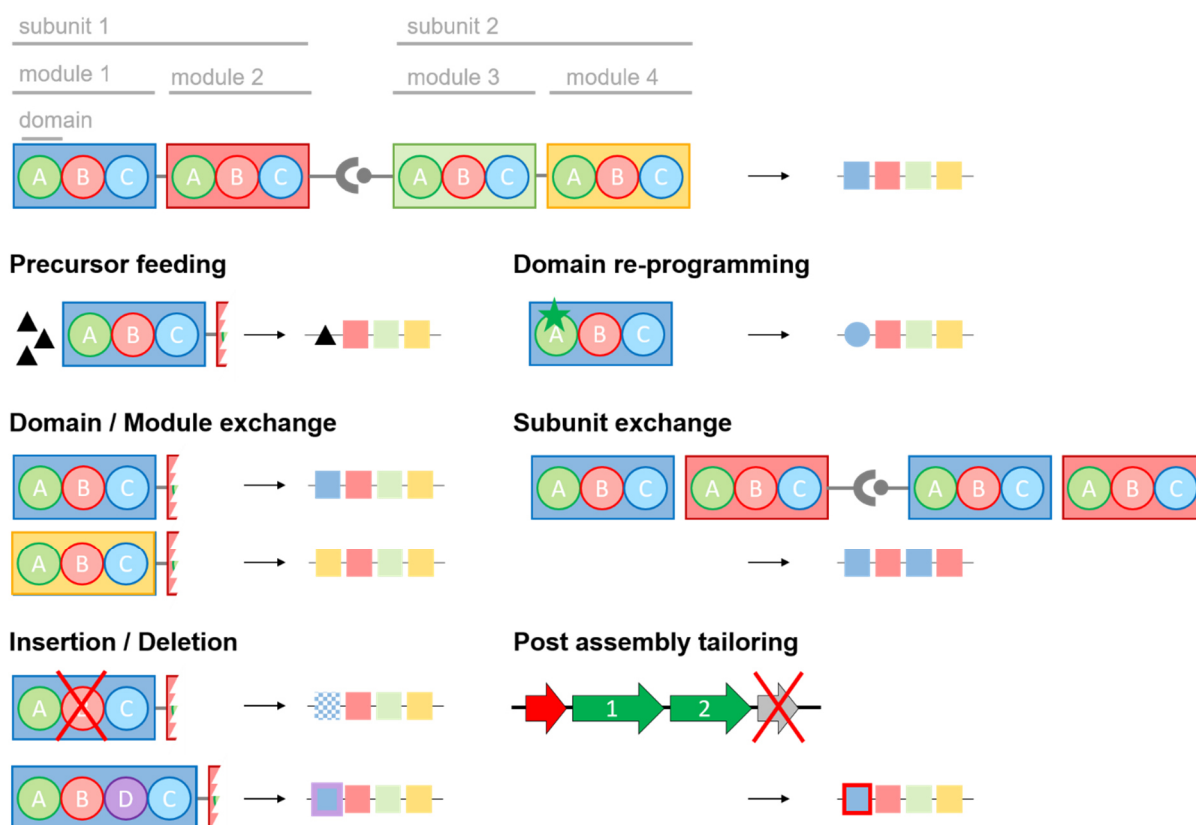


Figure 19. Overview of engineering strategies in NRPS and PKS systems. A, B, C can be NRPS or PKS domains. Rectangular box depicts one module. Subunit 1 and subunit 2 are two different polypeptides but communicate via docking domains (grey). Modified from Alanjary *et al.* 2019 and Beck *et al.* 2020.^{[143][142]}

Precursor directed biosynthesis and site-directed mutagenesis

The first strategies to modify the products of megasynthases included the use of alternative precursors which are endogenously added to the culture.^[142] The incorporation of substrates that are similar to the natural one is based on the observed substrate promiscuities of the megasynthases.^[143] A limitation of this precursor feeding is that the provided precursor is in competition with the natural substrates.^[142]

To overcome this limitation, site directed mutations can be conducted to alter the domain specificity beyond the natural promiscuity. For example, the site directed mutagenesis in one acyltransferase domain of erythromycin synthesising PKS allowed the incorporation of an alkyne group via the acceptance of propargyl-malonate as building block.^[144]

Domain/module exchange

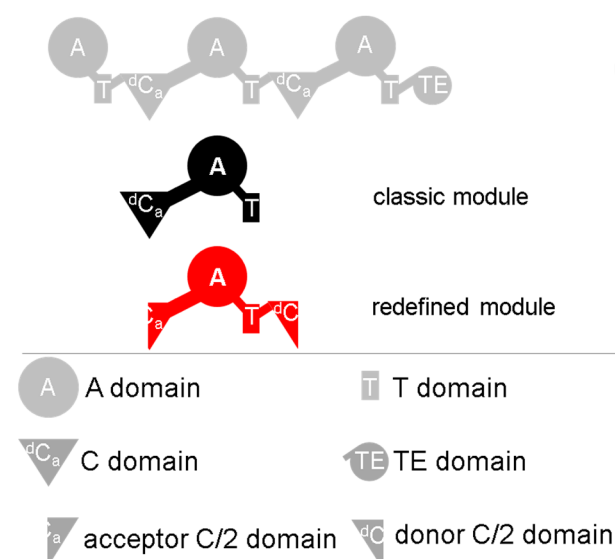
Another approach to diversify the product's structure is the exchange of single domains or modules and a variety of different strategies was described.^[142] Recently, a redefinition for both PKS and NRPS modules was proposed in respect to successful engineering experiments with

INTRODUCTION

exchange units that suggested an altered module perception (Fig. 20).^[143] In PKS the consideration of the downstream KS instead of the upstream KS resulted in the exchange of AT-T-KS units.^[143] In NRPSs, the identification of a fusion site within the C domain between the acceptor site (C_a) and donor site (dC) lead to the introduction of XUC units comprising C_a -A-T- dC .^{[143], [145]}

Classic and redefined modules of

A) an NRPS



B) a PKS

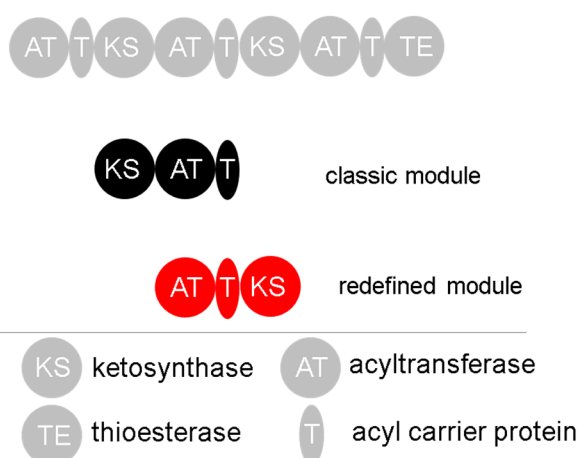


Figure 20. Classic definition and proposed redefinition of NRPS and PKS modules. **A** The redefined NRPS module comprises upstream half C domain with its acceptor site, A domain, T domain and downstream C domain with its donor site. **B** The redefined PKS module comprises AT domain, T domain and downstream KS domain. Taken from Alanjary *et al.* 2019.^[143]

Subunit exchanges

The exchange of subunits refers to the combination of different enzymatic units from separate polypeptide chains.^[142] This can be achieved by engineering of the docking domains of subunits to mediate the communication between them.^[120] In addition, docking domains can be introduced within an one protein NRPS and upon which an increased peptide yield was reported.^[146] Recently the possibilities of combining NRPS enzymes were enhanced with the introduction of synthetic zippers.^[147] Synthetic zipper pairs were shown to efficiently link two separate polypeptide chains, even allowing the generation of peptide libraries.^[147]

Insertion / deletion of domains

The insertion and/or deletion of modifying domains represents another engineering strategy. In PKSs, the insertion of DH and ER domains changes the configuration and oxidative state of the addressed building block.^[148] In NRPSs, A domains can be targeted to include additional

domains as was shown for the A domain of echinomycin synthetase in which an M domain was inserted.^[149]

Tailoring enzymes

Additionally, tailoring enzymes can be used to diversify product structure as they are shown to install further modifications on the product after it is released by the megasynthase.^[88] Common modifications include halogenation, acylation or glycosylation.^[88] Omics approaches can help to identify tailoring enzymes of NP pathways and once identified the enzymes can be applied in engineering.^[150]

Obviously, the different strategies can also be combined as was shown in the example of an alkyne polyketide product where docking domains as well as site directed mutagenesis of the carrier protein were employed.^[151] In the pursue of *de novo* assembly of specialised metabolites, bioinformatic tools such as ClusterCAD for PKS are employed and will in the future allow computer-aided re-engineering of NRPS and PKS assembly lines.^[143] In conclusion, the knowledge from these studies can help elucidate the biosynthesis of novel natural product and in turn every discovery of a novel mechanism increases the possibilities of the engineering toolbox.

1.7 Preliminary work on the BGCs involved in this work

The xenofuranone, xildivaline and glyoxpeptide BGCs were previously identified in the Bode lab.^[152–155] The following sections briefly summarise previous knowledge for these BGCs.

1.7.1 Xenofuranone biosynthesis encoded by the *xf*s BGC

The phenylpyruvate dimers xenofuranones A and B were identified from *X. szentirmaii* in our lab and xenofuranone A showed weak cytotoxic activity (IC₅₀ 1 µg/mL) against eukaryotic cells, L929.^[152] The proposed biosynthetic route suggests that condensation of two phenylpyruvic acid moieties leads to the formation of carboxylated furanone which is subsequently decarboxylated to yield xenofuranone B.^[152] S-Adenosylmethionine (SAM)-dependent O-methylation of xenofuranone B leads to xenofuranone A.^[152] In my master thesis, I was able to show that simultaneous deletion of two adjacent genes, *xf*sA and *xf*sB, coding for an NRPS-like enzyme and a predicted gluconolactonase results in a strain deficient in xenofuranone production. This work aims at further functional characterisation of the in *xf*sAB encoded enzymes and additional enzymes that may be involved in the xenofuranone biosynthesis.

INTRODUCTION

X. szentirmaii DSM16338

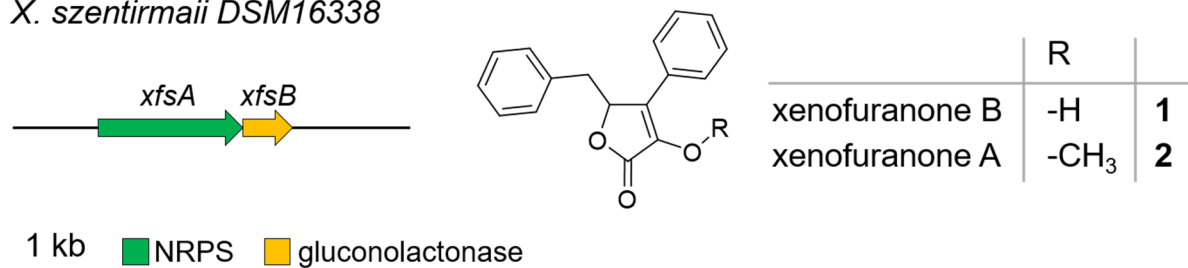


Figure 21. *xfs* BGC for xenofuranone biosynthesis. *xfs* BGC is involved in the biosynthesis of xenofuranones A and B in *X. szentirmaii*. Modified from [155].

1.7.2 Glyoxpeptide biosynthesis encoded by the *xgs* BGC

In previous work, the *xgs* gene cluster was activated via promoter exchange in *X. hominickii* Δhfq based on the easyPACId approach. This led to a product with m/z 292.2 $[M+H]^+$ for which a structure and biosynthesis were hypothesised.^[154] In following work, compound purification allowed structure elucidation by NMR and the elucidated structure led to a revised biosynthesis proposal (Fig.22).^[153] The compound is referred to as glyoxpeptide in this work.

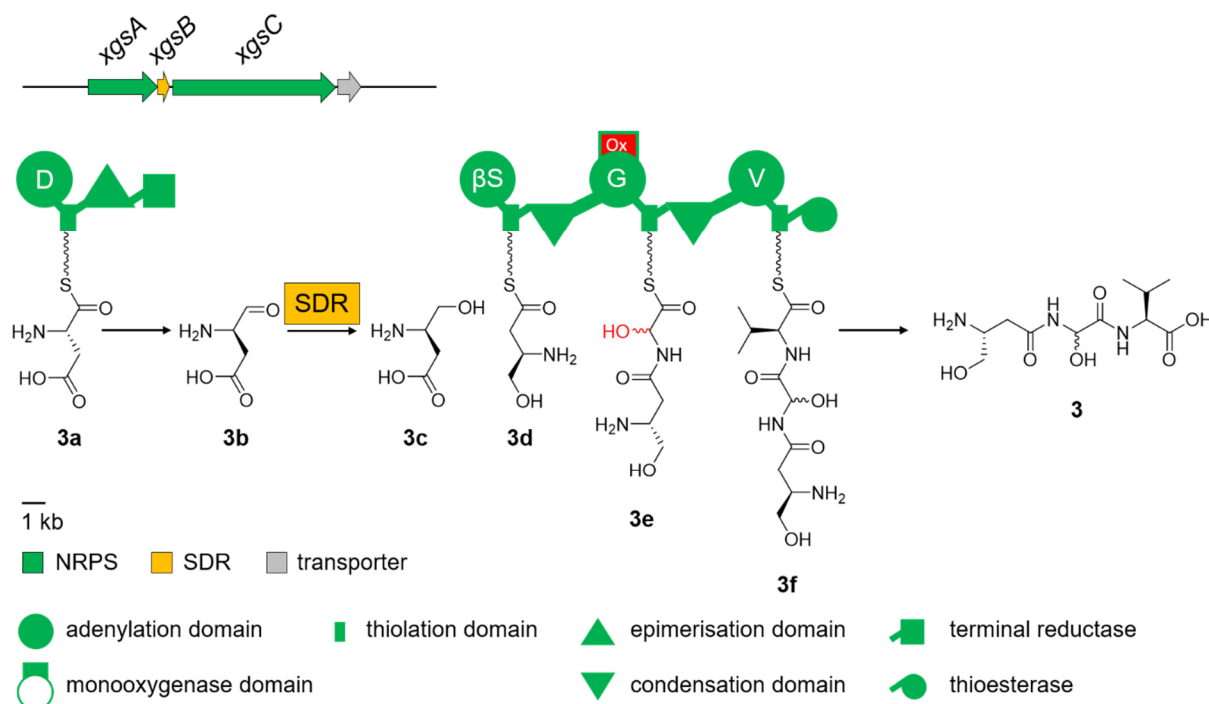


Figure 22. The *xgs* BGC and the proposed biosynthesis for glyoxpeptide in *X. hominickii*. The NRPS XgsA accepts L-aspartate (**3a**), epimerises the adenylate and releases 3-amino-4-oxobutanoic acid (**3b**). The SDR encoded by *xgsB* reduces **3b** to β -D-homoserine (**3c**). **3c** is accepted as first building block by the second NRPS XgsC and extended with α -hydroxyglycine and L-valine. XgsC releases the linear tripeptide glyoxpeptide (**3**). Modified from Janine Chekaiban's master thesis.^[153]

The first NRPS XgsA activates L-aspartate followed by epimerisation and reductive release to afford an aldehyde (**3b**). Next, the SDR reduces the aldehyde to an alcohol, yielding β -D-

homoserine (**3c**). This building block is then activated by the first A-domain of XgsC and loaded onto the T₁ domain (**3d**). This T₁ domain bound intermediate is condensed with T₂ domain bound glycine and subsequently the monooxygenase embedded in the A₂ domain hydroxylates the α position of the T₂ bound dipeptidyl intermediate (**3e**). **3e** is then condensed with an L-valine followed by hydrolytic release to yield the linear tripeptide glyoxpeptide (**3**).^[153,154] Nadine Keller and Janine Chekaiban both observed an additional mass signal (m/z 366.2 [M+H]⁺, C₁₁H₂₁N₃O₆) in the LC-MS analysis. Additionally, the deletion of the transporter encoding gene *xgsD* followed by heterologous expression of the modified BGC in *E. coli* did not affect the growth behaviour.^[153] This work aimed at a more detailed understanding of the biosynthesis and the observed compound derivative.

1.7.3 Xildivaline biosynthesis encoded by the *xis* BGC

In previous work, the easyPACId approach was used to specifically activate the *xis* gene cluster in *X. hominickii* Δhfq .^[153] The first gene of the cluster encodes a monomodular NRPS composed of an isoleucine specific A domain and adjacent T domain. The second gene encodes a PKS comprising KS-AT-KR-T1-T2-TD and *xisC* encodes an NRPS with A-T-C-A-T domains. The promoter activation led to the identification of two compound derivatives whose sum formulas were confirmed by HR-MS data and reverse feeding experiments (m/z 372.3 [M+H]⁺, C₁₉H₃₇N₃O₄. and m/z 358.3 [M+H]⁺, C₁₉H₃₅N₃O₃). Figure 23 depicts the proposed biosynthesis for compound JC372. The second derivative might be a Michael acceptor resulting from the spontaneous elimination of water. The culture supernatant of the induced promoter exchange mutant showed growth inhibition against *Micrococcus luteus* indicating a possible bioactivity against this strain. This work refers to these compounds as xildivalines and aims at a further characterisation of their biosynthesis.

INTRODUCTION

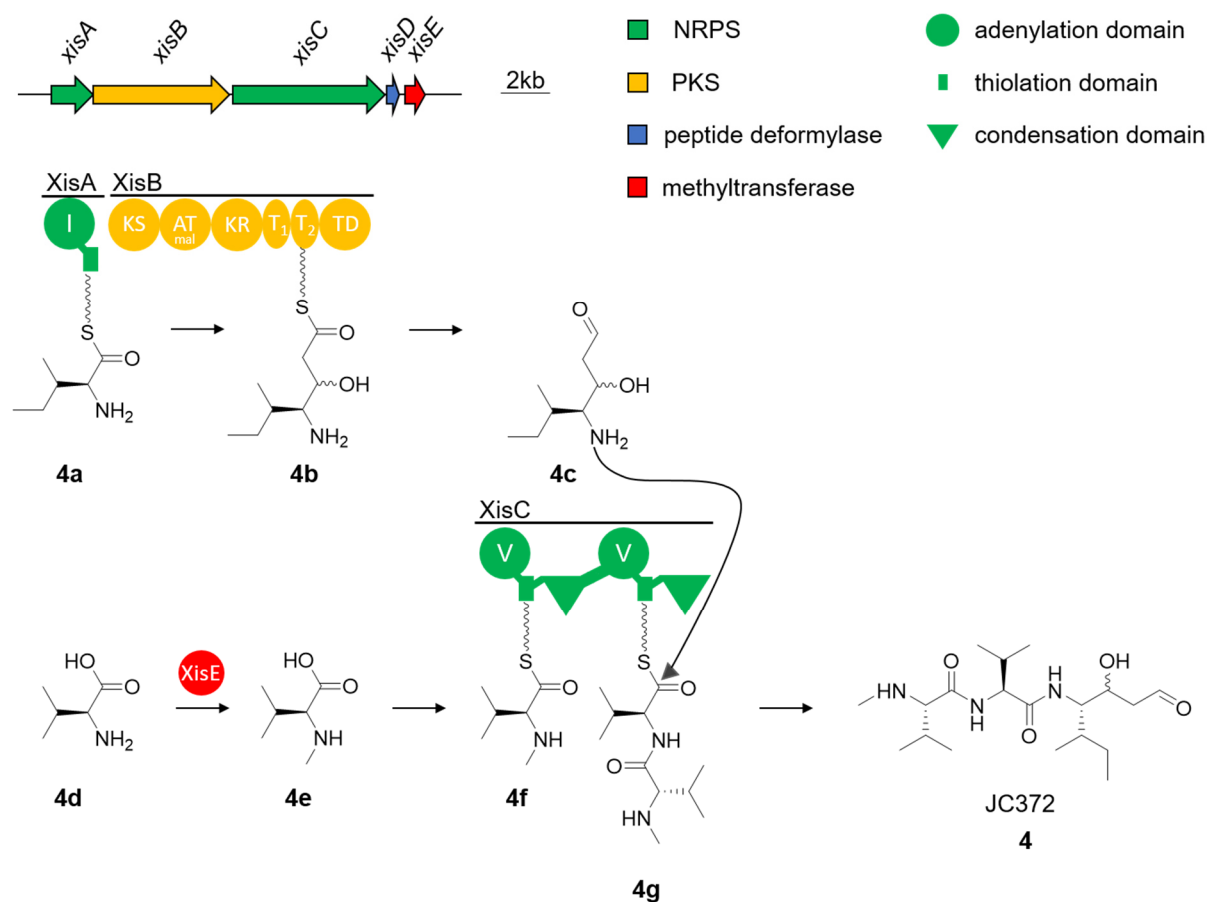


Figure 23. Proposed JC372 (4) biosynthesis in *X. hominickii*. The NRPS encoded by *xisA* activates isoleucine which is elongated by the PKS encoded by *xisB*. **4b** is reductively released to yield the NRP-PK hybrid **4c**. The methyltransferase encoded by *xisE* N-methylates valine which is then activated by the NRPS encoded in *xisC*. The NRP-PK hybrid attacks the dipeptidyl-val-val-S-T-Domain bound intermediate **4g** to yield **4**. Modified from Janine Chekaiban.^[153]

1.8 Aim of this work

This work aims at the functional characterisation of the enzymes involved in the biosynthesis of four different secondary metabolites from *Xenorhabdus*. The four compounds comprise xenofuranones produced by *X. szentirmaii* and glyoxpeptide, xildivaline and pyrrolizwilline investigated in *X. hominickii*, respectively.

Xenofuranones were identified from *X. szentirmaii* in previous work and this work aimed at a more detailed understanding of their biosynthesis. This included the deletion of the gluconolactonase encoding gene present in the *xf*s BGC as well as identification of the previously unknown methyltransferase.

The second BGC addressed in this work had not been described before as to my knowledge. Therefore, this work aimed at activation of this BGC and subsequent product purification and identification. The novel molecule was named pyrrolizwilline. Following structure elucidation, a

closer characterisation of the biosynthesis was pursued based on gene deletions, heterologous expression and an engineering strategy.

The following chapter in this work focuses on the *xgs* BGC responsible for glyoxpeptide production. This work aimed at the characterisation of the monooxygenase domain embedded in the XgsC_A2 domain and the identification of a previously detected glyoxpeptide derivative. With respect to the fourth BGC, single gene deletions and point mutations were conducted in order to investigate the xildivaline biosynthesis. This included the genes encoding for a methyltransferase and a peptide deformylase respectively, as well as two adjacent T domains exhibited by the PKS XisB.

2 Material and Methods

General methods which include cloning associated steps that apply to all four projects are described in chapter 2.1. Material and methods for specific experiments conducted for the respective projects are presented in sections 2.2-2.5.

2.1 General Methods

2.1.1 Genomic DNA

Genomic DNA was purified using the Genra Puregene Yeast/Bact. Kit (Qiagen) according to the manufacturer's instructions.

2.1.2 Plasmid DNA

For plasmid purification, either the alkaline lysis protocol modified from Birnboim and Doly 1979 or as an alternative the Invisorb® Plasmid Spin Mini *Two* kit (Stratec Molecular GmbH) was used.^[156] For the alkaline lysis protocol, 2 mL overnight culture were harvested at full speed for 30 seconds. The cell pellet was resuspended in 250 µL Buffer I (50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, add 40 µg/mL RNase A prior to use) by vortexing. Then, 250 µL Buffer II (200 mM NaOH, 1 % SDS) were added and the samples were inverted ten times. After adding 250 µL Buffer III (3 M KOAc, pH 5.5) the samples were inverted again ten times and centrifuged at full speed for 5 minutes. The supernatant was transferred to fresh tubes containing 650 µL isopropanol and the samples were inverted twenty times. The samples were centrifuged at full speed for 5 min, the supernatant was discarded and the pellet was washed with 500 µL 70 % ethanol at full speed for 2 min. The supernatant was discarded and the DNA pellets were air-dried prior to resuspension in 40 µL H₂O. Plasmid digest with restriction endonucleases was conducted according to manufacturer's instructions.

2.1.3 Polymerase Chain Reaction (PCR)

For amplification of DNA, polymerase chain reaction (PCR) was used.^[157] Polymerases S7 Fusion polymerase™ (MobiDiag), Q5® High Fidelity DNA polymerase (New England Biolabs), Phusion™ High Fidelity DNA polymerase (Thermo Fisher Scientific) were used according to

MATERIAL AND METHODS

manufacturer's instructions. PCR was conducted with Lab Cycler Gradient (Sensoquest GmbH) or peqSTAR 96X Universal (VWR Peqlab) thermocyclers.

Taq Colony PCR was used for verification of assembled plasmids in *E. coli* or plasmid insertions in *Xenorhabdus*. Self-made Taq DNA polymerase^[158] and 10xThermoPol™ Buffer (20 mM Tris-HCl, 10 mM (NH₄)₂SO₄, 10 mM KCl, 2 mM MgSO₄, 0.1 % Triton^R X-100, pH 8.8 @25 °C) were used. At first, the colony was resuspended in 50 µl 1xThermoPol™ Buffer and the suspension was incubated at 95 °C for 15 min in a thermocycler. 1 µL of the suspension was used for a 25 µl volume PCR with parameters as instructed for Taq DNA polymerase (Thermo Fisher Scientific). In case of PCR amplification of the full length plasmid backbone, a consecutive DpnI (Thermo Fisher Scientific) digest was conducted and the PCR product was purified from 1.5 % (w/v) agarose gel.

2.1.4 Purification from agarose gel and/or PCR

The MSB® Spin PCRapace (Stratec Molecular GmbH) was used to purify PCR products from the reaction mix. For purification from agarose gels, the Invisorb® Spin DNA Extraction (Stratec Molecular GmbH) kit was used according to manufacturer's instructions.

2.1.5 Hot Fusion cloning

In general for plasmid assembly, Hot Fusion cloning was used as described in Fu *et al.* 2014.^[159] Alternatively, NEBuilder® HiFi DNA Assembly mix (New England Biolabs) was used according to manufacturer's instructions.

2.1.6 *E. coli* transformation

E. coli cells were transformed via electroporation.^[160] For preparation of electrocompetent cells, a LB culture was inoculated 1:100 from an overnight culture and grown to an OD₅₉₅ = 0.6. The following steps were conducted on ice and with 4 °C cold 10% glycerol. The cells were harvested at 4000 rpm and 4 °C for 15 min and cells were washed with 4/5 culture volume 10 % glycerol at 4000 rpm and 4 °C for 15 min. This washing step was repeated with 1/25 culture volume and 1/50 culture volume 10 % glycerol. Finally, cells were resuspended in 1/500 culture volume and 50 µL aliquots were stored at -80 °C. Prior to electroporation, cell aliquots were thawed on ice and approx. 50-100 ng pDNA were added. For electroporation 1 m cuvettes and either GenePulser Xcell™ (Bio-Rad Laboratories GmbH) or GenePulser® II (Bio-Rad Laboratories GmbH) with 1250 V, 25 µF, 200 Ω was used. Following the electroshock, 1 mL LB medium was added and cells were incubated at 37 °C with shaking for 1 h. Next, cells were harvested at full speed for 5 min, resuspended in 200 µL LB, plated on LB agar plates with respective antibiotics and incubated overnight.

2.2 Xenofuranone project

2.2.1 Cultivation of strains

For liquid cultivation strains were grown in LB medium (10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl) with shaking at 37 °C for *E. coli* and at 30° C for *Xenorhabdus*. For cultivation on agar plates, 1.5 % agar was added to the LB medium. When appropriate, antibiotics kanamycin (50 µg/mL), chloramphenicol (34 µg/mL) and/or ampicillin (100 µg/mL) were added. For induction of promoter exchange strains, 0.2 % (w/v) L-arabinose was added to the culture at the beginning of cultivation.

2.2.2 Microorganisms

Table 1. Strains used for the work presented in section 3.1.

Strain	Genotype/Description	reference
<i>E. coli</i> DH10B	F- <i>mcrA</i> , Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80/ <i>lacZ</i> ΔM15,Δ <i>lacX74</i> , <i>recA1</i> , <i>endA1</i> , <i>araD139</i> , Δ(<i>araleu</i>)7697 <i>galU</i> , <i>galK</i> , <i>rpsL</i> , <i>nupG</i> , λ-	Invitrogen
<i>E. coli</i> S17-1 λpir	Tpr Smr <i>recA</i> , <i>thi</i> , <i>pro</i> , <i>hsdR</i> -M+RP4: 2-Tc:Mu:Km Tn7 λpir	[161]
<i>E. coli</i> ST18	<i>E. coli</i> S17-1 λpir Δ <i>hemA</i>	[162]
LP44	<i>E. coli</i> ST18 + pCEP_kan_xsze01901	Laura Penkert [163]
MW426	<i>E. coli</i> ST18 + pEB17_Km_Δxsze01902	This work
MW710	<i>E. coli</i> ST18 + pEB17_Km_Δxsze03490	This work
MW08	<i>X. szentirmaii</i> DSM 16338 WT	[164]
NN19	<i>X. szentirmaii</i> Δ <i>hfg</i>	[46]
LP53	<i>X. szentirmaii</i> Δ <i>hfg</i> + pCEP_kan_xsze01901	Laura Penkert [163]
MW19	<i>X. szentirmaii</i> Δxsze01902-1	Margaretha Westphalen [155]
MW88	<i>X. szentirmaii</i> + pCEP_kan_xsze01901	Margaretha Westphalen [155]
MW471	<i>X. szentirmaii</i> Δxsze01902	This work
MW503	<i>X. szentirmaii</i> Δxsze01902 + pCOLA_ara_xsze01902	This work
MW739	<i>X. szentirmaii</i> Δxsze03490	This work
MW742	<i>X. szentirmaii</i> Δxsze02284	This work
MW745	<i>X. szentirmaii</i> Δxsze01860	This work
MW748	<i>X. szentirmaii</i> Δxsze01783	This work
MW765	<i>X. szentirmaii</i> Δxsze034490 + pACYC_ara_xsze03490	This work

MATERIAL AND METHODS

2.2.3 Plasmids

Table 2. Plasmids used for the work presented in section 3.1

plasmid	description	reference
pCOLA_ara_tacl	ColA ori, <i>araC</i> -P _{BAD} , <i>tacl</i> , Km ^R	[165]
pACYC_ara_tacl	Modified from pACYC_tacl/I, contains ori p15A, <i>araC</i> -P _{BAD} , <i>tacl</i> , Cm ^R	Bode lab, [165]
pEB17	pDS132 based, R6K ori, oriT, Km ^R , <i>cipB</i> derivative with additional BglII site, <i>sacB</i>	[46]
pCk.CipB	pDS132 based, R6K ori, oriT, Km ^R , <i>cipB</i> derivative with additional BglII site, <i>sacB</i> Cm ^R	[166]
pCEP_kan	pDS132 based, R6K ori, oriT, Km ^R , <i>araC</i> , P _{BAD}	[31]
pCEP_kan_xfsA	<i>pCEP_kan</i> with first 645 bp of <i>xsze01901</i>	Laura Marie Pösche ^[167]
pCk.CipB_Δ <i>xsze01902-1</i>	R6K ori, oriT, <i>sacB</i> Cm ^R with 1013 bp upstream of <i>xsze01901</i> and 1035 bp downstream of <i>xsze01902</i>	This work
pEB17_Km_Δ <i>xsze01902</i>	R6K ori, oriT, Km ^R , <i>sacB</i> Kan ^R with 1114 bp upstream of <i>xsze01902</i> and 1026 bp downstream of <i>xsze01902</i>	This work
pCOLA_ara_x <i>xsze01902</i>	<i>pCOLA_ara_tacl</i> + <i>xsze01902</i> under P _{BAD} control	This work
pEB17_Km_Δ <i>xsze03490</i>	R6K ori, oriT, <i>sacB</i> , Kan ^R with 879 bp region (containing 861 bp upstream of <i>xsze03490</i> and 18 bp of <i>xsze03490</i>) and 996 bp downstream of <i>xsze03490</i>	This work
pEB17_Km_Δ <i>xsze02284</i>	R6K ori, oriT, <i>sacB</i> , Kan ^R with 956 bp region (containing 946 bp upstream of <i>xsze02284</i> and 10 bp of <i>xsze02284</i>) and 994 bp downstream region containing 979 bp downstream of <i>xsze02284</i> and 15 bp of <i>xsze02284</i>)	This work
pEB17_Km_Δ <i>xsze01860</i>	R6K ori, oriT, <i>sacB</i> , Kan ^R with 1011 bp region (containing 1005 bp upstream of <i>xsze01860</i>	This work

	and 6 bp of <i>xsze01860</i>) and 1011 bp downstream region containing 39 bp of <i>xsze03490</i> and 972 bp downstream of <i>xsze01860</i>	
pEB17_Km_Δ <i>xsze01783</i>	R6K ori, oriT, sacB, Kan ^R with 950 bp region (containing 935 bp upstream of <i>xsze01783</i> and 15 bp of <i>xsze01783</i>) and 1045 bp downstream region containing last 15 bp of <i>xsze01783</i> and 1030 bp downstream of <i>xsze01783</i>	This work
pACYC_ara_ <i>xsze03490</i>	<i>pACYC_ara_tacl</i> + <i>xsze03490</i> under P _{BAD} control	This work

2.2.4 Oligonucleotides

Oligonucleotides used for the work presented in section 3.1 are listed in Table 3.

Table 3. Oligonucleotides used for the work presented in section 3.1. Oligonucleotides are given in combination with their purpose, e.g. plasmid and insert amplification. The DNA template for the corresponding PCR is indicated. “Xsze” = *X. szentirmaii*.

Plasmid/purpose	Oligonucleotide Name	Sequence 5' -> 3' (overhang_template specific region)	PCR template
Verification for pCEP_Kan_ <i>xsze01901</i> insertion	V_pCEP_fw	GCTATGCCATAGCATTTCATCCATAAG	
	Mw_Xf_pexch_Ve_Rv	GTGAAGATCATCGGGCCATG	
pEB17_Δ <i>xsze01902</i>	MW170	CCTCTAGAGTCGACCTGCAG_GAAGCAGA AGAAATCATTGCC	Xsze gDNA
	MW171	CTCTATTTAGCTATCCGGTTTCAG_GTGTA TTTGAGCCTTCTATCC	Xsze gDNA
	MW172	GTTGGATAGAAGGCTCAAATACAC_CTGA AACCGGATAGCTAAATAG	Xsze gDNA
	MW173	TCCCGGGAGAGCTCAGATCT_GGAAGATG ATCCCATCAGG	Xsze gDNA
Verification for <i>X. szentirmaii</i> Δ <i>xsze01902</i>	MW170v	GAAGCAGAAGAAATCATTGCC	
	YS-Xf-D1_2-Ve-Rv-	CTGTCATGGAACATAGAGGTG	
pCOLA_ara_ <i>xsze01902</i>	MW60	CATGGAATTCCTCCTGTTAGC	pCOLA_ara_tacl
	MW61	TGCAGGAGCTGTTGACAATTAATC	
	MW174	GCTAACAGGAGGAATTCCATG_ATGAATG TTCAGATGGAAGCAAC	Xsze gDNA
	MW175	ATTGTCAACAGCTCCTGCAG_CTAACCTG CAAATGGTGTTG	
pEB17_Delta <i>xsze03490</i>	MW486	CCTCTAGAGTCGACCTGCAG_CGTCACAA ACAATATGGTGGG	Xsze gDNA

MATERIAL AND METHODS

	MW487	GAATATCATGCACAACGCAGT_GGCTATG TCTACTAACATTCATTC	
	MW488	GGCTATGTCTACTAACATTCATTC_ACTGC GTTGTGCATGATATTC	Xsze gDNA
	MW489	TCCCGGGAGAGCTCAGATCT_CCACCGTA AACAGCAACC	
Verification for <i>X. szentirmaii</i> Δ xsze03490	MW490	CCTGAATCAGGAGACTATCG	
	MW491	CGGTACAGGCAAGGAACTG	
pACYC_ara_ xsze03490	MW60	CATGGAATTCCTCCTGTTAGC	pACYC_a ra_tacl
	MW61	TGCAGGAGCTGTTGACAATTAATC	
	MW516	GCTAACAGGAGGAATTCC_ATGTTAGTAG ACATAGCCATAGC	Xsze gDNA
	MW517	ATTGTCAACAGCTCCTGCAG_CTATTTTTT CATTTCATAATATATTGAC	
pEB17_Delta xsze02284	MW492	CCTCTAGAGTCGACCTGCAG_CCATCGAA GTGTAAAGAAGG	Xsze gDNA
	MW493	GAACATTATTTGAGCTTTGTTTACT_GATG GTACATAAACTACCTCC	
	MW494	GGAGGTAGTTTATGTACCATC_AGTAAACA AAGCTGAAATAATGTTTC	Xsze gDNA
	MW495	TCCCGGGAGAGCTCAGATCT_CTGACTTG CAAGCTAAGGG	
Verification for <i>X. szentirmaii</i> Δ xsze02284	MW496	CAGAGCTAGGGAATATAAAGC	
	MW497	CCATCACTTACCCGTTTCAG	
pEB17_Delta xsze01860	MW498	CCTCTAGAGTCGACCTGCAG_CACGAAGG CGTAGCACG	Xsze gDNA
	MW499	CATCCGTACAAAACCCGG_TGACATTATTT CCCTCTTTTCATTC	
	MW500	GAATGAAAAGAGGGAAATAATGTCA_CCG GGTTTTGTACGGATG	Xsze gDNA
	MW501	TCCCGGGAGAGCTCAGATCT_CCAGCCCT GTCAGATCG	
Verification for <i>X. szentirmaii</i> Δ xsze01860	MW502	CGTTGACATTCTGGGATTATG	
	MW503	CCACAAAGAAACGCTCAGC	
pEB17_Delta xsze01783	MW504	CCTCTAGAGTCGACCTGCAG_GAGCTTCC CCAATGGATAG	Xsze gDNA
	MW505	CTTATTTCAATCATTTCATTTC_CATTT CCAATGACATCATTATTTCC	
	MW506	GGAATAATGATGTCATTGGAAATG_GAAA TGAGAAAATGATTGAAATAAG	Xsze gDNA
	MW507	TCCCGGGAGAGCTCAGATCT_CGACAATC ATACAAAGGCTC	
Verification for <i>X. szentirmaii</i> Δ xsze01783	MW508	CGGAGTCCCTATGAAATCAG	
	MW509	CGCCTTCGATTACTTTTAATGC	

2.2.5 Gene deletions via homologous recombination

For gene deletions, a protocol based on conjugation and homologous recombination was used.^[168] Approximately 1 kb long regions flanking the target gene were amplified via PCR and

thereby overhangs to the other flanking fragment and the target plasmid pEB17 were introduced. The length of flanking regions is given in Table 2 which lists the details of the plasmids used for deletion of *xsze01902* and *xsze03490*. The receiving plasmid pEB17 was digested with restriction enzymes BglII and PstI and the larger fragment was gel-purified. This pEB17 fragment was fused with the two flanking regions via Hot Fusion cloning.^[159] The confirmed plasmid was transformed into *E. coli* ST18 and this strain was then used as donor strain for conjugation with the recipient *Xenorhabdus* strain. 5 mL LB cultures were inoculated from overnight cultures in two different ratios: 1:50 and 1:100 for *E. coli* and 1:25 and 1:50 for *Xenorhabdus*, respectively. After growth to an OD₅₉₅ = 0.6, 2 mL of culture were harvested for each *E. coli* and *Xenorhabdus* by centrifugation. The harvested cells were washed twice with 1 mL fresh LB at 8000 rpm for 1 min and resuspended in 200 µL LB. Donor and recipient strain were mixed in a droplet on an LB agar plate and incubated overnight at 30 °C. The next day, the cell mass was scraped from the LB agar plate, resuspended in 2 mL LB and plated on two large LB agar plates containing ampicillin (100 µg/mL) and kanamycin (50 µg/mL). Following incubation at 30 °C for two days, overnight cultures with and without antibiotics were prepared for grown colonies. The next day, -80 °C long storage cultures were prepared from cultures grown in LB with antibiotics. From cultures grown in plain LB medium, serial dilutions were plated on LB agar plates containing ampicillin (100 µg/mL) and 6 % w/v sucrose. gDNA was purified (Gentra Puregene Yeast/Bact. Kit, QIAGEN GmbH) from grown colonies and the deletion was confirmed using PCR. Oligonucleotides used for verification bind upstream and downstream of the respective flanking regions used for homologous recombination. The oligonucleotides are listed in Table 3.

Cloning for plasmid-based complementation of *X. szentirmai* Δ*xsze03490*

pCOLA_*xsze03490* was assembled from two fragments via Hot Fusion cloning.^[159] The plasmid backbone was amplified from pCOLA_ara_tacl with primers (Table 3). The gene fragment was amplified with oligonucleotides. The plasmid was purified from *E. coli* DH10B and confirmed by restriction control digest prior to heat shock transformation into *X. szentirmai* Δ*xsze03490*.

2.2.6 Heat shock transformation of *Xenorhabdus*

X. szentirmai Δ*xsze03490* was transformed with pCOLA_*xsze03490* using a heat shock protocol, slightly modified than described previously.^[169] A fresh LB culture was inoculated 1:100 from an overnight culture and grown with shaking at 30 °C to an OD₅₉₅ = 0.6-0.8. From this culture, 1 mL was harvested for one transformation at 10000 rpm and 4 °C for 1 min. The cells were washed with 1 mL 4 °C cold XTB1 (100 µL 1 M Tris-HCl pH 6.5, 600 µL CaCl₂, 9.3 mL H₂O, filter sterilised) at 10000 rpm and 4 °C for 1 min. XTB1 was prepared fresh 30 min prior to cell harvest and buffer and cells were kept on ice for the following steps. Finally, cells

MATERIAL AND METHODS

were resuspended 200 μ L XTB1 and 7 μ L DMSO were added and the cells were kept on ice. After 5 min, another 7 μ L DMSO were added and after further 5 min incubation on ice, approx. 150 ng plasmid DNA were added. The cells were kept on ice for 30 min. For heat shock, cells were incubated at 37 °C for exactly 2 min. Next the cells were kept on ice for 2 min and then 800 μ L SOC medium (2 % tryptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM Glucose) were added for incubation with shaking at 30 °C for 1 h. Next, the entire culture volume was poured on an LB agar plate (50 μ g/mL kanamycin) and incubated for 2 days at 30 °C.

2.2.7 Feeding experiments with d₃-methionine or p-N₃-F

Deuterated (methyl-d₃)-methionine was added in 2 mM concentration at the beginning of cultivation. Cultures were grown with shaking at 30 °C in presence of 2 % XAD-16 (Amberlite®) resin. After 48 h cultivation, the resin was collected and extracted with one culture volume ACN while magnet stirring. The extract was filtered, centrifuged at full speed for 30 min and subjected to LC-HRMS analysis (section 2.2.9).

Para-azido-phenylalanine (p-N₃-F) was added in 2 mM final concentration at the beginning of cultivation. Cultures were grown with shaking at 30 °C for 72 h and samples for LC-MS analysis were prepared as described in section 2.2.9. L-tyrosine was added in 2 mM final concentration at the start of cultivation and the cultures were harvested after 24 h of cultivation with shaking at 30 °C. For reverse feeding of L-tyrosine, cultures were grown in Isogro® ¹³C medium.

2.2.8 Sample preparation for HPLC-MS analysis

For HPLC-MS sample preparation, 500 μ L of a 72 h culture were mixed 1:1 with ACN, diluted 1:10 and centrifuged at full speed for 30 min, if not stated otherwise.

For the extract comparison of cultures grown with or without XAD-16 (Amberlite®) resin, ethyl acetate was used for extraction. Cultures grown without XAD-16 resin, were extracted with ethyl acetate on a rolling shaker for 1 h. After phase separation, the organic solvent phase was transferred to a new flask and dried under reduced pressure. The dried extract was resuspended in ACN, centrifuged at full speed for 30 min and subjected to LC-MS analysis. The XAD-16 resin was collected from cultures grown in presence of XAD-16 and extracted with one culture volume of ACN for 1h whilst magnet stirring. Next, the extract was filtered through cellulose filters (3hw, 150 mm, 65 g/m², Ahlstrom Munktell) and the filtrate was dried under reduced pressure. The dried extract was resuspended in ACN, centrifuged at full speed for 30 min and subjected to LC-MS analysis.

2.2.9 HPLC-MS analysis

HPLC-MS analysis was conducted on a Dionex Ultimate 3000 (Thermo Fisher Scientific) RP-HPLC coupled to an AmaZonX (Bruker) ESI-IT mass spectrometer. 5 μ L sample were injected

and a 16 min gradient 5-95 % ACN/ 0.1% formic acid in H₂O/ 0.1 % formic acid (0-2 min 5 % ACN, 2-14 min 5-95 % ACN, 14-15 95 % ACN, 15-16 min 5 % ACN) with a flow rate of 0.4 mL/min was run on a C₁₈ Acquity UPLC BEH column (Waters). The mass spectrometer was set to positive ionisation mode with a mass range of m/z 100-1200 and cone voltage 4500 V.

HR data were generated using a Bruker Impact II™ ESI-Q-OTOF instrument in positive ionisation mode with a mass range m/z 100-1200, cone voltage 1500 V coupled to a Dionex Ultimate 3000 (Thermo Fisher Scientific) RP-HPLC. 5 µL sample were injected and a 16 min gradient 5-95 % ACN/ 0.1% formic acid in H₂O/ 0.1 % formic acid (0-2 min 5 % ACN, 2-14 min 5-95 % ACN, 14-15 95 % ACN, 15-16 min 5 % ACN) was used. The flow rate was 0.4 mL/min and a C₁₈ Acquity UPLC BEH column (Waters) was used. The data were analysed using the software Data Analysis 4.3 (Bruker). If not stated otherwise, the HPLC-UV-MS chromatograms are shown on the same scale in signal intensity when presented on the same x-axis.

2.2.10 Proteomics analysis

For sample preparation for proteome analysis, 5 mL LB culture were inoculated from an overnight culture at OD₅₉₅ = 0.1. Cells were grown with shaking at 30 °C until OD₅₉₅ = 4 and 750 µl were harvested at 10000 rpm for 1 min. The cell pellet was washed twice with 1xPBS at 10000 rpm for 1 min. The cell pellets were frozen in liquid nitrogen and stored at -80 °C. Samples were prepared in quadruplicates. Sample processing was conducted at the Max-Planck-Institute for Terrestrial Microbiology by Dr. Timo Glatter and the detailed method was previously described in [45].

2.2.11 Bioinformatic analysis

BlastP analysis was used to identify Xfs proteins in other *Xenorhabdus* strains and to compare them with those from *X. szentirmaii*.^[170]

2.3 Pyrrolizwilline project

2.3.1 Cultivation of strains

Strains were cultivated in LB medium (10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl) if not stated otherwise. Liquid cultivation was done with shaking. For cultivation on LB agar plates 1.5 % (w/v) agar was added to the medium. *E. coli* cells were grown at 37 °C and *Xenorhabdus* at 30 °C. Next to LB medium, M63 medium (2 g (NH₄)₂SO₄, 13.6 g KH₂PO₄, 0.5 mg FeSO₄ x 7 H₂O, 0.1 mL 0.5 % (w/v) thiamine, ad. 1L H₂O), Isogro® ¹³C medium and Gibco™ Sf-900™ SFM II medium were used. When appropriate, antibiotics kanamycin (50 µg/mL) and/or ampicillin (100 µg/mL) were added. For induction of promoter exchange strains, 0.2 % (w/v) L-arabinose was added to the culture at the start of cultivation. For the feeding experiment, d₈-L-valine (Sigma-Aldrich) was added in 2 mM final concentration.

MATERIAL AND METHODS

2.3.2 Microorganisms

The strains used for the work presented in section 3.2 are listed in Table 4.

Table 4. Strains used for the work presented in section 3.2.

Strain		Genotype/Description	reference
	<i>E. coli</i> DH10B	F- <i>mcrA</i> , Δ (<i>mrr-hsdRMS-mcrBC</i>) Φ 80 <i>lacZ</i> Δ M15, Δ <i>lacX74</i> , <i>recA1</i> , <i>endA1</i> , <i>araD139</i> , Δ (<i>araleu</i>)7697 <i>galU</i> , <i>galK</i> , <i>rpsL</i> , <i>nupG</i> , λ -	Invitrogen
	<i>E. coli</i> DH10B:: <i>mtaA</i>	<i>E. coli</i> DH10B with Δ <i>entD</i> :: <i>mtaA</i>	[171]
MW272	<i>E. coli</i> DH10b <i>mtaa</i> + pCOLA <i>ara_xhpA</i>	<i>E. coli</i> DH10B:: <i>mtaA</i> with pCOLA <i>ara_xhpA</i>	This work
MW277	<i>E. coli</i> DH10b <i>mtaa</i> + pCOLA <i>ara_xhpAB</i>	<i>E. coli</i> DH10B:: <i>mtaA</i> with pCOLA <i>ara_xhpAB</i>	This work
	<i>E. coli</i> S17-1 λpir	Tpr Smr <i>recA</i> , <i>thi</i> , <i>pro</i> , <i>hsdR</i> -M+RP4: 2-Tc:Mu:Km Tn7 λ pir	[161]
MW234	<i>E. coli</i> S17-1 λ pir + pCEP_Kan <i>pxaA</i>	<i>E. coli</i> S17-1 λ pir with pCEP_Kan <i>pxaA</i>	This work
MW532	<i>E. coli</i> S17-1 λ pir + pCEP_kan <i>pxaA_Xbov</i>	<i>E. coli</i> S17-1 λ pir with pCEP_kan <i>pxaA_Xbov</i>	This work
	<i>E. coli</i> ST18	<i>E. coli</i> S17-1 λ pir Δ <i>hemaA</i>	[162]
MW199	<i>E. coli</i> ST18 + pCEP_kan <i>xhpA</i>	<i>E. coli</i> ST18 with pCEP_kan <i>xhpA</i>	This work
MW672	<i>E. coli</i> ST18 + pCEP_kan_C _{starter} <i>xhpA_pxaA</i>	<i>E. coli</i> ST18 with pCEP_kan_C _{starter} <i>xhpA_pxaA</i>	This work
MW776	<i>E. coli</i> ST18 + pCEP_kan_C _{starter} <i>pxaA_xhpA</i>	<i>E. coli</i> ST18 with pCEP_kan_C _{starter} <i>pxaA_xhpA</i>	This work
MW484	<i>E. coli</i> ST18 + pEB17_Km_ Δ <i>xhpG</i>	<i>E. coli</i> ST18 with pEB17_Km_ Δ <i>xhpG</i>	This work
MW492	<i>E. coli</i> ST18 + pEB17_Km_ Δ <i>xhpF</i>	<i>E. coli</i> ST18 with pEB17_Km_ Δ <i>xhpF</i>	This work
MW558	<i>E. coli</i> ST18 + pEB17_Km_ Δ <i>pxaAB</i>	<i>E. coli</i> ST18 with pEB17_Km_ Δ <i>pxaAB</i>	This work
MW494	<i>E. coli</i> ST18 + pEB17_Km_ Δ <i>xhpE</i>	<i>E. coli</i> ST18 with pEB17_Km_ Δ <i>xhpE</i>	This work
MW450	<i>E. coli</i> ST18 + pEB17_Km_ Δ <i>ngrA</i>	<i>E. coli</i> ST18 with pEB17_Km_ Δ <i>ngrA</i>	This work
MW652	<i>E. coli</i> ST18 + pEB17_Km_ Δ <i>xhpB-G</i>	<i>E. coli</i> ST18 with pEB17_Km_ Δ <i>xhpB-G</i>	This work

	<i>X. hominickii</i> WT	<i>X. hominickii</i> DSM 17903 WT	[5]
	<i>X. hominickii</i> Δhfq	<i>hfq</i> is deleted	[172]
MW208	<i>X. hominickii</i> P _{BAD_xhpA}	pCEP_kan_xhpA is integrated, P _{BAD_xhpA}	This work
MW797	<i>X. hominickii</i> P _{BAD_} C _{starter} pxaA_xhpA	pCEP_kan_C _{starter} pxaA_xhpA is inserted in front of <i>xhpA</i>	This work
MW589	<i>X. hominickii</i> $\Delta pxaAB$ + P _{BAD_xhpA}	<i>pxaAB</i> is deleted; pCEP_kan_xhpA is integrated, P _{BAD_xhpA}	This work
MW260	<i>X. hominickii</i> Δhfq P _{BAD_xhpA}	<i>hfq</i> is deleted, pCEP_kan_xhpA is integrated, P _{BAD_xhpA}	This work
MW683	<i>X. hominickii</i> $\Delta xhpC-G$ P _{BAD_xhpA}	<i>xhpC-G</i> , are deleted, pCEP_kan_xhpA is integrated, P _{BAD_xhpA}	This work
MW520	<i>X. hominickii</i> $\Delta ngrA$ P _{BAD_xhpA}	<i>ngrA</i> is deleted, pCEP_kan_xhpA is integrated, P _{BAD_xhpA}	This work
MW518	<i>X. hominickii</i> $\Delta xhpE$ P _{BAD_xhpA}	<i>xhpE</i> is deleted, pCEP_kan_xhpA is integrated, P _{BAD_xhpA}	This work
MW515	<i>X. hominickii</i> $\Delta xhpF$ P _{BAD_xhpA}	<i>xhpF</i> is deleted, pCEP_kan_xhpA is integrated, P _{BAD_xhpA}	This work
MW500	<i>X. hominickii</i> $\Delta xhpG$ P _{BAD_xhpA}	<i>xhpG</i> is deleted, pCEP_kan_xhpA is integrated, P _{BAD_xhpA}	This work
MW829	<i>X. hominickii</i> $\Delta xhom00317$	<i>xhom00317</i> is deleted	This work
MW368	<i>X. hominickii</i> + pCOLA_ara_xhom00317	Contains plasmid pCOLA_ara_xhom00317	This work
	<i>X. szentirmaii</i> WT DSM 16338		[164]
MW91	<i>X. szentirmaii</i> P _{BAD_pxaA}	<i>X. szentirmaii</i> WT DSM 16338, pCEP_kan_pxaA_Xsze is integrated in front of <i>pxaA</i>	Margaretha Westphalen ^[155]

MATERIAL AND METHODS

MW677	<i>X. szentirmaii</i> $P_{BAD_C_{starter}xhpA_pxaA}$	pCEP_kan_C _{starter} <i>xhpA_pxaA</i> is inserted in front of <i>pxaA</i>	This work
	<i>X. bovienii</i> SS-2004		[173]
MW538	<i>X. bovienii</i> P_{BAD_pxaA}	<i>X. bovienii</i> SS-2004, pCEP_kan_ <i>pxaA_Xbov</i> is integrated in front of <i>pxaA</i>	This work

2.3.3 Plasmids

Plasmids used for the work presented in section 3.2 are listed in Table 5.

Table 5. Plasmids used for the work presented in section 3.2

plasmid	description	reference
pCOLA_ara_tacl	ColA ori, <i>araC</i> - P_{BAD} , <i>tacl</i> , Km ^R	[165]
pEB17	pDS132 based, R6K ori, oriT, Km ^R , <i>cipB</i> derivative with additional BglIII site, <i>sacB</i>	[46]
pCEP_kan	pDS132 based, R6K ori, oriT, Km ^R , <i>araC</i> , P_{BAD}	[31]
pCEP_kan_ <i>xhpA</i>	<i>pCEP_kan</i> with first 721 bp of <i>xhpA</i>	This work
pCEP_kan_ <i>pxaA_Xbov</i>	<i>pCEP_kan</i> with first 746 bp of <i>pxaA</i> (<i>X. bovienii</i>)	This work
pEB17_Δ <i>pxaAB</i>	R6K ori, oriT, <i>sacB</i> , Kan ^R with 1004 bp upstream of <i>xhpA</i> (including the first 18 bp of <i>xhpA</i>) and 1054 bp downstream of <i>xhpA</i>	This work
pCOLA_ara_ <i>xhpA</i>	pCOLA_ara_ <i>tacl</i> + <i>xhpA</i> inserted under P_{BAD} control	This work
pCOLA_ara_ <i>xhpAB</i>	pCOLA_ara_ <i>tacl</i> + <i>xhpAB</i> inserted under P_{BAD} control	This work
pCOLA_ara_ <i>xhom00317</i>	pCOLA_ara_ <i>tacl</i> + <i>xhom00317</i> inserted under P_{BAD} control	This work
pCEP_kan_C _{starter} <i>xhpA_pxaA</i>	pCEP_kan with first 1347 bp of <i>pxaA</i> (<i>X. szentirmaii</i>) fused to 2024 bp long region of <i>xhpA</i> (starting from 1348 th bp)	This work
pCEP_kan_C _{starter} <i>pxaA_xhpA</i>	pCEP_kan with first 1326 bp of <i>xhpA</i> (<i>X. szentirmaii</i>) fused to 2116 bp long region of <i>pxaA</i> (<i>X. szentirmaii</i> , starting from 1327 th bp)	This work

pEB17_Km_Δ <i>ngrA</i>	R6K ori, oriT, sacB, Kan ^R with 1044 bp upstream of <i>ngrA</i> fused to 1046 bp downstream of <i>ngrA</i>	This work
pEB17_Km_Δ <i>xhpE</i>	R6K ori, oriT, sacB, Kan ^R with 942 bp upstream of <i>xhpE</i> (including first 33 bp of <i>xhpE</i>) fused to 965 bp downstream of <i>xhpE</i> (including last 27 bp of <i>xhpE</i>)	This work
pEB17_Km_Δ <i>xhpF</i>	R6K ori, oriT, sacB, Kan ^R with 985 bp homologous region (958 bp thereof upstream of <i>xhpF</i>) and 1087 bp homologous region (970 bp thereof upstream of <i>xhpF</i>)	This work
pEB17_Km_Δ <i>xhpG</i>	R6K ori, oriT, sacB, Kan ^R with 981 bp region upstream of <i>xhpG</i> (including first 30 bp of <i>xhpG</i>) fused to 1120 bp downstream of <i>xhpG</i> (including last 27 bp of <i>xhpG</i>)	This work
pEB17_Km_Δ <i>xhpB-G</i>	R6K ori, oriT, sacB, Kan ^R with 1213 bp region upstream of <i>xhpB</i> (including first 60 bp of <i>xhpB</i>) fused to 1169 bp downstream of <i>xhpG</i> (including last 82 bp of <i>xhpG</i>)	This work
pEB17_Km_Δ <i>xhom00317</i>	R6K ori, oriT, sacB, Kan ^R 1017 bp region upstream of <i>xhom00317</i> and 985 bp region downstream of <i>xhom00317</i>	This work

2.3.4 Oligonucleotides

Oligonucleotides used for the work presented in section 3.2 are listed in Table 6.

Table 6. Oligonucleotides used for the work presented in section 3.2. Oligonucleotides are given in combination with their purpose, e.g. plasmid and insert amplification. Underscore marks border between template binding and introduced sequence. The DNA template for the corresponding PCR is indicated. “*Xhom*” = *X. hominickii*, “*Xbov*” = *X. bovienii*, “*Xsze*” = *X. szentirmaii*.

Plasmid/purpose	Oligonucleotide Name	Sequence 5' -> 3' (overhang_template specific region)	PCR template
pCEP_Kan_ <i>xhpA</i>	pCEP_b b_fw	ATGTGCATGCTCGAGCTC	pCEP_Kan
	pCEP_rv	ATGCTAGCCTCCTGTTAGC	

MATERIAL AND METHODS

	MW12	TTTGGGCTAACAGGAGGCTAGCAT_ATGAAT ATTGCTACTGAACTTGCATTTTC	<i>Xhom</i> gDNA
	MW8	TCTGCAGAGCTCGAGCATGCA_ACCCTCTAT CTATTCCGCGC	
Verification for pCEP_Kan_xhp A insertion	V_pCEP fw	GCTATGCCATAGCATTITTTATCCATAAG	
	MW9	GAGAAGTTGTTGCTCGCGAG	
pEB17_Km_Δp xaAB	MW290	CCTCTAGAGTCGACCTGCAG_GATACTGAT TCATCAGCGGG	<i>Xhom</i> gDNA
	MW291	GAGCTTCAGATATTTGCTGAACCATA_GTAC AATTTGTGATTTTTTCATAG	
	MW292	CTATGAAAAAATCACAAATTGTAC_TATGGTT CAGCAAATATCTGAAGC	<i>Xhom</i> gDNA
	MW293	TCCCGGGAGAGCTCAGATCT_GGCGTATGC ATGTAGTACC	
Verification for <i>X. hominickii</i> ΔpxaAB	MW300	GGCTAGTAACAAAGCAACCG	
	MW301	GGAATTTACGCTCTGCAG	
pCEP_kan_pxa A (<i>X. hominickii</i>)	pCEP_b b fw	ATGTGCATGCTCGAGCTC	pCEP_K an
	pCEP_rv	ATGCTAGCCTCCTGTTAGC	
	MW11	TTTGGGCTAACAGGAGGCTAGCAT_ATGAAA AAATCACAAATTGTACCTCTTAC	<i>Xhom</i> gDNA
	MW5	TCTGCAGAGCTCGAGCATGCA_ATACGACG GTTCAGCAGAGG	
Verification for pCEP_Kan_pxa A insertion	V_pCEP fw	GCTATGCCATAGCATTITTTATCCATAAG	
	MW6	GTGACATCGAACCATTGACG	
pCEP_kan_pxa A (<i>X. bovienii</i>)	pCEP_b b fw	ATGTGCATGCTCGAGCTC	pCEP_K an
	pCEP_rv	ATGCTAGCCTCCTGTTAGC	
	MW305	TTTGGGCTAACAGGAGGCTAGCAT_ATGAGT ATTGCTACCGATATTGG	<i>Xbov</i> gDNA
	MW306	TCTGCAGAGCTCGAGCATGCA_GCATTTGAC AGCAGAACTGC	
Verification for pCEP_Kan_pxa A insertion in <i>X.</i> <i>bovienii</i>	V_pCEP fw	GCTATGCCATAGCATTITTTATCCATAAG	
	MW307	CTGACGTAATTTCTGCATCG	
Verification for pCEP_Kan_pxa A insertion	V_pCEP fw	GCTATGCCATAGCATTITTTATCCATAAG	
	MW6	GTGACATCGAACCATTGACG	
pCOLA_ara_xh pA	MW60	CATGGAATTCCTCCTGTTAGC	pCOLA_ ara_tacl
	MW61	TGCAGGAGCTGTTGACAATTAATC	
	MW57	GGCTAACAGGAGGAATTCC_ATGAATATTGC TACTGAACTTGCATTTTC	<i>Xhom</i> gDNA
	MW58	ATTGTCAACAGCTCCTGCAG_TTACTCATTTT CCAGTACCTCAC	
pCOLA_ara_xh pAB	MW60	CATGGAATTCCTCCTGTTAGC	pCOLA_ ara_tacl
	MW61	TGCAGGAGCTGTTGACAATTAATC	
	MW57	GGCTAACAGGAGGAATTCC_ATGAATATTGC TACTGAACTTGCATTTTC	<i>Xhom</i> gDNA
	MW59	ATTGTCAACAGCTCCTGCAG_TTAATAAATAT TTTTCGTATTATAATTATCCTGAAC	

pCOLA_ara_xh om00317	MW60	CATGGAATTCCTCCTGTTAGC	pCOLA_ara_tacl XhoI gDNA
	MW61	TGCAGGAGCTGTTGACAATTAATC	
	MW73	GCTAACAGGAGGAATTCC_ATGATAGATAGA AACAAAGAAGAATTTGGTCC	
	MW74	ATTGTCAACAGCTCCTGCAG_TTATTTAGTA CTACCAGAGTTTGATGAGTTGC	
pEB17_Km_Δx hpG	MW234	CCTCTAGAGTCGACCTGCAG_GGTACTACA GTACTTTCTCTGG	XhoI gDNA
	MW254	GCACATCCTCCGACAAAAGAAA_CTGAGCAT CGCCAACCTTAAAC	
	MW255	GTAAAGTTGGCGATGCTCAG_TTTCTTTTG TCGGAGGATGTGCC	XhoI gDNA
	MW237	TCCCGGGAGAGCTCAGATCT_CATGTGCGC AGACTCAACG	
Verification of <i>X. hominickii</i> ΔxhpG	MW238	GAACGAATTGATCCAGATACAGG	
	MW239	GATGGTCGCTTATCTTCCCG	
pEB17_Km_Δx hpE	MW226	CCTCTAGAGTCGACCTGCAG_GTATATTGGT AACTGGTGCATCG	XhoI gDNA
	MW256	CATCTCCAAAAATATTAATGTC_TTTAAATGA CCATATTTGGAATGTTTC	
	MW257	CATTCCAAATATGGTCATTTAAA_GACATTAA TATTTTTGGAGATGATTG	XhoI gDNA
	MW229	TCCCGGGAGAGCTCAGATCT_GAAATGGTG ACCATGCAGCC	
Verification of <i>X. hominickii</i> ΔxhpE	MW242	CCAACGGCAGAAGAATTCGAG	
	MW243	GCTATAGACTGTGGAGCATTAGC	
pEB17_Km_ΔngrA	MW208	CCTCTAGAGTCGACCTGCAG_GCTGACCAC ATGCGTACTG	XhoI gDNA
	MW209	TTTTTTCCAATAAGGTAAGCACTC_CATCGG CTTTGTCATCGG	
	MW210	TTAAATCCGATGACAAAGCCGATG_GAGTGC TTACCTTATTGAAAAAAC	XhoI gDNA
	MW211	TCCCGGGAGAGCTCAGATCT_CGATGCGGG CGTGAAACTC	
Verification of <i>X. hominickii</i> ΔngrA	MW212	CTACGATAGTACGACTCGCC	
	MW213	GACTTATGCCGTACCTGCTACC	
pEB17_Δxhom0 0317	MW612	CCTCTAGAGTCGACCTGCAG_ACCCTCTATC TATTCCGCGC	XhoI gDNA
	MW613	CTGCGTTGTTAGATATTTTGGAAATT_CTACTA CCTCGTTATAAACCAAAATC	
	MW614	GATTTTGGTTTATAACGAGGTAGTAG_AATT CCAAAATATCTAACAACGCAG	XhoI gDNA
	MW615	TCCCGGGAGAGCTCAGATCT_GCATTGCAC CAGTGTCTGC	
Verification for <i>X. hominickii</i> Δxhom00317	MW616	CATTATCTTATCCCACTCAGCG	
	MW9	GAGAAGTTGTTGCTCGCGAG	
pEB17_Km_Δx hpF	MW230	CCTCTAGAGTCGACCTGCAG_GACTGAAAT AGATGAGTGGTTCG	XhoI gDNA

MATERIAL AND METHODS

	MW252	GGTAAAGTTCCAAAACAGCCG_AGAAAGTG GGCGATGTATATTAG	
	MW253	CCTAATATACATCGCCCACTTTCT_CGGCTG TTTTGGAACCTTACC	<i>Xhom</i> gDNA
	MW233	TCCCGGGAGAGCTCAGATCT_CTCCGCTCT GTTATTTTCAACG	
Verification for <i>X. hominickii</i> $\Delta xhpF$	MW240	GACTCATTATTATGTTGGAGATGATG	
	MW241	CCAAAATGGGGACATGATTATTC	
pEB17_Km_ Δx <i>hpF</i>	MW398	CCTCTAGAGTCGACCTGCAG_GCAAGGCTC AGTCAGCTG	<i>Xhom</i> gDNA
	MW399	GATTCTCTATATGTGCAAGATGCCCC_TGCA GCGGATAAGCCACC	
	MW400	GTTGGTGGCTTATCCGCTGCA_GGGGCATC TTGCACATATAG	<i>Xhom</i> gDNA
	MW401	TCCCGGGAGAGCTCAGATCT_GATATCGTTA ACCGAGCGCC	
Verification for <i>X. hominickii</i> $\Delta xhpF$	MW441	GGTCCATCCTATCTTGTGG	
	MW239	GATGGTCGCTTATCTTCCCG	
pCEP_kan_C _{star} ter $pxaA_xhpA$	pCEP_b b_fw	ATGTGCATGCTCGAGCTC	pCEP_K an
	pCEP_rv	ATGCTAGCCTCCTGTTAGC	
	MW422	TTTGGGCTAACAGGAGGCTAGCAT_ATGTAC ACATTACAGGCATTACC	<i>Xsze</i> gDNA
	MW432	GTAAGGTTTTATTGGACCATTCTC_AAACAG CGCCAATTGGGCC	
	MW459	GGCCCAATTGGCGCTGTTT_GAGAATGGTC CAATAAACCTTAC	<i>Xhom</i> gDNA
	MW425	TCTGCAGAGCTCGAGCATGCA_CGTGGGTC ACGTTGAGC	
Verification for pCEP_kan_C _{star} ter $pxaA_xhpA$ insertion	V_pCEP _fw	GCTATGCCATAGCATTATCCATAAG	
	MW9	GAGAAGTTGTTGCTCGCGAG	
pCEP_kan_C _{star} ter $xhpA_pxaA$	pCEP_b b_fw	ATGTGCATGCTCGAGCTC	pCEP_K an
	pCEP_rv	ATGCTAGCCTCCTGTTAGC	
	MW427	TTTGGGCTAACAGGAGGCTAGCAT_ATGAAT ATTGCTACTGAACTTGCATTTTC	<i>Xhom</i> gDNA
	MW433	CAGTGACTGGCCCCTGTTT_ATACTTTTTCA GGAAACTAATCTGTTTC	
	MW429	GTAATGCGCGGAATAGATAGAGGGT_GAAC AGGGGCCAGTCAC	<i>Xsze</i> gDNA
	MW430	TCTGCAGAGCTCGAGCATGCA_GCATTGAG GGAACAAAGTGC	
Verification for pCEP_kan_C _{star} ter $xhpA_pxaA$ insertion	V_pCEP _fw	GCTATGCCATAGCATTATCCATAAG	
	MW431	CGAAATATGAGACATCTACGG	

2.3.5 Promoter exchange

For the generation of the promoter exchange strains which carry the P_{BAD} in front of the target gene, a protocol well established in *Xenorhabdus* was used.^[31] The method is based on conjugative transfer of a designed promoter exchange plasmid from *E. coli* to *Xenorhabdus*. First, the respective promoter exchange plasmid was assembled from two fragments. The first approximately 750 bp of the target gene were amplified by PCR and assembled via Hot Fusion cloning with the pCEP_kan backbone which was generated from plasmid PCR. The used oligonucleotides are listed in Table 6. The assembled promoter exchange plasmid was verified and transformed into either *E. coli* S17-1 λ pir or *E. coli* ST-18. For conjugation with *Xenorhabdus*, overnight cultures were prepared for the donor *E. coli* and the recipient *Xenorhabdus* strain. The next day, fresh LB cultures were inoculated from overnight cultures in two different ratios: 1:50 and 1:100 for *E. coli* cultures and 1:25 and 1:50 from *Xenorhabdus* overnight cultures. The cultures were grown with shaking to an $OD_{595} = 0.6-0.8$ and harvested at 10000 rpm for 1 min. The cell pellets were washed with 1 mL LB, centrifuged at 10000 rpm for 1 min and resuspended in 200 μ L LB. 25 μ L *E. coli* cells were mixed with 75 μ L *Xenorhabdus* cells via pipetting in a droplet on an LB agar plate. The plate was incubated overnight at 30 °C and the next day the cells were scratched from the agar plate and resuspended in 2 mL LB medium. Serial dilutions of this suspension were plated on LB agar plates (50 μ g/mL kanamycin, 100 μ g/mL ampicillin) and incubated at 30 °C for two days. For obtained colonies, the correct insertion of the promoter exchange plasmid was verified by PCR from purified gDNA or by Colony PCR (section 2.1.3). One of the oligonucleotides used for verification PCR binds upstream of the P_{BAD} motif in the promoter exchange plasmid sequence and the corresponding reverse binding oligonucleotide binds downstream of the homologous region used for recombination.^[31]

2.3.6 $C_{starter}$ domain exchange and promoter exchange

The exchange of the promoter in front of *xhpA* and the exchange of the $C_{starter}$ domain sequence in *xhpA* was conducted by following the promoter exchange protocol (section 2.3.5). To perform both exchanges in one conjugation cycle, the conjugation plasmid carries the P_{BAD} in front of the heterologous $C_{starter}$ domain encoding sequence which is fused to the approx. 750 bp sequence in the recipient strain used for homologous recombination. The fusion site between $C_{starter}$ and A domain was determined in the C/A linker region based on the fusion point established in the XU concept.^[174] The plasmid was assembled from three fragments via Hot Fusion cloning: 1) PCR amplified pCEP_kan 2) PCR amplified $C_{starter}$ domain sequence c) PCR amplified approx. 750 bp long homologous region downstream of the fusion site in the *Xenorhabdus* recipient.^[159] The oligonucleotides are listed in Table 6. The conjugation and plasmid insertion verification steps were conducted as in section 2.3.5.

MATERIAL AND METHODS

2.3.7 Gene deletions

Gene deletions were conducted as described in section 2.2.5 and the respective oligonucleotides for deletion plasmid assembly and verification of deletion are listed in Table 6. In most cases, the gene deletion was followed by promoter exchange in front of *xhpA*. The method is explained in section 2.3.5.

2.3.8 Heterologous expression in *E. coli* DH10B::*mtaa*

For heterologous expression in *E. coli*, competent *E. coli* DH10B::*mtaa* cells were prepared and transformed via electroporation as explained in section 2.1.6. The transformed plasmids were assembled via Hot Fusion cloning and the respective oligonucleotides with their template DNA are listed in Table 6.^[159]

2.3.9 Heat shock transformation of *Xenorhabdus*

For transformation of plasmids into *X. hominickii*, the heat shock protocol was applied as described in section 2.2.6.

2.3.10 Sample preparation for LC-MS analysis

If not stated otherwise, production cultures were based on LB medium and incubated with shaking at 30 °C for 72 h. In general, 2 % (w/v) Amberlite® XAD-16 resin was added. For induction, 0.2 % L-arabinose was added at the beginning of cultivation. After 72 h, the resin was collected and extracted with one culture volume MeOH for 1 h with magnet stirring. The extract was filtered through cellulose filters (3hw, 150 mm, 65 g/m², Ahlstrom Munktell) and the filtrate was dried under reduced pressure. The dried extract was solved in 1/10 culture volume MeOH, diluted 1:10 with MeOH and subjected to LC-MS analysis.

2.3.11 LC-MS analysis

HPLC-MS and HPLC-HR-MS analysis were conducted as described in section 2.2.9. The data were analysed using the software Data Analysis 4.3 (Bruker). For relative quantification of pyrrolizwilline to compare production levels between the two respective strains, biological triplicates were prepared. The software Compass Target Analysis 1.3 (Bruker Daltonik GmbH) was used to determine the peak area for the pyrrolizwilline signal (C₂₁H₂₄N₄O₃H⁺+1, 381.1914, t_R = 7.5 min). The peak area values were normalised to the OD₅₉₅ of the respective culture measured at the end of cultivation. For statistical significance, two-sided unpaired t-test was conducted using GraphPad Prism version 8.1.2 for Windows, GraphPad Software (San Diego, California USA, www.graphpad.com).

2.3.12 Pyrrolizwilline purification

For pyrrolizwilline purification, *X. hominickii* P_{BAD}_*xhpA* was cultivated at 30 °C with shaking for 72 h in LB (50 µg/mL kanamycin, 0.2 % L-arabinose) with 2 % Amberlite® XAD-16 resin. The resin was harvested from 4x 1 L culture volume by filtration and extracted in three rounds with methanol and magnet stirring. The total of approx. 600 mL MeOH extracts was dried

under reduced pressure and the dried crude extract yield was 3.3 g. For purification, the crude extract was resuspended in 2 mL solvent (70% DMSO, 20% MeOH, 10% Isopropanol). Prefractionation was conducted with a 1260 Infinity II preparative HPLC (Agilent) instrument with ZORBAX Eclipse XDB C₁₈ column (80Å, 5 µm, 9.4 x 250 mm; Agilent) and a 20 min gradient 35-55 % ACN/ 0.1% formic acid in H₂O/ 0.1 % formic acid (0-15 min 35-55 % ACN, 15-20 min 35 % ACN). The prefractionation yield was 13.8 mg. For further purification, 1260 Infinity II semi-preparative LCMS instrument (Agilent) with ZORBAX Eclipse Plus C₁₈ column (95Å, 5 µm, 4.6 x 250 mm; Agilent) and a 20 min gradient 50-55 % ACN/ 0.1% formic acid in H₂O/ 0.1 % formic acid (0-15 min 50-55 % ACN, 15-20 min 50 % ACN) was used. The yield was 2.3 mg and the sample was solved in d₆-DMSO for ¹H (500 MHz, Bruker) and ¹³C (125 MHz, Bruker) NMR analysis. The NMR analyses and structure elucidation were conducted by Yi-Ming Shi. The software MestReNova (Mestrelab Research) was used for data analysis.

2.3.13 Agar diffusion assay

For bioactivity testing, a growth inhibition agar diffusion assay was conducted. Extracts were prepared as described in section 2.3.10 from *X. hominickii* WT as well as non-induced and arabinose induced *X. hominickii* P_{BAD_xhpA} strains. 50 µl extracts were pipetted onto paper disks and the cotton disks were dried under the hood. Kanamycin (0.05 µg) was used as positive control and MeOH (50 µl) served as negative control. For the investigated strains *Bacillus subtilis*, *Micrococcus luteus*, *Saccharomyces cerevisiae* and *E. coli* DH5α, overnight cultures were prepared and the next day, the grown cultures were diluted to an OD₅₉₅ = 0.5. Agar plates were inoculated from the diluted cultures by using a cotton swab. LB agar plates were used for all test strains, except for *S. cerevisiae* for which YPD agar (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose, 15 g/L agar) was used. Following plate inoculation, dried paper disks were placed on the agar and the plates were incubated at 30 C. The next day, growth and inhibition of growth was documented.

2.3.14 Bioinformatic and phylogenetic analysis

For initial bioinformatic characterisations of the BGC, an antiSMASH analysis (v 3.0.5) was conducted.^[175] Over time, further antiSMASH analyses with updated versions (most recent v5.1.0) followed.^[133] Clusterblast analysis from the antiSMASH 5.1.0 analysis was used for identification of related BGCs.^[176] In addition, BlastP analysis was conducted to collect PxaA sequences from other *Xenorhabdus* species. Genome sequences were retrieved from NCBI databases for *X. hominickii* DSM 17903 (NZ_NJAI01000003.1; WP_069318313.1 (PxaA), WP_069317110.1 (XhpA)), *Streptomyces cattleya* DSM 46488 (GCF_000237305.1; WP_078590884.1), *Actinosynnema mirum* DSM 43827 (NC_013093.1; WP_015803324.1), *Streptomyces* sp. Mg1, (GCF_000412265.2; WP_037789687.1), *Pseudomonas* sp. MIACH (GCF_001269925.1; WP_156338121.1), *Brenneria* sp. EniD312 (NZ_CM001230.1;

MATERIAL AND METHODS

WP_009113416.1), *Vibrio mimicus* VM223 (GCA_000176415.1; WP_000028413.1). In brackets first respective Refseq assembly accession number and second the respective NCBI protein ID for the protein sequences used for multiple sequence alignments (ClustalW) are given. *X. bovienii* (GCF_000027225.1; WP_012986920.1), *X. innexi* (GCF_002632485.1; WP_086953183.1), *X. stockiae* (GCF_002632825.1; WP_099124966.1), *X. mauleonii* (NITY01000015; PHM38685.1), *X. szentirmaii* (GCF_002632585.1; WP_051462298.1), *X. budapestensis* (GCA_002632465.1; WP_099135307.1), *Xenorhabdus* PB62.4 (GCA_014467285.1; WP_187654527.1), *X. japonica* (GCF_900115195.1; WP_092519303.1), *X. vietnamensis* (GCF_002127535.1; WP_167376261.1).^[177] For further characterisation of hits from BlastP analysis, antiSMASH analysis was used.^[133] Sequence alignments, dot plot analysis and tree model were generated using Geneious 6.1.8.^[178]

2.4 Glyoxpeptide project

2.4.1 Cultivation of strains

Microorganisms were grown in LB medium (10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl) for cloning, protein expression and overnight cultures. For cultivation on plates, 1.5 % (w/v) agar was added to the medium. Liquid cultivation was conducted on an orbital shaker. *E. coli* strains were cultivated at 37 °C and *Xenorhabdus* strains were cultivated at 30 °C. For compound production, *Xenorhabdus* strains were cultivated in XPP medium and variations thereof if not stated otherwise. The media compositions are given in Table 7. Induction of P_{BAD} exchange strains was done with 0.2 % (w/v) L-arabinose. When appropriate, antibiotics kanamycin (50 µg/mL) and/or ampicillin (100 µg/mL) were added.

Table 7. XPP medium preparation. XPP medium and two modified versions XPP-glycerol+Glc and XPP 3x glycerol.

XPP medium (1L)	10 g glycerol 20 mL salt A (M9) 20 mL salt B (M9) 2g L-amino acid mix 1g sodium pyruvate add H ₂ O, after autoclaving 2 mL vitamin solution 1 mL trace element solution
Salt A (M9) (1L)	350 g K ₂ HPO ₄ 100 g KH ₂ PO ₄
Salt B (M9) (1L)	29.4 g sodium citrate 50 g (NH ₄) ₂ SO ₄ 5 g MgSO ₄
Trace element solution (1L)	40 mg ZnCl ₂ 200 mg FeCl ₃ x 6 H ₂ O 10 mg CuCl ₂ x 2 H ₂ O 10 mg MnCl ₂ x 4 H ₂ O 10 mg Na ₂ B ₄ O ₇ x 10 H ₂ O 10 mg (NH ₄) ₆ Mo ₇ O ₂₄ x 4 H ₂ O

Vitamin solution (1L)	10 mg folic acid 6 mg Biotin 200 mg p-aminobenzoic acid 1 g thiamin hydrochloride 1.2 g pantothenic acid 2.3 g nicotinic acid 12 g pyridoxin hydrochloride 20 mg vitamin B12 Add dd H ₂ O
L-amino acid mix	2g L-alanine 2 g L-arginine 2 g L-aspartate 2 g L-asparagine 2g L-cysteine 2 g L-glutamate 2 g L-glutamine 2 g L-glycine 2 g L-histidine 2 g L-isoleucine 2g L-leucine 2 g L-lysine 2 g L-methionine 2 g L-phenylalanine 2 g L-proline 2 g L-serine 2 g L-threonine 2 g L-tyrosine
XPP – glycerol + Glc (1L)	XPP medium but glycerol is replaced by 0.2 % (w/v) glucose
XPP 3x glycerol (1L)	XPP medium but glycerol content is increased to 30 g

2.4.2 Microorganisms

Strains used in the work presented in section 3.3 are listed in Table 8.

Table 8. Strains used in the work presented in section 3.3.

	Strain	Genotype/Description	reference
	<i>E. coli</i> DH10B	F- <i>mcrA</i> , $\Delta(mrr-hsdRMS-mcrBC)$ $\Phi 80/lacZ\Delta M15,\Delta lacX74$, <i>recA1</i> , <i>endA1</i> , <i>araD139</i> , $\Delta(araleu)7697$ <i>galU</i> , <i>galK</i> , <i>rpsL</i> , <i>nupG</i> , λ -	Invitrogen
	<i>E. coli</i> DH10B :: <i>mtaA</i>	<i>E. coli</i> DH10B with $\Delta entD::mtaA$	[171]
	<i>E. coli</i> DH10B <i>mtaa</i> + pLZ58 + pFF1_ <i>pttBC</i>	<i>E. coli</i> DH10B <i>mtaa</i> containing pFF1_ <i>pttBC</i> and pLZ58	ref
MW864	<i>E. coli</i> DH10B <i>mtaa</i> + pLZ58 + pFF1_ <i>pttBC</i> _MonoOx	<i>E. coli</i> DH10B <i>mtaa</i> containing pFF1_ <i>pttBC</i> _ MonoOx and pLZ58	This work

MATERIAL AND METHODS

	BL21-Gold (DE3)	<i>E. coli</i> B F ⁻ <i>ompT hsdS</i> (r _B ⁻ m _B ⁻) <i>dcm</i> ⁺ <i>Tet</i> ^r <i>gal</i> λ(DE3) <i>endA Hte</i>	Agilent Technologies
MW800	BL21 Gold + pET11a_SMT3_MonoOx	BL21-Gold (DE3) contains pET11a_SMT3_MonoOx	This work
	<i>E. coli</i> S17-1 λpir	Tpr Smr <i>recA, thi, pro, hsdR</i> -M+RP4: 2-Tc:Mu:Km Tn7 λpir	[161]
	<i>E. coli</i> ST18	<i>E. coli</i> S17-1 λpir Δ <i>hemA</i>	[162]
SW187	<i>E. coli</i> ST18 + pCEP_Kan_xgsA	<i>E. coli</i> ST18 contains pCEP_Kan_xgsA	Sebastian Wenski
MW707	<i>E. coli</i> ST18 + pEB17_Km_ΔMonoOx	<i>E. coli</i> ST18 contains pCEP_Kan_xgsA	This work
MW705	<i>E. coli</i> ST18 + pEB17_Km_ΔxgsD	<i>E. coli</i> ST18 contains pCEP_Kan_xgsA	This work
	<i>X. hominickii</i> WT	<i>X. hominickii</i> DSM 17903 WT	[5]
	<i>X. hominickii</i> Δ <i>hfq</i>	<i>hfq</i> is deleted	[172]
PP108	<i>X. hominickii</i> Δ <i>hfq</i> P _{BAD} _xgsA	<i>hfq</i> is deleted, pCEP_kan_xgsA is inserted in front of <i>xgsA</i>	[154]
MW751	<i>X. hominickii</i> Δ <i>hfq</i> Δ <i>xgsD</i> P _{BAD} _xgsA	<i>hfq</i> is deleted, <i>xgsD</i> is deleted pCEP_kan_xgsA is inserted in front of <i>xgsA</i>	This work
MW757	<i>X. hominickii</i> Δ <i>hfq</i> ΔMonoOx P _{BAD} _xgsA	<i>hfq</i> is deleted, MonoOx domain sequence (in <i>xgsC</i>) is deleted pCEP_kan_xgsA is inserted in front of <i>xgsA</i>	This work

2.4.3 Plasmids

Plasmids used in for the work presented in section 3.3 are listed in Table 9.

Table 9. Plasmids used for the work presented in section 3.3.

plasmid	description	reference
pCEP_kan	pDS132 based, R6K ori, oriT, Km ^R , araC, P _{BAD}	[31]
pCEP_kan_SW452/3	<i>pCEP_kan</i> with first 1013 bp of <i>xgsA</i>	[154]
pEB17	pDS132 based, R6K ori, oriT, Km ^R , cipB derivative with additional BglII site, sacB	[46]
pEB17_Km_ΔMonoOx	R6K ori, oriT, sacB, Kan ^R with 1036 bp upstream of	This work

	SAAMDF MonoOx aa sequence and 1029 bp downstream of QLQEK MonoOx aa sequence	
pEB17_Km_ΔxgsD	R6K ori, oriT, sacB, Kan ^R with 1033 bp upstream of <i>xgsD</i> and 1112 bp (including 33 b of <i>xgsD</i> and 1079 bp downstream of <i>xgsD</i>)	This work
pET11a_SMT3	<i>T7</i> , <i>ampR</i> , <i>N-terminal his₆-smt3 tag</i> ,	[122]
pET11a_SMT3_MonoOx	pET11a_SMT3 with MonoOx domain sequence (4666 bp–5700 bp from <i>xgsC</i>)	This work
pFF1_	2μ ori, G418 ^R , P _{BAD} , pCOLA ori, MCS, Ypet-Flag, Km ^R	[174]
pLZ57	pFF1 with <i>pttBC</i> from <i>P. temperata</i> Meg1 under P _{BAD} control	[179]
pFF1_ <i>pttBC</i> _MonoOx	pLZ57 with MonoOx domain nt sequence (<i>xgsC</i> : 4666 nt – 5700 nt (<i>X. hominickii</i>) inserted after 2724 th nt in <i>pttB</i>)	This work
pCDF_ara_tacl	Modified pCDF_tacl/I that contains P _{BAD} and Sm ^R	[26]
pLZ58	pCDF_ara_tacl with inserted <i>pttA</i> from <i>P. temperata</i> Meg1 under P _{BAD} control	[179]

2.4.4 Oligonucleotides

Oligonucleotides used for the work presented in section 3.3 are listed in Table 10.

Table 10. Oligonucleotides used for the work presented in section 3.3. Oligonucleotides are given in combination with their purpose, e.g. plasmid and insert amplification. The DNA template for the corresponding PCR is indicated. “*Xhom*” = *X. hominickii*.

Plasmid/purpose	Oligonucleotide Name	Sequence 5' -> 3' (overhang_template specific region)	PCR template
Verification for pCEP_Kan_SW452/3 insertion	V_pCEP_fw	GCTATGCCATAGCATT TTTATCCATAAG	
	SW462	CTCCTGTGCGGTAATACAAACC	
pEB17_Km_Δ MonoOx	MW461	CCTCTAGAGTCGACCTGCAG_GCAACTGTGTTGATCACCTC	<i>Xhom</i> gDNA
	Mw462	CCGAGTACAGTGTCTTGGC_GGTGTGTTTCATGTTTCGGAAT	

MATERIAL AND METHODS

	MW463	ATTCCGAAACATGAACACACC_GCCAAGACACT GTA CTCCG	<i>Xhom</i> gDNA
	MW464	TCCCGGGAGAGCTCAGATCT_GCGCATTCTT GTTTATCCG	
Verification for <i>X. hominickii</i> Δ <i>MonoOx</i>	MW479	CGATTGATCCTGCCTACCC	
	MW480	CAGGGTTAGATAGATACTGGC	
pEB17_Km_ Δ <i>xgsD</i>	MW465	CCTCTAGAGTCGACCTGCAG_CGCACGCGTAT TGAGCG	<i>Xhom</i> gDNA
	MW466	CTAATTTAGTAGAATTATGCATGCG_AAGAGTA GCTGTGGGGC	
	MW467	CTGCCCCACAGCTACTCTT_CGCATGCATAATT CTACTAAATTAG	<i>Xhom</i> gDNA
	MW468	TCCCGGGAGAGCTCAGATCT_GTCACATAGTG GCCTATGC	
Verification for <i>X. hominickiii</i> Δ <i>xgsD</i>	MW477	CGCCTCCGGTAAACTGG	
	MW478	GTCGTCTTTCAAGTTGCAGC	
pET11a_SMT 3_MonoOx	pET11a- FW	TAAGGATCCGGCTGCTAAC	pET11 a
	pET11a_ smt3-RV	ACCACCAATCTGTTACGA	
	MW594	CATCGTGAACAGATTGGTGGT_TCAGCCGCAA TGGATTTTCAG	<i>Xhom</i> gDNA
	MW595	TTTGTTAGCAGCCGGATCCTTA_CGAGTACAGT GTCTTGGC	
pFF1_pttABC	MW667	CGTAGCGTTACTATTACCTCG	<i>P. tem</i> <i>perata</i> Meg1 gDNA
	MW668	AACTGAAATCCATTGCGGCTGA_AGAATTGGGT TTGTGCAAAAACAAC	
	MW669	GTTGTTTTGCACAAACCCAATTCT_TCAGCCGC AATGGATTTTCAG	<i>Xhom</i> gDNA
	MW670	GTTTTTTTCAGCCATTTATTTAATTCTTC_TTTCTC CTGCAATTGACCC	
	MW671	GATTATTTGGGTCAATTGCAGGAGAAA_GAAGA ATTAAATAAATGGCTGAAAAAAC	<i>P. tem</i> <i>perata</i> Meg1 gDNA
	MW672	CCGGTTGACCCAGAGGTA	

2.4.5 Deletion of *xgsD* and *XisC_MonoOX* domain encoding sequence

The deletion of genes was conducted as described in section 2.2.5. The recipient strain during conjugation was *X. hominickii* Δ *hfq*. For *xgsD* deletion the donor strain was *E. coli* ST18 + pEB17_Km_ Δ *xgsD*, while for deletion of the MonoOx domain sequence, *E. coli* ST18 + pEB17_Km_ Δ *MonoOx* was used (Table 7). The respective oligonucleotides for plasmid construction and deletion verification are listed in Table 9. Following confirmation of the deletion, conjugation with *E. coli* ST18 + pCEP_kan_*xgsA* was performed to exchange the promoter in front of *xgsA*. The protocol for conjugation is described in section 2.2.5.

2.4.6 Sample preparation for LC-MS analysis

If not stated otherwise, following cultivation for 48 h, culture samples were extracted 1:1 with ACN, centrifuged at full speed for 30 min and the supernatant was subjected to LC-MS analysis.

2.4.7 LC-MS analysis

HPLC-MS analysis was conducted on a Dionex Ultimate 3000 (Thermo Fisher Scientific) RP-HPLC coupled to an AmaZonX (Bruker) ESI-IT mass spectrometer. 5 μ L sample were injected and a 16 min gradient 5-95 % ACN/ 0.1% formic acid in H₂O/ 0.1 % formic acid (0-4 min 5 % ACN, 4-14 min 5-60 % ACN, 14-15 100 % ACN, 15-16 min 5 % ACN) was run with a C₁₈ Acquity UPLC BEH column (Waters) and a flow rate of 0.2 mL/min. The mass spectrometer was set to positive ionisation mode with a mass range m/z 100-1200 and cone voltage 4500 V. HR data were generated using a Bruker Impact IITM ESI-Q-OTOF instrument in positive ionisation mode with a mass range m/z 100-1200, cone voltage 1500 V coupled to a Dionex Ultimate 3000 (Thermo Fisher Scientific) RP-HPLC. 5 μ L sample were injected and a 16 min gradient 5-95 % ACN/ 0.1% formic acid in H₂O/ 0.1 % formic acid (0-4 min 5 % ACN, 4-14 min 5-60 % ACN, 14-15 100 % ACN, 15-16 min 5 % ACN) was used. The flow rate was 0.2 mL/min and a C₁₈ Acquity UPLC BEH column (Waters) was used. The data were analysed using the software Data Analysis 4.3 (Bruker). If not stated otherwise, the HPLC-UV-MS chromatograms are shown on the same scale in signal intensity when presented on the same x-axis.

The samples from experiments regarding MonoOx domain encoding sequence inserted in *pttB* were prepared and extracted as in described in sections 2.3.10 and 2.3.11.

2.4.8 MonoOx domain protein expression and purification

For protein expression of the XgsC_MonoOx domain, the 1035 nt sequence was cloned under control of T7 promoter and expressed as fusion protein with an N-terminal His₆-SMT3-tag. The used oligonucleotides to generate the expression plasmid pET11_SMT3_MonoOx are listed in Table 10. The protein was expressed in *E. coli* BL21 Gold DE3 cells. 500 mL LB with 100 μ g/mL ampicillin were inoculated from an overnight culture at OD₅₉₅ = 0.1 and incubated with shaking at 37 °C. When the culture reached an OD₅₉₅ of 0.7, 0.1 mM isopropyl- β -D-1-thiogalactopyranoside (IPTG) was added to induce protein expression and the cultivation temperature was changed to 20 °C. After 18 h of cultivation, the cells were harvested by centrifugation and stored at -20°C. For pellet workup, the cell pellet was thawed and resuspended in 42 mL lysis buffer (21 mL 2x lysis buffer: 100 mM Tris-HCl (pH 8.0), 600 mM NaCl, 2 mM EDTA, 6 μ L β -mercaptoethanol, 420 μ L 1 M MgCl₂, 1 ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail tablet, 0.5 μ L benzoase). The sample was kept on ice for the following steps. The lysate was clarified via centrifugation at 11000 rpm and 4 °C for 45 min and the supernatant was loaded onto a 5 mL

MATERIAL AND METHODS

His Trap™ HP column (GE Healthcare) and purified with an NGC Discover Chromatography System (ChromLab v6.0, Bio-Rad Laboratories GmbH). The column was equilibrated in buffer A (50 mM Tris-HCl (pH 8.0), 250 mM NaCl, 1% glycerol, 20 mM imidazole) and the protein was eluted with buffer B (50 mM Tris-HCl (pH 8.0), 250 mM NaCl, 1 % glycerol, 400 mM imidazole) by applying a linear gradient from 30% - 60 % buffer B. Subsequently, the His₆-SMT3-tag was cleaved off from the target protein by the protease Ulpl while the sample was dialysed (SpectraPor tubing (Spectrum Labs, MWCO: 5.000 Da)) into dialysis buffer (50 mM Tris-HCl (pH 8.0), 50 mM NaCl, 2 mM β-mercaptoethanol) overnight at 4°C. The next day, reverse NiNTA purification with a HisTrap™ HP column (GE Healthcare) was performed to remove uncleaved protein and the affinity tag and the flowthrough fraction was collected. The protein concentration was determined using photometer NanoDrop2000c (Thermo Fisher Scientific) and protein properties (extinction coefficient = 46.535, pI = 4.8, MW = 38.4 kDa) were extracted using Geneious 6.1.8. Sodiumdodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was conducted to check protein content, purity and size.

2.4.9 Sodiumdodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis

For SDS-PAGE analysis, a discontinuous Tris-glycine SDS-PAGE system was applied.^[180] The samples were mixed with 3x SDS loading dye buffer (20% glycerine, 4% SDS, 50 mM Tris-HCl pH6.8, 1.5 mM β-mercaptoethanol, 0.2% bromophenol blue) and incubated for 5 min at 95 °C prior to analysis. 8 µl sample mix and 4 µl size standard (PageRuler™ Plus Prestained Protein Ladder, Thermo Fisher Scientific) were loaded onto a SDS-polyacrylamide gel (stacking gel: pH 6.8, 5 % polyacrylamide (37.5:1 acrylamide: bisacrylamide); separating gel: pH8.8, 10 % polyacrylamide). SDS-Page was run for approx. 45 min with 40 mA per gel in SDS run buffer (1% (w/v) SDS, 1 M Tris-HCl (pH 8.45), 1 M Tricin). Coomassie Staining Solution (34% MeOH, 0.003% H₃PO₄, 0.00035% (w/v) Coomassie Brilliant Blue G-250) was used for staining.

2.4.10 Isothermal titration calorimetry

For ITC analysis MicroCal iTC200 (Malvern Instruments) was used. The purified MonoOx domain (50 µM) was provided in the sample cell and after a 120 sec delay time the respective cofactor (500 µM FAD or FMN) was added stepwise in 20 injections à 0.2 µL injections with 120 sec intervals. For cofactor reduction 5 mM DTT was added to the cofactor sample. The measurement was conducted at 20 °C with an 11µcal⁻¹ reference power setting, 750 rpm stirring speed and a high feedback mode. The software Origin7.0 (OriginLab) was used for automated peak integration and data visualisation. A one site binding model was assumed.

2.4.11 Cofactor analysis from denatured purified protein

A fraction of the purified supernatant sample containing the MonoOx domain was incubated at 95 °C for 5 min, centrifuged at full speed for 30 min and the supernatant was subjected to LC-HR-MS analysis (2.2.9).

2.4.12 Insertion of MonoOx domain in phototemtime synthetase B (PttB)

For insertion of the MonoOx domain encoding sequence in the phototemtime synthetase encoding BGC *pttABC*, the plasmids generated by Lei Zhao for heterologous expression of this BGC in *E. coli* were used. Plasmids pLZ57 and pLZ58 are listed in Table 10. pLZ57 was linearised using restriction enzymes RsrII and BstZ17I (both New England Biolabs) and the 21.7 kb fragment purified from agarose gel. The region left from the insertion site with overhangs to the RsrII restriction site was amplified with oligonucleotides MW667/MW668 from *P. temperata* Meg1 gDNA as well as the region right from the insertion site which was amplified with oligonucleotides MW671/2 with overhangs to the BstZ17I restriction site. The inserted MonoOx domain encoding sequence was amplified with primers MW669/MW760 from *X. hominickii* gDNA. The three PCR fragments contained appropriate overhangs and were assembled with the linearised pLZ57 backbone fragment via Hot Fusion cloning. *E. coli* DH10B ::*mtaA* was used for expression and co-transformed with the assembled pFF1_pttBC_MonoOx and pLZ58. For production, cells were cultivated in LB medium (50 µg/mL kanamycin, 50 µg/mL spectinomycin) at 30 °C with shaking for 72 h in presence of 2 % Amberlite™ XAD-16 and 0.2 % (w/v) L-arabinose. The extracts were prepared as described in section 2.3.10 and LC-MS analysis was done as described in section 2.3.11.

2.4.13 Growth experiments

Growth phase monitoring was conducted in 96-well microtiterplates pureGRADE™ BRANDplates® (BRAND Scientific GmbH) using a Spark® microplate reader (TECAN) with orbital shaking at 30 °C for 24 h. OD₅₉₅ was measured every 20 minutes. XPP medium (with 50 µg/mL kanamycin) was used for cultivation and culture wells were inoculated at OD₅₉₅ = 0.1 from overnight culture. For XPP medium control and analysed strains, triplicates were prepared. For induction, 0.2 % (w/v) L-arabinose was added at the beginning of cultivation.

2.4.14 Bioinformatic and phylogenetic analysis

Clusterblast analysis included in antiSMASH analysis was used for first analysis of related BGCs.^[133] BlastP analysis was used for identification of sequences similar to XgsC_A2 with integrated MonoOx domain.^[170] The sequences were accessed from NCBI databases for the following strains: *X. hominickii* DSM 17903 (NZ_NJAI01000003.1), *X. miraniensis* (GCF_002632615.1), *Xenorhabdus* PB62.4 (GCF_014467285.1), *Lonsdalea populi* (GCF_003730085.1), *Pseudomonas syringae* pv. *coryli* (GCF_001400235.1), *Chloroflexi bacterium* (VBJK01000441.1), *Noviherbaspirillum* sp. Root189 (GCF_001428545.1), *Burkholderia ubonensis* (GCF_001524455.1), *Janthinobacterium* sp. DG3

MATERIAL AND METHODS

(GCF_000344615.1).^[177] The respective Refseq codes are given in brackets. *Photorhabdus temperata* K122 genome sequence was not available at NCBI but was added manually to the data set from our lab's strain collection (personal communication, Bode Lab). AntiSMASH analysis was used to identify the predicted domain structure of protein hits from the BlastP analysis.^[133] For sequence alignments (ClustalW) Geneious 6.1.8 was used.^[178] Protein accession numbers for MonoOx and Ox domain alignments are given in brackets for the following: MonoOx domains: PuwE_MonoOx (AIW82282.1), Sky39_MonoOx (WP_097962190.1), MelG_MonoOx (CAD89778.1), MtaG_MonoOx (WP_002618700.1), LuxA_MonoOx (), BlmIII_Ox (AAG02365.1), EpoB_Ox (AAF62881.1), MelD_Ox (WP_095982063.1), MtaD_Ox (WP_013376004.1).

2.5 Xildivaline project

2.5.1 Cultivation of strains

Microorganisms were cultivated in LB medium for cloning purposes and overnight cultures. For cultivation on LB agar plates, 1.5 % w/v agar was added. If not stated otherwise, *E. coli* cells were grown at 37 °C while *Xenorhabdus* was incubated at 30 °C. Antibiotics kanamycin (50 µg/mL) or ampicillin (100 µg/mL) were added when appropriate. Production cultures were cultivated in XPP medium (2.4.1) at 30 °C for 72 h. For induction, 0.2 % (w/v) L-arabinose was added at the beginning of cultivation.

2.5.2 Microorganisms

Strains used for the work presented in section 3.4 (xildivaline project) are listed in Table 11.

Table 11. Strains used for the work presented in section 3.4.

	Strain	Genotype/Description	reference
	<i>E. coli</i> DH10B	F- <i>mcrA</i> , Δ (<i>mrr-hsdRMS-mcrBC</i>) Φ 80/ <i>lacZ</i> Δ M15, Δ <i>lacX74</i> , <i>recA1</i> , <i>endA1</i> , <i>araD139</i> , Δ (<i>araleu</i>)7697 <i>galU</i> , <i>galK</i> , <i>rpsL</i> , <i>nupG</i> , λ -	Invitrogen
	<i>E. coli</i> S17-1 λ pir	Tpr Smr <i>recA</i> , <i>thi</i> , <i>pro</i> , <i>hsdR</i> -M+RP4: 2- Tc:Mu:Km Tn7 λ pir	^[161]
	<i>E. coli</i> ST18	<i>E. coli</i> S17-1 λ pir Δ <i>hemA</i>	^[162]
SW189	<i>E. coli</i> ST18 + pCEP_kan_ <i>xisA</i>	<i>E. coli</i> ST18, contains pCEP_SW456/7	^[153]
MW706	<i>E. coli</i> ST18 + pEB17_ Δ MT	<i>E. coli</i> ST18, contains pEB17_ Δ MT	This work
MW708	<i>E. coli</i> ST18 + pEB17_ Δ PD	<i>E. coli</i> ST18, contains pEB17_ Δ PD	This work

MW582	<i>E. coli</i> ST18 + pEB17_T1	<i>E. coli</i> ST18, contains pEB17_T1	This work
MW584	<i>E. coli</i> ST18 + pEB17_T1	<i>E. coli</i> ST18, contains pEB17_T2	This work
	<i>X. hominickii</i> WT	<i>X. hominickii</i> DSM 17903 WT	[5]
MW696	<i>X. hominickii</i> P _{BAD_xisA}	pCEP_kan_xisA is integrated in front of xisA	This work
MW752	<i>X. hominickii</i> ΔxisE P _{BAD_xisA}	xisE is deleted, pCEP_kan_xisA is integrated in front of xisA	This work
MW754	<i>X. hominickii</i> ΔxisD P _{BAD_xisA}	xisD is deleted, pCEP_kan_xisA is integrated in front of xisA	This work
MW675	<i>X. hominickii</i> xisB_T1* P _{BAD_xisA}	point mutation in xisB at position 3973 (now encoding Ala instead of active site Ser in XisB_T1), pCEP_kan_xisA is integrated in front of xisA	This work
MW620	<i>X. hominickii</i> xisB_T1* P _{BAD_xisA}	point mutation in xisB at position 4213 (now encoding Ala instead of active site Ser in XisB_T2), pCEP_kan_xisA is integrated in front of xisA	This work
	<i>X. hominickii</i> Δhfq	<i>X. hominickii</i> Δhfq	[172]
JC97	<i>X. hominickii</i> Δhfq P _{BAD_xisA}	hfq is deleted, pCEP_kan_xisA is integrated in front of xisA	[153]
MW753	<i>X. hominickii</i> Δhfq ΔxisE P _{BAD_xisA}	hfq is deleted, xisE is deleted, pCEP_kan_xisA is integrated in front of xisA	This work

2.5.3 Plasmids

Plasmids used for the work presented in section 3.4 (xildivaline project) are listed in Table 12.

Table 12. Plasmids used for the work presented in section 3.4

plasmid	description	reference
pCEP_kan	pDS132 based, R6K ori, oriT, Km ^R , araC, P _{BAD}	[31]
pCEP_kan_SW456/7	pCEP_kan with first 1046 bp of xisA	[153]

MATERIAL AND METHODS

pEB17	pDS132 based, R6K ori, oriT, Km ^R , cipB derivative with additional BglII site, sacB	[46]
pEB17_Km_ΔMT	R6K ori, oriT, sacB, Kan ^R with 1098 bp upstream of <i>xisE</i> (including the first 3 nt of <i>xisE</i>) and 1063 bp downstream of <i>xisE</i> (including the last 6 nt of <i>xisE</i>)	This work
pEB17_Km_ΔPD	R6K ori, oriT, sacB, Kan ^R with 1010 bp upstream of <i>xisD</i> (including the first 12 nt of <i>xisD</i>) and 1093 bp downstream of <i>xisD</i> (including the last 75 nt of <i>xisD</i>)	This work
pEB17_Km_T1	R6K ori, oriT, sacB, Kan ^R with 838 bp upstream of T1 point mutation in <i>xisB</i> and 847 bp downstream of T1 point mutation in <i>xisB</i>	This work
pEB17_Km_T2	R6K ori, oriT, sacB, Kan ^R with 855 bp upstream of T2 point mutation in <i>xisB</i> and 855 bp downstream of T2 point mutation in <i>xisB</i>	This work

2.5.4 Oligonucleotides

Oligonucleotides used for the work presented in section 3.4 (xildivaline project) are listed in Table 13.

Table 13. Oligonucleotides used for the work presented in section 3.4. Oligonucleotides are given in combination with their purpose, e.g. plasmid and insert amplification. The DNA template for the corresponding PCR is indicated. Introduced point mutation is indicated in red. “*Xhom*” = *X. hominickii*.

Plasmid/purpose	Oligonucleotide Name	Sequence 5' -> 3' (overhang_template specific region)	PCR template
Verification for pCEP_Kan_ <i>xisA</i> insertion	V_pCEP_fw	GCTATGCCATAGCATTTCATCCATAAG	
	SW464	GCGAGTAAACCAGTAAGCATTCC	
pEB17_Km_ΔMT	MW473	CCTCTAGAGTCGACCTGCAG_CGCAGCTTAGCC AATTAATC	<i>Xhom</i> gDNA
	MW474	CTTCTTAAGTAGGTTCTATGC_CATTGTTTGA ATTCTAACAATC	
	MW475	GAAGTGTTAGGAATTCAAACAATG_GCATAGAAC CTACTAGTTAAGAAG	<i>Xhom</i> gDNA
	MW476	TCCCGGGAGAGCTCAGATCT_CGTATCGACTCAATGCTTGC	

Verification of <i>X. hominickii</i> $\Delta xisE$	MW483	CGCTTTGAACTGAGTTCCGG	
	MW484	GGCAAATAAGTCCTTTATGACG	
pEB17_Km ΔPD	MW469	CCTCTAGAGTCGACCTGCAG_TTGCGAACGAAG CAGAAGC	<i>Xhom</i> gDNA
	MW470	CGCTTTAGTGGAGATAGGTAATC_TCTAACTGTC ATTTATTTGCCTAATG	
	MW471	CATTAGGCAAATAAATGACAGTTAGA_GATTACC TATCTCCACTAAAGC	<i>Xhom</i> gDNA
	MW472	TCCCGGGAGAGCTCAGATCT_CGGGATGCGTTC ATATAGC	
Verification of <i>X. hominickii</i> $\Delta xisD$	MW481	CGATGAGCTGGTGATGATG	
	MW482	GCTGCAACCTACCGTCG	
pEB17_Km $\Delta T1$	MW325	CCTCTAGAGTCGACCTGCAG_GGTCCGTTAAAA GGCATTATTCAC	<i>Xhom</i> gDNA
	MW326	CGCTTGGATAGTTGTGTAAG_GATCAACGCGTCA CCACCTAC	
	MW327	GATGTAGGTGGTGACGCGTTGATC_TTTACACAA CTATCCAAGCGTATTC	<i>Xhom</i> gDNA
	MW328	TCCCGGGAGAGCTCAGATCT_GCAAACAGCCTTT TACAAACAACC	
pEB17_Km $\Delta T2$	MW329	CCTCTAGAGTCGACCTGCAG_GCAATATCTGATT GAGCAGACC	<i>Xhom</i> gDNA
	MW330	GATTTTCGCTATGAGCGAAGTAAACAT_GACAGC CTCACCGCCTAG	
	MW331	GACCTAGGCGGTGAGGCTGTC_ATGTTTACTTCG CTCATAGCG	<i>Xhom</i> gDNA
	MW332	TCCCGGGAGAGCTCAGATCT_GTCGACACAAAG TGTAATG	
Verification for T1 T2	MW355	CCAAGTGAATGAGCGC	
	MW356	CCAAGCAAGTGGCTGTAAC	

2.5.5 Promoter exchange

The promoter exchange mutant *X. hominickii* P_{BAD_xisA} was generated via conjugation between *X. hominickii* WT and donor strain *E. coli* ST18 + pCEP_kan_*xisA*. The detailed protocol is described in section 2.3.5. The donor *E. coli* ST18 + pCEP_kan_*xisA* was generated in previous work by Janine Chekaiban.^[153]

2.5.6 Deletions of *xisE* and *xisD*

For deletion of *xisD* and *xisE* respectively, deletion plasmids were constructed as described in section 2.2.5. The generated plasmids and used oligonucleotides are listed in Tables 12 and 13. The deletion mutants were obtained following the protocol described in section 2.2.5.x.x. Finally, P_{BAD} was inserted in front of *xisA* as described in section 2.3.5.

MATERIAL AND METHODS

2.5.7 T domain point mutations

In the two T domains of XisB the respective active site serine was mutated to alanine by the introduction of point mutations. In case of the T1 domain encoding sequence, the serine codon was altered from TCT -> GCT while it was changed from TCG -> GCG for the T2 domain encoding sequence. In both cases, the first position thymine was replaced by a guanine to change the codon specificity from serine to alanine. The point mutations were introduced with the help of oligonucleotides carrying the altered sequence (Table 13). Approximately 1 kb regions upstream and downstream of the position of the point mutation were amplified by PCR using the modified oligonucleotides. Amplicons were used to assemble the plasmid in analogy to deletion plasmids as explained in section 2.2.5. In the following, the protocol for conjugation and homologous recombination (section 2.2.5) was applied. Strains obtained after growth on 6% sucrose agar were screened for the point mutation. Therefore, gDNA was purified, the target sequence was amplified by PCR with the oligonucleotides MW355/MW356, in both cases. Sequencing (GATC Eurofins) was used to verify the sequence. In case of T2 point mutation, the restriction site of restriction endonuclease HinfI (New England Biolabs) is disturbed in the point mutation. Hence, restriction digest with HinfI according to the manufacturer's instruction was performed of the PCR products to exclude wild type sequences prior to sequencing.

2.5.8 LC-MS analysis

If not stated otherwise, 500 μ L culture were mixed 1:1 with 500 μ L ACN after 72 h of cultivation. The samples were then diluted 1:10 with ACN, centrifuged at full speed for 30 min and the supernatant was subjected to LC-MS analysis. For LC-MS analysis, instruments and applied methods are identical with section 2.2.9.

For relative quantification of xildivalines to compare production levels between the two respective T1 and T2 inactivation strains, biological triplicates were prepared. The software Compass Target Analysis 1.3 (Bruker Daltonik GmbH) was used for automated peak integration (xildivaline **26a** ($C_{19}H_{37}N_3O_4$ H^{+1} , 372.2852, t_R = 4.0 min), xildivaline **26b** ($C_{19}H_{37}N_3O_4$ H^{+1} , 372.2852, t_R = 4.5 min), xildivaline **27** ($C_{19}H_{35}N_3O_3H^{+1}$, 354.2746, t_R = 5.3 min), xildivaline **32** ($C_{19}H_{39}N_3O_3H^{+1}$, 358.3064, t_R = 4.9 min), xildivaline **31** ($C_{19}H_{39}N_3O_4H^{+1}$, 374.301, t_R = 3.9 min)). The peak area values were normalised to the OD_{595} of the respective culture measured at the end of cultivation. For statistical significance, two-sided unpaired t-test was conducted using GraphPad Prism version 8.1.2 for Windows, GraphPad Software (San Diego, California USA, www.graphpad.com).

2.5.9 4-bromothiophenol (BTP) addition

A 5 mL 72 h culture of *X. hominickii* $\Delta hfq + P_{BAD_xisA}$ was harvested at 4000 rpm for 20 min and the supernatant was stored at -20 °C. Non-inoculated LB was used as medium control.

For 4-bromothiophenol (BTP) addition, 50 µl 1mg/mL 4-bromothiophenol (MeOH) were added to 500 µL thawed culture supernatant. Samples were incubated overnight at 30 °C. As controls, samples without BTP addition and samples with addition of 500 µL MeOH were used. Prior to HPLC-MS analysis, samples were diluted 1:10 with ACN and centrifuged at full speed for 30 min.

2.5.10 Growth experiments

Growth phase monitoring was conducted as described in section 2.4.13.

2.5.11 Bioinformatic and phylogenetic analysis

Clusterblast analysis from antiSMASH was used for identification of related BGCs.^[176] Genome sequences were acquired from NCBI databases for *X. hominickii* DSM 17903 (NZ_NJAI01000003.1), *X. hominickii* ANU1 (NZ_CP016176.1), *Photorhabdus asymbiotica* DSM 15149 (NC_012962.1), *Vibrio tubiashii* ATCC 1909 (GCF_000772105.1).^[177] Refseq record numbers are given in brackets. Geneious 6.1.8 was used for plasmid map display.^[178] The different KR type sequences used for alignment with XisB_KR were retrieved from protein database (PDB). The PDB accession codes are given in brackets for amphotericin synthase KR2 AmpKR2 (4DIF), amphotericin synthase KR11 AmpKR11 (4L4X), tylosin synthase KR1 TyIKR1 (2Z5L), erythromycin synthase KR1 EryKR1 (2FR1), erythromycin synthase KR3 EryKR3 (3QP9). KR subtype characterisation was conducted using the protocol and characteristics described by Keatinge-Clay 2007.^[114] Multiple sequence alignments (ClustalW) were prepared with Geneious 6.1.8.^[178] The sequences were accessed from NCBI databases for VibH (AAD48879.1), Kj12C (ANG60373.1), XndB (WP_013184319), TqaA (ADY16697.1), PaxA (WP_012988600.1), PaxB (WP_012988599).^[177]

RESULTS

3 Results

3.1 Xenofuranone

3.1.1 *xfkB* deletion

Simultaneous deletion of the two adjacent genes *xfxA* and *xfxB* in *X. szentirmaii* resulted in a strain deficient to produce xenofuranones as shown in my master thesis.^[155] While *xfxA* encodes an NRPS-like enzyme, *xfxB* is predicted to encode a gluconolactonase by antiSMASH analysis.^[133] In this work, deletion of *xfxB* alone resulted in the detection of the carboxylated xenofuranone with m/z 311.1 $[M+H]^+$ (**5**) which is not observed in the wild type strain or in the induced plasmid-based complementation strain (Fig. 24). Additionally, xenofuranone B (**1**) is observed in the Δ *xfxB* strain and the non-induced plasmid-based complementation strain. When induced, plasmid-based complementation restores wild type xenofuranone production levels with xenofuranone A (**2**) as the main compound.

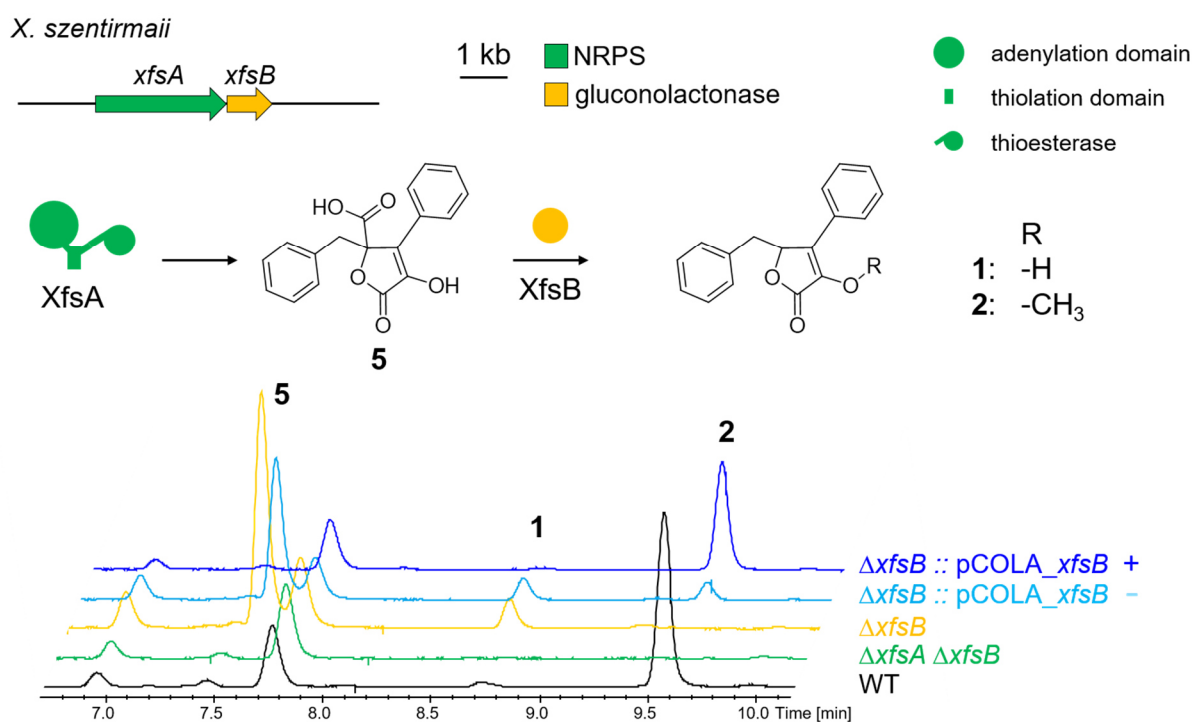


Figure 24. UV₂₈₀ chromatograms from extracts of the *xfxB* deletion mutant and plasmid-based complementation strain. The carboxylated furanone intermediate **5** is observed in *xfxB* deletion and non-induced complementation mutant but abolished in the induced plasmid-based complementation strain *X. szentirmaii* Δ *xfxB* ::pCOLA_*xfxB*. + arabinose-induction – no induction

The absence of a compound that is both carboxylated and O-methylated indicates that carboxylation occurs prior to methylation in the biosynthesis (Fig. 28). However, when XAD-16 is added to the cultures for subsequent compound extraction, and only then, a methylcarboxylated derivative **6** with m/z 325.1 $[M+H]^+$ is detected (Fig. 25). This compound was observed before by other lab members and their purification and structure elucidation revealed the carboxyfunction to be methylated (personal communication, Fig. S1). This compound is present in the wild type and in the Δ *xfxB* strain, but only when XAD-16 is added

during cultivation. Feeding with d_3 -methionine showed no incorporation of the d_3 -methyl group for **6** but is observed for xenofuranone A as described previously by Brachmann *et al.* 2006 (Fig. 25).^[152]

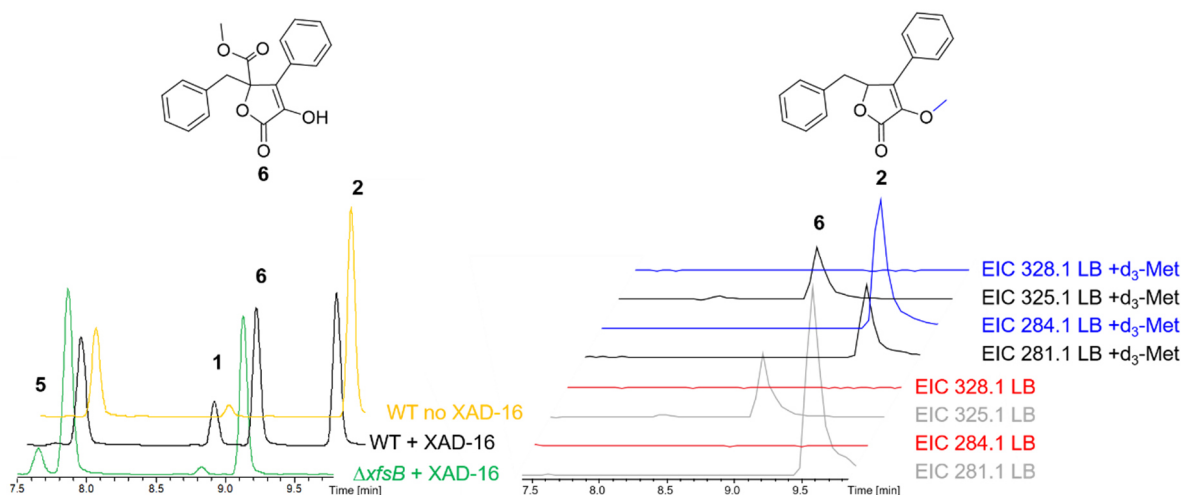


Figure 25. Compound **6** is exclusive in XAD-16 extracts and exhibits SAM-independent methylation. Compound **6** with m/z 325.1 $[M+H]^+$ is only observed upon cultivation with XAD-16, and then is found in *X. szentirmaii* WT and $\Delta xfsB$ (Left). Feeding with deuterated d_3 -methionine (d_3 -Met) shows incorporation for xenofuranone A (**2**) but not for compound **6**.

3.1.2 Promoter exchange based activation of the *xfS* BGC

Activation of the P_{BAD} promoter in front of *xfSA* (P_{BAD_xfSA}) leads to strong production of xenofuranone A and B in the *X. szentirmaii* wild type strain. When P_{BAD_xfSA} is activated in *X. szentirmaii* Δhfq strain, the non-methylated xenofuranone B is the major compound produced (Fig 26).

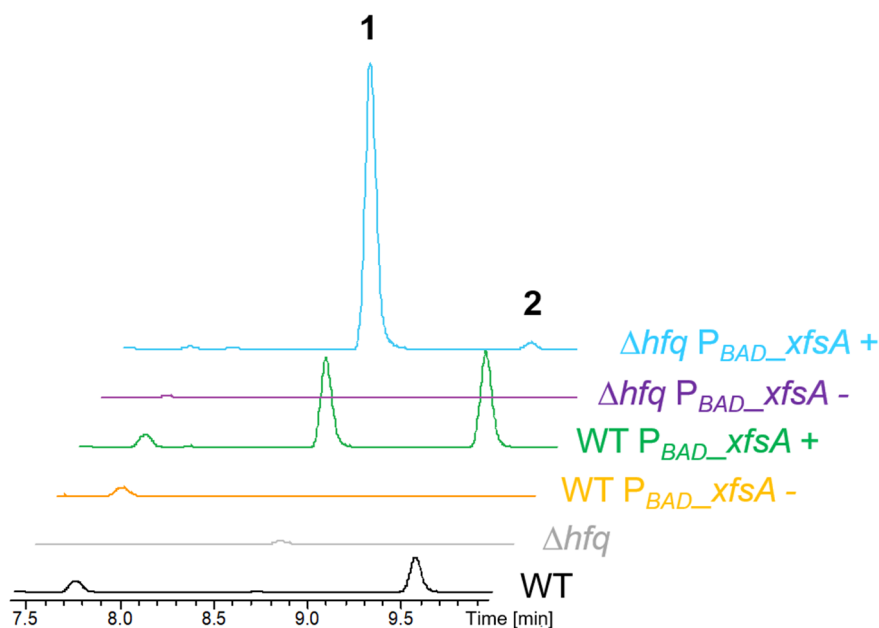


Figure 26. Distinct xenofuranone compound profiles for P_{BAD_xfSA} activation in *X. szentirmaii* WT and *X. szentirmaii* Δhfq . P_{BAD} activation in the *X. szentirmaii* WT strain shows production of **1** and **2** (green) and P_{BAD} activation in the *X. szentirmaii* Δhfq strain shows production of mainly **1** (light blue). UV₂₈₀ BPCs are shown. + induction with arabinose. – no induction.

RESULTS

The two distinct xenofuranone product profiles observed for promoter activation in *X. szentirmaii* WT or *X. szentirmaii* Δhfq strain were used to identify the xenofuranone methyltransferase. The xenofuranone BGC does not encode a methyltransferase and thus all SAM-dependent methyltransferases were initially regarded as potential candidates. However, the observation of two distinct xenofuranone compound profiles allowed a more targeted identification approach. It was assumed that the methyltransferase responsible for xenofuranone A biosynthesis was downregulated in the *X. szentirmaii* Δhfq strain. A proteome analysis for comparison of *X. szentirmaii* WT with *X. szentirmaii* Δhfq was conducted by Dr. Timo Glatter (Max-Planck-Institute for Terrestrial Microbiology, Marburg). From the processed data (see .xlsx file on attached CD-ROM), I identified four candidates that were predicted to be SAM-dependent methyltransferases and for which protein levels were less abundant in the *X. szentirmaii* Δhfq strain compared to the wild type (Fig. 27A). For all four candidates, single gene deletions were analysed and indeed deletion of one of the four candidate genes, *xfsC* (*xsze03490*), resulted in a strain deficient in xenofuranone A production (Fig. 27B). Plasmid-based complementation restores xenofuranone A production, when induced (Fig. 27C). The respective gene deletion of the three remaining candidate genes, namely *xsze01860*, *xsze01783* and *xsze02284*, had no effect on xenofuranone production and UV₂₈₀ BPCs showed no difference between extracts from these deletion strains and the wild type (Fig. 27B).

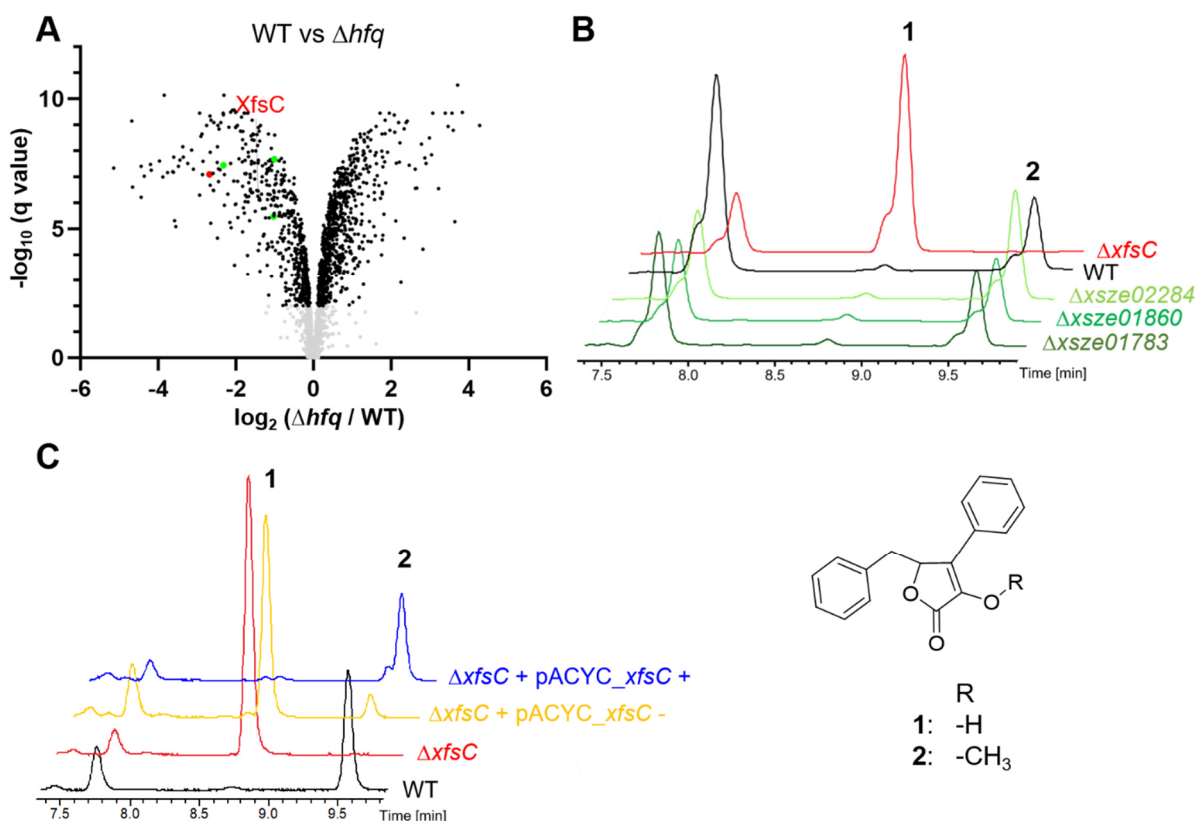

















Figure 27. Xenofuranone methyltransferase identification. **A** Comparative proteome analysis of *X. szentirmaii* WT with Δhfq shows three SAM-dependent methyltransferase candidates (green) and the identified xenofuranone methyltransferase XfsC (red). **B** UV₂₈₀ BPCs for extracts of *X. szentirmaii* WT and deletion strains of the four MT candidate genes. Deletion of *xfsC* (red) results in production of xenofuranone B (**1**) but no production of xenofuranone A (**2**) which is detected in the WT and the three candidate deletion strains (green). **C** Plasmid-based complementation of the *xfsC* deletion. Induction of *X. szentirmaii* $\Delta xfsC$ + pACYC_ *xfsC* restores wild type-like xenofuranone production. UV₂₈₀ BPCs are shown. + induction with arabinose. – no induction.

The identified methyltransferase XfsC belongs to class I SAM-dependent methyltransferases according to UniProt and thus shows a GxGxG conserved motif at the end of the first beta strand.^{[181][182]} All xenofuranone producing *Xenorhabdus* strains from our lab's strain collection show production of xenofuranone A and encode for a XfsC candidate (Table 14).

RESULTS

Table 14. Overview of xenofuranone biosynthesis genes in xenofuranone producing *Xenorhabdus* strains. Query cover and percentage identity are given for XfsC sequences. *X. szentirmaii* was used as reference sequence. All analysed *Xenorhabdus* strains produce xenofuranone A (**2**) and B (**1**).

	XfsA	XfsB	XfsC	query cover	percentage identity	xenofuranone A	xenofuranone B
<i>X. szentirmaii</i> DSM				100 %	100 %	√	√
<i>X. szentirmaii</i> US				100 %	100 %	√	√
<i>X. mauleonii</i>				100 %	80.56 %	√	√
<i>X. stockiae</i>				100 %	61.42 %	√	√
KJ12.1				100 %	61.73 %	√	√

 NRPS  gluconolactonase  methyltransferase

3.1.3 Xenofuranone biosynthesis

Figure 28 displays the proposed xenofuranone biosynthesis mediated by the NRPS-like enzyme XfsA, decarboxylase XfsB and methyltransferase XfsC. XfsA_A subsequently activates two phenylpyruvic acid building blocks of which the first is oxoester-bound by the TE domain while the second building block is thioestered to the Ppant arm of the T domain. Enolate formation of the first activated building block is followed by aldol-like condensation with the T domain bound intermediate. Cyclisation results in the T domain bound furanone core intermediate which is transferred to the TE domain. Next, the TE domain hydrolytically releases **5** which undergoes decarboxylation by XfsB. Finally, XfsC methylates xenofuranone B to xenofuranone A (Fig. 28).

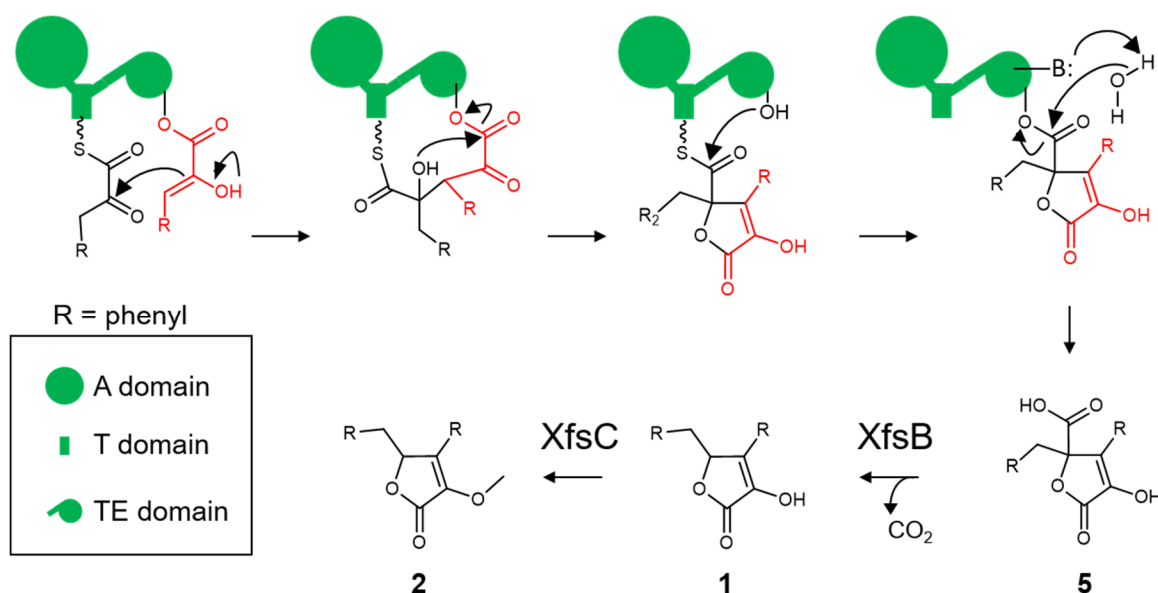


Figure 28. Proposed xenofuranone biosynthesis. NRPS-like mechanism modified from van Dijk *et al.* 2016.^[183] The first accepted phenylpyruvic acid (red) is oxoestered to the thioesterase domain while the second activated phenylpyruvic acid (black) is T-domain bound prior to Claisen-like condensation. The released intermediate **5** is decarboxylated by XfsB. Xenofuranone B (**1**) is methylated by XfsC to yield xenofuranone A (**2**).

3.1.4 4-Hydroxyphenylpyruvic acid incorporating xenofuranone derivatives

Two isomeric xenofuranone derivatives **7** and **8** with the same mass ratio m/z 283.1 $[M+H]^+$ and different retention times were already observed in the induced P_{BAD_xfsA} strain during my master thesis (Fig. 29).^[155] It was assumed that they are both derived from the incorporation of one tyrosine and one phenylalanine, respectively. In this work, reverse feeding of $^{12}C_9$ L-tyrosine to the strain grown in ^{13}C labelled medium was performed. This allowed the distinction of the two isomers. For **8** the incorporation of eight carbon atoms indicates that in this compound 4-hydroxyphenylpyruvic acid is the second activated building block of which the carboxyl group is lost during biosynthesis (Fig. 29). For **7**, the $\Delta m/z$ 9 lighter shift shows that here 4-hydroxyphenylpyruvic acid is the first building block accepted by the NRPS-like enzyme XfsA (Fig. 29).

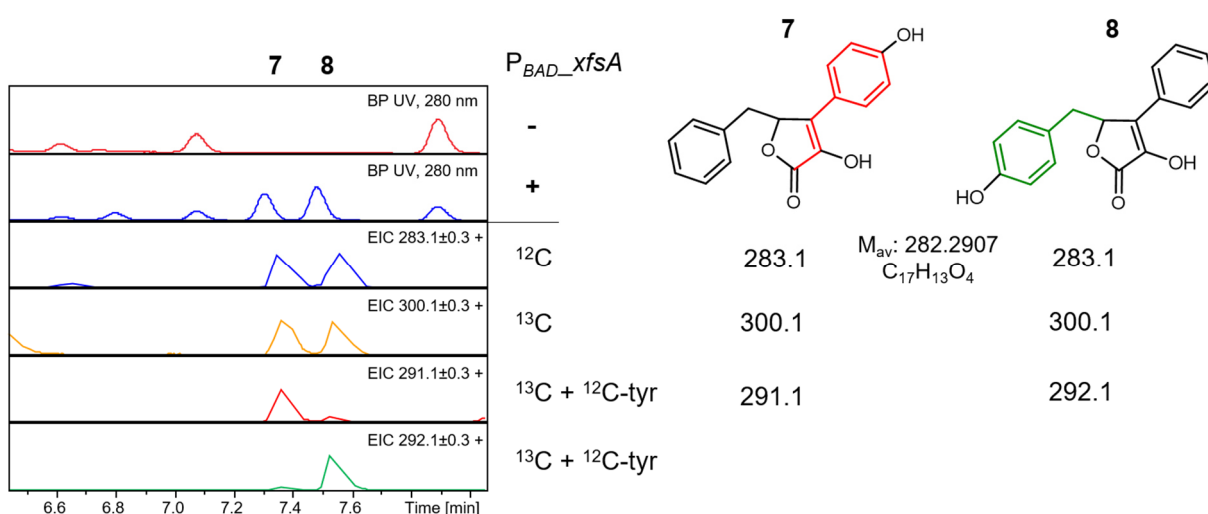


Figure 29. 4-hydroxyphenylpyruvic acid incorporating xenofuranone derivatives **5** and **6**. Reverse feeding of $^{12}C_9$ -L-tyrosine to cultures grown in ^{13}C labelled medium. UV₂₈₀ BPC for arabinose-induced (+) and non-induced (-) *X. szentirmaii* P_{BAD_xfsA} .

Table 15 shows the adapted specificity code for the activation of α -keto acids in NRPS-like enzymes introduced by Wackler *et al.* 2012 for XfsA in comparison with atromentin synthetase GreA, ralfuranone synthetase RalA and TdiA.^[96] The deduced ten letter code for XfsA differs in only one residue, serine (position 9) from the 4-hydroxyphenyl pyruvic acid activating NRPS-like enzyme GreA featuring cysteine in this position.

RESULTS

Table 15. Amino acid specific ten letter code for XfsA adenylation domain. Adapted from Wackler *et al.* 2012.^[96] Numbering (in brackets) is based on GrsA ten letter non-ribosomal code^[60]. GreA was used as a reference. Atromentin synthetase GreA, xenofuranone synthetase XfsA, ralfuranone synthetase RalA, terrequinone A synthetase TdiA. Accepted α -keto acid substrates are shown on the right.

NRPS-like enzyme	α -keto acid activating NRPS code										substrate
	1 (235)	2 (236)	3 (239)	4 (278)	5 (299)	6 (301)	7 (322)	8 (330)	9 (331)	10 (517)	
GreA	V	A	E	F	S	G	G	A	C	K	4-hydroxy-phenylpyruvic acid
XfsA	V	A	E	F	S	G	G	A	S	K	phenylpyruvic acid
RalA	V	A	E	F	S	G	A	A	S	K	phenylpyruvic acid
TdiA	V	A	H	F	T	G	S	A	C	K	indole-3-pyruvic acid

Since XfsA A domain displayed a broadened substrate specificity by activating 4-hydroxyphenylpyruvic acid next to phenylpyruvic acid it was investigated if XfsA might activate *para*-azido-phenylalanine (p -N₃-F) whose azide function can be used for click-labelling.^[184] p -N₃-F was fed to the *X. szentirmaii* P_{BAD_xfsA} strain and the culture extracts were analysed via HPLC-MS. Aside from the before observed xenofuranone compounds **1**, **2**, **7** and **8** no additional derivatives indicating the incorporation of p -N₃-F were observed in the UV₂₈₀ BPC or in the EICs for the expected derivatives with m/z 308.1 [M+H]⁺, m/z 322.1 [M+H]⁺, m/z 324.1 [M+H]⁺, m/z 349.1 [M+H]⁺, m/z 363.1 [M+H]⁺ (Figures 30 and S2).

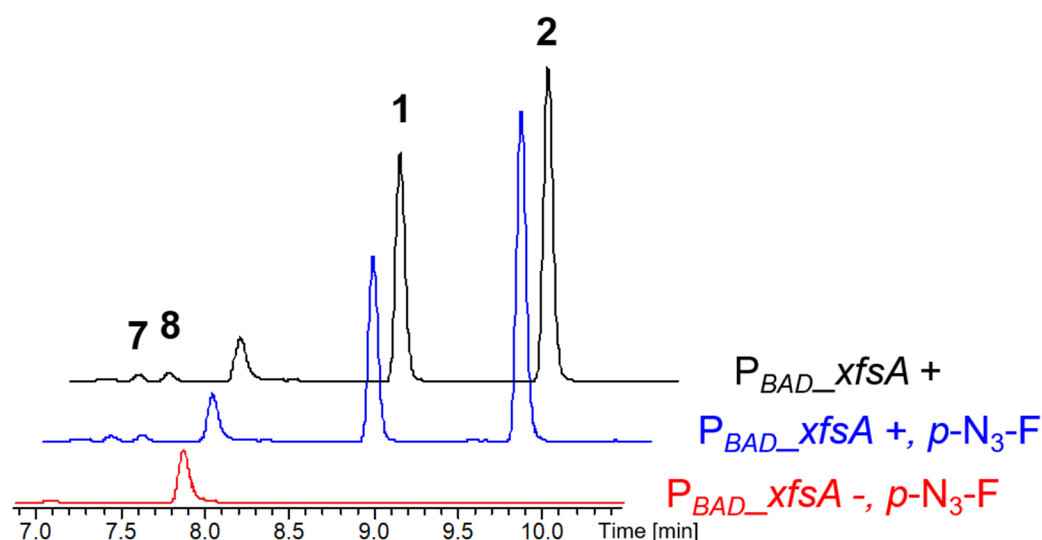


Figure 30. No incorporation of p -azido-phenylalanine (p -N₃-F) in xenofuranones. UV₂₈₀ BPCs from extracts of *X. szentirmaii* P_{BAD_xfsA} cultures that were cultivated in presence of p -N₃-F are shown. + induction with arabinose. – no induction.

3.2 Pyrrolizwilline

3.2.1 Pyrrolizwilline identification

The P_{BAD} promoter was genomically integrated in front of *xhpA* in *X. hominickii* WT by homologous recombination. Induction with 0.2 % L-arabinose leads to the production of compound **9** with m/z 381.2 $[M+H]^+$ and the chemical formula $C_{21}H_{24}N_4O_3$ (Fig. 31). The compound is only produced in the induced *X. hominickii* P_{BAD_xhpA} strain and not in the wild type strain. **9** is only produced in LB medium but not in M63 or Sf900 medium, nor in Isogro® ^{13}C or ^{15}N medium (Fig. S4). Compound purification was conducted to allow structure elucidation. 4 L cultivation with XAD-16 resins yielded 3.3 g dried extract. Preparative HPLC-MS-based pre-fractionation and subsequent semi-preparative HPLC-MS led to 2.3 mg of **9**. NMR analysis and structure elucidation was conducted by my colleague Dr. Yi-Ming Shi (Goethe University) (Table S2). The elucidated structure is depicted in Figure 31 and the compound was named pyrrolizwilline (**9**). Feeding with deuterated d_8 -L-valine confirmed incorporation of one valine (Fig. 31D).

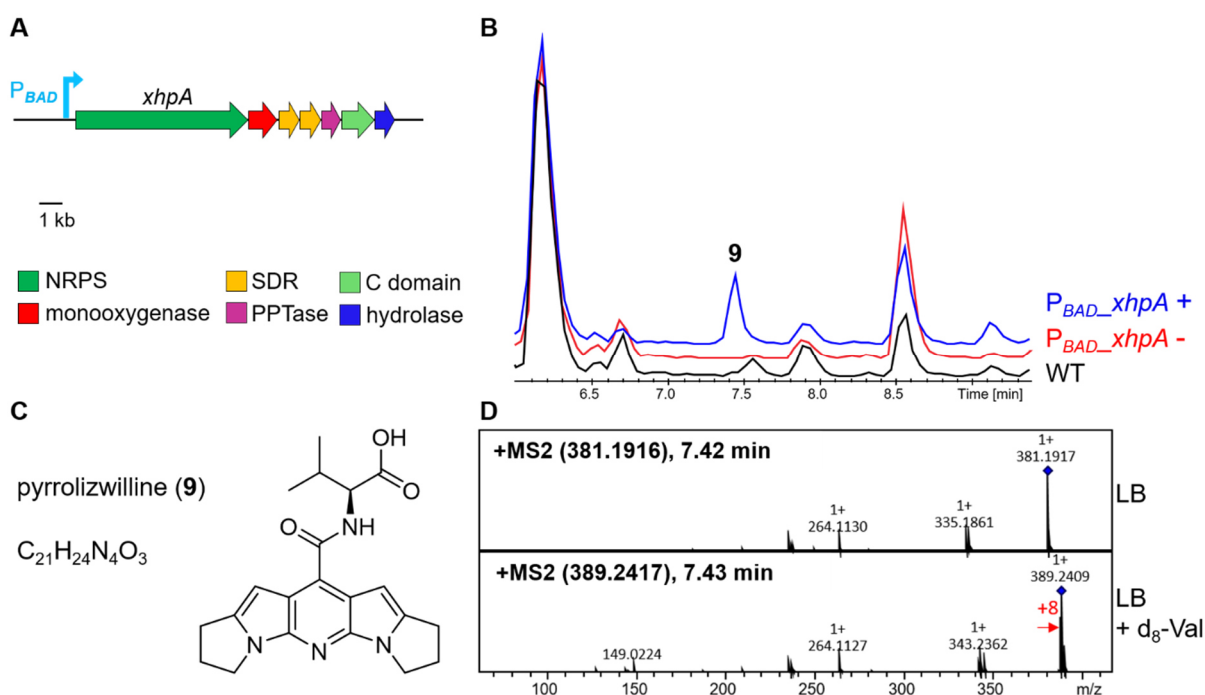


Figure 31. Promoter activation in front of *xhpA* results in pyrrolizwilline (**9**) production. **A** BGC *xhpA*-G. **B** BPC for *X. hominickii* WT, non-induced (-) and arabinose-induced (+) *X. hominickii* P_{BAD_xhpA} . **C** pyrrolizwilline structure. Structure elucidation was conducted by Dr. Yi-Ming Shi (AK Bode, Goethe University) (see Table S2). **D** MS² spectra for **9** in LB and when grown in LB with deuterated d_8 -L-valine.

3.2.2 Pyrrolizidine compounds in *Xenorhabdus*: Pyrrolizwilline and pyrrolizixenamide

Pyrrolizwilline exhibits two pyrrolizidine moieties, hence the name pyrrolizwilline as combination of pyrrolizidine and twin, based on the German translation of twin (“Zwilling”). The pyrrolizidine alkaloids pyrrolizixenamides also exhibit the pyrrolizidine moiety and were

RESULTS

identified from *X. stockiae* previously.^[185] The responsible *pxa* BGC encodes a bimodular NRPS PxaA with C_{starter}-A-T^DC_L-A-T-Te domain organisation followed by a Baeyer-Villiger monoxygenase PxaB (Fig. 32).

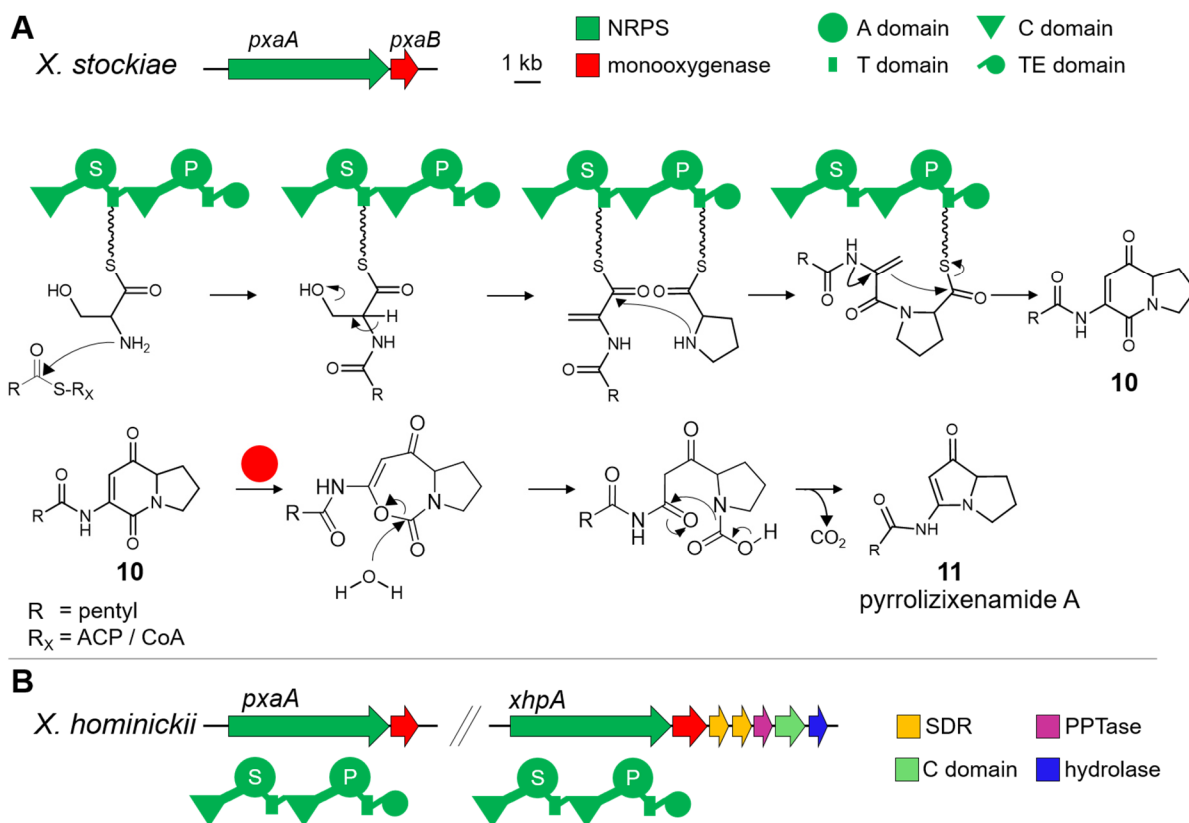


Figure 32. Pyrrrolizixenamide biosynthesis and pyrrrolizixenamide-like BGC organisation **A** Pyrrrolizixenamide biosynthesis. PxaA acylates the first activated amino acid serine, dehydrates serine, condenses proline and releases the 5,6-bicyclic intermediate **10**. Monoxygenase PxaB catalyses Baeyer-Villiger-oxidation followed by hydrolysis and ring contraction to yield pyrrrolizixenamide A (**11**). Modified from Schimming *et al.* 2015^[185] and ^[186]. **B** *X. hominickii* exhibits *pxaAB* and *xhpA-G* with PxaA and XhpA both featuring C_{starter}-A_{ser}-T^DC_L-A_{Pro}-T-Te NRPS domain organisation, each followed by a monoxygenase.

X. hominickii genomic analysis reveals two BGCs that each encodes a bimodular NRPS with this domain organisation followed by a monoxygenase, namely *pxaAB* and *xhpA-F* BGCs. Both NRPSs have the same domain architecture and A domain specificity according to an antiSMASH analysis (Fig.32).^[81] The overall amino acid sequence identity between *X. hominickii* PxaA and XhpA is 44 % and between PxaB and XhpB is 51.6 %. To investigate whether *pxaAB* are required for pyrrrolizwilline biosynthesis, genes *pxaAB* were deleted and the P_{BAD} inserted in front of *xhpA*. Pyrrrolizwilline production was detectable in *X. hominickii* Δ *pxaAB* P_{BAD}*_xhpA* even though in reduced amounts (Fig. 33). The pyrrrolizwilline production is reduced to 43% of the wild type production (Fig. 33).

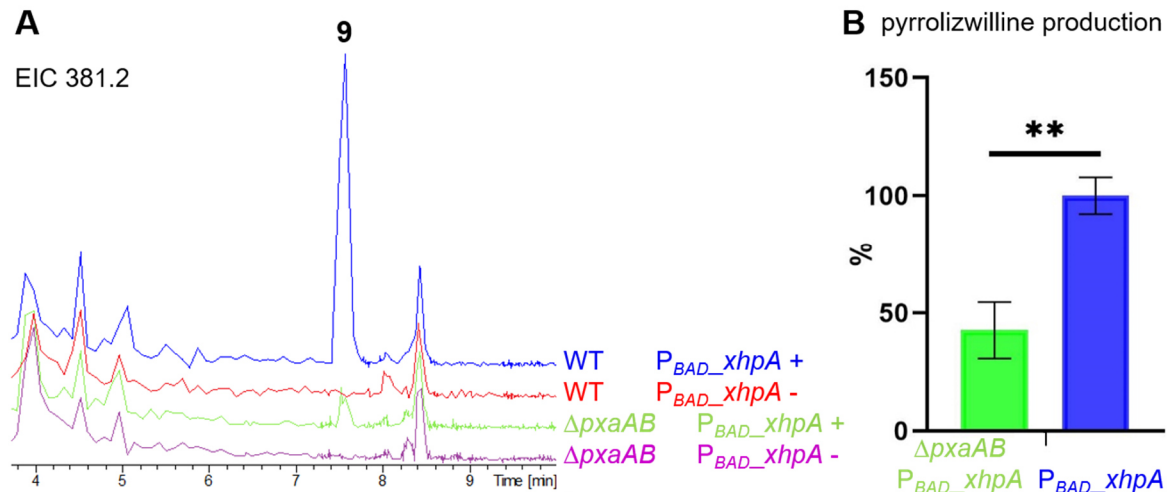


Figure 33. Reduced pyrrolizwilline production in *X. hominickii* $\Delta pxaAB$ P_{BAD_xhpA} (light green) when compared to *X. hominickii* P_{BAD_xhpA} upon induction (blue). **A** EICs for **9** with m/z 381.2 $[M+H]^+$ are shown. + arabinose-induction, - no induction. **B** relative pyrrolizwilline production in *X. hominickii* $\Delta pxaAB$ P_{BAD_xhpA} decreases to 43 % of mean WT level production. Peak areas for **9** were normalised to final OD_{595} . Bars show the mean of three samples and error bars indicate standard deviation. Asterisks mark statistical significance from two-tailed unpaired t-test ($p=0.0022$).

In *X. hominickii*, P_{BAD} transcriptional activation of *pxaA* did not allow pyrrolizixenamide production. Comparison of the LC-MS chromatograms of non-induced and induced *X. hominickii* P_{BAD_pxaA} samples did not show any differences (Fig. 34).

BPC

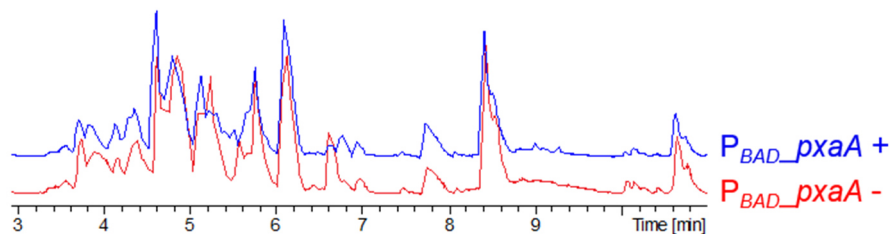


Figure 34. No product detection for promoter activation of *pxa* BGC in *X. hominickii*. No pyrrolizixenamide production is observed in the induced *X. hominickii* P_{BAD_pxaA} strain. BPCs are shown. + arabinose induction (blue), - no induction (red).

RESULTS

3.2.3 XhpAB

3.2.3.1 Heterologous expression in *E. coli*

For further investigation of the pyrrolizwilline biosynthesis, NRPS and monooxygenase encoding genes were heterologously expressed in *E. coli*. Plasmid-based expression of *xhpA* alone resulted in production of **12** (Fig. 35). The observed m/z 167.1 $[M+H]^+$ fragment ion indicates indolizidine core formation from serine and proline as observed in pyrrolizinenamide biosynthesis (Fig. 32). Expression of *xhpA* and *xhpB* together results in production of **13** with the pyrrolizidine characteristic fragment ion m/z 139.1 $[M+H]^+$ (Fig. 35). Both **12** and **13** show a C_3 acyl side chain.

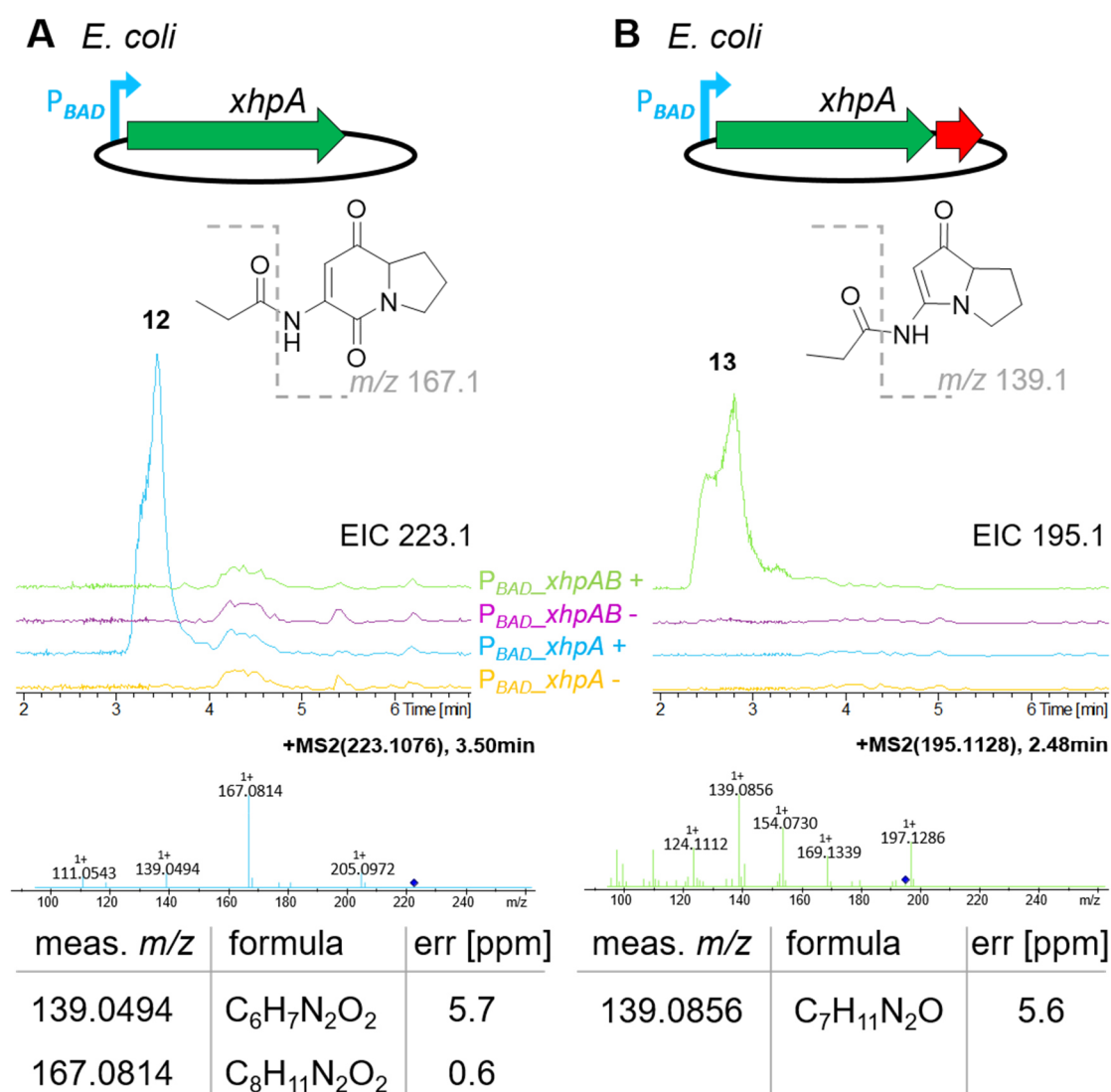
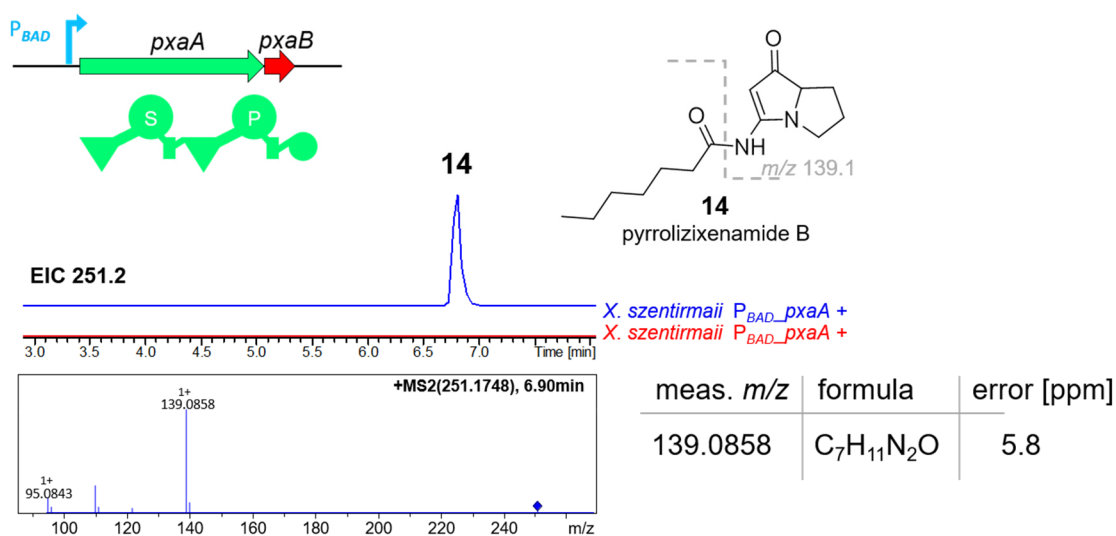


Figure 35. Heterologous expression of *xhpA(B)* in *E. coli*. **A** Expression of *xhpA* results in 5,6-bicyclic product **12** with characteristic fragment ion m/z 167.1 $[M+H]^+$ **B** Expression of *xhpAB* results in pyrrolizinenamide compound **13** with characteristic fragment ion with m/z 139.1 $[M+H]^+$. Grey dashed lines indicate characteristic fragment ions and the respective chemical formula is indicated for selected fragment ions.

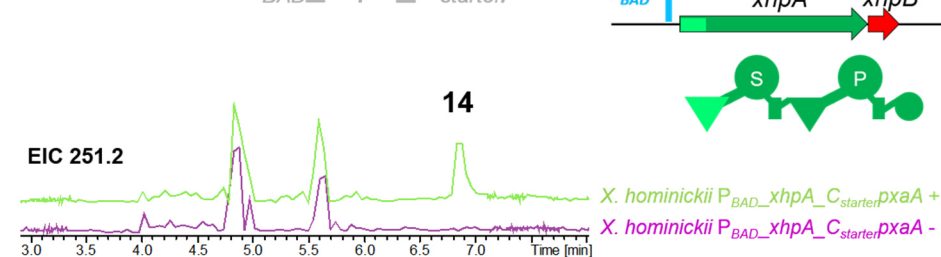
In contrast to the heterologous expression of *xhpA* alone, deletion of *xhpB-G* in *X. hominickii* and subsequent P_{BAD} -based activation of *xhpA* results in the production of **15** with m/z 223.1 $[M+H]^+$ and the predicted sum formula $C_{11}H_{15}N_2O_3$ (Fig. 36C). **15** does not show the expected fragment ion of m/z 167.1 $[M+H]^+$ but one with m/z 181.1 $[M+H]^+$. The increased mass difference of $\Delta m/z$ 14 might hint at an additional methylation of the 5,6-bicyclic core suggesting the incorporation of threonine instead of serine as observed during legonmycin biosynthesis.^[187] Another difference next to the first activated amino acid is observed for the acyl side chain incorporation. In *E. coli*, incorporation of a C_3 acyl side chain is observed for XhpA while in *X. hominickii* $\Delta xhpB-G$ the production of **15** hints at a C_2 acyl side chain. To further investigate the incorporated length of acyl side chains, $C_{starter}$ domain hybrids were generated between the pyrrolizixenamide producing *pxa* BGC in *X. szentirmaii* and pyrrolizwilline producing *xhp* BGC in *X. hominickii*. The fusion site in the C/A-linker region originally established for XU exchange units was used as fusion point and the hybrids were generated by homologous recombination in combination with simultaneous introduction of the P_{BAD} cassette.^[174] Thus the engineered NRPSs were investigated in their *Xenorhabdus* host under the control of P_{BAD} . In the modified *X. szentirmaii* encoding for PxaA with XhpA_ $C_{starter}$ domain, induction leads to production of **15**. Here, no MS² spectra were recorded but the signal shares the same m/z and t_R with **15** detected in *X. hominickii* $\Delta xhpB-G$ P_{BAD_xhpA} (Fig. 36). In *X. hominickii* $P_{BAD_xhpA_C_{starter_pxaA}}$, pyrrolizixenamide B (**14**) was produced upon P_{BAD} activation upstream of *xhpA* encoding the exchanged $C_{starter}$ domain of PxaA. **14** is also detected in *X. szentirmaii* P_{BAD_pxaA} and exhibits the pyrrolizidine characteristic fragment ion with m/z 139.1 $[M+H]^+$ (Fig. 36).

RESULTS

A *X. szentirmaii* P_{BAD}_pxaA



B *X. hominickii* P_{BAD}_xhpA_C_{starter}pxaA



C *X. hominickii* ΔxhpB-G P_{BAD}_xhpA

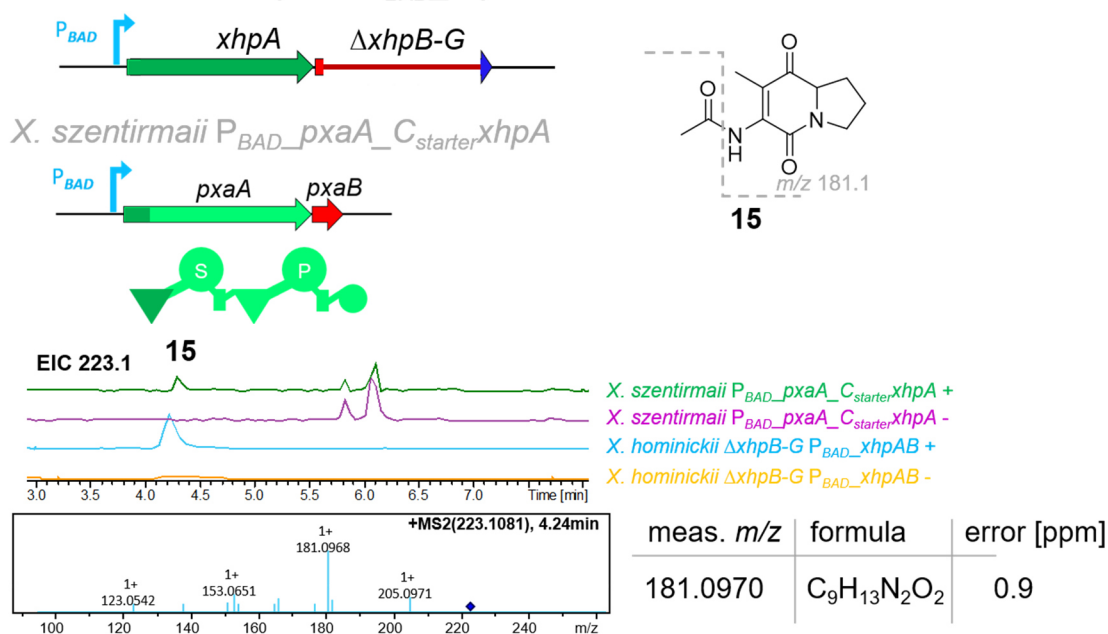


Figure 36. C_{starter} domain exchanges between P_{BAD} controlled pyrrolizixenamide BGC (light green) from *X. szentirmaii* and pyrrolizixenamide BGC (dark green) from *X. hominickii*. Pyrrolizixenamide B (**14**) production in induced *X. szentirmaii* P_{BAD}_pxaA (**A**) and *X. hominickii* P_{BAD}_xhpA with PxaA_C_{starter} domain (**B**). **C** Production of **15** in *X. szentirmaii* P_{BAD}_pxaA XhpA_C_{starter} domain and in *X. hominickii* ΔxhpB-G P_{BAD}_xhpA. + arabinose induction, - no induction, dashed lines indicate MS² fragmentation for the respective characteristic fragment ion.

3.2.4 P_{BAD_xhpA} activation in Δhfq background

For further analysis of pyrrolizwilline biosynthesis and possible facilitated compound identification, activation of the pyrrolizwilline BGC was performed in *X. hominickii* Δhfq . Promoter activation in front of *xhpA* in Δhfq did not result in pyrrolizwilline production (Fig. 37).

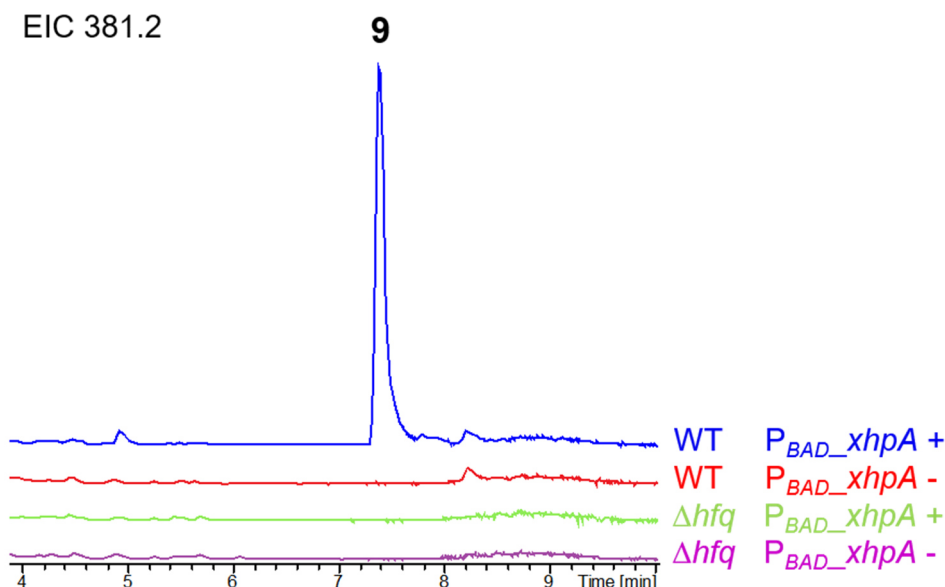


Figure 37. No pyrrolizwilline production in *X. hominickii* Δhfq P_{BAD_xhpA} upon induction. Pyrrolizwilline production is only observed in *X. hominickii* P_{BAD_xhpA} . + arabinose-induction, - no induction.

3.2.5 SDR1 XhpC and SDR2 XhpD

Within *xhp* BGC, *xhpC* and *xhpD* each encode an SDR family NAD(P) oxidoreductase. XhpC and XhpD share 68.2 % sequence similarity and antiSMASH predicts XhpC to be an inactive KR domain with C1 type stereochemistry.^[176] No deletion mutants could be obtained for either the single SDR genes or simultaneous deletion of both genes.

3.2.6 PPTase XhpE

Next to the two reductases, pyrrolizwilline BGC encodes a hypothetical protein, XhpE, according to antiSMASH.^[176] BlastP analysis reveals no conserved domains but similarities to protein sequences annotated as PPTases. XhpE shares 29.57 % identity (query cover 69 %) with protein annotated as PPTase from *V. mimicus* VM223. Overall, an approx. 150 aa C-terminal region of XhpE did not align with the 22 subject sequences identified in the BlastP analysis (Fig. S5).^[170] Deletion of *xhpE* still shows pyrrolizwilline production while *ngrA* deletion abolished pyrrolizwilline production in the arabinose induced *X. hominickii* P_{BAD_xhpA} . Although pyrrolizwilline production is detected in the *X. hominickii* $\Delta xhpE$ P_{BAD_xhpA} strain, the production is reduced to 28% of the production in *X. hominickii* P_{BAD_xhpA} (Fig. 38).

RESULTS

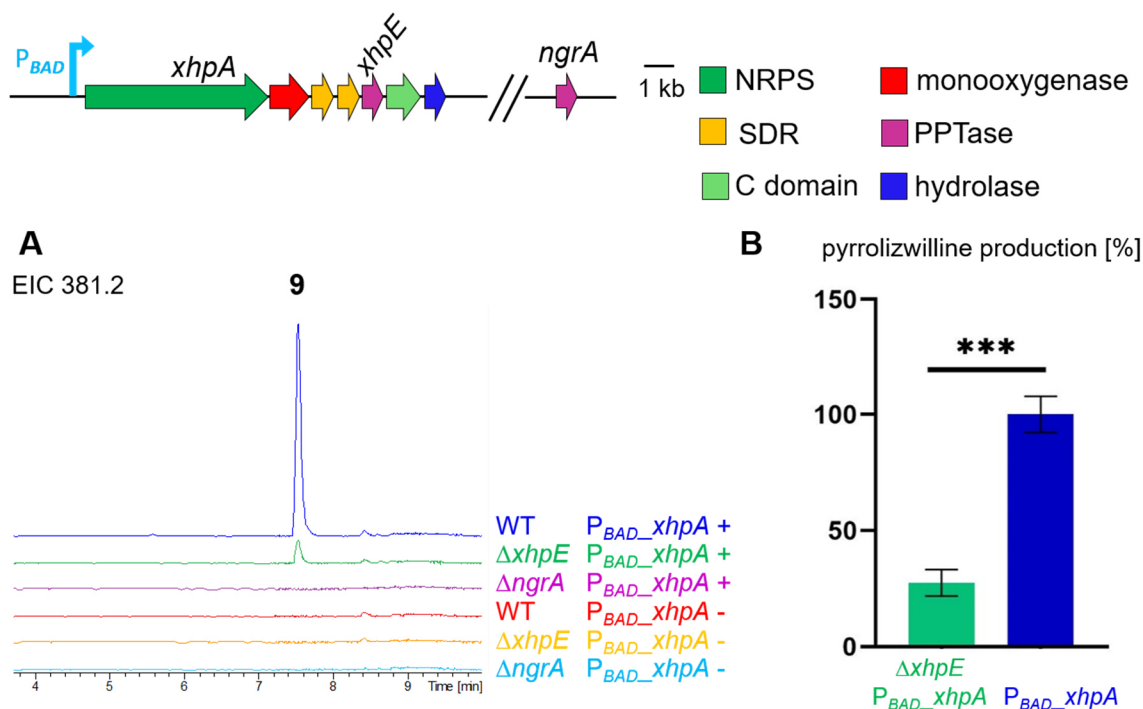


Figure 38. Reduced pyrrolizwilline production in *xhpE* deletion. Pyrrolizwilline is produced in reduced amounts upon activation of P_{BAD_xhpA} in *X. hominickii* $\Delta xhpE$ and not produced in *X. hominickii* $\Delta ngrA$. **A** EICs for **9** show production in *X. hominickii* $\Delta xhpE$ P_{BAD_xhpA} and in *X. hominickii* P_{BAD_xhpA} upon induction. **B** pyrrolizwilline production in *X. hominickii* $\Delta xhpE$ is 28 % of the mean WT production level. Peak areas for **9** were normalised to final OD₅₉₅. Bars show the mean of three samples and error bars indicate standard deviation. Asterisks mark statistical significance from two-tailed unpaired t-test ($p=0.0002$). + arabinose-induction, - no induction.

3.2.7 Condensation domain XhpF

XhpF is predicted to encode a free-standing ^LC_L enzyme according to antiSMASH and Natural product domain seeker (NaPDoS) analysis.^[81,176] When *xhpF* is deleted, no pyrrolizwilline production can be detected. Comparison of extracts from non-induced and induced *X. hominickii* $\Delta xhpE$ P_{BAD_xhpA} revealed no differences (Fig. 39).

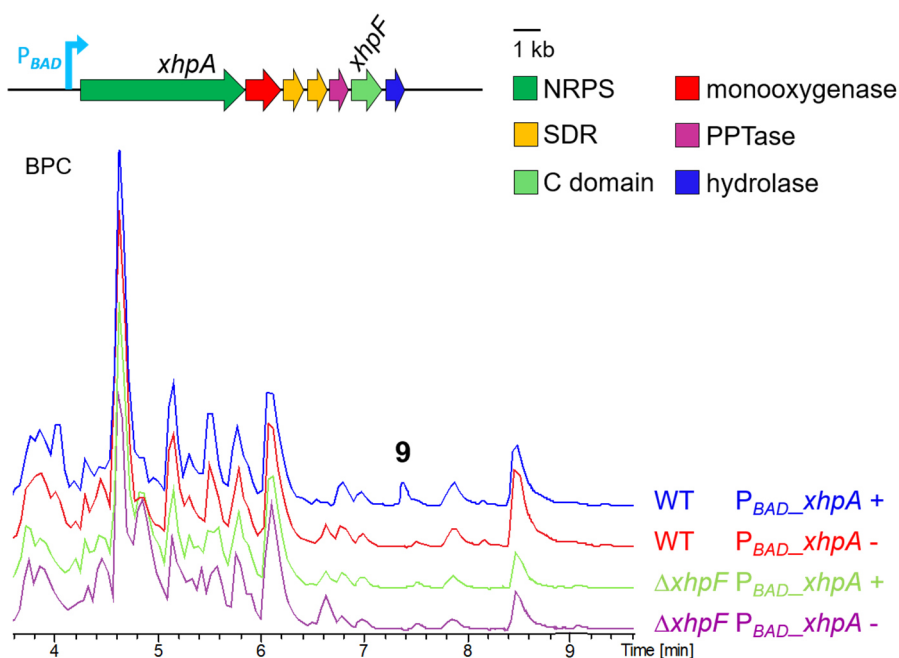


Figure 39. No *xhpA* related compound production in *xhpF* deletion. *X. hominickii* $\Delta xhpF$ P_{BAD_xhpA} shows no pyrrolizwilline (**9**) production upon induction. Comparison between non-induced and induced *X. hominickii* $\Delta xhpF$ P_{BAD_xhpA} shows no difference in the chromatograms. BPCs are shown. + arabinose-induction, - no induction.

3.2.8 α/β -Hydrolase XhpG

The terminal gene *xhpG* encodes a putative α/β -hydrolase according to an antiSMASH analysis.^[81] Promoter activation in front of *xhpA* in an *xhpG* deletion strain leads to production of **16** with m/z 181.1 $[M+H]^+$ and predicted chemical formula $C_9H_{13}N_2O_3$ (Fig. 40). **16** shows the pyrrolizixenamide characteristic fragment ion with m/z 139.1 $[M+H]^+$ and hence is built from serine, proline and acetic acid. Pyrrolizwilline production is not observed in $\Delta xhpG$ indicating that the hydrolysis is crucial for further biosynthesis steps.

RESULTS

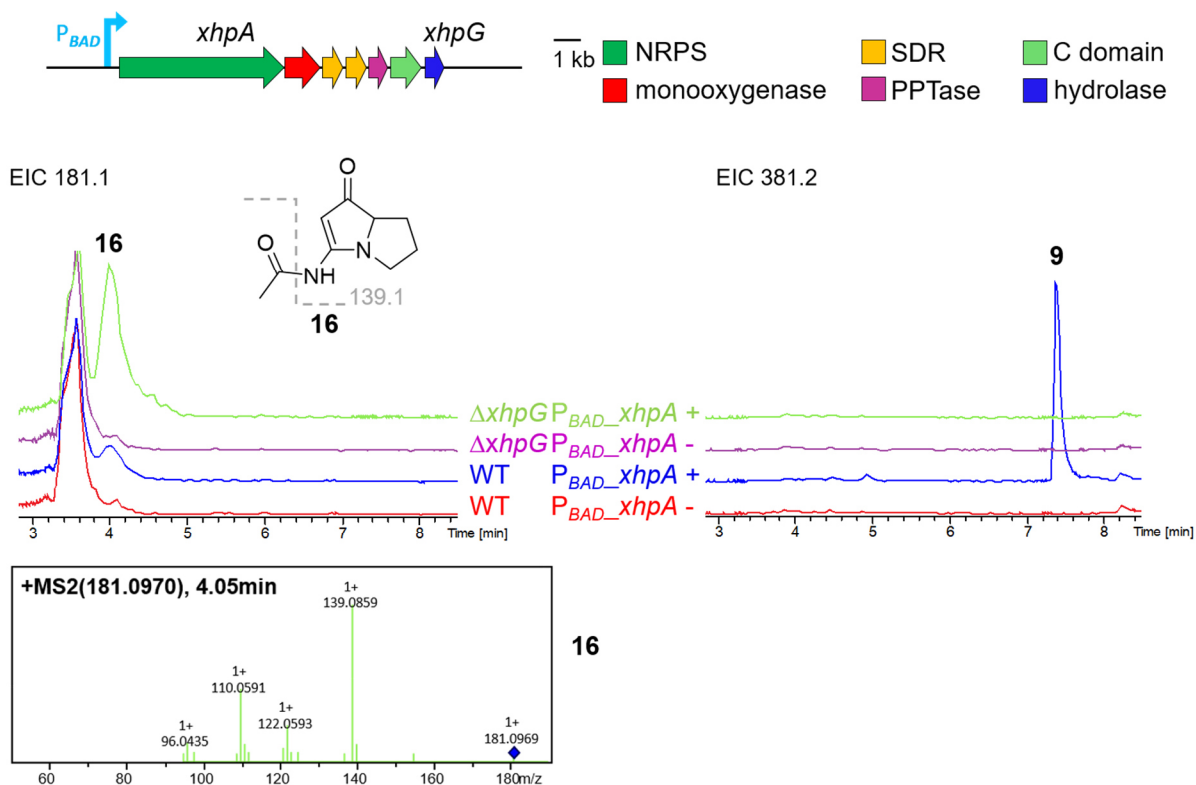


Figure 40. Pyrrolizinenamide-like compound production in *xhpG* deletion. *X. hominickii* $\Delta xhpG$ P_{BAD_xhpA} produces **16** and no pyrrolizwilline (**9**). Structure of **16** with indicated m/z 139.1 $[M+H]^+$ fragment as well as MS² spectrum for **16** is shown. + arabinose induction, - no induction.

3.2.9 Overview of observed intermediates and postulated pyrrolizwilline biosynthesis

Figure 41 summarises the results from the above presented experiments for investigation of pyrrolizwilline biosynthesis.

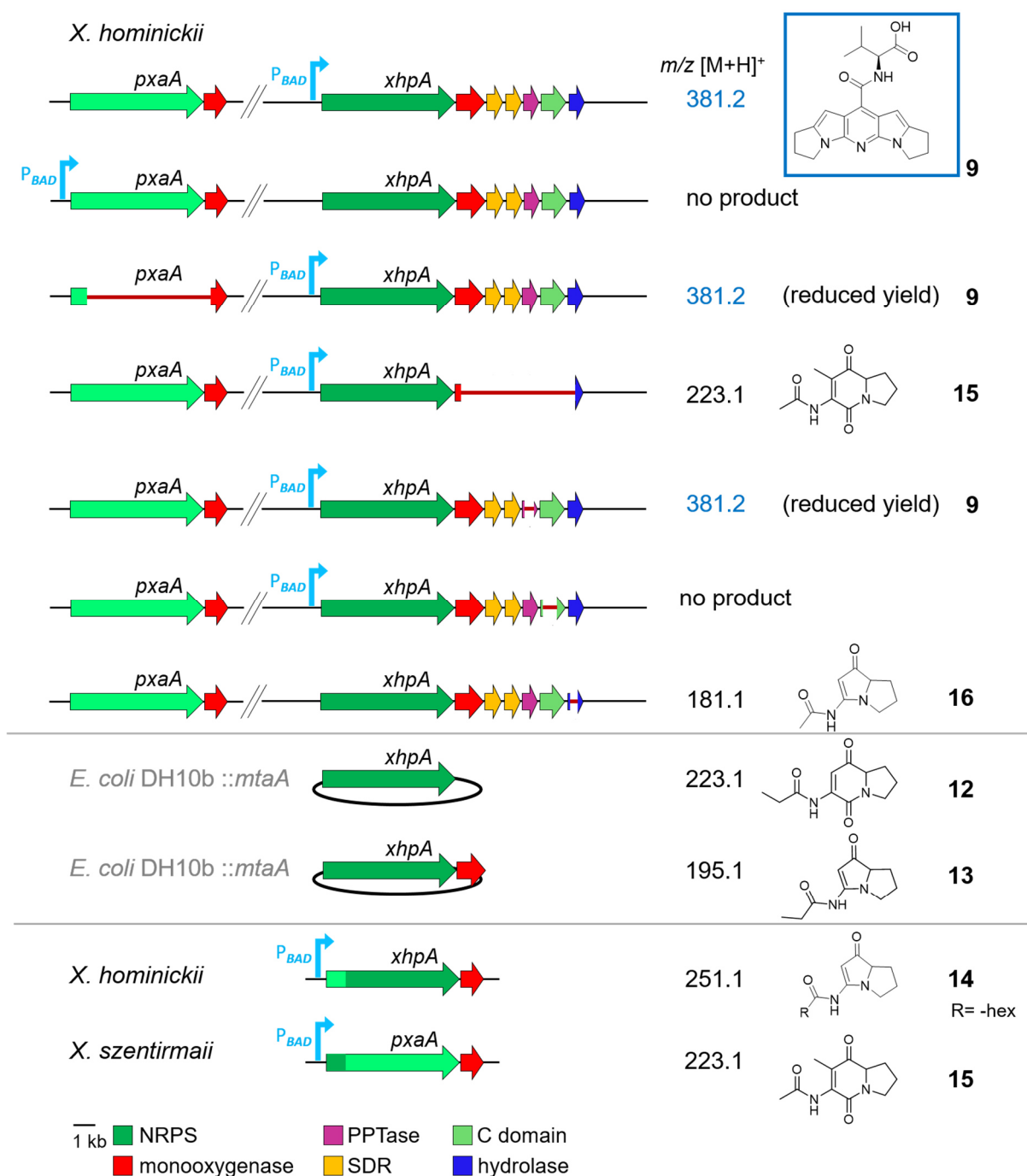


Figure 41. Overview of the constructed mutant strains and observed biosynthesis products. Dark red line indicates deleted region.

Based on the findings, pyrrolizwilline biosynthesis is proposed as follows (Fig. 42). The bimodular NRPS XhpA incorporates serine, proline and an acyl moiety to build a 5,6-bicyclic intermediate that undergoes oxygenation, hydrolysis and ring contraction, similarly to pyrrolizixenamide biosynthesis. Next, XhpG hydrolyses the acyl side chain to allow subsequent reduction of the keto-function to the unsaturated carbon-carbon bond by action of the two reductases. Two of these reduced pyrrolizidine building blocks (**16c**) are required for

RESULTS

pyrrolizwilline biosynthesis. One of the reduced pyrrolizidine cores is extended with glyoxylic acid via nucleophilic attack of the electron rich carbon in **16c** (Fig. 42). Elimination of water generates the carbenium which is attacked by a second pyrrolizidine building block (**16c**) for dimerisation. Deamination and oxidation then yield the six-membered pyridine ring in the core of **18d**. Next, **18d** is activated with coenzyme A, possibly by a CoA ligase. Finally, the free-standing condensation domain XhpF condenses L-valine and **18** to yield pyrrolizwilline (Fig. 42).

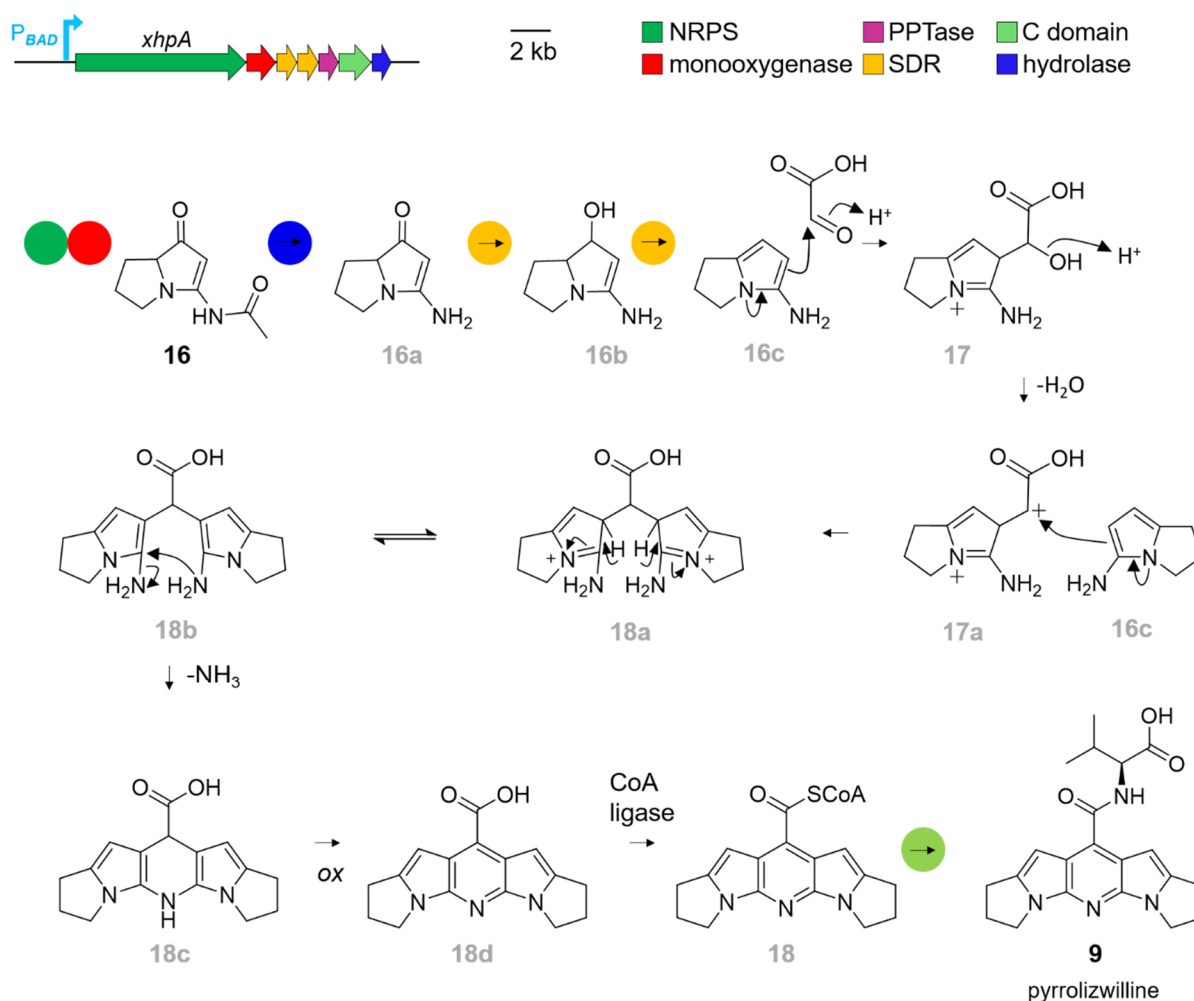


Figure 42. Postulated pyrrolizwilline biosynthesis. NRPS and monoxygenase provide the pyrrolizixenamamide-like compound **16**.^[185] XhpG hydrolyses the acyl moiety followed by the two subsequent two electron reductions of the carbonyl function by XhpC and XhpD. The pyrrolizidine moiety is extended with glyoxylic acid and dimerises with a second pyrrolizidine moiety to form **18a**. Deamination and ring oxidation yield **18d** which is activated with coenzyme A. XhpF catalyses the L-valine condensation step to yield pyrrolizwilline (**9**). Detected compounds are indicated by black numbers and proposed intermediates are indicated by grey numbers.

3.2.10 Phylogenetic analysis of pyrrolizwilline-like BGCs

Clusterblast analysis of *xhpA-G* revealed BGCs from *Brenneria* sp. EniD312, *Vibrio mimicus* VM223 and *Pseudomonas* sp. MIACH. Additionally, BlastP analysis using XhpA as query revealed further BGCs that contain genes encoding NRPSs predicted to show XhpA domain architecture in combination with further biosynthesis enzymes such as monooxygenases and reductases.^[170] Among them are BGCs from Gram-positive and Gram-negative bacteria and even predicted NRPS-PKS hybrids (Fig. 43).

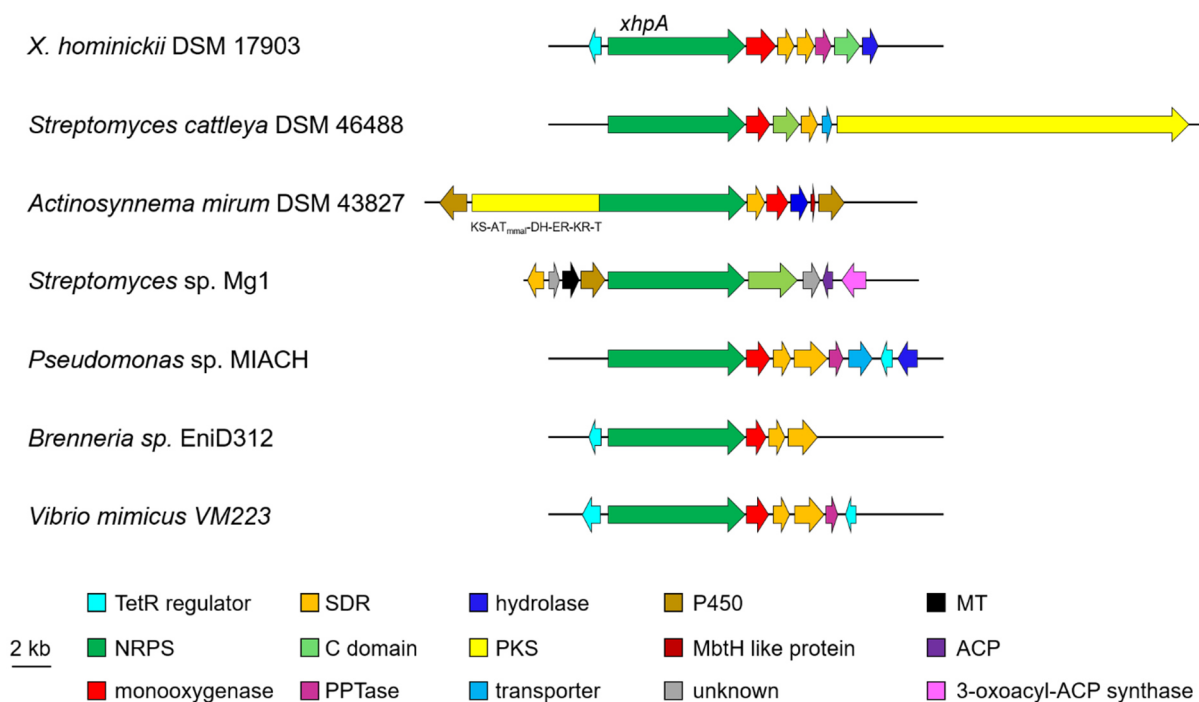


Figure 43. Selected BGCs with an organisation similar to *xhp* BGC. All presented BGCs exhibit XhpA domain architecture. Encoded proteins are indicated according to predictions from antiSMASH analysis.^[176]

Multiple sequence alignment with NRPS sequences from the above identified non-*Xenorhabdus* species as well as PxaA and XhpA sequences from *X. hominickii* and PxaA sequences from pyrrolizixenamide BGCs from *Xenorhabdus* species shows that XhpA does not group with the PxaA sequence from the known pyrrolizixenamide producers *X. szentirmaii* and *X. stockiae*. Dotplot analysis comparing XhpA and *X. szentirmaii* PxaA shows lowest similarity for the region of the C_{starter} domain (Fig. 44). XhpA and *X. szentirmaii* PxaA show an overall 41.3 % sequence identity and 55.1 % sequence identity for A₁, 47.2 % for A₂, 37.7 % sequence identity for ^DC_L and 23.1% identity for C_{starter} domain. The amino acid prediction by antiSMASH and the non-ribosomal code are identical for both A₁ and A₂ domain for XhpA and *X. szentirmaii* PxaA.^[176]

RESULTS

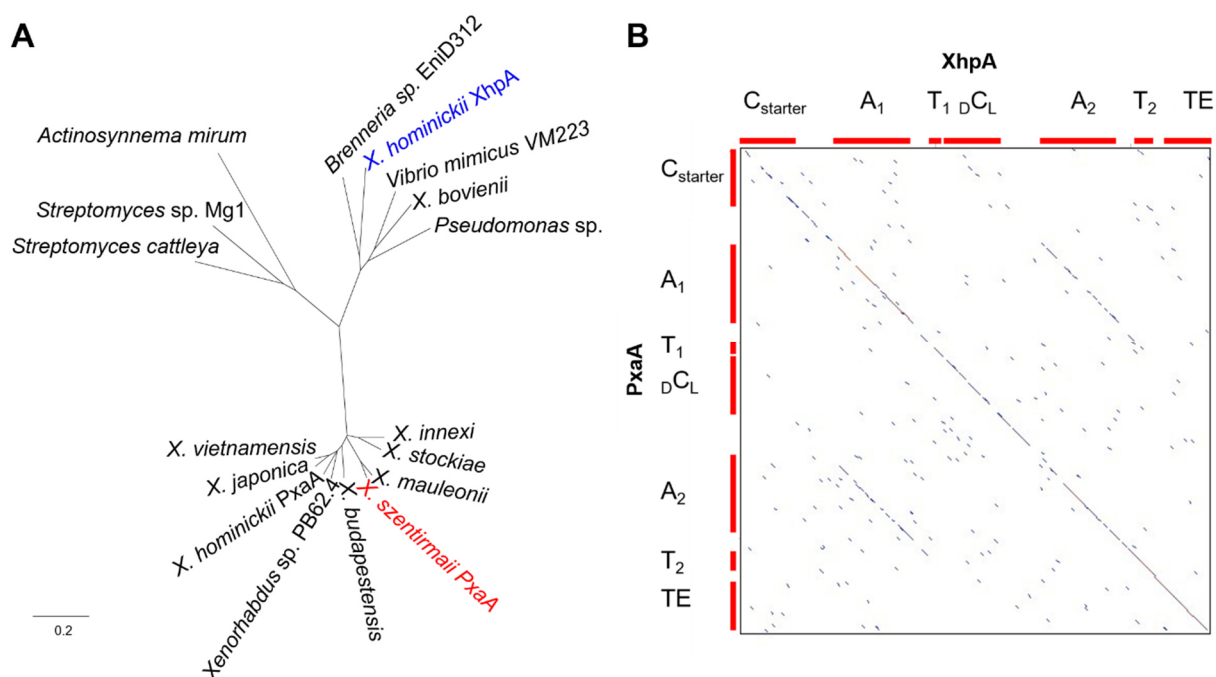


Figure 44. XhpA *in silico* analysis. **A** Unrooted tree for multiple sequence alignment of XhpA-like enzymes (see Fig. 43) results in three phylogenetic groups: NRPSs from Gram-positive bacteria (top left). Group that includes pyrrolizwilline synthesising NRPS XhpA (blue) (top right). Group that includes pyrrolizixenamide producing NRPS PxaA (red) (bottom). **B** Dot plot for *X. hominickii* XhpA and *X. szentirmaii* PxaA. Domain regions are indicated in red. Sequence alignments, tree and dot plot analysis were generated using Geneious v 6.1.8.

P_{BAD} activation in front of *pxaA* in *X. bovienii* resulted in production of pyrrolizixenamide A (**11**) with m/z 237.1 $[M+H]^+$ showing the characteristic m/z 139 $[M+H]^+$ fragment ion and exhibiting a C₆-acyl side chain (Fig. 45).

X. bovienii

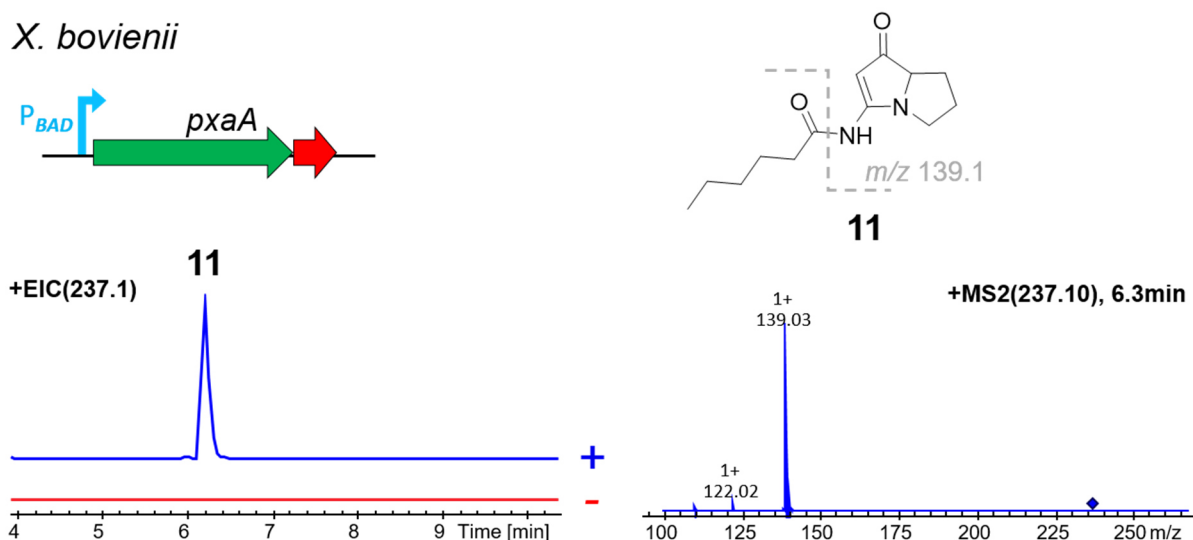


Figure 45. Pyrrolizixenamide A (**11**) production in *X. bovienii* P_{BAD_pxaA} . MS² fragmentation is shown for **11** with pyrrolizidine core fragment indicated by grey dashed line. + arabinose induction, - no induction.

3.2.11 TetR transcriptional regulator

The via clusterblast identified BGCs in *Brenneria* sp. EniD312, *Vibrio mimicus* VM223 and *Pseudomonas* sp. MIACH and the pyrrolizwilline BGC all show presence of a TetR family transcriptional regulator (TFR) encoding gene (Fig. 45). Deletion of the TetR transcriptional regulator encoding gene *xhom00317* did not show production of pyrrolizwilline and no differences in the LC-MS chromatogram when compared with the wild type strain (Fig. 46). The BPCs for plasmid-based expression of the TFR (pP_{BAD}*xhom00317*) in *X. hominickii* show no difference between induced and non-induced culture.

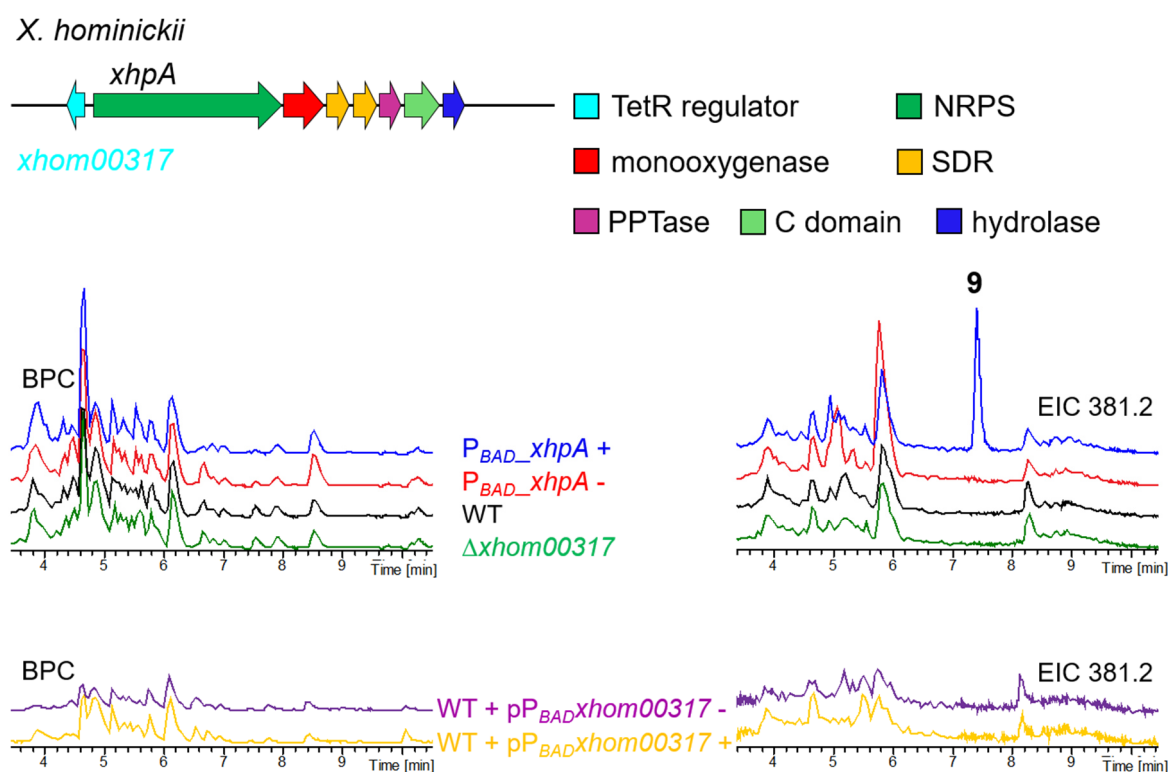


Figure 46. Deletion of TetR encoding gene *xhom00317* (TF) does not affect pyrrolizwilline (**9**) production. Plasmid-based expression of TetR shows no difference in induced and non-induced LC-MS sample (WT+pP_{BAD}TF). + arabinose induction, - no induction. BPCs and EICs for *m/z* 381.2 [M+H]⁺ are shown. TF stands for TetR transcriptional regulator encoded by *xhom00317*.

3.2.12 Bioactivity testing

For simple bioactivity screens, extracts were tested in a growth inhibition assay against Gram positive *Micrococcus luteus*, Gram-positive spore forming *Bacillus subtilis*, Gram-negative *E. coli* DH10b and yeast *Saccharomyces cerevisiae*. Kanamycin was used as positive control and methanol as negative solvent control. Growth inhibition was not observed for *E. coli* or *S. cerevisiae*. Observed growth inhibition for *M. luteus* and *B. subtilis* showed no difference between non-induced and arabinose induced *X. hominickii* P_{BAD_xhpA} sample (Fig. 47).

RESULTS

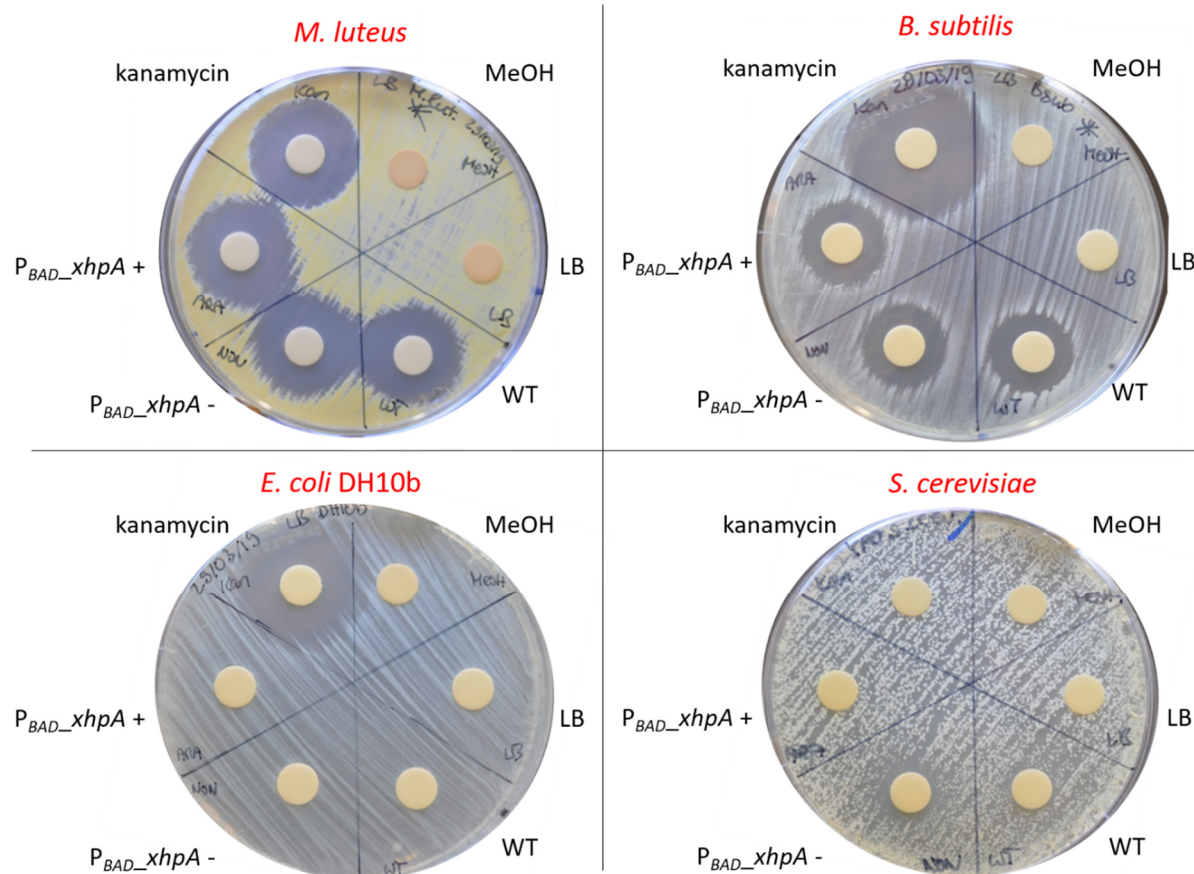


Figure 47. Pyrrolizwilline bioactivity assay. Growth inhibition assay with culture extracts of arabinose induced (+) and non-induced (-) *X. hominickii* P_{BAD_xhpA} against *M. luteus*, *B. subtilis*, *E. coli* DH10b and *S. cerevisiae*. Positive control: kanamycin, negative controls: LB medium, *X. hominickii* WT methanol (solvent control).

3.3 Glyoxpeptide

3.3.1 Glycerol adduct identification

In addition to the purified compound, glyoxpeptide **3** with m/z 292.1 $[M+H]^+$, another signal (**19**) with m/z 366.1 $[M+H]^+$ and the chemical formula $C_{11}H_{21}N_3O_6$ was observed in previous work upon P_{BAD} activation of *xgsA* in *X. hominickii* Δhfg . The chemical formula deviation $C_3H_6O_2$ between **3** and **19** hinted towards glycerol as glycerol is one main carbon source in the used culture medium, Xeno/Photo production medium (XPPM). Leaving glycerol out of the medium, no production of **19** is detected (Fig. 48B). Increasing the glycerol amount in the medium, the production levels of **3** and **19** approach a 1:1 ratio (Fig. 48B). In ^{13}C labelled medium no production of **19** is observed but can be restored by addition of non-labelled $^{12}C_3$ -glycerol (Fig. 48A). The detected signal with m/z 377.2 $[M+H]^+$ is explained by the incorporation of eleven ^{13}C carbon from production in ^{13}C labelled medium and introduction of three ^{12}C carbons from the added $^{12}C_3$ -glycerol (Fig 48). Comparison of MS² spectra and the b-ion fragmentation supports that the glycerol is attached at the tripeptide's C-terminal end (Fig. 48C).

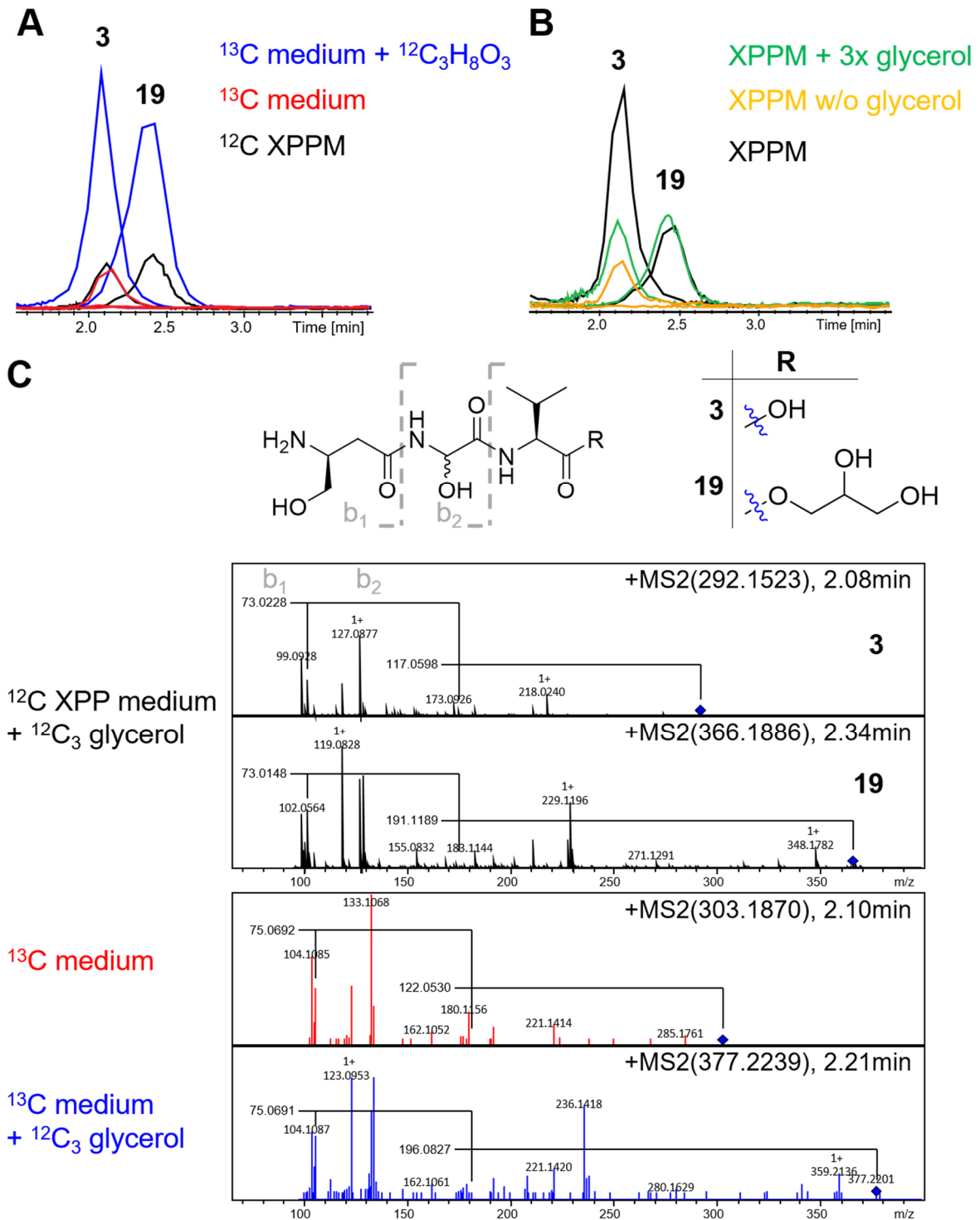


Figure 48. Glyoxpeptide (**3**) and its glycerol adduct (**19**) are observed in the induced *X. hominickii* Δhfg P_{BAD_xgsA} in XPPM (containing glycerol). **A** Feeding of ^{12}C -glycerol to ^{13}C labelled Isogro® medium. **B** **19** is not observed in XPPM without glycerol. Positive EICs for m/z 292.1, m/z 366.1, m/z 380.1 and m/z 377.1 are shown. **C** Schematic representation for observed b-ions in MS² spectra.

RESULTS

3.3.2 MonoOx domain characterisation

The glycine-specific A domain in XgsC harbours a monooxygenase (MonoOx) domain, assumed to be responsible for the α -hydroxylation of glycine.^[153] Sequence alignment of XgsC_A2 domain with well characterised phenylalanine-activating A domain GrsA (PDB: 1amu) revealed that the MonoOx domain is embedded in between the A8 and A9 core motifs of the XgsC_A2 domain (Fig. 49).^[59,60] Based on this sequence alignment, a 1035 nt long region was defined to encode the MonoOx domain (345 aa) and this complete region was deleted (Fig. 49).



Figure 49. Sequence alignment of XgsC_A2 with GrsA. The MonoOx domain is located within the XgsC_A2 domain in between core motifs A8 and A9. Deleted sequence in Δ MonoOx is shown in blue.

Subsequent activation of P_{BAD_xgsA} in *X. hominickii* Δ MonoOx allowed identification of the non-hydroxylated tripeptide **20** (Fig 50). The cultivation for samples shown in Figure 50 was conducted in XPP medium without glycerol. However, when cultivated in presence of glycerol, the non-hydroxylated glyoxlpeptide **20** as well as the non-hydroxylated glyoxpeptide glycerol derivative **21** are observed (Fig. S9, Fig. S10).

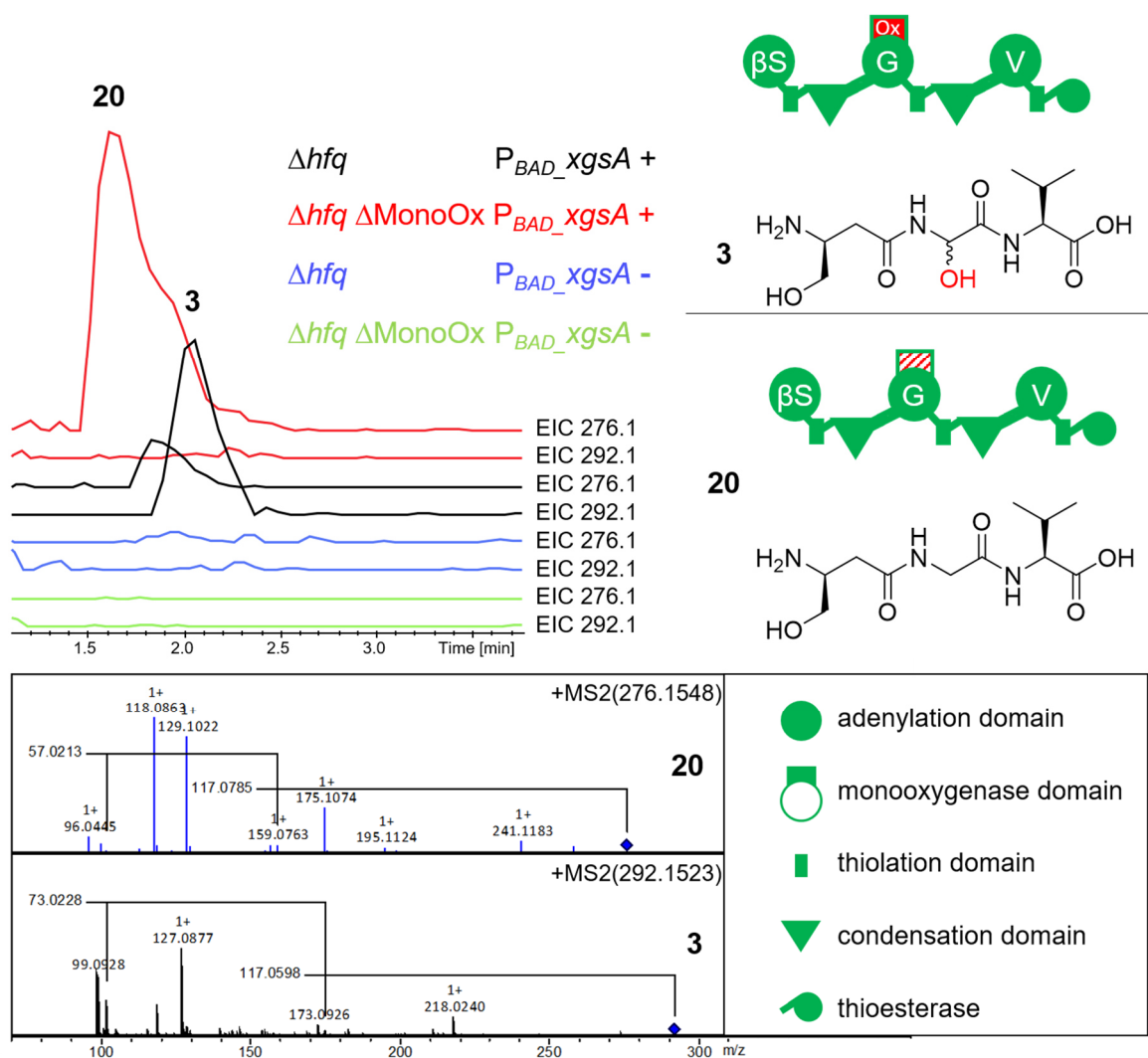


Figure 50. Production of non-hydroxylated glyoxypeptide derivatives in $\Delta MonoOx$. Removal of MonoOx domain (red) from XgsC_A2 results in the production of the non-hydroxylated compound **20**. B-ion fragmentation is indicated in the MS² spectrum for glyoxypeptide (**3**) and non-hydroxylated derivative **20**, respectively.

Sequence alignment (ClustalW, BLOSUM) with a set of oxidases and monoxygenases groups XgsC_MonoOx domain with the monoxygenase Sky39 and MonoOx domains PuwE, MtaG and MelG (Fig. 51A). The oxidase (Ox) domains in MelD, MtaD and EpoB are integrated in the A domain between core motifs A8 and A9 and the Ox domain in BlmIII is located next to the T domain.^[69,70] Within the phylogenetic group that includes XgsC_MonoOx domain, Sky39 is a lone-standing FAD-dependent monoxygenase, accompanied by reductase Sky40, and is responsible for α -hydroxylation of glycine in skyllamycin biosynthesis.^[188] MtaG and MelG exhibit MonoOx domains embedded in between motifs A4 and A5 of their glycine activating A domains in myxothiazol and melithiazol biosynthesis, respectively.^[69,189] PuwE monoxygenase domain is embedded in PKS-NRPS hybrid enzyme PuwE which is involved in puwainaphycin biosynthesis and exhibits the following domain organisation: the malonyl

RESULTS

elongating KS-A-ACP PKS module is followed by an aminotransferase (AmT) domain and a monooxygenase domain and finally shows a valine-incorporating C-A-T NRPS module (Fig. 51).^[190]

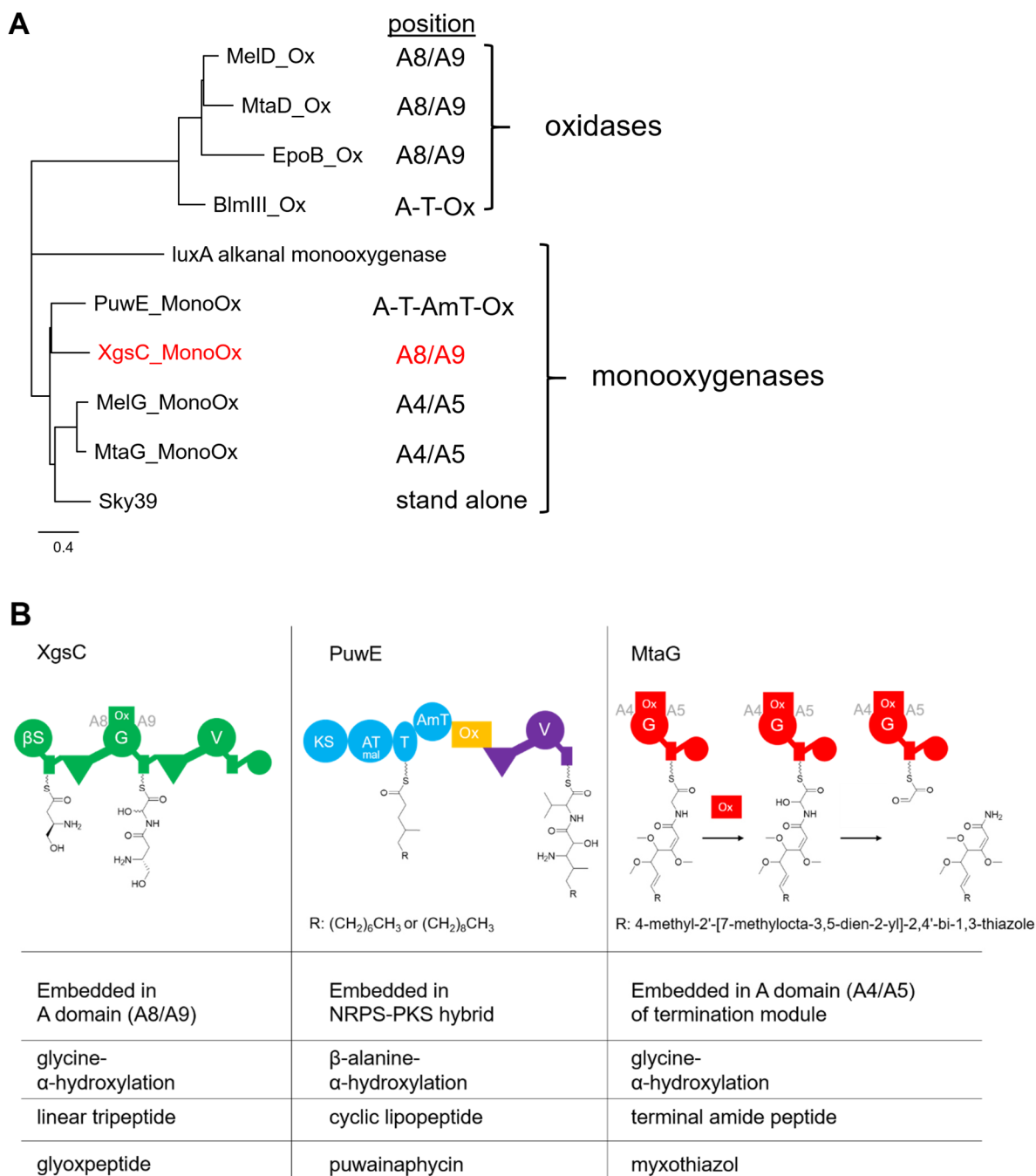


Figure 51. Overview of oxidases and monooxygenases involved in natural product biosynthesis. **A** Multiple sequence alignment groups XgsC_MonoOx domain with known monooxygenases. Alignment is shown in Fig. S12. **B** Comparison of XgsC_MonoOx domain with PuwE_MonoOx domain from puwainaphycin synthesising PKS-NRPS hybrid enzyme and with MtaG_MonoOx domain from myxothiazol synthesising NRPS.

An NCBI Conserved Domain Search for XgsC_MonoOx domain pointed towards flavin utilising monooxygenases super family (TIGR_04020).^[191] Molecular Modelling Database (MMDb) provided structural data for PuwE_MonoOx domain from this conserved domain family (PDB ID: 6KET). The sequence identity between MonoOx domains from PuwE and XgsC is 50.44 % (Swissmodel).^[192] The homology modelling server Swissmodel was used for homology modelling of XgsC_MonoOx domain with PuwE Monoox domain as template. The results are shown in Figure 52. In general, dark blue colouring indicates higher sequence identity.^[192] The local quality estimate gives an overview which parts of the model are more reliable than others. For residues approx. 256-281, which form the predicted α -helices coloured in orange and positioned at the outside, the prediction is less reliable. The structure for the PuwE_MonoOx domain depicts a homodimer and a TIM barrel-like fold.

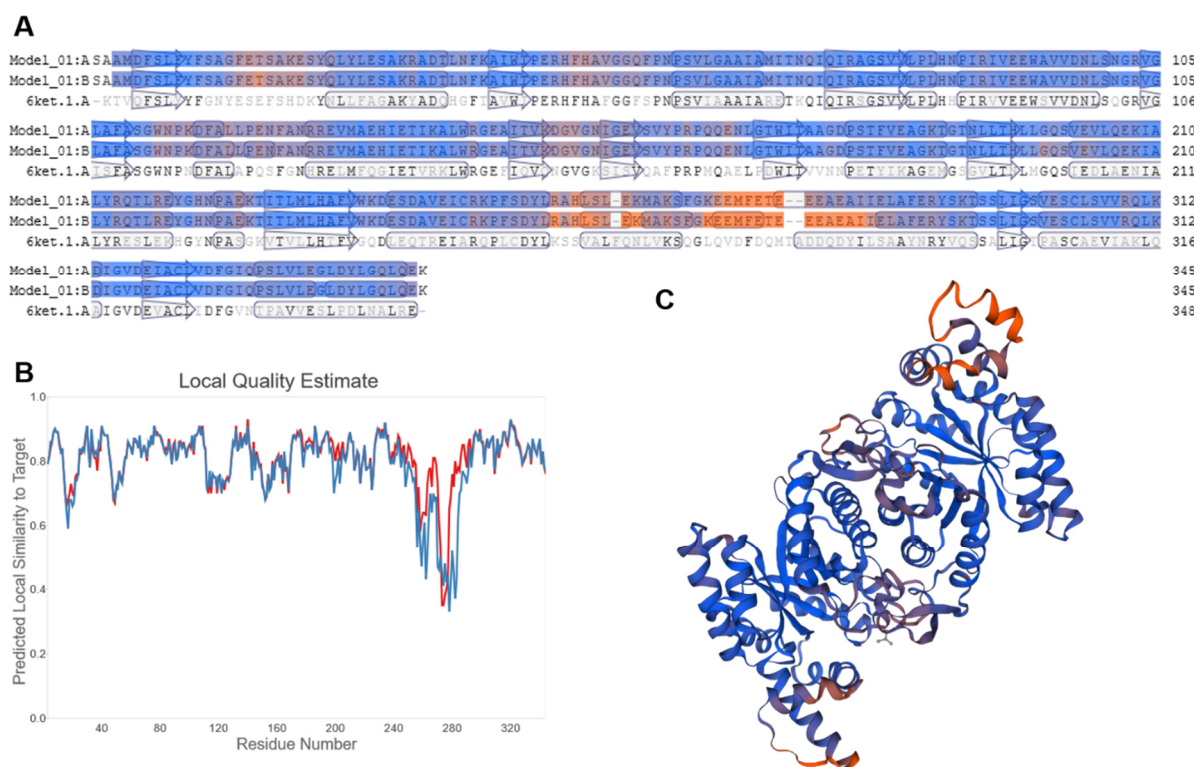


Figure 52. Results from homology modelling using Swiss Model. Sequence analysis of XgsC_MonoOx domain. Blue represents higher QMEAN, orange indicates region of lower QMEAN. **A** sequence alignment of XgsC_MonoOx domain with PuwE_Monoox domain (6ket.1A). Arrows show β -sheets, circles mark α -helices. **B** Local quality estimate plot. **C** Quaternary structure of homology model.

For further MonoOx domain characterisation, the domain was purified with N-terminal His₆tag and small ubiquitin-like modifier (SUMO) propeptide Smt3. Following Ulp1-mediated cleavage of Smt3, the protein was still soluble and was subjected to ITC experiments for flavine cofactor identification. No binding was observed for FAD or FMN (Fig. 53). To provide FAD and FMN in reduced form, 5 mM DTT was added. Moreover, the purified domain fractions show no

RESULTS

colour. A small sample fraction of the purified domain fraction was boiled and subjected to LC-MS analysis to check for presence of FAD or FMN. No signals corresponding to masses for FAD or FMN were detected (Fig. S13).

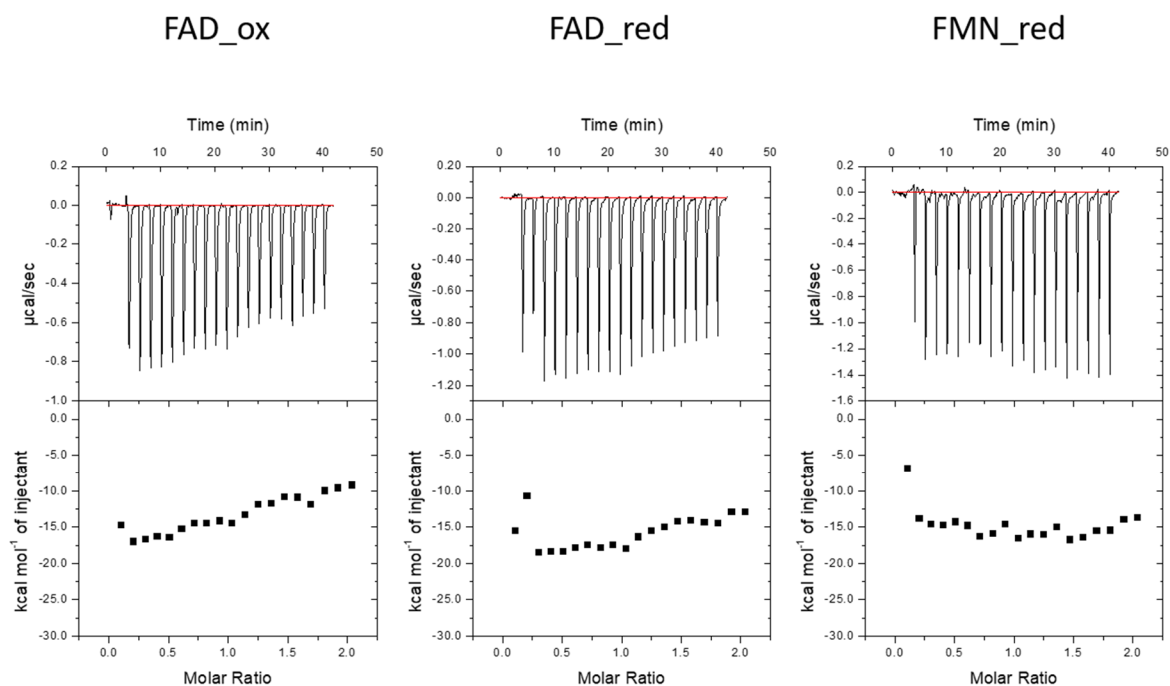


Figure 53. ITC measurements to investigate cofactor binding for XgsC_MonoOx domain. No binding is observed for the titrated cofactors FAD (FAD_ox, FAD_red) and FMN (FMN_red). 5 mM DTT was added for cofactor reduction.

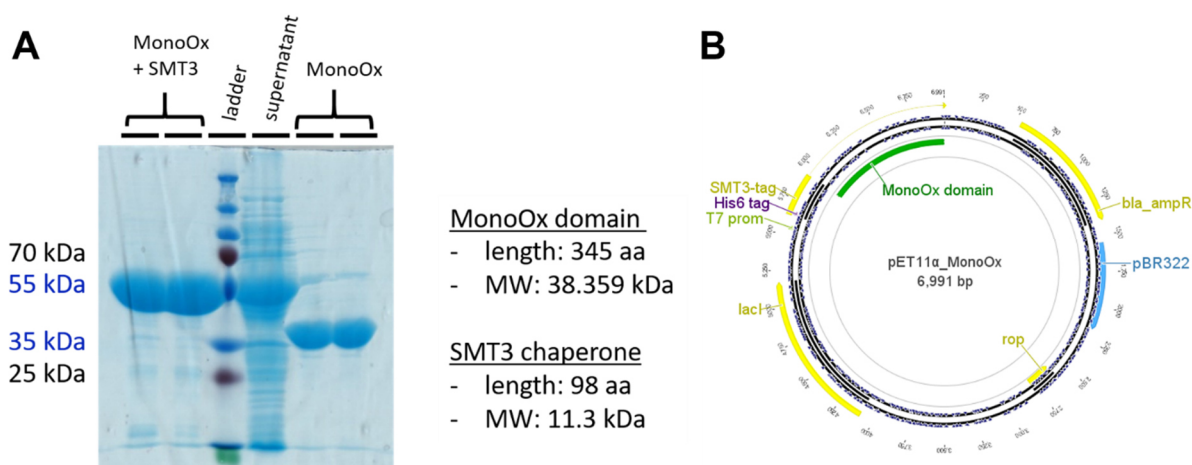


Figure 54. MonoOx domain protein purification. **A** SDS-page before and after Ulp1 mediated cleavage of SMT3. First N-terminal Poly-His₆ tag-based affinity purification, then reverse NiNTA purification was conducted. **B** Plasmid map pET11α_MonoOx for protein expression.

To investigate the possibility of inserting this MonoOx domain into other glycine specific A domains, phototemtidase PttB from *P. temperata* Meg1 strain was chosen. The *ptt*

BGC produces the cyclic lipopeptide phototemtide and plasmid-based expression of this BGC is possible in *E. coli* DH10b::*mtaA*.^[179] The MonoOx domain is positioned in the first A domain of PttB which activates glycine. Comparison of LC-MS chromatograms of extracts from non-induced and induced sample showed no difference (Fig. 55). As control, the non-modified *pttABC* expressing plasmid was used with detectable phototemtide production (Fig. 55). The constructed plasmid was verified by enzyme restriction digest but protein production via SDS-Page was not investigated.

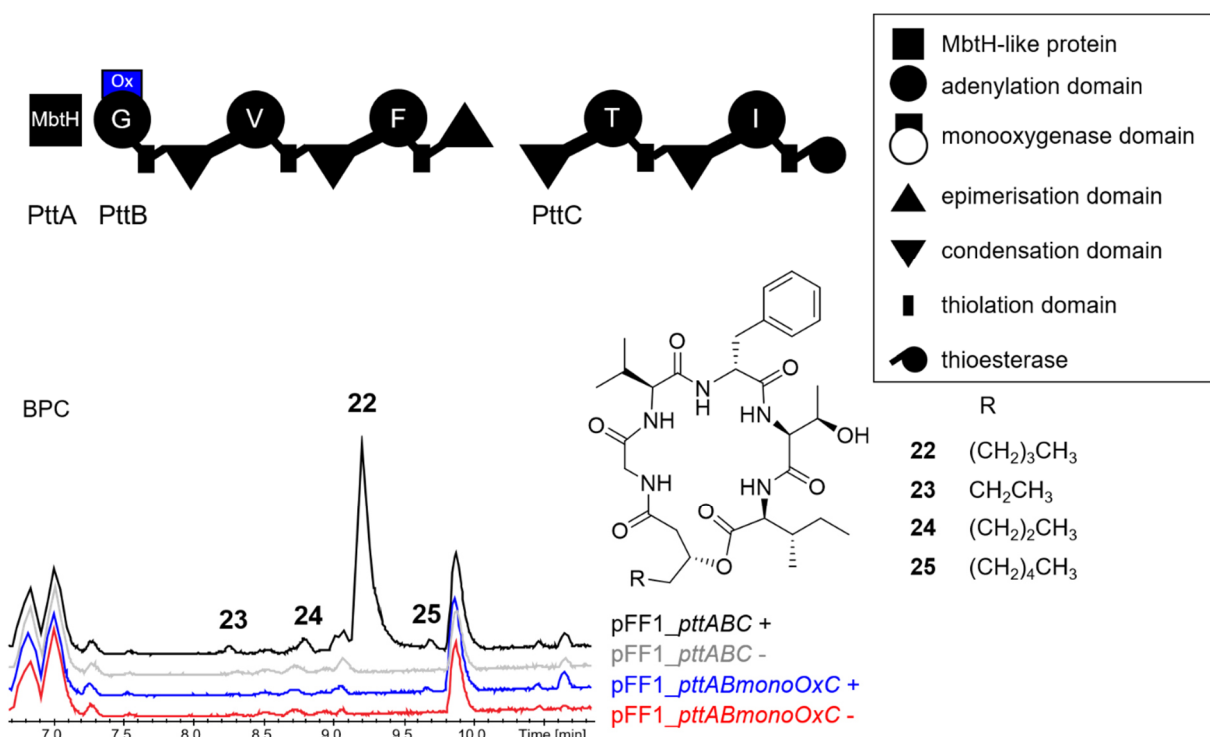


Figure 55. MonoOx domain is positioned in PttB_A1 domain. Induction of pFF1_*pttABC* leads to production of phototemtides A-D (**22-25**). No compound production upon induction of *pttABC* with MonoOx domain (pFF1_*pttABmonoOxC*). BPCs are shown. + arabinose induction, - no induction.

3.3.3 Transporter XgsD

The transporter encoding gene *xgsD* was deleted in *X. hominickii* Δ *hfq*. Similar to previous work in *E. coli*, *X. hominickii* Δ *xgsD* P_{BAD_xgsA} did not show an altered growth behaviour in comparison to *X. hominickii* P_{BAD_xgsA} (Fig. 56). The only observed difference in this experiment regards the finally reached OD₅₉₅ and is limited to a difference dependent on the absence or presence of L-arabinose (Fig. 56).

RESULTS

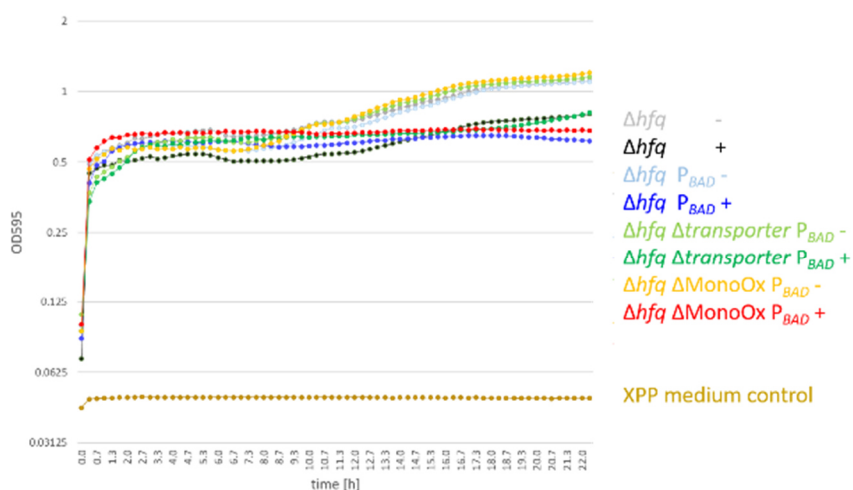


Figure 56. Growth behaviour of $\Delta xgsD$ and $\Delta MonoOx$ strains. Deletion of transporter encoding gene *xgsD* or deletion of *XgsC_MonoOx* domain encoding region do not affect cell growth. Growth curves for *X. hominickii* Δhfq , *X. hominickii* P_{BAD_xgsA} , *X. hominickii* $\Delta xgsD P_{BAD_xgsA}$ and *X. hominickii* $\Delta MonoOx P_{BAD_xgsA}$ are shown. + arabinose induction – no induction.

3.3.4 Distribution of XgsC_A2 domain with integrated MonoOx domain

BlastP analysis for *XgsC_A2* sequence was conducted to see whether similar domain organisation can be observed for other BGCs and strains. By this, the *xgs* BGC was found in other *Xenorhabdus* strains, namely *X. hominickii* ANU1, *X. miraniensis* and *Xenorhabdus* PB62.4 (Fig. 57).

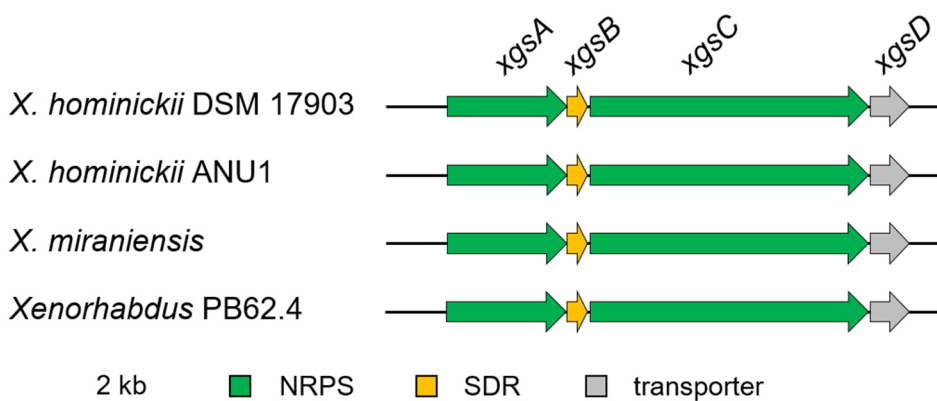


Figure 57. *xgs* BGC occurrence in *Xenorhabdus*. Glyoxypeptide BGC *xgsA-D* is observed in four *Xenorhabdus* strains: *X. hominickii* DSM 17903, *X. hominickii* ANU1, *X. miraniensis*, *Xenorhabdus* PB62.4.

Additionally, a variety of different BGCs from other proteobacteria was identified that harbour A domains with predicted glycine specificity and embedded MonoOx domain in between A8 and A9 motifs (Fig. 58). One exception is the free-standing A domain in *Chloroflexi bacterium* which, according to antiSMASH, activates β -alanine and not glycine.^[176] The observed BGCs in *Lonsdalea populi* and *Pseudomonas syringae* pv. *coryli* show the same domain organisation

of an NRPS that likely incorporates tyrosine and α -hydroxyglycine. In *Janthinobacterium* sp. DG3, the glycine specific A domain with embedded MonoOx domain is followed by an epimerisation domain in the second module of a trimodular NRPS which is part of a larger NRPS-PKS hybrid BGC (Fig. 39). This A_{Gly}-MonoOx-E domain feature is also found in the NRPS-PKS hybrid BGC in *P. temperata* K122 (Fig. 39). *P. temperata* K122 was not identified in the BlastP analysis as the genome was not accessible at NCBI at the time of the analysis but was added to this data set.

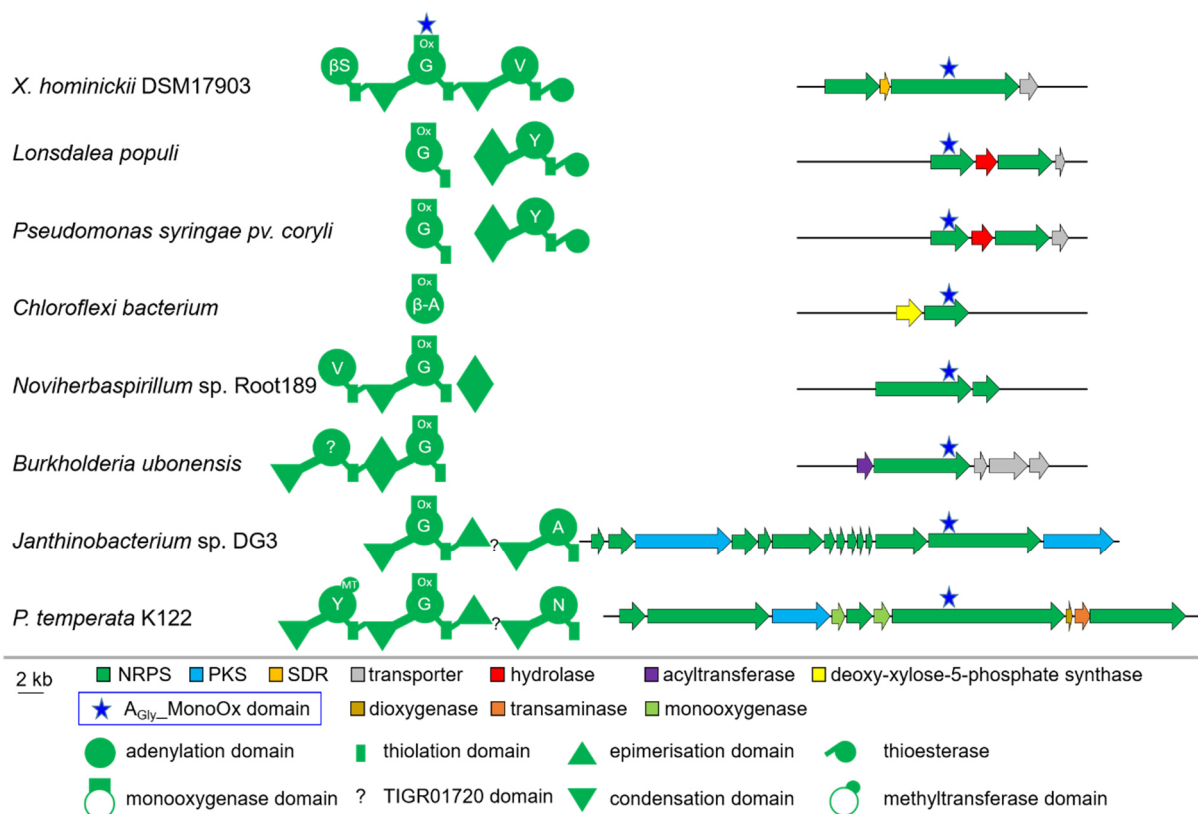


Figure 58. Occurrence of A_{Gly}_MonoOx domains. Glycine activating A domains with embedded MonoOx domain (A_{Gly}_MonoOx domain) in between A8/A9 core motifs are observed in different BGCs in proteobacteria. A_{Gly}_MonoOx domain position is indicated with blue star. Identification via BlastP analysis with XgsC_A2 domain as query sequence.^[170] AntiSMASH was used for BGC analysis.^[176]

RESULTS

3.4 Xildivaline

Analysis of the activated *X. hominickii* Δhfq P_{BAD_xisA} results in the detection of three main compounds, **27** with m/z 354.3 $[M+H]^+$ and **26a** and **26b** with m/z 372.3 $[M+H]^+$ as described previously.^[153] **26a** and **26b** show the same m/z 372.3 $[M+H]^+$ and the same MS² fragmentation but elute at different retention times (Fig. 59). This might indicate that the two signals correspond to two stereoisomers with different stereochemistry most likely for the hydroxyl function.

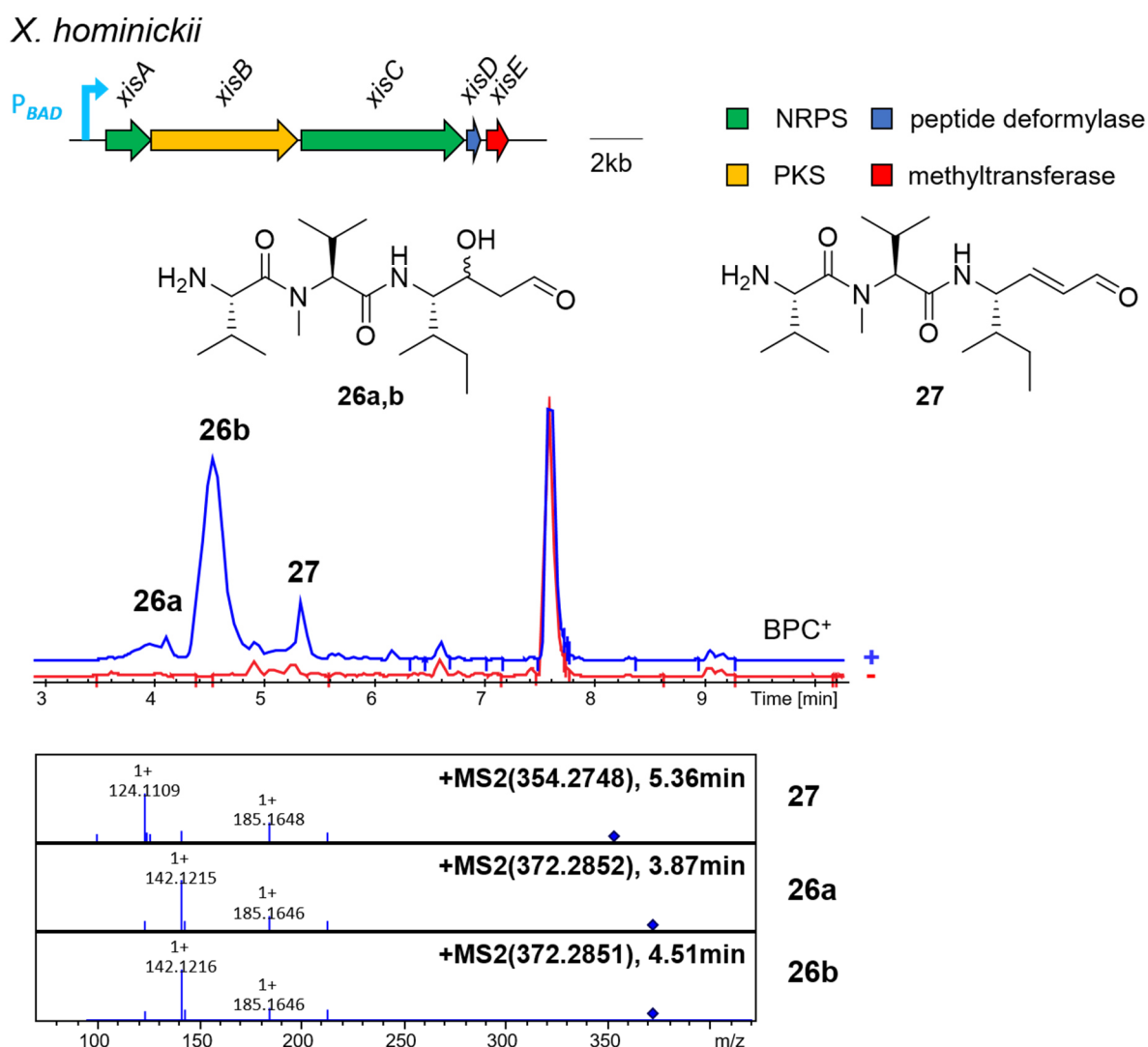


Figure 59. Compound overview for *X. hominickii* Δhfq P_{BAD_xisA} . P_{BAD} induction in *X. hominickii* Δhfq P_{BAD_xisA} leads to two signals with m/z 372.3 $[M+H]^+$ (**26a,b**) and one signal with m/z 354.3 $[M+H]^+$ (**27**). BPCs are shown, - no induction, + arabinose induction. MS² spectra and structures for **26a,b** and **27** are shown.

3.4.2 Methyltransferase XisE deletion

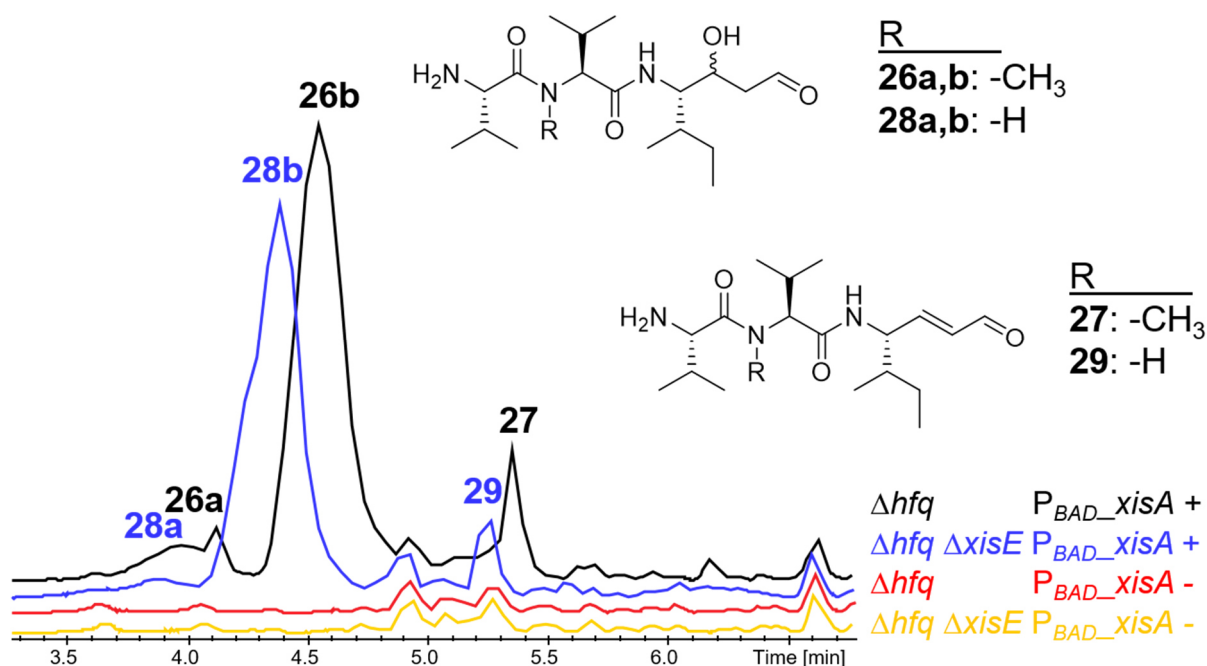


Figure 60. Production of non-methylated derivatives in *xisE* deletion strain. Arabinose induction in *X. hominickii* $\Delta hfq \Delta xisE P_{BAD_xisA}$ (blue) leads to production of **28a,b** and **29**, the non-methylated derivatives of **26a,b** and **27** produced in *X. hominickii* $\Delta hfq P_{BAD_xisA}$ (black). *xisE* encodes a methyltransferase. BPCs are shown. + arabinose induction, - no induction.

Deletion of the methyltransferase encoding gene *xisE* allowed identification of the non-methylated compounds **28a,b** and **29**. The comparison of the b-ion fragmentation pattern of methylated and non-methylated derivatives shows that they differ in the b_2 -fragment ion by $\Delta m/z$ 14 (Fig. 61). The analysis of data from d_3 -methionine feeding experiments performed by Janine Chekaiban shows the $\Delta m/z$ 3 shift for the b_2 -fragment ion of **26** and **27** (Fig. 61). This suggests that the second L-valine carries the *N*-methylation. The detected b_2 -ion Both A domains in XisB are predicted to incorporate valine according to antiSMASH and share the identical Stachelhaus code (Table 16).^[64,176]

Table 16. Stachelhaus code for XisA and XisC A domains. Based on alignment with GrsA (PDB: 1amu).^[59] Predicted (antiSMASH) and activated amino acids are stated in three letter code in the last two columns. X indicates no prediction from antiSMASH v5.1.0.^[176]

GrsA	D 235	A 236	W 239	T 278	I 299	A 301	A 322	I 330	C 331	K 517	Prediction antismash	In 26
XisA	D	A	F	F	F	G	C	I	W	K	X	Ile
XisC_A1	D	A	L	W	L	G	G	T	F	K	Val	Va I
XisC_A2	D	A	L	W	L	G	G	T	F	K	Val	Va I

RESULTS

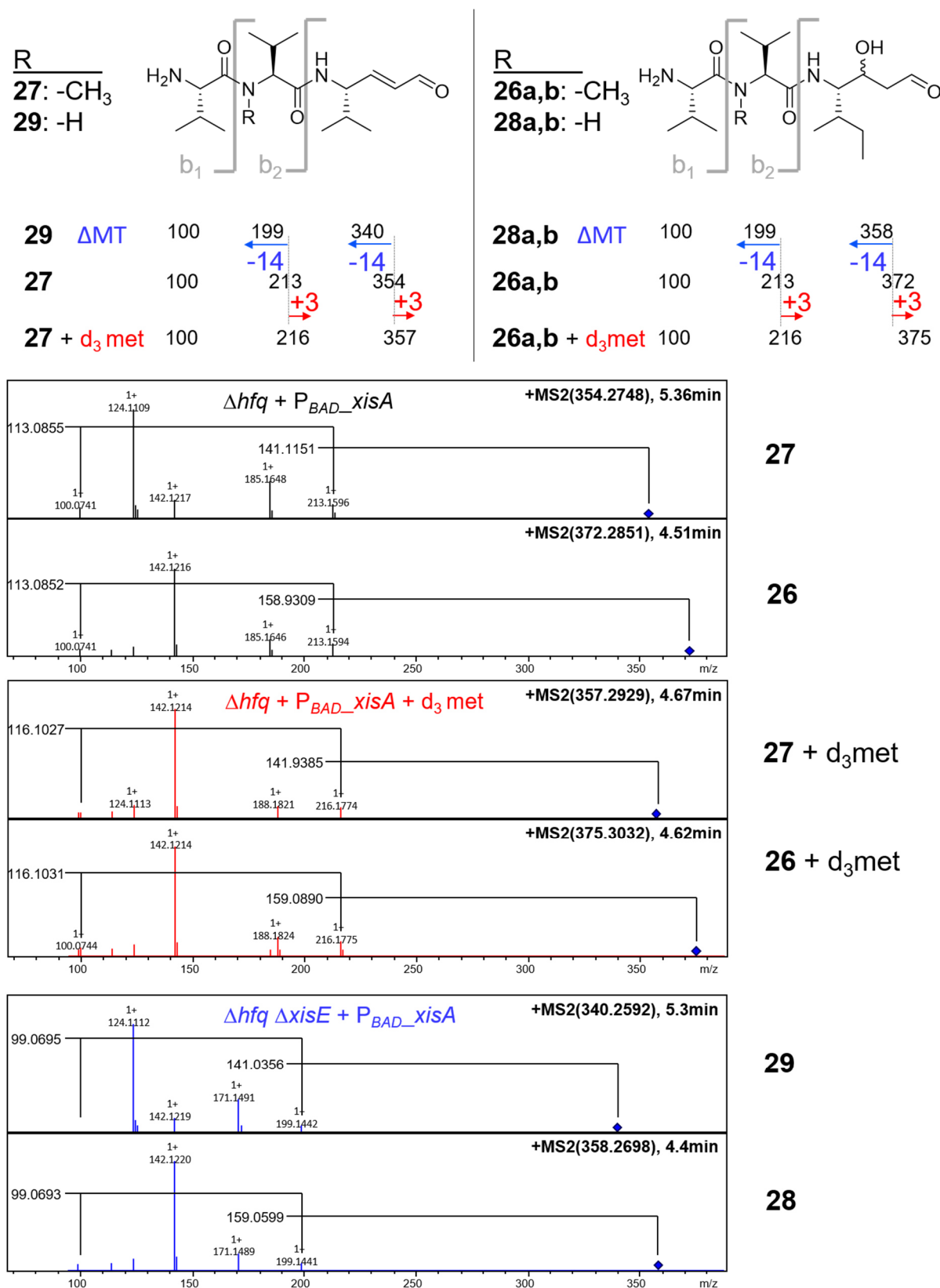


Figure 61. *N*-methylation of the second L-valine residue in xildivaline. b-ion fragmentation for non-methylated **28a,b** and **29** in comparison with methylated **26a,b** and **27** indicates position of the *N*-methylation in the b₂-fragment ion. Feeding experiment with d₃-methionine (d₃met) was conducted by Janine Chekaiban.^[153]

To investigate the presence of the unsaturated double bond, 4-bromothiophenol (BTP) was used as a strong nucleophile and added to culture supernatants of the induced and non-induced *X. hominickii* Δhfq P_{BAD_xisA} strain (Fig. 62). As result of the Michael addition of BTP to the Michael acceptor in **27**, the product **30** was detected exclusively in the induced culture sample. The EIC of the divalene fragment ion with m/z 213.1 $[M+H]^+$ at the same retention time supports this interpretation although no MS² spectrum was recorded for **30**.

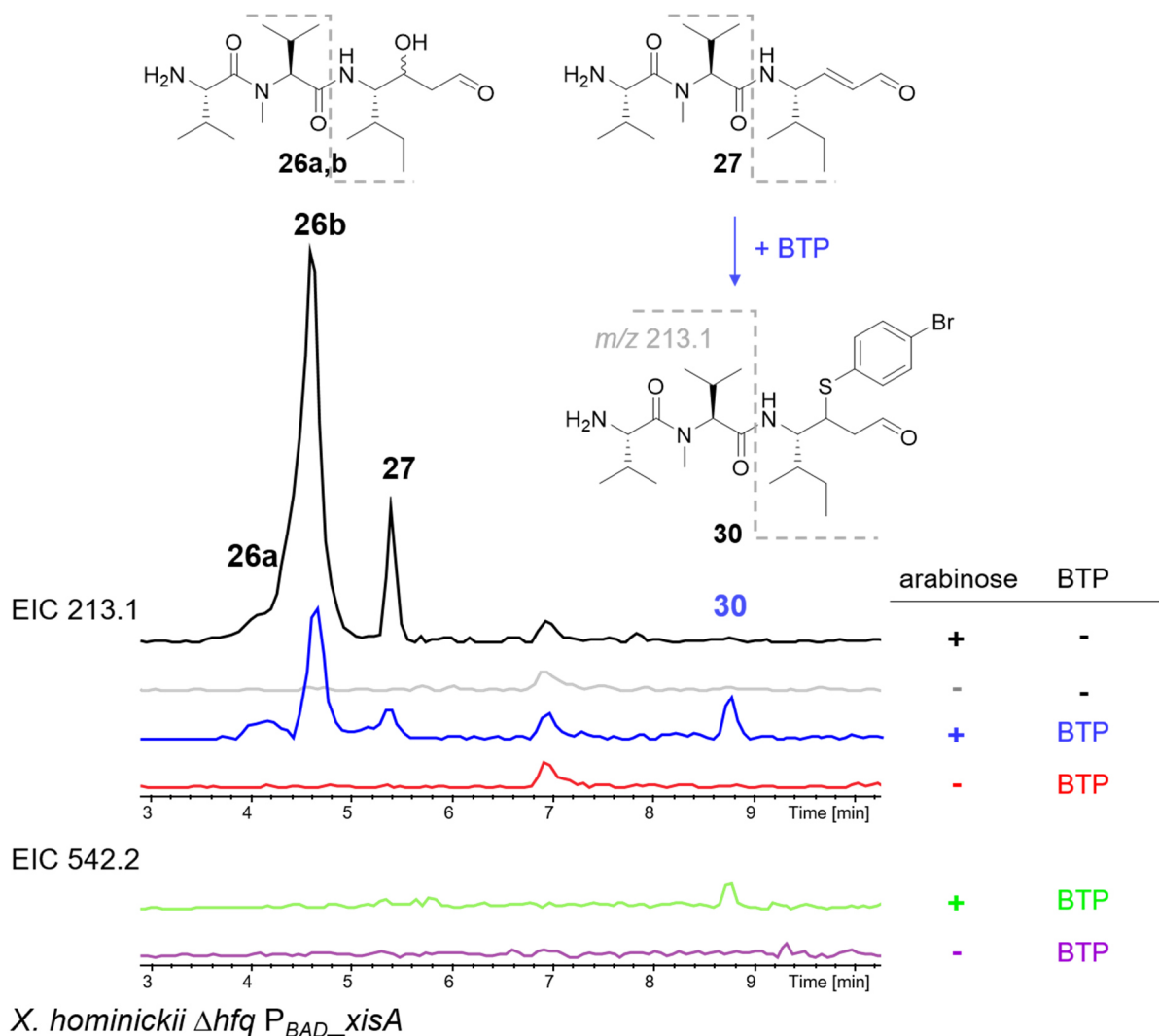


Figure 62. 4-bromothiophenol addition reveals a Michael acceptor in **27**. EICs for the L-Val-N-Me-L-Val-fragment ion (m/z 213.1 $[M+H]^+$) show presence of **26a,b**, **27** and **30** in the arabinose induced (+) sample of *X. hominickii* Δhfq P_{BAD_xisA} after incubation with 4-bromothiophenol (BTP). **30** is the product from Michael addition of BTP to the α,β unsaturated bond in **27**. The EICs for the expected product **30** with m/z 542.2 $[M+H]^+$, and the divalene fragment ion (m/z 213.1 $[M+H]^+$) are shown for extracts of the induced (+) and non-induced (-) *X. hominickii* Δhfq P_{BAD_xisA} sample.

3.4.3 P_{BAD_xisA} activation in *X. hominickii* WT and *X. hominickii* Δhfq

Comparison of the P_{BAD_xisA} activation in *X. hominickii* WT and *X. hominickii* Δhfq revealed two additional signals in *X. hominickii* P_{BAD_xisA} (Fig. 63). Signal **31** is observed at 3.8 min with m/z 374.3 $[M+H]^+$ and **32** with m/z 358.3 $[M+H]^+$. The observed b-ion fragmentation

RESULTS

indicates that both products differ from **26a,b** only in the C-terminal building block of the peptide that is likely installed by the PKS XisB of the BGC. HR-MS data support the sum formulas with $C_{19}H_{40}N_3O_4$ for **31** and $C_{19}H_{40}N_3O_3$ for **32** and corresponding structures are proposed (Fig. 63).

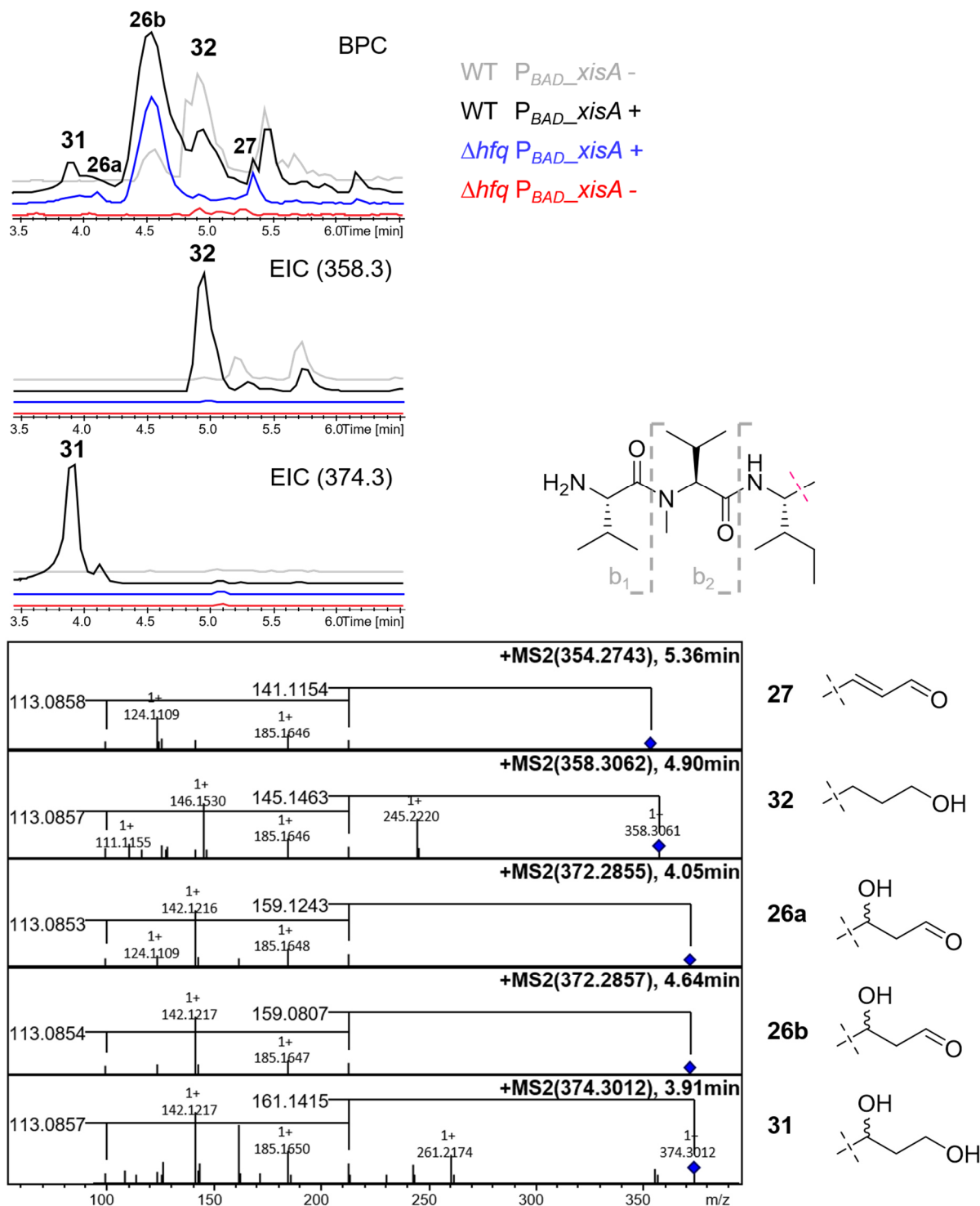


Figure 63. Comparison of P_{BAD_xisA} activation in *X. hominickii* WT (black) and Δhfq (blue). **31** and **32** are only observed upon promoter activation in the WT strain. BPCs and EICs for the observed masses of **31** and **32** are shown. MS² fragmentation for b-ion fragments (dashed grey lines) is shown. – no induction, + arabinose-induction

32 shares the similar m/z 358.3 $[M+H]^+$ with the non-methylated compounds **28a,b** but shows a different MS² fragmentation, different retention time and different sum formula prediction (Fig. S15). Performing *xisE* deletion in the *X. hominickii* WT strain, followed by P_{BAD_xisA} insertion and activation, compounds **33** and **34** are detected as the non-methylated derivatives of **29** and **30** respectively (Fig. 64).

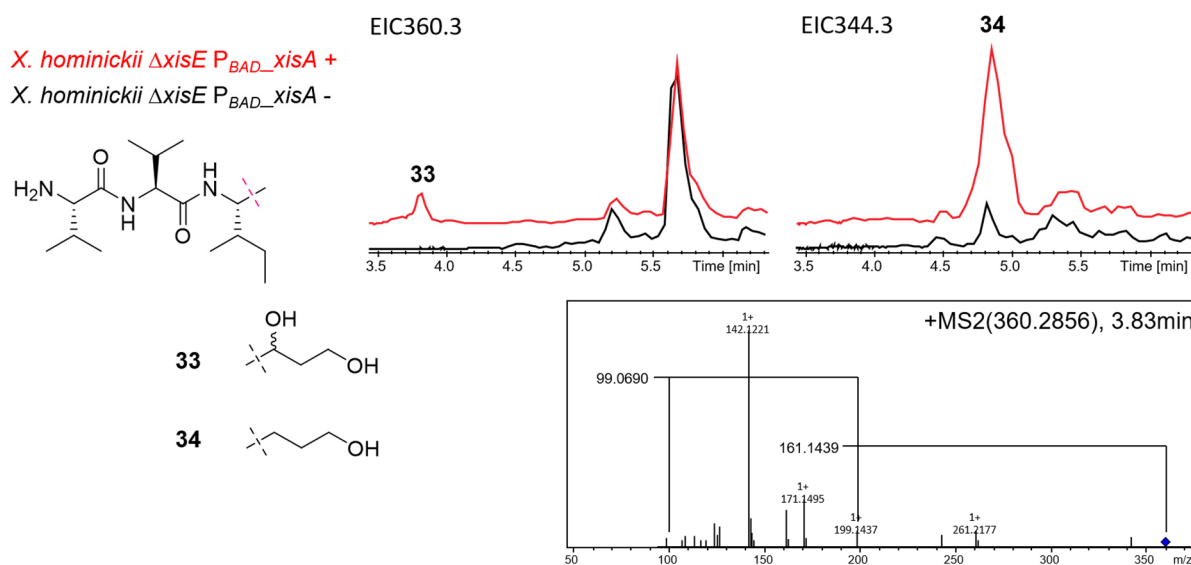


Figure 64. Non-methylated products **33** and **34** in *X. hominickii* $\Delta xisE$ P_{BAD_xisA} . MS² spectra was only recorded for **33**. B-ion fragmentation for **33** is indicated. EICs are shown for **33** and **34**. + arabinose induction, - no induction.

3.4.4 Peptide deformylase encoding gene *xisD* deleted in *X. hominickii* WT

The peptide deformylase (PD) encoding gene *xisD* was deleted in the *X. hominickii* WT strain, followed by promoter exchange in front of *xisA*. The deletion of *xisD* in *X. hominickii* Δhfq was not successful. Comparison of LC-MS data from P_{BAD_xisA} activation in WT and $\Delta xisD$ strain showed no differences (Fig. 65, Fig. S16). The previously observed signals for **31** and **32** from P_{BAD_xisA} activation in *X. hominickii* WT as well as signals for **26a,b** and **27** are detected in the induced *X. hominickii* $\Delta xisD$ P_{BAD_xisA} sample. EICs for m/z 381.3 $[M+H]^+$ and m/z 399.3 $[M+H]^+$ corresponding to the mass of formylated products of **26a,b** and **27** respectively, did not reveal differences between the samples (Fig. 65). In addition to the LC-MS analysis, the comparison of the cell growth also showed no differences between *X. hominickii* WT and *X. hominickii* $\Delta xisD$ (Fig. 66).

RESULTS

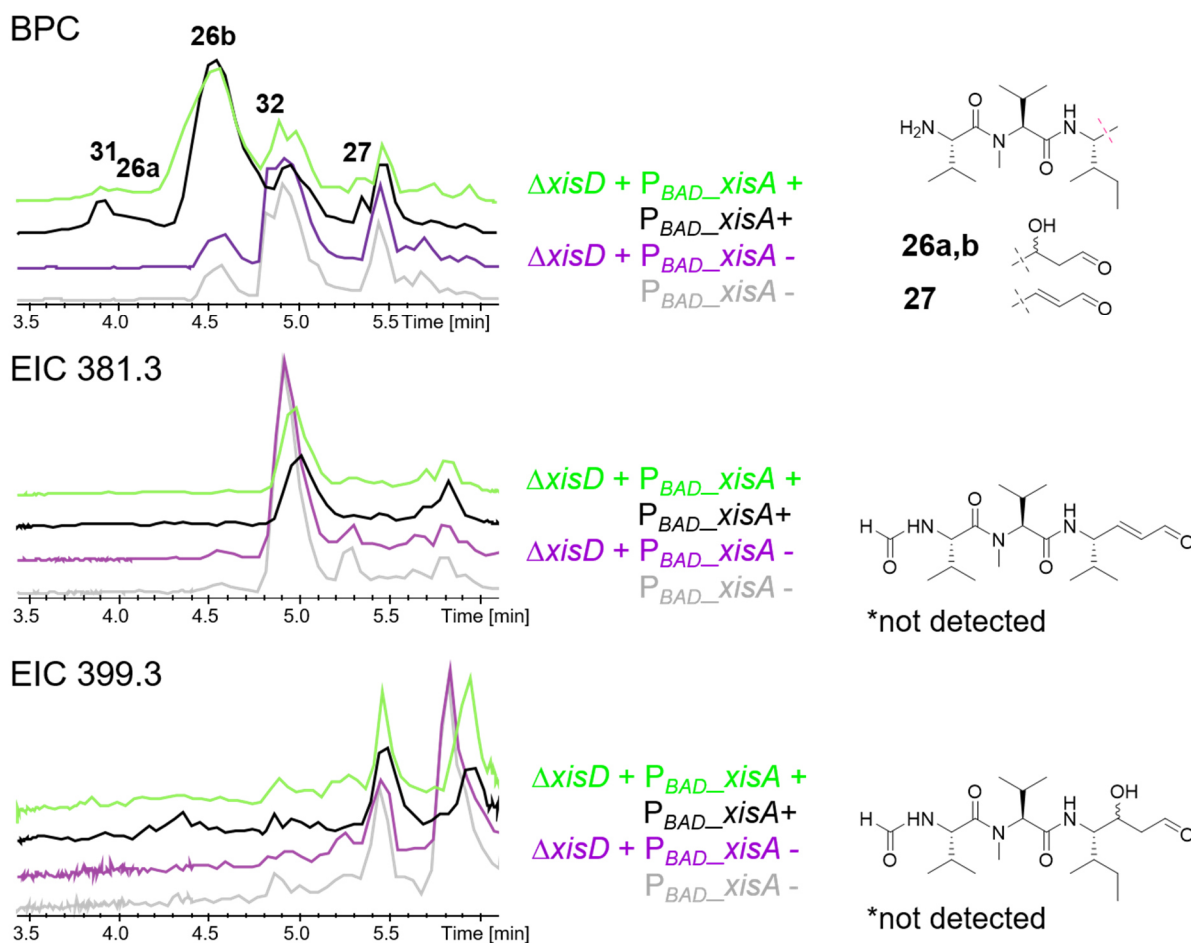


Figure 65. Compound production in peptide deformylase *xisD* deletion strain. P_{BAD_xisA} activation in *X. hominickii* WT (black) and $\Delta xisD$ (light green) leads to similar product formation. BPCs with signals for **26a,b**, **27**, **31** and **32** and EICs for the corresponding formylated masses of **26a,b** (EIC 399.3) and **27** (EIC 381.3) are shown. For EICs 399.3 and EICs 381.3 no signals are detected that would indicate presence of the hypothetical *N*-formylated derivatives of **26a,b** and **27** in the induced sample. – no induction, + arabinose induction. For EICs of masses corresponding to **26a,b**, **27**, **31** please see Fog. S16.

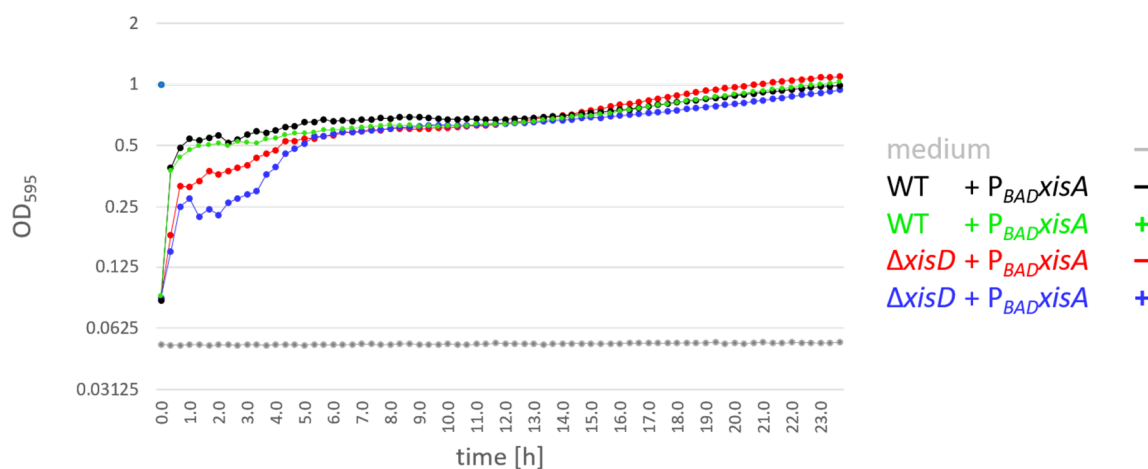


Figure 66. Growth behaviour for *X. hominickii* WT (black) and *X. hominickii* $\Delta xisD$ (blue).

3.4.5 XisB exhibits two adjacent thiolation domains

XisB features two adjacent T domains, now referred to as T1 and T2. To investigate their role in xildivaline biosynthesis, the active site serines, responsible for Ppant attachment, were exchanged to alanine, respectively in either the T1 or T2 domain. This was done by point mutation in the *X. hominickii* WT strain, followed by P_{BAD} insertion in front of *xisA*. The inactivation in both respective T domains still allowed the detection of the five xildivalines **26a,b**, **27**, **31** and **32** that are also observed in extracts of the induced *X. hominickii* WT P_{BAD_xisA} strain (Fig. 67A). However, the relative production for the detected xildivalines is decreased in both T domain mutant strains when compared to the wild type strain (Fig. 67B). For the main compounds **26a,b** and **27** the production levels decrease down to 39%, 47% and 27% in case of the point mutation in the T2 domain while the decrease is less strong for T1 inactivation with production levels of 65%, 69% and 57% in comparison to the wild type background. For xildivalines **31** and **32** the decrease is comparable between T1 and T2 domain point mutation which is around 70% production decrease compared to the wild type (Fig. 67B).

RESULTS

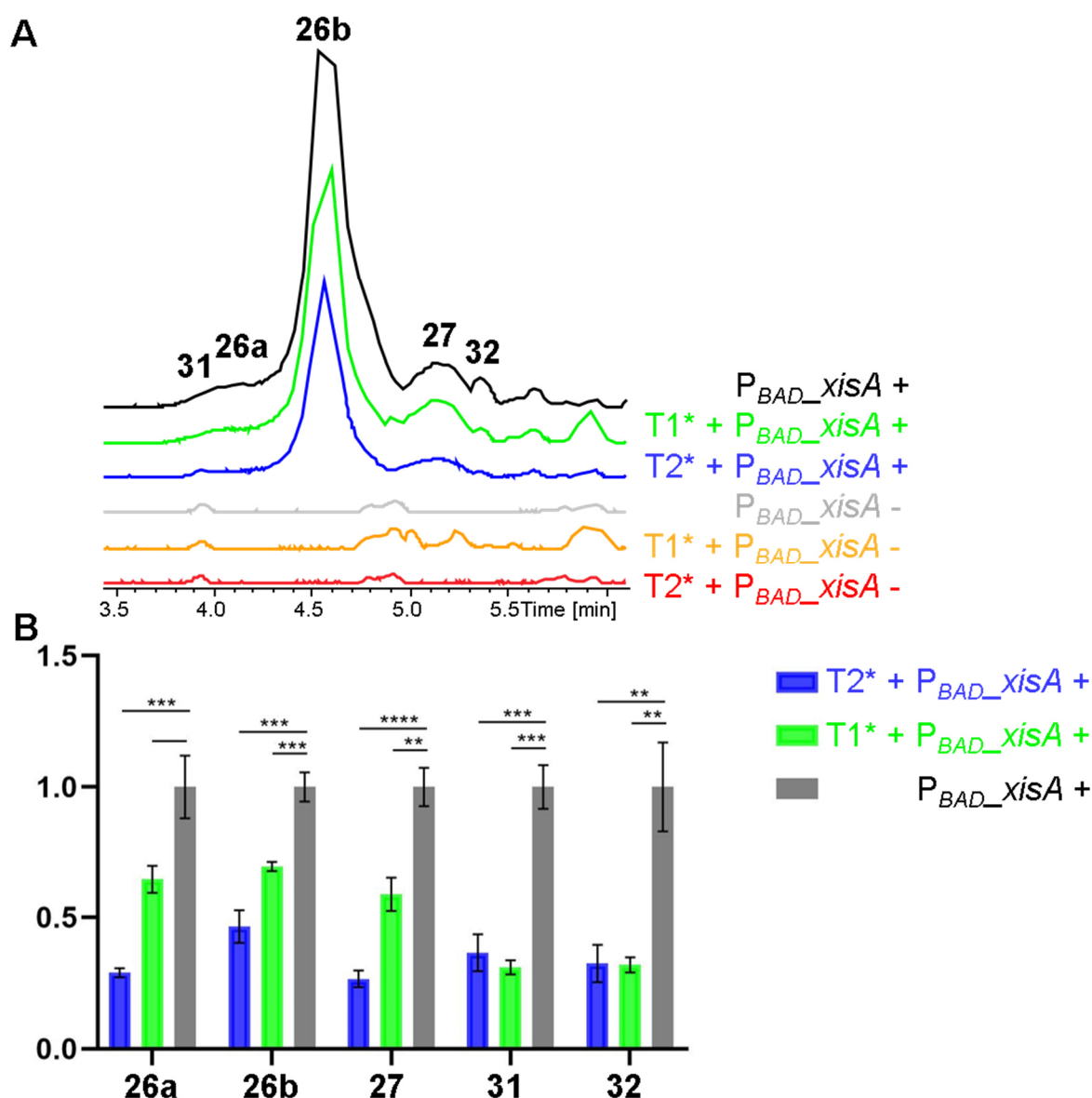


Figure 67. Influence of T domain point mutations on xildivaline production. **A** BPCs for P_{BAD_xisA} in *X. hominickii* wild type (WT), and point mutation in T1 and T2 domains in *xisB* respectively ($T1^*$, $T2^*$). BPCs are shown. + arabinose induction, - no induction. **B** Relative xildivaline production upon P_{BAD_xhpA} activation in *X. hominickii* WT, $T1^*$ and $T2^*$ strains. Production is relative compared to the wild type. For xildivalines **26a,b**, **27**, **31** and **32**. Peak areas were normalised to final OD_{595} . Bars show the mean of three replicates and error bars indicate standard deviation. Asterisks mark statistical significance from two-tailed unpaired t-test. ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. Exact p values are (from left to right) for **26a**: 0.0005, 0.009; **26b**: 0.0004, 0.0008, for **27**: <0.0001, 0.0018; for **31**: 0.0005, 0.0002; for **32**: 0.0031, 0.0023.

3.4.6 In silico analysis of XisB_KR

The two signals with the same m/z 372.3 $[M+H]^+$ at different retention times might represent two different stereoisomers for the hydroxyl function introduced by the ketoreductase domain of XisB (XisB_KR). This is supported by the presence of only one signal with m/z 354.3 $[M+H]^+$

in which the hydroxyl function is absent according to the prediction. The protein sequence based protocol to predict a ketoreductase product's stereochemistry introduced by Keatinge-Clay was used to further investigate XisB_KR.^[114] Sequence alignment with representatives from five different KR types reveals that XisB_KR shows the catalytic tyrosine indicating an active ketoreductase (Fig. 68). XisB displays no dehydratase (DH) domain and XisB_KR shows no LDD motif and no HXXY motif and hence, according to Keatinge-Clay's substituent flowchart, XisB_KR is a type A1 ketoreductase.

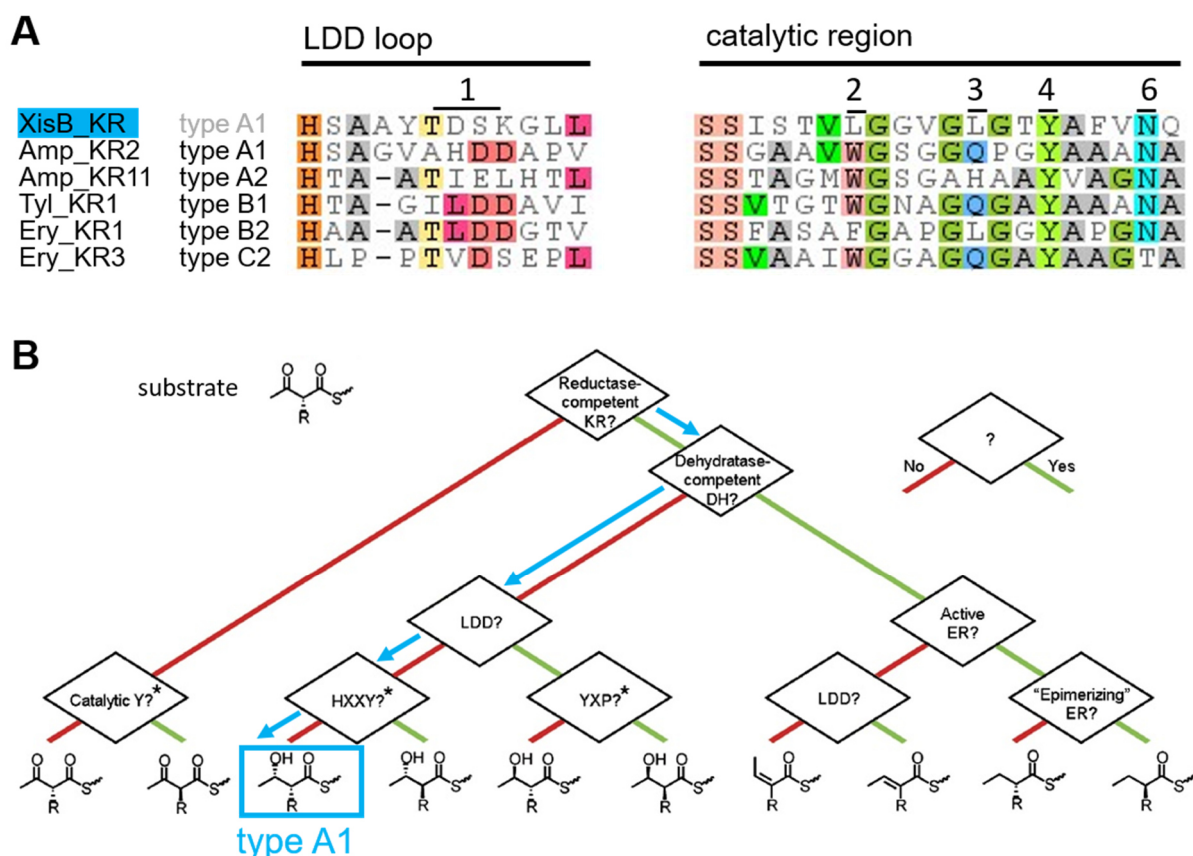


Figure 68. XisB_KR *in silico* analysis based on the protocol by Keatinge-Clay.^[114] **A** LDD loop and catalytic region are aligned with representatives from five different KR types. XisB_KR shows no LDD motif (1) present in B type ketoreductases and shows leucine (2) instead of conserved tryptophan found in A type ketoreductases. XisB_KR shows no H in position 3 and the catalytic tyrosine in position 4. **B** Substituent flowchart after Keatinge-Clay determines XisB_KR as type A1 KR. ^[114] XisB exhibits no DH domain, XisB_KR shows no LDD motif (1) and no HXXY motif (3-4).

Type A ketoreductases are described to display a tryptophan in position 2 to guide substrate entry in the active site from the left side resulting in a S-hydroxyl group (Fig. 68).^[114] XisB_KR exhibits a leucine instead of tryptophan in this position. The division of type A ketoreductases in type A1 and A2 is based on the occurrence of the epimerisation reaction in α -position. According to their classification, XisB_KR might be type A1 due to its LxxY motif which was previously observed for other type A1 KR.^[193] However, this is of less importance for compound **26a,b** as it does not show a stereocenter in α -position and thus is regarded as A0 type KR.^[194]

RESULTS

3.4.7 Two possible biosynthetic routes

Analysing the *xis* BGC organisation in combination with the identified product, two major biosynthetic routes are possible (Fig. 69). Route A follows the BGC order as first isoleucine is activated and then accepted by the KS of the PKS. NaPDoS identifies the KS as hybrid KS.^[81] The NRPS-PKS product is then reductively released and the amino group attacks the thioester-carbon of the dipeptide attached to NRPS XisC to generate the free tripeptide.

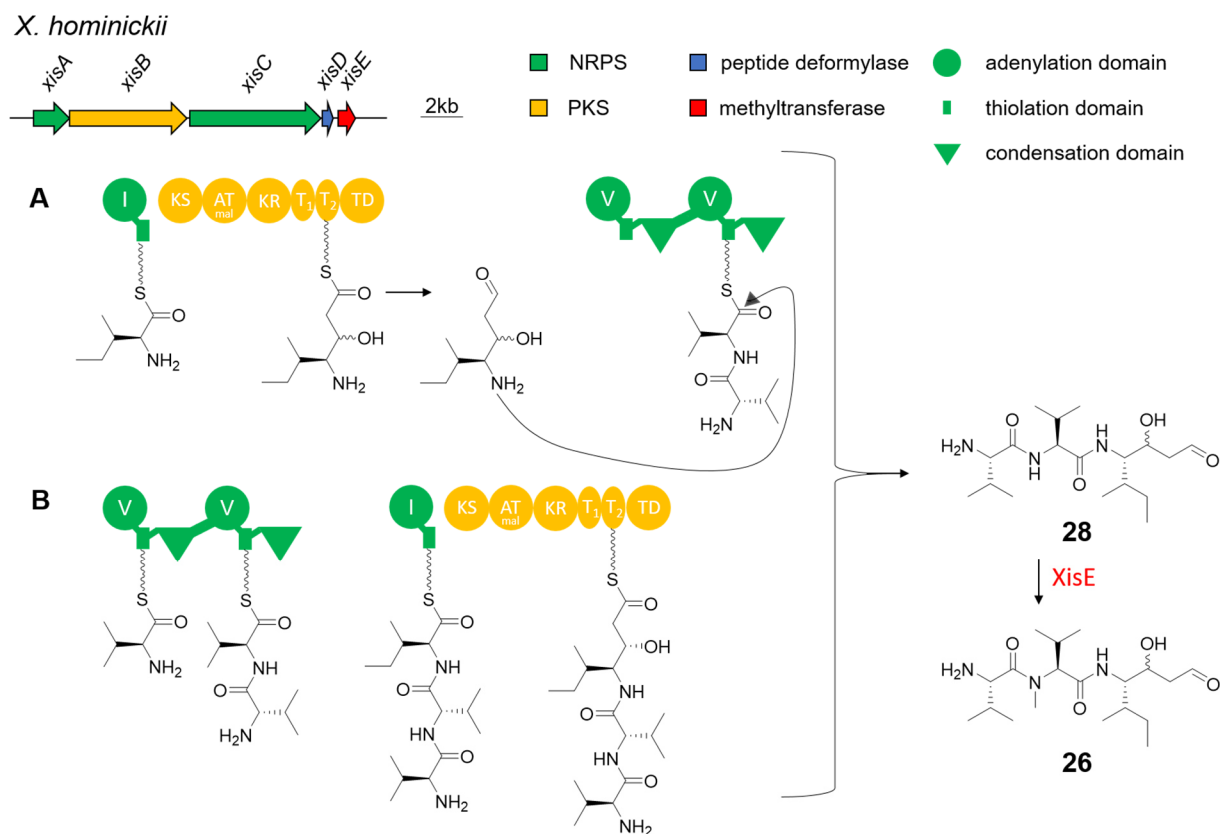


Figure 69. Two possible biosynthetic routes for xildivaline production. Route **A** follows the cluster organisation depicted above within the strain (modified from ^[153]) Route **B** starts with the bimodular NRPS XisC whose C-terminal C domain condenses the divaline peptide with isoleucine activated by XisA. XisB elongates the tripeptide and releases the non-methylated NRPS-PKS hybrid product **28**.

Alternatively, as depicted in Route B, XisC_C2 catalyses the condensation of the valylvaline bound to XisC_T2 with the isoleucine activated by XisA. The hybrid KS in XisB would then accept the tripeptide, elongate it and reductively release the non-methylated compound **28**. Finally, as suggested in both route A and B, XisE catalyses the *N*-methylation to yield compound **26**. Analyses using NaPDos and antiSMASH predict XisC_C2 as ^LC_L domain.^[81,176] The multiple sequence alignment of XisC_C2 with the free-standing C domains VibH and Kj12C, as well as with the terminal C domains from XndB and TqaA and with the ^LC_L domain XisC_C1 domain shows no additional C-terminal sequence for XisC_C2 that could be interpreted as a C-terminal docking domain for the communication with XisA (Fig. S18). Sequence alignment of XisA_T domain with XisC_T1 and XisC_T2 domains as well as with

PaxA_T1 and PaxB_T3 domains shows no C-terminal extension for XisA_T which could indicate a docking domain (Fig. 70). PaxA_T1 and PaxB_T3 exhibit C-terminal docking domains for communication with downstream C domains.^[122,195]



Figure 70. Multiple sequence alignment of XisA_T domain with standard T domains XisC_T1 and XisC_T2 and bridging T domains PaxA_T1 and PaxB_T3. PaxA and PaxB exhibit C-terminal regions following their respective terminal T domain which are involved in communication with downstream C domains in PAX peptide biosynthesis.^[122,195] Red arrows mark T domain sequence predicted from antiSMASH.^[133] Alignment was created with Geneious v 6.1.8.^[178]

Phylogeny of the *xis* cluster

In-depth sequencing results revealed that the *xis* BGC is located on a plasmid pXhom in *X. hominickii* DSM 17903 and not present in the circular bacterial chromosome (Lab Prof. Dr. Helge Bode, personal communication) (Fig. 71). Overall, *X. hominickii* DSM 17903 harbours four plasmids in addition to the circular chromosome. The approx. 96 kb plasmid exhibits a GC content of 43.3 %, similar to the chromosome's GC content of 43.4 %. For this plasmid, 90 coding sequences are annotated and antiSMASH predicts two additional PKS enzyme elements next to the investigated *xis* BGC.^[176] These are on the one hand a type I PKS (KS-AT-DH-KR-T) as well as a C2-type KR which is accompanied by a SDR (Fig. 71). Nucleotide BLAST analysis with the entire pXhom sequence as query showed that *X. hominickii* ANU1 exhibits a highly similar plasmid (100 % sequence identity and 99 % query cover) (Fig. S17).

RESULTS

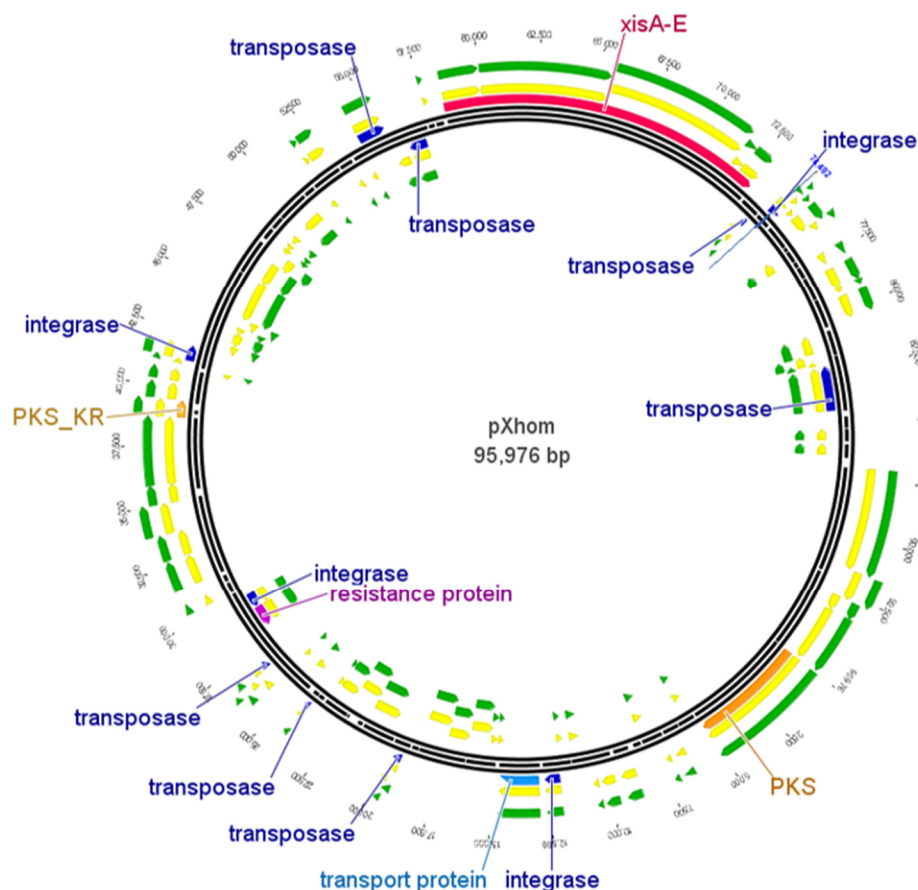


Figure 71. Plasmid map for natural plasmid pXhom in *X. hominickii* WT DSM 17903. pXhom carries the *xis* BGC and two additional PKS-related enzyme coding genes (orange). Transposase/integrase encoding genes are depicted in dark blue, transport related protein encoding gene in light blue and a predicted resistance protein encoding gene in purple.

Additionally, clusterblast analysis provided by the antiSMASH analysis revealed *Vibrio tubiashii* ATCC 19109 and *Photobacterium asymbiotica* ATCC 43949 to harbour a similar BGC (Fig. 72).^[176] The reference genomes available on NCBI indicate that therein the corresponding clusters are located on the bacterial chromosome.^[177]

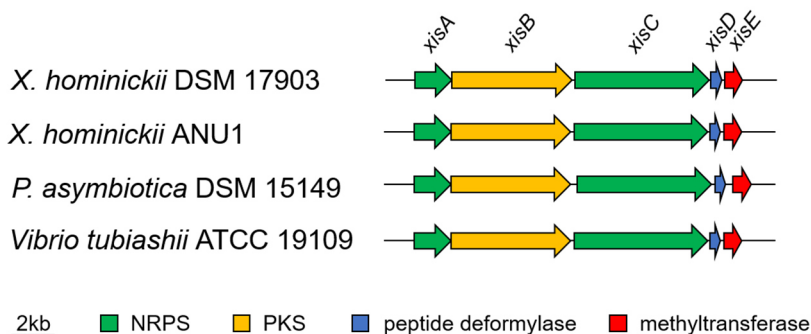


Figure 72. *xis* BGC occurrence. The *xis* BGC is found in two *X. hominickii* strains, *P. asymbiotica* DSM 15149 and *Vibrio tubiashii* ATCC 19109. Clusters were identified via clusterblast by antiSMASH v 5.1.0 analysis.^[176]

AntiSMASH predicts different amino acid specificities for the first NRPS (XisA) in the three strains.^[176] For *X. hominickii*, in which isoleucine incorporation was observed, no amino acid is specified by antiSMASH, while the program predicts leucine for the BGC in *V. tubiashii* and isoleucine in *P. asymbiotica*.^[153,176]

4 Discussion

4.1 Xenofuranone biosynthesis

In previous work, xenofuranones A and B were identified from *X. szentirmaii*.^[152] A biosynthetic route starting from two phenylpyruvate moieties was proposed and the corresponding *xf*s BGC was identified by gene deletion.^[152,155] This work aimed at a more detailed characterisation of enzymes responsible for xenofuranone biosynthesis. Single gene deletions led to the identification of the decarboxylase XfsB and the methyltransferase XfsC and provided further insights into xenofuranone biosynthesis (Fig. 73). Additionally, P_{BAD} activation of the *xf*s BGC resulted in the detection of two novel xenofuranone derivatives.

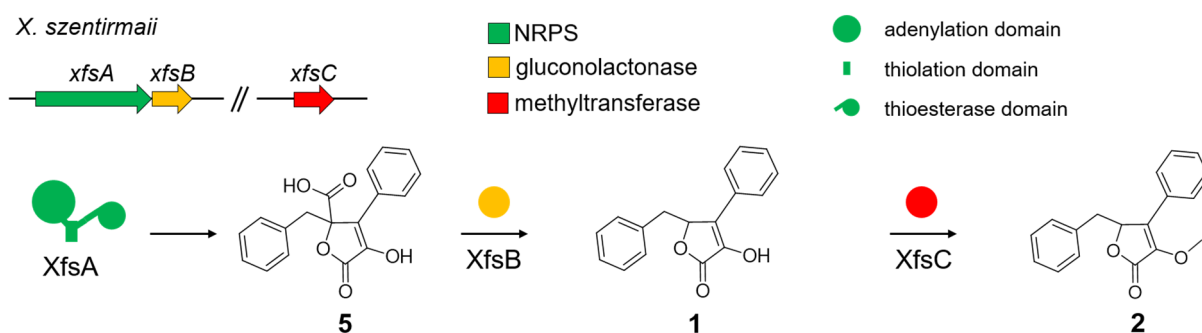


Figure 73. Xenofuranone biosynthesis in *X. szentirmaii*. The NRPS-like enzyme XfsA condenses two phenylpyruvic acids and releases **5**. XfsB decarboxylates **5** to yield xenofuranone B (**1**). The methyltransferase XfsC methylates **1** to produce xenofuranone A (**2**).

4.1.1 Xenofuranone decarboxylase XfsB

The *xf*sB deletion results in the production of the carboxylated furanone intermediate **5** with m/z 311.1 $[M+H]^+$ in *X. szentirmaii*. This compound was identified as phenylbutyrolactone IIa and was described previously to be produced by an engineered NRPS-like enzyme built from phenguignardic acid synthesising PgnA_A domain and butyrolactone producing BtyA_T domain and BtyA_TE domain (Fig. 74).^[183] Thus, the detection of **5** as product of XfsA alone is consistent with the literature that states NRPS-like enzymes alone are sufficient for furanone core biosynthesis which was demonstrated by heterologous expression studies.^[89,92]

DISCUSSION

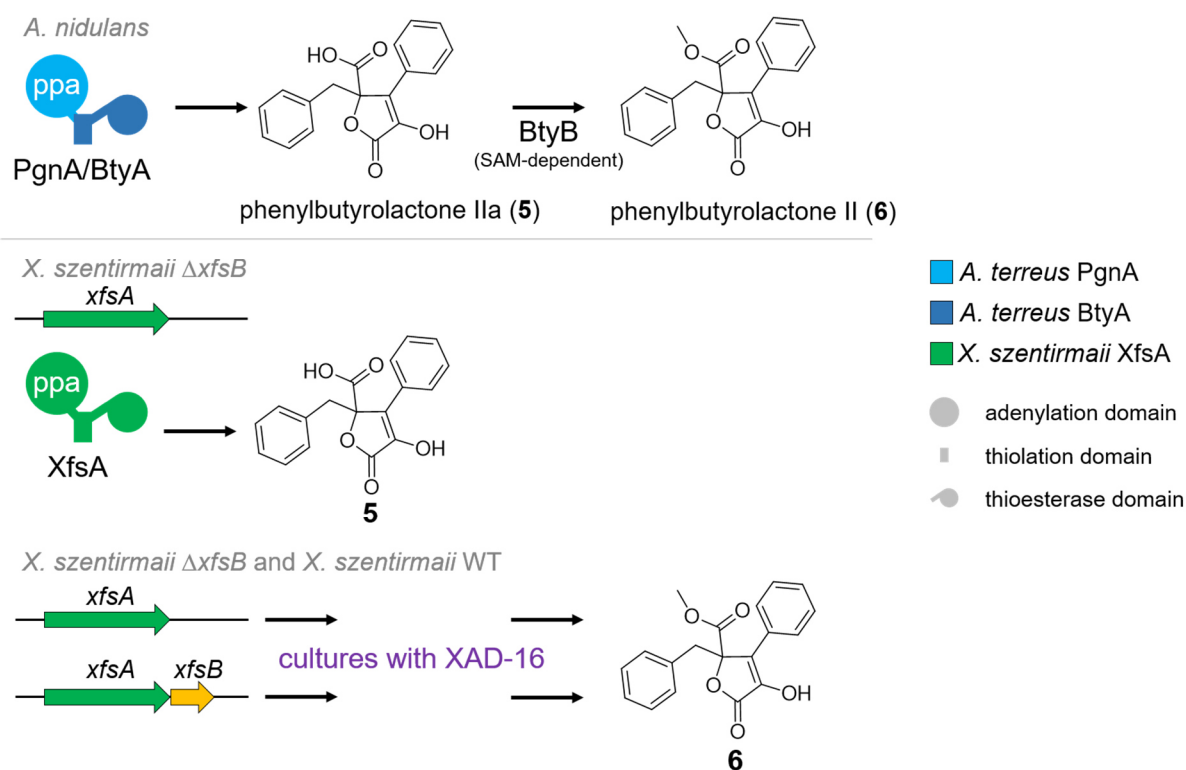


Figure 74. Detection of phenylbutyrolactone IIa and phenylbutyrolactone II in *A. nidulans* and *X. szentirmaii*. Phenylbutyrolactone IIa (**5**) is produced by PgnA/BtyA-hybrid protein in *A. nidulans* and is detected in *X. szentirmaii* $\Delta xfsB$. Phenylbutyrolactone II (**6**) is produced from **5** by BtyB in *A. nidulans* while **6** is detected in *X. szentirmaii* only when cultivated in presence of XAD-16. Taken and modified from van Dijk *et al.* 2018.^[196] *A. terreus* proteins were expressed in *A. nidulans* by van Dijk *et al.* 2018.^[196] Both PgnA_A and XfsA_A domains activate phenylpyruvic acid (ppa).

In addition, the data in this work support that XfsB promotes decarboxylation of **5** to yield xenofuranone B. The minor levels detected for **1** in the *xfsB* deletion and the non-induced complementation indicate that decarboxylation can occur spontaneously, but more slowly in comparison to the wild type strain or the induced complementation strain. XfsB is predicted to encode a gluconolactonase and the crystal structure obtained for XfsB exhibits a six-bladed β -propeller fold with two lid helices and a lid loop (Martin Schwalm, personal communication, data not shown). The overall fold resembles described senescence marker protein 30 (SMP30)/gluconolactonases which are involved in the hydrolysis of gluconolactone to gluconic acid.^[197] In addition, SMP30 proteins are able to hydrolyse diisopropyl phosphorofluoridate (DFP) and it was even postulated that the respective coordinated metal ions influence the reaction kinetics.^[197,198] The observed decarboxylation promoted by XfsB could be an additional catalytic activity previously not described for this enzyme type or even indicate that the XfsB fold represents a novel decarboxylase fold.^[199] This could be further addressed by challenging the purified XfsB with different substrates, e.g. xenofuranone B, gluconolactone or DFP, in presence of different ions. Altogether in xenofuranone biosynthesis, XfsA produces **5** which is subsequently decarboxylated by XfsB.

4.1.2 Methylcarboxylated compound **6**

The absence of a xenofuranone compound that is both carboxylated and O-methylated in the extracts of the *xfkB* deletion strain demonstrates that the decarboxylation takes place prior to methylation. Although this compound with carboxylation and O-methylation is not detected, the isomer **6** featuring a methylcarboxyl moiety is found in wild type strain extracts. This compound is phenylbutyrolactone II and was purified from *X. szentirmaii* cultures by former lab member Zakaria Cheikh Ali (Fig. S1). Interestingly, phenylbutyrolactone II is only observed for culture extracts when amberlite resin XAD-16 was added during cultivation (Fig. 74). It is possible that binding to the resin prevents initial decarboxylation. Subsequent methylation of the carboxyl function then prevents further decarboxylation and might explain why phenylbutyrolactone II can be observed in *X. szentirmaii* WT cultures while **5** is not observed. Van Dijk *et al.* 2018 identified the tailoring SAM-dependent methyltransferase BtyB responsible for methylation of the carboxyl function in phenylbutyrolactone II in *A. terreus* (Fig. 74).^[196] This is in contrast to the methyl group in **6** in *Xenorhabdus* which appears to be derived from another methyl donor than methionine. Possibly, methyltetrahydrofolate (MTHF) might function as the methyl group donor in this reaction since MTHF in cooperation with cobalamine is another known methyl-group provider for methyltransferases.^[200] Feeding with ¹³C labelled MTHF could be used to confirm or exclude MTHF as methyl group donor for **6**.

4.1.3 Xenofuranone methyltransferase XfsC

In a final step, xenofuranone B is methylated to yield xenofuranone A (Fig. 73). This methyltransferase had been unknown and was identified in this work based on the distinct compound profiles between WT and Δhfq strain (Fig. 75). P_{BAD_xfxA} activation in the wild type leads to production of xenofuranone A and B, while xenofuranone B is the main compound detected for P_{BAD_xfxA} activation in the Δhfq strain. Comparative proteome analysis of *X. szentirmaii* WT and Δhfq revealed four methyltransferase candidates and subsequent gene deletions confirmed XfsC to be responsible for xenofuranone methylation. The deletion of the remaining three methyltransferase candidates identified by the proteomics approach did not affect xenofuranone A production and therefrom it is postulated that XfsC is the only methyltransferase relevant in xenofuranone A production (Fig. 75).

DISCUSSION

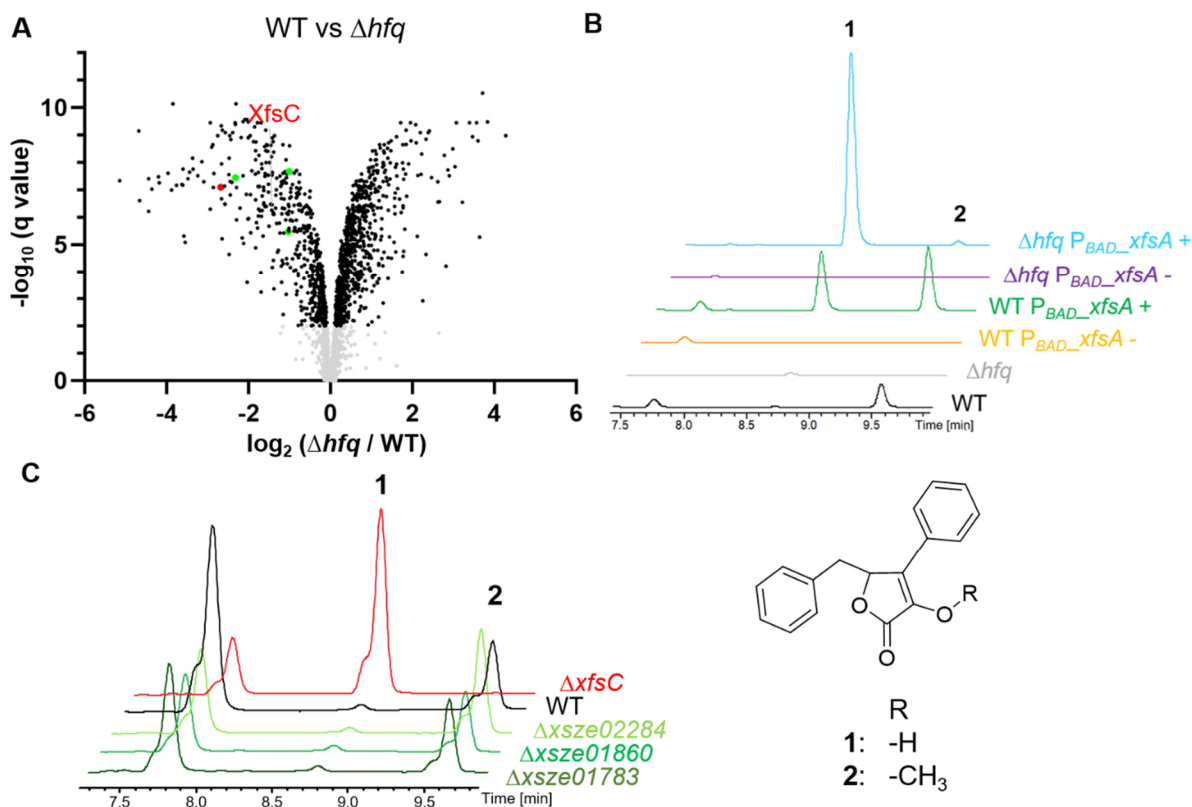


Figure 75. Xenofuranone methyltransferase identification. **A** Volcano plot comparing *X. szentirmaii* WT and *X. szentirmaii* Δhfq proteomes. XfsC is shown in red and three methyltransferase candidates are shown in green. **B** P_{BAD_xfsA} activation in *X. szentirmaii* WT and *X. szentirmaii* Δhfq . Xenofuranone B (1) is the main product in *X. szentirmaii* Δhfq P_{BAD_xfsA} . **C** Xenofuranone B (1) is the main product in *X. szentirmaii* $\Delta xfsC$.

Generally, methyltransferases are known for methylation of different substrates including nucleosides (DNA/RNA), proteins and small molecules.^[201] With respect to a partial structural similarity of xenofuranones with nucleosides a possible function of XfsC on RNA was assumed. But comparative analysis of *X. szentirmaii* WT and *X. szentirmaii* $\Delta xfsC$ on tRNA and rRNA modifications did not show any differences (Prof. Dr. Stefanie Kellner, personal communication). Several examples of methyltransferases with promiscuous substrate specificity are known.^[202] Data from this work show that XfsC methylates xenofuranone B but the possibility of further substrates cannot be excluded for XfsC.

Since the introduction of methyl groups can also alter a natural product's properties such as e.g. solubility, it would be interesting to study if the distinct production of non-methylated vs methylated xenofuranone production is used in *X. szentirmaii* in adaptation to a specific life cycle phase.^[203] Such a regulation could be mediated via Hfq to allow or suppress the action of the methyltransferase.

Furthermore, the applied proteomics-based identification of a tailoring enzyme could also be used in other examples where distinct compound profiles are observed for two different

conditions. These two different conditions are of course not limited to the comparison of two different mutant strains but might also be applicable to two different growth conditions such as growth phase, temperature, medium or other external stimuli. In case of xenofuranones with Hfq being an RNA-binding protein it was considered to study the RNA profiles but overall the proteome approach was regarded to be more straight forward as the protein level is more closely linked to the functional enzyme.

4.1.4 4-Hydroxyphenylpyruvic acid derivatives

Arabinose induction of *X. szentirmaii* P_{BAD_xfxA} led to the detection of two novel xenofuranone derivatives (**7** and **8**), which are not observed in the wild type strain. These two derivatives are each built from one phenylpyruvic acid moiety and one 4-hydroxyphenylpyruvic acid moiety, incorporated as first or second building block, respectively (Fig. 76).

The predicted α -keto acid specific code for 4-hydroxyphenylpyruvic acid and phenylpyruvic acid differs only in position 8 from cysteine to serine (XfsA) (Fig. 76). Both amino acids share a polar side chain function, which likely explains that both substrates are accepted by XfsA, but with a preference for the original substrate. Possibly, the strong P_{BAD} induction causes an overall high protein expression which allows detection of the two 4-hydroxyphenylpyruvic acid derivatives.

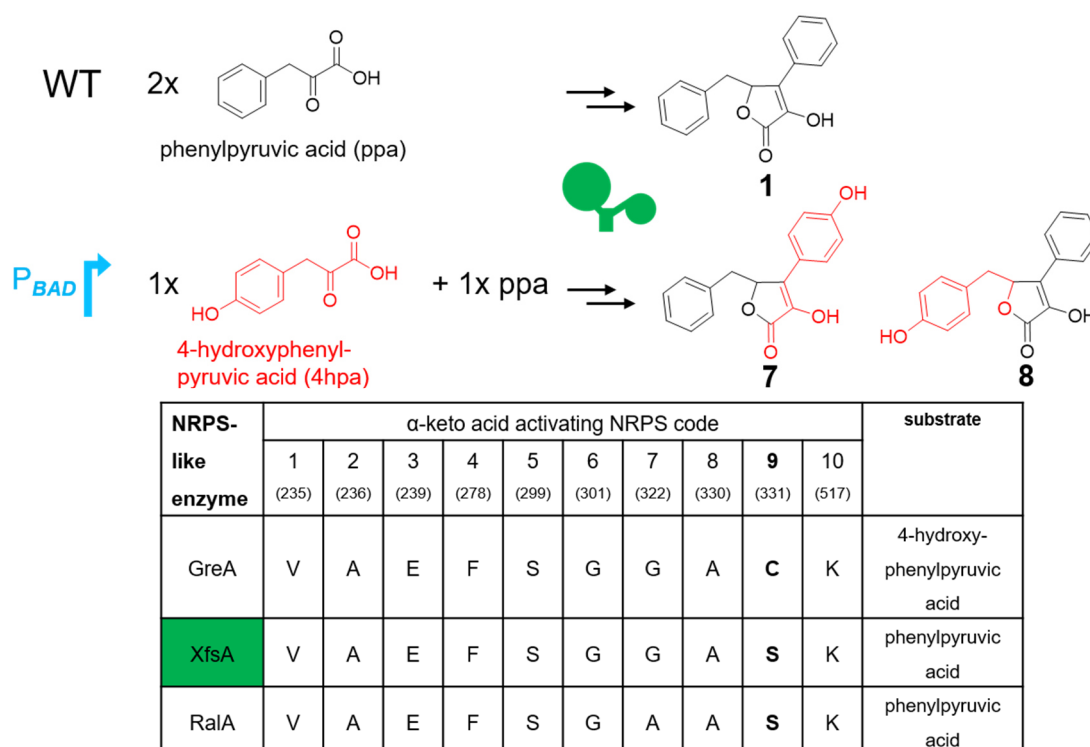


Figure 76. Xenofuranone derivatives **7** and **8** in *X. szentirmaii* P_{BAD_xfxA} . **7** and **8** are derived from condensation of 4-hydroxyphenylpyruvic acid (4hpa) and phenylpyruvic acid (ppa). The α -keto acid activating NRPS code is given for NRPS-like enzyme XfsA and compared with 4hpa activating GreA and ppa activating RalA.^[96]

DISCUSSION

A derivative built from two 4-hydroxyphenylpyruvic acids could not be detected. It might be that XfsA is not able to accept two of the larger building blocks with the 4-hydroxy-function at the same time due to sterical limitations within the active site. For p-N₃-F, that features an even bulkier substituent than 4-hydroxyphenylpyruvic acid, no incorporation could be achieved. Installation of the azide-moiety into xenofuranones could have allowed the introduction of dyes to the molecule via click-chemistry for e.g. localisation studies.^[204] Next to sterical limitations the aminotransferase reaction could be an issue for p-N₃-F incorporation. The transaminase is required for conversion of the amino acid p-N₃-F to the corresponding α -keto acid.^[205] For example, reduced production levels are observed for ralfuranones and terrequinones when aminotransferase encoding genes *ralD* and *tdiD* are genetically impaired, respectively.^[91,93] Direct feeding of p-N₃-phenylpyruvic acid would avoid the deamination step.

To circumvent these limitations, modification of the NRPS-released product could be more straightforward to introduce labels for subcellular localisation studies. Various modifications of butyrolactones appear to be installed subsequent to furanone formation.^[196,206] Moreover, a tRNA methyltransferase Trm1 was shown to transfer an alkyne side chain to its substrate from an SAM-analogue and thus the modified substrate was rendered susceptible for click-labelling.^[207] Investigations on subcellular localisation or temporal expression via click-labelling could help to identify the role of xenofuranones in the ecological context.

In summary, this work provided additional information on xenofuranone biosynthesis. The NRPS-like enzyme XfsA produces **1** from two phenylpyruvic acid moieties. In addition, XfsA is able to incorporate one 4-hydroxyphenylpyruvic acid instead of one of the phenylpyruvic acids as first or second building block but XfsA was not shown to condense two 4-hydroxyphenylpyruvic acids. The next step in the biosynthesis is the decarboxylation of **5** to yield xenofuranone B which is promoted by decarboxylase XfsB. Following decarboxylation, XfsC methylates xenofuranone B to yield xenofuranone A. The methylation step is downregulated in *X. szentirmaii* Δhfq .

4.2 Pyrrolizwilline

Activation of P_{BAD} in front of *xhpA* allowed identification, purification and structure elucidation of pyrrolizwilline. Based on mutant strains and bioinformatic predictions a biosynthetic pathway was postulated (Fig. 77).

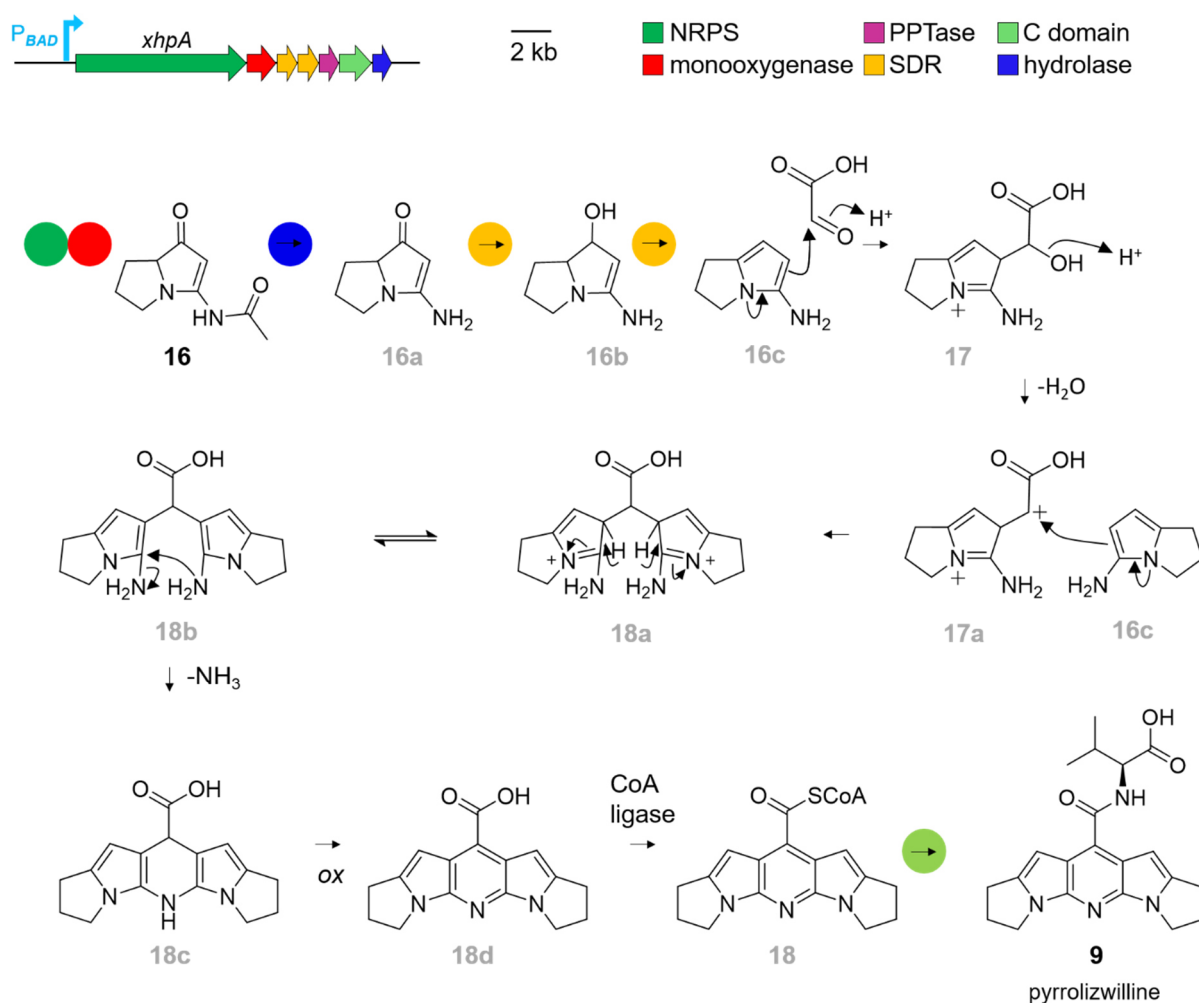


Figure 77. Postulated pyrrolizwilline biosynthesis. NRPS and monoxygenase provide the pyrrolizixenamamide-like compound **16**.^[185] XhpG hydrolyses the acyl moiety followed by the two subsequent two electron reductions of the carbonyl function by XhpC and XhpD. The pyrrolizidine moiety **16c** is extended by glyoxylic acid and following carbenium ion (**17a**) formation dimerises with a second pyrrolizidine moiety **16c** to form **18**. Deamination and ring oxidation yield **18d** which is thioestered to CoA by the action of a CoA ligase. XhpF catalyses the L-valine condensation step to yield pyrrolizwilline.

4.2.1 Two pyrrolizidine core units and the roles of XhpA and PxaA

Pyrrolizwilline exhibits two pyrrolizidine cores that resemble the pyrrolizidine moiety produced during pyrrolizixenamamide biosynthesis. Pyrrolizixenamides are produced by the bimodular NRPS PxaA and a Baeyer-Villiger monoxygenase PxaB.^[171] XhpA and XhpB show the same organisation as PxaA and PxaB. Next to the activated *xhp* BGC, *X. hominickii* provides another BGC that is homologous to *pxa* BGC. When the *pxa* BGC is deleted, pyrrolizwilline production is still detectable which suggests that *pxa* BGC is not crucial for pyrrolizwilline biosynthesis (Fig. 78). But since the pyrrolizwilline production levels are reduced in *X. hominickii* Δ *pxaAB* P_{BAD_xhpA} compared to the P_{BAD_xhpA} activation in the wild type it can be speculated that PxaAB activity can promote pyrrolizwilline production by supply of pyrrolizidine units. To test

DISCUSSION

this and to exclude that decreased production levels are caused by other effects than $\Delta pxaAB$, genomically integrated complementation of $\Delta pxaAB$ is required. The finding that P_{BAD_pxaA} activation did not lead to the detection of compound production weakens the argument of PxaAB contribution to pyrrolizwilline production. Heterologous expression of *X. hominickii* $pxaAB$ could be another method to study this BGC.

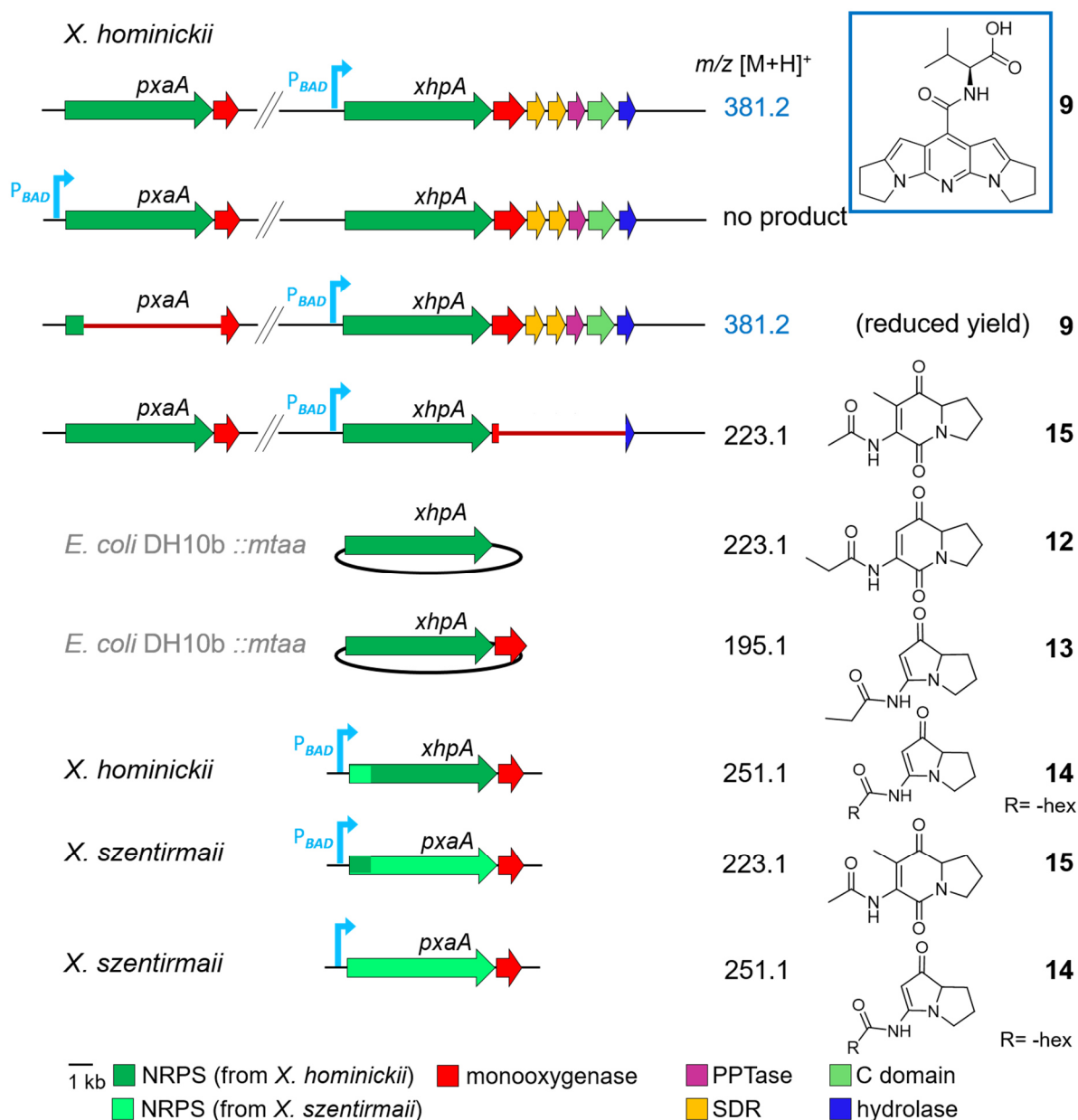


Figure 78. Overview of detected compounds from mutant strains and heterologous expression related to the *xhp* BGC. *X. hominickii* exhibits the *pxa* and the *xhp* BGC and deletion of *pxaAB* reduces pyrrolizwilline production. Heterologous expression of XhpA and XhpAB in *E. coli* results in the production of **12** and **13**, respectively. Activation of *pxaAB* in *X. szentirmaii* results in production of pyrrolizixenamide B (**14**). $C_{starter}$ domain exchanges lead to production of **14** and **15**. **15** is also detected in *X. hominickii* $\Delta xhpB-G$ P_{BAD_xhpA} . *X. szentirmaii* *pxaA* is depicted in light green. Red horizontal line indicates deleted region.

Heterologous expression of *xhpA* in *E. coli* confirmed that XhpA produces **12** from an acyl unit, serine and proline in analogy to PxaA in pyrrolizixenamide biosynthesis.^[185] In accordance, **12**

features the m/z 167.1 $[M+H]^+$ fragment ion. Contradictory, in *X. hominickii* $\Delta xhpB-G$ P_{BAD_xhpA} compound **15** is detected which shows the m/z 181.1 $[M+H]^+$ fragment ion. This is $\Delta m/z$ 14 larger than the pyrrolizidine characteristic fragment ion and the deduced sum formulas indicate a $-CH_2$ difference. The structure presented for **15** matches the observed MS data and could arise from the incorporation of threonine as a similar structure is detected for legonindolizidine, an intermediate in the legonmycin biosynthesis.^[187] The pyrrolizidine alkaloid legonmycin is built from proline, threonine and a branched fatty acid.^[187] But since amino acid specificities according to antiSMASH, the heterologous expression results and the results from $\Delta xhpG$ support that XhpA incorporates serine in analogy to PxaA, it is suggested that the methyl group in **15** is not caused by threonine incorporation. Possibly the methyl group is derived from methanol which was used as extraction solvent. **15** is also observed in *X. szentirmaii* P_{BAD_pxaA} in which the $C_{starter}$ domain was exchanged with *xhpA*. C domains are known to impact A domain specificity but again both XhpA and PxaA are mostly shown to incorporate serine and both show the same Stachelhaus code.^[82] Detection of **15** in the $C_{starter}$ domain engineered *X. szentirmaii* P_{BAD_pxaA} strain is also surprising as a product with the pyrrolizidine core unit (fragment ion m/z 139.1 $[M+H]^+$) was expected in this strain where PxaB is present. Possibly the methylation of the indolizidine core prevented further processing by the monooxygenase PxaB in the $C_{starter}$ domain engineered *X. szentirmaii* P_{BAD_pxaA} strain. XhpA incorporates shorter acyl side chains in comparison to pyrrolizixenamide producing PxaA. For XhpA, two different acyl moiety lengths, namely C_2 and C_3 carbon length, were found when expressed in *Xenorhabdus* and *E. coli*, respectively. Possibly the observed different phylogenetic groups for XhpA and PxaA based on the multiple sequence alignments correlate to the chain length of the incorporated acyl substrates. The difference for the phylogenetic groups is supported by the pairwise sequence alignment of XhpA with *X. szentirmaii* PxaA which identified the $C_{starter}$ domain as region with the lowest similarity between the two enzymes. The *X. bovienii* PxaA groups together with XhpA and induced *X. bovienii* P_{BAD_pxaA} produces **11** which is pyrrolizixenamide A, the shortest acyl rest to be incorporated in pyrrolizixenamides in *X. stockiae*.^[185] Still, more investigations on the *Xenorhabdus pxa* BGCs including $C_{starter}$ swaps are required to verify this hypothesis. Following incorporation of the acyl unit and serine, the second C domain in XhpA which belongs to the group of modifying C domains mediates the dehydration reaction of serine required for cyclisation.^[81,208] After release of the 5,6-bicyclic intermediate by XhpA, Baeyer-Villiger monooxygenase XhpB catalyses oxidative ring expansion which leads to hydrolysis and ring contraction as described in pyrrolizixenamide biosynthesis.^[185] The production of **11** from heterologous expression of *xhpAB* in *E. coli* support that XhpB operates as suggested in analogy to PxaB.

DISCUSSION

4.2.2 Gene deletions and further biosynthesis steps

Additional gene deletions were conducted to investigate pyrrolizwilline biosynthesis steps following formation of the pyrrolizidine core units. The *xhp* BGC encodes a PPTase and in general PPTases are required for post-translational modification, hence activation, of T domains. The deletion of *xhpE* led to the detection of decreased pyrrolizwilline levels in comparison to the wild type P_{BAD_xhpA} activation (Fig.79). This shows that XhpE is not essential for pyrrolizwilline formation but that XhpE contributes to the pyrrolizwilline production when expressed. In this case, complementation of $\Delta xhpE$ should restore WT production levels and should be investigated in future work. Furthermore, it could be interesting to study whether XhpE can influence the production of other natural products in *X. hominickii*, e.g. by expressing XhpE in *X. hominickii* $\Delta ngrA$ or *X. hominickii* $\Delta xhpE \Delta ngrA$. As no pyrrolizwilline production is detected in the strain where PPTase encoding *ngrA* is deleted, NgrA likely is the essential PPTase in pyrrolizwilline biosynthesis. It is possible that NgrA is relevant for PPant loading to the T domains in both PxaA and XhpA.

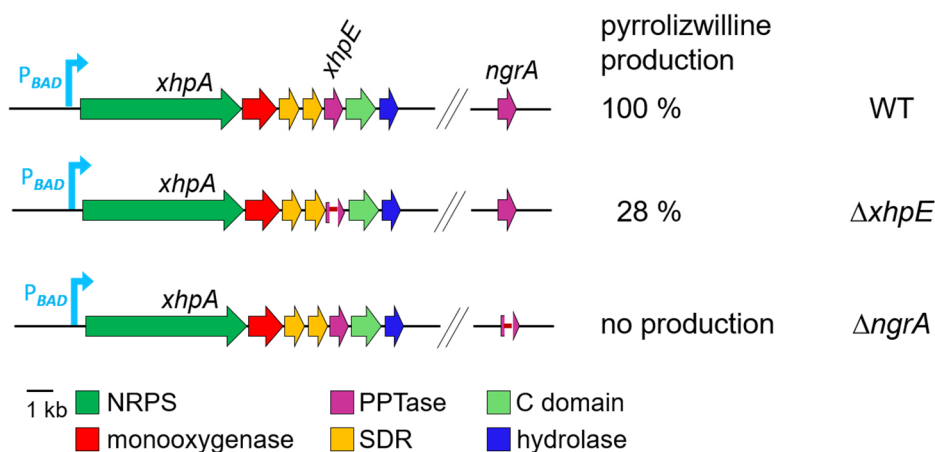


Figure 79. Overview of pyrrolizwilline production in *X. hominickii* WT and PPTase deletion mutants upon P_{BAD_xhpA} activation. Pyrrolizwilline is produced in *X. hominickii* WT and in reduced yield in *X. hominickii* $\Delta xhpE$ when compared to the wild type. No production is detected in *X. hominickii* $\Delta ngrA$. Red horizontal line indicates deleted region.

Deletion of the α,β -hydrolase encoding gene *xhpG* led to the detection of the intermediate **16** and indicates that XhpG is responsible for hydrolysis of the acyl side chain of the pyrrolizixenamide-like compounds. In addition, the observation of **16** suggests that the hydrolysis must occur prior to carbonyl reduction by the two SDRs (Fig. 80). As proof of concept, feeding of the pyrrolizidine core without acyl rest (**16a**) should restore pyrrolizwilline production in *X. hominickii* $\Delta xhpG P_{BAD_xhpA}$ (Fig. 80). Deletion of the SDR encoding genes *xhpC* and/or *xhpD* was not successful. But it is proposed that XhpC catalyses the ketoreduction followed by dehydration catalysed by XhpD which is consistent with antiSMASH predictions

for these enzymes. To support this, feeding of the chemically synthesised precursor **16c** to *X. hominickii* P_{BAD_xhpE} could be conducted and should result in pyrrolizwilline production (Fig. 80).

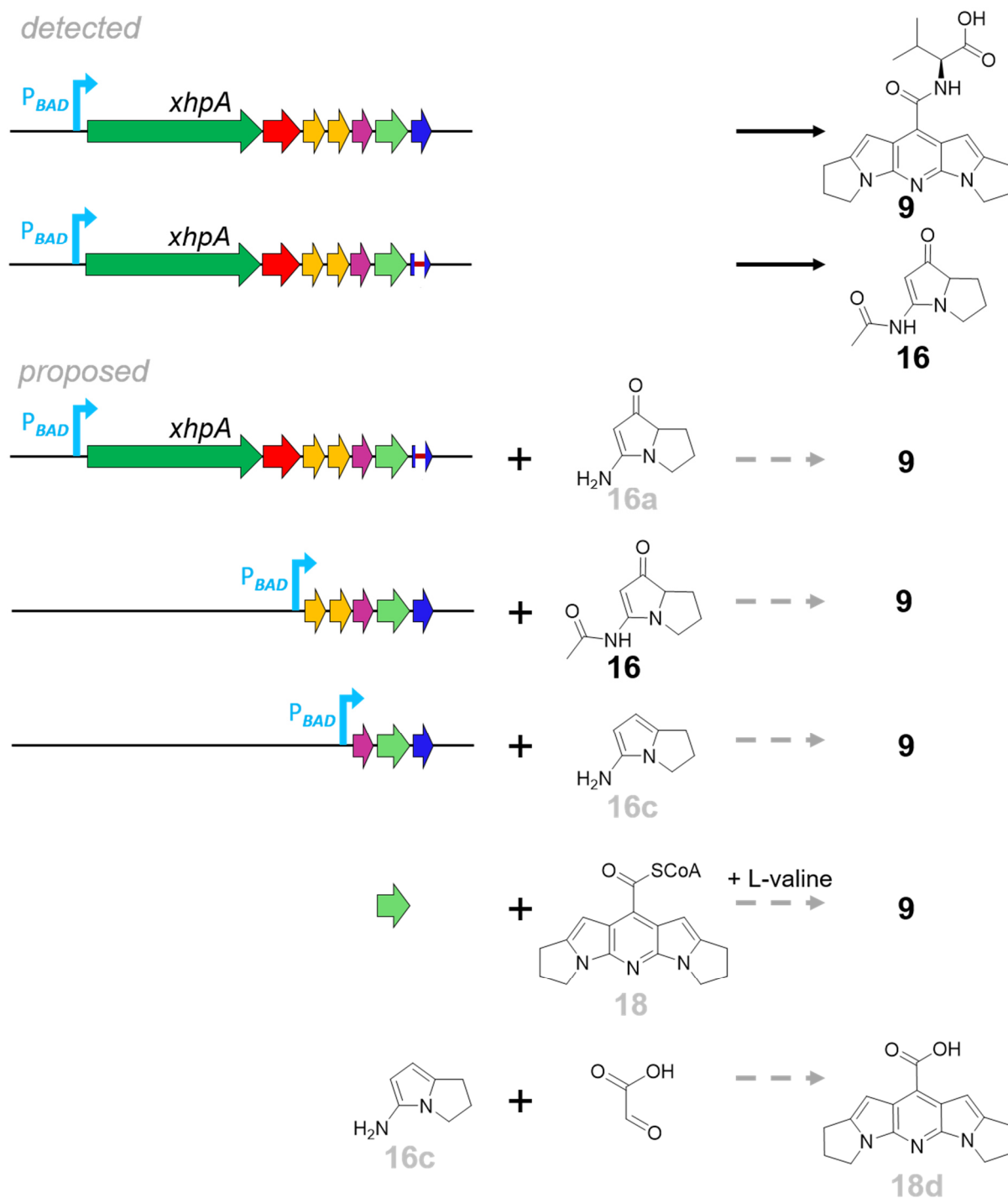


Figure 80. Detection of pyrrolizwilline intermediate **16** in *X. hominickii* $\Delta xhpB-G$ P_{BAD_xhpA} and proposed complementation experiments to restore pyrrolizwilline (**9**) production. Compounds numbered in black were detected in this work, compounds numbered in grey are proposed intermediates.

This work suggests that one of the reduced pyrrolizidine cores is extended by glyoxylic acid and then dimerises with another reduced pyrrolizidine core unit (**16c**) (Fig. 77). This is

DISCUSSION

postulated considering that a similar dimerisation mechanism was proposed for the formation of the dimeric diboheamine A. Diboheamine A is shown to be built from two molecules boheamine A (**37**) of which one is extended with formaldehyde in a Baylis-Hillman-reaction to yield **36** (Fig. 81).^[209] Protonation of the primary alcohol function is followed by a nucleophilic substitution to form diboheamine A (Fig. 81).^[209] The formation of diboheamine A occurs non-enzymatic.^[209] According to this, it should be tested if formation of **18d** can be obtained from **16c** and glyoxylic acid without any enzyme (Fig. 80).

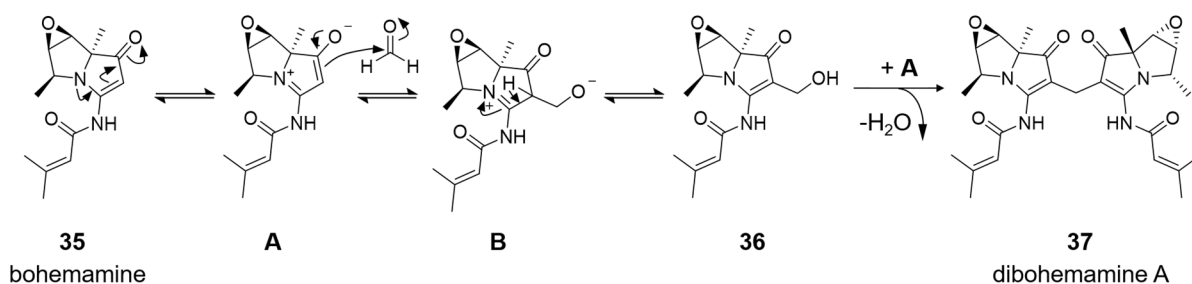


Figure 81. Proposed dimerisation mechanism for diboheamine formation. In a Baylis-Hillman reaction the electron-rich nucleophile in **A** attacks formaldehyde to form the intermediate **B**. Following protonation, the primary alcohol function is substituted with a second equivalent of **A**. Proposed by and taken from Fu *et al.* 2016.^[209]

The final step in the postulated pyrrolizwilline biosynthesis is the condensation of L-valine catalysed by the free-standing condensation domain XhpF. Since C domains usually require donor substrates which are either T-domain bound or thioestered to coenzyme A, it is suggested that **18d** is activated with CoA by the action of a CoA ligase to allow the final condensation step.^[52] Alternatively, if glyoxylyl-CoA instead of glyoxylic acid was used, no additional CoA ligation was needed to form **18**. While glyoxalate is produced by the isocitrate lyase within the glyoxalate cycle in bacteria, no evidence is found on the existence of glyoxylyl-CoA, as to my knowledge.^[210] Therefore, the postulated biosynthesis depicted in Fig. 77 uses glyoxylic acid as extender unit.

In addition to the main proposal (Fig. 77), two further alternative biosynthesis routes are presented in Fig. 82. Route B also uses glyoxylic acid to extend **16c** but instead of formation of the carbenium ion, the elimination of water generates a double bond. Next, Diels-Alder reaction ([4+2] cycloaddition) leads to ring formation between the two pyrrolizidine moieties. Similar to the postulated biosynthesis (Fig. 77), deamination, CoA ligation and L-valine condensation are needed to build pyrrolizwilline. Route A is based on a different bridging unit, namely oxalyl-CoA, to extend the reduced pyrrolizidine core **16c**. This leads to the formation of **39** whose α -carbonyl group is attacked from the nucleophile of the second pyrrolizidine core **16c** (Fig. 81). Subsequently, ring closure, deamination, CoA ligation and condensation with L-valine yield pyrrolizwilline as described above (Fig. 82).

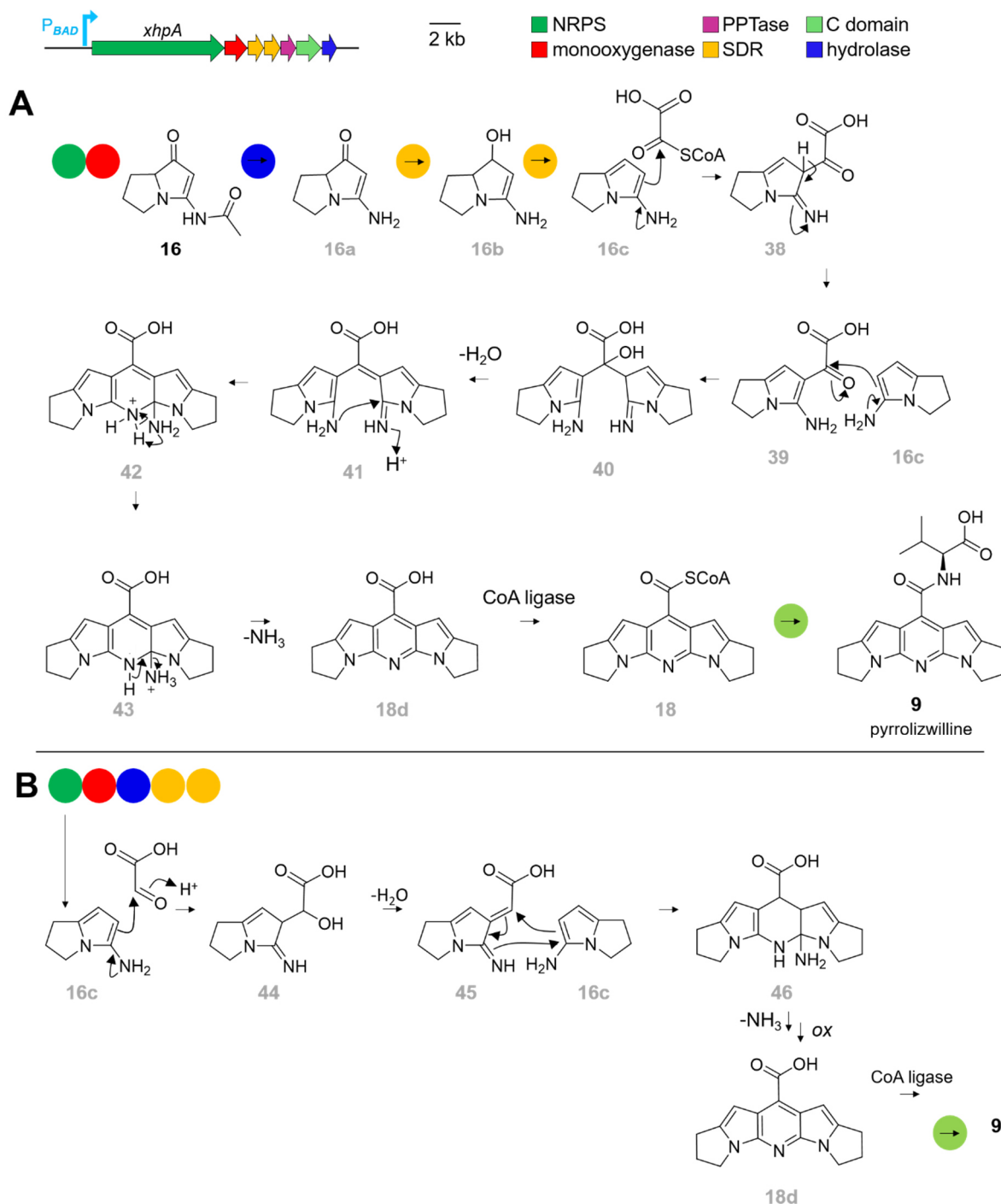


Figure 82. Two alternative biosynthesis routes for pyrrolizwilline biosynthesis. Route **A** uses oxalyl-CoA as building block to generate **38**. Nucleophilic attack of the α -carbonyl group establishes the C-C bond between **16c** and **39**. Ring closure, deamination, CoA ligation and condensation yield pyrrolizwilline (**9**). In route **B** glyoxylic acid is used as C₂-extender unit and, upon elimination of water, [4+2]-cycloaddition leads to formation of **46**. Deamination, building block activation with CoA and condensation with L-valine lead to pyrrolizwilline formation. Compounds numbered in black were detected in this work, compounds numbered in grey are proposed intermediates.

It is further proposed that XhpF catalyses the condensation between L-valine and **18** as final step in pyrrolizwilline biosynthesis. However, in *X. hominickii* $\Delta xhpF$ P_{BAD_xhpA} , no compound production was detected that could be linked to the BGC activation. It remains elusive whether

DISCUSSION

the L-valine is bound to a carrier protein or coenzyme or if it is available as free substrate. Usually, C domains catalyse the condensation of PCP-bound substrates but also accept free amines as acceptor substrate, e.g. norspermidine for VibH.^[75,211] No single T domain or freestanding A-T domains were identified in the antiSMASH analysis of *X. hominickii*. In case of CoA-bound substrate presentation, a CoA ligase might be required. CoA ligases are shown to produce aminoacyl-CoA species.^[212] It would be intriguing to study whether purified XhpF shows condensation of the proposed precursor **18** with L-valine or L-valine-SNAC *in vitro* (Fig. 80).

No pyrrolizwilline or intermediates, e.g. pyrrolizixenamide, were observed for P_{BAD_xhpA} activation in *X. hominickii* Δhfq . The absence of pyrrolizwilline in this Δhfq strain could be explained with respect to the above mentioned possible need of additional enzymes that are not present in Δhfq . However, the absence of intermediates of intermediates, e.g. **16** produced from XhpA and XhpB, is unexpected as pyrrolizixenamides are found to be produced upon P_{BAD} activation in front of *pxaA* in *X. szentirmaii* Δhfq .^[46] Complementation of the *hfq* deletion could be conducted to verify that the absence of pyrrolizwilline intermediates is dependent of Hfq.

4.2.3 Phylogenetic analysis and bioactivity

The pyrrolizwilline BGC was not found in other *Xenorhabdus* strains but *pxa* BGC organisation encoding for an NRPS and monooxygenase are observed in several *Xenorhabdus* strains. Thus, pyrrolizwilline is a characteristic metabolite for *X. hominickii* among *Xenorhabdus* species identified so far and possibly hints at a specialised function for this specific strain.

TetR family transcriptional regulators (TFRs) are known transcriptional regulators with an N-terminal DNA binding domain and a larger C-terminal domain, often involved in ligand binding.^[213] The majority of TFRs are type I TFRs being transcribed in opposite direction to the target gene and are predicted to be likely in a regulatory relationship with the target gene when the distance between TFR encoding gene and target gene is below 200 bp.^[214] The opposite transcription direction applies to *xhpA* and *xhom00317* and both genes are separated by 296 bp intergenic region which is longer than 200 bp mentioned above but according to Cuthbertson & Nodwell distances longer than 200 bp are not excluded, only less common.^[214] TFRs can function as activators and repressors and both as the same protein.^[214,215] Still, neither deletion nor plasmid-based expression of the TetR encoding gene did affect pyrrolizwilline production. Additionally, it might be that the required ligand to allow TetR-like protein binding is not present in the cell under laboratory conditions.

The observed inhibition zones of *M. luteus* growth showed no difference between the non-induced and induced P_{BAD_xhpA} sample indicating that pyrrolizwilline shows no bioactivity towards *M. luteus* nor any of the other tested strains, *S. cerevisiae*, *B. subtilis* and *E. coli*. In

plants, a variety of pyrrolizidine alkaloids are described and many of them show DNA-intercalating properties via their necine base causing genotoxic affects.^[216] An example for bacterial pyrrolizidine alkaloids are pyreudiones which are described to protect the producing *Pseudomonas* strain from amoebal predators.^[217] For pyrrolizwilline a possible bioactivity or ecological function remains to be determined.

Summing up, pyrrolizwilline was identified as product of *xhp* BGC by P_{BAD} activation and compound purification allowed structure elucidation. NRPS XhpA and Baeyer-Villiger monooxygenase XhpB contribute to formation of the pyrrolizidine core units and although *pxaAB* are not essential, their encoded enzymes could provide pyrrolizixenamides. Next, XhpG hydrolyses the acyl side chain in the pyrrolizixenamide-like structures followed by two subsequent two electron reductions of the carbonyl function catalysed by the two SDRs XhpC and XhpD. Proposedly, the generated pyrrolizidine moiety is extended with glyoxylic acid and then dimerises with another reduced pyrrolizidine core unit. Following activation by a CoA ligase, XhpF finally incorporates an L-valine to yield pyrrolizwilline. The chemical structure and the BGC organisation combined with a rare occurrence emphasise the interesting character of pyrrolizwilline and rise interest in the understanding of its biological function.

4.3 Glyoxpeptide

In previous work, glyoxpeptide was detected upon promoter activation in *X. hominickii* Δhfq P_{BAD_xgsA} and following structure elucidation the corresponding biosynthesis was postulated.^[153,154] This work aimed for a more detailed investigation of glyoxpeptide production including identification of the previously observed glyoxpeptide derivative and further characterisation of the MonoOx domain (Fig. 83).

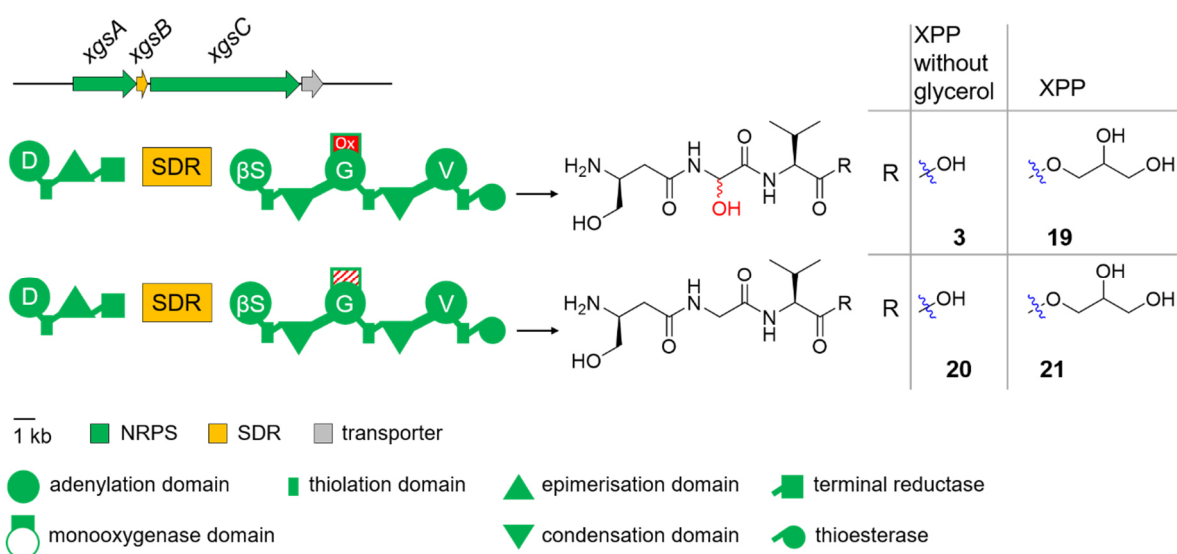


Figure 83. Glyoxpeptide (**3**) production from XgsA-C and production of the non-hydroxylated derivative (**20**) upon deletion of the XgsC_MonoOx domain. Glycerol derivatives **19** and **21** are detected in XPP medium which contains glycerol.

DISCUSSION

4.3.1 Glycerol derivative

The previously detected glyoxpeptide derivative is thought to be derived from glycerol addition based on reverse feeding experiments and altered medium composition. It has not been investigated whether the glycerol derivative is formed enzymatically. Lipases are known to esterify glycerol with free fatty acids and it cannot be excluded that in *X. hominickii* enzymes are involved in formation of the glycerol derivative.^[218] However, glyoxpeptide is likely the main compound produced by the investigated BGC. The glycerol derivative may be regarded as side product which is not relevant in the natural background as glycerol is an ingredient of the culture medium. In culture medium without glycerol, only glyoxpeptide production was observed (Fig. 83). It has to be noted that in the prepared medium in which glycerol was omitted, glucose was added to the medium. Glucose is a known repressor of the P_{BAD} system by reducing cAMP levels and in turn preventing the catabolite activator protein (CAP) from P_{BAD} activation^[219]. Apparently in this work glucose levels were low enough to still allow P_{BAD} activation and glyoxpeptide production.^[220]

4.3.2 MonoOx domain characterisation

Deletion of the MonoOx domain encoding region was successful and resulted in the production of a functional NRPS producing the non-hydroxylated product **20** (Fig. 83). In contrast, no product formation was detected for the induced *ptt* BGC where the MonoOx domain encoding region was genomically integrated into *pttB*. The genomic integration was confirmed but protein production was not checked. It might be possible that the position of the genetic insertion was not suitable and subsequently prevented proper protein folding. In this work only one attempt for insertion was conducted and maybe a selection of different fusion points should be addressed in the future. One promising approach could be to remain the short sequence encoding amino acid sequence CSS which was replaced in this work by the MonoOx domain encoding sequence (Fig. S11). In principle, it was shown before that M domains can be embedded successfully in A domains in between core motifs A8 and A9, as for example the serine O-methylating M domain from KtzH (A8-A9) was inserted into ecchinomycin NRPS module Ecm6.^[149] For successful integration of the MonoOx domain in the assembly line, limitations from the acceptance of downstream processing domains, e.g. C domains which are known to exhibit a gate keeping function, should be considered.^[76,82] Consequently, aside from testing alternative insertion sites in between A8/A9, the insertion of the α -hydroxyglycine incorporating module from XgsC could be investigated (Fig. 84). This could be achieved by using the established XU or XUC concept as presented in Fig. 84.^[76,145] Especially the use of the XUC unit could help to circumvent downstream substrate acceptance issues as the downstream C domain specificity can be neglected in the XUC approach, in contrast to the XU concept.^[145]

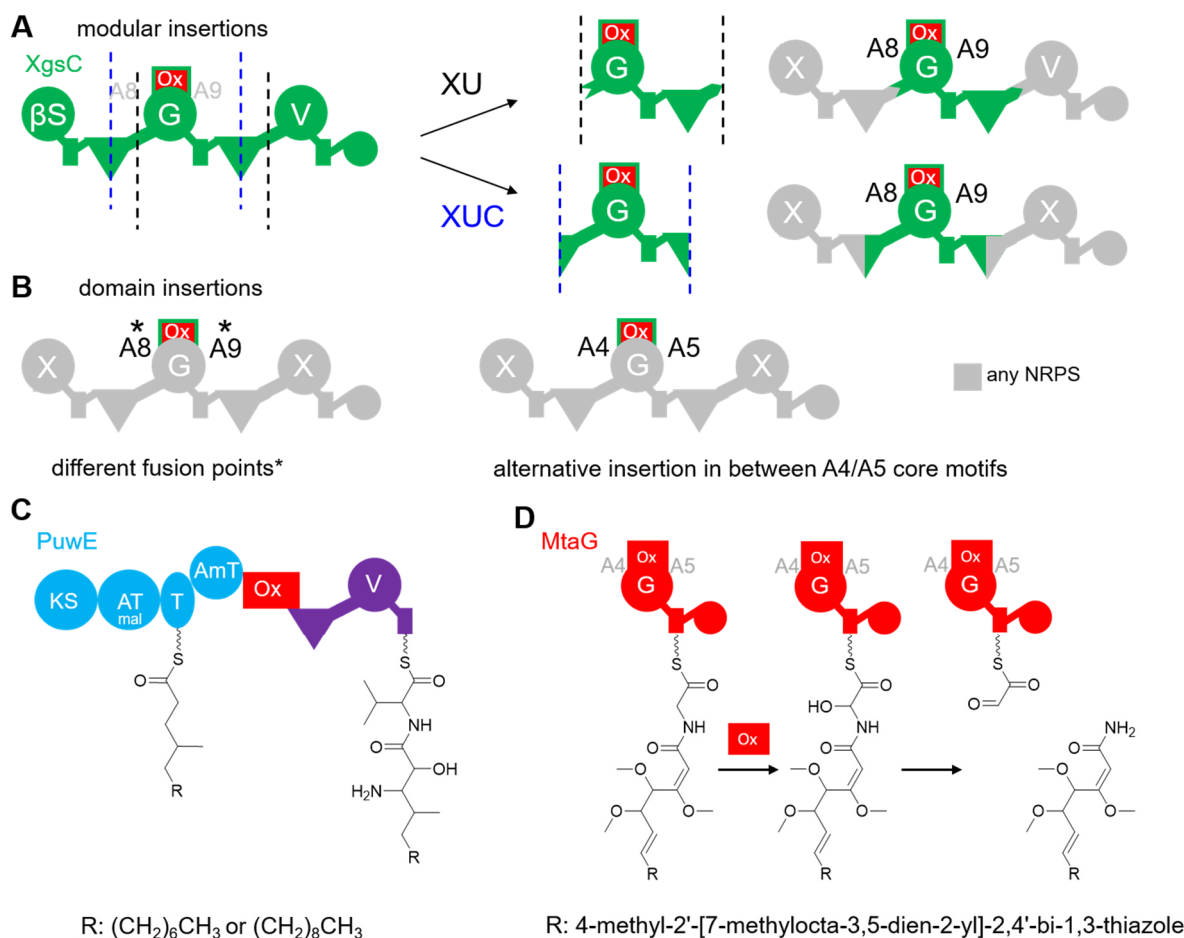


Figure 84. Overview of possible insertion sites for XgsC_MonoOx domain. **A** Modular insertion of XgsC_MonoOx domain based on the XU or XUC concept.^[145] Dashed lines indicate borders of XU (black) and XUC (blue). **B** Insertion of XgsC_MonoOx domain in between core motifs A8/A9 or A4/A5. **C** PuwE_MonoOx domain in puwainaphycin biosynthesis is inserted in between an aminotransferase (AmT) domain and a C domain.^[190] **D** MtaG_MonoOx domain is inserted in between core motifs A4/A5 in the A domain of the terminal module in myxothiazol biosynthesis. Following α -hydroxylation of glycine, the terminal amide myxothiazol is released while the glyoxal remains T domain bound.^[69]

The purification of the MonoOx domain did not provide further information on cofactor dependence, let alone structural information. Monooxygenases usually depend on NAD(P)H, a flavin cofactor and molecular oxygen. External flavoprotein monooxygenases require reduced NAD(P)H for reduction of the flavoprotein. Van Berkel and colleagues classified six types of external flavoprotein monooxygenases based on sequence and structure data.^[221] Thereafter type C monooxygenases comprise, among others, luciferases and alkanesulfonate monooxygenases and feature one or two monooxygenase components with a TIM barrel fold.^[221] They use reduced FMN as coenzyme which is provided by an accompanying reductase that accepts NADH and/or NADPH.^[221] Sky39, sharing 46.7 % sequence identity with XgsC_MonoOx domain, was identified to be a FAD-dependent free-standing monooxygenase, accompanied by the reductase Sky40 and supposedly exhibits dimeric catalytic behaviour.^[188,222] The structure for PuwE_MonoOx domain also shows homodimer

DISCUSSION

formation and similar to PuwE, NpnB_MonoOx domain from the nostophycin BGC follows an aminotransferase domain in a PKS-NRPS hybrid enzyme (Fig. 82).^[223] The dimeric character of both Sky39 and PuwE is interesting as originally some time ago, NRPSs were proposed to function as monomers while PKSs are assumed to be dimeric enzymes.^[224] It was suggested that in NRPS-PKS hybrid enzymes the PKS dimer could harbour the NRPS in a monomeric loop.^[224] However, the NRPS VibF was proposed to function as dimer.^[225] The overall matter of whether NRPSs are monomers or form dimers is of debate and require further investigation. Furthermore, it cannot be excluded that the observed homodimer of PuwE is an artefact from crystallisation conditions. For XgsC_MonoOx domain the generated data do not allow to exclude dimeric operation. In addition, more information on natural product biosynthesis luciferase-like monooxygenases is required to decipher the flavin cofactor and (in)dependence on a reductase. Albeit, two rare examples of monooxygenases are reported which do not rely on a flavin cofactor: quinone MonoOx YgiN and PKS MonoOx ActVa-Orf6.^[221] Structural evidence could also shed light on the stereochemistry of the hydroxyl function.

To the best of my knowledge, XgsC_MonoOx domain is the first example of this domain to be found in position A8/A9 within the A domain. MtaG and MelG possess glycine α -hydroxylating MonoOx domains, but embedded in position A4/A5 of their respective A domain (Fig. 84).^[69] For myxothiazol biosynthesis it was proposed that hydroxylation of glycine leads to cleavage of the carbinolamide releasing the myxothiazol peptide with its terminal amide function while the glyoxalate remains tethered to the T domain (Fig. 84).^[69] It remained elusive whether the cleavage during myxothiazol biosynthesis was spontaneous or enzyme catalysed.^[69] It would be interesting to study if the position of the MonoOx domain insertion (MtaG: A4/5 and XgsC: A8/A9) has an impact on the stability of the hydroxylated peptide for further processing by the NRPS and thus product outcome. In thiocoraline (*tio*) biosynthesis two different M domains, embedded in motifs A2-A3, and A8-A9 are observed.^[226] TioS_M domain is embedded in between A2-A3 motifs and catalyses *N*-methylation of cysteine while TioN_M domain, as most methyltransferases embedded between A8-A9 motifs, is responsible for *S*-methylation of cysteine.^[226] Engineering experiments with these two M domains even allowed the generation of a di-interrupted A domain.^[227] XgsC_MonoOx domain could be a useful addition to the pool of available domains for interrupted A domain engineering.

To elucidate if or how the insertion position (A4/A5 vs A8/A9) in the A domain influences MonoOx domains, XgsC_MonoOx domain could be inserted in between motifs A4/A5 within the XgsC_A2 domain while deleted in its original position in A8/A9. Alternatively, the replacement of the MtaG_MonoOx domain with XgsC_MonoOx domain could be attempted. With respect to the sequence similarity of 50 % between XgsC_MonoOx and PuwE_MonoOx domain, it could be tested whether PuwE_MonoOx domain can be exchanged for XgsC_MonoOx in puwainaphycin biosynthesis. This is especially interesting as PuwE is not

integrated in the A domain but instead found in between an aminotransferase and a C domain, bridging a PKS and NRPS module (Fig. 84).

Furthermore, XgsC_MonoOx domain could be applied to generate peptides with a C-terminal amide. During myxothiazol biosynthesis, it is assumed that hydrolysis of the carbinolamide results in the release of the amide product myxothiazol while the glyoxal remains T domain bound (Fig. 84).^[69] In addition, ring opening is reported for cyclic skyllyamycins at the position of the α -hydroxyglycine upon exposure to basic conditions.^[188] This cleavage provides an interesting application to split peptides and simultaneously generate both α -keto aldehyde peptides and C-terminal amide peptides (Fig. 85). To facilitate detection of the expected cleavage products, XgsC could be extended at the N- and/or C-terminal end to produce overall larger peptides with in turn longer split products. As indicated above, this could be achieved by NRPS engineering with XUC units and different building block combinations with and without original upstream and downstream modules of XgsC_module 2 (Fig. 85).

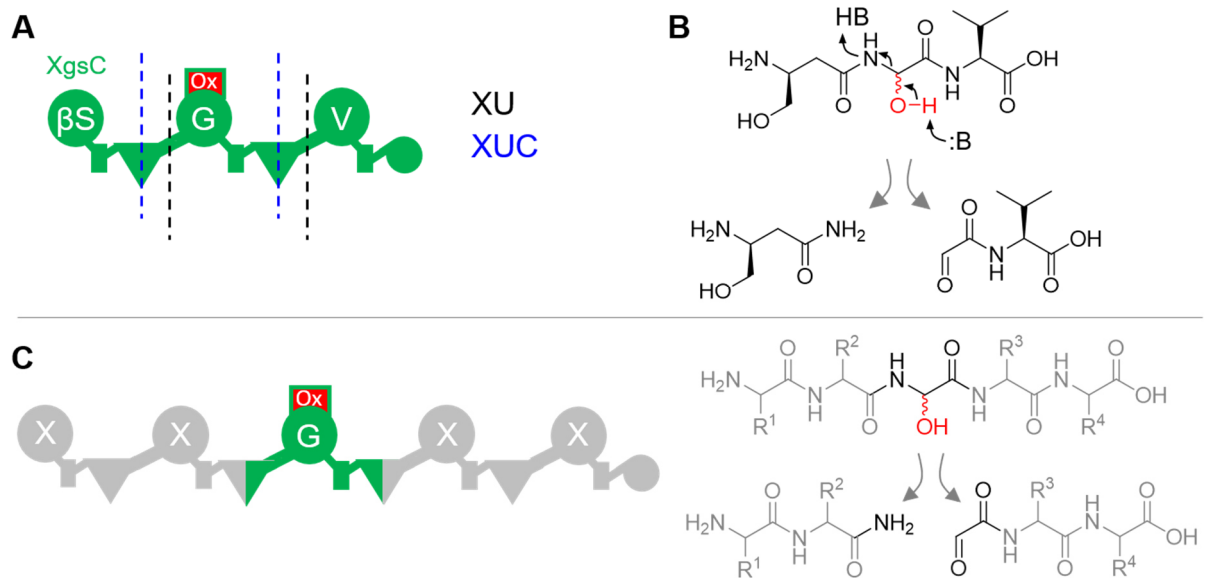


Figure 85. Proposed cleavage of glyoxpeptide and application using NRPS engineering. **A** XU and XUC borders are indicated by dashed lines.^[145] **B** Peptide splitting to generate peptide split products with C-terminal amide function and N-terminal glyoxal. **C** Schematic representation of a hypothetical engineered NRPS composed of five modules with module 3 inserted as XUC from XgsC.

C-terminal amide peptides are described to show increased bioactivity and higher metabolic stability.^[228] In mammalian cells, peptidylglycine- α -amidating monooxygenase (PAM) represents a common strategy to provide C-terminal amide peptides, but is not found in the same way in bacteria.^[229] However, for the production of C-terminal amide peptides, expression of human PAM enzymes were conducted in *E. coli*.^[230] Peptide keto aldehydes generated from split of the glyoxpeptides are also intriguing as peptide keto aldehydes are described to be specific and reversible proteasome inhibitors.^[231,232]

DISCUSSION

4.3.3 Phylogenetic distribution

From the BlastP analysis, glycine activating A domains with an embedded MonoOx domain have been found in different BGCs from other proteobacteria and one terrabacterium (*Chloroflexi bacterium*). This shows that this domain together with its assumed split mechanism is not exclusive to *Xenorhabdus* but more common in NRPSs in general.

4.3.4 Transporter XgsD

Apart from *X. hominickii*, the glyoxpeptide BGC was observed in three other *Xenorhabdus* strains, always exhibiting the presence of a transporter encoding gene. Deletion of this gene in *X. hominickii* did not change the growth behaviour hinting that glyoxpeptide can be transported out of the cell by other mechanisms and is not toxic to the producer itself as was already investigated by other lab members in heterologous expression experiments in *E. coli*.^[153,154]

In summary, the glyoxpeptide derivative is confirmed to be derived from glycerol, a culture medium component. To the best of my knowledge, the XgsC_MonoOx domain is the first reported example of this domain type to be present in between core motifs A8 and A9 within the A domain.

4.4 Xildivaline

Previous work identified NRP-PK hybrid products from promoter activation in *X. hominickii* Δhfq P_{BAD_xisA} and proposed a biosynthesis.^[153] This work aimed at further characterisation of the biosynthesis by the construction of gene deletion mutants in the producer strain. Thus, additional derivatives were detected, a different position for the *N*-methylation in xildivalines is proposed and two alternative routes for xildivaline biosynthesis are discussed.

4.4.1 Three main xildivalines

Among the three previously described compounds are the two stereoisomers **26a** and **26b** with different stereospecificity in their hydroxyl function and **27** as result of spontaneous elimination of water in this position (Fig. 86). It is assumed that the XisB_KR domain is responsible for the ketoreduction and thus the stereocontrol of the hydroxyl function. The protocol according to Keatinge-Clay discriminates between type A and type B KR's installing either *S*- or *R*-hydroxyl groups.^[114] The two described ketoreductase types allow the substrate to enter the active site either from the left side via a conserved tryptophan (type A) or the right side via an LDD motif (type B) determining the stereo outcome in a *S*-hydroxyl group (type A) or *R*-hydroxyl group (type B).^[114] XisB_KR lacks the LDD motif which according to Keatinge-Clay indicates XisB_KR to be type A. XisB_KR exhibits a leucine instead of the type A characteristic conserved tryptophan but in a few examples a deviation from the conserved tryptophan is observed for type A KR's, e.g. F instead of W in Rif_KR10.^[233] With respect to the

detection of **26a** and **26b**, XisB_KR likely catalyses the ketoreduction and shows a higher affinity to produce the *S*-hydroxyl function. Therefore, this work suggests that **26b** exhibits the *S*-hydroxyl function and **26a** displays the *R*-hydroxyl function (Fig. 86). Overall XisB_KR does not unambiguously match the described KR types which is in turn consistent with the antiSMASH prediction for XisB_KR to show unknown stereochemistry.

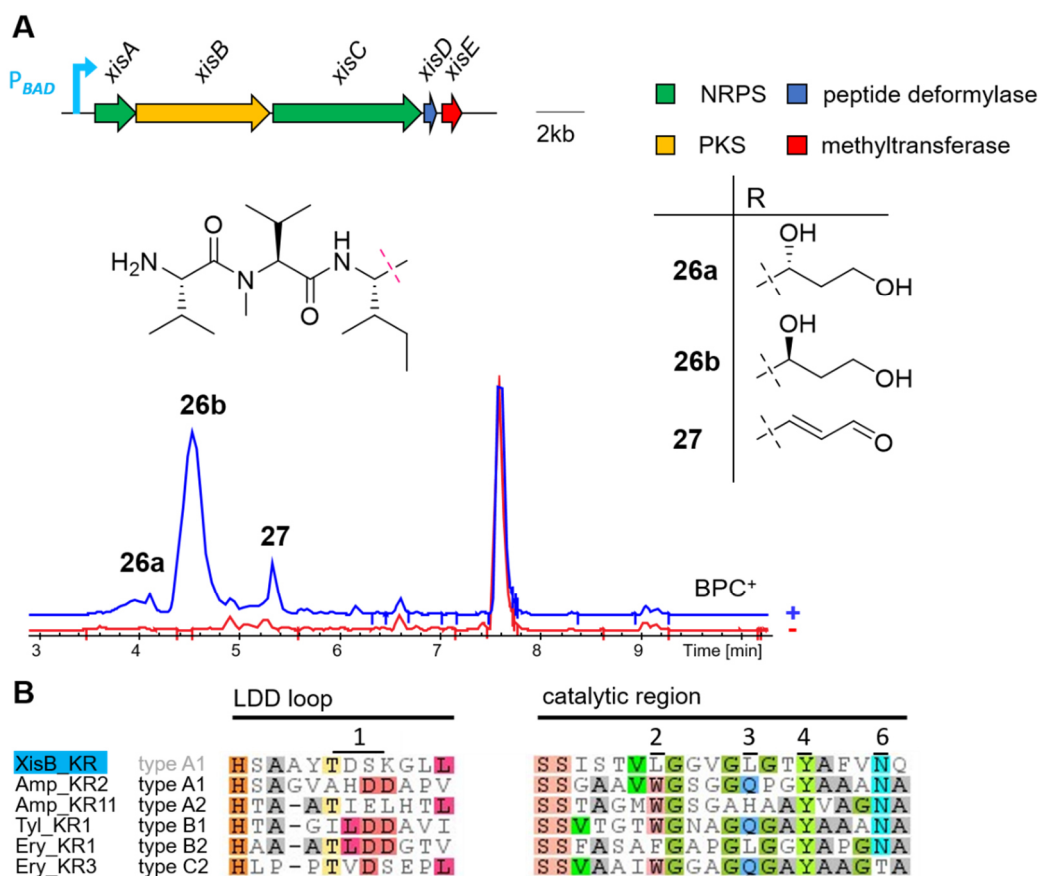


Figure 86. Xildivaline production in *X. hominickii* Δhfq P_{BAD_xisA} . **A** Xildivalines **26a**, **26b** and **27** are detected upon arabinose-induction. **26a** shows *R*-hydroxyl group and **26b** shows *S*-hydroxyl group. BPCs are shown in blue (arabinose induction (+)) and in red (no induction (-)). **B** Sequence alignment of XisB_KR with characterised KR. Absence of the LDD motif (1) suggests *S* configuration for the hydroxyl function in **26b**. Numbering and nomenclature after Keatinge-Clay.^[114]

In this work, 4-bromothiophenol (BTP) was used as strong nucleophile to reveal the presence of the Michael acceptor in **27**. The EIC for the BTP adduct with m/z 542.2 $[M+H]^+$ shows **30** to be only present in the arabinose-induced sample. Although no MS² spectra could be recorded the presence of the dipeptide fragment ion with m/z 213.2 $[M+H]^+$ at this retention time supports that the observed m/z 542.2 $[M+H]^+$ represents the **27**-BTP product. Low signal intensities did not allow to detect the prominent roughly 1:1 isotope pattern for bromine containing compounds with its isotopes Br⁷⁹ and Br⁸¹. Dehydration of the alcohol function is usually obtained by dehydration domains in PKSs.^[110] The PKS XisB in the studied BGC lacks such a domain. The dehydration step could be carried out by a tailoring post-PKS enzyme as was

DISCUSSION

shown for the generation of a third *cis* double bond during phospholactomycin biosynthesis or could also occur spontaneously.^[234]

4.4.2 Additional derivatives detected upon P_{BAD_xisA} activation in *X. hominickii* WT

Upon P_{BAD_xisA} activation in the *X. hominickii* wild type strain, two additional products were detected. Based on MS² fragmentation they only differ in the C-terminal NRP-PK building block from the main products. They are proposed to show a C-terminal alcohol instead of the aldehyde which requires another two-electron reduction (Fig. 87).

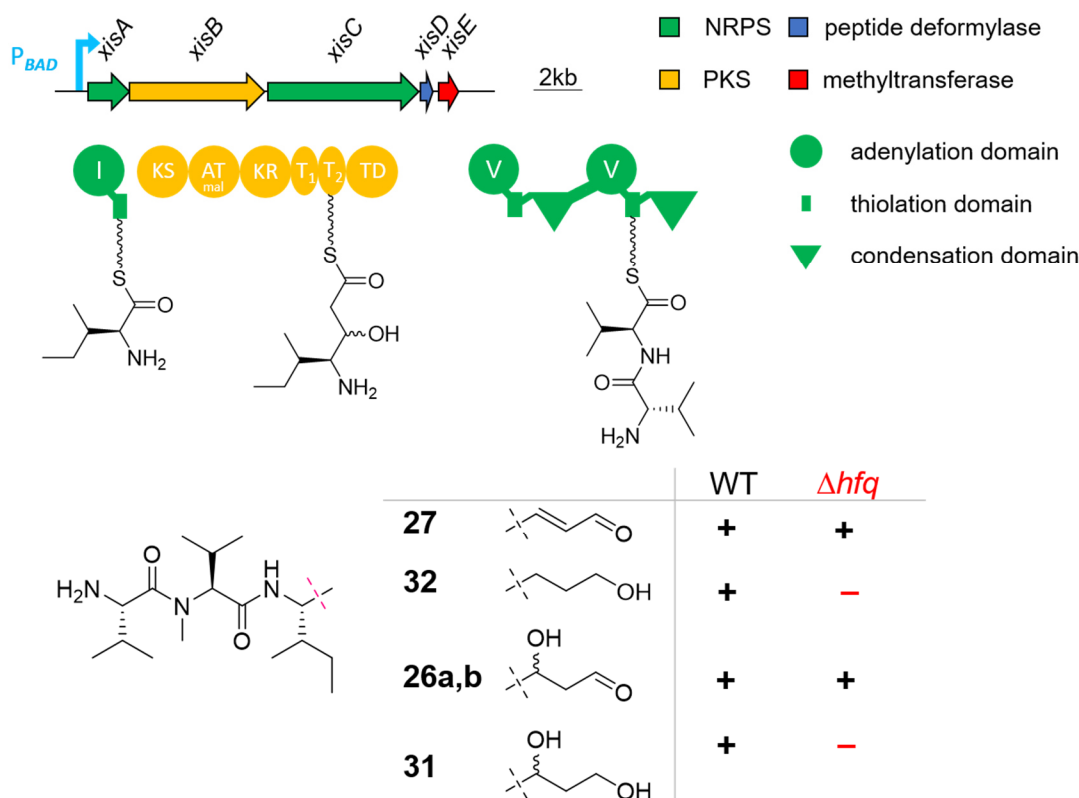


Figure 87. Different xildivaline product spectrum for P_{BAD_xisA} activation in *X. hominickii* WT and *X. hominickii* Δhfq . **31** and **32** are only produced in upon P_{BAD_xisA} activation in the wild type strain. + indicates production, - indicates no production.

In general, terminal reductases are shown to produce both aldehyde and alcohol products.^[84] Although the terminal reductase of the NRPS LtxA performs a four-electron reduction to the alcohol, Read & Walsh observed an aldehyde release intermediate.^[235] Since **31** and **32** are only detected in the wild type P_{BAD_xisA} activation, it appears more likely that the reduction to the alcohol is catalysed by an additional, tailoring enzyme that is not active in Δhfq . An additional SDR is for example required to conduct the four-electron reduction for linear gramicidin A production in *Bacillus*.^[236] **32** not only lacks the C-terminal aldehyde but also the hydroxyl function in β -position. In full-reducing PKSs, first a dehydratase and then an enoylreductase is required for this reduction but XisB lacks both domain types.^[97] The elimination of water resulting in the Michael acceptor could happen spontaneously as assumed

for formation of **27**. It might be that the required enzymes, namely an enoylreductase and a SDR to generate **31** and **32** are not expressed upon P_{BAD} activation and are downregulated in Δhfq background. Possibly, a proteomics-based approach as applied for xenofuranone MT identification in this work, could also be effective to identify the required reductases.

Overall the xildivaline structures presented in this work are predicted based on LC-MS data. The final proof for the proposed structures would require compound purification and subsequent NMR-based structure elucidation. In the course of her master thesis, Janine Chekaiban conducted compound purification but impurities and/or low yields did not allow structure elucidation by NMR.^[153] Another strategy might also be the LC-MS-based comparison of the natural product with the chemically synthesised standard.^[138] The synthesis could be achieved via solid-phase peptide synthesis from two Fmoc-protected L-valine building blocks and a Fmoc-protected statin-like building block. The use of different statin-like building blocks could then be used to determine the different structures including stereospecificity of the hydroxyl function and position of the *N*-methylation.

4.4.3 Methyltransferase XisE

The *N*-methylation in xildivaline is proposed to be present on the second L-valine building block based on the observed b-ion fragmentation pattern. In the $\Delta xisE$ strain, only non-methylated products are detected. Based on this, it is assumed that the *N*-methylation is installed on the released product and thus is not a crucial step during the assembly line biosynthesis. This is also in accordance with the deduced ten letter non-ribosomal code which is identical for both A domains to incorporate L-valine. Cycloaspeptides are *N*-methylated NRPs that are methylated by an additional *N*-methyltransferase which is also not embedded in the NRPS, like an M domain, but provides *N*-methylated phenylalanine for the cycloaspeptide NRPS.^[51,237] When in this case the *N*-methyltransferase is deleted, cycloaspeptides are no longer made.^[237] Taken together, XisE is predicted to methylate the fully assembled, NRPS released intermediate as presented in both biosynthetic routes (Fig. 88). To further support this proposition, XisE could be expressed, purified and presented with different substrates in an *in vitro* assay. There the different intermediates, e.g. L-valine, valyl-valine and valyl-valyl-isoleucine, could be analysed as free precursors or as SNAC-bound intermediates.

4.4.4. Peptide deformylase XisD and possible function of xildivaline

Deletion of the peptide deformylase encoding gene *xisD* showed no altered product formation and no signals corresponding to *m/z* ratios of formylated xildivalines. From this, it is assumed that XisD is not involved in the biosynthesis of xildivaline. In prokaryotes, peptide deformylases are engaged in the maturation of ribosomally synthesised peptides by deformylation of the

DISCUSSION

formylated start methionine.^[238] Peptide deformylase function is assumed to be essential when polypeptide formylation occurs.^[239] All four strains which were identified to harbour the *xis* BGC exhibit a peptide deformylase encoding gene (*def*) in proximity to the methionyl-tRNA-formyltransferase encoding *fmt*. In *X. hominickii*, *def* and *fmt* are encoded on the bacterial chromosome and not on pXhom which carries the *xis* BGC. The finding that *X. hominickii* harbours an additional peptide deformylase is congruent with the observation that *xisD* deletion showed no severe growth effects and suggests that *xisD* is not essential. Yet, *xisD* could encode a functional peptide deformylase and play a role as resistance gene in the *xis* BGC. In a peptide deformylase inhibitor study it was shown that expression of heterologous peptide deformylases can save the producer from the peptide deformylase inhibitor.^[240] Possibly, XisD could function in a similar way and thus represent a resistance mechanism for the xildivaline producer. This could hint that xildivaline exhibits peptide deformylase inhibitory activity and should be investigated in the future. To test whether XisD is functional and can protect the cell from peptide deformylase inhibitors, XisD could be expressed heterologously in *E. coli* which is then treated with a known peptide deformylase inhibitor, e.g. actinonin.

In general, xildivaline shows similarities with other aldehyde-containing peptides which are known proteasome inhibitors as already pointed out by Janine Chekaiban, as for example leupeptin, fellutamide B or MG-132.^[84,231] Peptide aldehydes represent reversible serine or cysteine protease inhibitors.^[231] Furthermore, xildivaline was previously shown to exhibit growth inhibitory activity towards *Micrococcus luteus*.^[153] Taken together, these findings suggest to test for protease inhibitory activity as well as peptide deformylase inhibitory activity of xildivaline.

4.4.5 Tandem T domains

In the investigated BGC, XisB exhibits two adjacent T domains. Point mutations, respectively in the first or second T domain, showed that product formation of all five xildivalines was possible with only one active T domain. Relative quantification showed an overall decreased production for all xildivalines in both the T1 and T2 mutant strain when compared to the wild type-like promoter exchange strain. The production decrease is in accordance with literature examples.^[241,242] For some polyunsaturated fatty acid (PUFA) synthases, point mutations within several of up to six tandem ACPs, e.g. in *Shewanella japonica* did not alter the product spectrum but affected the product titer.^[242] In an engineering approach DEBS module 6 was equipped with two additional ACP domains to its native one, resulting in increased product yield.^[243] It might be interesting to study if duplication of T domains in XisB could enhance xildivaline production in order to facilitate compound purification for structure elucidation.

4.4.6 Proposed biosynthesis

Two possible biosynthetic routes were proposed from the generated data (Fig. 88). The terminal C domain of XisC is predicted to be a ^LC_L domain pointing at the possibility to condense the XisC_T2 tethered dipeptidyl with the XisA_T-S-isoleucine as suggested in route B. Route B lines the enzymes in the order of XisC, XisB and XisA based on the ^LC_L prediction of XisC_C2. This sequence would be different from the order in which genes are organised in the *xis* BGC. Glidobactin is an example where an NRPS-PKS hybrid operates in an order that is distinct from the genetic organisation.^[244]

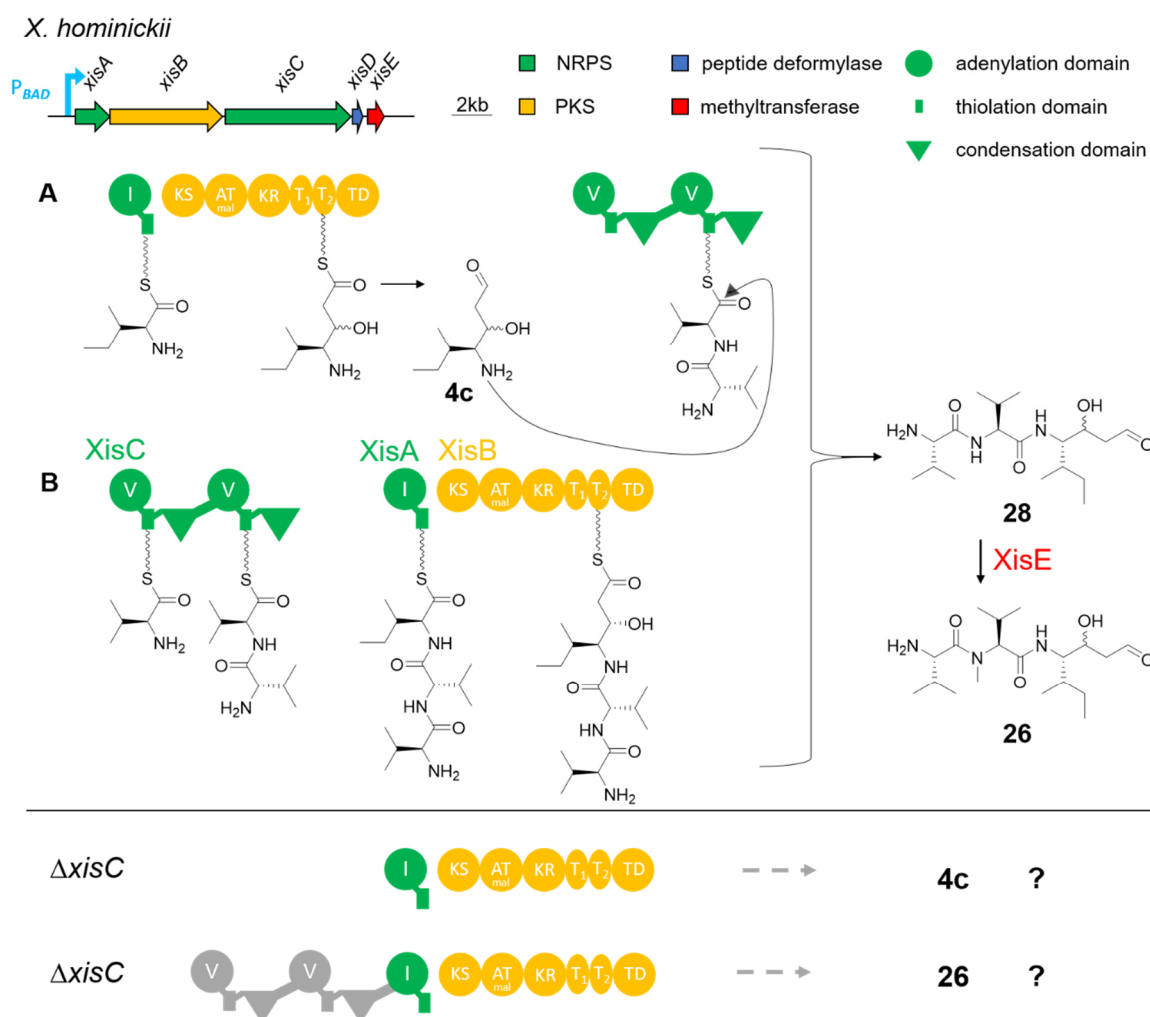


Figure 88. Two possible routes for xildivaline biosynthesis and possible future experiments. In Route A XisC_C2 condenses the XisC_T2-bound divaline with free **4c**. In Route B XisC_C2 condenses the XisC_T2 bound divaline with XisA-bound isoleucine. The final step in both routes is the *N*-methylation of L-valine in **28** to yield xildivaline **26**. Outlook: Deletion of *xisC* could result in accumulation of **4c**. Engineering the N-terminus of XisA could result in production of **26** if XisA and XisB accept the substrates according to route B.

In route A, the enzyme order follows the BGC organisation and suggests that XisC_C2 condenses the dipeptidyl chain with the free NRP-PK as terminal C domains are also described to condense their donor substrate with amines of free substrates.^[76,211] To the best of my

DISCUSSION

knowledge, it is not possible to discriminate these two condensation domain types based on the amino acid sequences so far. In the structure of the terminal C-T didomain of TqaA, Zhang *et al.* observed an additional α -helix instead of the N-terminal lobe known from standard C domains.^[80] They observed a reduced accessibility of the acceptor site which does not require an acceptor substrate as the release mechanism functions by intramolecular cyclisation of fumiquinazoline F.^[80] The sequence alignments of XisC_C2 with other condensation domains did not indicate a clear presence of an additional N-terminal region indicating an α -helical region seen in the TqaA_C_{term} domain. In addition, the sequence alignments did not allow the identification of a possible C-terminal docking domain to mediate interaction with XisA. Since many different docking domain pairs with very diverse structural features were discovered the presence of a (novel) docking domain cannot be excluded. Targeted truncations or deletions in the C-terminal region of XisC could help to investigate a possible docking domain interaction. Dehling *et al* 2019 used a genetically encoded photocrosslinker for DD investigation in GrsA C/E interface.^[245]

Provided xildivaline biosynthesis follows route A with XisC_C2 accepting the free NRP-PK-hybrid product **4c**, the deletion of *xisC* would result in **4c** production (Fig. 88). This could be addressed in the future and O-(2,3,4,5,6-Pentafluorobenzyl)hydroxylamine (PFBHA) could be used to derivatise **4c** to facilitate the detection.^[246] To analyse whether XisA and XisB are able to accept and process the divalyl-substrate provided by XisC as suggested in route B, XisA could be extended at its N-terminal end by NRPS engineering (Fig. 88). Therefore, a divalyl incorporating NRPS unit could be covalently fused to XisA_A domain via a C/A-linker and in case XisA and XisB accept the longer substrate, xildivaline **26** should be produced by this engineered system in the $\Delta xisC$ strain. Provided that this engineered system following route B is functional, different amino acid building blocks could be used to produce structural diversity. This could be an interesting strategy with respect to the above mentioned biological activity of xildivaline.

Another strategy to study whether XisB accepts the single isoleucyl-S-acyl or the tripeptidyl chains as substrate could be the use of chain terminators which mimic the downstream ACP-bound intermediate of the KS (mal/ohmal) and when incorporated cause offloading of the intermediate from the PKS enzyme.^[140] Thereby chain terminators allow the investigation of the upstream accepted substrate by the KS.^[140,247]

4.4.7 Phylogeny

The fact that BGC *xisA-E* is located on a plasmid is interesting in respect to the plasmid paradox whereafter e.g. its replication is an energy consuming process while its genetic content has to make up for it, e.g. antibiotic resistance, detoxification of substances or host

invasion.^[248] Generally, this might hint that xildivalines exhibit beneficial properties in the *X. hominickii* life style. The cluster blast analysis via antiSMASH shows that the BGC *xisA-E* is not widespread in the so far available *Photorhabdus* and *Xenorhabdus* strains and is found in one *Vibrio* strain. This indicates that this BGC is not required for the majority of so far known *Xenorhabdus/ Photorhabdus* species. *Vibrio tubiashii* represents a marine bacterium that is ubiquitously found and both *Xenorhabdus* and *Vibrio* belong to the family of γ -proteobacteria.^[249] Clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 mediated plasmid curing could be used to remove the plasmid entirely and to study bacterial growth or the ability to kill the model insect *Galleria melonella* comparing the *X. hominickii* strain with and without the *xisA-E* encoding plasmid.^[250] Finally, promoter exchanges or heterologous expression could be conducted to investigate whether XisA really activates leucine in the *Vibrio tubiashii* BGC as the antiSMASH prediction suggests.

In conclusion, xildivalines show structural diversity in their PKS-derived building block for which tailoring enzymes might be responsible which in turn may be regulated by Hfq. It might be worth investigating to increase product titers, e.g. by T domain duplication, to facilitate compound purification for structure elucidation and bioactivity testing.

4.5 Concluding remarks

In general, P_{BAD} activation in combination with gene deletions allowed the investigation of selected secondary metabolites and their biosynthesis in *Xenorhabdus*. In addition to the previously demonstrated use to facilitate compound identification, the application of easyPACId in combination with a comparative proteomics approach allowed the identification of the xenofuranone methyltransferase in this work. Overall, different compound profiles were observed for P_{BAD} activation in WT vs Δhfq strains. With Hfq being a global regulatory protein, the differences of product formation are not surprising as tailoring biosynthesis enzymes can be affected by Hfq as shown for xenofuranone methyltransferase. The distinct compound profiles and the tight regulation might hint that *Xenorhabdus* displays different compound profiles in nature at different life cycle stages. One example from *Streptomyces lunaelactis* is the production of bagremycins and ferroverdins, which are both produced by the same BGC but each of them exclusively in dependence of iron availability.^[251] In general, the combination of data from RNA, proteome and metabolome studies could shed more light on these networks. Additionally, generated designated mutant strains that display certain compound profiles could be used to mimic specific life cycle phases in laboratory experiments.

REFERENCES

5 References

- [1] G. M. THOMAS, G. O. POINAR, *Int. J. Syst. Bacteriol.* **1979**, *29*, 352.
- [2] S. Forst, K. Neelson, *Microbiol. Rev.* **1996**, *60*, 21.
- [3] E. Sajnaga, W. Kazimierczak, *Symbiosis* **2020**, *80*, 1.
- [4] K. Lengyel, E. Lang, A. Fodor, E. Szállás, P. Schumann, E. Stackebrandt, *Syst. Appl. Microbiol.* **2005**, *28*, 115.
- [5] P. Tailliez, S. Pagès, N. Ginibre, N. Boemare, *Int. J. Syst. Evol. Microbiol.* **2006**, *56*, 2805.
- [6] T. A. Ciche, C. Darby, R.-U. Ehlers, S. Forst, H. Goodrich-Blair, *Biol. Control* **2006**, *38*, 22.
- [7] Y.-M. Shi, H. B. Bode, *Nat. Prod. Rep.* **2018**, *35*, 309.
- [8] G. R. Richards, H. Goodrich-Blair, *Cell. Microbiol.* **2009**, *11*, 1025.
- [9] H. Snyder, S. P. Stock, S.-K. Kim, Y. Flores-Lara, S. Forst, *Appl. Environ. Microbiol.* **2007**, *73*, 5338.
- [10] E. C. Martens, K. Heungens, H. Goodrich-Blair, *J. Bacteriol.* **2003**, *185*, 3147.
- [11] H. Goodrich-Blair, D. J. Clarke, *Mol. Microbiol.* **2007**, *64*, 260.
- [12] M. Sicard, K. Brugirard-Ricaud, S. Pagès, A. Lanois, N. E. Boemare, M. Brehélin, A. Givaudan, *Appl. Environ. Microbiol.* **2004**, *70*, 6473.
- [13] L. Batalla-Carrera, A. Morton, D. Shapiro-Ilan, M. R. Strand, F. García-del-Pino, *J. Nematol.* **2014**, *46*, 281.
- [14] S. E. Brown, A. T. Cao, E. R. Hines, R. J. Akhurst, P. D. East, *J. Biol. Chem.* **2004**, *279*, 14595.
- [15] T. Baiocchi, G. Lee, D.-H. Choe, A. R. Dillman, *Sci. Rep.* **2017**, *7*, 6270.
- [16] M. F. Brivio, M. Mastore, *Insects* **2018**, *9*.
- [17] D. R. Sugar, K. E. Murfin, J. M. Chaston, A. W. Andersen, G. R. Richards, L. deLéon, J. A. Baum, W. P. Clinton, S. Forst, B. S. Goldman et al., *Environ. Microbiol.* **2012**, *14*, 924.
- [18] M. Cao, H. Goodrich-Blair, *Environ. Microbiol.* **2020**, *22*, 5433.
- [19] a) J. M. Chaston, G. Suen, S. L. Tucker, A. W. Andersen, A. Bhasin, E. Bode, H. B. Bode, A. O. Brachmann, C. E. Cowles, K. N. Cowles et al., *PloS one* **2011**, *6*, e27909; b) N. J. Tobias, H. Wolff, B. Djahanschiri, F. Grundmann, M. Kronenwerth, Y.-M. Shi, S. Simonyi, P. Grün, D. Shapiro-Ilan, S. J. Pidot et al., *Nat. Microbiol.* **2017**, *2*, 1676.
- [20] a) C. T. Calderone, S. B. Bumpus, N. L. Kelleher, C. T. Walsh, N. A. Magarvey, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 12809; b) C. T. Walsh, A. M. Gehring, P. H. Weinreb, L. E. N. Quadri, R. S. Flugel, *Curr. Opin. Chem. Biol.* **1997**, *1*, 309.
- [21] K. Ciezki, S. Wesener, D. Jaber, S. Mirza, S. Forst, *Microbiology (Reading, Engl.)* **2019**, *165*, 538.
- [22] N. J. Tobias, Y.-M. Shi, H. B. Bode, *Trends Microbiol.* **2018**, *26*, 833.

- [23] J. C. Castillo, S. E. Reynolds, I. Eleftherianos, *Trends Parasitol.* **2011**, *27*, 537.
- [24] I. González-Santoyo, A. Córdoba-Aguilar, *Entomol. Exp. Appl.* **2012**, *142*, 1.
- [25] J. M. Crawford, C. Portmann, X. Zhang, M. B. J. Roeffaers, J. Clardy, *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 10821.
- [26] X. Cai, S. Nowak, F. Wesche, I. Bischoff, M. Kaiser, R. Fürst, H. B. Bode, *Nat. Chem.* **2017**, *9*, 379.
- [27] H. Tunaz, Y. Park, K. Büyükgüzel, J. C. Bedick, A. R. Nor Aliza, D. W. Stanley, *Arch. Insect Biochem. Physiol.* **2003**, *52*, 1.
- [28] a) M. Sadekuzzaman, Y. Park, S. Lee, K. Kim, J. K. Jung, Y. Kim, *J. Invertebr. Pathol.* **2017**, *145*, 13; b) S. Shrestha, Y. P. Hong, Y. Kim, *J. Asia-Pac. Entomol.* **2010**, *13*, 55.
- [29] M. D. Manniello, A. Moretta, R. Salvia, C. Scieuzo, D. Lucchetti, H. Vogel, A. Sgambato, P. Falabella, *Cell. Mol. Life Sci.* **2021**, *78*, 4259.
- [30] R. S. Khush, W. D. Cornwell, J. N. Uram, B. Lemaitre, *Curr. Biol.* **2002**, *12*, 1728.
- [31] E. Bode, A. O. Brachmann, C. Kegler, R. Simsek, C. Dauth, Q. Zhou, M. Kaiser, P. Klemmt, H. B. Bode, *ChemBioChem* **2015**, *16*, 1115.
- [32] L. Pantel, T. Florin, M. Dobosz-Bartoszek, E. Racine, M. Sarciaux, M. Serri, J. Houard, J.-M. Campagne, R. M. de Figueiredo, C. Midrier et al., *Mol. Cell* **2018**, *70*, 83-94.e7.
- [33] M. Gualtieri, A. Aumelas, J.-O. Thaler, *J. Antibiot.* **2009**, *62*, 295.
- [34] T. D. Vo, C. Spahn, M. Heilemann, H. B. Bode, *ACS Chem. Biol.* **2021**, *16*, 447.
- [35] S. W. Fuchs, F. Grundmann, M. Kurz, M. Kaiser, H. B. Bode, *ChemBioChem* **2014**, *15*, 512.
- [36] K. McGary, E. Nudler, *Curr. Opin. Microbiol.* **2013**, *16*, 112.
- [37] E. P. C. Rocha, *Annu. Rev. Genet.* **2008**, *42*, 211.
- [38] I. Bervoets, D. Charlier, *FEMS Microbiol. Rev.* **2019**, *43*, 304.
- [39] T. M. Gruber, C. A. Gross, *Annu. Rev. Microbiol.* **2003**, *57*, 441.
- [40] J. J. Bijlsma, E. A. Groisman, *Trends Microbiol.* **2003**, *11*, 359.
- [41] D. F. Browning, S. J. W. Busby, *Nat. Rev. Microbiol.* **2016**, *14*, 638.
- [42] A. S. N. Seshasayee, K. Sivaraman, N. M. Luscombe, *Subcell. Biochem.* **2011**, *52*, 7.
- [43] Y. Engel, C. Windhorst, X. Lu, H. Goodrich-Blair, H. B. Bode, *Front. Microbiol.* **2017**, *8*, 209.
- [44] N. J. Tobias, A. K. Heinrich, H. Eresmann, P. R. Wright, N. Neubacher, R. Backofen, H. B. Bode, *Environ. Microbiol.* **2017**, *19*, 119.
- [45] N. Neubacher, N. J. Tobias, M. Huber, X. Cai, T. Glatter, S. J. Pidot, T. P. Stinear, A. L. Lütticke, K. Papenfort, H. B. Bode, *Nat. Microbiol.* **2020**, *5*, 1481.
- [46] E. Bode, A. K. Heinrich, M. Hirschmann, D. Abebew, Y.-N. Shi, T. D. Vo, F. Wesche, Y.-M. Shi, P. Grün, S. Simonyi et al., *Angew. Chem. Int. Ed. Engl.* **2019**, *58*, 18957.
- [47] M. A. Fischbach, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3468.

REFERENCES

- [48] J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel, M. D. Burkart, *Nat. Prod. Rep.* **2014**, *31*, 61.
- [49] L. Du, C. Sánchez, B. Shen, *Metab. Eng.* **2001**, *3*, 78.
- [50] T. Stachelhaus, H. D. Mootz, V. Bergendahl, M. A. Marahiel, *J. Biol. Chem.* **1998**, *273*, 22773.
- [51] R. D. Süssmuth, A. Mainz, *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 3770.
- [52] K. Bloudoff, T. M. Schmeing, *Biochim Biophys Acta Proteins Proteom* **2017**, *1865*, 1587.
- [53] H. D. Mootz, D. Schwarzer, M. A. Marahiel, *ChemBioChem* **2002**, *3*, 490.
- [54] A. Wiest, D. Grzegorski, B.-W. Xu, C. Goulard, S. Rebuffat, D. J. Ebbole, B. Bodo, C. Kenerley, *J. Biol. Chem.* **2002**, *277*, 20862.
- [55] S. Götze, P. Stallforth, *Nat. Prod. Rep.* **2020**, *37*, 29.
- [56] M. Hahn, T. Stachelhaus, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 15585.
- [57] M. Hahn, T. Stachelhaus, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 275.
- [58] A. M. Gulick, *ACS Chem. Biol.* **2009**, *4*, 811.
- [59] E. Conti, T. Stachelhaus, M. A. Marahiel, P. Brick, *EMBO J.* **1997**, *16*, 4174.
- [60] M. A. Marahiel, T. Stachelhaus, H. D. Mootz, *Chem. Rev.* **1997**, *97*, 2651.
- [61] H. Yonus, P. Neumann, S. Zimmermann, J. J. May, M. A. Marahiel, M. T. Stubbs, *J. Biol. Chem.* **2008**, *283*, 32484.
- [62] A. Scaglione, M. R. Fullone, L. C. Montemiglio, G. Parisi, C. Zamparelli, B. Vallone, C. Savino, I. Grgurina, *FEBS J.* **2017**, *284*, 2981.
- [63] K. J. Labby, S. G. Watsula, S. Garneau-Tsodikova, *Nat. Prod. Rep.* **2015**, *32*, 641.
- [64] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol. (Oxford, U. K.)* **1999**, *6*, 493.
- [65] a) G. L. Challis, J. Ravel, C. A. Townsend, *Chem. Biol.* **2000**, *7*, 211; b) C. Rausch, T. Weber, O. Kohlbacher, W. Wohlleben, D. H. Huson, *Nucleic Acids Res.* **2005**, *33*, 5799.
- [66] A. Stanišić, A. Hüsken, H. Kries, *Chem. Sci.* **2019**, *10*, 10395.
- [67] J. M. Crawford, C. Portmann, R. Kontnik, C. T. Walsh, J. Clardy, *Org. Lett.* **2011**, *13*, 5144.
- [68] C. T. Walsh, R. V. O'Brien, C. Khosla, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 7098.
- [69] S. Weinig, H.-J. Hecht, T. Mahmud, R. Müller, *Chem. Biol.* **2003**, *10*, 939.
- [70] T. L. Schneider, B. Shen, C. T. Walsh, *Biochemistry* **2003**, *42*, 9722.
- [71] M. Singh, S. Chaudhary, D. Sareen, *J. Biosci.* **2017**, *42*, 175.
- [72] B. R. Miller, A. M. Gulick, *Methods Mol. Biol. (N. Y.)* **2016**, *1401*, 3.
- [73] T. A. Keating, C. T. Walsh, *Curr. Opin. Chem. Biol.* **1999**, *3*, 598.
- [74] T. Izoré, M. J. Cryle, *Nat. Prod. Rep.* **2018**, *35*, 1120.
- [75] T. A. Keating, C. G. Marshall, C. T. Walsh, A. E. Keating, *Nat. Struct. Biol.* **2002**, *9*, 522.
- [76] S. Dekimpe, J. Masschelein, *Nat. Prod. Rep.* **2021**.
- [77] C. Rausch, I. Hoof, T. Weber, W. Wohlleben, D. H. Huson, *BMC Evol. Biol.* **2007**, *7*, 78.

- [78] L. Zhong, X. Diao, N. Zhang, F. Li, H. Zhou, H. Chen, X. Bai, X. Ren, Y. Zhang, D. Wu et al., *Nat. Commun.* **2021**, *12*, 296.
- [79] X. Gao, S. W. Haynes, B. D. Ames, P. Wang, L. P. Vien, C. T. Walsh, Y. Tang, *Nat. Chem. Biol.* **2012**, *8*, 823.
- [80] J. Zhang, N. Liu, R. A. Cacho, Z. Gong, Z. Liu, W. Qin, C. Tang, Y. Tang, J. Zhou, *Nat. Chem. Biol.* **2016**, *12*, 1001.
- [81] N. Ziemert, S. Podell, K. Penn, J. H. Badger, E. Allen, P. R. Jensen, *PLoS one* **2012**, *7*, e34064.
- [82] S. Meyer, J.-C. Kehr, A. Mainz, D. Dehm, D. Petras, R. D. Süßmuth, E. Dittmann, *Cell Chem. Biol.* **2016**, *23*, 462.
- [83] L. Du, L. Lou, *Nat. Prod. Rep.* **2010**, *27*, 255.
- [84] M. W. Mullowney, R. A. McClure, M. T. Robey, N. L. Kelleher, R. J. Thomson, *Nat. Prod. Rep.* **2018**, *35*, 847.
- [85] M. McErlean, J. Overbay, S. van Lanen, *J. Ind. Microbiol. Biotechnol.* **2019**, *46*, 493.
- [86] G. Schoenafinger, N. Schracke, U. Linne, M. A. Marahiel, *J. Am. Chem. Soc.* **2006**, *128*, 7406.
- [87] K. Haslinger, M. Peschke, C. Brieke, E. Maximowitsch, M. J. Cryle, *Nature* **2015**, *521*, 105.
- [88] M. Winn, J. K. Fyans, Y. Zhuo, J. Micklefield, *Nat. Prod. Rep.* **2016**, *33*, 317.
- [89] W.-W. Sun, C.-J. Guo, C. C. C. Wang, *Fungal Genet. Biol.* **2016**, *89*, 84.
- [90] a) P. Schneider, S. Bouhired, D. Hoffmeister, *Fungal Genet. Biol.* **2008**, *45*, 1487; b) C.-J. Guo, B. P. Knox, J. F. Sanchez, Y.-M. Chiang, K. S. Bruno, C. C. C. Wang, *Org. Lett.* **2013**, *15*, 3562.
- [91] C. J. Balibar, A. R. Howard-Jones, C. T. Walsh, *Nat. Chem. Biol.* **2007**, *3*, 584.
- [92] H.-H. Yeh, Y.-M. Chiang, R. Entwistle, M. Ahuja, K.-H. Lee, K. S. Bruno, T.-K. Wu, B. R. Oakley, C. C. C. Wang, *Appl. Microbiol. Biotechnol.* **2012**, *96*, 739.
- [93] B. Wackler, P. Schneider, J. M. Jacobs, J. Pauly, C. Allen, M. Nett, D. Hoffmeister, *Chem. Biol.* **2011**, *18*, 354.
- [94] E. Hühner, K. Öqvist, S.-M. Li, *Org. Lett.* **2019**, *21*, 498.
- [95] E. Geib, F. Baldeweg, M. Doerfer, M. Nett, M. Brock, *Cell Chem. Biol.* **2019**, *26*, 223-234.e6.
- [96] B. Wackler, G. Lackner, Y. H. Chooi, D. Hoffmeister, *ChemBioChem* **2012**, *13*, 1798.
- [97] C. Hertweck, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 4688.
- [98] B. Shen, *Curr. Opin. Chem. Biol.* **2003**, *7*, 285.
- [99] A. T. Keatinge-Clay, *Nat. Prod. Rep.* **2012**, *29*, 1050.
- [100] T. Robbins, Y.-C. Liu, D. E. Cane, C. Khosla, *Curr. Opin. Struct. Biol.* **2016**, *41*, 10.

REFERENCES

- [101] R. Broadhurst, D. Nietlispach, M. P. Wheatcroft, P. F. Leadlay, K. J. Weissman, *Chem. Biol.* **2003**, *10*, 723.
- [102] B. J. Dunn, D. E. Cane, C. Khosla, *Biochemistry* **2013**, *52*, 1839.
- [103] E. M. Musiol-Kroll, W. Wohlleben, *Antibiotics* **2018**, *7*.
- [104] C. Khosla, Y. Tang, A. Y. Chen, N. A. Schnarr, D. E. Cane, *Annu. Rev. Biochem.* **2007**, *76*, 195.
- [105] Y. A. Chan, A. M. Podevels, B. M. Kevany, M. G. Thomas, *Nat. Prod. Rep.* **2009**, *26*, 90.
- [106] T. Robbins, J. Kapilivsky, D. E. Cane, C. Khosla, *Biochemistry* **2016**, *55*, 4476.
- [107] L. Zhang, T. Hashimoto, B. Qin, J. Hashimoto, I. Kozono, T. Kawahara, M. Okada, T. Awakawa, T. Ito, Y. Asakawa et al., *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 1740.
- [108] A. T. Keatinge-Clay, *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 4658.
- [109] G. J. Dodge, F. P. Maloney, J. L. Smith, *Nat. Prod. Rep.* **2018**, *35*, 1082.
- [110] J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **2001**, *18*, 380.
- [111] M. Grininger, *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 8680.
- [112] R. Reid, M. Piagentini, E. Rodriguez, G. Ashley, N. Viswanathan, J. Carney, D. V. Santi, C. R. Hutchinson, R. McDaniel, *Biochemistry* **2003**, *42*, 72.
- [113] A. T. Keatinge-Clay, R. M. Stroud, *Structure* **2006**, *14*, 737.
- [114] A. T. Keatinge-Clay, *Chem. Biol.* **2007**, *14*, 898.
- [115] Y. Li, G. J. Dodge, W. D. Fiers, R. A. Fecik, J. L. Smith, C. C. Aldrich, *J. Am. Chem. Soc.* **2015**, *137*, 7003.
- [116] J. F. Barajas, R. M. Phelan, A. J. Schaub, J. T. Kliever, P. J. Kelly, D. R. Jackson, R. Luo, J. D. Keasling, S.-C. Tsai, *Chem. Biol.* **2015**, *22*, 1018.
- [117] C. Olano, C. Méndez, J. A. Salas, *Nat. Prod. Rep.* **2010**, *27*, 571.
- [118] C. D. Richter, D. Nietlispach, R. W. Broadhurst, K. J. Weissman, *Nat. Chem. Biol.* **2008**, *4*, 75.
- [119] F. Liu, S. Garneau, C. T. Walsh, *Chem. Biol.* **2004**, *11*, 1533.
- [120] H. G. Smith, M. J. Beech, J. R. Lewandowski, G. L. Challis, M. Jenner, *J. Ind. Microbiol. Biotechnol.* **2021**, *48*.
- [121] C. Hacker, X. Cai, C. Kegler, L. Zhao, A. K. Weickhmann, J. P. Wurm, H. B. Bode, J. Wöhnert, *Nat. Commun.* **2018**, *9*, 4366.
- [122] J. Watzel, C. Hacker, E. Duchardt-Ferner, H. B. Bode, J. Wöhnert, *ACS Chem. Biol.* **2020**, *15*, 982.
- [123] K. M. Fisch, *RSC Adv.* **2013**, *3*, 18228.
- [124] C. Bender, D. Palmer, A. Peñaloza-Vázquez, V. Rangaswamy, M. Ullrich, *Arch. Microbiol.* **1996**, *166*, 71.

- [125] A. G. Atanasov, S. B. Zotchev, V. M. Dirsch, C. T. Supuran, *Nat. Rev. Drug Discovery* **2021**, *20*, 200.
- [126] H. B. Bode, B. Bethe, R. Höfs, A. Zeeck, *ChemBioChem* **2002**, *3*, 619.
- [127] K. Scherlach, C. Hertweck, *Org. Biomol. Chem.* **2009**, *7*, 1753.
- [128] A. Marmann, A. H. Aly, W. Lin, B. Wang, P. Proksch, *Mar. Drugs* **2014**, *12*, 1043.
- [129] J. H. Kim, N. Lee, S. Hwang, W. Kim, Y. Lee, S. Cho, B. O. Palsson, B.-K. Cho, *J. Ind. Microbiol. Biotechnol.* **2021**, *48*.
- [130] C. T. Walsh, M. A. Fischbach, *J. Am. Chem. Soc.* **2010**, *132*, 2469.
- [131] M. H. Medema, R. Kottmann, P. Yilmaz, M. Cummings, J. B. Biggins, K. Blin, I. de Bruijn, Y. H. Chooi, J. Claesen, R. C. Coates et al., *Nat. Chem. Biol.* **2015**, *11*, 625.
- [132] P. J. Rutledge, G. L. Challis, *Nat. Rev. Microbiol.* **2015**, *13*, 509.
- [133] K. Blin, H. U. Kim, M. H. Medema, T. Weber, *Briefings Bioinf.* **2019**, *20*, 1103.
- [134] M. Myronovskyi, A. Luzhetskyy, *Nat. Prod. Rep.* **2016**, *33*, 1006.
- [135] W. F. Reynolds, E. P. Mazzola, in: *Progress in the Chemistry of Organic Natural Products, vol 100 (Eds.: Kinghorn A., Falk H., Kobayashi J.)*. Springer, Cham. **2015**, 223.
- [136] N. Stoll, E. Schmidt, K. Thurow, *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1692.
- [137] H. B. Bode, D. Reimer, S. W. Fuchs, F. Kirchner, C. Dauth, C. Kegler, W. Lorenzen, A. O. Brachmann, P. Grün, *Chemistry* **2012**, *18*, 2342.
- [138] D. Reimer, F. I. Nollmann, K. Schultz, M. Kaiser, H. B. Bode, *J. Nat. Prod.* **2014**, *77*, 1976.
- [139] Y. Tsunematsu, *J. Nat. Med.* **2021**, *75*, 261.
- [140] M. Tosin, L. Betancor, E. Stephens, W. M. A. Li, J. B. Spencer, P. F. Leadlay, *ChemBioChem* **2010**, *11*, 539.
- [141] K. A. Bozhüyük, J. Micklefield, B. Wilkinson, *Curr. Opin. Microbiol.* **2019**, *51*, 88.
- [142] C. Beck, J. F. G. Garzón, T. Weber, *Biotechnol. Bioprocess Eng.* **2020**, *25*, 886.
- [143] M. Alanjary, C. Cano-Prieto, H. Gross, M. H. Medema, *Nat. Prod. Rep.* **2019**, *36*, 1249.
- [144] U. Sundermann, K. Bravo-Rodriguez, S. Klopries, S. Kushnir, H. Gomez, E. Sanchez-Garcia, F. Schulz, *ACS Chem. Biol.* **2013**, *8*, 443.
- [145] K. A. J. Bozhüyük, A. Linck, A. Tietze, J. Kranz, F. Wesche, S. Nowak, F. Fleischhacker, Y.-N. Shi, P. Grün, H. B. Bode, *Nat. Chem.* **2019**, *11*, 653.
- [146] C. Kegler, H. B. Bode, *Angew. Chem. Int. Ed. Engl.* **2020**, *59*, 13463.
- [147] K. A. J. Bozhueyuek, J. Watzel, N. Abbood, H. B. Bode, *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 17531.
- [148] A. Nivina, K. P. Yuet, J. Hsu, C. Khosla, *Chem. Rev.* **2019**, *119*, 12524.
- [149] T. A. Lundy, S. Mori, S. Garneau-Tsodikova, *ACS Synth. Biol.* **2018**, *7*, 399.

REFERENCES

- [150] T. Dugé de Bernonville, N. Papon, M. Clastre, S. E. O'Connor, V. Courdavault, *Trends Pharmacol. Sci.* **2020**, *41*, 142.
- [151] W. B. Porterfield, N. Poenateetai, W. Zhang, *iScience* **2020**, *23*, 100938.
- [152] A. O. Brachmann, S. Forst, G. M. Furgani, A. Fodor, H. B. Bode, *J. Nat. Prod.* **2006**, *69*, 1830.
- [153] J. Chekaiban, *Identifizierung und Charakterisierung neuer Naturstoffe aus Xenorhabdus hominickii*. MSc thesis. Goethe University Frankfurt/Main (Germany), **2019**.
- [154] N. Keller, *Untersuchung verschiedener Naturstoffe in Xenorhabdus*. MSc thesis. Goethe University Frankfurt/Main (Germany), **2019**.
- [155] M. A. Westphalen, *Analysis and regulation of natural product biosynthesis in Xenorhabdus*. MSc thesis. Goethe University Frankfurt/Main (Germany), **2017**.
- [156] H. C. Birnboim, J. Doly, *Nucleic Acids Res.* **1979**, *7*, 1513.
- [157] M. L. Metzker, C. T. Caskey in *Encyclopedia of life sciences*, Wiley, Chichester, **2005**.
- [158] U. J. Desai, P. K. Pfaffle, *BioTechniques.* **1995**, *19*, 780-2, 784.
- [159] C. Fu, W. P. Donovan, O. Shikapwashya-Hasser, X. Ye, R. H. Cole, *PLoS one* **2014**, *9*, e115318.
- [160] W. J. Dower, J. F. Miller, C. W. Ragsdale, *Nucleic Acids Res.* **1988**, *16*, 6127.
- [161] R. Simon, U. Prierer, A. Pühler, *Nat Biotechnol* **1983**, *1*, 784.
- [162] S. Thoma, M. Schobert, *FEMS Microbiol. Lett.* **2009**, *294*, 127.
- [163] L. Penkert, *Promoter exchange in regulatory mutants for the production of selected natural products in Xenorhabdus*. MSc thesis. Goethe University Frankfurt/Main (Germany), **2016**.
- [164] M. Gualtieri, J.-C. Ogier, S. Pagès, A. Givaudan, S. Gaudriault, *Genome Announc.* **2014**, *2*.
- [165] W. Lorenzen, T. Ahrendt, K. A. J. Bozhüyük, H. B. Bode, *Nat. Chem. Biol.* **2014**, *10*, 425.
- [166] F. I. Nollmann, A. K. Heinrich, A. O. Brachmann, C. Morisseau, K. Mukherjee, Á. M. Casanova-Torres, F. Strobl, D. Kleinhans, S. Kinski, K. Schultz et al., *ChemBioChem* **2015**, *16*, 766.
- [167] L. M. Pöschel, *Aktivierung von Naturstoff-Biosynthese Genclustern in Xenorhabdus bovienii SS-2004 und Xenorhabdus szentirmaii*. BSc thesis. Goethe University Frankfurt/Main (Germany), **2014**.
- [168] A. O. Brachmann, S. A. Joyce, H. Jenke-Kodama, G. Schwär, D. J. Clarke, H. B. Bode, *ChemBioChem* **2007**, *8*, 1721.
- [169] J. Xu, S. Lohrke, I. M. Hurlbert, R. E. Hurlbert, *Appl Environ Microbiol* **1989**, *55*, 806.
- [170] W. Gish, D. J. States, *Nat. Genet.* **1993**, *3*, 266.

- [171] O. Schimming, F. Fleischhacker, F. I. Nollmann, H. B. Bode, *ChemBioChem* **2014**, *15*, 1290.
- [172] S. L. Wenski, N. Berghaus, N. Keller, H. B. Bode, *J. Ind. Microbiol. Biotechnol.* **2021**, *48*.
- [173] R. J. Akhurst, N. E. Boemare, *J. Gen. Microbiol.* **1988**, *134*, 1835.
- [174] K. A. J. Bozhüyük, F. Fleischhacker, A. Linck, F. Wesche, A. Tietze, C.-P. Niesert, H. B. Bode, *Nat. Chem.* **2018**, *10*, 275.
- [175] T. Weber, K. Blin, S. Duddela, D. Krug, H. U. Kim, R. Bruccoleri, S. Y. Lee, M. A. Fischbach, R. Müller, W. Wohlleben et al., *Nucleic Acids Res.* **2015**, *43*, W237-43.
- [176] K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S. Y. Lee, M. H. Medema, T. Weber, *Nucleic Acids Res.* **2019**, *47*, W81-W87.
- [177] NCBI Resource Coordinators, *Nucleic Acids Res.* **2018**, *46*, D8-D13.
- [178] *Geneious 6.1.8*, Biomatters, Ltd.
- [179] L. Zhao, T. D. Vo, M. Kaiser, H. B. Bode, *ChemBioChem* **2020**, *21*, 1288.
- [180] B. J. DAVIS, *Ann. N. Y. Acad. Sci.* **1964**, *121*, 404.
- [181] H. L. Schubert, R. M. Blumenthal, X. Cheng, *Trends Biochem. Sci.* **2003**, *28*, 329.
- [182] The UniProt Consortium, *Nucleic Acids Res.* **2021**, *49*, D480-D489.
- [183] J. W. A. van Dijk, C.-J. Guo, C. C. C. Wang, *Org. Lett.* **2016**, *18*, 6236.
- [184] A. M. Saleh, K. M. Wilding, S. Calve, B. C. Bundy, T. L. Kinzer-Ursem, *J. Biol. Eng.* **2019**, *13*, 43.
- [185] O. Schimming, V. L. Challinor, N. J. Tobias, H. Adihou, P. Grün, L. Pöschel, C. Richter, H. Schwalbe, H. B. Bode, *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 12702.
- [186] A. Tietze, *Engineering and characterisation of non-ribosomal peptide synthetases. Dissertation. Goethe University Frankfurt/Main (Germany)*, **2020**.
- [187] S. Huang, J. Tabudravu, S. S. Elsayed, J. Travert, D. Peace, M. H. Tong, K. Kyeremeh, S. M. Kelly, L. Trembleau, R. Ebel et al., *Angew. Chem.* **2015**, *127*, 12888.
- [188] S. Pohle, C. Appelt, M. Roux, H.-P. Fiedler, R. D. Süssmuth, *J. Am. Chem. Soc.* **2011**, *133*, 6194.
- [189] B. Silakowski, H. U. Schairer, H. Ehret, B. Kunze, S. Weinig, G. Nordsiek, P. Brandt, H. Blöcker, G. Höfle, S. Beyer et al., *J. Biol. Chem.* **1999**, *274*, 37391.
- [190] J. Mareš, J. Hájek, P. Urajová, J. Kopecký, P. Hrouzek, *PloS one* **2014**, *9*, e111904.
- [191] A. Marchler-Bauer, Y. Bo, L. Han, J. He, C. J. Lanczycki, S. Lu, F. Chitsaz, M. K. Derbyshire, R. C. Geer, N. R. Gonzales et al., *Nucleic Acids Res.* **2017**, *45*, D200-D203.
- [192] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli et al., *Nucleic Acids Res.* **2018**, *46*, W296-W303.
- [193] J. Zheng, C. A. Taylor, S. K. Piasecki, A. T. Keatinge-Clay, *Structure.* **2010**, *18*, 913.

REFERENCES

- [194] J. Zheng, A. T. Keatinge-Clay, *Med. Chem. Commun.* **2013**, 4, 34.
- [195] J. Watzel, E. Duchardt-Ferner, S. Sarawi, H. B. Bode, J. Wöhnert, *Angew. Chem. Int. Ed. Engl.* **2021**, 60, 14171.
- [196] J. W. A. van Dijk, C. C. C. Wang, *Org. Lett.* **2018**, 20, 5082.
- [197] S. Chakraborti, B. J. Bahnson, *Biochemistry* **2010**, 49, 3436.
- [198] Y. Kondo, A. Ishigami, S. Kubo, S. Handa, K. Gomi, K. Hirokawa, N. Kajiyama, T. Chiba, K. Shimokado, N. Maruyama, *FEBS Lett.* **2004**, 570, 57.
- [199] M. P. Schwalm, *Structural study on enzymes participating in biosynthetic pathways of bioactive compounds in Xenorhabdus. Protokoll. Goethe University Frankfurt/Main (Germany)*, **2020**.
- [200] S. W. Ragsdale, *Vitam. Horm.* **2008**, 79, 293.
- [201] A.-W. Struck, M. L. Thompson, L. S. Wong, J. Micklefield, *ChemBioChem* **2012**, 13, 2642.
- [202] a) J. Aranda, M. Roca, I. Tuñón, *Org. Biomol. Chem.* **2012**, 10, 5395; b) P. Parajuli, R. P. Pandey, T. H. T. Nguyen, D. Dhakal, J. K. Sohng, *Appl. Biochem. Biotechnol.* **2018**, 184, 1404.
- [203] E. J. Barreiro, A. E. Kümmerle, C. A. M. Fraga, *Chem. Rev.* **2011**, 111, 5215.
- [204] S. Li, L. Wang, F. Yu, Z. Zhu, D. Shobaki, H. Chen, M. Wang, J. Wang, G. Qin, U. J. Erasquin et al., *Chem. Sci.* **2017**, 8, 2107.
- [205] F. Guo, P. Berglund, *Green Chem.* **2017**, 19, 333.
- [206] A. Fan, J. Winkelblech, S.-M. Li, *Appl. Microbiol. Biotechnol.* **2015**, 99, 7399.
- [207] Y. Motorin, J. Burhenne, R. Teimer, K. Koynov, S. Willnow, E. Weinhold, M. Helm, *Nucleic Acids Res.* **2011**, 39, 1943.
- [208] S. Wang, Q. Fang, Z. Lu, Y. Gao, L. Trembleau, R. Ebel, J. H. Andersen, C. Philips, S. Law, H. Deng, *Angew. Chem. Int. Ed. Engl.* **2021**, 60, 3229.
- [209] P. Fu, A. Legako, S. La, J. B. MacMillan, *Chemistry* **2016**, 22, 3491.
- [210] H. L. Kornberg, *J. Biochem.* **1966**, 99, 1.
- [211] D. Reimer, K. N. Cowles, A. Proschak, F. I. Nollmann, A. J. Dowling, M. Kaiser, R. ffrench-Constant, H. Goodrich-Blair, H. B. Bode, *ChemBioChem* **2013**, 14, 1991.
- [212] M. J. Koetsier, P. A. Jekel, H. J. Wijma, R. A. L. Bovenberg, D. B. Janssen, *FEBS Lett.* **2011**, 585, 893.
- [213] H. Bhukya, R. Anand, *J Indian Inst Sci* **2017**, 97, 245.
- [214] L. Cuthbertson, J. R. Nodwell, *Microbiol. Mol. Biol. Rev.* **2013**, 77, 440.
- [215] P. Le Nguyen Minh, S. de Cima, I. Bervoets, D. Maes, V. Rubio, D. Charlier, *FEBS open bio* **2015**, 5, 76.
- [216] T. Chen, N. Mei, P. P. Fu, *J. Appl. Toxicol.* **2010**, 30, 183.

- [217] M. Klapper, S. Götze, R. Barnett, K. Willing, P. Stallforth, *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 8944.
- [218] R. Schuch, Mukherjee. Kumar D., *Appl. Microbiol. Biotechnol.* **1989**, 332.
- [219] E. L. Kline, C. S. Brown, V. Bankaitis, D. C. Montefiori, K. Craig, *Proc. Natl. Acad. Sci. U. S. A.* **1980**, *77*, 1768.
- [220] L. M. Guzman, D. Belin, M. J. Carson, J. Beckwith, *J. Biotechnol.* **1995**, *177*, 4121.
- [221] W. J. H. van Berkel, N. M. Kamerbeek, M. W. Fraaije, *J. Biotechnol.* **2006**, *124*, 670.
- [222] S. Uhlmann, *Studien zur α - und β -Hydroxylierung des Cyclodepsipeptides Skyllamycin aus Streptomyces sp. Acta 2897. Dissertation. Technical University of Berlin (Germany)*, **2014**.
- [223] D. P. Fewer, J. Osterholm, L. Rouhiainen, J. Jokela, M. Wahlsten, K. Sivonen, *Appl. Environ. Microbiol.* **2011**, *77*, 8034.
- [224] S. A. Sieber, U. Linne, N. J. Hillson, E. Roche, C. T. Walsh, M. A. Marahiel, *Chem. Biol.* **2002**, *9*, 997.
- [225] N. J. Hillson, C. T. Walsh, *Biochemistry* **2003**, *42*, 766.
- [226] S. Mori, A. Garzan, O. V. Tsodikov, S. Garneau-Tsodikova, *Biochemistry* **2017**, *56*, 6087.
- [227] T. A. Lundy, S. Mori, S. Garneau-Tsodikova, *Org. Biomol. Chem.* **2019**, *17*, 1169.
- [228] D. J. Merkler, *Enzyme Microb. Technol.* **1994**, *16*, 450.
- [229] R. M. F. Attenborough, D. C. Hayward, M. V. Kitahara, D. J. Miller, E. E. Ball, *Mol. Biol. Evol.* **2012**, *29*, 3095.
- [230] S. Handa, T. J. Spradling, D. R. Dempsey, D. J. Merkler, *Protein Expression Purif.* **2012**, *84*, 9.
- [231] A. F. Kisselev, W. A. van der Linden, H. S. Overkleeft, *Chem. Biol.* **2012**, *19*, 99.
- [232] M. A. Gräwert, N. Gallastegui, M. Stein, B. Schmidt, P.-M. Kloetzel, R. Huber, M. Groll, *Angewandte Chemie (International ed. in English)* **2011**, *50*, 542.
- [233] a) Y.-O. You, C. Khosla, D. E. Cane, *J. Am. Chem. Soc.* **2013**, *135*, 7406; b) J. Zheng, A. T. Keatinge-Clay, *J. Mol. Biol.* **2011**, *410*, 105.
- [234] N. Palaniappan, M. M. Alhamadsheh, K. A. Reynolds, *J. Am. Chem. Soc.* **2008**, *130*, 12236.
- [235] J. A. Read, C. T. Walsh, *J. Am. Chem. Soc.* **2007**, *129*, 15762.
- [236] N. Schracke, U. Linne, C. Mahlert, M. A. Marahiel, *Biochemistry* **2005**, *44*, 8507.
- [237] K. M. J. de Mattos-Shiple, C. Greco, D. M. Heard, G. Hough, N. P. Mulholland, J. L. Vincent, J. Micklefield, T. J. Simpson, C. L. Willis, R. J. Cox et al., *Chem. Sci.* **2018**, *9*, 4109.
- [238] C. Giglione, M. Pierre, T. Meinel, *Mol. Microbiol.* **2000**, *36*, 1197.
- [239] D. Mazel, S. Pochet, P. Marlière, *EMBO J.* **1994**, *13*, 914.

REFERENCES

- [240] C. M. Apfel, H. Locher, S. Evers, B. Takács, C. Hubschwerlen, W. Pirson, M. G. Page, W. Keck, *Antimicrob. Agents Chemother.* **2001**, *45*, 1058.
- [241] a) L. Gu, E. B. Eisman, S. Dutta, T. M. Franzmann, S. Walter, W. H. Gerwick, G. Skiniotis, D. H. Sherman, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 2795; b) A. S. Rahman, J. Hothersall, J. Crosby, T. J. Simpson, C. M. Thomas, *J. Biol. Chem.* **2005**, *280*, 6399.
- [242] H. Jiang, R. Zirkle, J. G. Metz, L. Braun, L. Richter, S. G. van Lanen, B. Shen, *J. Am. Chem. Soc.* **2008**, *130*, 6336.
- [243] Z. Wang, S. R. Bagde, G. Zavala, T. Matsui, X. Chen, C.-Y. Kim, *ACS Chem. Biol.* **2018**, *13*, 3072.
- [244] B. Schellenberg, L. Bigler, R. Dudler, *Environ. Microbiol.* **2007**, *9*, 1640.
- [245] E. Dehling, J. Rüschenbaum, J. Diecker, W. Dörner, H. D. Mootz, *Chem. Sci.* **2020**, *11*, 8945.
- [246] C. Deng, X. Zhang, *Rapid Commun. Mass Spectrom.* **2004**, *18*, 1715.
- [247] R. Sigrist, H. Luhavaya, S. M. K. McKinnie, A. Ferreira da Silva, I. D. Jurberg, B. S. Moore, L. Gonzaga de Oliveira, *ACS Chem. Biol.* **2020**, *15*, 1067.
- [248] A. C. Carroll, A. Wong, *Can. J. Microbiol.* **2018**, *64*, 293.
- [249] F. L. Thompson, B. Hoste, K. Vandemeulebroecke, J. Swings, *Syst. Appl. Microbiol.* **2001**, *24*, 520.
- [250] I. Lauritsen, A. Porse, M. O. A. Sommer, M. H. H. Nørholm, *Microb. Cell Fact.* **2017**, *16*, 135.
- [251] L. Martinet, A. Naômé, B. Deflandre, M. Maciejewska, D. Tellatin, E. Tenconi, N. Smargiasso, E. de Pauw, G. P. van Wezel, S. Rigali, *mBio* **2019**, *10*.
- [252] B. Pang, Y. Chen, F. Gan, C. Yan, L. Jin, J. W. Gin, C. J. Petzold, J. D. Keasling, *J. Am. Chem. Soc.* **2020**, *142*, 10931.

6 Supplementary Information

Table S1. Overview of compounds detected in this work. Asterisk marks compounds for which no HR data were recorded.

compound	name	Chemical formula	found [M+H] ⁺	calculated [M+H] ⁺	Δ ppm	retention time t _R [min]	reference
1	xenofuranone B	C ₁₇ H ₁₄ O ₃	267.1016	267.1016	0.0	8.7	[152]
2	xenofuranone A	C ₁₈ H ₁₆ O ₃	281.1170	281.1172	0.8	9.5	[152]
3	glyoxpeptide	C ₁₁ H ₂₁ N ₃ O ₆	292.1504	292.1503	-0.1	2.1	[153]
5	phenylbutyrolactone IIa	C ₁₈ H ₁₄ O ₅	311.0914	311.0915	-0.4	7.6	This work
6	phenylbutyrolactone II	C ₁₉ H ₁₆ O ₅	325.1071	325.1071	-0.3	9.1	This work
7		C ₁₇ H ₁₄ O ₄	283.0969	283.0965	-1.5	7.3	[155]
8		C ₁₇ H ₁₄ O ₄	283.0969	283.0965	-1.4	7.5	[155]
9	pyrrolizwilline	C ₂₁ H ₂₅ N ₄ O ₃	381.1914	381.1921	1.2	7.8	This work
11	pyrrolizixenamide A	C ₁₃ H ₂₁ N ₂ O ₂	237.1	237.1598	*	6.3	[185]
12		C ₁₁ H ₁₅ N ₂ O ₃	223.1076	223.1077	-0.3	3.6	This work
13		C ₁₀ H ₁₅ N ₂ O ₂	195.1128	195.1128	-0.5	2.5	This work
14	pyrrolizixenamide B	C ₁₄ H ₂₃ N ₂ O ₂	251.1753	251.1754	0.6	6.8	[185]
15		C ₁₁ H ₁₅ N ₂ O ₃	223.1081	223.1077	-1.6	4.2	This work
16		C ₉ H ₁₃ N ₂ O ₂	181.0972	181.0972	0.0	4.0	This work
19		C ₁₄ H ₂₇ N ₃ O ₈	366.1871	366.1871	1.5	2.4	[153]
20		C ₁₁ H ₂₁ N ₃ O ₅	276.1548	276.1554	2.3	1.6	This work
21		C ₁₄ H ₂₇ N ₃ O ₇	350.1912	350.1922	2.8	1.8	This work
22	phototemtide A	C ₃₄ H ₅₃ N ₅ O ₈	660.4	660.3967	*	9.2	[179]
23	phototemtide B	C ₃₂ H ₄₉ N ₅ O ₈	632.4	632.3654	*	8.2	[179]
24	phototemtide C	C ₃₃ H ₅₁ N ₅ O ₈	646.4	646.3810	*	8.8	[179]
25	phototemtide D	C ₃₅ H ₅₅ N ₅ O ₈	674.4	674.4123	*	9.7	[179]
26a		C ₁₉ H ₃₇ N ₃ O ₄	372.2852	372.2857	1.2	3.9	[153]
26b	xildivaline	C ₁₉ H ₃₇ N ₃ O ₄	372.2851	372.2857	1.5	4.5	[153]
27		C ₁₉ H ₃₅ N ₃ O ₃	354.2746	354.2751	0.9	5.4	[153]
28a		C ₁₈ H ₃₅ N ₃ O ₄	358.2699	358.2700	0.5	3.8	This work

SUPPLEMENTARY INFORMATION

28b		$C_{18}H_{35}N_3O_4$	358.2699	358.2700	0.4	4.4	This work
29		$C_{18}H_{33}N_3O_3$	340.2592	340.2595	0.4	5.3	This work
30		$C_{35}H_{40}BrN_3O_3S$	542.19	542.2046	*	8.8	This work
31		$C_{19}H_{39}N_3O_4$	374.3010	374.3013	0.8	3.9	This work
32		$C_{19}H_{39}N_3O_3$	358.3064	358.3064	0.5	5.0	This work
33		$C_{18}H_{37}N_3O_4$	360.2856	360.2857	0.2	3.8	This work
34		$C_{18}H_{37}N_3O_3$	344.2907	344.2908	0.1	4.7	This work

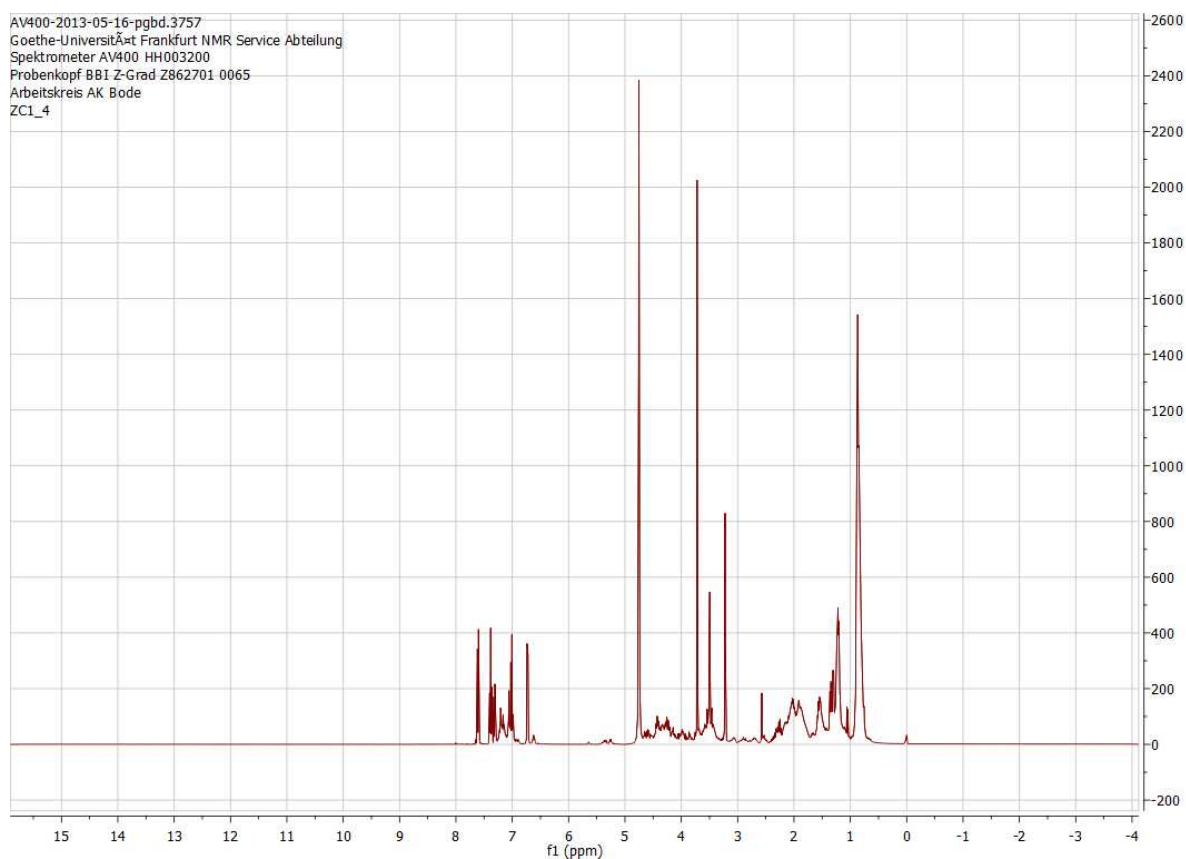


Figure S1. 1H spectrum for **6**. Purification and NMR analysis was conducted by Dr. Zakaria Cheikh Ali (Lab Prof. Dr. Helge B. Bode, Goethe University Frankfurt am Main).

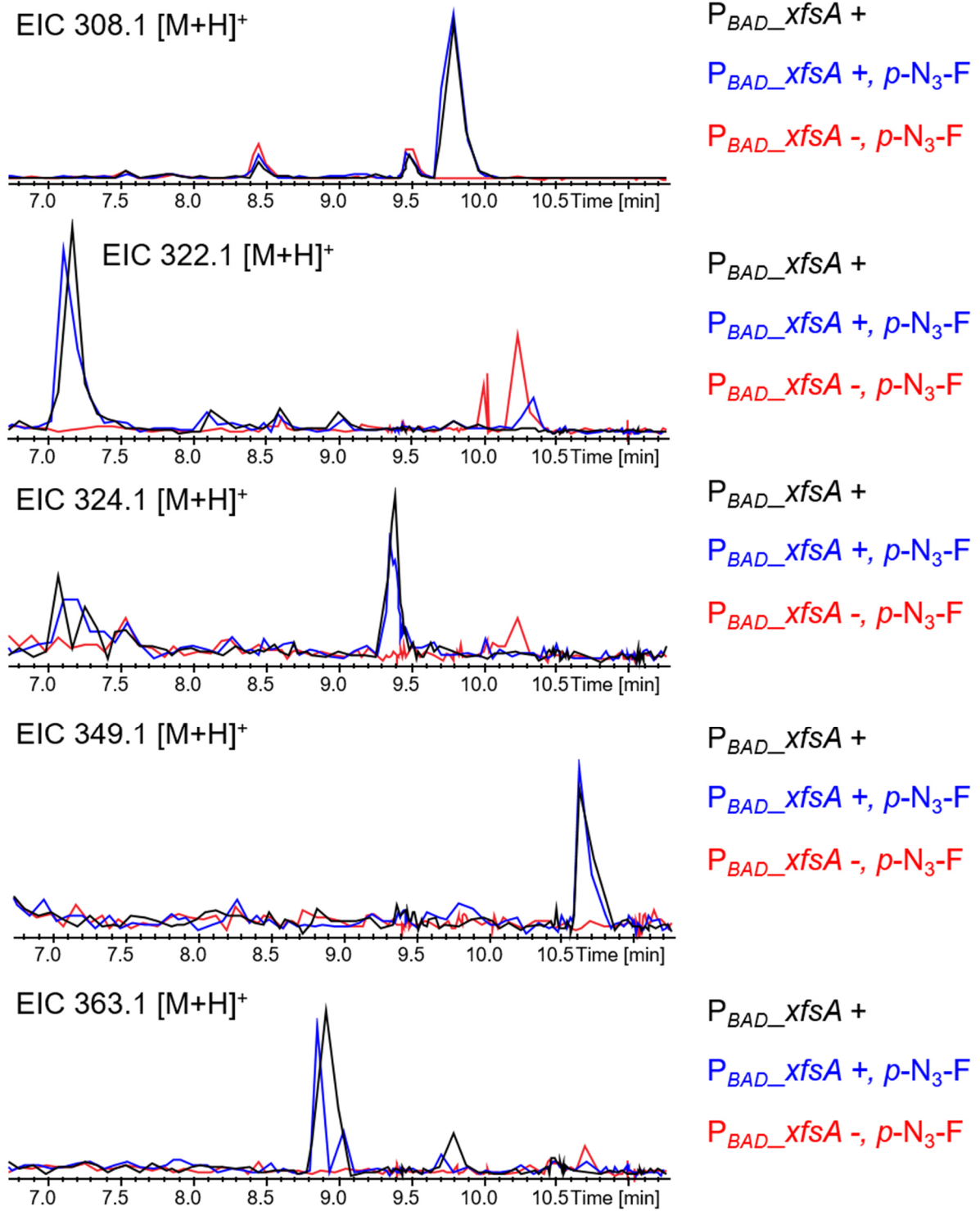


Figure S2. Feeding of para-azido-phenylalanine ($p-N_3-F$) to *X. szentirmaii* P_{BAD_xfsA} showed no incorporation into xenofuranones. EICs for the expected xenofuranone derivatives are shown. + arabinose induction, - no induction.

SUPPLEMENTARY INFORMATION



Figure S3. Multiple sequence alignment of A domains of xenofuranone synthetase XfsA, atromentin synthetase GreA, xenofuranone synthetase XfsA, ralfuranone synthetase RaiA, atromentin synthetase InvA5, terrequinone A synthetase TdiA and phenylalanine-activating A domain from gramicidin S synthetase GrsA. Residues involved in α -keto acid specificity as predicted by [96] are depicted in grey below the sequence. Non-ribosomal code conferring residues identified by [60] are indicated in dark blue.

Table S2. ^1H (500 MHz) and ^{13}C (125 MHz) NMR spectroscopic data for pyrrolizwilline (**9**) in d_6 -DMSO (δ in ppm and J in Hz). Analysis was conducted by Dr. Yi-Ming Shi (Lab Prof. Dr. Helge B. Bode, Goethe University Frankfurt am Main).

No.	δ_{C} (δ_{H} (mult., J)
1	18.3, CH ₃	0.95 (d, 6.7)
2	19.3, CH ₃	0.99 (d, 6.7)
3	30.0, CH	2.22 (dt, 6.5, 13.2)
4	27.2, CH ₂	2.57 (m)
5	24.1, CH ₂	3.02 (t, 7.3)
6	42.9, CH ₂	4.13 (t, 6.9)
7	57.9, CH	4.41 (br s)
8	89.8, CH	6.34 (s)
9NH	-	7.91 (br s)
10	118.1, C	-
11	141.4, C	-
12	143.7, C	-
13	166.5, C	-
14	173.0, C	-
15	124.6, C	-

SUPPLEMENTARY INFORMATION

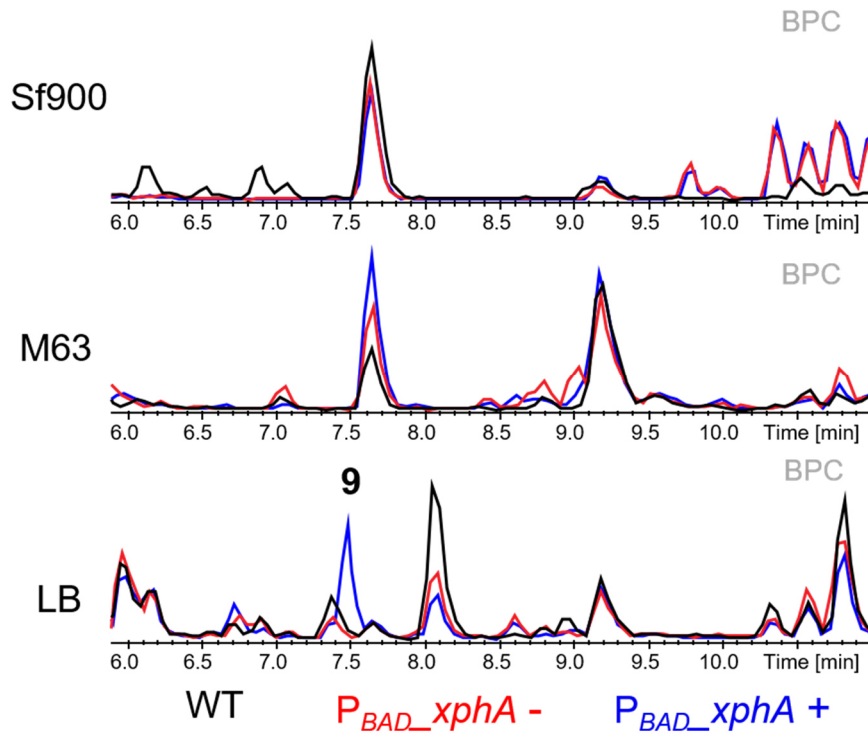


Figure S4. Production for pyrrolizwilline (**9**) is only observed for the arabinose-induced (+) *X. hominickii* P_{BAD_xhpA} in LB medium. BPCs are shown.

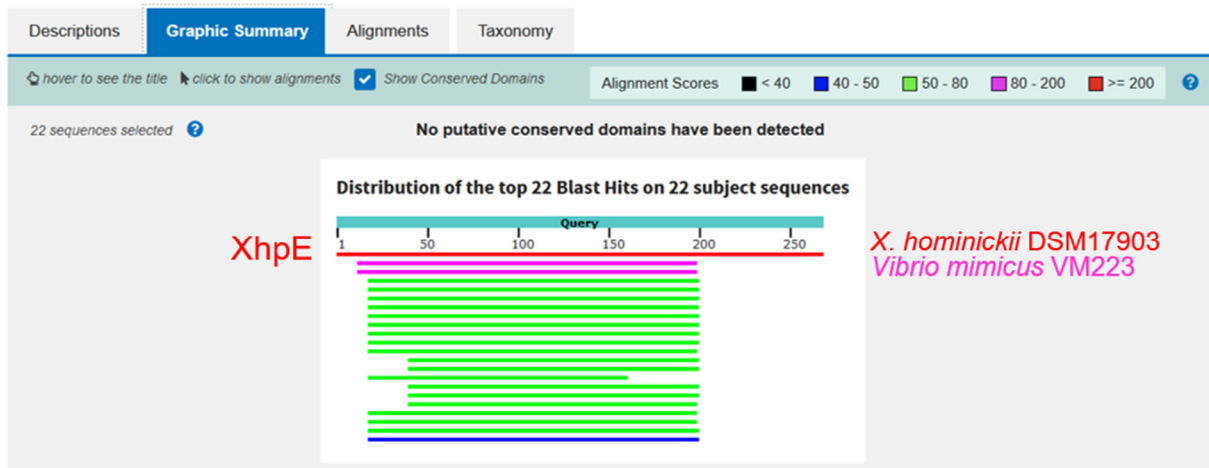


Figure S5. Graphic summary for BlastP analysis of XhpE (query, red). Highest E value for predicted PPTase in *Vibrio mimicus* VM223 (pink). Overall 22 subject sequences were identified and XhpE shows C-terminal region with no sequence similarity to these sequences. No putative conserved domains have been detected.

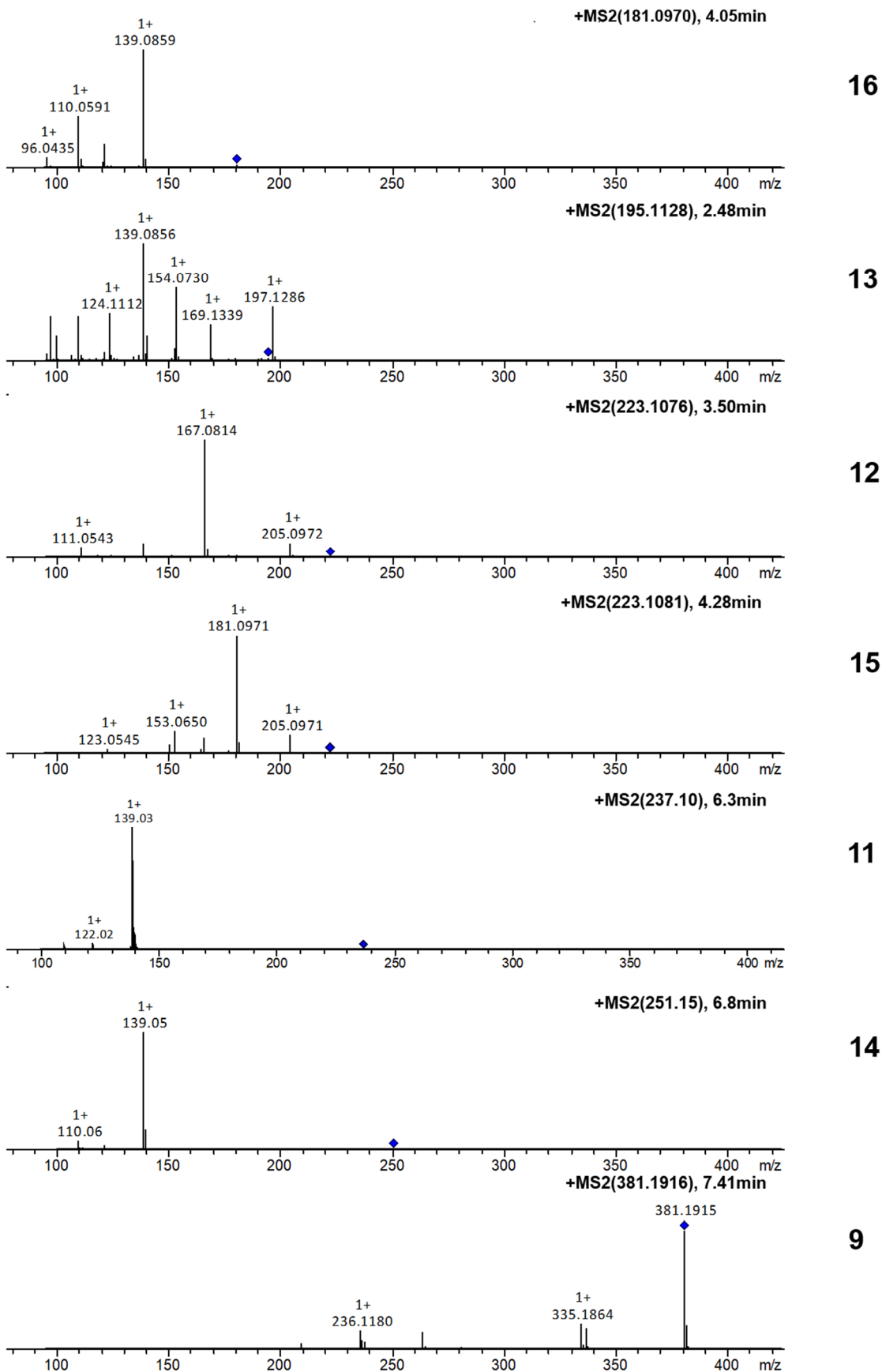


Figure S6. MS² spectra for pyrrolizwilline **9** and observed products **11-16** during investigation of pyrrolizwilline biosynthesis.

SUPPLEMENTARY INFORMATION



Figure S7. Multiple sequence alignment to show fusion site for C^{starter} hybrid engineering. XhpA_C^{starter}/A1 linker region was aligned with PxaA_C^{starter}/A1 linker, GxpS_C^{Dual}/A5 linker and GxpS_LC₁/A3 linker. Dashed line indicates fusion site previously described by Bozhüyük *et al.* 2018.^[174]

<i>X. japonica</i>	<i>X. vietname...</i>	<i>X. hominicki...</i>	<i>X. sp PB62.4</i>	<i>X. budapest...</i>	<i>X. mauleonii</i>	<i>X. szentirmai</i>	<i>X. innexi</i>	<i>X. stockiae</i>	<i>Brenneria sp...</i>	<i>X. hominicki...</i>	<i>V. mimicus V...</i>	<i>X. bovienii</i>	<i>Pseudomonas...</i>	<i>Streptomyces...</i>	<i>Streptomyces...</i>	<i>Actinosyrme...</i>
89.3%	89.3%	84.2%	84.1%	81.6%	73.8%	74.6%	74.7%	75.8%	43.2%	42.8%	43.2%	43.3%	42.1%	37.4%	38.3%	33.5%
<i>X. vietnamensis</i>		85.8%	85.3%	82.6%	73.8%	74.3%	74.1%	75.7%	42.8%	42.4%	43.0%	43.0%	42.1%	37.6%	38.1%	33.4%
<i>X. hominickii Pxa4</i>		84.2%	85.1%	83.1%	73.5%	74.1%	74.6%	75.8%	43.0%	43.0%	43.4%	43.2%	42.1%	37.6%	37.8%	33.6%
<i>X. sp PB62.4</i>		81.1%	85.3%	82.6%	74.2%	74.9%	74.9%	76.9%	42.8%	42.8%	42.9%	42.9%	42.0%	37.8%	38.3%	33.7%
<i>X. budapestensis</i>		81.8%	82.8%	83.1%	73.6%	75.6%	75.6%	77.2%	42.9%	42.9%	42.9%	42.3%	41.5%	37.9%	37.9%	33.3%
<i>X. mauleonii</i>		73.8%	73.8%	82.6%	75.6%	89.2%	89.2%	74.1%	41.2%	40.4%	41.1%	41.7%	40.5%	36.7%	37.5%	32.7%
<i>X. szentirmaii</i>		74.6%	74.3%	74.1%	74.9%	75.6%	74.9%	74.9%	41.9%	40.6%	41.3%	41.5%	40.7%	36.8%	37.5%	32.9%
<i>X. innexi</i>		74.7%	74.6%	74.9%	72.4%	73.3%	73.3%	74.9%	42.5%	41.7%	42.9%	42.3%	41.6%	37.1%	37.9%	33.2%
<i>X. stockiae</i>		75.8%	75.1%	75.6%	77.4%	74.1%	74.1%	82.9%	43.1%	42.7%	42.9%	43.3%	42.7%	38.1%	38.6%	33.5%
<i>Brenneria sp. EniD312</i>		42.8%	42.8%	43.0%	42.8%	41.7%	41.9%	43.1%	64.1%	64.1%	60.8%	62.7%	57.6%	36.3%	37.7%	34.1%
<i>X. hominickii XhpA</i>		42.8%	42.4%	42.8%	42.3%	40.4%	40.6%	42.7%	61.4%	60.8%	60.8%	70.1%	57.6%	38.3%	37.7%	32.9%
<i>V. mimicus VM223</i>		43.2%	43.0%	43.4%	42.9%	41.1%	41.3%	42.9%	60.8%	60.8%	60.8%	70.1%	57.6%	38.3%	37.7%	34.5%
<i>X. bovienii</i>		43.3%	43.0%	43.2%	42.9%	41.7%	41.8%	43.3%	60.8%	62.7%	62.7%	70.1%	57.6%	37.5%	38.1%	33.4%
<i>Pseudomonas sp. MIACH</i>		42.1%	42.1%	42.0%	40.9%	40.7%	40.7%	41.8%	57.4%	57.6%	57.6%	62.1%	62.1%	37.7%	39.3%	33.4%
<i>Streptomyces cattleya</i>		37.4%	37.6%	37.6%	36.7%	36.8%	36.8%	38.1%	39.1%	38.3%	38.3%	37.5%	37.7%	37.7%	39.3%	43.7%
<i>Streptomyces sp. Mg1</i>		38.3%	38.1%	37.6%	38.3%	37.5%	37.5%	38.6%	39.5%	37.2%	39.6%	38.1%	37.7%	54.1%	54.1%	43.0%
<i>Actinosyrmenarium</i>		33.5%	33.4%	33.6%	33.7%	32.7%	32.9%	33.5%	34.1%	32.2%	34.5%	33.4%	33.4%	43.7%	43.0%	43.0%

Figure S8. Multiple protein sequence alignment of NRPS sequences from pyrrolizinenamide(-like) BGCs from *X. japonica*, *X. vietnamensis*, *X. hominickii* (PxaA, XhpB), *Xenorhabdus* PB62.4, *X. budapestensis*, *X. mauleonii*, *X. szentirmaii*, *X. innexi*, *X. stockiae*, *Brenneria* sp. EniD312, *Vibrio mimicus* VM223, *X. bovienii*, *Pseudomonas* sp. MIACH, *Streptomyces cattleya*, *Streptomyces* sp. Mg1. Sequence identity is given in percent.

SUPPLEMENTARY INFORMATION

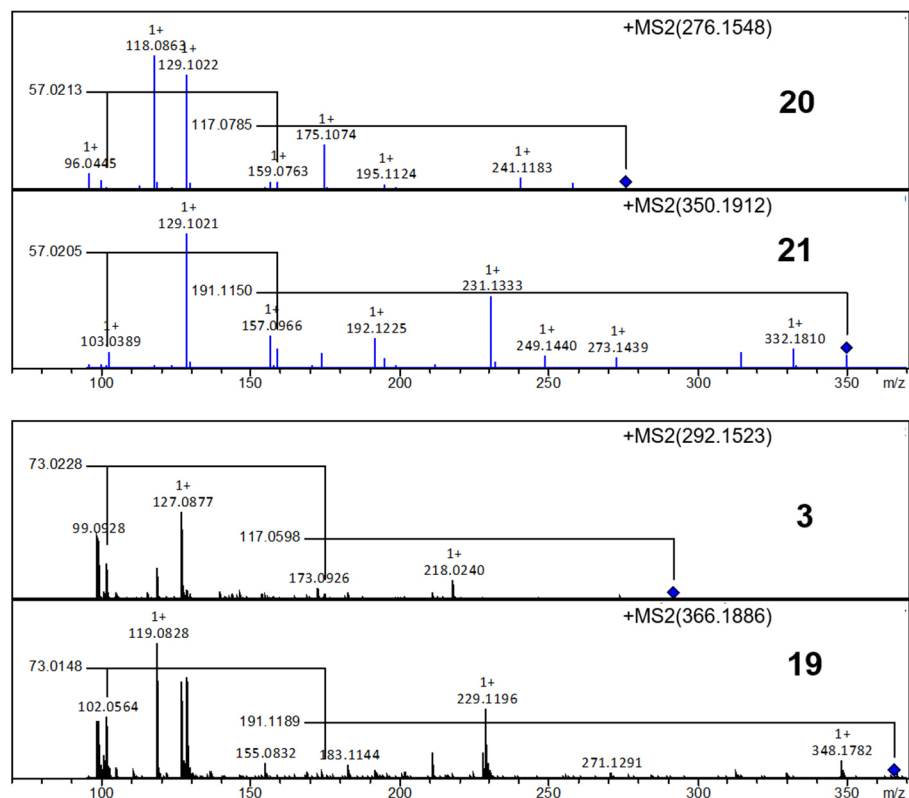
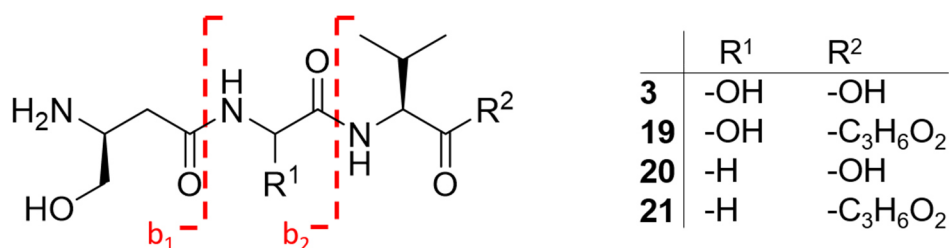


Figure S9. MS² spectra for glyoxypeptide (**3**), glyoxypeptide glycerol derivative (**19**) and non-hydroxylated glyoxypeptide (**20**) and non-hydroxylated glyoxypeptide glycerol derivative (**21**). Fragmentation for b-ions is indicated.

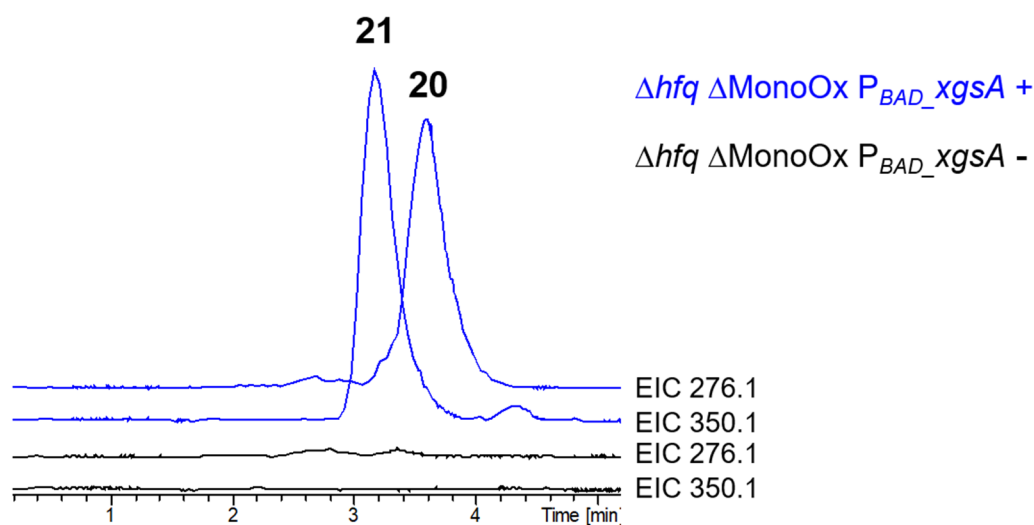


Figure S10. EICs for non-hydroxylated glyoxypeptide (**3**) (m/z 276.1 [M+H]⁺) and the glycerol-derivative of the non-hydroxylated glyoxypeptide (**21**) (m/z 350.1 [M+H]⁺) observed for the induced *X. hominickii* $\Delta hfq P_{BAD_xgsA}$ in standard XPP medium. Used flow rate for the LC-MS analysis was here 0.1 mL/min.



Figure S11. Multiple protein sequence alignment for XgsC_A2_MonoOx domain insertion into PttB_A1 domain. Wild type sequences are PttB_A1 and XgsC_A2. Engineered sequences are PttB_A1_MonoOx and XgsC_A2_ΔMonoOx.

SUPPLEMENTARY INFORMATION

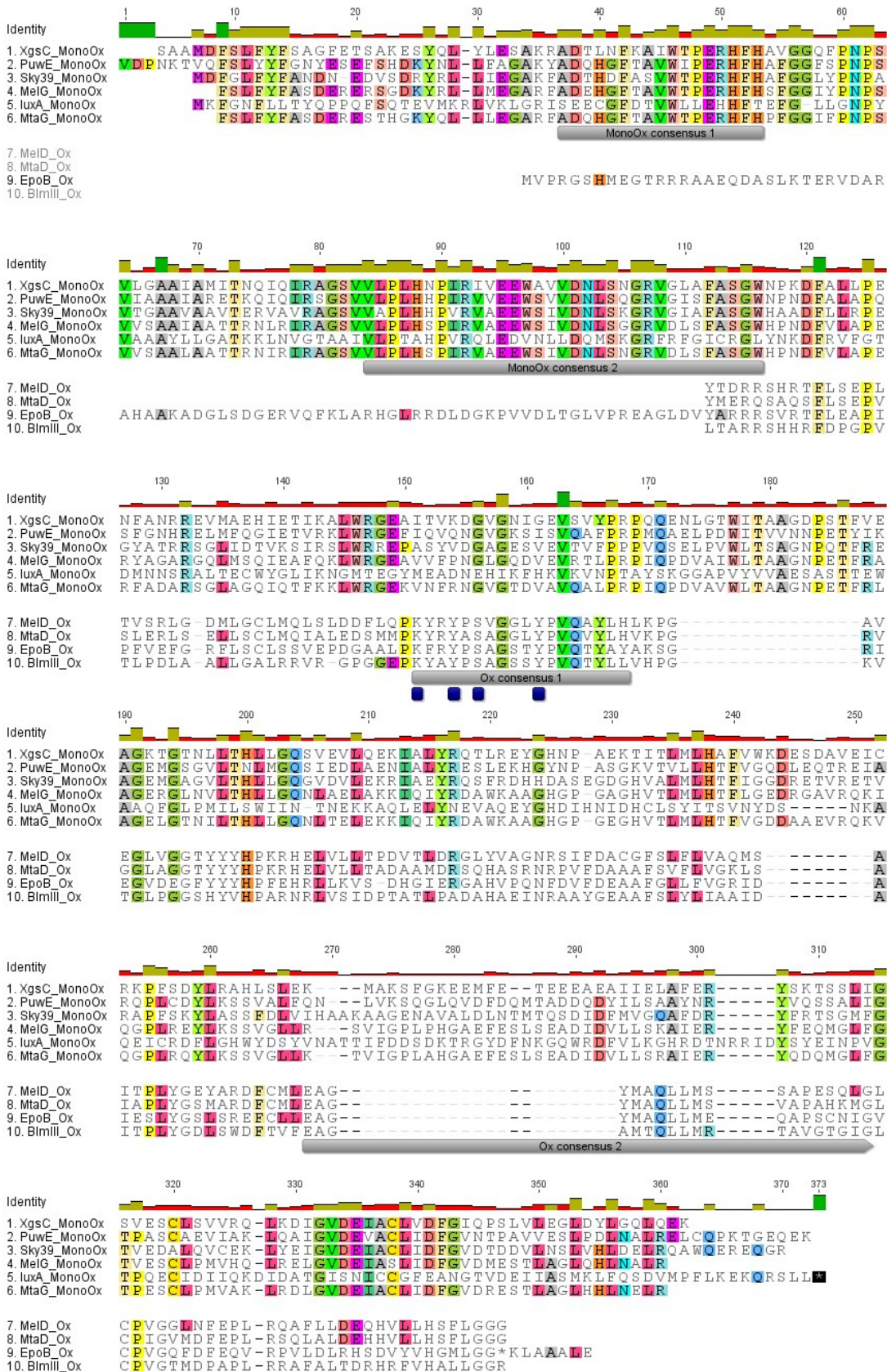


Figure S12. Multiple sequence alignment of selected MonoOx and Ox domains. Grey bars show consensus sequences according to Silakowski *et al.* 1999 and Pang *et al.* 2020.^[189,252] Blue squares indicate putative FMN binding sites in the Ox domains according to Pang *et al.* 2020.^[252] Alignment was generated using Geneious 6.1.8. Please see section 3.5.14 for protein accession numbers.

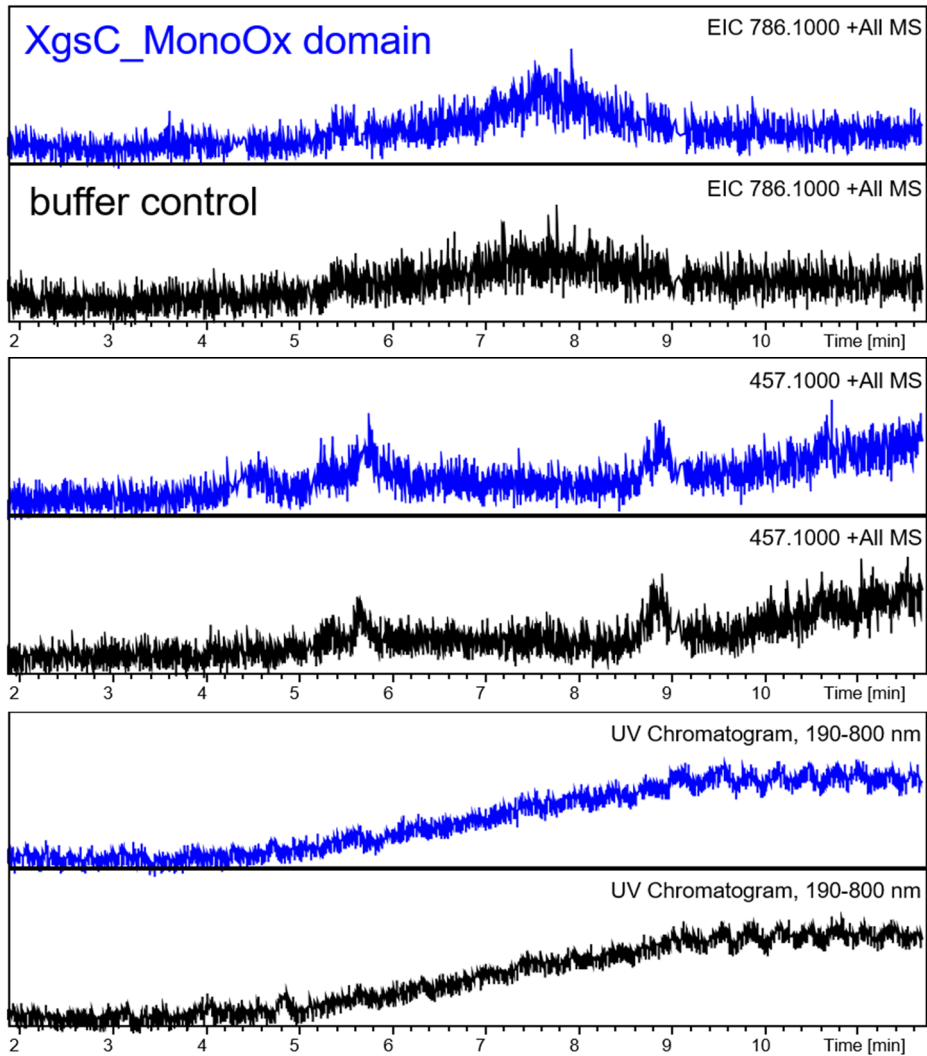


Figure S13. For flavin cofactor investigation, the purified XgsC_MonoOx domain sample was boiled and the supernatant was subjected to HPLC-MS analysis. Buffer A from reverse NiNTA purification was used as a negative control. EICs for FAD (786.1) and FMN (457.1) show no distinct signals and no difference between sample and control. UV-BPC also shows no difference.

SUPPLEMENTARY INFORMATION

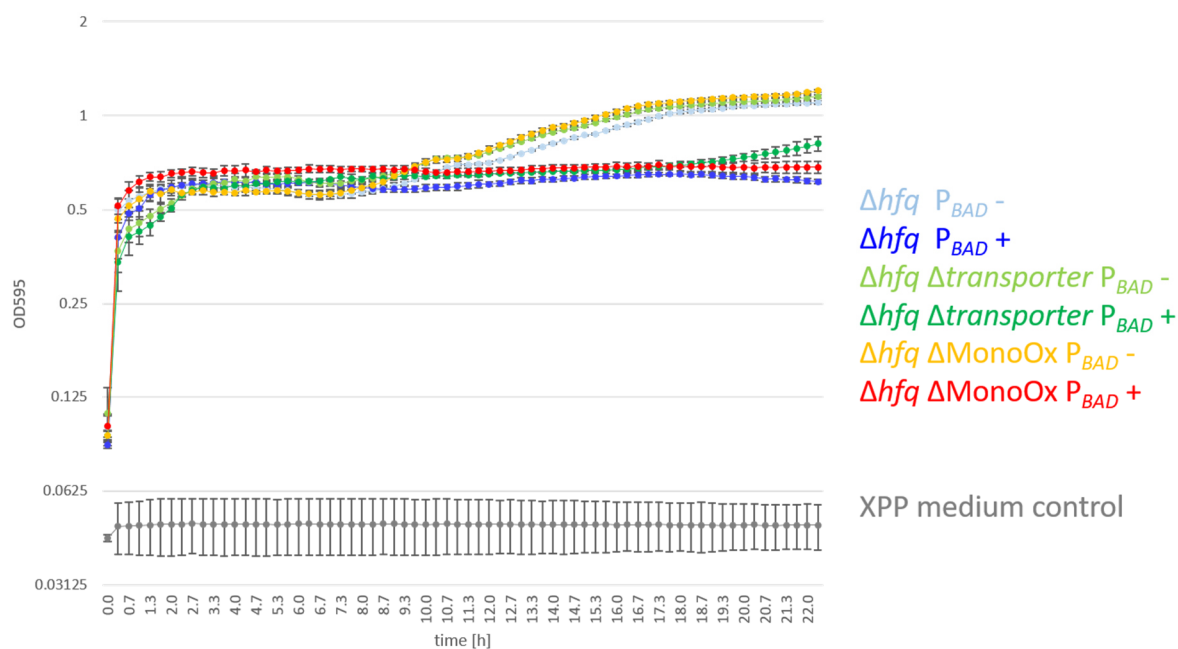


Figure S14. Growth curve with included standard deviation. In triplicates.

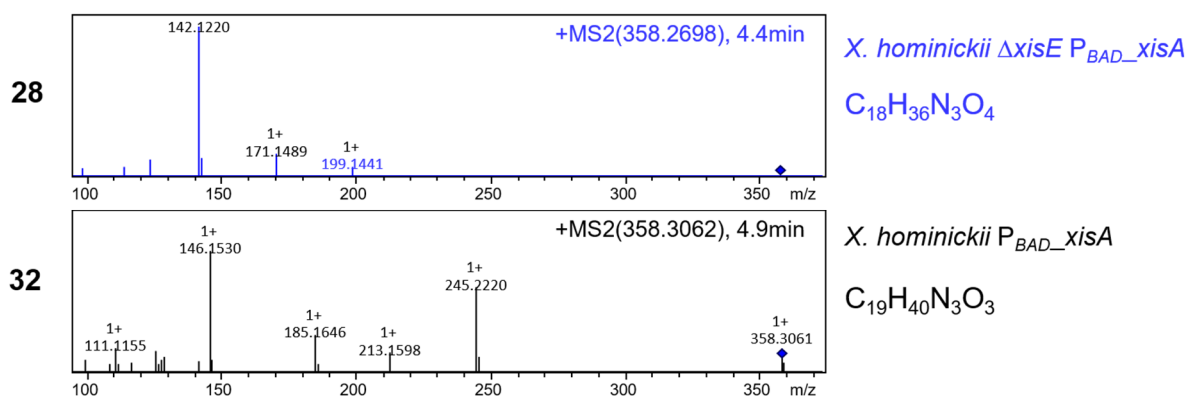


Figure S15. MS² spectra for **28** and **32**. **28** is detected at 4.4 min, shows the fragment ion with m/z 199.1 [M+H]⁺ and is produced in *X. hominickii* $\Delta hfq \Delta xisE P_{BAD_xisA}$ while **32** is detected at 4.9 min, shows the fragment ion with m/z 213.2 [M+H]⁺ and is produced in *X. hominickii* P_{BAD_xisA} .

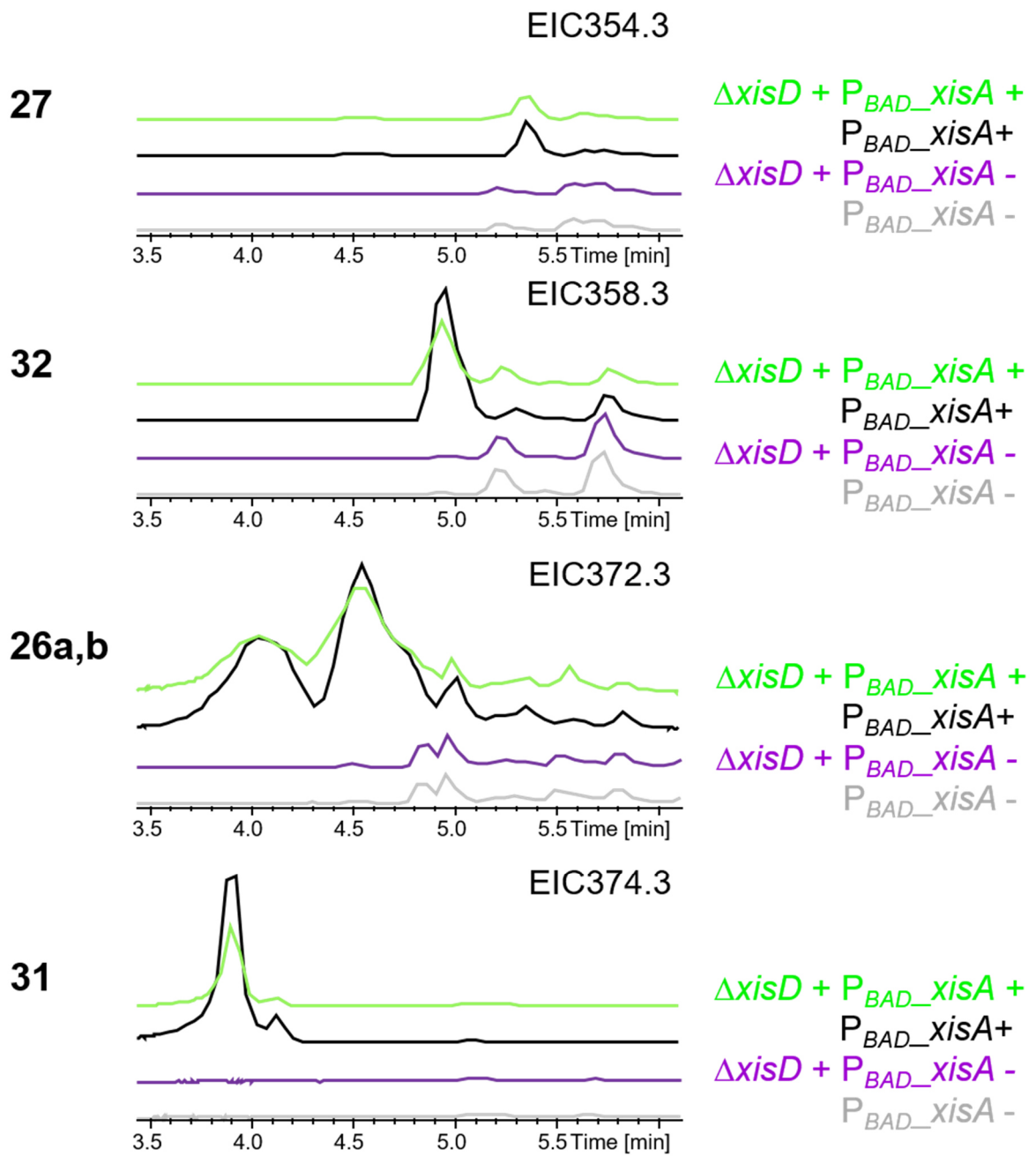
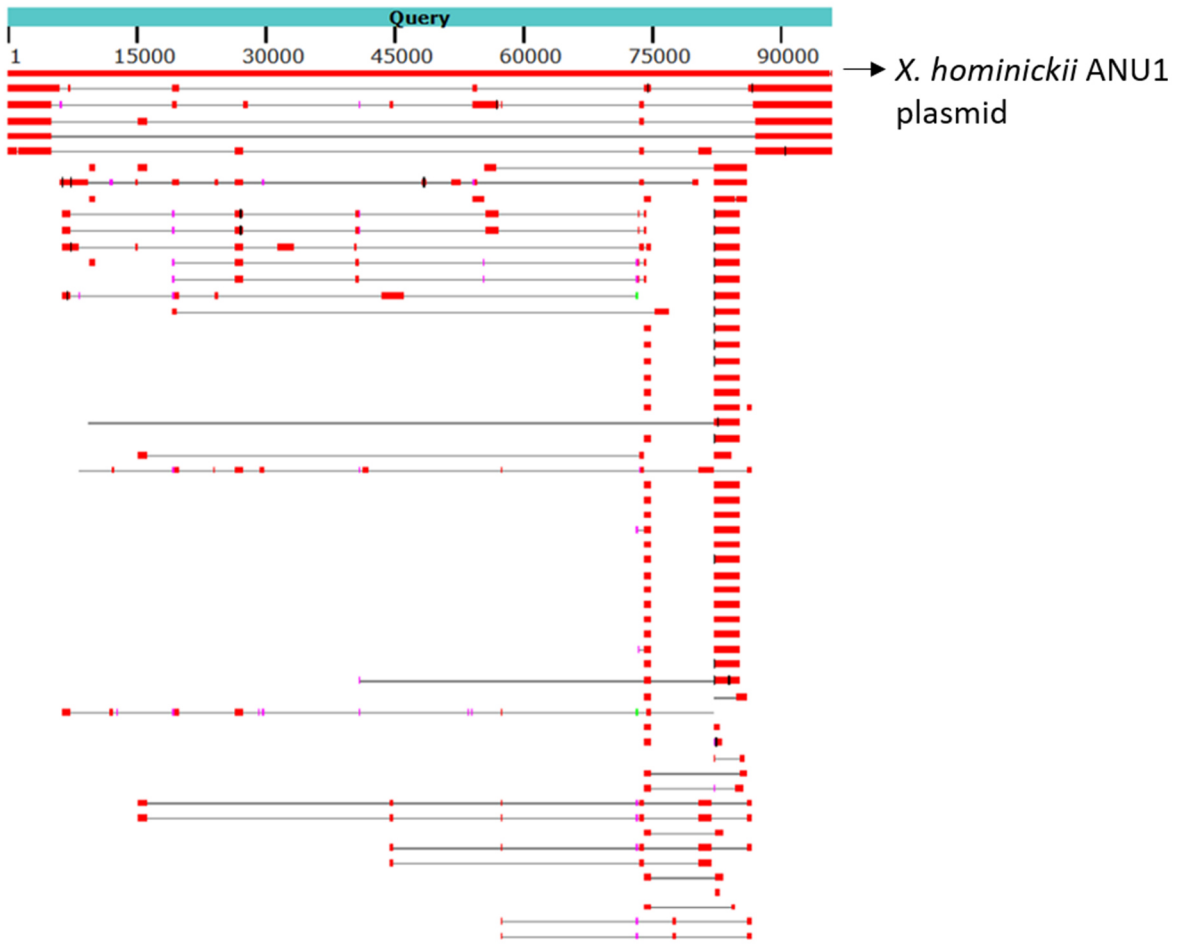


Figure S16. EICs for 26a, 26b, 27, 31 and 32 for P_{BAD} activation in *X. hominickii* WT (black) and *X. hominickii* $\Delta xisD$ (light green). + arabinose induction, - no induction.

Distribution of the top 521 Blast Hits on 100 subject sequences



Xenorhabdus hominickii strain ANU101 plasmid unnamed2 sequence

Sequence ID: [KX517799.1](#) Length: 95915 Number of Matches: 20

Range 1: 287 to 95915 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
1.766e+05 bits(95626)	0.0	95629/95630(99%)	1/95630(0%)	Plus/Plus
Query 1	ATGATGAGAAGCGATAAATATGCCGATTGCCATTATCGGTATGAGTTGCCGTTTCCCGGGG	60		
Sbjct 287	ATGATGAGAAGCGATAAATATGCCGATTGCCATTATCGGTATGAGTTGCCGTTTCCCGGGG	346		
Query 61	GATGTTAGCTCTCCAGAGGAATTCTGGCAGCTGTTGTGGTTCAAGTGACGCAATTACC	120		
Sbjct 347	GATGTTAGCTCTCCAGAGGAATTCTGGCAGCTGTTGTGGTTCAAGTGACGCAATTACC	406		

Figure S17. Snapshot of BlastN analysis with pXhom as query.^[170]

SUPPLEMENTARY INFORMATION

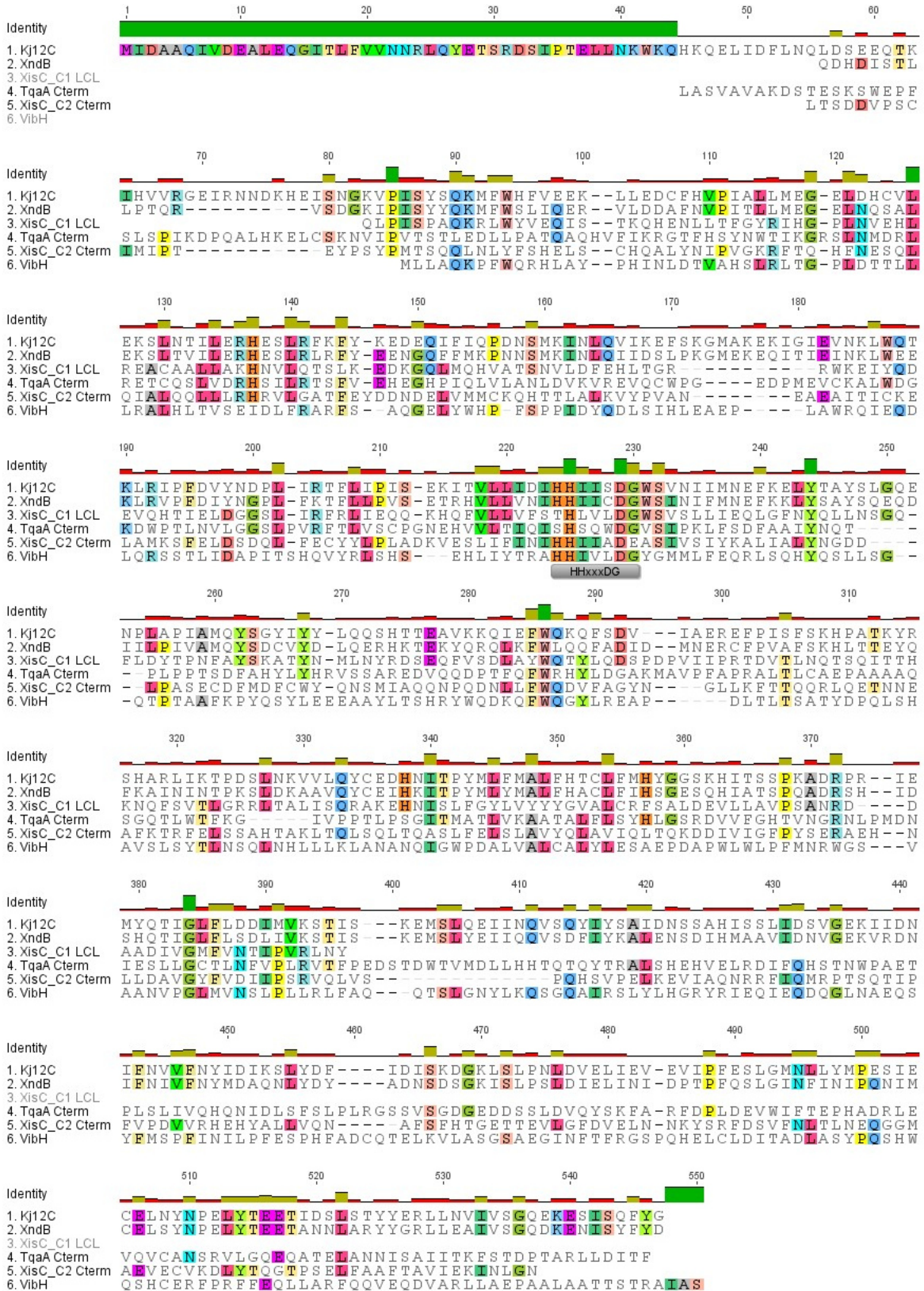


Figure S18. Protein sequence alignment of XisC C domains (XisC_C1_LCL and XisC_C2_Cterm) with sequences of stand alone C domain VibH (vibriobactin biosynthesis), of the terminal condensation domains of XndB (xenortide biosynthesis) and of TqaA (fumiquinazoline biosynthesis) and of the terminal monomodular C domain Kj12C (rhabdopeptide biosynthesis). Active site motif HHxxxDG is indicated in grey. Sequences were aligned using Geneious v 6.1.8.^[178] Please see Material and Methods section 2.5.11 for sequence accession references.

CONFERENCE PARTICIPATION

7 Conference participation

- 09/2018 Gene regulation in prokaryotes VAAM, poster presentation. Margaretha A. Westphalen, Yi-Ming Shi, Helge B. Bode Investigation of possible regulatory effects of natural products in entomopathogenic bacteria
- 03/2019 VAAM annual meeting, poster presentation. Margaretha A. Westphalen, Sebastian L. Wenski, Nadine Keller, Janine Chekaiban Yi-Ming Shi, Helge B. Bode Investigation of natural product biosynthesis in entomopathogenic bacteria
- 09/2019 Biology of microorganisms producing natural products, poster presentation. Margaretha A. Westphalen, Janine Chekaiban Yi-Ming Shi, Helge B. Bode Investigation of natural product biosynthesis in the entomopathogenic bacterium *Xenorhabdus hominickii*
- 10/2020 MegaSyn Symposium on Megasynthases, poster presentation. Identification of two novel NRPS and NRPS/PKS hybrid products in *Xenorhabdus hominickii*
- 01/2021 Advances in Chemical Biology 2021, poster presentation. Identification and biosynthesis of a non-ribosomal tripeptide with unusual amino acid building blocks in the entomopathogenic bacterium *Xenorhabdus hominickii*

8 Acknowledgement

First of all, I would like to thank Prof. Dr. Helge Bode for giving me the opportunity to prepare this thesis. I am very grateful for the supervision, the numerous ideas to advance the individual research projects and the continuous support.

I would like to thank Prof. Dr. Martin Grininger for the assessment of this thesis.

I would like to acknowledge Dr. Yi-Ming Shi and Dr. Carsten Kegler for their constant support and the proof-reading of this thesis. You have encouraged me and this work by many conversations.

I would like to express my gratitude to the former lab members Nadine Kaller, Janine Chekaiban and Dr. Sebastian Wenski who have conducted previous work in *X. hominickii* for characterisation of glyoxypeptide and xildivaline.

I would like to thank Dr. Timo Glatter for the proteomics analysis of *X. szentirmaii* which enabled identification of the xenofuranone methyltransferase.

All present and former lab members I would like to thank for the fantastic work environment and in particular members of lab 0.24, office 0.04, the LC-MS team and the coffee round table for support, inspiration and wonderful memories.

Finally, I would like to say a special thanks to my friends, former and present flatmates and especially to my family for their love and support.

9 Declaration

I herewith declare that I have not previously participated in any doctoral examination procedure in a mathematics or natural science discipline.

Frankfurt am Main,(Date)(Signature)

10 Author's declaration

I herewith declare that I have produced my doctoral dissertation on the topic of

.....
.....
.....

independently and using only the tools indicated therein. In particular, all references borrowed from external sources are clearly acknowledged and identified.

I confirm that I have respected the principles of good scientific practice and have not made use of the services of any commercial agency in respect of my doctorate.

Frankfurt am Main,(Date)(Signature)

11 Curriculum vitae**Personal data**

Name: Margaretha Westphalen

Date of birth: [REDACTED]

Birthplace: [REDACTED]

Nationality: [REDACTED]

Contact data

Address: [REDACTED]

Email: [REDACTED]

Education

01/2018-present PhD student in Prof. Dr. Helge Bode's Lab at Goethe University Frankfurt am Main

10/2015-10/2017 Master of Science: Molecular Biotechnology at the Johann Wolfgang Goethe University Frankfurt, MSc thesis: *Analysis and regulation of natural product biosynthesis in Xenorhabdus* (Institute for Molecular Biosciences, Prof. Dr. Helge Bode's Lab)

10/2012-09/2015 Bachelor of Science: Molecular Biosciences at the Westfälische Wilhelms University Münster, Germany
 BSc thesis: *Interaction studies of Glucose-6-Phosphate/Phosphate Translocator 1 (GPT1) from Arabidopsis thaliana with Thioredoxin h7 and Glutaredoxin c1* (Institute for Plant Biology and Biotechnology, University of Münster)

09/2003-03/2012 A-level, "Abitur", Gymnasium am Kaiserdom, Speyer, Germany

Study abroad

09/2014-12/2014 ERASMUS semester at Université François Rabelais Tours, France

Work experience

12/2016-02/2017 research internship at Hans-Knöll-Institute Jena, Germany. Chemical Biology of Microbe-Host Interactions, Dr. Christine Beemelmann's Lab

05/2012-07/2012 internship in the department of terrestrial ecotoxicology at the BASF SE agricultural centre in Limburgerhof, Germany

Frankfurt am Main, 30th July 2021

Margaretha Westphalen