Proteasome Inhibitors from Actinobacteria
Biosynthesis of the Cyclopropylalanine Moiety of Belactosin A and the Role of Proteasome Inhibitors from *Nocardia* in Pathogen Survival

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Alicia Laura Engelbrecht
aus Filderstadt

Tübingen
2023

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

Tag der mündlichen Qualifikation:	04.08.2023
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter/-in:	Prof. Dr. Leonard Kaysser
2. Berichterstatter/-in:	Prof. Dr. Harald Groß
3. Berichterstatter/-in	PD Dr. Bertolt Gust

Erklärung

Ich erkläre hiermit, dass ich die zur Promotion eingereichte Arbeit mit dem Titel: „Proteasome Inhibitors from Actinobacteria Biosynthesis of the Cyclopropylalanine Moiety of Belactosin A and the Role of Proteasome Inhibitors from *Nocardia* in Pathogen Survival“ selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und Zitate als solche gekennzeichnet habe. Ich erkläre, dass die Richtlinien zur Sicherung guter wissenschaftlicher Praxis der Universität Tübingen (Beschluss des Senats vom 25.5.2000) beachtet wurden. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

Wien, den

Content

Publications and Presentations

Publications:

Engelbrecht, A., Saad, H., Gross, H., & Kaysser, L. (2021). „*Natural Products from Nocardia and Their Role in Pathogenicity". Microbial Physiology"*, 1-16.

Shimo, S., Ushimaru, R., Engelbrecht, A., Harada, M., Miyamoto, K., Kulik, A., Kaysser L. & Abe, I. (2021). „Stereodivergent Nitrocyclopropane Formation during Biosynthesis of Belactosins and Hormaomycins", *Journal of the American Chemical Society*, *143*(44), 18413-18418.

Engelbrecht A., Wolf F., Esch A., Kulik A., Kozhuskov SI, de Meijere A, Hughes CC and Kaysser L. „*Discovery of a cryptic nitro intermediate in the biosynthesis of the 3-(trans-2'-aminocyclopropyl)alanine moiety of belactosin A " Organic Letters* 24.2 (2022): 736-740.

Beller P., Fink PJ, Helmle I, Männle D, Unterfrauner D, Wolf F, Engelbrecht A, Staudt N, Kulik A, Groß H and Kaysser L (2023). „*Characterization of the Cystargolide Biosynthetic Gene Cluster."* Manuscript, in preparation

Scientific Presentations (selected):

GRK1708 Final Symposium, Tuebingen, October 2020: *"Biosynthesis of the cyclopropyl moiety of belactosin A"*

International VAAM Workshop 2019, Biology of Bacteria Producing Natural Products (VAAM), Jena, September 2019: *„Biosynthesis of the cyclopropyl moiety of the proteasome inhibitor Belactosin A"*

IGIM 3rd Summer School "Microbes, Host and Infection", Bad Urach, July 2018: *„Proteasome Inhibitors from Nocardia spp. and their Role in Pathogen Survival"*

CERMEL, Gabon, Center of Medical Research Lambaréné, Gabun, Afrika, March 2018: "*The Discovery of Natural Products and their Development as anti-infectives."*

Poster presentations:

GIM, Genetics of industrial Microorganisms, Pisa, September 2019: *"Biosynthesis of the cyclopropyl moiety of belactosin A*"

2nd minisymposium of the GRK1708 "bacterial storage compounds", Tuebingen, July 2019: *"Proteasome Inhibitors from Nocardia spp. and their Role in Pathogen Survival":*

VAAM Biology of Bacteria Producing Natural Products Frankfurt, September 2018: "*Proteasome Inhibitors from Nocardia spp. and their Role in Pathogen* Survival"

Annnual Conference of the Association for General and Applied Microbiology, Wolfsburg, 2018: *"Proteasome Inhibitors from Nocardia spp. and their Role in Pathogen Survival"*

30. Irseer Naturstofftage, Kloster Irsee, February 2018: "*Biosynthesis of the cyclopropyl moiety of belactosin A*"

VAAM Biology of Bacteria Producing Natural Products, Tuebingen, September 2017: *"Proteasome Inhibitors from Nocardia spp. and their Role in Pathogen Survival"*

1st Minisymposium of the RTG1708: "Stress Induced Metabolic Pathways", Tuebingen, July 2017:" *"Proteasome Inhibitors from Nocardia spp. and their Role in Pathogen Survival"*

Contributions of other scientists to this work

Felix Wolf:

- Construction and screening of fosmid library for *Streptomyces* sp. UCK14
- Generation and analysis of heterologous border cluster knock-out strains *S. albus* bel01/Δorf-1, ΔbelA-E, belO-V, ΔbelQ-V, ΔbelR-V, ΔbelT-V
- Generation of heterologous expression strains *S. albus* bel01, *S. albus* bel01/Δ*belK, S. albus* bel01/Δ*belL, S. albus* bel01/Δ*belM, S. albus* bel01/Δ*belN.*
- Cloning and conjugation of knock-out strains *Streptomyces* sp. UCK14/Δ*belK*, UCK14/Δ*belL,* UCK14/Δ*belN*
- Cloning of pUWL_belN

Chambers Hughes and Annika Esch:

- Chemical synthesis, derivatization and analysis of lactone intermediate
- LC-MS analysis of dansylchloride derivatized compounds

Abbrevations

°C	degree Celsius
µ	micro
α-KG	α-ketoglutarate
aa	amino acids
aac(3)IV	apramycin resistance gene
ACP	acyl carrier protein
Acpa	3-(*trans*-2'-aminocyclopropyl)alanine
Amp	ampicillin
Apra	apramycin
ASST	arylsulfate sulfotransferase
ATP	adenosine triphosphate
BACs	Bacterial Artificial Chromosome
BHI	Brain Heart Infusion broth
BGC	biosynthetic gene cluster
bla	ampicillin/carbenicillin resistance gene
bp	base pair
BPC	base peak chromatogram
CFU	colony forming unit
Cm	chloramphenicol
d_2-Acpa	3,3-d_2-3-(*trans*-2'-aminocyclopropyl)alanine
d_2-Ncpa	3,3-d_2-3-(*trans*-2'-nitroocyclopropyl)alanine
Da	Dalton
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside 5´-triphosphate
dsDNA	double-stranded DNA
dTDP	deoxythymidine triphosphate
DTT	1,4-dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylendiamine tetra-acetic acid
EIC	extracted ion chromatogram
ESI	electrospray ionization
EtoAc	Ethylacetate
FRT	FLP recognition target
h	hour
HCl	hydrochloric acid
HCOOH	formic acid
HIS	histidine
HPLC	high performance liquid chromatography
IPTG	isopropyl-β-thiogalactoside
IMAC	Immobilized metal affinity chromatography
k	kilo
KAc	potassium acetate
Kan	kanamycin
kb	kilo base pairs
kDa	kilo Dalton

l	litre
LC	liquid chromatography
M	molar
M.	*Mycobacterium*
m	milli
min	minute
MS	mass spectrometry
MW	molecular weight
m/z	mass-to-charge ratio
n	nano
NaOH	sodium hydroxide
Ncpa	3-(*trans*-2'-nitroocyclopropyl)alanine
neo	neomycin/kanamycin resistance gene
nt	nucleotide
NP	natural product
OD_{600}	optical density at 600 nm
ORF	open reading frame
oriT	origin of transfer
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethylsulfonyl fluoride
PPM	parts per million
^{R}	resistant
RNA	ribonucleic acid
RNase	ribonuclease
RP	reverse phase
rpm	rounds per minute
s	second
S.	*Streptomyces*
SAM	S-adenosylmethionine
SDS	sodium dodecyl sulfate
sp.	Species
TAR	Transformation-Associated Recombination
TEMED	N,N,N′,N′-tetramethylethylenediamine
TES	N-Tris-(hydroxymethyl)-methyl-2-aminoethanesulfonic acid
TIC	total ion chromatagram
Thr	threonin
Tris	Tris(hydroxymethyl)aminomethan
U	unit
S. sp. UCK14	*Streptomyces* sp UCK14
UCK14	*Streptomyces* sp UCK14
UV	ultraviolet
VLC	vaquum liquid chromatography
WT	wild-type
×*g*	ground acceleration
X-gal	5-bromo-4-chloro-3-indolyl-*β*-D-galactopyranoside

Summary

Natural product research is a fruitful area of investigation for the discovery of new drugs and bioactive molecules. The biosynthesis of proteasome inhibitor associated natural products has gained attention in recent years, including the class of β-lactone-containing peptide natural product proteasome inhibitors. Belactosin A and hormaomycins are such peptide natural products containing 3-(2-aminocyclopropyl)alanine (Acpa) and 3-(2- nitrocyclopropyl)alanine (Ncpa) residues, respectively.

In the first project of this thesis I shed light on the formation of the unique Acpa moiety of belactosin A. I was able to establish heterologous pathway expression of the biosynthetic gene cluster of the belactosin series from *Streptomyces* sp. UCK14, including border cluster knock-outs. We could thereby show that the biosynthetic gene cluster of belactosins reaches from genes *belF* to *belV* and the four gene operon *belK-N* is responsible for building the Acpa moiety of belactosin A. Construction of a set of gene deletion and feeding experiments for chemical complementation that include incorporation of stable isotope-labeled precursors clearly demonstrated that in the biosynthesis of the Acpa building block, a cryptic nitrocyclopropylalanine (Ncpa) intermediate is generated from L-lysine. Additionally, I could demonstrate that the subsequent reduction of the *N*-oxygenated precursor Ncpa to the corresponding amine compound Acpa is mediated by the molybdopterin-dependent enzyme BelN. Feeding of double deuterium labeled Acpa to *Streptomyces* sp. UCK14/Δ*belN* fully restored belactosin A production. These results were further verified by genetic complementation of *Streptomyces* sp. UCK14/*belN* with pUWL_*belN.* Additionally feeding of $^{13}C_6$-L-lysine to the *Streptomyces* sp. UCK14/*belN* mutant resulted in the accumulation of the Ncpa intermediate. The determination of the actual enzymatic mechanism for the reduction of Ncpa to Acpa will be an intriguing subject for future investigations and further studies are currently ongoing.

In parallel to our studies, we started a close collaboration with the research group around Prof. Dr. Abe from the University of Tokyo. In agreement with our initial in vivo feeding studies, they could enlighten the function of heme oxygenase-like enzyme BelK and nonheme iron enzyme BelL in detail.

The second half of this work focuses on natural products from *Nocardia* and their potential connection to the virulence of pathogenic strains. *Nocardia* spp. are filamentous actinobacteria and can cause localized and systemic infections in humans. However, the

virulence mechanisms of human pathogenic *Nocardia* and the progression of nocardiosis are poorly understood. As more *Nocardia* strains are genome sequenced, genome analysis tools, such as genome mining approaches, provide access to the potential for the production of novel and unique natural products. The genome search of the human pathogenic strain *N. cyriacigeorgica* GUH-2 provided us with a total of 19 biosynthetic gene clusters. A particular gene cluster from *N. cyriacigeorgica* GUH-2 has spiked our interest because this cluster contains a highly conserved acyl-CoA dehydrogenase (ACAD) with homologies to the epoxyketone synthase from the proteasome inhibitor-associated biosynthetic pathways of epoxomicin. In order to gain access to the biosynthesis cluster declared as ACAD cluster in this thesis and to the genome of the *N. cyriacigeorgica* GUH strain in general, we created and systematically screened a fosmid-based genomic library. After successfully identifying the ACAD gene cluster, we expressed it heterologously in a suitable recombinant host systems. This heterologous expression followed by purification revealed a potentially novel compound with mass *m/z* 487.1599 $[M+H]^+$. I was able to purify about 1.3 mg pure substance of this compound for further NMR structure elucidations.

Zusammenfassung

Die Naturstoffforschung ist ein ertragreiches Forschungsgebiet für die Entdeckung neuer Wirkstoffe und bioaktiver Moleküle. Die Biosynthese von Proteasom-Inhibitor assoziierten Naturstoffen hat in den letzten Jahren an Aufmerksamkeit gewonnen, einschließlich der Klasse der β-Lacton-haltigen Peptid-Naturstoff-Proteasom-Inhibitoren. Belactosin A und Hormaomycine sind Peptid-Naturstoffe, die 3-(2-Aminocyclopropyl)alanin- (Acpa) beziehungsweise 3-(2-Nitrocyclopropyl)alanin- (Ncpa) Reste enthalten.

Im ersten Projekt dieser Arbeit habe ich die Biosynthese der einzigartigen Acpa-Einheit von Belactosin A beleuchtet. Ich konnte die heterologe Pathway-Expression des biosynthetischen Genclusters der Belactosin-Serie von *Streptomyces* sp. UCK14 nachweisen, einschließlich Knock-outs zur Bestimmung der Clustergrenzen. Wir konnten zeigen, dass das biosynthetische Belactosin Gencluster von den Genen *belF* bis *belV* reicht und dass, das vier-Gen-Operon *belK-N* für die Synthese der Acpa-Einheit von Belactosin A verantwortlich ist. Die Konstruktion einer Reihe von Gen-Deletions- und Fütterungsexperimenten zur chemischen Komplementierung, welche den Einbau von stabilen Isotopen markierten Vorläufemolekülen umfassen, zeigte deutlich, dass bei der Biosynthese des Acpa-Bausteins aus L-Lysin ein kryptisches Nitrocyclopropylalanin (Ncpa)-Zwischenprodukt entsteht. Darüber hinaus konnte ich zeigen, dass die anschließende Reduktion des *N*-oxygenierten Vorläufers Ncpa zum entsprechenden Amin der Acpa Untereinheit durch das Molybdopterin-abhängige Enzym BelN vermittelt wird. Ich konnte deutlich zeigen, dass durch die Fütterung von doppelt Deuterium markiertem Acpa in *Streptomyces* sp. UCK14/Δ*belN* die Belactosin A Produktion vollständig wiederhergestellt werden konnte. Diese Ergebnisse wurden durch genetische Komplementierung von. *Streptomyces* sp. UCK14/Δ*belN* mit pUWL_*belN* untermauert. Zudem konnten wir zeigen, dass die Fütterung von $^{13}C_6$-L-Lysin in *Streptomyces* sp. UCK14/*belN* zur Akkumulation des Ncpa-Zwischenprodukts führt. Die Bestimmung des tatsächlichen enzymatischen Mechanismus für die Reduktion von Ncpa zu Acpa wird ein interessantes Thema für zukünftige Untersuchungen sein, und weitere Studien laufen derzeit.

Parallel zu unseren Untersuchungen gingen wir eine enge Zusammenarbeit mit der Forschungsgruppe um Prof. Abe von der Universität Tokio ein. In Übereinstimmung mit den Ergebnissen unserer initialen in-vivo-Fütterungsstudien konnten sie die Funktion des

Hämoxygenase-ähnlichen Enzyms BelK und des Nicht-Häm-Eisenenzyms BelL im Detail aufklären.

Die zweite Hälfte dieser Arbeit konzentriert sich auf Naturstoffe aus *Nocardia* und ihrer potentiellen Verbindung zur Virulenz pathogener Stämme. *Nocardia* spp. sind filamentöse Actinobakterien und können beim Menschen lokalisierte und systemische Infektionen verursachen. Die Virulenzmechanismen der humanpathogenen *Nocardia* und das Fortschreiten der Nokardiose sind jedoch kaum verstanden. Da immer mehr vollständige Genomsequenzierungen von *Nocardia* Stämme zur Verfügung stehen, ermöglichen Genomanalysen, wie zum Beispiel, Genome-Mining-Ansätze, Zugang zum Produktionspotential neuartiger und besonderer Naturstoffe. Die Genomanalyse des humanpathogen Stammes *N. cyriacigeorgica* GUH-2 lieferte uns insgesamt 19 biosynthetische Gencluster. Ein bestimmtes Gencluster von *N. cyriacigeorgica* GUH-2 hat unser Interesse geweckt, da eine hoch konservierte Acyl-CoA-Dehydrogenase (ACAD) mit Homologien zu der Epoxyketone Synthase aus den Proteasom-Inhibitor-assoziierten Biosynthesewegen von Epoxomicin in diesem Cluster kodiert ist. Um Zugang zu in dieser Thesis als ACAD-Cluster deklariertes Biosynthese Cluster und zum Genom des *N. cyriacigeorgica* GUH Stammes im Allgemeinen zu erhalten, habe ich eine Fosmid-basierte Genombibliothek erstellt und systematisch gescreent. Nach erfolgreicher Identifizierung des ACAD-Genclusters, exprimierte ich dieses Cluster heterolog in geeigneten rekombinanten Wirtssystemen. Diese heterologe Expression mit anschließender Aufreinigung lieferte uns eine potentiell neuartige Verbindung mit der Masse *m/z* 487.1599 $[M+H]^+$. Es ist mir gelungen etwa 1.3 mg Reinstoff dieser Verbindung für weiterführende NMR Strukturaufklärungen zu reinigen.

1 Introduction

Introduction I: Biosynthesis of the cyclopropylalanine (Acpa) moiety of the proteasome inhibitor belactosin A

1.1 Ways of uncovering potential in natural product discovery

Natural products have been one of the fundamental source for the discovery of novel compounds with diverse biological activities.[1, 2] Historically, natural products played a key role in drug discovery, especially for infectious diseases and cancer and are therefore highly valuable for medical applications or the pharmaceutical industry.[3, 4] From 1981 to date 64.9% of the small-molecule chemicals which were introduced as new chemical entity drugs are derived from natural products.[5]

However, in the last decade the pharmaceutical industry switched its‘ interest towards combinatorial chemistry, high- throughput screening of synthetic chemical libraries and computer assisted drug design, as screening of natural product libraries and strain collections is time intensive and often not profitable.[6] However, natural products still represent unprecedented structural diversity and are regarded as privileged scaffolds for drug lead development.[7] Several current technological and scientific developments are still revitalizing interest in natural products as drug leads as which include improved analytical tools, genome mining and genetic engineering and are opening up new opportunities.[1]

The development of novel drugs is crucially needed, as infectious diseases caused by bacteria, parasites, fungi or viruses are still a major cause of death worldwide.[8] Interestingly, the therapeutic agents that are used to combat such infections are also predominantly derived from microorganisms. Thereby, actinobacteria are one of the most important producers of small molecules with pharmaceutical value. They show antibacterial, antiviral, anticancer, antiparasitic, or immunosuppressing activities.[9-11]

The biosynthesis of these natural products is directed by defined sets of genes that are organized in biosynthetic gene clusters (BGCs). Those BGCs are a spatial grouping of all structural genes that are responsible for the assembly of a natural product and are encoding for e.g. polyketide synthases (PKS), non-ribosomal peptide synthetases (NRPS), and hybrid NRPS/PKS machineries. Those biosynthetic genes are usually clustered together with genes that mediate pathway-specific regulation, export and self-resistance.[12, 13] Such compounds,

produced as secondary metabolites, are often crucial elements in bacterial survival strategies e.g. as "defense weapons" against other microorganisms or within a host organism. Additionally, these molecules can mediate inter-species and intra-species communication or provide an advantage in nutrient acquisition.[14, 15]

As more and more species are whole-genome sequenced the number of predicted clusters of genes for natural product biosynthesis is permanently increasing. To fully utilize the potential of novel bioactive compounds and chemical entities, it is necessary to link the natural products to the genes that are responsible for producing them. Strategies for triggering activation of silent biosynthetic gene clusters (e.g. environmental stimuli, global regulators, and promotor exchange), gene deletion or disruption strategies, heterologous gene expression, comparative genomics or retro *in silico* biosynthesis are just a few approaches to uncover potential in natural product discovery.[1, 16-19] Genome mining can be used to search microbial genome sequences for cryptic or orphan biosynthetic gene clusters, which may have the potential to direct the production of novel and structurally complex natural products.[20-22]

Identification of such novel natural products can lead to the discovery of promising drug leads with applications in cancer, inflammatory infectious and neurodegenerative disease or cardio metabolic associated disease, to name the most prominent. Natural product derived compounds with their chemical diversity and biochemical specificity thereby have very diverse molecular therapeutic targets.[23]

Within this thesis the focus will be on natural product derived molecules that target the eukaryotic ubiquitin proteasome system. Compounds with proteasome inhibitory activity are an important class of drugs for the first line treatment of diseases like multiple myeloma or mantle cell lymphoma. However, they are also being investigated for other diseases.[24]

1.2 The ubiquitin proteasome system

The 26S proteasome is the main non-lysosomal large enzymatic complex for proteolytic protein degradation in eukaryotes.[25, 26] As the central protease of the ubiquitin-dependent signaling pathway, the proteasome degrades an enormous variety of proteins that contain specific degradation signals.[27, 28] Thereby, unneeded or misfolded proteins are marked with an ubiquitin tag, which is recognized by the 19S regulatory particle (19S cap) and subsequently funneled into the proteolytic active core of the proteasome (20S core) (Figure 1).

The ubiquitination of proteins is carried out by sequential activities of specific ubiquitin activating enzymes (E1), ubiquitin conjugating enzymes (E2), and ubiquitin protein ligases (E3) (Figure 1). The E1 enzyme activates the highly conserved, 76 amino acid polypeptide ubiquitin in an ATP-dependent manner. In this process, E1 forms a covalent bond between the cysteine residue in the active site of the enzyme and C-terminus of ubiquitin. Therefore, ubiquitin is thioesterified and passaged to the ubiquitin-conjugating enzyme E2. In the last step of the cascade, E3, an ubiquitin ligase, binds the ubiquitin onto the protein substrate. [29, 30] Besides poly-ubiquitin as the main tag for degradation, the 26S proteasome recognizes diverse E3 ligases substrates harboring various degradation signals and contribute to selectivity and specificity of the ubiquitin proteasome system.[31]

The 26S proteasome is a large protein complex that can be divided into two sub-complexes: The catalytic core particle (20S) of approximately 700 kDa and the 19S cap regulatory particle of approximately 900 kDa. Both are built of a set of multiple distinct subunits.[32-34] The 19S cap consists of at least 19 different subunits and can generally be divided into the two sub-complexes base and lid.[35] The 20S core consists of four heteroheptameric axial stacked rings, built out of two outer α-rings and two inner β-rings, each comprising seven structurally similar α- and β-subunits, respectively. Thereby, the β-rings form the proteolytic active chamber, and the α-rings serve as a gate for the entry into the chamber. Out of all β-subunits, β1, β2 and β5 have hydrolytic activity and act as Threonin (Thr) proteases. These three active sites are located on the inner surface of the 20S proteolytic chamber and their activities are referred to as caspase-like, trypsin-like and chymotrypsin-like, respectively.[34] However, the maturation of the active β-subunits is independent from the presence of other active subunits and occurs by intra-subunit threonine-dependent posttranslational autolysis of the *N*-terminal propeptide, resulting in freely accessible active sites.[36, 37]

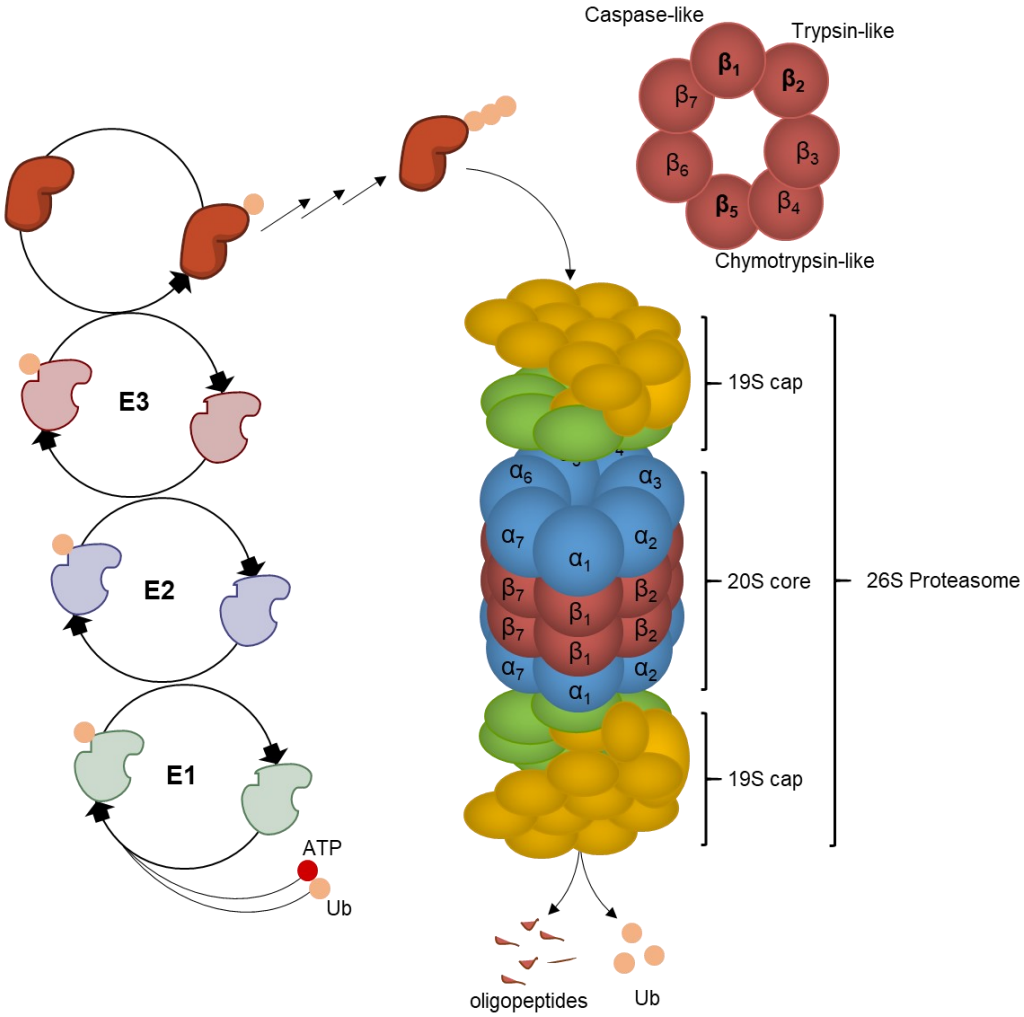


Figure 1: The ubiquitin proteasome system: First ubiquitin is activated by **E1** (ubiquitin-activating enzyme) followed by ubiquitin delivery to **E2** (ubiquitin-conjugating enzyme). The second step involves complex formation by **E3** (ubiquitin ligase) and the substrate. In the third step, ubiquitin molecules are transferred to the substrate to mark the substrate with a polyubiquitin chain. Next, the polyubiquitinated substrate is released from E3. The proteasome recognizes the polyubiquitinated protein, unfolds the substrate in an ATP-dependent manner and removes the ubiquitin chain through a proteasome-associated ubiquitin hydrolase activity. Subsequently, the unfolded protein is funneled into the proteasome chamber with the proteolytic active sites. The recycled ubiquitin molecules and generated oligopeptides are released. The **26S proteasome** consists of the catalytic core particle (20S) and the two regulatory particles (19S cap). The 20S core consists of two outer α- and two inner β-rings, which are stacked in a heteroheptameric manner. Subunits **β1**, **β2** and **β5** have caspase-like, trypsin-like and chymotrypsin-like activity, respectively.

By not only degrading misfolded proteins, but also important regulators in cell metabolism, the 26S proteasome plays an essential role in many cellular processes, including protein quality control and regulation of cell-cycle progression by regulating cyclin and cyclin-dependent kinases. This system also plays an important role in the regulation of key transcription factors like NF-κB or cellular tumor antigen p53. Furthermore, DNA repair,

apoptosis, immune response or even inflammation are regulated via the proteasome ubiquitin system.[24] It thereby precisely catalyzes and regulates the level of key factors, ensuring unidirectional progression of those biological processes, as eukaryotic proteostasis is very important for cell development, cellular ageing, and protection of cells against disease.[38, 39]

This mediation is essential in cellular processes that are deregulated in cancer cells and central elements in carcinogenesis and tumorigenesis. Thus, the inhibition of the proteasome specifically targets heavily proliferating cells over dormant cells, which has made the proteasome an established target in cancer therapy.[40-42]

1.3 Proteasome inhibitor molecules

Due to the versatile application possibilities of proteasome inhibitors in medical treatments or even biotechnology, the investigation of the biosynthesis of small molecules is a promising field of research. Interestingly, of the major structural classes of inhibitors of eukaryotic proteasomes such as aldehydes, β-lactones, epoxyketones, syrbactins, and cyclic peptides either were discovered as natural products or have natural products among them. To facilitate their bioactivity and inhibitory nature, the key feature of proteasome inhibitor molecules are specialized moieties. For example, proteasome inhibitors often possess electrophilic moieties which specifically bind to the proteolytic subunits of the proteasome, resulting in inhibition and loss of activity. Such electrophilic groups cover Michael acceptor systems, ring-strained scaffolds β-lactones or epoxides and various other groups. However, for many of those natural product derived compounds the biosynthetic principal of such moieties remains elusive. Their very promising bioactivity has boosted ongoing approval of proteasome inhibitors as cancer drugs and the associated research including development of molecules with improved properties is an ongoing effort.[43-47]

The first proteasome inhibitor bortezomib (Figure 2) that came to the market as Velcade® by Millenium Pharmaceuticals, was approved by the US Food and Drug Administration (FDA) in 2003 (EMA approval in April, 2004). It is currently applied as a first-line treatment for multiple myeloma and mantle cell lymphoma.[48, 49]

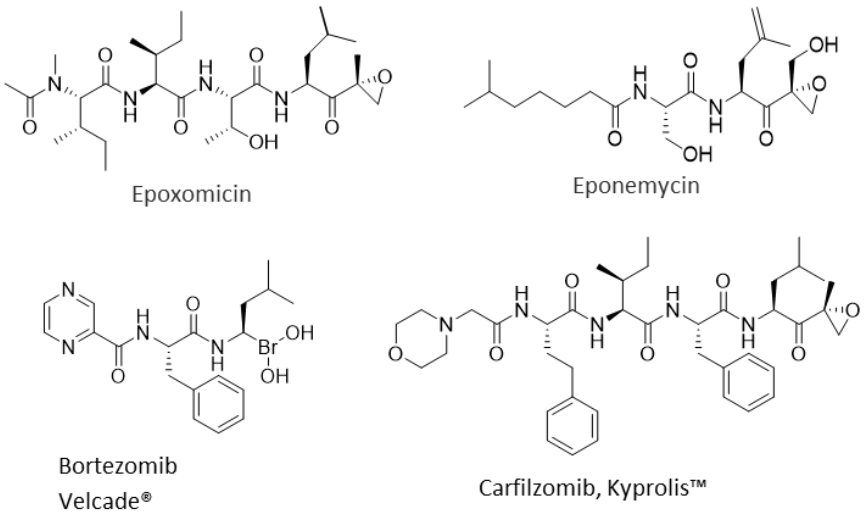


Figure 2: Chemical structures of representative proteasome inhibitor molecules and their marketed drugs

A prominent class of proteasome inhibitors are peptide epoxyketones which include epoxomicin[50] and eponemycin.[51] Epoxomicin was the lead structure for the development of

carfilzomib (Kyprolis™, Onyx Pharmaceuticals), which was granted accelerated approval by the FDA in July 2012 (EMA approval in November, 2015) for the treatment of refractory and relapsed multiple myeloma. Kyprolis was designated an orphan medicine, which resembles medicine used in rare diseases.[52, 53] The drug appears to be better tolerated with less side effects by patients than bortezomib and can therefore be applied in higher and more effective doses.[54, 55] Besides their usage as anticancer drugs, epoxyketone containing compounds have also shown to have excellent activity against parasites[56] such as *Plasmodium falciparum*, one of the causative agents of malaria, which is vitally affected by this class of proteasome inhibitors at different stages of its life cycle.[57]

1.4 β-lactone containing proteasome inhibitors

To date literature describes more than thirty known core scaffolds of β-lactone containing natural products with interesting bioactivities. In this thesis the focus manly lies on the inhibition of the 26S proteasome, but β-lactone-containing molecules can show a wide variety of different biological functions. Many compounds show potent bioactivity against bacteria, fungi, or human cancer cell lines.[58] An interesting example for such novel compounds are lipase inhibitors, like lipstatin (Figure 3), which was isolated from *Streptomyces toxytricini*, as an irreversible inhibitor of human pancreatic lipase. Lipstatin is the precursor of a saturated derivative, orlistat, an established FDA approved drug for the treatment of obesity. [59, 60] The also β-lactone containing ebelactones, produced by *S. aburaviensis* ATCC 31860, have shown to be inhibitors of esterases, lipases, cutinases, cathepason, homo serin tranysacetlyases and additionally to be able to modulate the mTOR pathway.[61] Another example are ClpP peptidases, which are able to catalyze intracellular protein turnover in bacteria, including human pathogens such as *Mycobacterium tuberculosis*.[62]

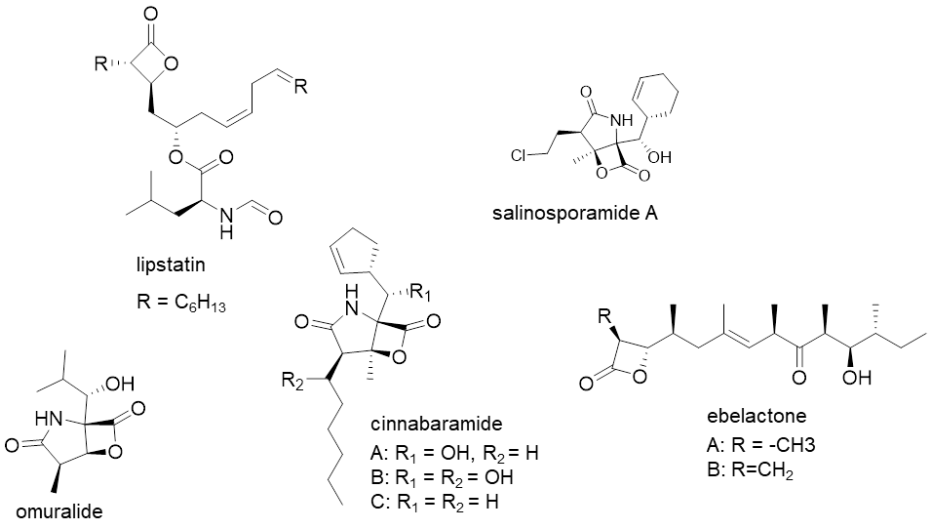


Figure 3: Chemical structure of selected natural products containing a β-lactone moiety: R= residue

The structurally related β-lactone containing compounds salinosporamide A, omuralide and cinnabaramides (Figure 3) are all potent proteasome inhibitors that form covalent bindings to the catalytic Thr1 in the proteasome active subunits. Thereby, the binding mechanism by salinosporamide A to the proteasome is irreversible, whereas reversible but stable for omuralide.[45, 63, 64]

Salinosporamide A (Figure 3) was discovered and isolated from the marine bacteria *Salinispora tropica.*[63] The drug is a structurally and pharmacologically unique β-lactone-γ-lactam irreversible proteasome inhibitor that binds to all catalytic active subunits of the proteasome with IC50 values in the low to mid nanomolar range.[65] Marizomib is a very promising candidate as a potentially first-in-class agent against intrinsic brain tumors, most importantly, because it is able to cross the blood-brain barrier. While preclinical studies of the treatment with marizomib have demonstrated significant anti-glioma activity, its clinical benefit has yet to be proven. Exploiting the biological effects of proteasome inhibitors in combination with other therapeutic strategies may represent a next key step in their clinical development.[66, 67]

1.5 Cystargolides and belactosins

Their versatile potential for biological activities and functions makes β-lactone-containing molecules of particular interest as probes for the identification of novel substances and drug leads.[68] Besides the listed compounds above, many other promising molecules that are in the discovery pipeline could be addressed. However, the main focus of this thesis lays on the proteasome inhibitors belactosins and cystargolides. They are build up out of a linear dipeptide backbone which are *N*-acylated harboring a unique *trans*-disubstituted β-lactone moiety (Figure 4).[69, 70] Belactosin A and C are produced by *Streptomyces* sp. UCK14 and it could be shown for both belactosin derivatives that they inhibit the 26S proteasome in the sub-micromolar range. The chymotrypsin-like inhibitory activity of belactosins in a rabbit 20S proteasome based *in vitro* assay showed inhibition values of IC_{50} 0.21 µM for belactosin A and C. Furthermore, inhibition of growth in HeLa S3 cells showed readings about 51 µM for belactosin A and 200 µM for belactosin C.[71] Cystargolide A and B were isolated from *Kitasatospora cystareginea* NRRL B16505. For those molecules, inhibition of the human 20S proteasome could be demonstrated in vitro with IC_{50} of 0.93 µM for cystargolide A and 0.35 µM for cystargolide B, respectively.[70, 72]

Consequently, to facilitate the inhibitory nature of those two compounds the β-lactone moiety allows stable covalent and irreversible binding of cystargolides and belactosins to the proteolytic active chymotrypsin-like β_5-subunit of the proteasome. The binding mode has been elucidated in great detail for homo belactosin C and in direct comparison with the binding mode of omuralide, it could be shown that it takes place by acylation of the free hydroxyl group of the catalytic *N*-terminal nucleophilic Thr1 residue in the active side of the proteasomal β_5-subunit (Figure 4).[73]

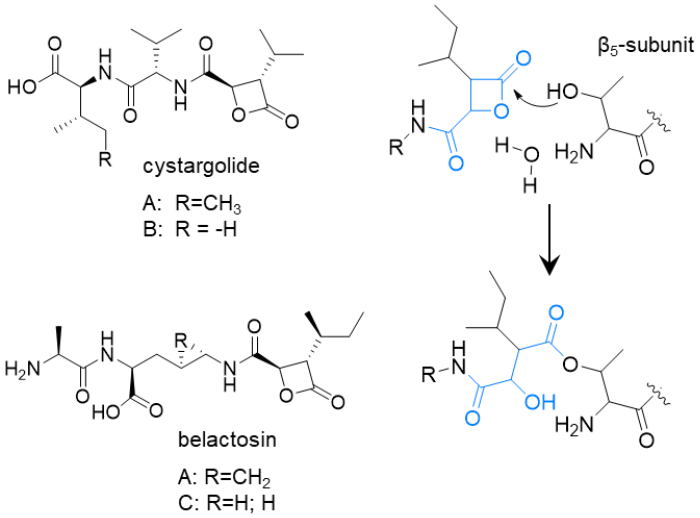


Figure 4: Chemical structures of cystargolides and belactosins and the binding mechanism of the β-lactone warhead to the β5-subunit of the proteasome: R = residue

As shown in the last chapters, the chemical nature of β-lactone containing natural products is very diverse and has high potential for clinical application. However, the production of such compounds and derivatized drug leads has been largely restricted to chemical synthesis. Partly this is associated with the fact, that there are still significant gaps in biochemical knowledge about β-lactone biosynthesis mechanisms. The future development of anticancer drugs and other biological active agents requires innovative biochemical and biological driven approaches like engineered genome guided biosynthesis studies. To this date the biosynthesis of several natural products, containing ring-strained β-lactone moieties, has been extensively studied, e.g. in the structurally related omuralide and previously described cinnabaramides or ebelactone and salinosporamides (all structures shown in Figure 3).[61, 64, 74-76]

However, the biosynthesis of belactosins and cystargolides, including the formation of the β-lactone warhead, is understudied and biosynthesis mechanisms have only been elucidated for single parts of the molecules. This encouraged us to explore those electrophilic proteasome inhibitors in more detail.[77]

1.6 Biosynthesis of cystargolides and belactosins

Although the chemical synthesis of the β-lactone containing natural products cystargolides and belactosins had been studied in detail [72, 78-83], their biosynthesis pathways remain speculative and understudied. Within our research group, we could recently report on the identification of the biosynthetic gene clusters for cystargolides and belactosins in their respective producer strains, *Kitasatospora cystarginea* NRRL B-16505 and *Streptomyces* sp. UCK14, with some astonishing findings.[84]

As a fundamental basis of this thesis, Felix Wolf and colleagues were able to analyze and characterize both belactosins and cystarogolides BGCs in great detail. The first insight and further first draft of a possible biosynthesis routes are elaborated in the next chapter. A detailed description of the complete composition and annotation of genes in the belacatosin and cystargolide biosynthetic gene cluster is shown in the preliminary results section in chapter 1.8.

Identification of both gene clusters was not as straight forward as anticipated for similar natural products containing a peptidic backbone. Sequencing the genome of the cystargolide producer strain and applying a general genome guided approach using the antiSMASH[85] platform to evaluate possible biosynthesis gene clusters, did not lead to any conclusive hits, as none of the identified clusters could be correlated with the size or specificity to synthesize cystargolide. As an alternative approach feeding studies with stable isotope precursor molecules were conducted to possibly get first insight into a biosynthesis route towards cystargolides and the responsible biosynthetic genes. In a first step, ^{13}C5-valine was fed to *K. cystarginea* NRRLB16505 cultures and LC-MS analysis revealed the incorporation of ten and fifteen carbon atoms in the cystargolides A and C molecules, respectively (Figure 5A). Subsequently, $^{13}C_2$-labelled acetate was fed to the cultures and showed incorporation of two carbon atoms in both cystargolide A (Figure 5A) and C. Further MS/MS experiments showed that five carbon atoms from valine and two carbon atoms from acetate are incorporated into the β-lactone moiety of both cystargolide A and C. In order to specifically predict the exact position of the respective labeled carbon atoms NMR studies were conducted with isolated extracts from 1,2-$^{13}C_2$- and 1-^{13}C- acetate marked cystargolides from *K. cystarginea* NRRLB16505. NMR analysis showed that the β-lactone moiety was marked at the C1'' for 1-^{13}C- acetate and C-2'' for 1,2-$^{13}C_2$-acetate (Figure 5A). Those results suggested that the β-lactone moiety in cystargolides may be constructed via condensation of a valine derived

metabolite with acetate. To further proof this theory similar feeding studies were conducted in the belactosin producer strain *Streptomyces* sp. UCK14 (*S.* sp. UCK14), as an analogous process could take place in the synthesis of this molecule, as well. However, instead of labeled valine, 1-$^{13}C_1$-isoleucin was fed to the cultures, taking account of the *sec* butyl group of the β-lactone moiety in the belactosin molecule. Additionally, 1,2-$^{13}C_2$-acetate was fed to the cultures. Indeed, LC-MS and MS/MS analysis showed incorporation of two carbon atoms of 1,2-$^{13}C_2$-acetate and one carbon atom of 1-$^{13}C_1$-isoleucin into the β-lactone moiety of belactosin (Figure 5A).[77] Those findings strongly suggest that the resulting product is built up via a condensation reaction between a valine metabolite and acetate.

This gave us a first hint, that the carbon skeleton of the decorated β-lactone moiety may be constructed by a pathway reminiscent of leucine biosynthesis. In leucine biosynthesis, the enzyme isopropylmalate synthase (IPMS) catalyzes a Claisen-type condensation of a-ketoisovalerate (α-KIV) and acetyl-CoA to generate 2-isopropylmalate (2-IPM). Very interestingly, a-KIV is a degradation product of valine. Subsequently, 2-IPM gets isomerized to 3-IPM by an isopropylmalat-dehydratase. In leucine biosynthesis 3-IPM gets oxidized, decarboxylated and amidated to generate leucine.[86] However, in cystargolide and belactosin biosynthesis such an IPMS enzyme could be involved in the generation of an intermediate towards β-lactone formation. When screening the genome of both cystargolides and belactoins producer strains we could successfully identify two distinct homologs of the IPMS in both cystargolide and belactosin biosynthetic gene clusters, CysA and BelJ, respectively. Indeed, we were able to verify the functionality of the BelJ IPMS enzyme in an *in vitro* assay. BelJ was able to generate the desired specific product 2-(sec-butyl)-2-hydroxysuccinic acid from the alternative substrates acetyl-CoA and 2-keto-3 methylvalerate (2-KMV), an intermediate from isoleucine metabolism, in an aldol reaction.[77] IPMS are essential enzymes in primary metabolism and in general show strong substrate specificity towards α-KIV.[86] However, it could be shown that in certain cases, like for example in glucosinolate biosynthesis, alternative substrates are also accepted by the enzyme.[87] However, a similar functionalization of an bacterial IPMS from primary to secondary metabolism has not been reported yet.[88]

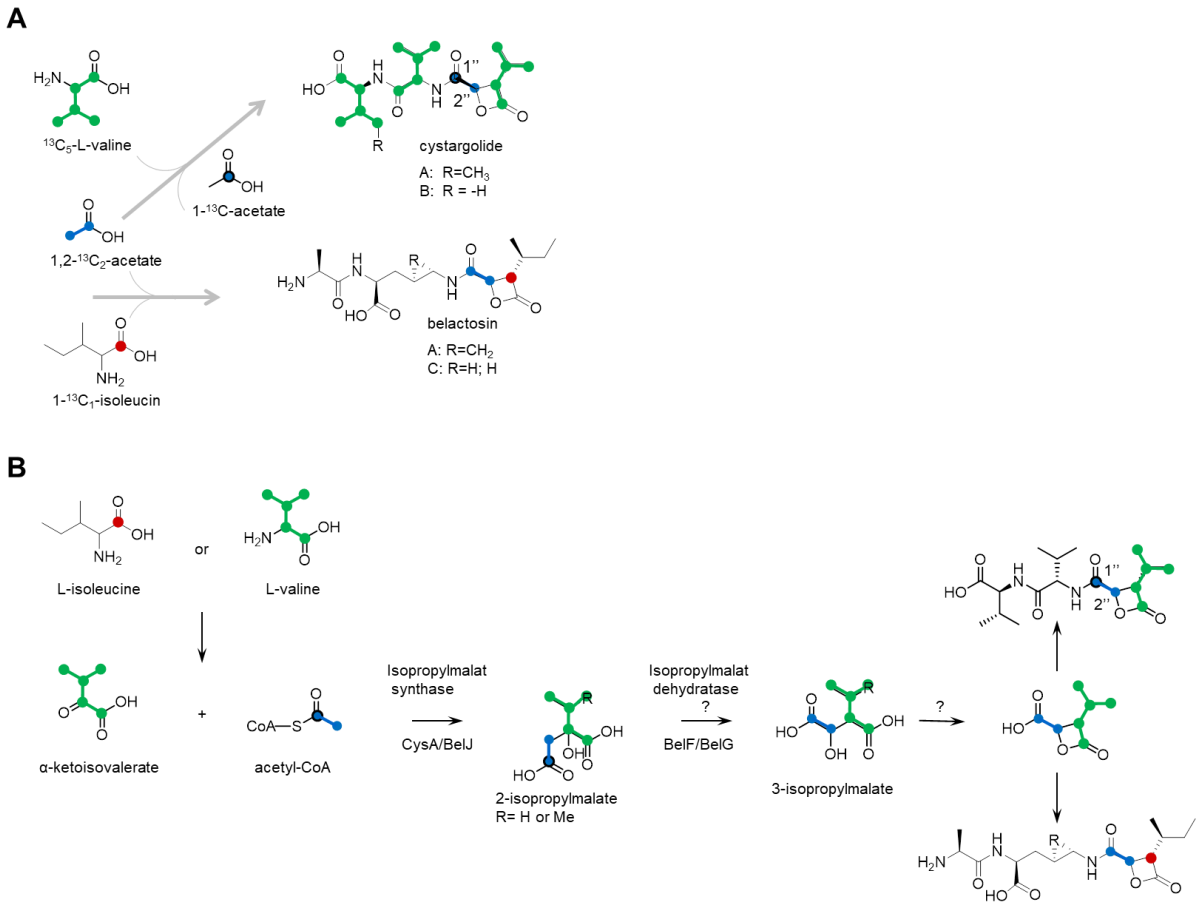


Figure 5: Proposed biosynthesis steps towards the β-lactone ring formation in cystargolide A and belactosin C: A: Feeding studies with stable isotope precursor molecules, $^{13}C_5$-valine and, $^{13}C_2$, which showed incorporation of ten and two carbon atoms, respectively. Analysis of 1,2-$^{13}C_2$- and 1-^{13}C- acetate fed cultures extracts showed that the β-lactone moiety was marked at the C1'' for 1-^{13}C- acetate and C2'' for 1,2-$^{13}C_2$-acetate. The belactosin producer strain was fed with 1-$^{13}C_1$-isoleucin and 1,2-$^{13}C_2$-acetate, showing incorporation of one and two carbon atoms, respectively. **B:** Starting from the degradation product of valine, α-KIV together with acetyl-CoA, CysA would generate 2-IPM. For belactosin biosynthesis the IPMS homologue BelJ would use the isoleucine metabolism intermediate 2-keto-3- methylvalerate to generate 2-(*sec*-butyl)-2-hydroxysuccinic acid (not shown in this figure). Following, 2-IPM would then be isomerized to 3-IPM, most likely by primary metabolic enzymes, in belactosin formation we suggest that this reaction could be provided by BelF and/or BelG. Parts of this figure are taken and adapted from Wolf et al.[77]

Within the same operon as the IPMS homologs, we could co-localize genes for putative AMP-forming synthetases and ATP-grasp enzymes. It has been shown that discrete amide ligases are able to catalyze the formation of peptide bonds and could therefore possibly generate the peptidic backbone of cystargolides and belactosins, in accordance with an NRPS independent biosynthesis mechanism for the generation of the backbone of these molecules. Discrete amide ligases, mostly found in peptidic or pseudo peptide natural products, are adenylate-forming amide synthetases and carboxylate-amine ligases of the ATP-grasp

enzyme family. They show homology to free standing A domains or aminoacyl-tRNA synthetase and thereby catalyze the ATP-dependent activation of carboxylic acids. Following, they generate a acylO-AMP, which then serves as the activated precursor for a thiolationin-dependent amide bond formation as the acyl group is directly transferred to an amine substrate.[89] The involvement of ATP-dependent carboxylate-amine ligases in natural product biosynthesis has been increasingly recognized including e.g. in the pathways for the generation of ketomemicins, microviridins, dapdiamide, resorcinomycin or pheganomycin.[89-94] These findings further supported our theory that the identified clusters may be responsible for the synthesis of the small peptidic molecules belactosins and cystargolides.

Following, based on analysis with a combination of *in silico* and biochemical data, we preliminary proposed two possible analogous biosynthesis routes towards the β-lactone ring containing cystarogolides and belactosins (Figure 5B). In a first step, the IPMS homologue CysA would generate 2-IPM from α-KIV and acetyl-CoA, as described above. For belactosin biosynthesis the IPMS homologue BelJ would use the isoleucine metabolism intermediate 2-keto-3- methylvalerate, instead of *α*- ketoisovalerate, to generate 2-(*sec*-butyl)-2-hydroxysuccinic acid. In the cystargolide pathway 2-IPM would then be isomerized to 3-IPM, presumably by primary metabolic enzymes, whereas in belactosin formation, this function could maybe be provided by the isopropylmalate dehydratase homologues BelF and/or BelG.

Further biosynthesis steps from the generated intermediates to be able to afford the formation of the β-lactone ring were highly speculative at the start of this thesis. The newest findings will be further discussed in 5.6.

1.7 Cyclopropyl moieties in natural product biosynthesis

Cyclopropane-containing molecules are widely distributed in biological systems.[95] The cyclopropane ring is an important and common structural element in synthetic therapeutics or in natural products such as terpenoids but can also be found e.g. in polyketides, non-ribosomal peptides and fatty acid-derived natural products (Figure 6). They are quite common in plant derived natural products but can also have bacterial origin. The record for the number of cyclopropane rings in one molecule is held by the fatty acid U-106305, isolated from *Streptomyces* sp. UC-11136 and was shown to be a new cholesteryl ester transfer protein (CETP) inhibitor.[96] Cyclopropyl moieties contribute to configuration, conformational stability, and reactivity of a bioactive compound. The cyclopropane may undergo ring opening by reacting with nucleophilic agents contributing to the biological and pharmacological activities.[97-101] Because of their inherent ring strain and stereo defined functionalization in complex molecules chemical synthesis is challenging and has led to the development of a multitude of strategies to access this chemistry.[102] Until today, many currently available methods for asymmetric cyclopropanation give low yields, show poor diastereo- and enantioselectivity or require harsh reaction conditions.[103]

In natural product synthesis, the formation of cyclopropane rings is generally enabled either via carbocationic or carbanionic intermediates (or transition states) and often involves sophisticated biochemistry.[101, 104, 105] This includes, for example, PLP-mediated S_N2-type α,γ-elimination to afford 1-aminocyclopropane-1-carboxylate (ACC) (Figure 6) from S-adenosylmethionine (SAM).[106] Another example are the antifungal and antimicrobial kutznerides 1-9 (kutzneride 1, Figure 6) cyclic hexadepsipeptides isolated from the soil actinomycete *Kutzneria* sp. 744, which, like hormaomycins (Figure 6), are synthesized non-ribosomally. All kutznerides harbor the cyclopropyl containing functionalized amino acid 2-(1-Methylcyclopropyl)glycine (MecPGly)).The identification of the biosynthetic gene cluster of kutznerides[97], by Walsh et al. led to the postulation that genes *ktzA–D* might be involved in construction of the remarkable MecPGly moiety. KtzD, which shows high homology to mononuclear nonheme iron halogenase proteins[107, 108], might generate the chlorinated δ-chloro-(allo)Ile-S-KtzC intermediate , that is bound to KtzC, which is then subsequently converted to the MecPGly moiety.[109, 110]

Similar intramolecular nucleophilic displacement mechanisms seem also to be employed in the biosynthesis of coronatine and curacin A (Figure 6) using chlorinated carrier protein-

linked substrates and the halogen as a leaving-group. [107, 110-112] In the biosynthesis of the pseudomonal phytotoxin coronatine (Figure 6), the aminocarboxycyclopropane-forming enzyme CmaC-catalyzes the formation of coronamic acid by a thioester-stabilized C_α-carbanion.[112] However, in kutznerides biosynthesis, the formation of the cyclopropane ring requires intramolecular replacement of the δ-chloride leaving group by C_β, which could be fulfilled by KtzA, an acyl-CoA dehydrogenase-like protein. A similar dehydrogenation of T domain-bound thioesters is catalyzed by flavoprotein dehydrogenases in the biosynthesis of pyrrole moieties in pyoluteorin and undecylprodigiosin (22, 23).[113, 114] Analogous dehydrogenation of δ -chloro-(allo-)Ile-S-KtzC would yield an enamine intermediate, where C_β could serve as an intramolecular nucleophile in cyclopropane formation. In a final step an imine reduction would produce MecPGly.[110]

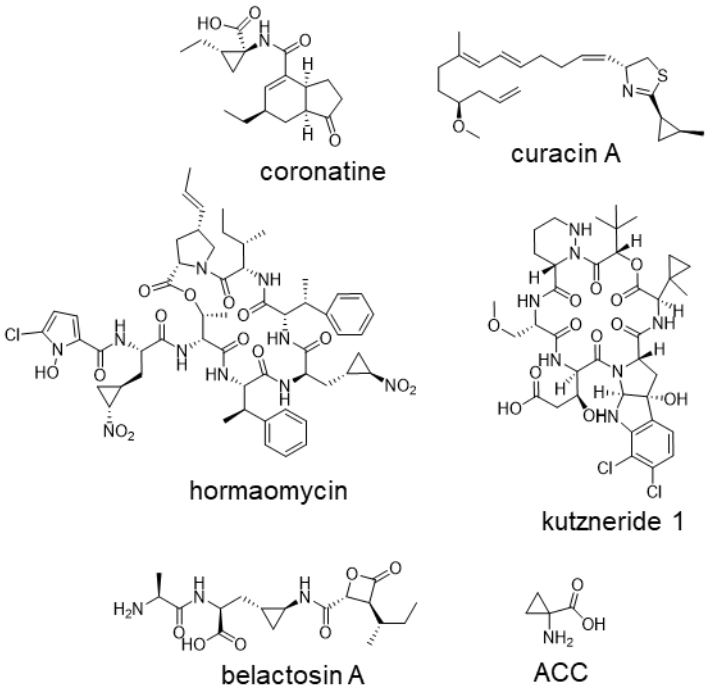


Figure 6: Selected natural products containing cyclopropyl moieties:ACC = aminocyclopropyl carboxylic acid.

As described in literature and in the chapter before, till date the elucidation of the biosynthesis of several cyclopropyl moieties has been anticipated for several compounds. However, the biosynthesis of belactosins (Figure 6), the main compounds of this thesis, which shows interesting chemical features, including a cyclopropyl containing unique amino acid building block, was understudied at the start of this thesis. Belactosin A contains a unique 3-(*trans*-2'-aminocyclopropyl)alanine (Acpa) moiety. It represents a novel non-proteinogenic amino acid with potential applications as a building block in NRPS engineering approaches. Very

interstingly, among the cyclopropyl containing natural products, the peptide lactone hormaomycins, whose structure has been confirmed by total synthesis[115] contains several unprecedented moieties. Those moieties include two of the previously unknown 3-(*trans*-2-nitrocyclopropyl)alanine (Ncpa) building blocks as a key substituent. Hormaomycins have been isolated from *Streptomyces griseoflavus* strain W 384, which were identified as unique among signal metabolites among streptomycetes. They show selective antibacterial activity, and additionally some antimalarial activity.[116,117, 118] The unique structural features of belactosins and hormaomycins have attracted much attention regarding their biosynthetic mechanisms, however only limited information on the biosynthesis of those moieties existed.[119] Taking into account the unique chemical features of these moieties, we speculated that so far undescribed biochemistry might be involved in the generation of the aminocyclopropyl group. This prompted us to investigate the formation of Acpa in the biosynthesis of belactosin A in more detail.

1.8 Preliminary Work regarding the biosynthetic gene cluster of belactosins

1.8.1 Identification and sequence analysis of the biosynthetic gene cluster of belactosins

During his PhD thesis Felix Wolf conducted fundamental experiments towards the identification of the biosynthetic gene cluster of belactosins, which I carried on working during the duration of this thesis, and will be briefly described in the next sections.

In order to get more insights into the biosynthesis of natural products, it is inevitable to have the sequence information of the responsible biosynthetic genes. To identify a possible BGC, which is able to synthesize the belacostin derivatives A and C and the responsible genes involved in the biosynthesis of the Acpa moiety, we conducted a genome search. Because of their peptidic nature, we assumed that both belactosin derivatives could be synthesized via a non-ribosomal peptide synthetase mediated biosynthesis machinery. Surprisingly, analysis of the whole genome sequence of the belactosin WT producer strain *Streptomyces* sp. UKC14 with web based bioinformatic applications, like antiSMASH database[85], BLAST homology searches[120, 121], and MultiGeneBlast[122] did not reveal a putative BGC that correlated to the predicted size and specificity of a NRPS assembly line for the production of belactosins.

The prediction of the BGC was not as straight forward as we anticipated, but by feeding and knock-out studies in the cystargolide producer strain *Kitasatospora cystarginea* NRRL-B16505, Felix Wolf could show that the 2-Isopropylmalat synthase (IPMS) CysA is involved in the β -lactone warhead assembly (see chapter 1.6). We assigned genes *cysA* to *cysH* to the biosynthetic gene cluster of cystargolides, coding for an MFS multidrug transporter (cysB), two AMP-dependent synthetases (cysC and cysF), an ATP-grasp enzyme (cysD), amethylesterase (cysE), a SAM-dependent methyltransferase (cysG), and a LysR-like transcriptional regulator (cysH) (Figure 7).

Furthermore, when applying *cysA* as a genome-mining probe to the genome of *Streptomyces* sp. UCK14, we could identify a second putative IPMS-encoding gene *belJ,* that is flanked by genes encoding potential export, amide bond formation, methylation, and regulatory functions.[77] Together with the IPMS, the enzymes possibly required for amide bond formation and the amide ligase are likely responsible for the formation of the β-lactone

and the assembly of the peptide backbone. In total, based on protein similarity and operon organization, we assigned 22 genes to the putative belactosin BGC (Figure 7, Table 1).

When analyzing the sequence of the cystargolides and belactosins gene clusters in more detail, we could identify several homologous gene/protein pairs, most likely catalyzing similar biosynthesis steps (Figure 7). For instance, BelH and CysF are both putative AMP-dependent synthases, showing 54 % identity (ID) on amino acid sequence level. There are also two putative SAM-dependent methyltransferases, CysG and BelI, with 65% sequence ID in the clusters. Additionally, *cysE* and *belR* both encode for homologous putative methylesterases of the α/β -hydrolase superfamily (52% sequence ID). Moreover, two genes, *belF* and *belG*, possibly encoding for the subunits of an isopropylmalate dehydratase homologue, were also found in the operon. Furthermore, genes *belO-belQ* encode for potential transcriptional regulators.

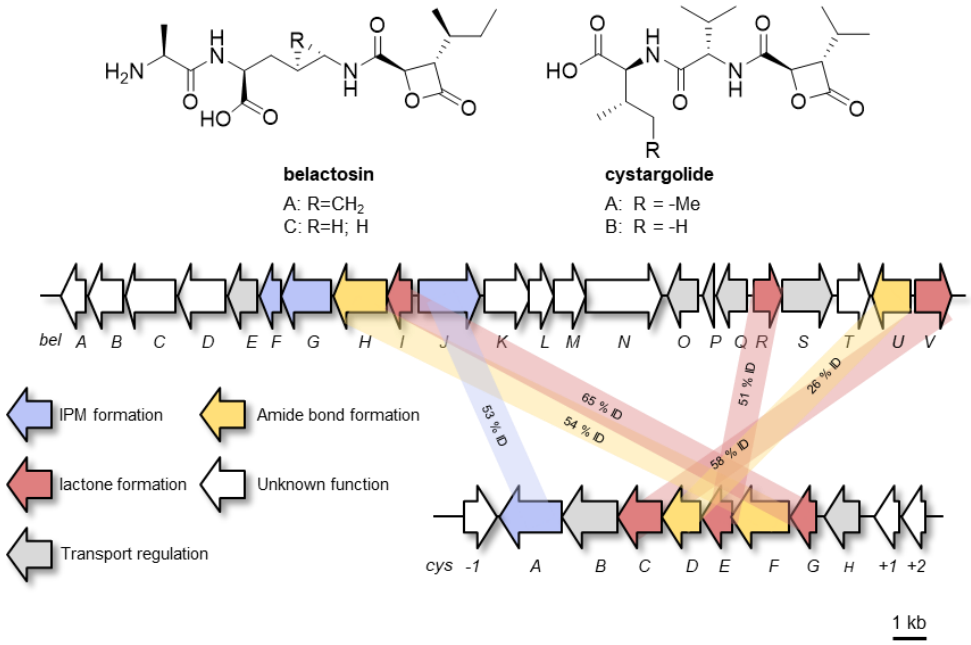


Figure 7: Biosynthetic gene clusters of belactosins and cystargolides: Gene cluster organization of belactosins (*belA* to *belV*) and cystargolides (*cysA* to *cysH*). Genes for IPM, lactone and amid formation, as well as for transport and other so far unknown functions are colored and depicted by the arrows. Homologous genes between the *cys* and *bel* gene cluster (with ID in %) are indicated. ID = identity, R = residue

Table 1: Biosynthetic gene cluster annotation table for *bel* gene cluster: Homologous genes from *cys* cluster are indicated in grey. The sequences of the *bel c*luster are available at NCBI under the accession number KY249118.

Gene	Size (AA)	Protein homolog (Genb. Ac. No.)	ID/sim[a]	Predicted function
orf-1	116	*S. misionensis* DSM 40306 (WP_146466243.1)	98/97	glyoxalase
belA	195	*S. misionensis* DSM 40306 (WP_146466242.1)	97/97	acetyltransferase
belB	308	*S. misionensis* DSM 40306 (WP_146466241.1)	99/99	peptidase U61
belC	650	*S. misionensis* DSM 40306 (WP_146466240.1)	97/98	acyl-peptide hydrolase
belD	435	*S. misionensis* DSM 40306 (WP_146466239.1)	98/99	acetyl-ornithine deacetylase
belE	278	*S. misionensis* DSM 40306 (WP_146466238.1)	99/99	ABC transporter
belF	199	*St. anmyonensis NBRC 103185* (WP_042425392.1)	87/92	3-isopropylmalate dehydratase SS
belG	464	*St. anmyonensis NBRC 103185* (WP_084700861.1)	90/95	3-isopropylmalate dehydratase LS
belH	492	*K.cystarginea* (ARO49564.1)	54/63	AMP-dependent synthetase
belI	229	*K.cystarginea* (ARO49565.1)	65/76	SAM-dependent methyltransferase
belJ	566	*K.cystarginea* (ARO49559.1)	53/64	2-isopropylmalate synthase
belK	359	*St. anmyonensis NBRC 103185* WP_052441272.1	85/92	Oxygenase
belL	230	*St. anmyonensis NBRC 103185* WP_042425400.1	91/96	Oxygenase
belM	291	*St. anmyonensis NBRC 103185* WP_042425403.1	79/84	Epimerase
belN	696	*St. anmyonensis NBRC 103185* WP_042425531.1	92/96	Molybdopterin oxidoreductase
belO	290	*St. anmyonensis NBRC 103185* WP_042425408.1	86/90	LysR family TR
belP	107	*St. anmyonensis NBRC 103185* WP_084700865.1	82/86	LuxR family transcriptional regulator
belQ	292	*St. anmyonensis NBRC 103185* WP_052441275.1	75/81	XRE family transcriptional regulator
belR	270	*K.cystarginea* (ARO49563.1)	51/59	Methylesterase
belS	469	*K.cystarginea* (ARO49560.1)	58/73	MFS transporter
belT	351	*S. misionensis* DSM 40306 WP_146466268.1	97/98	1,4-beta-xylanase
belU	421	*K. cystarginea* (ARO49562.1)	26/43	ATP-dependent amide ligase ATP-grasp conserved domain
belV	390	*K. cystarginea* (ARO49561.1)	58/71	AMP-dependent synthetase
orf+1	240	S. misionensis DSM 40306 WP_146466223.1	99/99	SAM-dependent methyltransferase

[a] amino acid sequence homology [%] from blastp analysis

AA protein length in amino acid

ID amino acid identity

Sim amino acid similarity

SS small subunit

LS large subunit

S. *Streptomyces*

St. *Streptacidiphilus*

K. *Kitasatospora*

1.8.2 Construction and screening of a genomic library to isolate the belactosin biosynthetic gene cluster

To further analyze the BGC, elucidate the biosynthesis mechanism of belactosin A and C as well as making the BGC accessible for molecular manipulation and genetic engineering, we aimed to integrate and isolate the complete *bel* cluster on a fosmid vector. Therefore, high quality genomic DNA from *Streptomyces* sp. UCK 14 was extracted and sheared into approximately 40 kb fragments in order to be ligated in to the pCC1FOS shuttle vector, and then transfected into the EPI300™-T1^{R} host cells to generate a fosmid based genomic library. In total, 2016 clones were subjected to several rounds of screening by PCR. To identify and isolate the BGC of belactosin (*bel* cluster), perfect matching primer pairs, up and downstream of the BGC, were applied on the genomic library of *Streptomyces* sp. UCK14. Fosmid clone 6D1 with an insert size of 42826 bp contained the complete *bel* cluster, which was confirmed by restriction analysis and end-sequencing.[77] The sequences of the *bel* cluster are available at NCBI under the accession numbers KY249118. The putative annotation table of the bel cluster is depicted in Table 1.

Introduction II: Proteasome inhibitors from *Nocardia* and their Role in Pathogen Survival

1.9 Natural products from *Nocardia*: Introduction of the genus *Nocardia*

The second part of this thesis is dedicated to the topic of exploration of natural products deriving from *Nocardia* strains and studying their potential role and impact in pathogen survival. Overall, *Nocardia*, which are described as opportunistic human pathogens, are an under-investigated actinobacterial genus in regard of secondary metabolites discovery. In order to get access to this quite under-discovered source for novel compounds, orphan biosynthetic gene clusters can be recombinantly expressed in heterologous expression systems. Such strategies can help to awaken such silent biosynthetic gene clusters and assign them to the production of potentially novel natural product compounds with interesting bioactivity.

Nocardia spp. are gram-positive, aerobic, filamentous Actinobacteria and belong to the order Corynebacteriales. They are closely related to the genera *Corynebacterium,* and *Mycobacterium* but particularly show similarities to the genus *Rhodococcus*. Most of the time they form mycelia, but can also fragment into rod- to coccoid-shaped bacteroid non-motile elements.[123, 124] Additionally, a characteristic feature of *Nocardia* spp. is, that they are partially acid fast, containing tuberculostearic acids in their cell wall, similar as seen in members of the genus *Mycobacterium*. Nonetheless, the mycolic acid produced by *Nocardia* spp., in contrast to M*ycobacteria*, are short chains (40–60 chain) and usually exhibit branching on Gram staining.[125]

This genus has worldwide distribution and often colonizes organic-rich soil as ubiquitous saprophyte, but can also be found in dust, sand or stagnant water. Nevertheless, they also show pathogenic life style and therefore are most prominently recognized as human opportunistic pathogens. .[124 126] The pathogenicity of *Nocardia* was first described by Edward Nocard in 1888 as causative agents of severe infections of the skin, lung or central nervous system in humans and animals causing nocardiosis.[127] Since then several *Nocardia* strains have been extensively studied for their pathogenicity by Beaman and others. [128, 129] In general, when infected with *Nocardia*, immunocompromised patients are affected by nocardiosis with a high mortality rate. However, cases of severe nocardiosis in immunocompetent patients have also been reported.[130, 131] Until today a variety of different strains of *Nocardia* have been identified, many of which are pathogens and described as infectious to humans and animals.

To date, out of the more than 86 described strains, around 50 *Nocardia* species are considered clinically significant.[132-134] The most prevalent human pathogens are *N. farcinica*, *N. asteroides* type VI (*N. cyriacigeorgica*)[135], *N. brasiliensis*, the *N. nova* complex, *N. abscessus* and the *N. transvalensis* complex.[125] Usually, *Nocardia* cause localized cutaneous or pulmonary infections; thereby, most *Nocardia* infections are acquired through inhalation or by traumatic introduction of organisms. From the site of infection bacteria can disseminate throughout the whole body, leading to systemic forms of nocardiosis, which often includes severe brain infections.[131, 136-138]

Nocardia primarily infect the lungs or cutaneous and lymphocutaneous tissues of immunocompromised patients. Such opportunistic behavior is also known from pathogenic non-tuberculous mycobacteria (NTM) including the *Mycobacterium avium* complex (MAC)*, Mycobacterium kansasii, Mycobacterium abscessus* and *Mycobacterium ulcerans.* The common aspects of *Nocardia* and *Mycobacterium* pathogenesis have been attributed to a shared cell wall physiology containing layers of mycolic acid and arabinogalactan as well as more general virulence factors. For example, catalase and superoxide dismutase enzymes, invasion-like proteins, phospholipase C, haemolysin and the cord factor are possible virulence factors of those pathogens.[128, 139-141] However, major parts of *Nocardia* infections remain poorly understood, including strain-specific virulence and progression of nocardiosis. [142] Raising the question of natural products of *Nocardia* might contribute to the virulence of these organisms.

1.10 Virulence associated natural products from *Nocardia*

Pathogenic bacteria in general have already been recognized as a rich source for natural products in drug discovery. Some of those isolated small molecule compounds or metabolites could be successfully linked to the pathogenic behavior of these organisms. However, the genus *Nocardia* still remains largely untapped and understudied in this area of research.[125, 128, 143] In particular because of their clinical importance, many *Nocardia* genomes were sequenced within the last decade, showing an extraordinary potential for the production of diverse secondary metabolites for most *Nocardia* strains.[22] Moreover, some *Nocardia* strains have recently moved into focus as producers of compounds with pharmaceutical value, facilitating immunosuppressant, antimicrobial, cytotoxic or antifungal bioactivity, i.e. *N. brasiliensis*, *N. abscessus*, *N. transvalensis*, *N. terpenica*, and *N. pseudobrasiliensis*.[126, 144-146] It has been speculated that cytotoxic and immunosuppressive properties of *Nocardia*-derived secondary metabolites may play a role in infection, immune evasion and clinical manifestation of the disease. How the encoded chemistry exactly contributes to the general and strain-specific pathogenicity of *Nocardia* spp. is still a focus of ongoing research. However, in the last decade a selection of novel compounds with interesting bioactivities has been identified and a connection of those compounds to bacterial immunoevasive behaviors was enlightened to certain points. The next paragraphs will give an overview of selected interesting compounds including siderophores, cell wall associated compounds, molecules with immunosuppressive activity and further novel *Nocardia* derived natural products.

Siderophores are considered as important virulence factors for diverse pathogens including *Mycobacterium* spp. Pathogenic bacteria thereby secrete iron-chelating siderophores into iron-limiting environments of their human hosts to compete for ferric iron, as iron is an essential element for pathogenic bacteria to survive and grow inside the host organism. Most of the iron is tightly bound to specific proteins such as hemoglobin, ferritin, or transferrin, thus severely restricting iron acquisition by the bacteria. [144, 147, 148]

The most prominent known siderophore of *Nocardia* are nocobactins, which are structurally alike to mycobactins deriving from the closely related human pathogen *Mycobacterium tuberculosis*.[149] Mycobactin-like siderophores have been disclosed from different pathogenic and non-pathogenic bacteria. However, the structures of siderophores yersiniabactin and vibrionbactin, from the bacteria that cause plague and cholera, respectively, are analogous to that of mycobactin.[22, 144, 150, 151]

The architecture of this nocobactin like siderophores commonly frames two units of lysine, a C-terminal cyclic and a central acyclic one, usually decorated with N^{ε}-OH and N^{ε}-acyl functional groups. In addition, the 2-OH-phenyl-(5/6)-methyl-oxazol(in)e moiety defines a further scaffold backbone unit. The variation in the substitution patterns of the oxazol(in)e, the phenyl residues and N^{ε} of both lysine units in tandem with the fatty acid nature affords a broad panel of molecular congeners supporting their cognate producers to cope with their environments.[144, 152] Biosynthetically, a nonlinear assembly mechanism is hired to generate such entities recruiting three NRPS modules and one PKS module. It could be shown that diverse subfamilies of the nocobactin-like siderophores originate from one gene cluster family which is highly conserved in almost all *Nocardia* strains.[22]

Nocobactin-like siderophores are assumed to be correlated with pathogenicity during the progression of nocardiosis. [152] Recently, nocobactin NA (Figure 8A) was identified in a screening campaign, searching for small-molecule immune suppressors secreted by *Nocardia* species. It could be shown that this compound inhibits Notch signaling in a dose-dependent manner. Notch is an important regulator of immune cell development and function, and inference by the pathogen may serve as an immune evasion mechanism.[153]

Nocardia spp. encode an outstanding number of comparably simple pathways driven by ketosynthase biochemistry. An example is the gene cluster for the biosynthesis of nocardiolactones (Figure 8B) which has been found in 23 *Nocardia* isolates. Besides, the associated genome mining approach revealed that the nocardiolactone gene cluster is enriched in human pathogenic strains. However, the responsible biosynthetic gene cluster was not found in any closely related genera such as *Mycobacterium*, *Rhodococcus*, or *Streptomyces*, suggesting that nocardiolactone biosynthesis may be specific to the genus *Nocardia*.[154, 155] In general such ketosynthase biochemistry related pathways which are also present in *M. tuberculosis* were determined to participate in the synthesis of diverse surface-exposed glycolipids including the assembly of mycolic acids, mycocerosic acids and pthiocerols.[156] The most abundant of such glycolipids are trehalose-6,6'-dimyco-late (TDM) and the pthiocerol containing phenolic glycolipid (PGL) (Figure 8B). Those compounds are known to affect the host immune system, by stimulation of cellular and humoral immunity, granuloma formation and anti-tumor activity.[157] TDM, also named the cord factor, has been linked to intracellular survival of virulent *Nocardia* strains. They might play a role in favored uptake into macrophages by activating the alternative complement pathway without triggering respiratory burst.[158, 159]

Immune evasion of human host cell defense mechanisms is a well-known survival strategy by pathogenic bacteria. It could be shown, that *Noardia brasiliensis* is able to induce an immunosuppressive microenvironment by triggering an upsurge in specialized regulatory T-cells (Treg). This leads to a downregulating of other T-cell subpopulations to assure survival in the host.[160]A prominent set of molecules that were discovered and isolated from *N. terpenica* IFM 0406 (*Nocardia brasiliensis* IFM-0406) are brasilicardins (A-D) from which inherit remarkable immunosuppressive entities. Those compounds feature a diterpene skeleton decorated with *N*-acetylglucosamine (GlcNAc), L-rhamnose, pyruvate and 3-hydroxybenzoate (3-HBA) units.[146] Outstandingly, identification of brasilicardin A (Figure 8C), revealed an antitumor activity with a striking differential cytotoxicity (IC50 = 0.22-100 µg/mL) against a broad panel of various tumor cell lines.[161, 162] Even more, using a mixed lymphocyte reaction (MLR) based assay as a model for allograft rejection, brasilicardin A exhibited comparable immunosuppressive readings to cyclosporine A, a cyclic peptide and potent immunosuppressive agent mainly used in transplantation medicine to prevent graft rejection[163] Alongside with brasilicardins, the retrieval of brasilinolides (A-C) (Figure 8C) as bioactive macrolides from *N. terpenica* IFM 0406 provided further proof of the fascinating molecular chemotypes that *Nocardia* can assemble. Structurally, they frame a glycosylated 32-membered polyhydroxy macrolactone with a tetrahydropyran hemiketal ring entrenched within the macrocycle besides 2-deoxy-L-fucose (2dF) as a deoxy sugar decoration. Biologically, this molecular family was found to represent immunosuppressive, antitumor, and antifungal bioactivities.[164]

Driven by the pathogenicity of several members of the genus *Nocardia*, co-culture experiments of *N. tenerifensis* IFM 10554^T with the mouse macrophage-like cell line J774.1 were conducted. Interestingly, nocarjamide (Figure 8D), a cyclic nonapeptide with Wnt signal-alleviating activities, was exclusively expressed under this exceptional conditions.[165] Such Wnt signal-modulating secondary metabolites might contribute to untangling *Nocardia*'s pathogenicity counting on the involvement of these signaling cascades in numerous crucial processes such as the proliferation of eukaryotic cells and various disease-underlying mechanisms.[165]

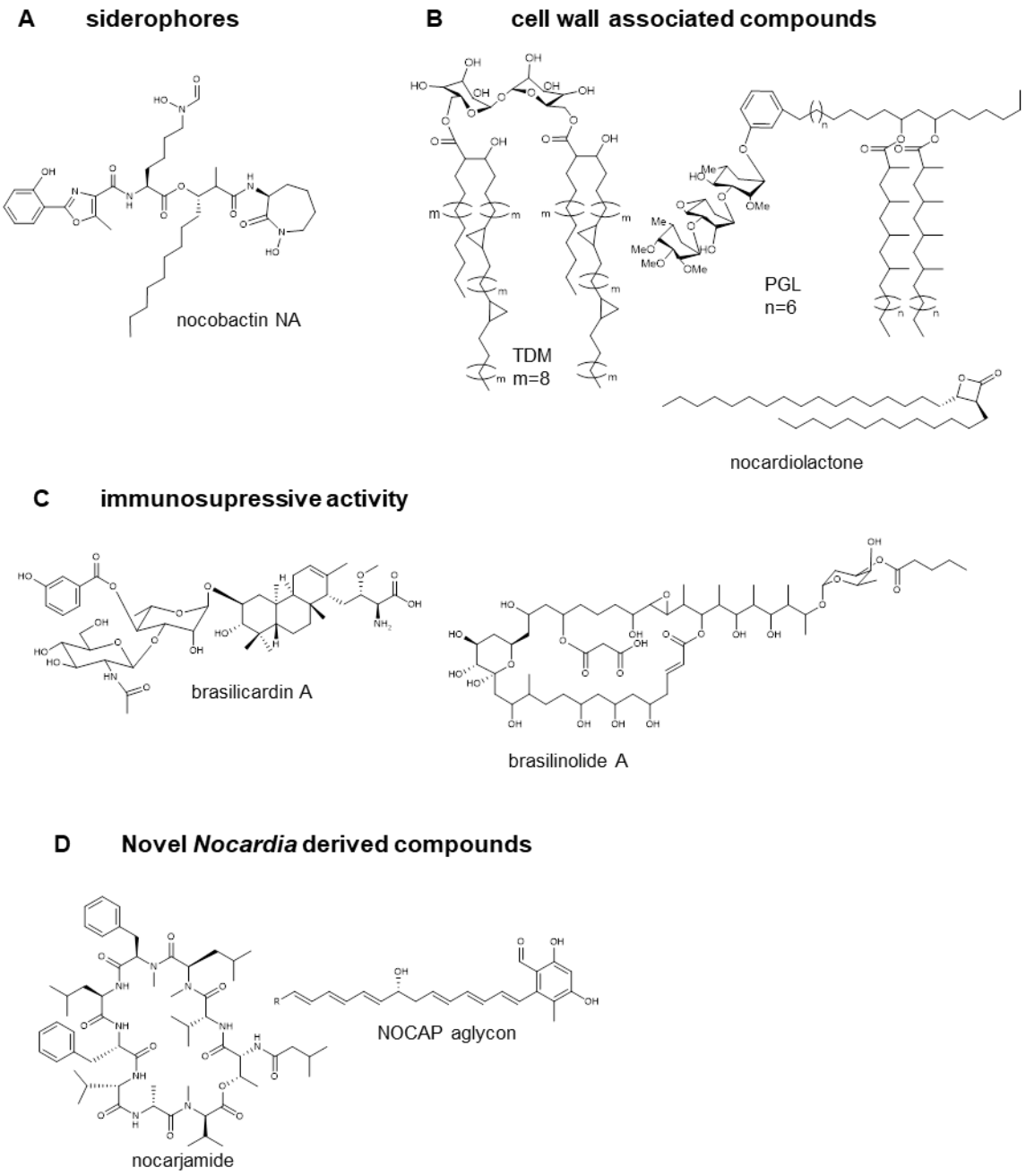


Figure 8: Chemical structures of virulence associated natural products produced by *Nocardia* spp.: Siderophore Nocobactin NA **(A)**; cell wall associated compounds including surface-exposed glycolipids **(B);** compounds with immunosuppressant or cytotoxic activities **(C);** recently identified novel compounds **(D).** NOCAP = nocardiosis-associated polyketide synthase; TDM = Trehalose-6,6'-dimyco-late and PGL= (phenolic glycolipid).

Another novel approach to unlock the hidden potential of *Nocardia* chemistry and revealing virulence associated compounds was followed by complete reconstitution and deorphanization of a 3 MDa trans-AT PKS system that is predominantly found in clinical *Nocardia* isolates. This PKS was hence termed nocardiosis-associated polyketide synthase (NOCAP).[166] This molecule likely represents the aglycon of a yet to be discovered natural

product which would be decorated with a 6-deoxy sugar (Figure 8D). Further investigation of its biological role is underway.[167] The relatively large number of biosynthetic gene clusters identified in *Nocardia* genomes implies that this genus represents a rich source of novel secondary metabolites which are yet to be explored. Furthermore, large-scale genome driven analysis of biosynthetic gene clusters allows the distinction and prediction of natural product structural variations, as well as possible crosslinking to virulence association of those molecules within human or animal pathogenic strains. Serval genome driven analyses and gene cluster characterization studies are underway or already finalized within the Kaysser research group.[22, 142]

Previous to this thesis, starting from a genome mining analysis of strains associated with the production of epoxyketone containing proteasome inhibitors, which was executed by Leonard Kaysser and Michele Schorn, revealed a very intriguing biosynthetic gene cluster from the human pathogenic *Nocardia* strain *Nocardia cyriacigeorica* GUH-2. This particular cluster sparked our interest, because when analyzing the identified gene cluster in more detail, it revealed that this cluster encodes for a hybrid PKS-NPRS system and a highly conserved acyl-CoA dehydrogenase (ACAD) gene, with homology to the epoxyketone synthase from the proteasome inhibitor-associated biosynthetic pathways of epoxomicin. In epoxomicin biosynthesis the ACAD gene is responsible for building the pharmacological active epoxyketone warhead. Further genome mining and detailed genome analysis showed that this cluster is also present in other human pathogenic *Nocardia* strains.[168-170] Acyl-CoA dehydrogenase enzymes are known to generate electrophilic groups for the production of potent small-molecule protease and proteasome inhibitors.[46] We thus speculated that the ACAD, might modify a CoA- or ACP-activated fatty acyl product of the encoded PKS, accordingly. Additionally, and to our interest, *N. cyriacigeorgica* GUH-2 has been shown to inhibit proteasome activity and induce apoptosis in cells by a yet unidentified secreted small-molecule.[171, 172] In this thesis I aimed to identify the compound produced by the ACAD associated gene cluster and analyze if the reported proteasome inhibitory activity correlates with the compound synthesized by the detected orphan gene cluster.

1.11 Aims of this study

Biosynthesis of the cyclopropylalanine moiety of belactosin A

At the beginning of this thesis little was known about the biosynthesis of belactosins and in particular how the cyclopropylalanine moiety of belactosin A is constructed. In the course of this work, we wanted to address and enlighten the following research points:

- Analysis and verification of the biosynthetic gene cluster of belactosins, including determination of the gene cluster borders
- Heterologous expression of the belactosin biosynthetic gene cluster
- Detailed investigation of the biosynthesis of the Acpa moiety of belactosin A, using gene deletion studies, isotope labeled precursor feeding and chemical complementation

The role of proteasome inhibitors from *Nocardia* in pathogen survival

Until today, virulence mechanisms of human pathogenic *Nocardia* and progression of nocardiosis is poorly understood. Furthermore, *Nocardia* show huge potential for the production of diverse secondary metabolites. In the scope of this thesis, we wanted to explore the potential of the human pathogenic strain *Nocardia cyriacigeorgica* GUH-2 for the production of diverse natural products and shed light on the putative associated virulence of these compounds. In this context we aimed the following research goals:

- Genome wide analysis of the secondary metabolite potential of *N. cyriacigeorgica* GUH-2 and comparison with other genome mining approaches within our working group
- Generation and screening of a fosmid based genomic library of the complete *N. cyriacigeorgica* GUH-2 genome
- Heterologous expression of the identified orphan putative acyl-CoA dehydrogenase (ACAD) containing cluster (ACAD cluster) from *N. cyriacigeorgica* GUH-2
- Purification and characterization of the putative novel compound of the ACAD cluster

2 Material and Methods

Materials

2.1 Devices and instruments

Table 2: Devices used in this study

Device	Specification	Manufacturer
Autoclaves	Systec VX-100	Systec GmbH, Linden, Germany
	Systec DE-23	
Centrifuges	Sorvall RC6 Plus (rotor SLA-1500)	Thermo Fisher Scientific, Waltham, Massachusetts, USA
	VWR MicroStar AR Ch500057 PP75003424	Thermo Fisher Scientific, Waltham, Massachusetts, USA
	Centrifuge 5415R (rotor F 45-24-11)	Eppendorf, Hamburg, Germany
	Centrifuge 5810R (rotor A-4-62)	Eppendorf, Hamburg, Ger many
	Multifuge 1 S-R 75002005 G Sorvall Heraeus	Thermo Fisher Scientific, Waltham, Massachusetts, USA
Cleanbench	Heraeus	Thermo Fisher Scientific, Waltham, Massachusetts, USA
	Die zwei anderen	
	S2	
Electroporator	MicroPulserTM Electroporator	Bio-Rad Laboratories, Inc., Hercules California, USA-
FPLC	Äkta start	Cytivia, Marlborough, Massachusetts, USAn
Freeze Dryer	Christ	Martin Christ Gefriertrock nungsanlagen GmbH, Osterode am Harz, Germany
French Press	SLM Aminco French Press	
HPLC systems (analytical and semi preparative)	Agilent 110 Series	Agilent Technologies, Santa Clara, California, USA
	Agilent 1260 Infinity II	

HPLC systems (preparative)	Waters 2545 Quaternary Gradient Module Waters 2489 UV/Visible Detector Waters Prep Inject Waters Prep Degasser	Waters
HPLC columns	Unisol C18(2)	Bonna-Agela Technologies (Tianjin, China)
	Reprospher C18	Dr. Maisch GmBH, Ammerbuch, Germany
	Luna C18(2)	Phenomenex Inc., Torrance, California, USA
Incubators	Innova® 44 Incubator Shaker Series	Eppendorf, Hamburg, Ger many
	Certomat HK	Sartorius AG, Göttingen, Germany
	Heraeus Instruments Function Line	Thermo Fisher Scientific, Waltham, Massachusetts, USA
Laboratory Shaker	TS 100 Thermo Shaker	PEGLAB Biotechnologie GmbH, Erlangen, Germany
Magnetic stirrer	Heidolph MR Hei-Standard	Heidolph Instruments GmbH &Co. KG, Schwabach, Germany
Manual Injector Syringe	2.5 mL Gastight Syringe Model 1002 TLL	Hamilton Company, Reno, Nevada, USA
Microscale	Sartorius MC210 P ISO9001	Sartorius AG, Göttingen, Germany
Microwave	Sharp Inverter	SHARP, Sakai, Japa
NanoDrop	NanoDrop 1000	Thermo Fisher Scientific, Waltham, Massachusetts, USA
NMR spectrometer	BRUKER Avance III HD 400 MHz Nano	Bruker, Rheinstetten, Germany; Fällanden, Switzerland

PCR electrophoresis chambers	PerfectBlue Gelsystem Midi ExW	PEQLAB Biotechnologies, Erlangen, Germany
	PerfectBlue Gelsystem Mini L	
Photographic agarose gel documentation system	VILBER	PEQLAB Biotechnologie GmbH, Erlangen, Germany
	Modell?	
pH meter	HANNA Instruments pH 211 Mi- croprocessor pH-Meter	Sigma-Aldrich®, St. Louis, Mis souri, USA
Rotary evaporator	Rotavapor R-210	Büchi Deutschland, Essen, Germany
	Vacuum Controller V-855	
	Vacuum Pump V-700	
	Recirculating Chiller F-108	
	Heating Bath B-491	
Scale	Kern 572	KERN & SOHN GmbH, Balingen, Germany
SDS electrophoresis chambers	Mini-PROTEAN Tetra Vertical Electrophoresis Cell	Bio-Rad Laboratories, Inc., Hercules, California, USA
SpeedVac-Vacuum Concentrator	Concentrator plus	Eppendorf, Hamburg, Germany
Thermocyclers	iCycler BioRad	Bio-Rad Laboratories, Inc., Hercules, California, USA
	peqSTAR	PEQLAB Biotechnologie GmbH, Erlangen, Germany
Vortex mixer	VF2	IKAR-Werke GmbH& Co.KG, Staufen, Germany
UV-Vis	SpectraMax M2	MolecularDevices, San Jose, California, USA

2.2 Consumables

Table 3: Consumables used in this study

Device	Specification	Manufacturer
Amicon Ultra – 4, 10K, 30K		Merck Millipore, Burlington, Massachusetts, USA
Centriprep – 10K, 30K		Merck Millipore, Burlington, Massachusetts, USA
Centrifuge tubes	50 mL, 15 mL	Greiner bio-one, Frickenhausen, Germany
Durapore membrane filter	0,22 µM PVDF	Merck Millipore, Burlington, Massachusetts, USA
Durapore membrane filter	0,45 µM HV For Proteins	Merck Millipore, Burlington, Massachusetts, USA
Electroporation cuvettes	Peqlab brand, 0.2 cm gap	PEQLAB Biotechnologies, Erlangen, Germany
Eppendorf® Safe-Lock microcentrifuge tubes	2 mL, 1.5 mL	Eppendorf, Hamburg, Germany
HisTrap™ HP (5 mL)	Ni-NTA column (*IMAC*)	GE Healthcare
Parafilm		Bemis®, Neenah, WI, USA
Pasteur pipettes		Hirschmann GmbH & Co. KG, Eberstadt, Germany
Pipette tips	2, 50, 200, 1000 µl	Eppendorf, Hamburg, Germany
Sterile filters	0,2 µM	Sartorius AG, Göttingen, Germany
Petri dish		Sarstedt AG & Co. KG, Nümbrecht, Germany

2.3 Chemicals

Table 4: Chemicals used in this study

Chemical	Chemical formula	Manufacturer
α-Ketoglutarate(α-KG)	$C_5H_4O_5^{-2}$	Sigma-Aldrich®, St. Louis, Mis- souri, USA
Acetic acid	$C_2H_4O_2$	Sigma-Aldrich®, St. Louis, Mis- souri, USA
Acetone	C_3H_6O	J.T.Baker®Chemicals, Center Valley, Pennsylvania, USA
Acetonitrile	C_2H_3N	J.T.Baker®Chemicals, Center Valley, Pennsylvania, USA
Agarose	LE 9012-30-6	Genaxxon bioscience GmbH, Ulm, Germany
Ammonium molybdate tetrahy- drate	$(NH_4)_6Mo_7O_{24}\cdot 4H$	Santa Cruz Biotechnology, Dallas, Texas, USA
Boric acid	H_3BO_3	Merck Chemicals GmbH, Darm- stadt, Germany
BSA		Merck Chemicals GmbH, Darm- stadt, Germany
Calcium chloride	$CaCl_2$	Merck Chemicals GmbH, Darm- stadt, Germany
Copper sulfate	$CuSO_4\cdot 5H_2O$	Merck Chemicals GmbH, Darm- stadt, Germany
Deuterated chloroform	$CDCl_3$	Sigma-Aldrich®, St. Louis, Mis- souri, US
Dichlormethane	CH_2Cl_2	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Dithiothreitol	$C_4H_{10}O_2S_2$	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Dipotassiumhydrogenp hos phate	K_2HPO_4	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
D-Lysine	$C_6H_{14}N_2O_2$	Sigma-Aldrich®, St. Louis, Mis- souri, USA
DMSO	C_2H_6OS	Sigma-Aldrich®, St. Louis, Mis- souri, USA
dNTPs		Genaxxon bioscience GmbH,Ulm, Germany
EDTA	$C_{10}H_{16}N_2O_8$	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Ethanol	C_2H_6O	
Ethidium bromide	$C_{21}H_{20}BrN_3$	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Ethylacetate	$C_4H_8O_2$	Carl Roth GmbH & Co. KG, Karlsruhe, Germany,
Ferrous sulfate	$FeSO_4\cdot 7H_2$	Merck Chemicals GmbH, Darm stadt, Germany

Glucose	$C_6H_{12}O_6$	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Glycerol (99 % +)	$C_3H_8O_3$	Thermo Fisher Scientific, Waltham, Massachusetts, USA
Isopropanol	$C_3H_8O_3$	Thermo Fisher Scientific, Waltham, Massachusetts, USA
IPTG	$C_9H_{18}O_5S$	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
L-arabinose	$C_5H_{10}O_5$	Sigma-Aldrich®, St. Louis, Missouri, USA
L-lysine	$C_6H_{14}N_2O_2$	Sigma-Aldrich®, St. Louis, Mis- souri, USA
Magnesium sulfate	$MgSO_4 \cdot 7H_2O$	Merck Chemicals GmbH, Darm- stadt, Germany
Magnesium chloride	$MgCl_2 \cdot 6H_2O$	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Manganese sulfate	$MnSO_4 \cdot H_2O$	Merck Chemicals GmbH, Darm stadt, Germany
Mannitol	C6H14O6	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Methanol	CH_4O	Honeywell, Thermo Fisher Scientific, Waltham, Massachusetts, USA
MOPS	$C_7H_{15}NO_4S$	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
peqGREEN		PEQLAB Biotechnologie GmbH, Erlangen, Germany
Phenol/Chloroform/ Isoamyl alcohol	25:24:1 (*v/v*)	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
PMSF	$C_7H_7FO_2S$	Sigma-Aldrich®, St. Louis, Mis- souri, USA
Polygoprep	60-50 C_{18}	Macherey-Nagel, Düren, Germany
Potassium acetate	$C_2H_3KO_2$	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Potassium chloride	KCl	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Potassiumdihydrogenor thophosphate	KH_2PO_4	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Rubidium chloride	RbCl	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
SDS	$NaC_{12}H_{25}SO_4$	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Sodium chloride	NaCl	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Sodium hydroxide	NaOH	Sigma-Aldrich®, St. Louis, Mis- souri, USA
Soluble starch		Difco Becton Dickinson, Franklin Lakes, New

		Jersey, USA
Soy flour		SOBO Naturkost, Köln, Germany-
Trichloroacetic acid	$C_2HCl_3O_2$	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Trifluoroacetic aci	$C_2HF_3O_2$	Sigma-Aldrich®, St. Louis, Missouri, USA
Triton-X-100	$C_{14}H_{22}O(C_2H_4O)_n$ (n=9-10)	Fluka, Thermo Fisher Scientific, Waltham, Massachusetts, USA
Trizma base	$NH_2C(CH_2OH)_3$	Sigma-Aldrich®, St. Louis, Missouri, USA
Tryptone		Difco Becton Dickinson, Franklin Lakes, New Jersey, USA
TSB		Difco Becton Dickinson, Franklin Lakes, New Jersey, USA
X-Gal		Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Yeast extract		Difco Becton Dickinson, Franklin Lakes, New Jersey, USA
Zinc sulfate	$ZnSO_4 \cdot 7\ H_2O$	Merck Chemicals GmbH, Darm stadt, Germany

2.4 Isotope labeled amino acids and deuterium labeled precursors

Table 5: Labeled amino acids and precursor molecules used for feeding studies

Labeled AA	Chemical formula	Manufacturer
$^{13}C_6$- L-lysine Hydrochlorid	$^{13}C_6H_{14}N_2O_2$. Hcl	Silantes GmbH, Munich, Germany
$^{13}C_5$-L-ornithine Hydrochlorid	$^{13}C_6C_5H_{12}N_2O_2$	Sigma-Aldrich®, St. Louis, Missouri, USA
3,3-d_2-3-(*trans*-2'-nitrocyclopropyl)alanine	C6H8D2N2O4	Prof. Dr. Armin de Meijere, University of Göttingen [173, 174]
3,3-d_2-3-(*trans*-2'-aminocyclopropyl)alanine	C6H10D2N2O2	Prof. Dr. Armin de Meijere, University of Göttingen Brandl, Kozhushkov et al. 2005, Kozhushkov, Zlatopolskiy et al. 2005)
5,5-d_2-4-hydroxylysine lactone	C6H10D2N2O2	Chambers Hughes, Annika Esch

2.5 Enzymes

Table 6: Enzymes used in this study

Enzyme	Specification	Manufacturer
Alkaline Phosphatase	Calf Intestinal Phosphatase (CIP)	New England BioLabs® Inc., Ipswich, Massachusetts, USA
*Bam*I HI	CutSmart® Buffer, 37 °C	New England BioLabs® Inc., Ipswich, Massachusetts, USA
*Eco*RI	CutSmart® Buffer, 37 °C	New England BioLabs® Inc., Ipswich, Massachusetts, USA
Gibson Assembly® Master Mix	1 h, 50 °C	New England BioLabs® Inc., Ipswich, Massachusetts, USA
*Hind*III	CutSmart® Buffer, 37 °C	New England BioLabs® Inc., Ipswich, Massachusetts, USA
Lysozyme	from chicken egg white	Genaxxon Bioscience, Ulm, Germany
*Nde*I	CutSmart® Buffer, 37 °C	New England BioLabs® Inc., Ipswich, Massachusetts, USA

Phusion® High-Fidelity DNA polymerase	See PCR	New England BioLabs® Inc., Ipswich, Massachusetts, USA
Proteinase K	Carl Roth GmbH & Co. KG, Karlsruhe, Germany	New England BioLabs® Inc., Ipswich, Massachusetts, USA
Q5® High-Fidelity DNA polymerase		New England BioLabs® Inc., Ipswich, Massachusetts, USA
RNase	stock solution: 100 mg/mL	New England BioLabs® Inc., Ipswich, Massachusetts, USA
*Spe*I	CutSmart® Buffer, 37 °C	New England BioLabs® Inc., Ipswich, Massachusetts, USA
Taq:Pfu polymerase (9:1)	prepared in-house	Department of Pharmaceu tical Biology
T4 DNA Ligase	T4 DNA ligase buffer, 4 °C	New England BioLabs® Inc., Ipswich, Massachusetts, USA
*Xba*I	CutSmart® Buffer, 37 °C	New England BioLabs® Inc., Ipswich, Massachusetts, USA
*Xho*I	CutSmart® Buffer, 37 °C	New England BioLabs® Inc., Ipswich, Massachusetts, USA

2.6 Kits

Table 7: Kits used in this study

Kit	**Specification**	**Manufacturer**
Cycle Pure Kit	purification of PCR products	PEQLAB Biotechnologies, Erlangen, Germany
QIAquick PCR Purifica tion Kit	purification of PCR products	QIAGEN, Hilden, Germany
QIAquick Gel Extraction Kit	purification of DNA from agarose gels	QIAGEN, Hilden, Germany

2.7 Solutions and Buffers

2.7.1 Buffers and solutions for plasmid isolation from E. coli

Table 8: Buffers and solutions for plasmid isolation from *E. coli*.

Solution	Components	Final concentration	Comments
Sol I	Glucose	50 mM	Adjust to pH 8.0. Add RNase directly bevor use
	EDTA	10 mM	
	Tris (pH 8)	5 mM	
	RNase A	100 µg/ml	
Sol II	NaOH	0,2 N	Store at room temperature
	SDS	1% (w/v)	
Sol III	KAc	3 M	Adjust to pH 4,8 with glacial acetic acid

2.7.2 Buffers for PCR

***Taq* polymerase buffer (*taq* buffer)**

Table 9: Components of *taq* buffer

Solution	Components	Final concentration	Comments
Taq-Puffer (10x)	TRIS	100 mM	pH 9,0
	KCl	500 mM	
	Nonidet P-40	0,8% (w/v)	
	$MgCl_2$	15 mM	

2.7.3 Buffers and solutions for DNA gel electrophoresis

Table 10: Buffers and solutions for DNA gel electrophoresis

Solution	Components	Final concentration	Comments
50x TAE	Tris Base	2 M	Adjust to pH 8.0 with glacial acetic acid
	EDTA	50 mM	
Loading buffer	Bromphenolblau Xylanol FF	0,03%	Store at 4 °C
	Glycerin 98% EDTA	0,03%	
	Tris-HCl (pH=7,6)	60%	
		60 mM	
		10 mM	
Ethidium bromide staining solution	Ethidium bromide	1 µg/ml	
PeqGreen staining solution	20.000x stock solution	1 μL/mL	50x aliquots in water

2.7.4 Buffers for protein purification by nickel affinity chromatography

Buffers for IMAC:

Table 11: Buffers for protein purification by nickel affinity chromatography

Solution	Components	Final concentration
Lysis buffer	Tris-HCl pH 8	50 mM
	NaCl	500 mM
	Tween® 20	1 % (v/v)
	Imidazol	20 mM
	β -mercaptoethanol	10 mM
	Lysozyme	0.5 mg/mL
	PMSF	0.5 mM
Washing buffer	Tris-HCl pH 8	50 mM
	NaCl	500 mM
	Imidazol	20 mM
	β -mercaptoethanol	10 mM
Elution buffer	Tris-HCl pH 8	50 mM
	NaCl	500 mM
	Glycerin	10 % (v/v)
	Imidazol	250 mM
	β -mercaptoethanol	10 mM

Imidazole was added before use. If needed, lysozyme, PMSF and DTT were directly added to the lysis buffer before use. All buffers were filtered (2 µm) before use.

Buffers for the preparation of SDS Gels:

Before the proteins could be separated and analysed by SDS gel, the gels must first be prepared. An SDS gel consists of a separation and a stacking gel. The compounds of each gel are described in Table 12.

Table 12 SDS gel electrophoresis: Composition of buffers and gels for SDS gel electrophoresis

Solution B (200 ml) **Solution C (200 ml)**

Components	**Quantity**	**Components**	**Quantity**
2 M Tris	150 ml	1 M Tris	50 ml
pH 8.8		pH 6.8	
20 % SDS	4 ml	20% SDS	2 ml
ddH_2O	46 ml	ddH_2O	48 ml

12 % separation gel (10 ml) **Stacking gel (5 ml)**

Components	**Quantity**	**Components**	**Quantity**
solution B	2.5 ml	solution C	1.25ml
Polyacrylamide gel	4 ml	Polyacrylamide gel	0.65 ml
dH2O	3.45	ddH2O	3.07
10% APS	80 µl	10 % APS	25 µl
TEMED	8 µl	TEMED	5 µl

Table 13: SDS gel: Composition of loading dye, running buffer, coomassie staining and destaining solution.

	compound	**quantity per liter**
6x Laemmli SDS-Page loading dye	375 mM Tris	2,271 g
	6 % SDS	15 ml
	48 % Glycerol	24 ml
	0,03 % Bromophenol blue	4,5 ml
	9 % 2-Mercaptoethanol	0,015 g
	filled up with ddH2O to 50 ml	
SDS-Page running buffer	Tris	15 g
	Glycine	72 g
	20 % SDS-solution	25 ml
	filled up with dH2O to 5l	
Coomassie staining solution	Methanol	300 ml
	Acetic acid	100 ml
	dH2O	600 ml
	Brillant Blue R-250	1 g
Coomassie destaining solution	Methanol	400 ml
	Acetic acid	100 ml
	dH2O	500 ml

2.7.5 Buffer for 4-OH-Lys in vitro enzyme assay (2.18.6)

Table 14: Enzyme assay kPi buffer

Solution	**Components**	**Quantitiy**	**Comments**
100 ml 50 mM kPi buffer pH8	K_2HPO_4	814,2 mg	pH 8,0
	KH_2PO_4	44,4 mg	

2.7.6 Buffer for genomic DNA isolation from *Nocardia*

Table 15: Lysozyme solution

Solution	Components	Quantitiy	Comments
100 ml Lysozyme solution	Sucrose	10,3 g	10,3 %
	Tris HCl	0,3 g	25 mM
	EDTA	0,73 g	25 mM
	Lysozyme	0,3 g	3 mg/ml

Dissolve in ddH_20 and autoclave, the lysozyme is added directly before use.

2.8 Bacterial growth Media

Unless otherwise stated, all media, solutions and buffers were prepared using the required volume of ddH_2O to fully dissolve the components. Afterwards, they were sterilized for at least 20 minutes at 121 °C and a pressure of 2 bar. Solutions were sterile-filtered (0.2 μm pore size) if temperature sensitive. For the preparation of solid media, 20 g/L agar was added to the media before autoclaving. Antibiotic for selection was added after sterilization and cooling of the media to around 50 °C, if necessary. Liquid cultures were supplemented with the appropriate antibiotic or supplements prior to inoculation.

2.8.1 Cultivation of *Escherichia coli*

LB–Medium [175])

NaCl	10.0 g
Trypton	10.0 g
Yeast extract	5.0 g
Agar (for solid medium)	20.0 g

Ingredients were dissolved in about 900 ml water and the pH adjusted to pH 7. Afterwards, the total volume was adjusted to 1 L with water. Sterilized by autoclaving.

SOB Medium [175]

Difco Bacto tryptone	20.0 g
Difco Bacto yeast extract	5.0 g
NaCl	0.5 g
KCl	0.186 g

Ingredients were dissolved in about 900 mL water and the pH adjusted to pH 7. Afterwards, the total volume was adjusted to 1 L with water. Sterilized by autoclaving.

2x YT Medium: [176]

Tryptone	16.0 g
Yeast extract	10.0 g
NaCl	5 g

Ingredients were dissolved in 100 mL water. Sterilized by autoclaving.

TB Medium

Tryptone	12.0 g
Yeast extract	24.0 g
Glycerol (87%)	4.6 mL

Ingredients were dissolved in 900 mL water. Sterilized by autoclaving.

Phosphate buffer:

KH_2PO_4	2.31 g
K_2HPO_4	12.54 g

Ingredients were dissolved in 100 mL water, sterilized by autoclaving and added to the medium before use.

2.8.2 Cultivation of *Streptomyces*

MS (Mannitol Soya flour) Agar [176]

Mannitol	20.0 g
Soy flour	20.0 g
Agar	20.0 g

Mannitol was dissolved in 1 L tap water and 100 ml of the solution was poured into 300 mL Erlenmeyer flasks, each already containing 2 g of agar and 2 g of soy flour. Sterilized by autoclaving twice at 115 °C for 15 min.

TSB (Tryptone Soy Broth) Medium (Kieser et al. 2004)

Tryptone Soy Broth 30.0 g

Ingredients were dissolved in 1 L water and sterilized by autoclaving.

SM19 Medium: Belactosin Production Medium [77]

Tomato paste	40 g
Oat flour	15 g
Glucose	2 g

All ingredients were dissolved in 1 L tab water and autoclaved afterwards.

R5 production Medium:

Sucrose	103.0 g
K_2SO_4	0.25 g
$MgCl_2$ x 6 H_2O	10.12 g
Glucose	10.0 g
Difco Casaminoacids	0.1 g
Trace elements solution [1)]	2.0 mL
Difco Yeast extract	5.0 g
TES buffer	5.73 g

[1)] Trace elements solution:

$ZnCl_2$	40 mg
$FeCl_3$ x 6 H_2O	200 mg
$CuCl_2$ x 2 H_2O	10 mg
$MnCl_2$ x 4 H_2O	10 mg
$Na_2B_4O_6$ x 10 H_2O	10 mg
$(NH_4)_6Mo_7O_{24}$ x 4 H_2O	10 mg

Dissolved in 1 L water and sterilized by autoclaving.

50 ml of R5 media were prepared in 300 ml baffled Erlenmeyer flask equipped with a metal spiral and the following supplements were added right before inoculation:

KH_2PO_4 (0,5%)	0.5 mL
$CaCl_2$ x 2 H_2O	0.2 mL
L-proline (20%)	0.75 mL
NaOH (1N)	0.35 mL

2.8.3 Cultivation of *Nocardia*

BHI (Brain Heart Infusion broth: (Sigma-Aldrich®, St. Louis, Mis souri, USA):
BHI medium was used for the standard cultivation of Nocardia strains. Therefor 30 g of the ready to use powder were dissolved in 1 L ddH_2O and autoclaved.

Hoshino Medium: [152]

Glucose	30 g
Pharmamedia	6,0 g
Soybean flour	7,5 g
K2HPO4	1,0 g
CaCO3	5,5 g
(NH4)2SO4	5,0 g

2.9 Primer

Table 16: Primers used for biosynthesis studies of belactosin A

Name	Nucleotidsequence (5' → 3')
KOBelN_FP	CACGTCGCGGCGGCCGGTACCGAGCCCGTGCGCGCGCTAATTCCGGGGATCCGTCGACC
KOBelN_RP	CTTCGTGGAGGTCGGATGTCAATGTGCGCCGCGGGCGCGCGCGTCGCCGTGCTTG
screenBelN_FP	TGTGATGCCGCTCCATCCAG
screenBelN_RP	CGTACTGGTGCCAGTGGCGG
belUS_FP	AGATTATGGGCGCCTCGCTG
belUS_RP	TCGCCGGGTTCATCAACATC
belDS_FP	ACGAGACGACGCTCCATCTG
belDS_R	CTTGTTCGTCCTCGCGATAC
pUWL_belN_FP	GGTTGGTAGGATCGACGGTATCGATAAGCTTATGTGAACGGGGACACCCG
pUWL_belN_RP	CACCGGCGCCGGCTCATCGTGATCAAGATCTTCGAGACGTAATTACTTAG
pUWL_test_fwd	ACGCCTGGTCGATGTCGGAC
pUWL_new_rev	GAGCGAGGAAGCGGAAGAGC
KOBelK_FP	CACGTCGCGGCGGCCGGTACCGAGCCCGTGCGCGCGCTAATTCCGGGGATCCGTCGACC
KOBelK_RP	GGGGAAACGCCGAATCCCTGTCGGCGGACCGAGGAAATGTGTAGGCTGGAGCTGCTTC
screenBelK_FP	TGTGATGCCGCTCCATCCAG
screenBelK_RP	CCAGACAAATCCCTGACAAG
KOBelL_FP	CCAGCGTGTGGTACTTCACGACGGTGGTCATCCGCGGCCATTCCGGGGATCCGTCGACC
KOBelL_RP	CCGCCACCCCGAAAGGGATCGACCGTGACTCTCAACGCGTGTAGGCTGGAGCTGCTTC
screenBelL_FP	TGAACACCCGTACCCGGAAC
screenBelL_RP	CCGGCTGGTACAAGAACGTG

KOBelM_FP	CCCGTCCCGCCGGGCCGCGGGTGTCCCCGTTCACCGCGGATTCCGGGGATCCGTCGACC
KOBelM_RP	GGAGGGCGCCATCGCCGAGGGCCGCGGATGACCACCGTCTGTAGGCTGGAGCTGCTTC
screenBelM_FP	ACCTTCACGCACAGGGTTCC
screenBelM_RP	ATCACGGTGGTGTCGTTCTC
KO_orf-1_FP	CCGTGACCGGCGCCCGGGCCGTCCGCGGGCGGAGCGTCAATTCCGGGGATCCGTCGACC
KO_orf-1_RP	AAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACTGTAGGCTGGAGCTGCTTC
screen_orf-1_FP	ACCGGCTGGAGATCAACGTC
screen_orf-1_RP	CTCGTATGTTGTGTGGAATTGTGAGC
KOBelA-E_FP	AGGAAGTTTCCCCTCCGGCGCGTAACGTGGCGCAGCATGATTCCGGGGATCCGTCGACC
KOBelA-E_RP	AAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACTGTAGGCTGGAGCTGCTTC
screenBelA-E_FP	TCGACGTGGACGTGGCAAAG
screenBelA-E _RP	CTCGTATGTTGTGTGGAATTGTGAGC
KOBelO-V_FP	ATTCCGGGGATCCGTCGACC
KOBelO-V_RP	TGTAGGCTGGAGCTGCTTC
screenBelO-V_FP	CGTGGAGTGGTCCAATAATG
screenBelO-V _RP	ACCTCCTTCGGCTTCAACG
KOBelQ-V_FP	CTATAGGGCGAATTCGAGCTCGGTACCCGGGGATCCCACATTCCGGGGATCCGTCGACC
KOBelQ-V_RP	CAAACGCTACGTTCGATCGGCGGACTCAGCCACGCACTATGTAGGCTGGAGCTGCTTC
screenBelQ-V_FP	GGATGTGCTGCAAGGCGATTAAGTTGG
screenBelQ-V _RP	TCGGCCTGAAACCGGACAAC
KOBelR-V_FP	CTATAGGGCGAATTCGAGCTCGGTACCCGGGGATCCCACATTCCGGGGATCCGTCGAC
KOBelR-V_RP	ATATTTTGATCATCTCACCCGGGAGGAGCTCTCATCATGTGTAGGCTGGAGCTGCTTC

screenBelR-V_FP/	GGATGTGCTGCAAGGCGATTAAGTTGG
screenBelR-V _RP	GTGAATCCAATCCGGTTCTC
KOBelT-V_FP	AAGCCGCGGTGGCCGGCCCGCTCGGTCAGGGCCGGTTCAATTCCGGG GATCCGTCGACC
KOBelT-V_RP	CAGGCCCTCCACACCCGCACAGGGAGGAACACCCCCATGTGTAGGCT GGAGCTGCTTC
screenBelT-V_FP	ACCCGAGTACATCCGGTACG
screenBelT-V _RP	TGTCATGCCCATGCCATTTC
pET28a_belL_EcoRI_f	CTCTCTGAATTCGTGACTCTCAACGCGCAG
pET28a_belL_HindIII_r	CTCTCTAAGCTTTCATCCGCGGCCCTCGGC

Table 17: Primers used for *Nocardia* project

Name	**Nucleotidsequence (5' → 3')**
pCC1FOS_FP	GGATGTGCTGCAAGGCGATTAAGTTGG
pCC1FOS_RP	CTCGTATGTTGTGTGGAATTGTGAGC
ACAD-cluster-left_f	CGATACTCGCCAGCAATCCG
ACAD-cluster_left_r	GCAGCCGATTCCGATGATCG
ACAD-cluster-right_f	CTACGGCATCTGGCCGATAT
ACAD-cluster_right_r	CGTACACGTCGGTTTGGAGG
screen_ACAD_GUH-2_f	TGGCACGAGTACCCAAGAGC
screen_ACAD_GUH-2_r	GCTGGAGCCTTCGAAGATGC

2.10 Antibiotics and Inductors

Antibiotics were dissolved in appropriate solvents as stock solutions and kept at -20 °C. The aqueous solutions were sterilized by passing through a 0.22 µm filter. The solutions in ethanol and DMSO were auto sterile. For antibiotic selection, the required antibiotics were added to the cooled media (room temperature to 60 °C) in appropriate concentration.

Table 18: Antibiotics used in this study

Antibiotic	**stock solution (mg/mL)**	**media (mg/mL)**	**Solvent**
Apramycin	50	15-50[a]	H_2O
Carbenicillin	50-100	50-100	H_2O
Chloramphenicol	25-50	25-50	ethanol
Kanamycin	50	15-50[a]	H_2O
Tetracycline	12	12	ethanol
Nalidixic acid	25	25	0.3 M NaOH

2.11 Vector, plasmids and fosmids

Table 19:Vectors used in this study

Name	Description	Source or reference
pBluescript II SK(-)	$Carb^R$, 2958 bp , cloning vector	Stratagene Inc., La Jolla, California, USA
pHis8	Expression vector; pET-28a(+) (Merck) derivative, N-terminal His8-tag and C-terminal His6-tag, T7 promotor, KanR	[177]
pUWL-apra-oriT	E. coli - Streptomyces shuttle vector, pUWL201 derivative, ermE* promotor, colE1, rep; AmpR, ApraR	Andreas Günther, Universität Frankfurt
pET28a (+)	Expression vector (Merck), N- and C-terminal His6-tag, T7 promotor, KanR	
pCC1FOS		

Table 20: Plasmids used in this study

Name	Description	Source or reference
pIJ773	pBS SK(+)-derivative, P1-FRT-oriT-aac(3)IV-FRT-P2	(Gust et al. 2003)
pUZ8002	tra, RP4 ; KanR	(Paget et al., 1999)
pUB307	„Driver"-Plasmid for triparental conjugation; tra oriT ; KanR, CmlR	(Flett et al., 1997)
pKD46	λ-RED (exo, bet, gam), araC, bla, repA101ts; CarbR	Datsenko and Wanner, 2000

Table 21: Fosmids used in this study for belactosin studies

Name	Description	Source or reference
merLK01	attP, int, oriT, neo, neoP), Kan^R	[178]
6D1	pCC1FOS containing the belactosin biosynthetic gene cluster from *Streptomyces* sp. UCK14, Cm^R	Felix Wolf
belFeW01	from 6D1, chloramphenicol resistance gene replaced by cassette from merLK01 (attP, int, oriT, neo, neoP), Kan^R	Felix Wolf
belFeW02	from belFeW01, *belK* is replaced by an apramycin resistance cassette from pIJ773, $Apra^R$, Kan^R	Felix Wolf
belFeW03	from belFeW01, *belL* is replaced by an apramycin resistance cassette from pIJ773, $Apra^R$, Kan^R	Felix Wolf
belFeW04	from belFeW01, *belN* is replaced by an apramycin resistance cassette from pIJ773, $Apra^R$, Kan^R	Felix Wolf
belFeW05	from 6D1 *belK* is replaced by an apramycin resistance cassette from pIJ773, $Apra^R$,Cm^R	Felix Wolf
belFeW06	from 6D1 *belL* is replaced by an apramycin resistance cassette from pIJ773, $Apra^R$,Cm^R	Felix Wolf
belFeW07	from 6D1 *belM* is replaced by an apramycin resistance cassette from pIJ773, $Apra^R$,Cm^R	Felix Wolf
belFeW08	from 6D1 *belN* is replaced by an apramycin resistance cassette from pIJ773, $Apra^R$,Cm^R	Felix Wolf
belFeW09	Δ*belK* from belFeW05, flipped apramycin resistance cassette at FRT sites, 81 bp scar, chloramphenicol resistance gene is replaced by apramycin resistance cassette, $Apra^R$	Felix Wolf
belFeW10	Δ*belL* from belFeW06, flipped apramycin resistance cassette at FRT sites, 81 bp scar, chloramphenicol resistance gene is replaced	Felix Wolf

	by apramycin resistance cassette, Apra^R	
belFeW11	Δ*belM* from belFeW07, flipped apramycin resistance cassette at FRT sites, 81 bp scar, chloramphenicol resistance gene is replaced by apramycin resistance cassette, Apra^R	Felix Wolf
belFeW12	Δ*belN* from belFeW08, flipped apramycin resistance cassette at FRT sites, 81 bp scar, chloramphenicol resistance gene is replaced by apramycin resistance cassette, Apra^R	Felix Wolf
belFeW13	Δ*orf-1* from belFeW01, *orf-1* is replaced by an apramycin resistance cassette from pIJ773flipped apramycin resistance cassette at FRT sites, 81 bp scar, Kan^R	Felix Wolf
belFeW14	Δ*belA-E* from belFeW01, *belA-E,* is replaced by an apramycin resistance cassette from pIJ773, flipped apramycin resistance cassette at FRT sites, 81 bp scar, Kan^R	Felix Wolf
belFeW15	Δ*belO-V* from belFeW01, *belO-V* is replaced by an apramycin resistance cassette from pIJ773, flipped apramycin resistance cassette at FRT sites, 81 bp scar, Kan^R	Felix Wolf
belFeW16	Δ*belQ-V* from belFeW01, *belQ-V* is replaced by an apramycin resistance cassette from pIJ773, flipped apramycin resistance cassette at FRT sites, 81 bp scar, Kan^R	Felix Wolf
belFeW17	Δ*belR-V* from belFeW01, *belR-V* is replaced by an apramycin resistance cassette from pIJ773, flipped apramycin resistance cassette at FRT sites, 81 bp scar, Kan^R	Felix Wolf
belFeW18	Δ*belT-V* from belFeW01, flipped apramycin resistance cassette at FRT sites, 81 bp scar, Kan^R	Felix Wolf

Table 22: Fosmids use in this study in the Nocardia project

Name	Description	Source or reference
21C6	pCC1FOS containing the ACAD biosynthetic gene cluster from Nocardia cyriacigeorgica GUH-2, CmR	this study
guhAE01	from 21C6, chloramphenicol resistance gene replaced by cassette from merLK01 (attP, int, oriT, neo, neoP), KanR	this study

2.12 Bacterial expression strains

Table 23: *E. coli* strains used in this study

Name	Description	Reference
E. coli XL1 blue MRF^{-}	*rec*A1 *end*A1 gyrA96 *hsd*R17 *sup*E44 *thi*- 1 *rel*A1 lac [F'proAB lacI^{q}ZΔM15, Tn10, (Tet^{R})]	Stratagene
E. coli ET 12567	*dam, dcm hsdM, hsdS, hsdR,* Tet^{R}, Cm^{R}	[179]
E. coli DH5*α*	*huA2, lacΔU169,* phoA, glnV44, Φ 80', lacZΔM15, gyrA96, recA1, relA1, endA1, thi-1, hsdR17	[180]
E. coli Top10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^{R}) endA1 λ^{-}	Invitrogen (Waltham, Massachusetts, USA) [181]
E. coli BW25113/pKD46	K-12 Derivat: *ΔaraBAD, ΔrhaBAD* Kan^{R} recombination plasmid: λ/RED (*gam, bet, exo*), *araC, rep101ts* ,Amp^{R}	[182 183]
E. coli BT340	DH5α/pCP20	[184]
E. coli ET 12567/pUB307	*dam, dcm hsdM, hsdS, hsdR,* Tet^{R}, Cm^{R}; conjugation plasmid: *tra*, RP4, *oriT* Kan^{R}	[185]
E. coli Rosetta(DE3) pLysS	Host for the heterologous expression of His8 tagged proteins (F- *omp*T *hsd*SB(rB- mB-) *gal dcm* λ(DE3 [*lacI lacUV5-T7* gene 1 *ind1 sam7 nin5*]) pLysSRARE); Cm^{R}	Invitrogen
E. coli EPI300^{TM}	T1R Phage T1-resistenter *E. coli* [F– *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) (StrR) Φ80d*lacZ*ΔM15, Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara,leu*)7697 *gal*U *gal*K λ– *rps*L *nup*G *trf*A *ton*A *dhfr*]	Epicentre®

Table 24: *Actinobacteria* strains used in this study

Name	Description	Reference
Amycolatopsis japonicum MG417-CF17	EDDS producer strain	[186, 187]
Streptomyces albus J1074	*S. albus* G derivative: Defect in SalGI-mediated Restriction, valine and isoleucine auxotroph	[188, 189] gift from the John-Innes-Centre, Norwich, UK
Streptomyces sp. UCK14	Belactosin producer, wilde type strain	International Patent Organisms Depositary, NITE as FERM BP-5203
Nocardia cyriacigeorgica GUH-2	Wild type strain	
S. albus J1074 bel01	Heterologous expression of the belactosin BGC in *S. albus* J1074	Felix Wolf
S. albus J1074 bel01/Δ*belK*	Δ*belK* mutant in the belactosin heterologous expression strain S. albus J1074 *bel01*	Felix Wolf
S. albus J1074 bel01/Δ*belL*	*ΔbelL* mutant in the belactosin heterologous expression strain *S. albus* J1074 *bel01*	Felix Wolf
S. albus J1074 bel01/Δ*belN*	*ΔbelN* mutant in the belactosin heterologous expression strain *S. albus* J1074 *bel01*	Felix Wolf
S. albus J1074 bel01/Δ*orf-1*	cluster border elucidation strain *Δorf-1* in *S. albus* J1074	Felix Wolf
S. albus J1074 *bel01/ΔbelA-E*	cluster border elucidation strain *ΔbelA-E* in *S. albus* J1074	Felix Wolf
S. albus J1074 *bel01/ΔbelO-V*	cluster border elucidation strain *ΔbelO-V* in *S. albus* J1074	Felix Wolf

S. albus J1074 *bel01/ΔbelQ-V*	cluster border elucidation strain *ΔbelQ-V* in *S. albus* J1074	Felix Wolf
S. albus J1074 *bel01/ΔbelR-V*	cluster border elucidation strain *ΔbelR-V* in *S. albus* J1074	Felix Wolf
S. albus J1074 *bel01/ΔbelT-V*	cluster border elucidation strain *ΔbelT-V* in *S. albus* J1074	Felix Wolf
S. sp. UCK14*/ΔbelK*	*ΔbelK* mutant in the belactosin wild type producer srain *S.* sp. UCK14	This work
S. sp. UCK14*/ΔbelL*	*ΔbelL* mutant in the belactosin wild type producer srain *S.* sp. UCK14	This work
S. sp. UCK14*/ΔbelM*	*ΔbelM* mutant in the belactosin wild type producer srain *S.* sp. UCK14	This work
S. sp. UCK14*//ΔbelN*	*ΔbelN* mutant in the belactosin wild type producer srain *S.* sp. UCK14	This work
Amycolatopsis japonicum MG417-CF17/ ACAD GUH-2	Heterologous expression of ACAD GUH-2 BGC in *Amycolatopsis japonicum* MG417-CF17, *Amycolatopsis japonicum* MG417-CF17 + guhAE01	This work
S. albus J1074/ ACAD GUH-2	Heterologous expression of ACAD GUH-2 BGC in *S. albus* J1074, *S. albus* J1074 + guhAE01	This work

2.13 Size standards for gel electrophoresis and SDS Page

Table 25: Size standards used in this study

Size standard	Specification	Manufacturer
GeneRuler 1 kb DNA Ladder	250-10000 bp	Thermo Fisher Scientific, Waltham, Massachusetts, USA
GeneRuler High Range DNA Ladder	10171-48502 bp	Thermo Fisher Scientific, Waltham, Massachusetts, USA
Color Prestained Protein Standard, Broad Range	(10-250 kDa)	New England BioLabs® Inc., Ipswich, Massachusetts, USA

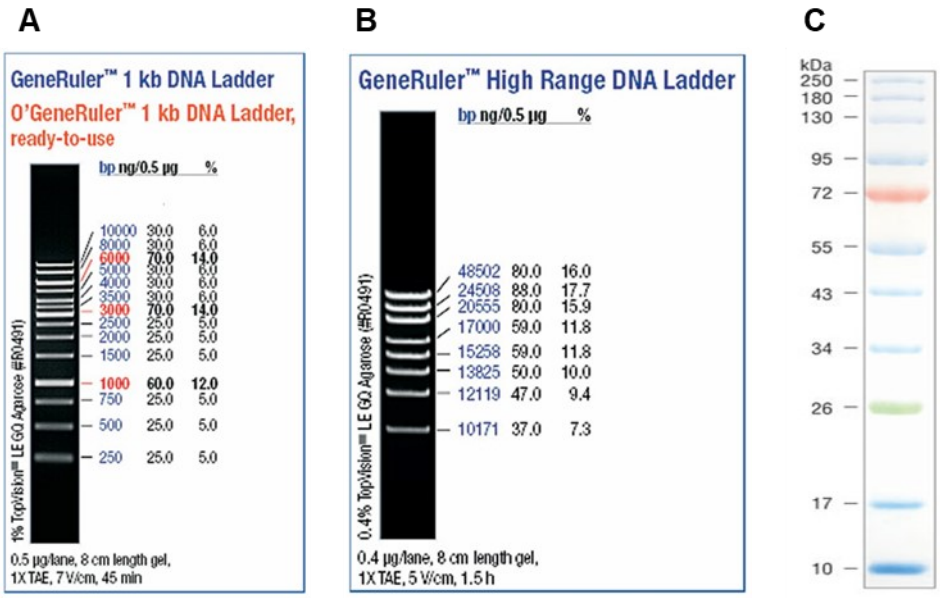


Figure 9: DNA and protein size standards:A: Thermo Scientific™ GeneRuler 1 kb DNA Ladder. **B:** Thermo Scientific™ GeneRuler High Range DNA Ladder. **C:** Color Prestained Protein Standard, Broad Range (10-250 kDa)

2.14 Software

Table 26: Software used in this study

Name	Specification	Manufacturer
ChemBio Draw	Ultra, Version 14.0.0.117	PerkinElmer, Waltham, Mas-
DataAnalysis		Agilent Technologies, Santa Clara, California USA
DataAnalysis	6300 Series Ion Trap LC-MS Software 6.1	BRUKER DALTONIK GmbH Life Sciences, Bremen, Germany
OpenLAB CDS	C.01.10	Agilent Technologies, Santa Clara, California USA
PyMol	PyMol 2.4	Schrödinger
Äkta Unicorn	Unicorn 7	Cytivia, Marlborough, Massachusetts, USA
Chemstation	B.04.03	Agilent Technologies, Santa Clara, California, USA

Methods

2.15 Cultivation and strain maintenance

2.15.1 Cultivation and preservation of Escherichia coli

For cloning experiments, *E. coli* cultures were grown over night at 37 °C in liquid or solid LB-medium. Antibiotics were added in appropriate concentration to the medium. Cultivation in liquid medium was performed while shaking at 220 rpm.
For recombineering experiments with *E. coli* BW12567 strains, the incubation temperature was 30 ºC.
E. coli strains were preserved in 20 % Glycerol stock solutions at – 80 °C

2.15.2 Cultivation of *Actinobacteria*

Cultivation of *Actinobacteria* strains was carried out in liquid TSB or on solid MS medium. Liquid cultures were performed in 300 mL baffled Erlenmeyer flasks containing 50-mL medium and a stainless steel spring at 200 rpm and 30 °C for 2-3 days. Appropriate antibiotics were added to the medium.

For production of belactosin A, 5 ml of a two-day-old TSB culture of the respective strains were added to 50 mL of SM19 production medium and grown at 30 °C and 220 rpm for 5 days. Alternatively, 10 µl of spores were inoculated onto a SM19 Agar plate (10 ml, small petri dishes, 6 cm diameter).

For heterologous expression of the belactosin and ACAD gene cluster, 50 µl of a two-day TSB pre culture were inoculated into 50 ml of R5 medium and incubated for 5 days at 200 rpm.

2.15.3 Preparation of spore suspensions of Actinobacteria

For preparation of spore suspensions, the respective Streptomyces strains were spread on MS agar plates and incubated at 30 °C for approximately 5-7 days until sporulation of the cultures was observed. The spores were harvested by adding 3 mL of sterile water to each plate and scraping of the spores from the top of the plates. The resulting suspension was transferred into a falcon tube and vortexed vigorously. Afterwards, the spores were separated from the mycelium by passing the through a sterile cotton plugged syringe. Spores were collected by

centrifugation (4000 rpm, 10 min, 4 °C) and resuspended in 500 µL of 20% glycerol, or accordingly to density of spores. The Spore suspensions were stored at -80 °C.

2.16 Methods of molecular biology and biochemistry

2.16.1 Isolation of genomic DNA from *Streptomyces spp. UCK14*

For isolation of genomic DNA, 25 mL of a 3-day-old culture in TSB medium were harvested by centrifugation (5000 rpm, 4 °C, 1 min). The cells were resuspended in 10 ml solution I () and a spatula tip of lysozyme and 100 µL of proteinase K (20 mg/mL) were added. The suspension was incubated for 1 h at 37 °C. Subsequently, 1 mL of a 10% SDS solution was added, followed by another incubation period of 30-60 min at 60 °C. Afterward 5 mL phenol/chloroform/isoamyl alcohol (25:24:1) were added. The mixture was gently shaken and centrifuged at 5000 rpm and 4 °C for 10 min. The upper layer was recovered and extracted multiple times with 5 mL phenol/chloroform/isoamyl alcohol, until the supernatant appeared clear. The aqueous phase was separated by centrifugation and genomic DNA was precipitated by addition of 0.1 volume of NaAc solution and 0.7 volume of isopropanol. After centrifugation (5000 rpm, 4 °C, 3 min). The DNA was washed in ice cold 70% ethanol twice. After a short drying period at RT, the DNA was resuspended in ddH_2O and the DNA solution stored at -20 °C.

2.16.2 Isolation of plasmid or cosmid DNA of *E. coli* – mini-, midi- and maxi preparation

For the isolation of plasmids, vectors and fosmids out of E. coli, alkaline lysis was performed. Depending on the required amount of DNA, mini-, midi- or maxi preparations were performed. For minipreparations, 2 mL of an overnight culture (10 ml for midi, 50 ml for Maxi) of the strain containing the respective plasmid was spun down (13 000 rpm, 2-5 min, 4 °C). The resulting pellet was resuspended in 250 µL (2; 5 mL) solution I (Sol I) (2.7.1, Table 8) containing 1 µL/mL RNase and incubated for 5 minutes at room temperature. After incubation, 250 µL (2; 5 mL) solution II (Sol II) (2.7.1, Table 8) was added and mixed by gently inverting the vessel several times. In the next step 250 µL (2; 5 mL) ice-cold solution III (Sol III) (2.7.1, Table 8) was added and the contents were mixed again by inverting the vessel. Afterwards, the supernatant containing plasmids was obtained by centrifugation for 10

minutes at 13 000 rpm, 4 °C (6000 rpm, 4 °C). The supernatant was transferred into a new vessel. Subsequently, the DNA was purified by the phenol chloroform method (2.16.4) or directly precipitated (2.16.3). The resulting DNA pellet was dissolved in 50 μL (500 μL) ddH_2O and the DNA was stored at -20 °C until further use.

2.16.3 Precipitation of DNA

For the precipitation of DNA with alcohol, the DNA solution was thoroughly mixed with one-tenth volume of 3 M sodium acetate and one volume of isopropanol. After incubation on ice for 10 minutes, the mixture was centrifuged for 30 minutes at 4 °C (13.000 rpm). The supernatant was removed and the pellet was washed with 500 μL 70 % ethanol. After a second centrifugation step (13000 g, 5 min, 4 °C), the supernatant was discarded and the pellet was dried at RT. Afterwards, the pellet was dissolved in an appropriate volume of ddH_2O. The DNA solution was stored at -20 °C.

2.16.4 Purification/extraction of DNA with phenol/chloroform/isoamyl alcohol

First the DNA solution was transferred into a 2 mL safety lock tube and mixed with an equal volume of phenol/chloroform/isoamyl alcohol by vortexing for 1 minute. After centrifugation of the mixture (13.000 x g, 10 min, 4 °C), the aqueous phase containing the DNA was pipetted of and transferred into a new vessel. The DNA was subsequently precipitated as described in 2.16.3.

2.16.5 Purification of DNA using the Cycle Pure Kit or QIAquick PCR purification Kit

After PCR reactions or DNA restriction DNA was purified using the Cycle Pure or QIAquick PCR purification Kit. The purification was performed according to the manufacturer instructions. Briefly, the DNA solution was mixed with 1:1 volume of binding buffer and applied to the column. After centrifugation (13.000 rpm, 1 min), the flow-through was discarded and the column was washed twice with wash buffer. After an empty centrifugation step, the DNA was eluted with 50 ddH_2O or less, regarding desired DNA concentration (5000 rpm, 3 min).

2.16.6 Polymerase Chain Reaction (PCR) DNA amplification

PCR amplifications were carried out with thermocyclers iCycler BioRad (Bio-Rad Laboratories, Inc., Hercules, California, USA) or peqSTAR (QIAGEN, Hilden, Germany). Annealing temperature and elongation time were adjusted to the individual primers and template size, respectively. Pipetting schemes and cycle conditions for PCR with *taq:pfu* DNA polymerase are displayed in and for phusion DNA polymerase in, respectively. *Taq:pfu* is mixed in ratio 9:1.

Taq:Pfu-Polymerase:

Table 27: Reaction mix for taq:pfu PCR

Component	50 µl reaction
taq buffer	5 µl
10 mM dNTPs	2 µl
10 µM forward primer	1 µl
10 µM reverse primer	1 µl
DMSO	2.5 µl
Taq:pfu DNA polymerase	1 µl
Template DNA	variable
ddH_2O	up to 50 µl

Table 28: Cycle conditions for taq:pfu PCR

	Temperature [°C]	Length [s]	Cycles
Hotstart	98 °C	60	1
Denaturation	98 °C	30	30-35
Annealing	55 °C	30	
Elongation	72 °C	60/kb	
Terminal elongation	72 °C	300-600	1
Cooling	4 °C	∞	1

Phusion DNA Polymerase:

Annealing temperatures required for use with Phusion DNA polymerase tend to be higher than with other polymerase and were calculated with the NEB T_m calculator for each primer pair individually, ranging between 55 °C and 82 °C.

Table 29: Reaction mix for phusion PCR

Component	25 µl reaction
5X Phusion HF or GC buffer	5 µl
10 mM dNTPs	0.5 µl
10 µM forward primer	1.25 µl
10 µM reverse primer	1.25 µl
DMSO	0.75 µl
Phusion DNA polymerase	0.25 µl
Template DNA	variable
ddH_2O	up to 25 µl

Table 30: Cycle conditions for phusion PCR

	Temperature [°C]	Length [s]	Cycles
Hotstart	98°C	30	1
Denaturation	98 °C	10	30-35
Annealing	55-80 °C	*10-30*	
Elongation	72 °C	*30/kb*	
Terminal elongation	72 °C	*300*	1
Cooling	4 °C	∞	1

2.16.7 DNA manipulation with restriction enzymes

Restriction of DNA was carried out to generate DNA inserts for cloning or confirmation of generated plasmids and fosmids. Restriction dsDNA was applied via restriction endonucleases. The following reaction mix was used. The corresponding restriction buffer and incubation conditions were applied as indicaded by the manufacturer (NEB Biolabs).

Table 31: Reaction mixture for restriction digestion

DNA	20–100 ng
Restriction endonuclease	1 µl (of each RE)
10x restriction buffer	2 µl
dH_2O	x µl
total volume	**20 µl**

The whole restriction mix was incubated at 37 °C for 30 min to 1 h and the restriction endonucleases were heat inactivated at 65 °C for 10 min afterwards.

2.16.8 Dephosphorylation of 5´ and 3´ ends of DNA

To minimize intramolecular religation of vectors with compatible ends, the vector DNA was treated with an alkaline phosphatase (CIP). Therefore, 1U CIP per 1 µL DNA was added directly to the performed restriction digest and the mixture was incubated for 30 min at 37 °C.

2.16.9 Ligation of DNA

Ligation of DNA fragments was achieved by using T4-DNA ligase (Thermo Scientific). The ligation preparation contained 1U T4 DNA ligase, 1x ligation buffer and about 100 ng DNA in a 5:1 or 3:1 ratio mixture of linearized vector and insert in a final volume of 10 µL. The mixture was incubated for 1 h at room temperature or over night at 4 °C.

Table 32: Ligation reaction mixture

DNA fragment	x µl
Linearized vector	x µl
T4 DNA Ligase	1 µl
10x T4 DNA Ligase reaction buffer	2 µl
dH_2O	x µl
total volume	**10 µl**

The mixture was incubated for 1 h at room temperature or over night at 4 °C. The mixture was incubated for 1 h at room temperature or over night at 4 °C. The T4-Ligase was heat inactivated at 65 °C for 10 min.

2.16.10 Transformation of DNA

Preparation of electrocompetent cells

For fresh electrocompetent *E. coli* cells, 100 mL LB medium (with antibiotics if necessary) was inoculated with a 50 μL of an over-night culture and incubated at 200 rpm, 37 °C until OD_{600} between 0.4 and 0.6. Then cells were transferred into pre-cooled vessels (2x 50 mL) and stored on ice for 15 minutes. Cells were harvested by centrifugation (4000 rpm, 10 min, 4 °C) and washed twice with 25 mL ice-cold 10 % glycerol. Following the last centrifugation step, the supernatant was discarded and cells were resuspended in the residual 10 % glycerol. The suspension was aliquoted into 50 μL portions and stored at -80 °C.

Electroporation

For the transformation of DNA, electrocompetent *E. coli* cells were thawed on ice and mixed with 1-2 μL of DNA to be transformed. The DNA-*E. coli*-mixture was transferred into a pre-cooled electroporation cuvette (0.2 cm gap) and the mixture was pulsed at 200 **Ω**, 25 μF and 2.5 kV. Immediately afterwards, 500 μL of pre-cooled LB medium was added to the cuvette and cells were transferred into a 1.5 mL tube. The cells were regenerated for one hour at 37 °C and 300 rpm. Cells were subsequently plated onto LB agar plates containing appropriate antibiotics for selection and incubated at 37 °C overnight if not stated otherwise.

Preparation of chemically competent cells

For the preperation of chemically competent cells, 100 ml LB medium complemented with 2 ml of magnesium solution was inoculated with 500 µl of an over-night culture of *E. coli* cells. The cells were grown up to an OD_{600} of 0.4 to 0.6 at 37 C and 200 rpm. The logarithmic culture was then centrifuged at 4 000 rpm. for 10 min and 4 °C. The supernatant was discarded and the pellet was resuspended in 5 ml TMF buffer (Table 33). After adding 20 % glycerol to the suspension, the competent cells were aliquoted in 100 µl fractions and stored at -80 °C.

Buffer for competent cells preparation:

Table 33: Composition of magnesium solution and TMF-buffer

magnesium solution	**TMF-buffer**
500 mM MgCl2	100 mM CaCl2·2H2O
500 mM MgSO4	40 mM MnCl2·2H2O
	50 mM RbCl

Heat shock

100 μL chemically competent cells were thawed on ice and mixed with around 10 μL of DNA to be transformed or the complete ligation sample. The mixture was incubated on ice for 30 minutes. The heat shock was then carried out at 42 °C for 1 min. Subsequently, cells were placed on ice for another 10 minutes and afterwards 1 ml of LB medium was added. Regeneration of the cells was carried out at 37 °C, 300 rpm for at least one hour. Cells were plated on LB agar plates with appropriate antibiotics and incubated at 37 °C over night.

2.16.11 Agarose gel electrophoresis

Size accordingseparation of DNA fragments was achievd by gel electrophoresis with 0.8-1.5% (w/v) agarose . Agarose in the desired concentration was boiled up in 1x TAE buffer. Before pouring the solution in a trav of a gel chamber the RT cooled agarose was stained with peqGREEN in a final concentration of 1 μL/mL. The agarose gel, equipped with a comb and was cooled down till complete polymerization. After filling up the chamber with 1x TAE as running buffer and removing the comb, the gel pockets were loaded with sample, that ha been mixed with 6 x loading dye in a 6:1 ratio. By default, 6 μL (0.5 μg) of the GeneRuler 1 kb DNA ladder was used as DNA size standard. Gels were run at 100-120 V for 45 min. DNA was visualized under UV light at 312 nm and photographed by using Eagle Eye II gel documentation system. DNA concentration was determined by comparison to the GeneRuler size standard. For preparative agarose gel DNA fragments were separated at lower voltage at low voltage and the desired DNA fragment was excised from the gel with a scalpel under UV light. DNA was extracted from the gel using a gel extraction kit according to the manual. If necessary, gels were alternatively stained with the fluorescent dye ethidium bromide.

2.16.12 Colony PCR

Generated mutant strains where analyzed for harboring the respective plasmids or genetic modifications where teste by colony PCR for quick verification.

Colony PCR with *E. coli* cells was carry out according to the standard protocol for *taq:pfu* polymerase (). Instead of template DNA, ddH_2O was added and single colonies of the *E. coli* clones were picked with a sterile toothpick and resuspended in the respective PCR tube.

2.16.13 Gibson Assembly

Gibson Assembly is generally used for the assembly of multiple DNA fragments, independent of fragment length or end compatibility (New England BioLabs®Inc., Ipswich, Massachusetts, USA). Thereby, multiple overlapping DNA fragments can be joined in a single-tube isothermal reaction. The master mix contains an exonuclease, a DNA polymerase and a DNA ligase.

The sequences of the DNA fragments to be fused were amplified by PCR using Phusion DNA polymerase. Primers were design considering at least 20 bp overlapping nucleotide sequences with the respective upstream or downstream DNA fragment. Additionally, restriction sites were included at both ends of the homology domain as well as homologous sequences to the used cloning or expression vector. The vector was linearized and mixed in equal concentration with the DNA fragments to be assembled in a total volume of 10 μL. The DNA mixture was supplemented with 10 μL of Gibson Assembly Master Mix and the solution was incubated at 45 °C for 90 minutes in a thermocycler. Afterwards, 2 μL of the assembly mixture was transformed into electrocompetent *E. coli* XL1-Blue cells.

2.16.14 Blue-White-Screening

Blue-White-Screening is a useful cloning approach for screening for correct clones. Vectors used for Blue-White-Screening carry the gene for the ß-galactosidase (*lacZ*) within the multiple cloning site and this allows to use the gene as reporter gene. After insertion of a gene into the multiple cloning site, the ß-galactosidase is inactivated and can no longer cleave the yellow dye X-Gal which is part of the growth medium into the blue dye (5,5'-dibromo-4,4'-dichloro-indigo) and galactose. However, host strains with an inactivated ß-galactosidase stay uncolored and can be identified and isolated.

Blue-White-Screening requires the application of mutant strains not expressing any further gene for the ß-galactosidase. In this work, *E. coli* XL1-Blue was used as host strain, which is not expressing any further gene for the β-galactosidase, and transformed with 1-2 μL of the respective plasmid. After regeneration, cells were plated on LB agar plates. Those plates contained the appropriate antibiotic and were overlayed with 40 µl of 100 mM IPTG and 120 µl of 20 mg/ml X-Gal (dissolved in DMSO). After incubation at 37 °C over night, white clones could be distinguished from blue clones and used for further screening.

2.16.15 PCR targeting: Red/ET-mediated recombination in *E. coli*

Gene knock-outs and introduction of promotor cassettes were achieved by PCR-targeted gene replacement (REDIRECT© technology kit for PCR targeting). Red/ET- mediated recombination was applied to introduce linear DNA fragments into the respective cosmids. This approach takes advantage of the *λ*-Red (*exo, bet, gam*) proteins that promote a greatly enhanced rate of recombination with linear DNA in the targeted strain. Thereby a DNA sequence is replaced by a PCR-generated selectable marker, which is flanked with homology arms (39 nucleotides) to the DNA region intended to be modified. An origin of transfer in the resistance cassette enables conjugation of the processed construct into *Streptomyces* strains.

The required fragments were amplified by PCR using primers, which are equipped with 39 nucleotides matching the sequence adjacent to the gene to be inactivated and at their 3' end matching the right or left end of the apramycin disruption cassette of vector pIJ773, respectively (19 nucleotides or 20 nucleotides).The pipetting and cycle conditions are displayed in Table 34 and Table 35.

Table 34: Reaction mix for PCR targeting

Component	Amount	Specification
Taq Buffer (10x)	5 µL	1X
Template DNA (100 ng/µL)	0.5 µL	50 ng (0.06 pmol)
Primer FW (100 pmol/µL)	0.5 µL	50 pmol
Primer RV (100 pmol/µL)	0.5 µL	50 pmol
dNTPs (2.5 mM each)	4 µL	50 µM each
Taq:pfu polymerase (9:1)	1 µL	
DMSO (100 %)	2.5 µL	5 %
ddH_2O	36 µL	

Table 35: Cycle conditions for PCR targeting

	Temperature [°C]	Length [s]	Cycles
Hotstart	94 °C	120	1
Denaturation	94 °C	45	10
Annealing	50 °C	45	
Elongation	72 °C	90	
Denaturation	94 °C	45	15
Annealing	55 °C	45	
Elongation	72 °C	90	
Terminal elongation	72 °C	300	1
Cooling	4 °C	∞	

PCR products were purified as described in 2.16.5 and finally eluted from the column with 12 μL of ddH_2O.

In order to express the λ-RED genes, the plasmid pKD46 had to be introduced into *E. coli* BW25113 via electroporation (2.16.10). The recombination plasmid pKD46 has a temperature-sensitive origin of replication and therefore *E. coli* BW25113 harboring this plasmid has to be grown at 30 °C as long as the plasmid is required. In a second transformation, the construct that contains the target sequence that should be altered is introduced into *E. coli* BW25113/pKD46. A positive clone is cultured in 10 mL SOB medium, supplemented with the appropriate antibiotics and 10 mM *L*-arabinose to induce the *red* genes (30 °C, 200 rpm). At an OD_{600} of around 0.6, cells were harvested by centrifugation and prepared for electroporation by washing twice with 10 mL of ice cold 10 % glycerol. After the final washing step, glycerol was decanted and 50 μL of the remaining cell suspension was mixed with 2-4 μL of PCR product (extended resistance cassette). Electroporation was carried out as described in 2.16.10). After overnight incubation at 37 °C or 30 °C if further targeting was planned, single colonies were picked and inoculated for liquid overnight cultures for mini preparations (2.16.2). The isolated DNA was verified by restriction analysis, test PCRs and finally sequenced to confirm the correct PCR targeting.

Excision of the apramycin resistance marker located between the two FRT sites, flanking the resistance cassette from pIJ773, was achieved via use of a FLP-recombinase. To this end, *E.*

coli BT340 was transformed with the respective knock-out construct and incubated for two days at 30 °C on solid LB medium containing the appropriate antibiotics. Single colonies were picked, streaked again on LB medium without antibiotics and incubated at 42 °C over night (temperature sensitive expression of the flippase). Apramycin sensitive clones were screened via colony PCR and successful excision of the resistance cassette was verified via restriction analysis and sequencing.

2.16.16 Triparental intergeneric conjugation of DNA from *E. coli* to *Streptomyces*

Intergeneric conjugation was performed for the transfer of modified DNA into *Actionmycetes* strains.[185] In this study fosmids were transferred for gene cluster knock-out mutant strains or heterologous expression of BGCs. First of all, the fosmid (containing an *ori*T) was introduced into the non-methylating strain *E. coli* ET12567, and selected for ET12567 with chloramphenicol and the incoming fosmid with kanamycin. As preperation for triparental mating an overnight culture of ET12567 containing the fosmid and the strain ET12567/puB307, harbouring the *tra*-genes required for mobilization and transfer of circular DNA, was inoculated. Of these cultures, 200 µl were inoculated into 10 ml fresh LB plus antibiotics as above and grow for ~ 3 h at 37 °C to an OD_{600} of 0.4 - 0.6. The cell were washed twice with 10 ml of LB to remove antibiotics that might inhibit the growth of actinomycete host strains, and resuspened in 1 ml of LB. Meanwhile, for each conjugation 10 µl (~10^8) *Streptomyces* spores were added to 500 µl 2 × YT broth. Heat shock at 50 °C for 10 min and then cooled by leaving at room temperature for 15 minutes. Next, 0.25 ml of each *E. coli* cell suspension (*E. coli* ET12567 + fosmid and *E. coli* ET12567 pUB307) were mixed to 0.5 ml heat-shocked spores and spun briefly. The supernatant was discarded and the pellet was resuspended in the 50 µl residual liquid. A dilution series from 10^{-1} to 10^{-4} each step in a total of 100 µl of water was prepared and then 100 µl of each dilution was plated out on MS agar + 10mM $MgCl_2$ (without antibiotics) and incubate at 30 °C for 16-20 h. The next day, the plates were overlayed with 1 ml water containing 0.5 mg nalidixic acid (20 µl of 25 mg/ml stock; selectively kills *E. coli*) and 1.25 mg kanamycin (25 µl of 50 mg/ml stock) and further incubated at 30 °C. When single colonies appeared after 1 to 3 weeks, they were stroke out on MS agar plates for confluent growth until appearance of spores. Spores were harvested as described in 2.15.3. These spore were used for a first test cultivation, isolation of genomic DNA and verification by restticition analysis and sequencing. Further selection was made by

dilution series and plating on MS Agar. If markerless mutants were generated and selection on antibiotic markers was not possible, screening for double cross over mutants was speeded up by directly separating and streaking out single sporulating colonies to replica plates, for multiple rounds.

2.16.17 Sequencing

Sequencing was performed with Eurofins Genomics sequencing service (Ebersberg, Germany). For sequencing, DNA samples were diluted with ddH_2O to the required DNA concentration of 100 ng/μL (plasmids), 1000 ng/μL (cosmids) or 5 ng/μL (purified PCR products in the range of 300-1000 bp) in a final volume of 15 μL and, if necessary, were supplemented with 2 μL of a specifically designed sequencing primer. The obtained sequences were analyzed using Clone manager or BLAST alignment tool.

2.17 Construction of heterologous expression and *Streptomyces* mutant strains

2.17.1 Heterologous expression of the *bel* gene cluster

For heterologous expression of the belactosin gene cluster in *S. albus* J1074, first, an int_neo cassette was purified from merLK01 by XbaI restriction digestion. [178] The int_neo cassette encodes for an origin of transfer (oriT), a neo resistance marker for selection and the phage ϕC31 components, namely an attachment site (attP) and integrase (int), that allow stable and site-specific chromosomal integration in the host strain.[190] The purified int_neo cassette was then transferred into *E. coli* BW25113/pKD46 containing the fosmid 6D1 (fosmid backbone pCC1FOS, Lucigen, Middleton, Wisconsin, USA), which carries the complete bel gene cluster. [77] This fosmid was then refurbished by replacing the chloramphenicol resistance cassette with int_neo enabled by λ-Red-mediated recombination.[182] The resulting fosmid belFW01 was verified by restriction analysis and subsequently transferred into *E. coli* ET12567 for ongoing triparental conjugation. To yield strain *S. albus* J1074 bel01, the fosmid was introduced in to the host strain *S. albus* J1074 by triparental mating with the help of *E. coli* ET12567/pUB307 as described in 2.16.16.[185] After an incubation period of 5-10 days at 30 °C, kanamycin resistant (Kan^R) positive exconjugants were selected. Genomic DNA of several *S. albus* J1074 bel01 clones was prepared and screened by PCR and sequencing with

primer pairs belUS_FP/RP and belDS_FP/RP (Table 16). To evaluate the heterologous expression of belactosins in the verified heterologous mutants, 10 µL of *S. albus* J1074 spore suspensions (WT and three clones with intact cluster) were inoculated in 50 mL TSB. After cultivation for 2 days at 30 °C and 200 rpm with appropriate antibiotics, 2 % of the TSB culture were transferred to 50 ml R5 production medium (see 2.8.2) and were cultivated for 5 days at 30 °C and 200 rpm. Subsequently the cultures were extracted and analyzed by LC-MS as described in 2.23.1.

2.17.2 Generation of individual *belK belL and belN* gene knock-outs in the heterologous producer

The *belK, belL,* and *belN* gene knock-outs in the heterologous host were also generated using a PCR targeting system as described in 2.16.15.[183] To this end, an apramycin resistance cassette (Apra^R) was isolated from plasmid pIJ773 by XbaI restriction digestion. The 773 Apra resistance cassette was elongated with primer pairs KOBelK_FP/RP, KOBelL_FP/RP, KOBelN_FP/RP, respectively (Table 16). The PCR products were purified and transferred into *E. coli* BW25113/pKD46 containing belFW01. Target regions were replaced by the Apra^R resistance cassette with the help of λ-Red-mediated recombination.[182] Restriction analysis confirmed the resulting fosmids belFeW02, belFeW03 and belFeW04 for knock-out of *belK, belL and belN*, respectively. The fosmids were then conjugated into *S. albus* J1074 with the help of *E. coli* ET12567/pUB307, as described above and in 2.16.16. Apra^R/Kan^R resistant exconjugants were selected and three clones of the resulting mutant strains *S. albus* J1074 bel01/Δ*belK, S. albus* J1074 bel01/Δ*belL and S. albus* J1074 bel01/Δ*belN* were further verified by PCR and sequencing with screening primers screenBelK_FP/RP, screenBelL_FP/RP and screenBelN_FP/RP (Table 16). In the next step, cultures of those three individual knock-out strains were extracted and analyzed by LC-MS for the production of belactosins, as described in 2.23.1.

2.17.3 Generation of gene cluster border knock-outs in *S. albus* J1074

The gene knock-outs for analysis of the gene cluster borders were generated using the PCR targeting system as described above and in 2.16.15. The 773 Apra resistance cassette was elongated with primer pairs KO_orf-1_FP/RP and KOBelA-E_FP/RP for upstream cluster border analysis. PCR products for targeting the down stream cluster borders were generated by primer pairs KOBelO-V_FP/RP, KOBelQ-V_FP/RP, KOBelR-V_FP/RP and KOBelO-T_FP/RP. The PCR products were purified and transferred into *E. coli* BW25113/pKD46

containing belFeW01. . The apramycin resistance cassette in all mutant fosmids was removed in *E. coli* BT340 by taking advantage of the flanking FRT site, leaving behind a 81 bp "scar" sequence in the preferred reading frame and lacking stop codons. To ensure in-frame deletion of the respective genes, the scar sequence was verified by PCR and sequencing with the help of the primer pairs screen_orf-1_FP/RP, screenBelA-E_FP/RP, screenBelO-V_FP/RP, screenBelQ-V_FP/RP, screenBelR-V_FP/RP and screenBelT-V_FP/RP. Further sequencing of all gene cluster knock-outs confirmed the resulting fosmids belFeW13 for *Δborf-1*, belFeW14 for Δ*belA-E,* belFeW15 for Δ*belO-V,* belFeW16 for Δ*belQ-V,* belFeW17 for Δ*belR-V,* belFeW18 for Δ*belT-V*. Finally, the fosmids were then conjugated into *S. albus* J1074 with the help of *E. coli* ET12567/pUB307, as described above and in 2.16.16. Kan^R resistant exconjugants were selected. They were cultivated, screened and analyzed as describe before. Finally yielding, strains *S. albus* J1074 *bel01/Δorf-1, S. albus* J1074 *bel01/*Δ*belA-E, S. albus* J1074 *bel01/*Δ*belO-V, S. albus* J1074 *bel01/*Δ*belQ-V, S. albus* J1074 *bel01/*Δ*belR-V and S. albus* J1074 *bel01/*Δ*belT-V.*

2.17.4 Generation of unmarked in-frame deletion of *belKLMN* in *S.* sp. UCK14

To obtain double crossover mutants in the UCK14 WT producer strain, the PCR targeting system was taken advantage of, as described before. The apramycin resistance cassette was elongated using primer pairs, KOBelK_FP/RP for Δ*belK*, KOBelL_FP/RP for Δ*belL*, KOBelM_FP/RP for Δ*belM* and KOBelN_FP/RP for Δ*belN* (Table 16). The PCR products were purified and transferred into *E. coli* BW25113/pKD46 containing 6D1. After homologous recombination has taken place, restriction analysis confirmed the resulting fosmids belFeW05 for Δ*belK*, belFeW06 for Δ*belL*, belFeW07 for Δ*belM* and belFeW08 for Δ*belN*. To avoid polar effects on the expression of adjacent genes transcriptionally downstream, the apramycin resistance cassette was removed in *E. coli* BT340 by taking advantage of the flanking FRT site, leaving behind a 81 bp "scar" sequence in the preferred reading frame and lacking stop codons. To ensure in-frame deletion of the respective genes, the scar sequence was verified by PCR and sequencing with the help of the primer pair screenBelK_FP/RP, screenBelL_FP/RP, screenBelM_FP/RP and screenBelN_FP/RP (Table 16). Additionally, the chloramphenicol resistance gene was replaced by an $Apra^R$ and oriT, yielding the final fosmid constructs belFeW09 for *ΔbelK*, belFeW10 for *ΔbelL*, belFeW11 for *ΔbelM* and belFeW12 for *ΔbelN*, which were conjugated into *S.* sp. UCK 14, as described above. Exconjugants were selected for $Apra^R$, indicating a successful single crossover. Clones were tested for double crossover by replica plating and loss of $Apra^R$. Ongoing, the genomic

DNA of these clones was isolated and tested by PCR and sequencing with the screening primers mentioned before (Table 16). Finally, three individual clones of *S.* sp. UCK14/Δ*belK, S.* sp. UCK14/Δ*belL, S.* sp. UCK14/Δ*belM* and *S.* sp. UCK14/Δ*belN* were identified as successful double crossover mutants. The verified mutants were cultivated, extracted, and tested for belactosin production by LC–MS analysis, as described in 2.23.2.

2.17.5 Generation of the genetic complementation plasmid pUWL_*belN*

Cloning of the pUWL_*belN* construct was achieved using Gibson Assembly® Master Mix – Assembly (New England BioLabs® Inc., Ipswich, Massachusetts, USA). The *belN* gene was amplified by PCR using the Q5®High-Fidelity DNA polymerase. Primers pUWLbelN_FP/pUWLbelN_R include 18 bp 5´ and 3` sequences of the *belN* gene fragment, HindIII and SpeI restriction sites, as well as up- and downstream homologous 32 bp sequences to the pUWL_apra_oriT cloning vector. The pUWL_apra_oriT vector was linearized using an HindIII, SpeI restriction digest and subsequently the vector was mixed with the elongated *belN* DNA fragment in equal concentrations in a total volume of 10 µL. The DNA mixture was supplemented with 10 µL of Gibson Assembly Master Mix and incubated at 45 °C for 90 minutes in a thermocycler. Afterwards, 1 µL of the assembly mixture was transformed into electrocompetent *E. coli* XL1-Blue cells. Positive clones were selected on apramycin containing LB plates, plasmid DNA of selected mutant clones were isolated and verified by sequencing using primer pair pUWL_test_fwd/ pUWL_new_rev (Table 16). Furthermore, after conjugation of pUWL_*belN* in *S.* sp. UCK14, the genomic DNA of three selected exconjugants was isolated, introduced into *E. coli* XL1-Blue via electroporation and subsequently following plasmid preparation verified by PCR using primer pair pUWL_test_fwd/ pUWL_new_rev.

2.18 Heterologous overproduction and purification of recombinant proteins from *E. coli* for biotransformation assays

Heterologous overproduction and purification of 6×His-tagged proteins from *E. coli* were carried out using expression vector pET28a (+) and *E. coli* Rosetta2 (DE3) pLys (Novagen, Darmstadt, Germany) as the recombinant host strain. The vector provides the sequence encoding an N-terminal as well as a C-terminal His_6-tag

2.18.1 Cloning of the expression vector pET28a_glbB

glbB was ordered as a codon optimized gBlock (BioCat GmbH, Heidelberg, Germany):

DNA sequence of GlbB: (codon optimized using IDT Codon Optimization Tool as suggested by [191])

ATGCAGCTGGGTGTGCCGTTTCTGAAGCGTGAGAGCATGAGCAAAATGACCGGCCAAGAATG
GGCGGCGGCGGCGCCGGAGACCGAACCGGGTGATGTTCGTGCGGCGCTGCAGCAACGTGGCT
GGGCGCGTTTTGACGCGACCGATATGCAGGTGGCGGTTGACGAGGCGGCGGATCTGCAACGT
CTGACCGAATATGCGCGTAGCCTGCCGGTGGACCGTTTTGGTACCGGTGGCCGTCACCGTAGC
TATGCGGAGGGCATCCTGACCCCGCGTCGTAAGACCATTGCGTGGAAGGCGGGTGCGCGTAC
CCCGGATGGTCGTGTGGAAATCGCGTACGTTCAGCACAGCGAGTTCCAACCGGAACACGGTG
GCGTGGTTCGTAACTTTGCGCGTACCCGTGAGGACATCCTGGCGCTGCCGCTGGTTCACCGTC
TGATTTGGTATGATCTGAGCCTGACCCCGATGTTCGACGCGGAAGGTGATCTGCTGTGCGGCT
TTCACATGATTCGTATGCAGGCGACCCCGGGTGCGGTGGCGCGTATTACCCCGGACTGCCTGC
ACCAGGATGGCCAACCGTTTACCGCGGTGCACCTGGTTGAGCGTAGCCACGCGGAAGGTGGC
GTTAACTTCATCGCGCCGCCGCGTTACACCGGTCGTCAATTTGATGAGGTGCCGAGCCACCTG
CTGAGCGCGTTCGTTCTGGGTAGCCCGCTGCAGAGCTATATCATTGACGATGCGGCGATTTGC
CACCAAGTTACCGCGGTTAGCTGCAGCCCGGGTGCGAGCCATGGTACCCGTACCGTGATCCTG
ATTGATTTCAGCCCGCTGAACCCGGCGAGCAGCGCGCAGCCGACCTGA

For the construction of expression vector, the above sequence was inserted between NdeI on the 5' restriction site and XhoI on the 3'restriction sites within the commercial pET28a(+) vector. The resulting plasmid pET28a_glbB was confirmed via sequencing and transferred into *E. coli* Rosetta2 (DE3) pLysS via electroporation. The cells containing pET28a_glbB were selected by their kanamycin resistance.

2.18.2 Cloning of the expression vector pET28a_belL

The *belL* gene was amplified by PCR from genomic DNA of S. sp. UCK14 with primers pET28a_belL_EcoRI_f and pET28a_belL_HindIII_r (Table 16). The 689 bp PCR product was subcloned via HindIII and EcoRI restriction sites into the pBluescript (SK-) cloning vector (Promega, Mannheim, Germany) first. Candidate clones were selected by Blue-White-Screening (2.16.14) and the constructed plasmid was verified by restriction analysis and sequencing. After expression of pBlueacript_belL in *E. coli* XL1 blue and subsequent plasmid preparation, belL was cloned in the expression vector pET28a (+) taking advantage of the HindIII/EcoRI restriction sites. The resulting plasmid pET28a_belL was confirmed via

sequencing and transferred into *E. coli* Rosetta (DE3) pLysS via electroporation. The cells containing pET28a-belL were selected by their kanamycin resistance.

2.18.3 Recombinant protein production of GlbB and BelL in *E. coli*

Different conditions for cultivation and purification were tested (e.g. different cultivation temperatures, amount of IPTG used for induction). The purification method described below yielded the best results.

The respective *E. coli* Rosetta2 (DE3) pLysS strains containing pET28a-glbB or pET28a_belL were cultivated in 50 mL of over-night culture and transferred into 500 ml and up to 1 L TB broth, both supplemented with 25 µg/mL chloramphenicol and 50 µg/mL kanamycin at 37 °C and 220 rpm. At OD_{600} 0.6, the temperature was adjusted to 20 °C and isopropylgalactoside (IPTG) was added to a final concentration of 0.5 mM. After additional cultivation over 16 h the culture was harvested (4000 rpm, 20 min, 4 °C) and 2.5 ml per gram cells of lysis buffer containing 0.5 mg/mL lysozyme, 0.5 mM PMSF and 1 mM DTT were added. Cells were disrupted by sonification (Sonifier Generator, Branson) or French press (SimAminc). For sonifictaion On the French press cells were lysed 3-5 times at 1000 psi. The lysate was used for the further purification of the enzyme.

2.18.4 Purification of GlbB and BelL by Ni^{2+} affinity chromatography

The cell lysate was centrifuged at 18000 rpm for 45 min and then the supernatant was filtered through a 0.45 µm filter. The flow through was applied to affinity chromatography using an ÄKTAstart™ platform (GE Healthcare) equipped with a 5 mL His-Trap™ HP column (GE Healthcare). The His-tagged protein was eluted from the column using a linear gradient from 0-100 % of elution buffer over 45 to 60 min and collected by a Frac-30 system (GE Healthcare). Fractions were tested for the presence of the respective proteins by SDS-PAGE. GlbB and BelL containing fractions were concentrated using an Amicon Ultra centrifugal filter (*M*r cutoff of 10,000; Millipore). Concentrations of the purified proteins were measured with a Spectrophotometer at 280 nm with UV vis spectrometer spectraMax M2 (Molecular Devices) and calculated using associated extinction coefficients. M^{-1} cm^{-1}, respectively (calculated with http://web.expasy.org/protparam/) [192].The purified proteins were stored in elution buffer in aliquots at −80 °C. For long time storage proteins were stored in buffer A, 20 % glycerol and adapted salt concentrations, as GlbB as well as BelL have a tendency to precipitate under described conditions.

2.18.5 Denaturing Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The discontinuous SDS-PAGE was carried out according to the method of [193] with small adjustments.

Samples were mixed with 6x Laemmli sample buffer in ratio 1:5 and denatured at 95 °C for 10 min. After briefly spinning the probes down, 10 µl were subsequently applied to the SDS-Gel. The separation gel consisted of 12.5 % polyacrylamide and 4 % acrylamide in the stacking gel. The first and last gel pocket on the SDS-gel was loaded with 5 µl of protein marker (Color Prestained Protein Standard, Broad Range P7712, New England Biolabs) to determine the approximate molecular mass of the proteins. Gel electrophoresis was carried out with a constant voltage of 150 mA in 1X SDS running buffer, using a Mini- PROTEAN® II Electrophoresis Cell (Bio-Rad, München, Germany). After the gel was completely run, the separated proteins were visualized with the Coomassie staining solution for 30min. Gels were then destained with the Coomassie destainig solution for about 1 hour or until good visualization of the protein bands was achieved.

2.18.6 Generation of 4-OH-Lys with GlbB and BelL in a biotransformation enzyme assay

In order to generate 4-OH-lys for feeding studies or protein characterization of belL an in vitro enzyme assay was performed. In a first step 50 mM kPi buffer was added to a vial (pH 8.0) according to the desired volume, molarity and given protein stock concentration, followed by 15 µL of FeSO4 solution in H2O (1 mM final concentration). Depending on the reaction volume a 5 ml for BelL assays, 20 ml or 50 ml for GlbB assays, scintillation vial were used and charged with L-lysine (20 mM final concentration), L-ascorbic acid (10 mM final concentration), and α-ketoglutaric acid (50 mM final concentration). All assay components, despite the protein stock solution, were dissolved in ddH_20. Beforehand GlB and BelL were buffer exchanged to 50 mM kPi buffer with centrifugal filter units after manufacturer recommendations (Amicon Ultra-4, 10K cut off). The reaction was started by the addition of GlbB or BelL stock solution. The final concentration was 10 µM for 4-OH-Lys preparation for feeding studies and 5, 10 or 20 µM for BelL assays. For subsequent feeding studies within UCK14 mutants, stable isotope labeled $^{13}C_6$- L-lysine (Silantes GmbH, Munich, Germany) was used to be able to detect successful incorporation into the belactosin molecule.

The assays were shaken for 12 h at 20 ºC, 250 rpm under air. After incubation time, the reaction mixture was centrifuged (13,000 rpm, 10 min) and the supernatant was submitted directly to LC-MS measurement. For further feeding studies the reaction mixture was stored at -20 °C or at 4 °C over night.

2.19 Chemical synthesis of labeled precursor

2.19.1 Chemical synthesis of 3,3-d_2-3-(trans-2'-nitrocyclopropyl)alanine (3,3-d_2-Ncpa) and of 3,3-d_2-3-(trans-2'-aminocyclo-propyl)alanine (3,3-d_2-Acpa)

The double deuterium labeled Ncpa and Acpa intermediates were kindly provided by our collaboration partners from the working group of Prof. Dr. Armin de-Meijere (Universität Göttingen). Twofold deuterium-labeled racemic 3,3-d_2-3-(*trans*-2'-nitrocyclopropyl)alanine (3,3-d_2-Ncpa)[174] and 3,3-d_2-3-(*trans*-2'-aminocyclopropyl)alanine (3,3-d_2-Acpa)[173, 174] were prepared according to the previously published protocols.

2.20 Chemical synthesis of the 5,5-d_2-4-hydroxylysine lactone

Chemical synthesis of the 5,5-d_2-4-hydroxylysine lactone was carried out by Chambers Hughes and Annika Esch based on the chemical synthesis protocol byMelanie Brandl et al.[173]

2.21 Chemical feeding and complementation studies

2.21.1 Feeding of $^{13}C_5$-L-ornithine, $^{13}C_6$- L-lysine, d_2-Ncpa and d_2-Acpa

For feeding studies, stable isotope labeled-precursor molecules, as well as deuterium labeled intermediates of the belactosin biosynthesis pathway were fed to the wild type belactosin producer strain *S.* sp. UCK14 and the *S.* sp. UCK14/Δ*belK, S.* sp. UCK14/Δ*belL, S.* sp. UCK14/Δ*belN* mutants. To this end cultures were grown on SM19 production medium agar plates as described above and cultures were supplemented with $^{13}C_5$-L-ornithine (Merck KgaA, Darmstadt, Germany), $^{13}C_6$- L-lysine (Silantes GmbH, Munich, Germany), double deuterium-labeled racemic 3-(*trans*-2'--nitrocyclopropyl)-alanine (d_2-Ncpa) and double deuterium-labeled racemic 3-(*trans*-2'-aminocyclopropyl)alanine (d_2-Acpa) in a final concentration of 2.5 mM. Therefore labeled precursors were dissolved in MeOH at a final concentration of 250 mM. 100 µl of the stock were spread on 60 x 15 mm petri dishes

(Sarstedt, Newton, North Carolina, USA), containing 10 ml SM19 agar and inoculated with 10 µl spore solution. After incubation of 5 days at 30 °C, cultures were extracted as described in 2.23.2 and subsequently analyzed with LC-MS.

2.21.2 Feeding of $^{13}C_6$-4-hydroxy-lysine generated by GlbB

In order to generate $^{13}C_6$-4-hydroxy-lysine in sufficient amounts for feeding and chemical complementation studies it was prepared in an in vitro assay as described in 2.18.6. The raction volume was scaled up to 10 ml in a 50 ml scintillation vial. After incubation time a sample of the assay was analsed by HILIC-tandem Mass Spectrometry as described in 2.23.7. The complete assay mixture was freezed at -80 °C for an hour and freezed dried over night. Subsequently the lyophilized sample was dissolved in 200 µl ddH_2O. 100 µl were spread over an SM19 agar plat, air dried and inoculated with spore suspension as described in 2.21.1. After incubation of 5 days at 30 °C, cultures were extracted as described in 2.23.2 and subsequently analyzed with LC-MS.

2.21.3 Feeding of the lactone

The lactone was obtained as a freeze-dried powder and dissolved into a approximately 250 mM ddH_2O solution. As the lactone was obtained as a racemic mix, we aimed calculated weighed compound accordingly. At the end, as described before .100 µl were spread over an SM19 agar plat, air dried and inoculated with spore suspension. After incubation of 5 days at 30 °C, cultures were extracted and subsequently analyzed with LC-MS.

2.22 Methods for identification and isolation of a new compound from *Nocardia cyriacigeorgica* GUH-2

2.22.1 Isolation of genomic DNA from *Nocardia* spp.

For isolation of genomic DNA a 50 ml BHI culture was harvested by centrifugation, 10 min 400o rpm. The pellet was then wased in TSE buffer twice (5 min, 4000 rpm). In the next step the cell pellet was solved in 10 ml lysozyme solution an incubated at 37 °C for at least one hour up to 4 hours and cell lysis becomes visible. Then 8 ml of a 10 % SDS solution was added, carefully inverted a couple of times and incubated at 60 °C for at least 30 min and up to 2 hours. For further purification 1 volume of phenol/chloroform/isoamyl alcohol (25:24:1) was added, invert carefully and centrifugation at 4000 rpm, 10 min. The upper aqueous phase is transferred to a new vessel. This step was repeated until the aqueous phase appears clear. Then 0.1 volumes NaAcetate (3 M) solution and 0.8 volume of ice cold isopropanol were added and centrifugated at 4 °C, 4000 rpm, 30 min. The DNA pellet was washed with 70 % Ethanol (0.5 volumes) twice. The DNA pellet was then air dried at RT and dissolved in ddH_2O

2.22.2 Construction of a fosmid based genomic library

The fosmid library for the genomic library of *N. cyriacigeorgica* GUH-2 was constructed using the CopyControl^{TM} Fosmid Library Production Kit (Epicentre^{R}, now distributed by Lucigen) following the manufacturer's instruction with minor modifications.[194]

In a first step the genomic DNA of strain *N. cyriacigeorgica* GUH-2 was sheared by pipetting up and down with a 200 µl tip for about 30 times to generate about 40 kb DNA fragments. Those DNA fragments were end-repaired by generating blunt-ended, 5'-phosphorylated DNA following the kits instructions. Subsequently, the end-repaired DNA was directly ligated into the cloning ready pCC1FOS^{TM} vector at a ratio of 10:1 (vector:insert). The ligation reaction was carried out for 2 hours. The ligated DNA was then packaged using the MaxPlax Lambda packaging extracts, mixed with phage T1-resistant Epi300 cells and plated out on LB-medium containing chloramphenicol (12.5 µL/mL). After calculating the titer, the cells were plated out on 250 ml agar plates and in total 2016 clones were picked and cultivated in 96er well plates containing LB medium and chloramphenicol (12.5 µL/mL). This number of clones ensures that the complete *N. cyriacigeorgica* GUH-2 genome is represented in the library. The plates

were then incubated at 37 °C for 16-20 h. Following incubation, the clones were preserved with glycerol and stored at -80 °C.

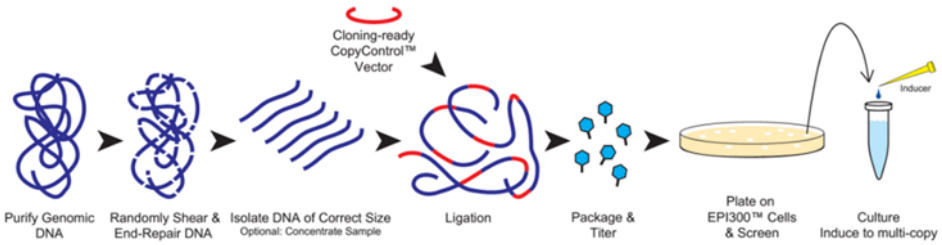


Figure 10: Construction of a pCC1FOS based genomic library: Purified genomic DNA was sheared, blunt end repaired and optional size separated. Followed by ligation of ~40kb fragment into a CopyControlTM vector, and packaged into a phage. EPI300^{TM} cells were infected with the library and selected on LB Chloramphenicol (12,5 µg/ml) overnight at 37 °C. (Figure adapted from CopyControl™Fosmid Library Production Kit Manual [194])

2.22.3 Screening of the genomic library

To screen the fosmid library for positive clones that contain the gene cluster of interest, the 96well plates were plated out on LB-agar plates. To transfer bacterial colonies from 96-well plates a stainless steel replica plater (48 circular prongs) was used. Following, every 96-well plate was pooled into two 48 pools, resulting in an A and B pool for every plate. The fosmids of this DNA pools were isolated (2.16.2). Following, a PCR with all 48 DNA pools was performed to identify the gene cluster of interest, using primer pairs ACAD-cluster-left_f/r and ACAD-cluster-right_f/r (Table 17) to detect the borders up- and down-stream of the ACAD cluster. Furthermore, a PCR was performed to screen for the ACAD gene within the cluster using primer pair screen_ACAD_GUH-2_f/r. If a pool was tested positive, a colony PCR (2.16.12) was performed with the respective plate, to determine which exact well was containing the positive clone. This clone was then analyzed with restriction digestion and end-sequencing using primer pair pCC1FOS_FP/RP (Table 17) for final determination and review.

2.22.4 Heterologous expression of the ACAD biosynthetic gene cluster

For heterologous expression of the ACAD gene cluster *from N. cyaricigeorgica* GUH-2, first, an int_neo cassette was purified from fosmid merLK01 by XbaI restriction digestion. [178] The purified int_neo cassette was then transferred into *E. coli* BW25113/pKD46 containing the fosmid 21C6, which carries the complete ACAD gene cluster. This fosmid was then

refurbished by replacing the chloramphenicol resistance cassette with int_neo enabled by λ-Red-mediated recombination The resulting fosmid guhAE01 was verified by restriction analysis and subsequently transferred into *E. coli* ET12567 for ongoing conjugation. The fosmid guhAE01 was introduced in to the host strain *S. albus* J1074 and *Amycolatopsis japonicum* M6417-CF17 by triparental mating with the help of *E. coli* ET12567/pUB307 as described in 2.16.16. After an incubation period of 5-10 days at 30 °C, kanamycin resistant (Kan^R) positive exconjugants were selected. Genomic DNA of several clones was prepared and screened by PCR and sequenced with primer pairs ACAD-cluster-left_f/r and ACAD-cluster-right_f/r (Table 17). Yielding heterologous expression strains *S. albus* J1074/ ACAD GUH-2 and *Amycolatopsis japonicum* MG417-CF17/ ACAD GUH-2. To evaluate if a putative novel ACAD compound is heterologously expressed, the verified heterologous mutants, were inoculated in 50 mL TSB. After cultivation for 2 days at 30 °C and 200 rpm with appropriate antibiotics, 2 % of the TSB culture were transferred to 50 ml R5 production medium (see 2.8.2) and were cultivated for 5 days at 30 °C and 200 rpm. Subsequently the cultures were extracted and analyzed by LC-MS as described in 2.23.1.

2.22.5 Purification of the ACAD compound

In order to produce the novel ACAD GUH-2 compound in higher amounts and consequently purify the compound. The heterologous expression system *S. albus* J1074/ ACAD GUH-2 was up scaled in a first attempt from 50 ml standard cultivation in R5 medium (2.8.2) to a 2 L culture volume in 5 L baffled Erlenmeyer flasks equipped with a metal spiral. Thereby 2 L of R5 production media were inoculated with 10 ml of a TSB pre culture. In total 12 L of culture were grown at 30 °C, 220 rpm for 5 days. In a next step, the whole culture broth was extracted with one volume of EtoAc and evaporated under reduced pressure. The gained crude extract was weighed and applied to vacuum liquid chromatography as described in 2.23.3, using C-18 RP polygoprep stationary phase and the different MeOH gradient steps as described in Table 36. The VLC fractions were measured in LC-MS. In most cases a second round of VLC was conducted to reach better purity. ACAD containing fractions were further purified using preparative HPLC as described in 2.23.5 using a gradient as depicted in Table 38. In a next step fractions were finally purified over a semi preparative column on a analytic HPLC device as described in 2.23.4 using gradient as described in Table 37. Fractions containing compound with good purity were finally analyzed in NMR. The general applied workflow for purification of the ACAD GUH-2 compound is described in Figure 11.

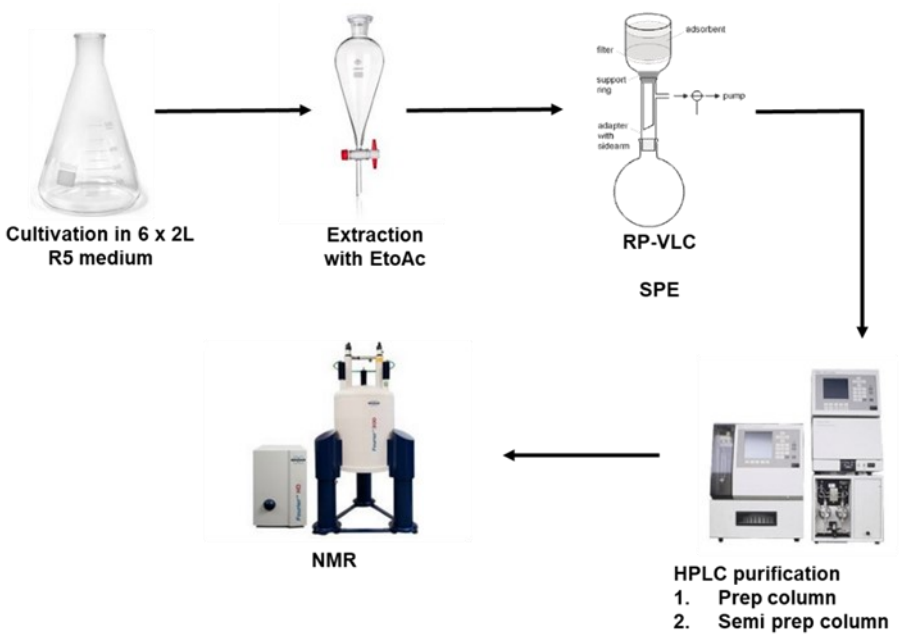


Figure 11: General workflow for the purification of the ACAD compound:

2.23 Methods of chemistry and analytical methods

2.23.1 Extraction of liquid bacterial cultures with solvent

After the appropriate cultivation time liquid cultures were extracted three times with ethyl acetate, methanol or butanol in a ratio of 1:1 using a separating funnel or shaking in an erlenmeyer flask. The aqueous phase was discarded after successful separation of phases or with the help of centrifugation at 4000 rpm for 10 min in 50 ml falcon tubes. The organic phase was dried using a rotary evaporator (Flawil, Schweiz).

2.23.2 Extraction of bacterial cultures on agar plates with solvent

To extract compounds from bacterial cultures grown on agar plates, 3 ml butanol were added directly to the plate and mixed with the help of a cotton swap. The suspension was centrifuged in 2 ml reaction tubes at 13 0000 rpm for 10 min and the supernatant was dried using a rotary evaporator.

2.23.3 Vacuum liquid chromatography (VLC)

The vacuum liquid chromatography was carried out with a column with a diameter of 65 mm, which was filled with Polygoprep 60-50- C18 material. The filling level of the stationary phase making up about half to two thirds of the column. It was then first equilibrated with the appropriate methanol-water mixture, which should be used as the starting condition.

The crude extract was dissolved in a small volume of the starting methanol-water mixture and applied to the column bed. Elution then takes place with increasing proportions of solvent under applied vacuum, if necessary. As a standard the following gradient, as shown in Table 36, was applied and adjusted to specific compound characteristics if necessary.

Table 36: Gradient fractions for VLC

Fraction	Solvent	Mixing ratio (v/v)
1	MeOH/H_2O	10/90
2	MeOH/H_2O	30/79
3	MeOH/H_2O	50/50
4	MeOH/H_2O	60/40
5	MeOH/H_2O	70/30
6	MeOH/H_2O	90/10
7	MeOH	100
8	CH_2Cl_2	100

2.23.4 HPLC analysis and purification of compounds

Standard analytic HPLC (Agilent 1200 series; Waldbronn, Germany) analysis was performed using a Reprospher 100 C18, 100 Å (5 µm, 250 × 4 mm) column (Dr. Maisch, Ammerbuch, Germany) at a flow rate of 0.5-1 mL/min.

Semi-preparative HPLC for compound purification and isolation of the ACAD compound was carried out on a Luna 5 µM C18 (2), 100 Å, 250 x 10 mm column (Phenomenex Inc., Torrance, California, USA) at a flowrate of 3 ml/min. A linear gradient as described in Table 37 was used (A is 0.1% formic acid in water and B is 0.06% formic acid in acetonitrile). Detection was carried out at 220 and 254 nm. Fractions were collected automatically with the

connected fraction collector (for ACAD compound retention time min 6 to 15 min every 0.25 min).

Table 37: Gradient Semi-Prep HPLC ACAD compound

Time (min)	% B	Flow (ml/min)
0	40	3
35	100	3
37	100	3
40	40	3

2.23.5 Preparative HPLC

In order to isolate higher quantities of pure compound for structure elucidation preparative HPLC was applied on an Unisol C18(2), 5 µM, 110, 21.2x250mm column (Bonna-Agela Technologies) on a waters instrument (see 2.1). A linear gradient as described in Table 38 was used (A is 0.1% formic acid in water and B is 0.06% formic acid in acetonitrile). Samples of 2 ml were injected manually with a gastight syringe and fractions were collected manually as well. Detection was carried out at 210 nm.

Table 38: Gradient: Prep HPLC ACAD compound

Time (min)	% B	Flow (ml/min)
0	30	10
30	100	10
40	100	10
42	30	10

2.23.6 HPLC-ESI/MS and MS/MS analysis

For routine LC-MS and MS/MS analysis of culture extracts, 2.5 µL of sample were injected on to a Nucleosil 100 C18 3 µm, 100 × 2 mm column fitted with a pre-column 10 × 2 mm at a flow rate of 0,4 mL/min and a linear gradient of t_0 = 0 % solvent B (solvent B: acetonitrile (ACN) with 0,6% formic acid; solvent A: water with 0,1% formic acid) to t_{15} = 100 % B. HPLC-ESI-MS measurements were performed on a LC-MSD Ultra Trap System XCT 6330, which was coupled to a Agilent 1200 series HPLC (Agilent Technologies, Waldbronn). MS

analysis was performed by ESI (positive and negative ionization) in Ultra Scan mode with a capillary voltage of 3.5 kV and drying gas temperature of 350 °C.

2.23.7 HILIC-tandem Mass Spectrometry for analysis of 4-hydroxy-lysine

For detection of L-lysine and 4-OH- L-lysine a SeQuant ZIC-HILIC column (100 x 4.6 mm, 5 µm, 200 Å) as a stationary phase with a flow rate of 1.0 mL/min and buffer A (9:1 Acetonitrl: 50 mM NH_4HCO_2 pH 3.2 (so 5 mM NH_4HCO_2 effective concentration) and buffer B: 5:4:1 Acetonitril:water:50 mM NH_4HCO_2 pH 3.2 (so 5 mM NH_4HCO_2 effective concentration) as mobile phase was used. The method was chamstddualHILIClysine2, which is 50-100% B over 15 min.

2.23.8 High resolution HPLC-ESI/MS and MS/MS analysis

High-resolution LC-ESI/MS and MS^2 measurements were carried out on a Bruker Daltonics MaXis 4G connected to a Thermo Scientific Ultimate 3000 system using a reversed-phase Reprosil 100 C18 column (3µm, 100 mm x 3 mm) at a flow rate of 0.5 mL/min.

For separation of culture extracts a linear gradient from 10 to 100 % of solvent B for 35 min was used (solvent A: 0.1 % formic acid in water; solvent B: 0.06 % formic acid in methanol). The injection volume was 5 µl.

The acquisition parameters for the positive ion polarity were a capillary voltage of 4.5 kV and an end plate offset of -500 V. The acquisition parameters for the negative ion polarity were a capillary voltage of 3.2 kV and an end plate offset of -500 V. The nebulizer pressure was set to 2.0 bar and dry gas flow to 8.0 L/min at a dry heater temperature of 200 °C. The measurements were internally calibrated using sodium formate as a reference.

2.23.9 NMR methods and structural characterization

Methods for NMR analysis were developed in cooperation with Harald Groß and Irina Helmle. The analysis was routinely carried out by Irina Helmle.

For the NMR analysis of the ACAD GUH-2 compound, 1,34 mg pure compound were dissolved in 400 µl MeOH-d_4 and filled into an NMR vial.

NMR spectra were recorded on a BRUKER Avance III HD 400 MHz NanoBay NMR spectrometer (^{1}H, 400 MHz; ^{13}C, 100 MHz) at 293 K. Chemical shifts were determined relative to the solvent as internal standard (CD3OD, δH/δC 3.35/49.0).

All ^{13}C NMR spectra of ^{13}C enriched versions of conexipyrone A were recorded by using inverse-gated decoupling (zgig pulse program (Bruker); d1=2 s).

UV spectra were obtained in methanol on a Lambda 25 UV-Visible Spectrophotometer (Perkin Elmer). Infrared spectra were recorded on a Jasco FT/IR-4100 spectrometer. Optical rotation values were determined with a Jasco P-2000 Polarimeter (d = 1 cm).

2.23.10 Derivatization of *d_2*-Ncpa, $^{13}C_6$-Ncpa and Acpa with dansyl chloride

To an aqueous solution of the extract (200 µL, ~2 mg mL^{-1}) was added an aqueous solution of $NaHCO_3$ (200 µL, 100 mM, pH 9.5) followed by a solution of dansyl chloride in MeCN (200 µL, 10 mM). After 1 h, the reaction mixture was analyzed on a Kinetex C18 column (100 x 4.6 mm, 5 µm, 1 mL min^{-1}) using a 10 min gradient from 10-100% MeCN in water containing 0.1% formic acid.

3 Results I

Results I: Studies on the biosynthesis of the cyclopropylalanine (Acpa) moiety of the proteasome inhibitor belactosin A

3.1 Heterologous expression of the belactosin biosynthetic gene cluster

First, based on the findings by Felix Wolf and his PhD thesis, we wanted to verify by heterologous gene cluster expression, the integrity of the cluster and if this BGC synthesizes belactosins and to identify the elements that are necessary to build the Acpa moiety. To this end, we transferred the fosmid 6D1 containing the belactosin pathway into *Streptomyces albus* J1074 for heterologous expression. To achieve stable chromosomal integration and expression of the *bel* cluster in the heterologous host system, the fosmid needed to be refurbished. Therefore, the chloramphenicol resistance cassette of the 6D1 fosmid was replaced by the so called, int_neo cassette by λ-RED mediated recombination. This cassette was digested and purified from merLK01[178] and encodes for an origin of transfer (*oriT*), a *neo* resistance marker for selection and the phage ϕC31 components, an attachment site (attP) and an integrase (*int*), to allow stable and site-specific chromosomal integration in the host strain.[190] The detailed cloning procedure is described in 2.17.1. After successful conjugation and chromosomal integration of the fosmid belFeW01, which carries the complete belactosin gene cluster, into the *S. albus* J1074 host system, we conducted cultivation and production experiments, with subsequent compounds isolation by whole liquid culture solvent extraction. To our satisfaction, LC-MS analysis of those extracts showed that the generated heterologous expression mutant *S. albus*/belFeW01, further referred to as *S. albus J1074 bel01* (*S. albus* + *bel* cluster), was able to produce two new peaks with *m/z* 370.2 $[M+H]^+$ and *m/z* 358.2 $[M+H]^+$ (Figure 12 *(ii)*). In comparison the host without cluster (*S. albus* w/o cluster, Figure 12 *(iii)*) did not produce such a compound. Those peaks were exhibiting retention times corresponding exactly to those of belactosin A and belactosin C as produced in the wild type producer strain *Streptomyces* sp. UCK14 (Figure 12 *(i)*), respectively. Furthermore, when analyzing *m/z* 370.2 $[M+H]^+$ and *m/z* 358.2 $[M+H]^+$ from *S. albus* bel01 in MS/MS spectra, we could correlate all fragments to the belactosin C (Figure 13) and belactosin A (Figure 14) molecule. This confirmed that the introduced BGC encodes for all genetic elements required to produce both belactosins, including the Acpa moiety.

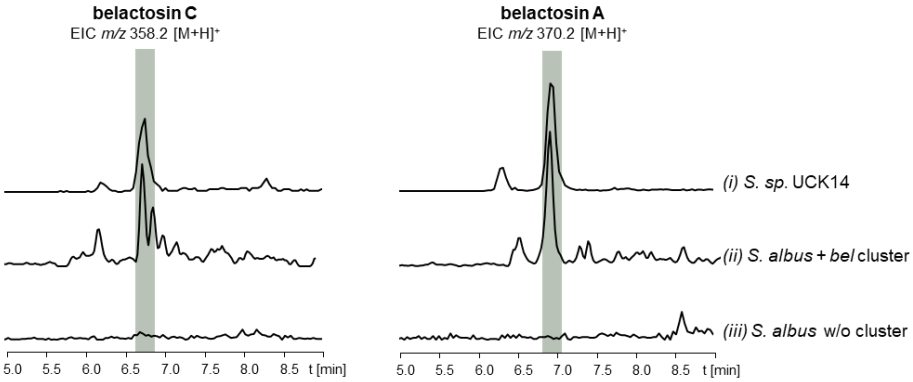


Figure 12: LC-MS analysis of the heterologous expressed belactosin gene cluster:*(i)* LC-MS analysis of the belactosin wild type producer strain *S.* sp. UCK14, EIC of belactosin C *m/z* 358.2 $[M+H]^+$ and EIC *m/z* 370.2 $[M+H]^+$ of belactosin A. *(ii)* Heterologous expression of belactosin C *m/z* 358.2 $[M+H]^+$ and belactosin A m/z 370.2 $[M+H]^+$ in *S. albus* J1074 *bel01* (*S. albus* + *bel* cluster). *(iii)* EIC *m/z* 358.2 $[M+H]^+$ and *m/z* 370.2 $[M+H]^+$ of culture extracts of the heterologous host without cluster (*S. albus* w/o cluster).

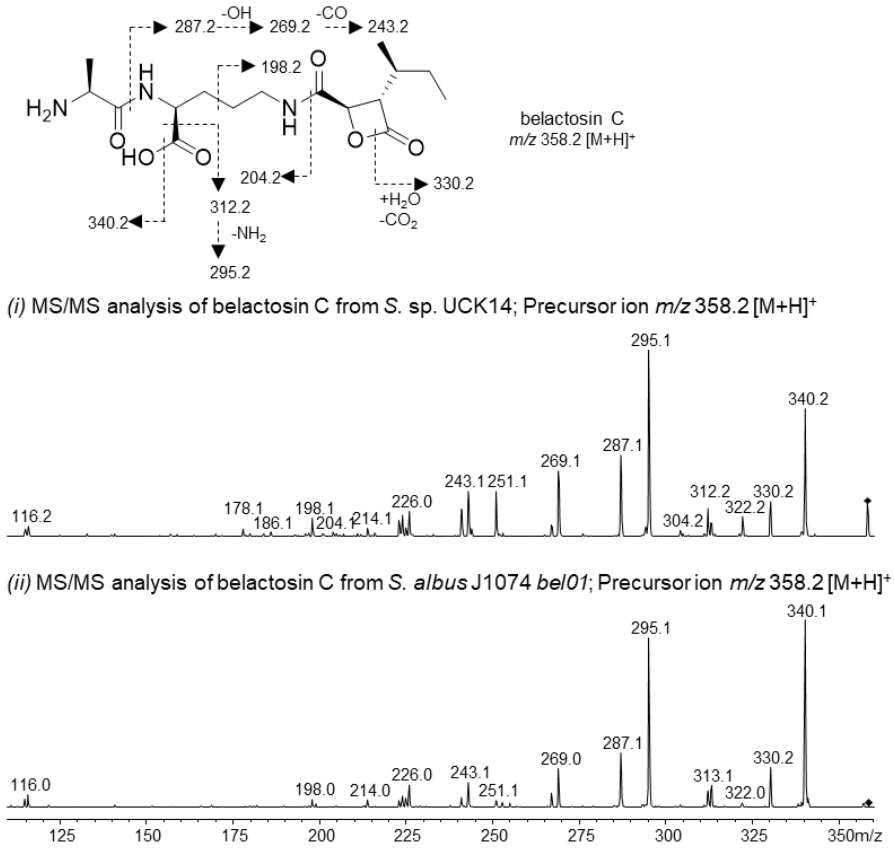


Figure 13: MS/MS analysis of belactosin C: MS/MS analysis of belactosin C (*m/z* 358.2 $[M+H]^+$) in *S.* sp. UCK14 *(i)* and *S. albus* J1074 *bell01* (*S. albus* + *bel* cluster) (*ii).* Precursor mass is marked with a black diamond and the expected MS/MS fragments for belactosin C are depicted above.

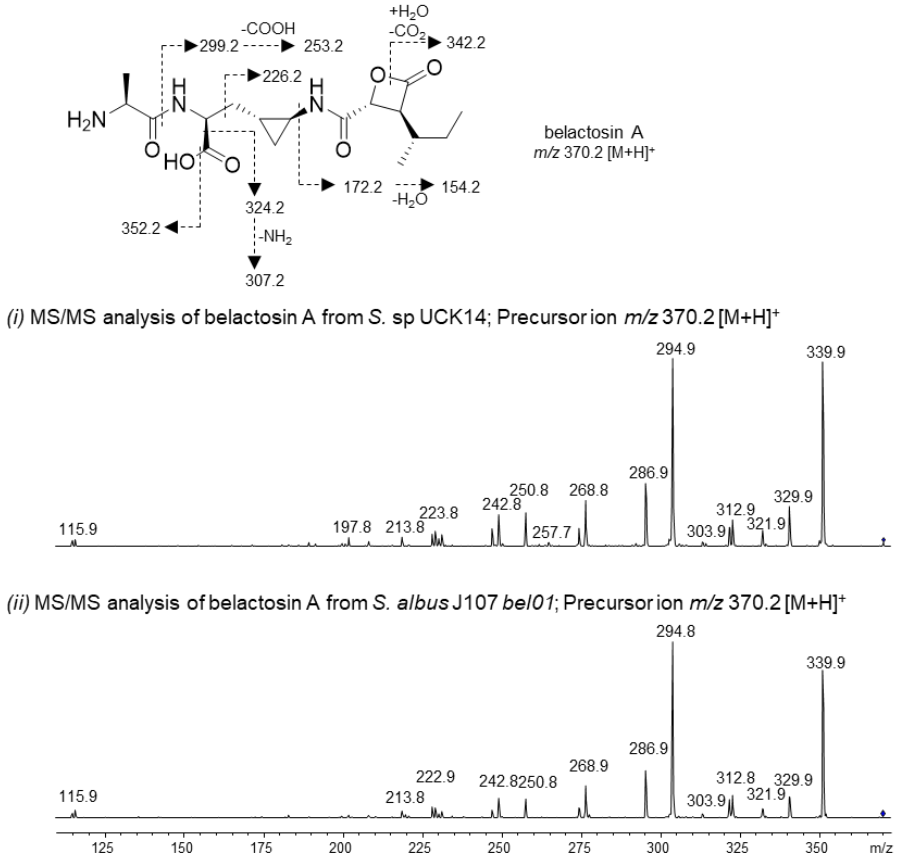


Figure 14: MS/MS analysis of belactosin A: MS/MS analysis of belactosin A (*m/z* 370.2 $[M+H]^+$) in *S. sp.* UCK14 *(i)* and *S. albus* J1074 *bell01* (*S. albus* + *bel* cluster) *(ii)*. Precursor mass is marked with a black diamond and the expected MS/MS fragments for belactosin A are depicted above.

3.2 Border analysis of the belactosin biosynthetic gene cluster

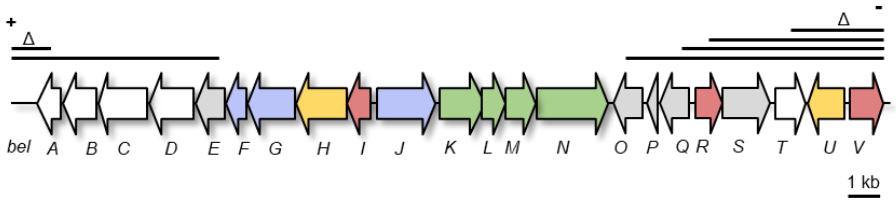
One of the main goals of this thesis was to identify the enzymes, which participate in the generation of the Acpa moiety within belactosin biosynthesis. Hence, we aimed to determine the borders of the corresponding BGC we had identified by heterologous expression. Initially, based on protein similarity, operon organization and comparison with characterized and uncharacterized homologous gene clusters we assigned 22 genes to the production of belactosins (Figure 15A).[77]

To validate the gene cluster borders up and down stream of the *bel* cluster and identify the genes essential for belactosin and Acpa biosynthesis, we generated a set of gene deletion mutants in the generated heterologous expression system. On the upstream border we deleted *belorf-1* and *belA-E* (Figure 15C, *(ii)* and *(iii)* and genes *belO-V, belQ-V*, *belR-V,* and *belT-V* (Figure 15 *(iv) – (vii)*), on the downstream border, respectively. The different gene knock-outs are also indicated in Figure 15A in the respective BGC, as well as in a gene annotation table in Figure 15B. Surprisingly, the production of both belactosin derivatives A and C was not affected in the two downstream mutants. Both EICs *m/z* 358.2 $[M+H]^+$ for belactosin C and *m/z* 370.2 $[M+H]^+$ for belactosin A could still be detected, as seen in the *S. albus* J1074 *bell01* (*S. albus* + *bel* cluster) control strain (Figure 15C (*i*)-(*iii*)).

This indicates that genes *orf-1-belE* are most likely not part of the BGC. On the other hand, knocking out genes *belO-V* on the downstream border resulted in complete loss of production for both belactosin derivatives in all generated mutants. Neither EIC *m/z* 358.2 $[M+H]^+$ for belactosin C nor EIC *m/z* 370.2 $[M+H]^+$ for belactosin A could be detected any more (Figure 15C(*i*)-(*iii*)), proving them to be essential for synthesis of these compounds. We thus reassigned the belactosin biosynthetic gene cluster to the genes *belF-belV*.

Concluding from the cluster border analysis and our previous studies on the biosynthesis of belactosins, this leaves two putative operons as candidate sub-clusters to direct the synthesis of the unique Acpa moiety in belactosin A. One of them consists of the genes *belA-belE,* which are encoding for a putative acetyltransferase, two peptidases, a deacetylase and an ABC-type transporter. The other candidate is represented by the four gene operon *belK-N* (indicated in green, Figure 15B), encoding for two oxygenases, an epimerase and a molybdopterin dependent oxidureductase. As the border knock-outs have already demonstrated, that genes *belA-E* are not essential for belactosin production, bel*K-N* remain the most promising gene candidates to synthesize the Acpa moiety.

A



B

Gene	Predicted function	Gene	Predicted function
orf-1	glyoxalase	***belK***	Iron-containing redox enzyme
belA	acetyltransferase	***belL***	BsmA domain containing protein
belB	peptidase U61	***belM***	diaminopimelate epimerase
belC	acyl-peptide hydrolase	***belN***	molybdopterin oxidoreductase
belD	acetyl-ornithine deacetylase	***belO***	transcriptional regulator
belE	ABC transporter	***belP***	transcriptional regulator
belF	3-isopropylmalate dehydratase SS	***belQ***	transcriptional regulator
belG	3-isopropylmalate dehydratase LS	***belR***	methylesterase
belH	AMP-dependent synthetase	***belS***	MFS transporter
belI	SAM-methyltransferase	***belT***	1,4-beta-xylanase
belJ	2-IPM synthase	***belU***	ATP-dependent amide ligase
		belV	AMP-dependent synthetase

C

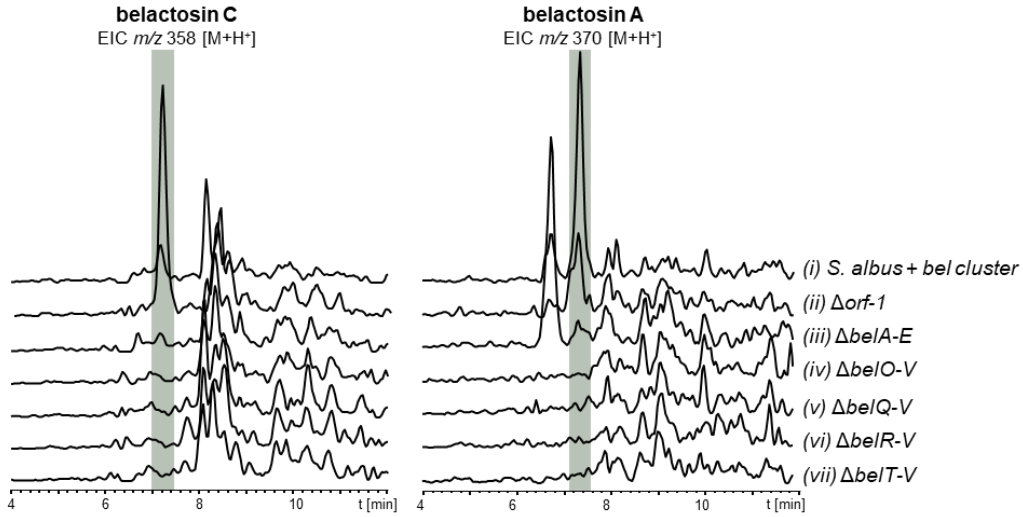


Figure 15: Gene cluster border analysis of the belactosin BGC:A: Illustration of the belactosin biosynthetic gene cluster. The gene cluster border knock-outs on the upstream (*orf-1*; *belA-E*) and the downstream border (*belO-V*; *belQ-V*; *belR-V*; *beltT-V*) of the cluster are indicated by black lines and Δ. Additionally, + and - indicate, whether belactosins are still produced in the generated knockout strains or not. **B:** Gene cluster annotation table of the demonstrated belactosin gene cluster in A. **C:** LC-MS analysis of gene cluster border knock-outs. Extracted ion chromatograms *m/z* 370.2 $[M+H]^+$ (belactosin A) and *m/z* 358.2 $[M+H]^+$ (belactosin C) of *orf-1* and *belA-E* (*ii*), (*iii*), and Δ*belO-V*, Δ*belQ-V*, Δ*belR-V* and Δ*belT-V* (*iv*) – (*vii*). The positive control *S. albus* + *bel* cluster (*S. albus* J1074 *bel01*), showing EIC chromatograms *m/z* 370.2 $[M+H]^+$ (belactosin A) and *m/z* 358.2 $[M+H]^+$ (belactosin C) at the expected retention time, is shown in *(i)*. This figure was partly taken and adjusted from Engelbrecht et al.[195]

3.3 Gene knock-out studies in the heterologous expression system: Assignment of the putative *belK-N* operon for the biosynthesis of Acpa

To identify and clearly demonstrate the responsible enzymes, which participate in the generation of the unique 3-(2-aminocyclopropyl)alanine (Acpa) residue within belactosin A biosynthesis, we decided to knock-out genes *belKLN* in the heterologous expression system. Therefore, we first constructed knock-outs of genes *belK*, *belL* and *belN* individually, on the belFeW01 fosmid containing the belactosin pathway and the *int_neo* cassette (cloning procedure is depicted in 2.17.2). At this point of the project we did not knock-out, gene *belM*, as it encodes for an putative esterase enzyme that is most likely involved in supply for the precursor lysine from primary metabolism and might not be directly involved in synthesis of Acpa.[117] In the next step, these altered knock-out fosmids were transferred into *Streptomyces albus* J1074 by conjugation for heterologous expression. The mutant strains, together with the wild type producer strain *Streptomyces* sp. UCK14 and *S. albus* J1074 bel01 (*S. albus*+ *bel* cluster) were cultivated in R5 production medium. Gratifyingly, LC-MS analysis of the culture extracts showed that the generated mutants *S. albus* J1074 bel01Δ*belK, S. albus* J1074 bel01, *S. albus* J1074 bel01/Δ*belL, S. albus* J1074 bel01/Δ*belN* were not able to produce belactosin A anymore. In all three mutants, the expected peaks for EIC *m/z* 370.2 $[M+H]^+$ could not be detected anymore. In contrast, belactosin C production was not impaired in all three mutant strains (Δbel*K/L/N*) and distinct peaks could be detected for EIC *m/z* 358.2 $[M+H]^+$ (Figure 16 *(iii-v)*) as seen in *Streptomyces* sp. UCK14 and *S. albus* J1074 bel01 *(S. albus + bel cluster*) (Figure 16 *(i-ii)*). Deduced from the fact, that only belactosin A carries a cyclopropyl containing residue and in combination with the knock-out experiments of the *belK-N* operon, it could be clearly shown, that the enzymes BelK/L/N are required for the formation of the 3-(2-aminocyclopropyl)alanine moiety.

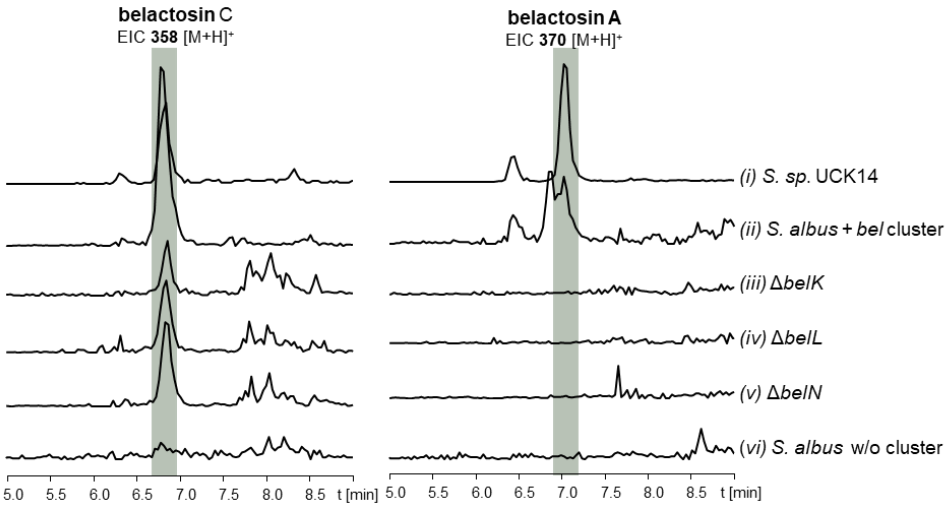


Figure 16: Individual gene cluster knock-outs of genes *belK, belL and belN* in the heterologous host *S. albus* J1074: LC-MS analysis for verification of the production of belactosin C (EIC *m/z* 358.2 $[M+H]^+$) and belactosin A (EIC *m/z* 370.2 $[M+H]^+$) in the wild type producer strain *(i)*, the heterologous host with *(ii)* and without belactosin cluster *(vi)* and the three mutants Δ*belK (iii), ΔbelL (iv), and ΔbelN (v).*
Δ*belK* = *S. albus* J1074 bel01/Δ*belK*; Δ*belL* = *S. albus* J1074 bel01/Δ*belL*; Δ*belN* = *S. albus* J1074 bel01/Δ*belN; S. albus + bel cluster* =*S. albus* J1074 *bel01; w/o = without*

3.4 Biosynthesis hypothesis proposal towards the Acpa moiety of belactosin A: Feeding studies with stable isotope labeled precursors

Based on the knowledge that genes *belK/L/N* are involved in the biosynthesis of the unique Acpa moiety of belactosin A, we wanted to further elucidate the specific catalytic roles of the encoded enzymes in more detail and proposed a theoretical biosynthesis route to the Acpa intermediate. Taking a closer look at this four gene operon it became quite obvious that homologous genes to *belK, belL* and *belM* but not *belN* are encoded in the biosynthetic gene cluster of hormaomycin from *Streptomyces griseoflavus* W-384 (Figure 17A).[117] These homologous proteins show identities of BelK to HrmI of 51 %, BelL to HrmJ of 49 % and BelM to HrmT of 43 % on amino acid level. Thereby, *belK*, *belL* encode for putative oxidoreductases and *belM* for a putative epimerase (Figure 17B).

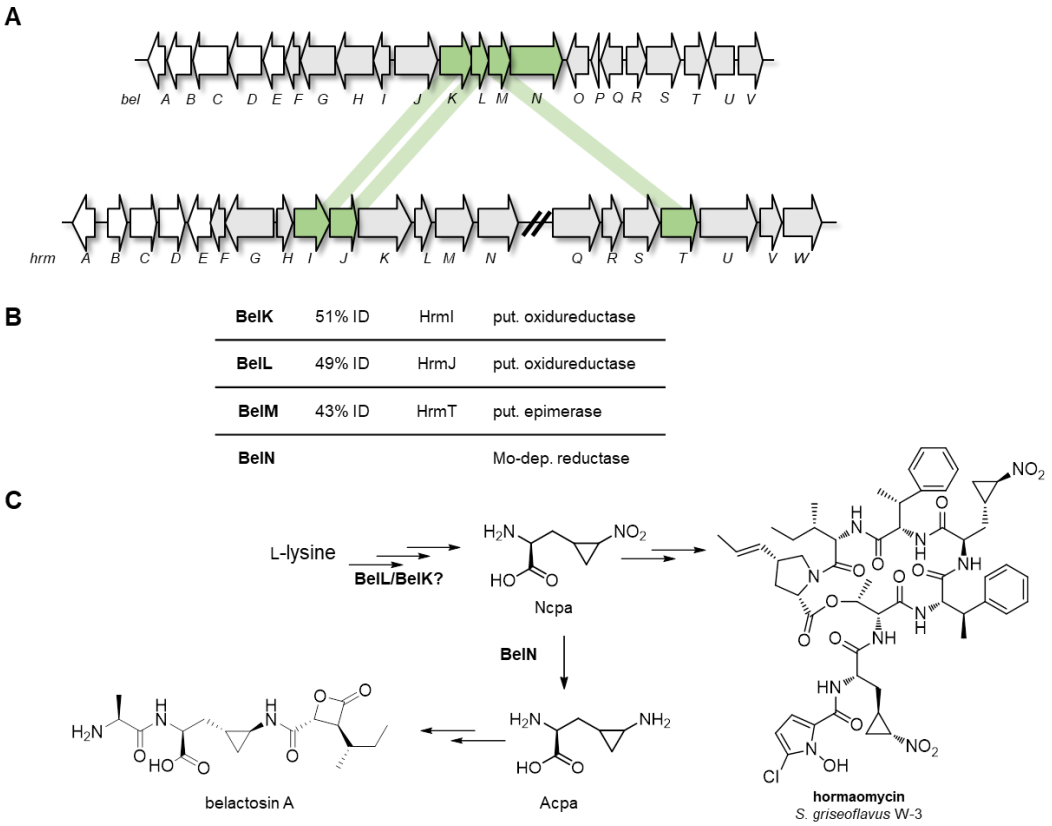


Figure 17: Biosynthesis model towards formation of the Acpa moiety of belactosin A: A: BGCs of belactosin (*bel*) and hormaomycin *(hrm*), homologous genes are indicated by green bars. **B:** Identities (ID, in %) and gene annotation of homologous *bel* and *hrm* genes. **C:** Biosynthesis model for generation of the Acpa intermediate to belactosin A, starting from L-lysine and via the Ncpa intermediate, which is part of hormaomycin biosynthesis and incorporated into the hormaomycin molecule.

Based on their detailed studies on the biosynthesis of hormaomycin, the group around Prof. Jörn Piel (ETH Zürich) suggested that HrmI and HrmJ might be involved in the formation of the unique amino acid nitrocyclopropylalanine (Ncpa) which is incorporated in hormaomycin.[117] As both, Ncpa and the belactosin A intermediate Acpa contain *trans*-disubstituted cyclopropanes, we speculated that they could be constructed via the same biosynthetic strategy.

Based on those findings we postulated a biosynthesis hypothesis for the formation of the Acpa moiety towards belactosin A (Figure 17C). To test the postulated biosynthesis hypothesis we first aimed to determine whether the Acpa moiety derives from an L-ornithine precursor by methylation of a double bond or originates from, as proposed for the Ncpa intermediate in hormaomycin biosynthesis, L-lysine via intramolecular cyclopropanation (Figure 18A). Therefore, we fed separate cultures of the belactosin producer strain with stable labeled $^{13}C_5$-L-ornithine and $^{13}C_6$- L-lysine, respectively, and analyzed the culture extracts by LC-MS and MS/MS in comparison with untreated cultures.

Analysis by LC-MS (Figure 18B) and MS/MS (Figure 19) revealed the incorporation of all five carbon atoms from $^{13}C_5$-L-ornithine as the central amino acid in belactosin C, demonstrated by an isotopic mass shift from *m/z* 358.2 $[M+H]^+$ to *m/z* 363.2 $[M+5+H]^+$ (Figure 18B *(ii)*). Furthermore, when analyzing masses of fed and unfed cultures *m/z* 358.2 $[M+H]^+$ and *m/z* 363.2 $[M+5+H]^+$ with MS/MS, the incorporation of all five labeled carbon atoms of $^{13}C_5$-L-ornithine and a 5 Da shift could nicely be followed in all expected MS/MS fragments (Figure 19 *(i)-(iii)*). Notably, the production of belactosin C was greatly diminished in the $^{13}C_6$- L-lysine supplemented cultures, but the isotope pattern of the compound was unchanged ((Figure 18B *(iii)*). These results clearly indicate that the Acpa moiety in belactosin A is deriving from L-lysine.

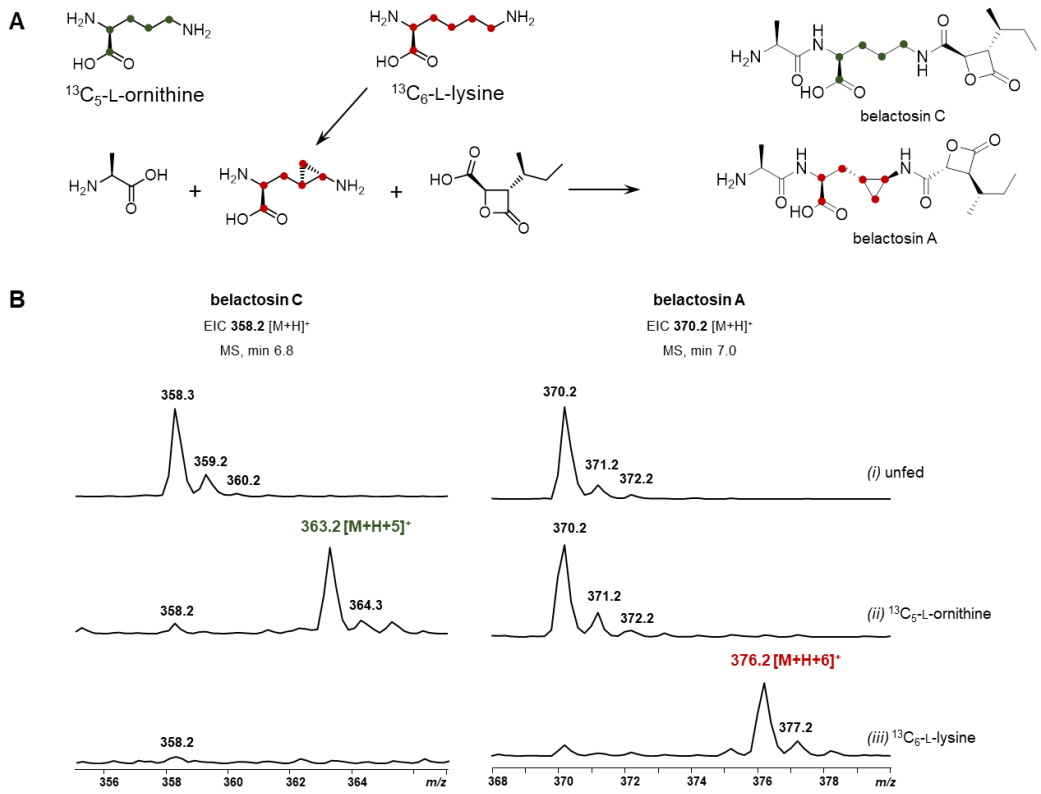


Figure 18: Feeding experiments with stable isotope labeled precursors: A: To investigate if the cycloporopyl containing moiety Acpa in belactosin A is deriving by ornithine by methylation of a double bond or from L-lysine via intramolecular cyclopropnation, cultures of wild type producer strain *S. sp.* UCK14 were fed with $^{13}C_5$ L-ornithine and $^{13}C_6$ L-lysine, respectively. Incorporation of the labeled carbon atoms from with $^{13}C_5$ L-ornithine and $^{13}C_6$ L-lysine in belactosin C and beleactosin A are indicated by green and red dots, respectively. **B:** Isotopic peak pattern analysis of the precursor feeding studies for masses *m/z* 358.2 $[M+H]^+$ for belactosin C and *m/z* 370.2 [M+H]+ for belactosin A in unfed cultures *(i), m/z* 363.2 $[M+5+H]^+$ for $^{13}C_5$-belactosin C in $^{13}C_5$-L-ornithine fed *S. sp.* UCK14 WT cultures *(ii),*and *m/z* 376.2 $[M+6+H]^+$ for $^{13}C_6$-belactosin A in $^{13}C_6$-L-lysine fed *S. sp.* UCK14 WT cultures *(iii).*

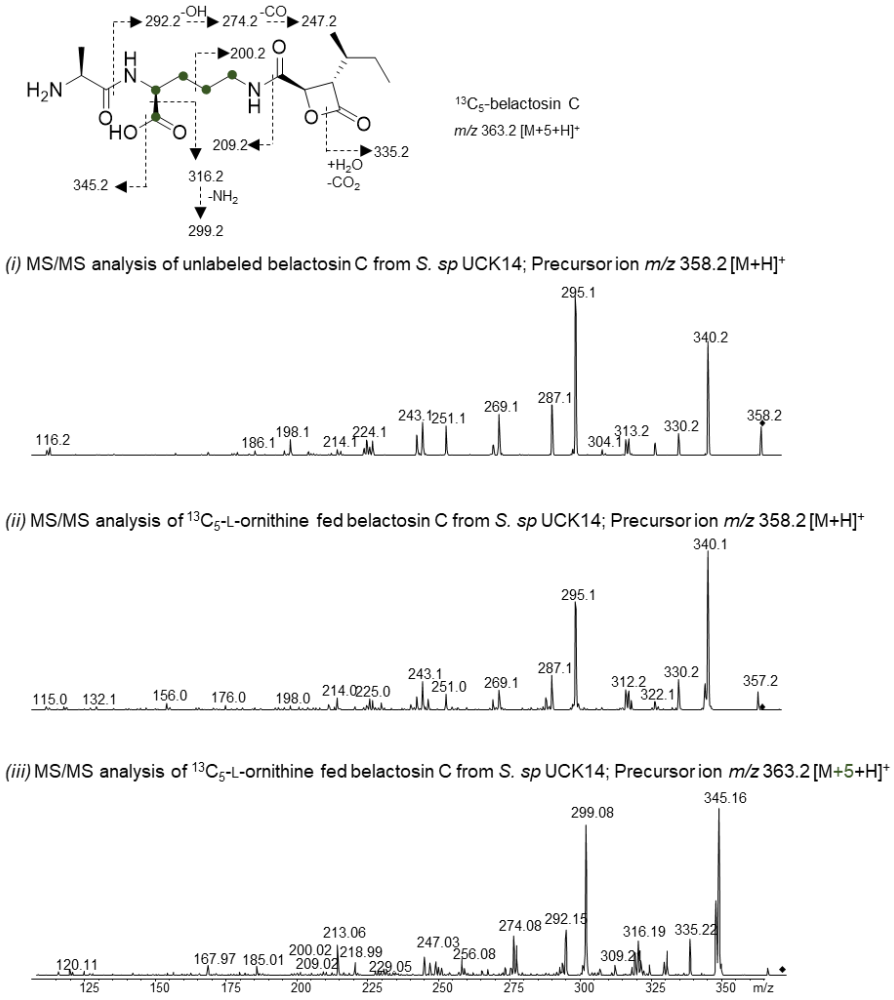


Figure 19: MS/MS analysis of $^{13}C_5$-L-ornithine feeding studies: Proposed MS/MS fragments for $^{13}C_5$-enriched belactosin C, labeled carbon atoms are marked in green. MS/MS profiles for belactosin C (precursor mass 358.2 $[M+H]^+$) from unfed *(i)* and $^{13}C_5$-L-ornithine fed *(ii)* cultures. MS/MS profile of $^{13}C_5$-enriched belactosin C (precursor mass 363.2 $[M+5+H]^+$) *(iii)*, all in *S.* sp. UCK14 respectively. Precursor mass is marked with a black diamond.

On the other hand, the isotopic peak pattern of belactosin A was not affected in the presence of the labeled $^{13}C_5$-L-ornithine precursor and no mass shift could be observed (Figure 18B *(ii)*). In contrast, feeding of stable isotope labeled $^{13}C_6$-L-lysine to the UCK14 producer strain, resulted in an isotopic mass shift of six from *m/z* 370.2 $[M+H]^+$ to *m/z* 376.2 $[M+6+H]^+$ and a substantial increase of the *m/z* 376.2 $[M+6+H]^+$ isotope signal for belactosin A (Figure 18B *(iii)*). MS/MS analysis of isotope enriched belactosin A showed that all six ^{13}C-carbons from the precursor are incorporated in the belactosin A molecule, including the Acpa moiety, as all expected fragments could be nicely detected (Figure 20).

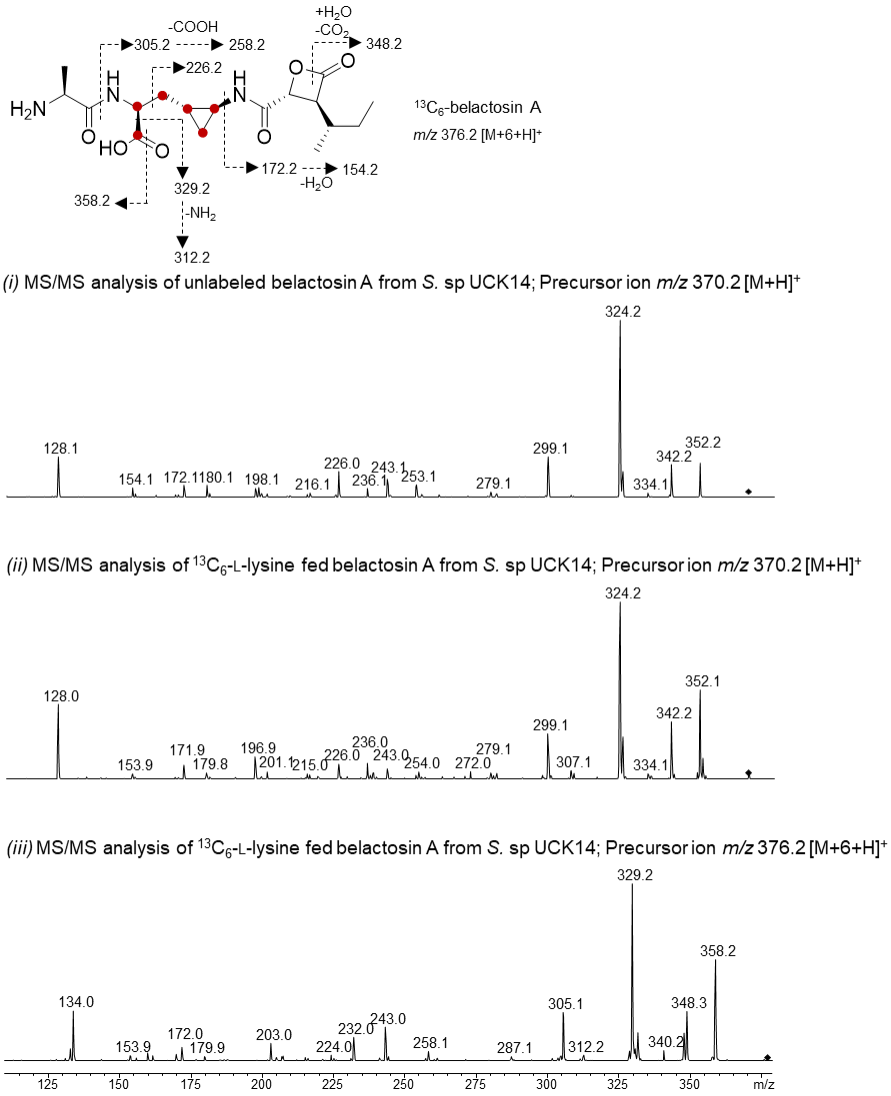


Figure 20: MS/MS analysis of $^{13}C_6$-L-lysine feeding studies: Proposed MS/MS fragments for $^{13}C_6$-enriched belactosin A, labeled carbon atoms are marked in red. MS/MS profiles for belactosin A (precursor mass *m/z* 370.2 $[M+H]^+$) in unfed *(i)* and $^{13}C_6$- L-lysine fed wild type cultures *(ii)* as well as $^{13}C_6$-enriched belactosin A (precursor mass *m/z* 376.2 $[M+6+H]^+$) *(ii)* in *S. sp.* UCK14. Precursor mass is marked with a black diamond.

Our results from the labeled precursor feeding studies clearly demonstrated that the Acpa moiety derives from L-lysine. So we suggest that the same biochemistry directs the Acpa and Ncpa biosynthesis pathway, presumably via the generation of a common *N*-oxidized intermediate, as descried in 3.4 (Figure 17C). In order to explore the nature of this proposed intermediate in the biosynthetic pathway of belactosin A, we supplemented cultures of *Streptomyces* sp. UCK14 with synthetic 3,3-d_2-Ncpa (3,3-d_2-3-(*trans*-2'-nitrocyclopropyl)alanine), which was synthesized and provided by the group of Prof. Armin de Meijere from the University of Göttingen.[196] For additional quality and integrity control of the synthetic 3,3-d_2-Ncpa we conducted ^{1}H-NMR analysis in collaboration with Dr. Chambers Hughes (University of Tübingen). The measured ^{1}H-NMR spectrum (Appendix,

Figure 51) showed complete accordance to the published data.[197] Intriguingly, LC-MS and MS/MS analysis showed the successful incorporation of the double deuterium-labeled precursor in the belactosin A molecule. This incorporation could be nicely followed by a drastic increase of the peak for EIC *m/z* 372.2 $[M+2+H]^+$ isotope signal in comparison to *m/z* 370.2 $[M+H]^+$ for unlabeled belactosin A of unfed cultures and a simultaneous decrease of the respective monoisotopic signal (Figure 21*(i, ii)*).

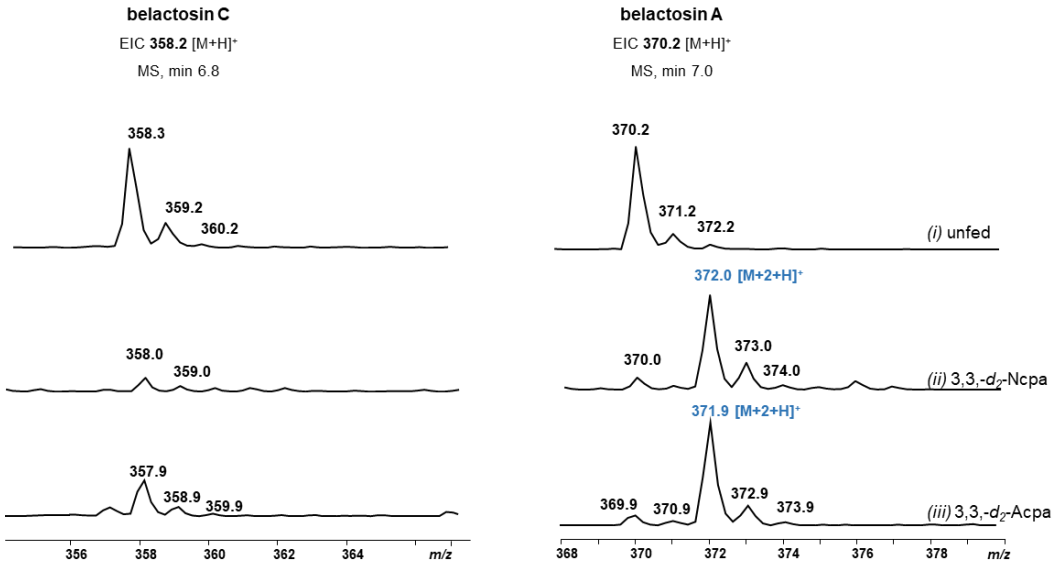


Figure 21: Feeding of double deuterium-labeled intermediates of belactosin A biosynthesis: MS isotope peak pattern analysis of precursor mass *m/z* 358.2 $[M+H]^+$ and *m/z* 370.2 $[M+H]^+$ for unfed *(i),* 3,3-d_2-Ncpa fed *(ii),* and the 3,3-d_2-Acpa fed *(iii) S.* sp. UCK14 wild type producer strain. The successful incorporation of 3,3-d_2-Ncpa *(ii)* and 3,3-d_2-Acpa *(iii)* in the belactosin A molecule, can nicely be followed by a 2 Da shift in the fragmentation patterns of precursor mass *m/z* 370.2 $[M+H]^+$ in comparison to all expected fragments from precursor mass *m/z* 372.0 $[M+2+H]^+$ and *m/z* 371.9 $[M+2+H]^+$, respectively.

This 2 Da shift and the incorporation of the 3,3-d_2-Ncpa could also be nicely followed in all predicted MS/MS fragments of belactosin A in comparison with the fragmentation pattern of precursor ion EIC *m/z* 372.2 $[M+H]^+$ with EIC *m/z* 370.2 $[M+H]^+$ in the unfed UCK14 WT sample (Figure 22*(i, ii)*). By feeding 3,3-d_2-Ncpa the production of belactosins is almost completely channeled towards the belactosin A derivative, similar to cultures supplemented by L-lysine. This decrease is indicated by a drastic decline of the *m/z* 358.2 $[M+H]^+$ signal in direct comparison to the signal of unfed cultures (Figure 21*(i,ii)*). These results strongly implicate that Ncpa is a genuine intermediate of the belactosin A biosynthetic pathway.

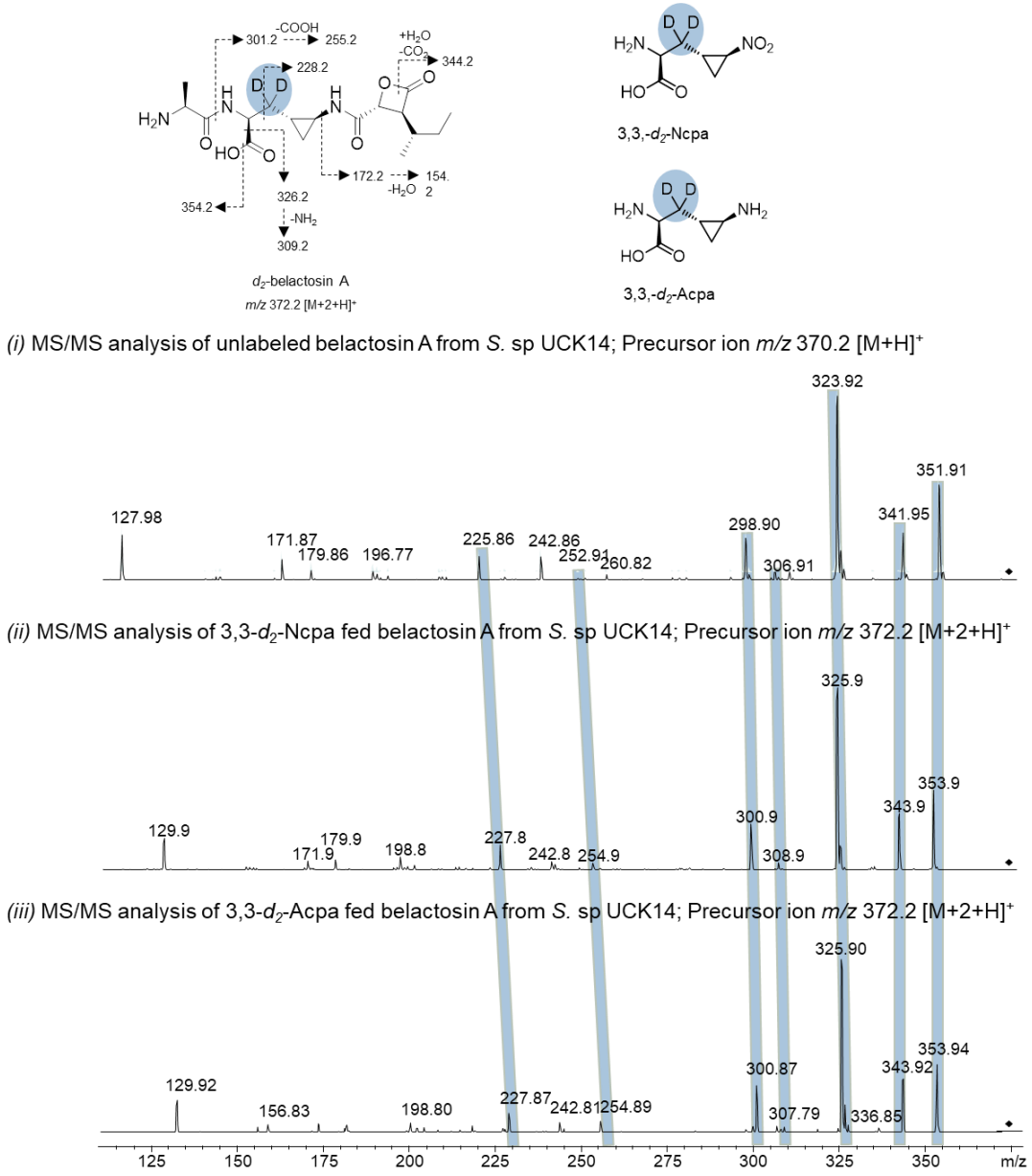


Figure 22: MS/MS analysis of 3,3-d_2-Ncpa and 3,3-d_2-Acpa feeding studies: Proposed MS/MS fragments for double deuterium-labeled belactosin A and molecule structure of the 3,3-d_2-Ncpa and 3,3-d_2-Acpa intermediates, position of the deuterium label is marked in blue. MS/MS profiles for belactosin A (*m/z* 370.2 $[M+H]^+$) in unfed *(i)* and 3,3-d_2-Ncpa fed *S.* sp. UCK14 cultures (precursor mass *m/z* 372.2 $[M+2+H]^+$) *(ii)* as well as 3,3-d_2-Acpa fed cultures (precursor mass *m/z* 372.2 $[M+2+H]^+$) *(iii)*. Precursor mass is marked with a black diamond.

Following our postulated biosynthesis hypothesis towards belactosin A, the Ncpa intermediate would have to undergo further reduction to afford the terminal amine in the Acpa moiety of belactosin A. To test if the Acpa moiety is formed as an intermediate towards the belactosin A molecule, we conducted feeding studies with stable isotope labeled Acpa. Analogous to the 3,3-d_2-Ncpa feeding studies, we supplemented *Streptomyces* sp. UCK14

bacterial cultures with synthetic 3-(*trans*-2'-aminocyclopropyl)alanine (3,3-d_2-Acpa), which was chemically synthesized[197, 198] and kindly provided to us by the working group of Prof. Armin de Meijere (University of Göttingen). Before conducting feeding studies with synthetic 3,3-d_2-Acpa, we checked the integrity of the chemically synthesized intermediate by LC-MS. To optimize the analysis regarding the retention time on reversed phase RP-HPLC, we performed derivatization of the molecule with dansyl chloride and analyzed the samples by LC-MS. This analysis showed that the provided sample indeed contains 3,3-d_2-Acpa (Appendix, Figure 52).

To our satisfaction after feeding with 3,3-d_2-Acpa, we could observe a +2 Da shift in the isotopic peak pattern of belactosin A from *m/z* 370.2 $[M+H]^+$ to *m/z* 372.2 $[M+2+H]^+$ (Figure 21 *(iii)*). This shift could also be determined by comparison of all expected MS/MS fragment patterns deriving from precursor masses *m/z* 370.2 $[M+H]^+$ and *m/*z 372.2 $[M+2+H]^+$ of the unlabeled belactosin A molecule and the double deuterium-labeled belactosin A of 3,3-d_2-Acpa fed cultures, respectively (Figure 22 *(i,iii)*). The successful incorporation of the 3,3-d_2-Acpa precursor into the belactosin A molecule confirmed its' role as a crucial intermediate in the belactosin A biosynthesis. These findings prompted us to investigate the formation of this Acpa intermediate in more detail, and we decided to carry out chemical complementation studies in comprehensive gene knockout strains.

3.5 Chemical complementation studies – Generation and analysis of markerless *belK*, *belL belM* and *belN* gene knock-out mutants

Now that we could confirm, that both Ncpa and Acpa are genuine intermediates of the belactosin A biosynthetic pathway, we aimed to further investigate the role of the genes *belK-N* within the Acpa formation in more detail. Hence, we wanted to knock-out all four genes of the *belK-N* operon individually and subsequently subject them to chemical complementation feeding studies. First attempts of conducting feeding studies with double deuterium-labeled Ncpa (3,3-d_2-Ncpa) in the heterologous expression strains *S. albus* J1074 bel01 and mutants *S. albus* J1074 bel01/Δ*belK, S. albus* J1074 bel01/Δ*belL. S. albus* J1074 bel01/Δ*belN*, did not lead to robust and reliable results (data not shown). Thus, we decided to generate individual double-crossover markerless knock-out mutants of the genes *belK-N* (Figure 23) in the UCK14 WT producer strain. All four genes of the operon were individually deleted from fosmid GD1, which contains the complete gene cluster of interest and was identified by screening a genomic library as described in 221.8.2, using Red/ET-mediated recombination.[183] To create in-frame deletions, the disruption cassette from plasmid pIJ773 was subsequently removed by the use of FLP-recombinase[183, 199] generating fosmids belFeW09, belFeW10, belFeW11, belFeW12 (see Table 21) for constructs *ΔbelK, ΔbelL, ΔbelM and ΔbelN,* respectively. The detailed cloning strategy is described in 2.17.4. After introduction of the generated fosmids into the UCK14 producer strain via triparental conjugation and alternating replica plating, mutant strains were screened by PCR for successful homologous recombination. Positive clones were finally confirmed by sequencing. The resulting markerless double cross over mutants, UCK14/Δ*belK,* UCK14/Δ*belL,* UCK14/Δ*belN* were subsequently cultivated in belactosin production medium SM19 and culture extracts were analyzed by LC-MS. This analysis showed, that the production of the belactosin A derivative was completely abolished in the *Streptomyces* sp. UCK14/*ΔbelK,* UCK14/*ΔbelL,* UCK14/*ΔbelN* mutants, as no signal for EIC *m/z* 370.2 $[M+H]^+$ could be detected anymore (Figure 23 *(ii, iii, v)*). On the other hand, in the *Streptomyces* sp. UCK14/*Δbel*M mutant, belactosin A (EIC *m/z* 370.2 $[M+H]^+$) was still produced as seen in the WT, implying that the putative epimerase enzyme BelM might not catalyze a crucial biosynthesis step in the construction of belactosin A. In contrast, the formation of the belactosin C derivative (*m/z* 358.2 $[M+H]^+$) was not affected in all four mutant strains *Streptomyces* sp. UCK14/Δ*belK,* UCK14/Δ*belL,* UCK14/Δ*belN* (Figure 23 *(ii, iii, iv, v)*). This further strengthened our previous findings from the analysis of the heterologous gene cluster

mutants *S. albus* J1074 bel01/Δ*belK*, *S. albus* J1074 bel01/Δ*belL,* *S. albus* J1074 bel01/Δ*belN*, that genes *belK, belL,* and *belN* seem to play an essential role for belactosin A formation.

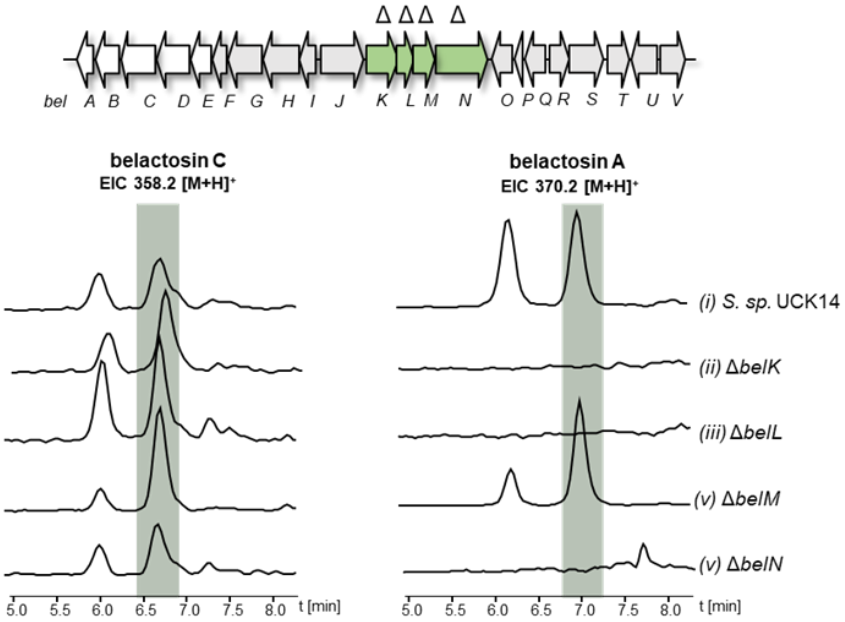


Figure 23: LC-MS analysis of markerless gene cluster knock-outs: Genes *belK*, *belL belM* and *belN* were individually knocked out in the wild type producer strain *S.* sp. UCK14 and culture extracts of these strains were subsequently analyzed by LC-MS: EIC *m/z* 358.2 $[M+H]^+$ for belactosin C and *m/z* 370.2 $[M+H]^+$ for belactosin A in *S.* sp. UCK14 *(i)* and the bel*K-N* mutants *(ii-v)*. Δb*elK* = *S.*sp. UCK14/ Δb*elK;* Δb*elL* = *S.* sp. UCK14/ Δb*elL*; Δb*elM* = *S.* sp. UCK14/ Δb*elM;* Δb*elN* = *S.* sp. UCK14/ Δb*elN*.

Intriguingly, when we carefully examined the LC-MS analysis of a dansyl-chloride treated UCK14/Δ*belN* extract we could discover the Ncpa accumulation shown by appearance of a mass signal matching dansyl chloride-derivatized Ncpa with *m/z* 408.4 $[M+H]^+$ (Appendix, Figure 54), whereas no Ncpa accumulation could be observed in the UCK14 wild type. The specific accumulation of Ncpa could be further confirmed in the *Streptomyces* sp. UCK14/Δ*belN* mutant by feeding the stable isototpic labeled biosynthesis specific precursor molecule $^{13}C_6$-L-lysine. After derivatization of the culture extract with danysl chloride we could detect the production of $^{13}C_6$-Ncpa, by the significant accumulation of *m/z* 414.4 $[M+6+H]^+$ (Appendix, Figure 55). This further strongly indicates that Ncpa is a substrate for the BelN enzyme.

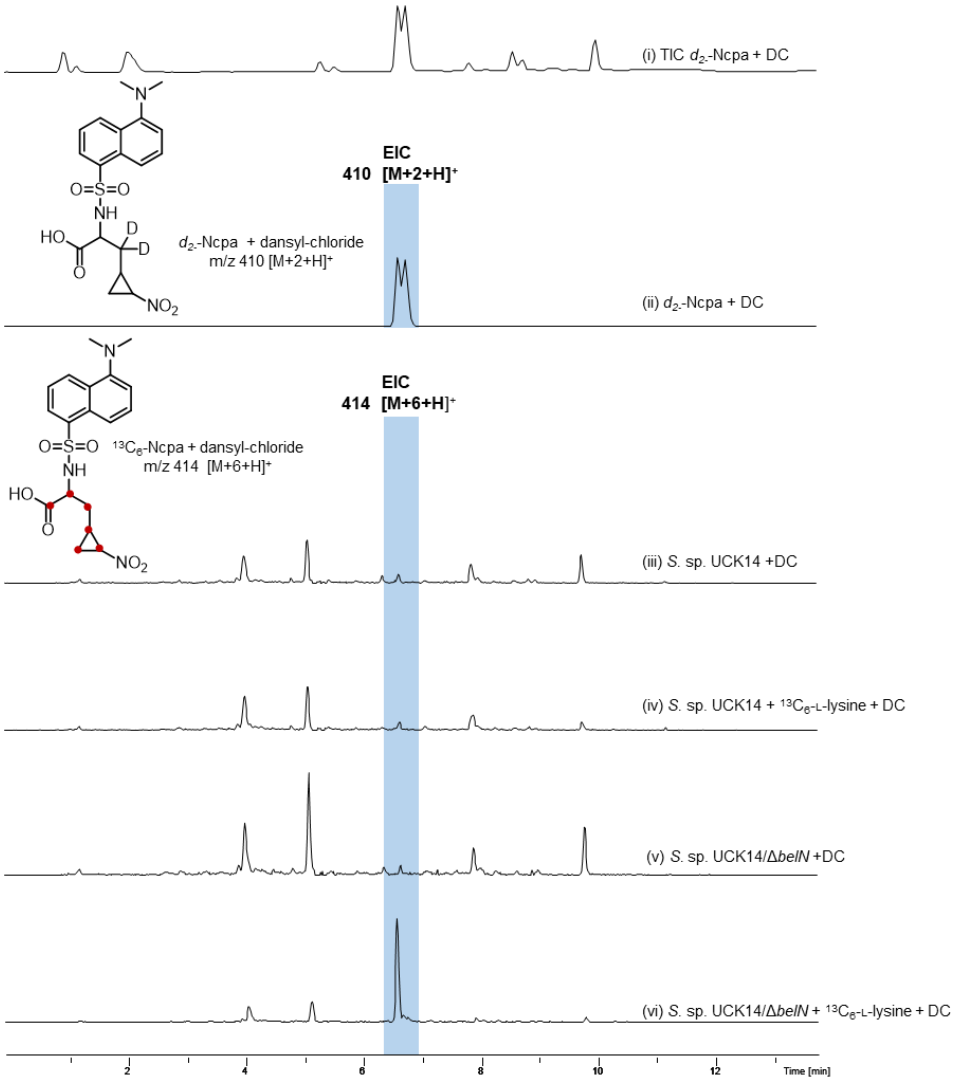


Figure 24: LC-MS analysis of dansyl chloride derivatized d_2-Ncpa:TIC (i) and EIC *m/z* 410 $[M+2+H]^+$ (ii) LC-MS analysis of culture extracts of *S. sp.* UCK14 and *S. sp.* UCK14/*ΔbelN* which were unfed (iii, v) or supplemented with $^{13}C_6$-L-lysine (iv,vi). Accumulation of $^{13}C_6$ labeled and dansyl chloride derivatized Ncpa intermediate was followed by EIC 414 $[M+6+H]^+$. DC = dansyl chloride

To study the possible involvement of genes *belK, belL,* and *belN* in the formation of the Acpa moiety in more detail, we conducted further chemical complementation studies, which are described in the next chapter.

3.6 Chemical complementation studies with deuterium-labeled 3,3-d_2-Ncpa and 3,3-d_2-Acpa intermediates

Aim of the chemical complementation experiments was to further confirm and characterize the involvement of genes *belK/L/N* in the reduction step from the Ncpa to the Acpa moiety in belactosin A synthesis. As stated, Ncpa would have to be reduced to afford the terminal amine in Acpa. Given the co-localization with the genes BelK, BelL and BelM, a compelling candidate for such a reaction is the putative molybdopterin-dependent oxidoreductase BelN (Figure 25A). This class of enzymes is known to catalyze two-electron reductions of diverse inorganic and organic substrates including nitrate, sulfate, DMSO and arsenate.[200] To evaluate the possible role of BelN in the reduction from Ncpa to Acpa, we decided to carry out chemical complementation studies. The theoretical and expected incorporation scheme for d_2-Ncpa and d_2-Acpa into the belactosin A molecule is depicted in Figure 25B.

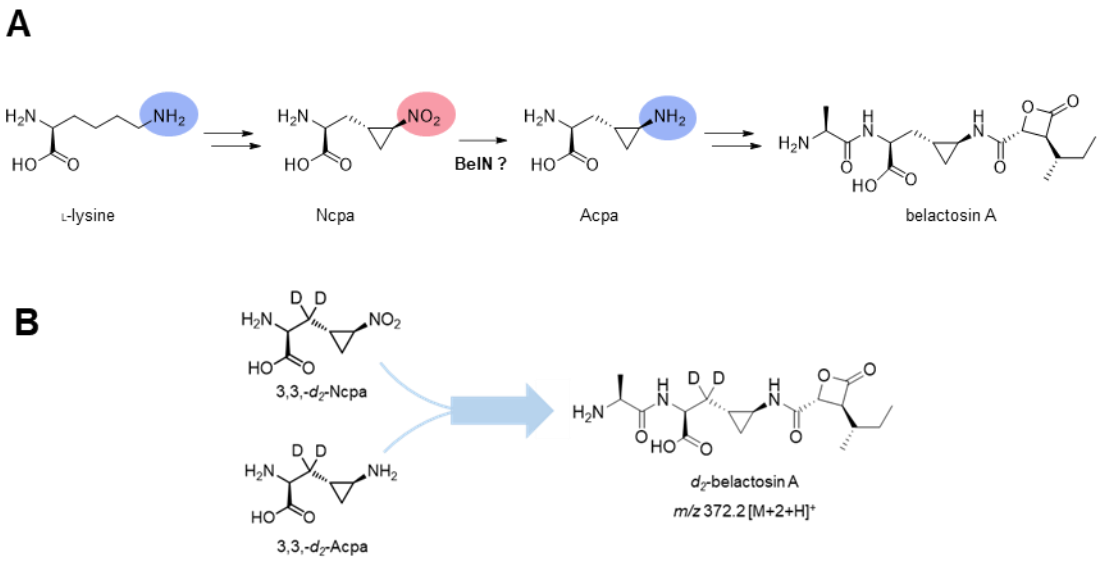


Figure 25: Chemical complementation studies: A: Reaction scheme for the biosynthesis pathway from L-lysine towards belactosin A via the postulated Ncpa and Acpa intermediates. **B:** Incorporation scheme of 3,3-d_2-Ncpa and 3,3-d_2-Acpa into belactosin A.

We first fed double deuterium-labeled 3,3-d_2-Ncpa to *Streptomyces* sp. UCK14 and the *Streptomyces* sp. UCK14/Δ*belK,* UCK14/Δ*belL,* UCK14/Δ*belN* mutants and subsequently analyzed the culture extracts by LC-MS (Figure 26A). First, the *Streptomyces* sp.UCK14 WT control (Figure 26A *(i)*) shows, that previous to feeding with 3,3-d_2-Ncpa EIC *m/z* 372.2 $[M+2+H]^+$ could not be detected, but after successful incorporation of the labeled precursor into the molecule, we were able to detect EIC *m/z* 372.2 $[M+2+H]^+$ in the culture extracts.

Furthermore, feeding of 3,3-d_2-Ncpa sucessfully restored the belactosin A production in the UCK14/Δ*belK* and UCK14/Δ*belL* mutants (Figure 26A *(ii, iii)*). In extracts of both mutant strains we were able to detect EIC *m/z* 372.2 $[M+2+H]^+$ again. On the other hand, feeding of 3,3-d_2-Ncpa did not lead to complementation of the UCK14/Δ*belN* mutant. Production of d_2-belactosin A could not be observed by LC-MS analysis (Figure 26A *(iv)*).

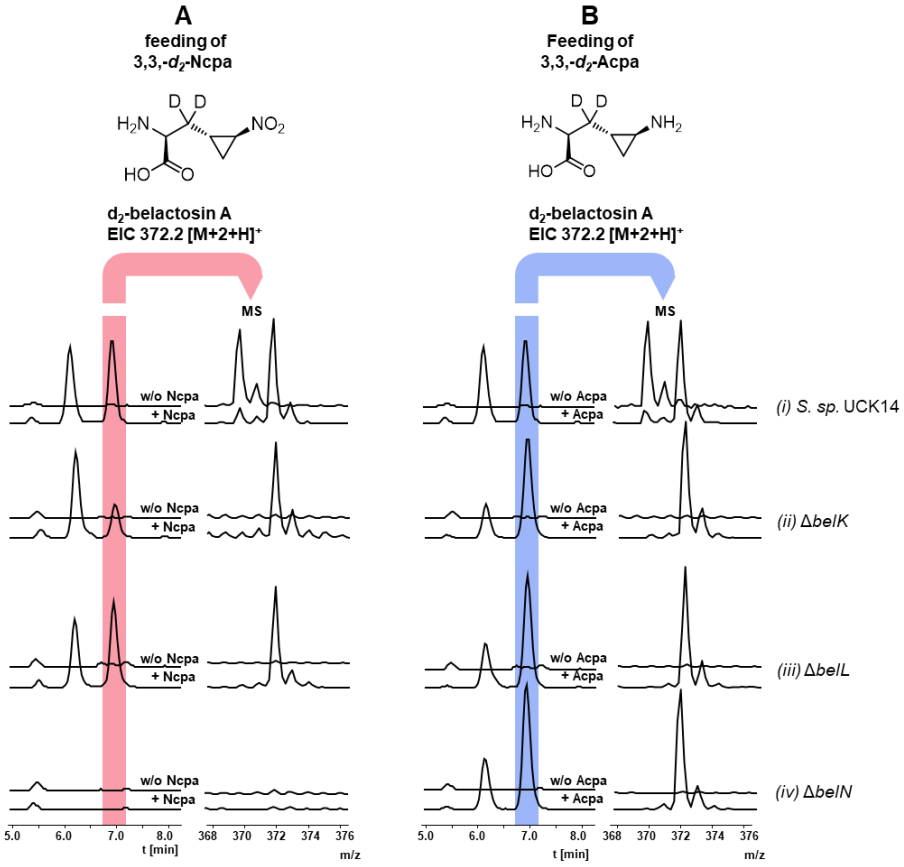


Figure 26: LC-MS analysis of chemical complementation studies:A: Chemical complementation studies with 3,3-d_2-Ncpa. EIC *m/z* 372 $[M+2+H]^+$ for fed (+ Ncpa) and unfed (w/o Ncpa) cultures of *S.* sp. UCK14*(i)* and mutant strains Δ*belK*/Δ*belL*/Δ*belN (ii-iv).* **B:** Chemical complementation studies with 3,3-d_2-Acpa. EIC *m/z* 372 $[M+2+H]^+$ for fed (+ Ncpa) and unfed (w/o Ncpa) cultures of *S.* sp. UCK14 *(i)* and mutant strains Δ*belK*/Δ*belL*/Δ*belN (ii-iv)*. The respective MS profiles for each EIC chromatogram at retention time 7 min is indicated with a red or blue arrow (A, B). Δ*belK* = *S.* sp. UCK14/Δ*belK;* Δ*belL* = *S.* sp. UCK14/Δ*belL;* Δ*belN* = *S.* sp. UCK14/Δ*belN*. Parts of this figure are taken and adjusted from Engelbrecht et al. and Shimo et al.[195, 201]

The successful incorporation of the 3,3-d_2-Ncpa precursor into the belactosin A molecule in the complemented mutant strains UCK14/Δ*belK* and UCK14/Δ*belL,* could also be nicely followed by MS/MS analysis. MS/MS spectra of the culture extracts of 3,3-d_2-Ncpa fed UCK14/Δ*belK* and UCK14/Δ*belL* with the precursor mass *m/z* 372.2 $[M+2+H]^+$ for d_2-belactosin A, showed a continuous 2 Da shift in all expected fragments (Figure 27*(iii,iv)*) in comparison to unfed wild type cultures (Figure 27 *(i)*). Additionally, these MS/MS results

more reliably showed, that the compound produced by the chemically complemented knock-out strains represents the belactosin A molecule.

Subsequently, we supplemented cultures of *Streptomyces* sp. UCK14 and the three individual mutant strains *Streptomyces* sp. UCK14/Δ*belK,* UCK14/Δ*belL,* UCK14/Δ*belN* with synthetic 3,3-d_2-Acpa for chemical complementation. The generated enriched d_2-belactosin A molecule could be detected with *m/z* 372.2 $[M+2+H]^+$ and correlating retention time in LC-MS analysis as known for belactosin A from unfed cultures (Figure 26B *(i)*). In contrast to feeding with 3,3-d_2-Ncpa, when supplementing cultures with 3,3-d_2-Acpa (+ Acpa), we were able to compensate the gene deletions in all mutant strains, including the *Streptomyces* sp. UCK14/*ΔbelN* mutant. This complementation was demonstrated by the presence of a d_2-belactosin A peak for the EIC *m/z* 372.2 $[M+2+H]^+$ at wild type intensity level in all three mutants *Streptomyces* sp. UCK14/*ΔbelK,* UCK14/*ΔbelL,* UCK14/*ΔbelN*, in comparison to unfed mutant cultures (w/o Acpa) where belactosin A production was completely abolished (Figure 26B *(ii-iv)*). Furthermore, when analyzing the MS/MS spectra of the *m/z* 372 $[M+2+H]^+$ precursor mass from the 3,3-d_2-Acpa-complementation studies, revealed that all fragments found in the *Streptomyces* sp.UCK14/Δ*belK,* UCK14/Δ*belL,* UCK14/Δ*belN* extracts perfectly correlate with the predicted fragments of the d_2-belactosin A molecule produced by *Streptomyces* sp. UCK14 (Figure 28 *(ii-v)*). Additionally, they also show the expected 2 Da shift in all predicted fragments in comparison to the unfed cultures (Figure 28 *(i)*). Those results clearly support our hypothesis that BelN is the responsible enzyme to catalyze the reduction of the Ncpa to the Acpa intermediate in belactosin biosynthesis.

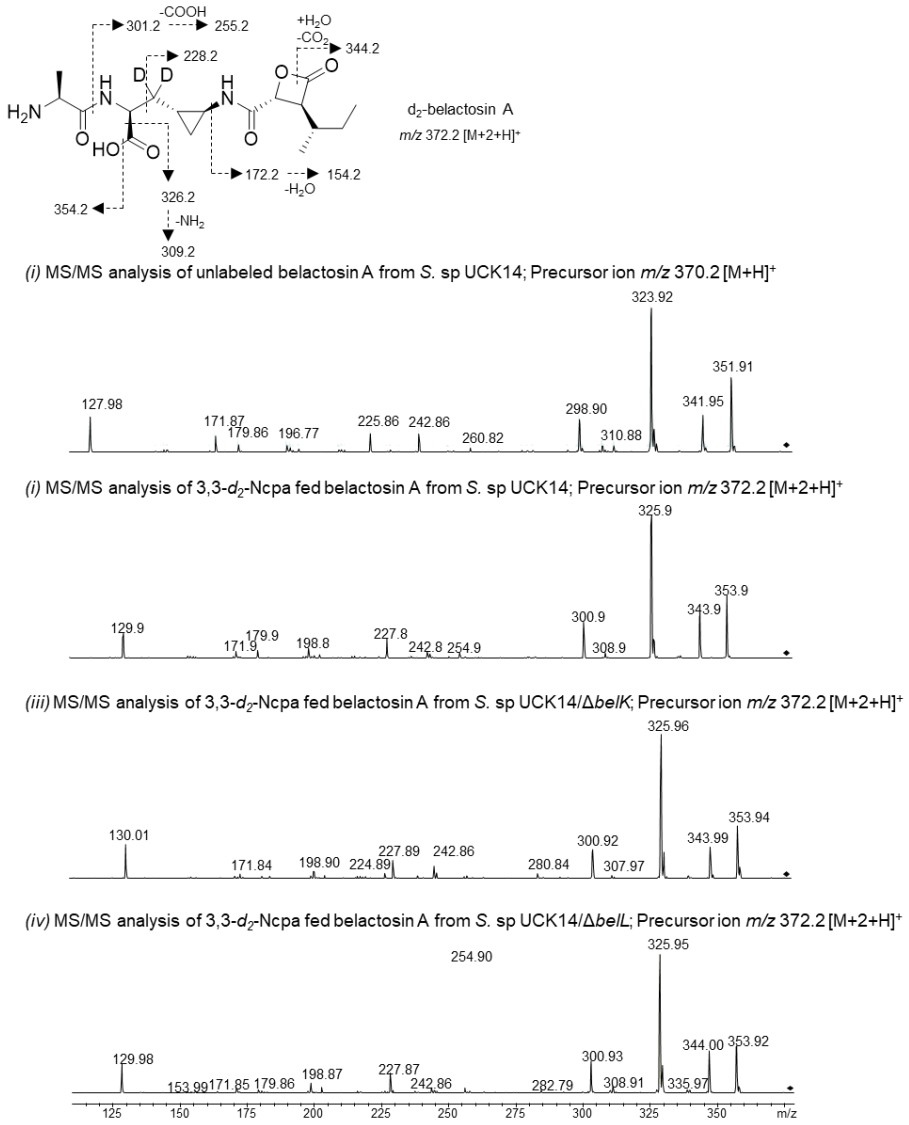


Figure 27: MS/MS analysis of chemical complementation studies with 3,3,-d_2-Ncpa: Proposed MS/MS fragments for d_2-enriched belactosin A. MS/MS profiles for belactosin A in unfed (precursor mass *m/z* 370.2 $[M+H]^+$) *(i),* and 3,3,-d_2-Ncpa fed *S.* sp UCK14 cultures (precursor mass *m/z* 372.2 $[M+2+H]^+$) *(ii),* as well as 3,3,-d_2-Ncpa fed *S.* sp UCK14/Δ*belK* (precursor mass *m/z* 372.2 $[M+2+H]^+$) *(iii*) and 3,3,-d_2-Ncpa fed UCK14/Δ*belL* (precursor mass *m/z* 372.2 $[M+2+H]^+$) *(iv)*. Precursor mass is marked with a black diamond.

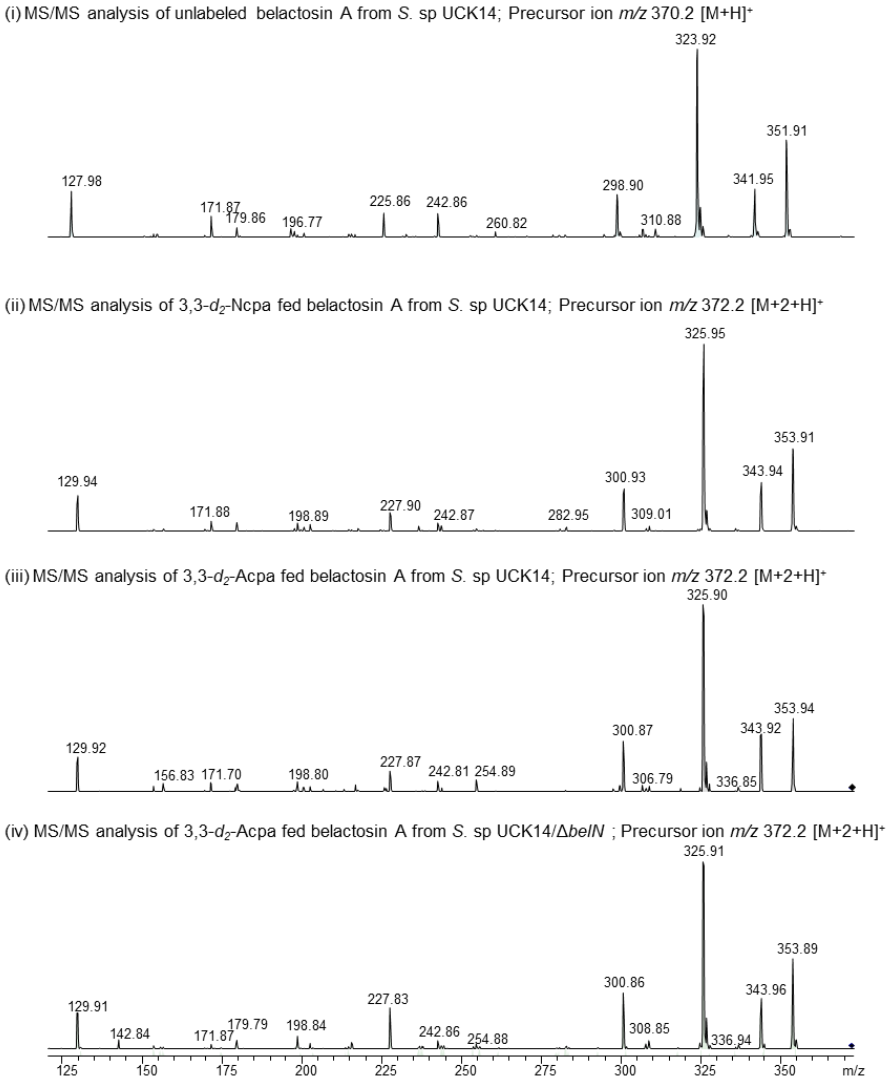


Figure 28: MS/MS analysis of chemical complementation studies with 3,3,-d_2-Acpa: MS/MS profiles for belactosin A in unfed (precursor mass *m/z* 370.2 $[M+H]^+$) *(i),* and 3,3,-d_2-Acpa fed *S.* sp UCK14 cultures (precursor mass *m/z* 372.2 $[M+2+H]^+$) *(ii),* as well as 3,3,-d_2-Acpa fed UCK14/Δ*belK* (precursor mass *m/z* 372.2 $[M+2+H]^+$) *(iii*), 3,3,-d_2-Ncpa fed UCK14/Δ*belL* (precursor mass *m/z* 372.2 $[M+2+H]^+$) *(iv)* and 3,3,-d_2-Acpa fed UCK14/Δ*belN* (precursor mass *m/z* 372.2 $[M+2+H]^+$) *(v)*. Precursor mass is marked with a black diamond.

3.7 Genetic complementation of *Streptomyces* sp. UCK14/Δ*belN*

Chemical complementation showed us, that BelN is most likely the responsible enzyme to catalyze the reduction of the Ncpa nitro group to the amine of the Acpa. To further prove the genetic integrity of the UCK14/Δ*belN* mutant, besides the chemical complementation, we decided to further conduct genetic complementation studies. Therefore, we introduced *belN* into the UCK14/Δ*belN* mutant under a constitutive promotor. A copy of the *belN* gene was introduced into the expression vector pUWL_apra_oriT (cloning procedure described in 2.17.5). The pUWL_*belN* construct was then subsequently introduced into *Streptomyces* sp UCK14 via triparental conjugation to receive the complementation strain *Streptomyces* sp UCK14/pUWL_*belN*. The production of belactosin A in *Streptomyces* sp. UCK14 could be confirmed by detection of EIC *m/z* 370.2 $[M+H]^+$ (Figure 29 *(i)*). Production levels of belactosin in *Streptomyces* sp. UCK14/*belN* without introduced genetic complementation vector (w/o pUWL_*belN*), was directly compared to belactosin A production in the complemented strain (+ pUWL_*belN).* LC-MS analysis clearly showed that the mutant was successfully complemented, as detection of a major peak for EIC *m/z* 370.2 $[M+H]^+$ (Figure 29 *(ii)*) indicated complete restoration of belactosin A production. Production levels were as high as determined in the *Streptomyces* sp UCK14 wild type producer strain (Figure 29). These results gave us additional validation that BelN seems to be responsible for catalyzing the proposed Ncpa to Acpa reduction.

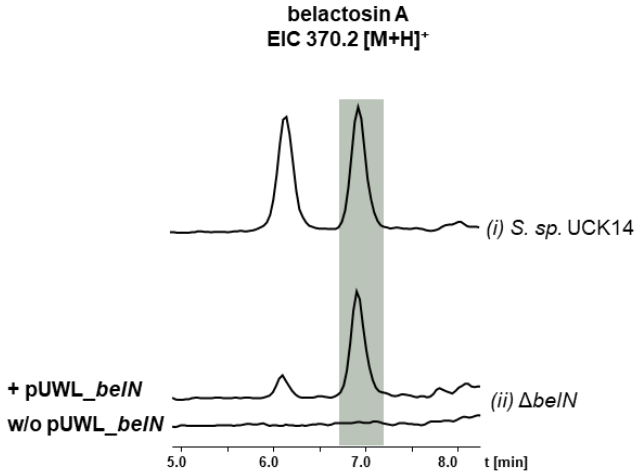


Figure 29: Genetic complementation of the UCK14/*belN* mutant: Belactosin A production in the WT producer strain *S.* sp. UCK14 (EIC *m/z* 370.2 $[M+H]^+$) *(i),* belactosin A production in the *S.* Sp. UCK14/Δ*belN* mutant before (w/o pUWL_*belN*) and after introduction of pUWL-*belN* (+ pUWL_*belN*) indicated by EIC *m/z* 370.2 $[M+H]^+$ *(ii).*

3.8 Dioxygenases BelL and HrmJ

With the previous experiments described in this thesis we could show that BelN is the responsible enzyme for the reduction of Ncpa to Acpa. This prompted us to further investigate the earlier steps towards the cyclopropyl moiety within belactosin A biosynthesis in more detail. Zeeck and co-workers performed precursor feeding studies in *Streptomyces griseoflavus* using racemic double deuterium-labeled 3,3-d_2-L-lysine and a chemical synthesized γ-lactone, in order to obtain more information of the unknown biosynthetic pathway towards the unusual amino acid 3-Ncpa building block of hormaomycin. By feeding studies they could show that Ncpa in hormaomycin derives from L-lysine. Moreover, they suggested that L-lysine might be activated at its C4 carbon atom by hydroxylation to generate 4-hydroxylysine which likely then could form a γ-lactone prior to *N*-oxygenation. This intermediate could then further be oxidized at its γ-amino group by the HrmI homolog BelK. This would allow intramolecular cyclization to form the 3-cyclopropyl containing intermediate. The by such a pathway generated Ncpa would subsequently undergo final reduction to Acpa possibly catalyzed by BelN, inconceivable via sequential two-electron reduction of the nitro group.[117, 173, 196]

Recently, it could be shown that the BelL and HrmJ homolog, GlbB from the glidobactin gene cluster is a regio- and stereo-selective lysine 4-hydroxylase.[191] Thereby BelL shows 37% and HrmJ 38 % similarity on amino acid sequence level towards GlbB (see Table 39).

Table 39: Comparison of GlbB to homologous αKG/Fe-dep. dioxygenases of the *bel* and *hrm* gene cluster

Protein (AA) (Genb. Ac. No.)	Protein homolog (AA) (Genb. Ac. No.)	ID/sim (%)	Protein homolog (AA) (Genb. Ac. No.)	ID/sim (%)	Predicted function
glbB (286) A8KCI8	BelL (229) AEO49580.1	25/37	HrmJ (227) F8S6W1	26/37	αKG/Fe-dep. dioxygenase

AA protein length in amino acid
ID amino acid identity
sim amino acid similarity

All three enzymes belong to the iron- and α-ketoglutarate-dependent (Fe/αKG) superfamily. Members of this enzyme family are capable of hydroxylating free-standing aliphatic amino acids. GlbB shows quite narrow substrate specificity, but is able to catalyze the hydroxylation of L-lysine with excellent total turnover number and complete regio- and diastereo-

selectivity.[191] Properties of these hydroxylases also demonstrated that they are α-ketoglutaric acid (α-KG) dependent and require α-KG, ferric ion, and ascorbic acid as electron donator for their maximum activity. The hydroxylation of such a C-H bond by a αKG dependent dioxygenase was also shown to take part in the biosynthesis of diverse other natural products.[202-205]

Whether BelL is able to catalyze a similar reaction by activating lysine at its C4 position to afford the proposed 4-hydroxy lysine intermediate in the biosynthesis of the Acpa moiety is a question we wanted to address and to shed light on by enzymatic biotransformation assays. The generated derivatized intermediates should then be subsequently applied in chemical complementation feeding studies with corresponding gene knock-out strains.

3.9 Expression and purification of BelL, GlbB and analysis of the potential function of BelL as a lysine 4-hydroxylase

For the biochemical investigation of the potential activation of L-lysine by hydroxylation and speculated formation of a γ-lactone prior to *N*-oxygenation (Figure 30A), the putative lysine 4-hyroxylases BelL and GlbB, genes *belL* and *glbB* were cloned into a pET28a vector for the expression as *N*-terminal His-tagged proteins and introduced into *E. coli* Rosetta2™ (DE3) pLys. The respective cloning procedures are described in 2.18.1 and 2.18.2. Protein overexpression was induced with IPTG. After cell lysis, protein with ~57.7 kDa, matching the calculated molecular mass of His_6-BelL, and 33,6 kDa, matching the calculated molecular mass of His_6-GlbB, could be obtained as soluble proteins with good purity approximately over 90 % (SDS Page, Figure 30B). After Ni^{2+} affinity chromatography fractions which contained mostly pure BelL or GlbB enzyme, were pooled individually. The final protein concentration of purified BelL was 0.65 mg/ml resulting in a total yield of 19,5 mg BelL out of 2 L of bacterial culture, and 2 mg/ml for GlbB, resulting in a total yield of 25 mg GlbB out of 2 L of bacterial culture. The obtained protein solutions were stored at -80 °C or directly buffer exchanged for further applications in biochemical in vitro characterization or biotransformation assays.

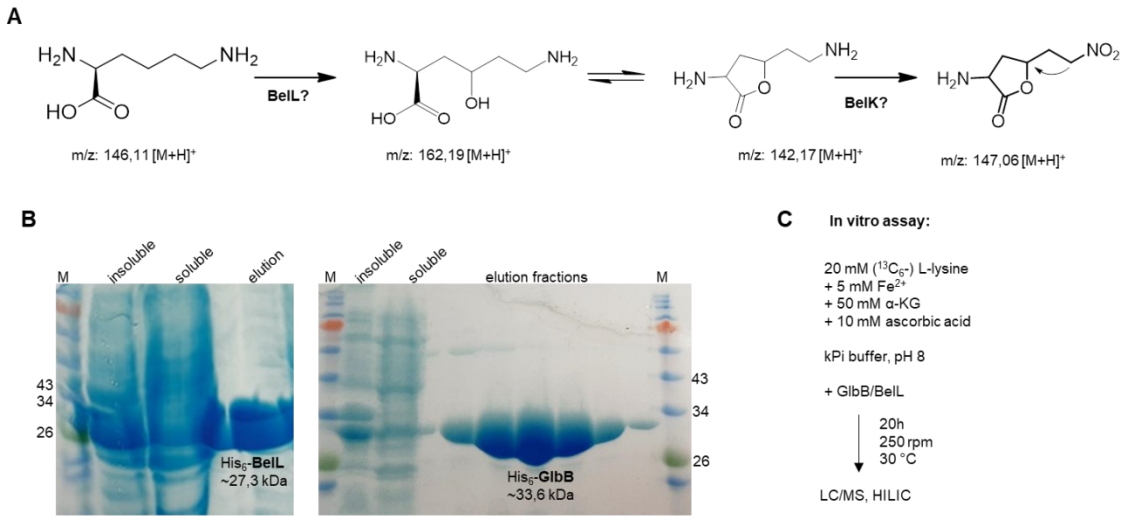


Figure 30: Biotransformation assay for the generation of a 4-hydroxy-L-lysine intermediate: A: Chemical reaction scheme for potential activation of L-lysine by hydroxylation and speculated formation of a γ-lactone prior to *N*-oxygenation. **B:** SDS-PAGE analysis of expression and purification of His8-BelL and His8-GlbB. M: protein size standard; insoluble: protein sample from the lysed cell debris; soluble: sample of the protein slurry after cell lysis; elution fractions: purified protein after IMAC. **C:** Reaction conditions of the 4-OH-Lys in vitro assay.

To investigate the enzymatic function of BelL in vitro, an enzyme assay based on the approach described by Renata et al,[191] was conducted. For that, we charged a reaction vessel with the desired amino acid L-lysine and started the reaction with the addition of BelL. Because BelL as well as GlbB, belong to the family of iron- and a-ketoglutarate-dependent dioxygenases (Fe/αKGs)[204] the assay was performed in presence of 5 mM Fe^{2+} and 50 mM α-KG, as co- factors. Ascorbic acid acts as electron donor and is known to support activity of αKGs[206, 207] and so was supplemented to the assay reaction mix in a 10 Mm final concentration. The assay was incubated under air for 20 h, and then submitted to LC-MS analysis (Figure 30C).

Chromatographic analysis and detection of amino acids in LC-MS can be burdensome. They are not directly suitable for RP-HPLC-MS analysis, due to one or more of their physicochemical properties, Chemical labeling of amino acids in means of derivatization, has been a favored method to detect small molecules and assay product outcomes in metabolomics.[208] Because we planned to use the possibly produced 4-OH-L-lysine in subsequent chemical complementation feeding studies, we aimed to detect the amino acid without any pre- or post-column derivatizations. In cooperation with Chambers Hughes we decided to analyse the polar underivatized amino acids with hydrophilic interaction chromatography (HILIC) -tandem Mass Spectrometry (MS) (as described in 2.23.7).[209, 210]

Assay conditions contained different BelL enzyme concentrations (5 and 10 µM final concentration, Figure 31 (*iii, iv*)) and a substrate control (L-lysine substrate, (Figure 31 (*i*)) as well as a negative control with heat shock inactivated BelL enzyme (hs BelL, (Figure 31 (*ii*)). However, the subsequent results of the LC-MS measurements and analysis of the TIC of all samples could not demonstrate that BelL was able to generate a conversion of the L-lysine substrate (*m/z* 147 $[M+H]^+$) to the hydroxylated 4-hydroxy-L-lysine (*m/z* 162 $[M+H]^+$), as no new peak arised. At a retention time around 7 min, only the L-lysine substrate could be detected. In this reaction conditions, BelL at a final enzyme concentration of 5 µM and 10 µM (Figure 31 (*iii, iv*)) was not able to hydroxylate the L-lysine substrate. In further attempts we tried different approaches to optimize the assay conditions to trigger enzyme activity of BelL in vitro, e.g., different pH values, buffer systems, incubation time and substrate to enzyme ratio (Data not shown). We also tested the enantiomer D-lysine instead of L-lysine as an alternative substrate in case BelL needs to catalyze a substrate with different stereochemistry (data not shown): BelL activity could not be observed in any of the mentioned approaches.

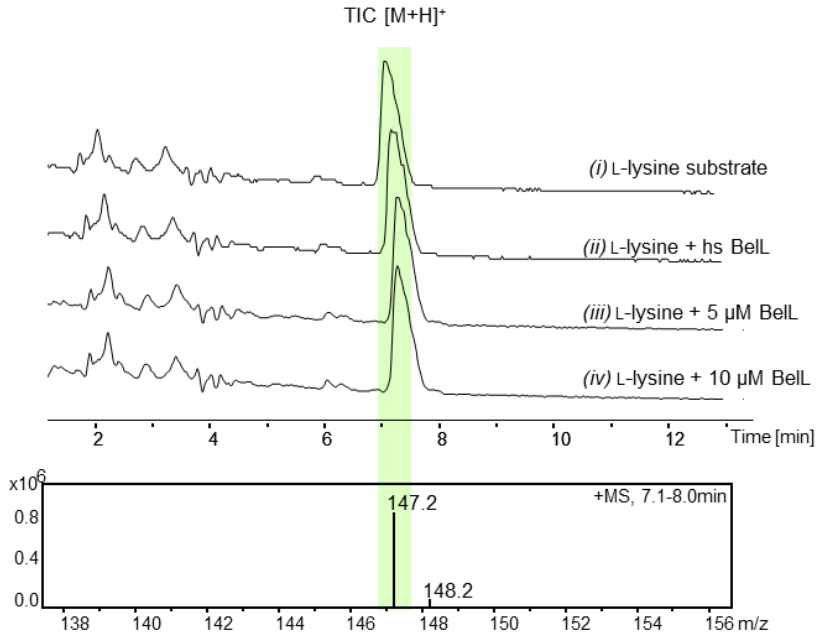


Figure 31: LC-MS analysis of BelL enzyme assay:*(i)* TIC of the substrate control without any enzyme at RT around 7 min. *(ii)* TIC of the substrate incubated with heat shocked BelL. *(iii, iv)* TIC of enzyme assay carried out with 5 µM and 10 µM BelL, respectively**. Bottom:** MS spectrum at around RT 7 min of *(iv) m/z* 147 $[M+H]^+$. hs = heat shocked

3.10 Generation of $^{13}C_6$-4-hydroxy-L-lysine by an GlbB catalyzed in vitro assay

To bypass the difficulties trying to characterize the enzyme activity of BelL in vitro and determine its role in the biosynthesis of the Ncpa intermediate, we got the idea to use an established in vitro based assay system to generate the 4-OH-Lys intermediate and subsequently feed it to the *Streptomyces* sp. UCK14 wild type and the *Streptomyces* sp. UCK/Δ*belL* mutant, in means of possible chemical complementation. In a first step, we attempted to reproduce the published in vitro conversion of L-lysine to 4-hydroxy-lysine by GlbB.[191] Employing the exact same reaction conditions as described (Figure 30C), we could obtain comprehensive data which showed the hydroxylation of L-lysine by GlbB to form OH-lysine. The conversion could nicely be followed by HILIC LC-MS analysis (data not shown). Because we planned to complement the *S.* sp. UCK/Δ*belL* mutant with 4-hydroxy-lysine and needed a robust control for the incorporation of the molecule in the belactosin A molecule, we applied the stable isotope labeled $^{13}C_6$-L-lysine as a substrate in the assay. To our satisfaction, GlbB was able to facilitate the conversion of $^{13}C_6$-L-lysine (*m/z* 153 $[M+H]^+$) to a new peak with the mass *m/z* 169 $[M+H]^+$, most likely corresponding to $^{13}C_6$-4-hydroxy-L-lysine. However, to really proof this, a reference standard should be measured or further experiments with NMR should be conducted. The TIC chromatogram in Figure 32 *(i)* shows two peaks at retention time around 7.3 and 7.7 min with only partial conversion of $^{13}C_6$-L-lysine to the hydroxylated $^{13}C_6$-4-hydroxy-L-lysine with 5 µM GlbB enzyme in the in vitro assay. Whereas 10 µM final concentration of GlbB are sufficient to complete convert the labeled lysine substrate to its hydroxylated derivative. This complete conversion can also be nicely ascertained by the loss of the substrate peak at retention time 7.3 min and the fact that only the second peak at retention time 7.7 min could be detected in the TIC (Figure 32 *(ii)*). Inferring from this we scaled up the assay volume up to 10 ml in order to generate higher amounts of $^{13}C_6$-4-hydroxy-L-lysine, to ensure enough substance for subsequent feeding, using a final enzyme concentration of 10 µM GlbB. After HILIC LC-MS analysis, samples with high conversion rates were freeze dried and submitted to further feeding studies.

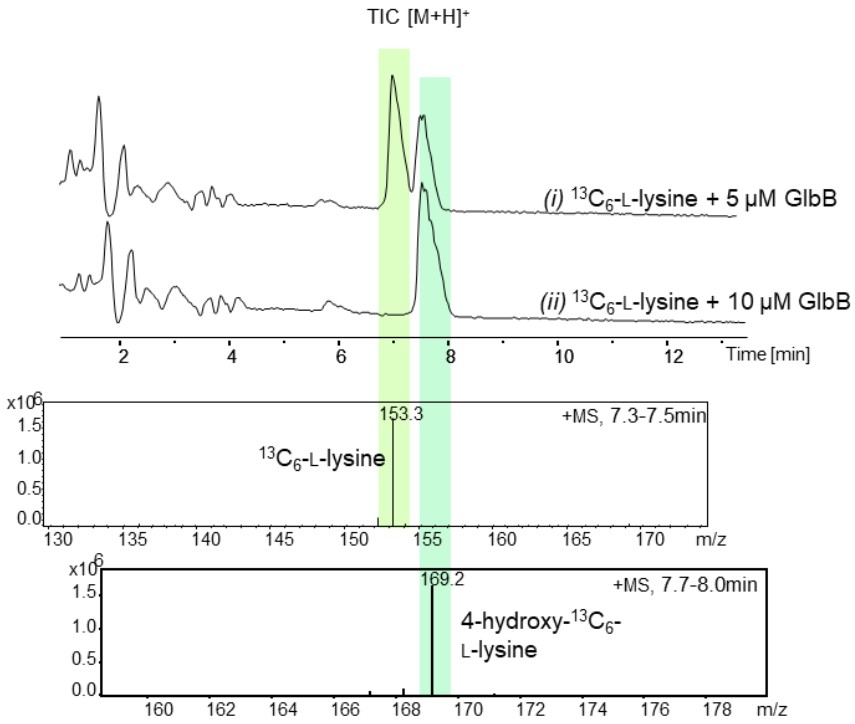


Figure 32: Enzymatic conversion assay of $^{13}C_6$-L-lysine with GlbB in vitro: TIC of enzyme assays carried out with 5 µM and 10 µM GlbB with peaks at 7.3 and 7.7 min retention time, respectively (*i.ii*). **Bottom:** MS spectra of RT 7.3 min *m/z* 153 $[M+H]^+$ for $^{13}C_6$-L-Lysine and RT 7.7 min *m/z* 169 $[M+H]^+$ for most likely $^{13}C_6$-4-hydroxy-L-Lysine.

3.11 Chemical complementation studies with $^{13}C_6$-4-hydroxy-L-lysine

After we successfully generated the GlbB biocatalyst conversion product (which will be further be reffered to as $^{13}C_6$-4-hydroxy-L-lysine), we aimed to utilize this stable isotope-labeled compound to conduct feeding studies and chemical complementation in the UCK14 producer strain and the UCK14/Δ*belL* mutant (Figure 33).

In a first step, we included a positive control of non-hydroxylated $^{13}C_6$-L-lysine, which already showed successful incorporation in previous feeding studies. As we could show before, after feeding stable isotope labeled $^{13}C_6$-L-lysine, we could detect a + 6 Da shift from belactosin A 370.2 $[M+H]^+$ to $^{13}C_6$ -labeled belactosin A *m/z* 376.2 $[M+6+H]^+$ (Figure 33 (*i, ii*)).

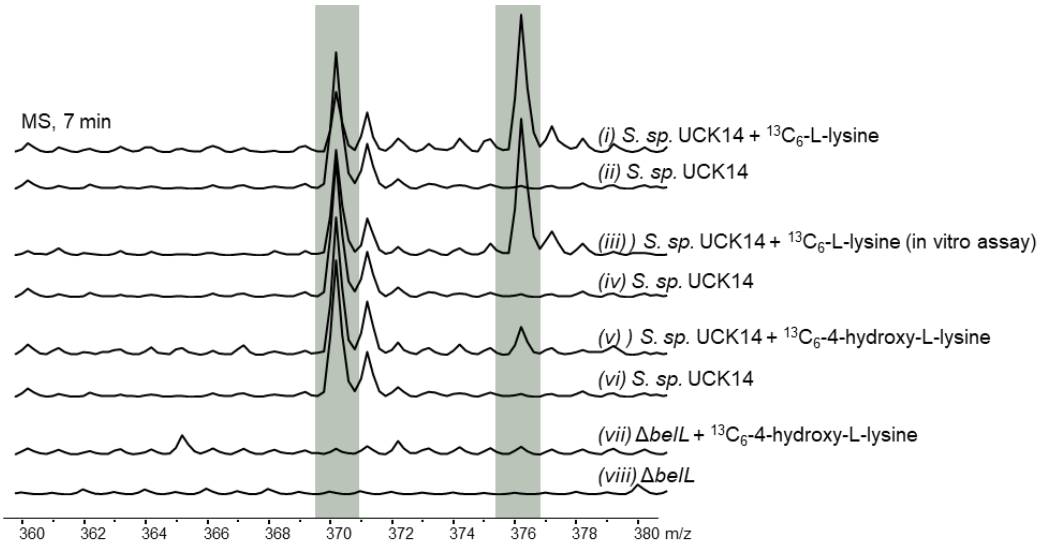


Figure 33: Chemical complementation studies with $^{13}C_6$-4-hydroxy-L-lysine: MS analysis of culture extracts at a retention time of 7 min and positive ion mode scan: (*i, ii*): Substrate control with $^{13}C_6$ labeled-L-lysine in comparison with unfed culture. (*iii, iv*): Substrate control with $^{13}C_6$-L-lysine out of the in vitro assay control in comparison to unfed culture. (*v, vi*): MS profile after feeding of $^{13}C_6$-4-hydroxy-L-lysine to *S.* sp.UCK14in comparison to unfed culture. *(vii, viii)*: MS profile of the UCK14/Δ*bel*L mutant after feeding of $^{13}C_6$-4-hydroxy-L-lysine in comparison to unfed UCK14/Δ*belL*.

When feeding the $^{13}C_6$ L-lysine substrate control without the GlbB enzyme, it also resulted in an isotopic mass shift of six from *m/z* 370.2 $[M+H]^+$ to *m/z* 376.2 $[M+6+H]^+$ and a substantial increase of the *m/z* 376.2 $[M+6+H]^+$ isotope signal (Figure 33 (*iii, iv*)). As a next step we wanted to see if feeding the generated $^{13}C_6$-4-hydroxy-L-lysine could also lead to such 6 +Da shift. This would be an indication that this substrate is accepted by the UCK14 producer strain as an intermediate and incorporated into the belactosin molecule. We could observe an *m/z*

376.2 $[M+6+H]^+$ isotope signal and a +6 Da shift. However, the detected peak was substantially smaller in comparison to the control feedings (Figure 33 (*v, vi*)). Finally, we wanted to achieve chemical complementation to restore belactosin A production in the UCK14/Δ*belL* mutant with feeding of $^{13}C_6$-4-hydroxy-L-lysine. Unfortunately, after feeding$^{13}C_6$-4-hydroxy-L-lysine to the UCK14/ΔbelL mutant we could not detect any *m/z* 376.2 $[M+6+H]^+$ isotope signal (Figure 33 (*vii, viii*)) and chemical complementation was unsuccessful with this labeled intermediate.

3.12 Chemical complementation studies with 5,5-d_2-4-hydroxylysine lactone

Unfortunately, we could not show chemical complementation of the UCK14/Δ*belL* mutant strain with 4-hydroxy-L-lysine or confirm the a biochemical function of BelL to hydroxylate lysine at its' C4, position. Whilst $^{13}C_6$-4-hydroxy-L-lysine generated by the BelL homolog GlbB seems not to be a suitable intermediate within belactosin A biosynthesis, a γ-lactone, which might be spontaneously formed from the activated amino acid 4-hydroxy-L-lysine, could be a more suitable incorporation candidate. In previous feeding studies in the hormaomycin producer strain, it could already been shown that a deuterium-labeled lactone of 4-hydroxylysine (5,5-d_2-4-hydroxylysine lactone) is an acceptable substrate in *Streptomyces griseoflavus* W-384, and thus may be an intermediate in the biosynthesis towards Ncpa.[173] Hence, we decided to conduct a chemical synthesis approach to obtain deuterium-labeled lactone for subsequent chemical complementation studies. The chemical synthesis of the lactone was kindly implemented by Chambers Hughes and Annika Esch (Interfaculty Institute of Microbiology and Infection Medicine, Microbial Bioactive Compounds, University of Tuebingen) following a protocol inspired by the former synthesis approach as described by Brandl et al.[173] Starting from the attempted preparation of 5,5-d_2-4-hydroxylysine, the respective γ-lactone was obtained in five steps with overall about 10 % yield (data not shown).

After successful synthesis of the γ-lactone, we obtained around 40 mg of racemic 5,5-d_2-4-hydroxylysine lactone. LC-MS analysis also showed that the racemic γ-lactone can also exist in an open-chain confirmation (Figure 34A, Appendix: Figure 56, Figure 57). For the lactone feeding study a 250 mM lactone stock solution, solved in ddH_2O, was prepared and 200 µl were spread on a 10 ml SM19 agar plate, resulting in 10 mM final concentration.

Most likely because of the possible low final lactone concentration in the media, analysis and detection of the signals of the LC-MS measurement in extracted ion mode for belactosin A, was not as straight forward as anticipated. So we decided to include d_2-Acpa fed samples as a positive control for comparison regarding retention time, formation of the MS isotopic peak pattern and MS/MS analysis (Figure 34B *(ii)*) as well as comparison with *Streptomyces* sp.UCK14 (Figure 34B *(i)*). First of all, we needed to determine if the 5,5-d_2-4-hydroxylysine lactone intermediate is accepted as a substrate in the *Streptomyces.* sp. UCK14 producer strain and incorporated into the belactosin A molecule.

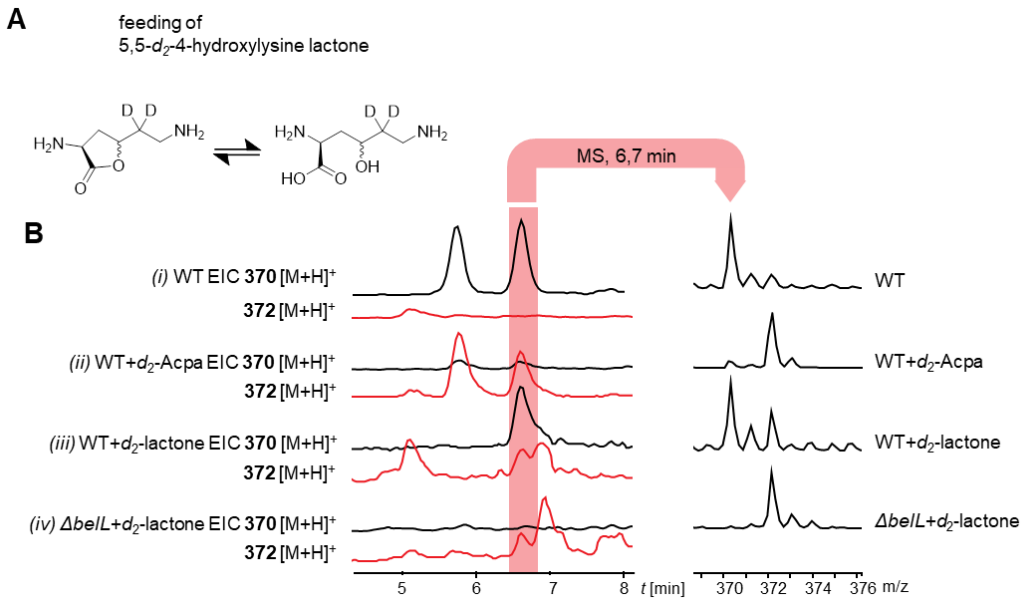


Figure 34: Chemical complementation studies with 5,5-d_2-4-hydroxylysine lactone: A: Synthetic 5,5-d_2-4-hydroxylysine lactone and synthetic 5,5-d_2-4-hydroxylysine. **B:** EIC for belactosin *m/z* 370.2 $[M+H]^+$ and double deuterium-labeled belactosin A *m/z* 372.2 $[M+2+H]^+$, as well as MS analysis for the wild type without feeding *(i)*, d_2-Acpa fed wild type *(ii)*, d_2-lactone fed wild type *(iii)* and d_2-lactone fed *S.* sp.UCK14/ΔbelL *(iv)*.

Analysis of the culture extract of *Streptomyces.* sp. UCK14 fed with the d_2-lactone showed successful incorporation of the deuterium-labeled intermediate, indicated by the detection of EIC *m/z* 372.2 $[M+2+H]^+$, at a retention time comparable to the d_2-Acpa fed belactosin A. Additionally, we could detect an isotopic mass shift in the MS analysis at a retention time of 6.7 min of + 2 Da from *m/z* 370.2 $[M+H]^+$ to *m/z* 372.2 $[M+2+H]^+$ (Figure 34B *(iii)*) for this compound. The incorporation of the d_2-lactone was also analyzed in MS/MS measurements (Figure 35). Thereby, the successful incorporation of the 3,3-d_2-Ncpa precursor into the belactosin A molecule in *Streptomyces.* sp. UCK14 could nicely be followed in MS/MS analysis, showing a +2 Da shift in all expected fragments (Figure 35 (*i,ii*)). Because we could show first hints about a possible incorporation of the d_2-lactone within the *Streptomyces.* sp. UCK14 producer strain, we further wanted to determine if feeding of the 5-d_2-4-hydroxylysine lactone could chemically complement *Streptomyces.* sp. UCK14/Δ*belL* and recover belactosin A production. The labeled lactone was fed to the bacterial cultures as described before, and the culture crude extracts were analyzed with LC-MS and MS/MS. Comparing retention times of the EICs, MS isotopic peak patterns and MS/MS profiles of the chemical complementation with d_2-Acpa and d_2-lactone in the *Streptomyces.* sp. UCK14 strain, as well as in the *Streptomyces.* sp. UCK14/Δ*belL*, showed that feeding of the 5-d_2-4-hydroxylysine lactone could at least to a certain level, restore belactosin A production in the

Streptomyces. sp. UCK14/Δ*belL* mutant. In contrast to the unfed *S.* sp. UCK14/Δ*belL* culture extracts, we could detect EIC *m/z* 372.2 $[M+2+H]^+$, as well as an isotopic peak shift of + 2 Da in the MS spectrum, after feeding the d_2-lactone (Figure 35B *(iv)*). Moreover, when analyzing the MS/MS spectra, the most abundant fragments of precursor mass *m/z* 372.2 $[M+2+H]^+$ of belactosin A (labeled in grey) were present (Figure 35 *(iii)*).

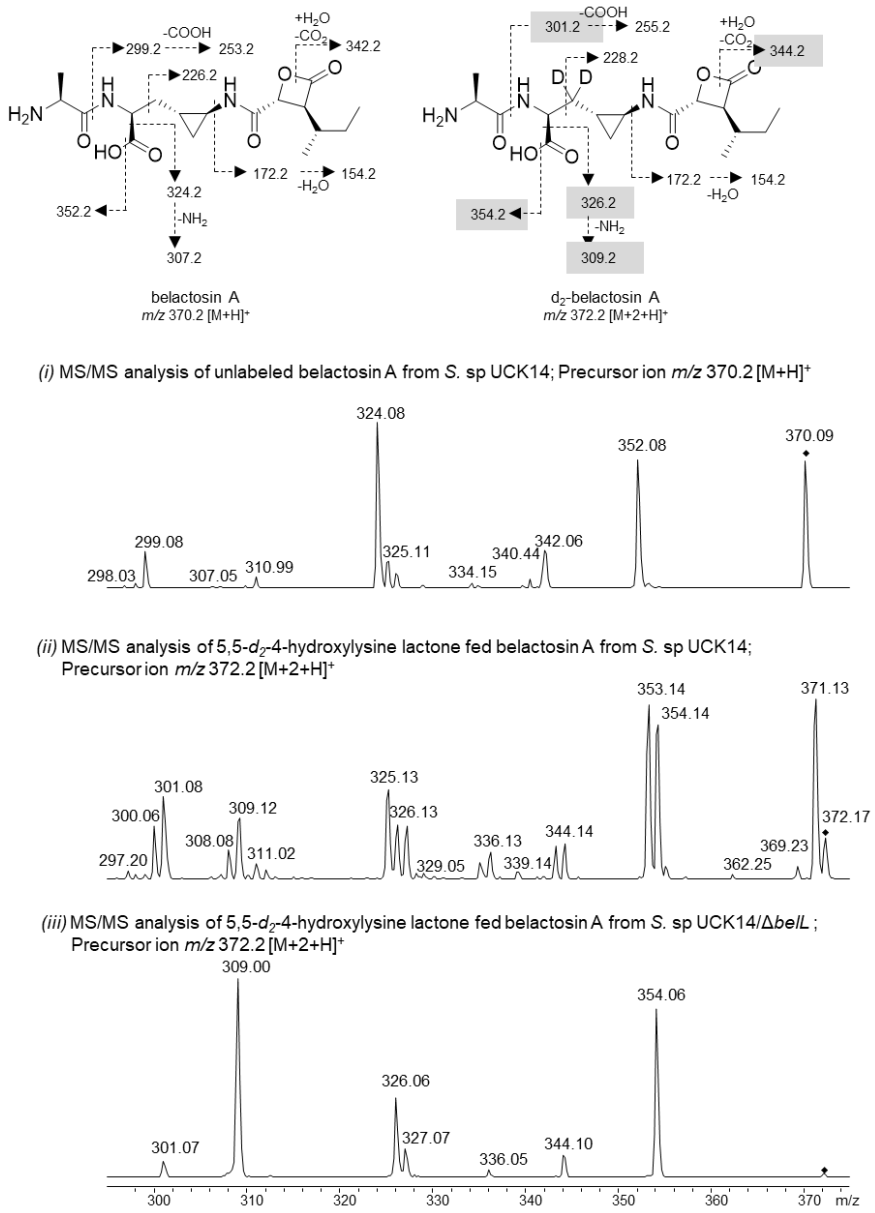


Figure 35: MS/MS analysis:MS/MS analysis of chemical complementation studies with 5,5-d_2-4-hydroxylysine lactone: Proposed MS/MS fragments for unlabeled and double deuterium-labeled belactosin A. MS/MS profiles for belactosin A in unfed (precursor mass *m/z* 370.2 $[M+H]^+$) *(i),* and 5,5-d_2-4-hydroxylysine lactone fed wild type cultures (precursor mass *m/z* 372.2 $[M+2+H]^+$) *(ii),* as well as 5,5-d_2-4-hydroxylysine lactone fed *S.* sp.UCK14/Δ*belL* (precursor mass *m/z* 372.2 $[M+2+H]^+$) Precursor mass is marked with a black diamond.

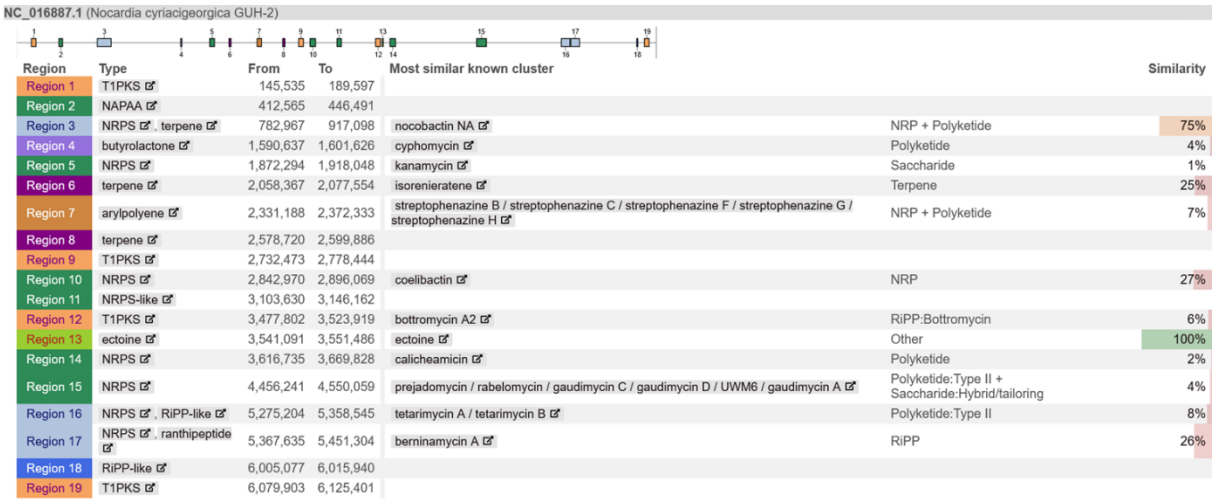
4 Results II

Results II: Studies on the role of proteasome inhibitors from *Nocardia* in pathogen survival

4.1 Analysis of the genome of *Nocardia cyriacigerogica* GUH-2 as potent producers of diverse and potential virulence associated secondary metabolites

Nocardia, which are known to comprise opportunistic human pathogens, are still an under-investigated actinobacterial genus regarding their metabolic potential. Primarily because of their clinical significance, more and more *Nocardia* genomes have been sequenced within the last decade, showing a huge potential for the production of diverse secondary metabolites. Furthermore, under-investigated bacterial taxa in general were shown to comprise prolific producers of novel bioactive chemistry.[211] A very detailed computational bioinformatics-based analysis within the genus *Nocardia* by Daniel Männle from our group revealed a plethora of putative biosynthetic gene clusters of various classes, including polyketide (PKS), non-ribosomal peptide (NRPS), and terpenoid pathways. Hereby, a maximum likelihood tree created with the web tool autoMLST, based on 63 housekeeping genes within 103 *Nocardia* strains, revealed several distinct gene cluster families.[22]

Within this thesis, in order to unravel the individual natural product production potential of the human pathogenic *N. cyriacigeorgica* GUH-2 strain, an automated secondary metabolite analysis with antiSMASH was employed. This tool thereby allows the rapid genome-wide identification, annotation, and analysis of secondary metabolite biosynthetic gene clusters by integrating and cross-linking with a large number of in silico secondary metabolite analysis tools.[85] The investigation employed in this thesis revealed 19 secondary metabolite gene clusters (Figure 36, Region 1 to 19, referred to as cluster 1.1 to 1.19) for the production of putative for example PKS, NRPS, PKS-NRPS hybrid or terpene derived natural products.



NC_016887.1 (Nocardia cyriacigeorgica GUH-2)

Region	Type	From	To	Most similar known cluster		Similarity
Region 1	T1PKS	145,535	189,597			
Region 2	NAPAA	412,565	446,491			
Region 3	NRPS, terpene	782,967	917,098	nocobactin NA	NRP + Polyketide	75%
Region 4	butyrolactone	1,590,637	1,601,626	cyphomycin	Polyketide	4%
Region 5	NRPS	1,872,294	1,918,048	kanamycin	Saccharide	1%
Region 6	terpene	2,058,367	2,077,554	isorenieratene	Terpene	25%
Region 7	arylpolyene	2,331,188	2,372,333	streptophenazine B / streptophenazine C / streptophenazine F / streptophenazine G / streptophenazine H	NRP + Polyketide	7%
Region 8	terpene	2,578,720	2,599,886			
Region 9	T1PKS	2,732,473	2,778,444			
Region 10	NRPS	2,842,970	2,896,069	coelibactin	NRP	27%
Region 11	NRPS-like	3,103,630	3,146,162			
Region 12	T1PKS	3,477,802	3,523,919	bottromycin A2	RiPP:Bottromycin	6%
Region 13	ectoine	3,541,091	3,551,486	ectoine	Other	100%
Region 14	NRPS	3,616,735	3,669,828	calicheamicin	Polyketide	2%
Region 15	NRPS	4,456,241	4,550,059	prejadomycin / rabelomycin / gaudimycin C / gaudimycin D / UWM6 / gaudimycin A	Polyketide:Type II + Saccharide:Hybrid/tailoring	4%
Region 16	NRPS, RiPP-like	5,275,204	5,358,545	tetarimycin A / tetarimycin B	Polyketide:Type II	8%
Region 17	NRPS, ranthipeptide	5,367,635	5,451,304	berninamycin A	RiPP	26%
Region 18	RiPP-like	6,005,077	6,015,940			
Region 19	T1PKS	6,079,903	6,125,401			

Figure 36: Genome analysis of strain *Nocardia cyriacigeorgica* GUH-2 with the antiSMASH web tool:The GUH-2 genome compromises 19 BGCs in total, which mostly encode for potential NRPS, PKS, Ripp-like or terpene derived natural products. This analysis also includes similarity comparisons to known compounds (Ref.seq. NCBI genome sequence: NC_016887.1). AntiSMASH analysis conducted as described in Blin et al.[85], figure is a screenshot from https://antismah.secondarymetabolites.org (search for Acc. Nbr. NC_016887.1).

Furthermore, genome mining approaches using LC-MS and NMR guided metabolic profiling of selected *Nocardia* strains and GNPS (Global Natural Product Social molecular networking) showed that biosynthetic gene cluster families (GCF) above a BiGSCAPE[212] threshold of 70% can be assigned to distinct structural features of natural products.[22] Under the lead of Daniel Männle from our research group, we analyzed this network in more detail. The analysis revealed that cluster 1.19 of *Nocardia cyriacigeorgica* GUH-2 clusters togethter with GCF14 in the BiGSCAPE similarity network. This GCF14 represents 49 strains out of the total 103 analyzed which comprise a homologous gene cluster. In our approach to link natural products of *Nocardia* to their potential virulence associated importance, this particular PKS gene cluster 1.19 of *N. cyriacigeorgica* GUH-2 sparked our interest. It did so, because this BGC encodes a putative acyl-CoA dehydrogenase (ACAD) with homology to the epoxyketone synthase from the proteasome inhibitor-associated biosynthetic pathways of epoxomycin. A genome mining guided analysis by the Kaysser group in collaboration with Schorn et al, using the ACAD gene EpnF from the eponemycin-BGC as a specific probe, revealed that the ACAD gene is highly conserved in many homologous gene clusters from different bacterial strain backgrounds. These identified strains also show involvement in biosynthesis of epoxyketone containing compounds.[168-170]

The so far orphan ACAD containing gene cluster of *N. cyriacigeorgica* GUH-2 (following named ACAD cluster), encodes for a hybrid NRPS/PKS biosynthesis machinery and a highly conserved acyl-CoA dehydrogenase enzyme (ACAD) (Figure 37). The ACAD gene cluster architecture shows similarity to the biosynthetic gene clusters of the two epoxyketone proteasome inhibitors epoxomicin (epx) and eponemycin (epn), which were identified in the actinobacterial producer strains ATCC 53904 and *Streptomyces hygroscopicus* ATCC 53709, respectively.[213]

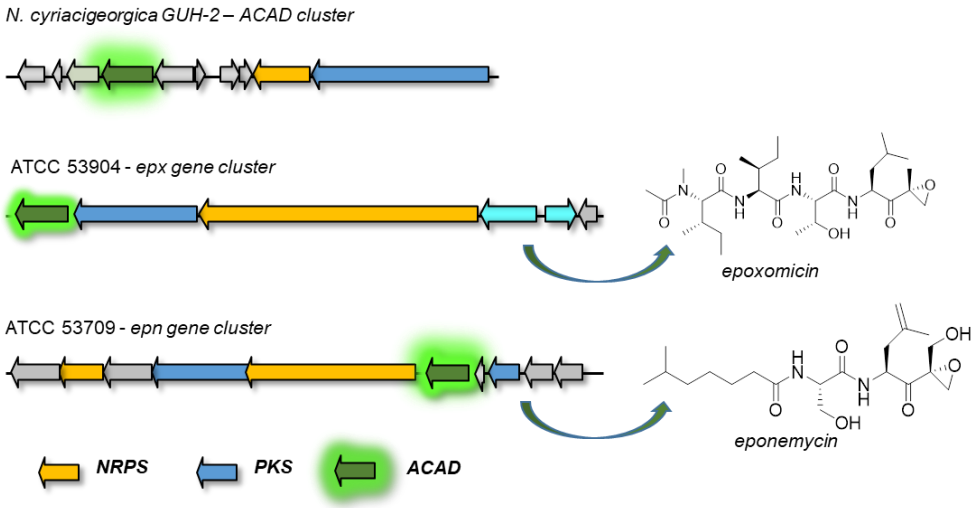


Figure 37: Biosynthetic gene clusters of epoxomicin and eponemycin: All three gene clusters encode for a hybrid NRPS/PKS system and a highly conserved ACAD gene. ACAD = acyl-CoA dehydrogenase

Proteasome inhibitors of the epoxyketone-family are an established class of therapeutic agents, and their unique epoxyketone pharmacophore allows specific binding to the catalytic proteasome β-subunits of the proteasome. Additionally, Liu et al. could show that only the putative ACAD enzyme in combination with the hybrid NRPS/PKS system is required to build the epoxyketone warhead.[214] The formation of electrophilic groups thereby in general play an essential role for the production of potent small-molecule protease and proteasome inhibitors.[46] Following we speculated that the ACAD enzyme from cluster 1.19 might modify a CoA- or ACP-activated fatty acyl product of the encoded PKS. Furthermore, Beaman and co-workers could show that *N. cyriacigeorgica* GUH-2 cells culture filtrates as well as living cells are able to inhibit proteasome activity and induce apoptosis in cells by a yet unidentified secreted small-molecule.[172, 215] However, if this activity correlates with the product from cluster 1.10 will have to be determined.

In this work we further analyzed the prevalence of this highly conserved ACAD gene in other *Nocardia* genomes. This analysis showed the wide distribution of such PKS/NRPS system in the presence of an ACAD gene in the most prevalent human pathogenic *Nocardia* strains (Figure 38). Notably, those gene clusters are in many cases co-located to a putative invasion associated gene (Figure 38, highlighted in red). Interestingly, there are reports of such invasion gene associated compounds to be involved in cell surface organization, adhesion, and internalization in epithelial host cells and thereby playing an important part in virulence associated mechanisms.[216-218] However if the ACAD compound fulfills such a virulence associated role has to be proven.

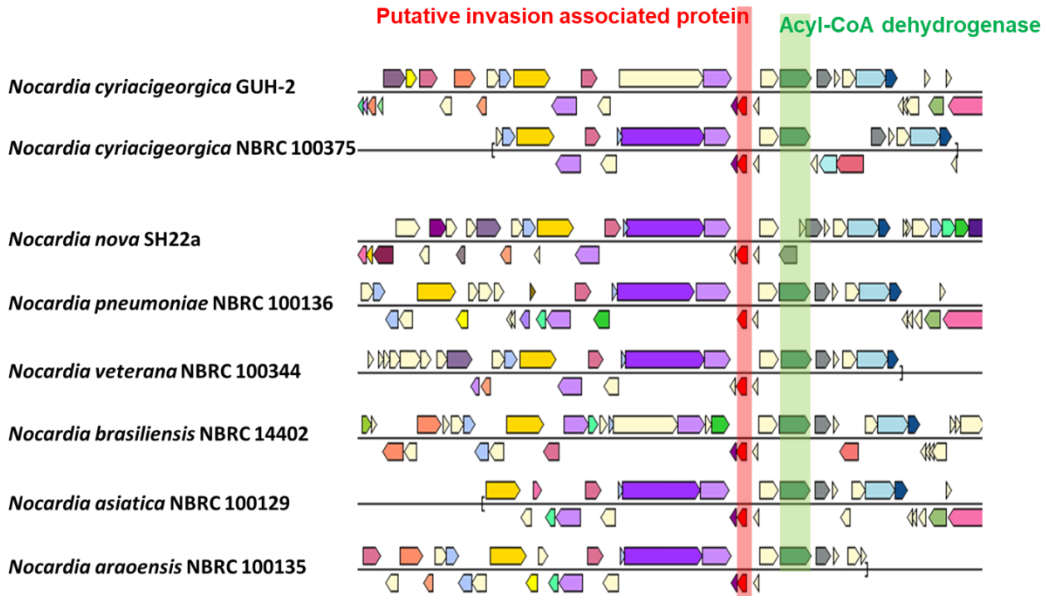


Figure 38: Prevalence of the ACAD cluster in human pathogenic *Nocardia* strains:A selection of human pathogenic *Nocardia* strains harboring the ACAD biosynthetic gene cluster containing the acyl-CoA-dehydrogenase gene (depicted in green), as well as the associated putative invasion associated protein (depicted in red).

In this work we want to address if the ACAD cluster from *N. cyriacigeorgica* GUH-2 is producing a novel compound with a putative epoxyketone warhead and further more if this compound may play a role in the pathogenicity of this *Nocardia* strain. In order to solve this challenges, we first aimed to identify the clone that carries the complete biosynthetic gene cluster from a fosmid based genomic library in order to heterologously express the complete cluster in different *Streptomyces* host systems for cultivation and further purification, as well as characterization of the new molecule. Building on that, we planned to establish eukaryotic cell guided assays to evaluate the virulence mechanism of this putative virulent compound.

4.2 The ACAD biosynthetic gene cluster: Construction and screening of a genomic fosmid library

In order to characterize and identify the ACAD biosynthetic gene cluster and the putative new epoyketone compound, we first aimed to isolate the BGC on a fosmid, modify it accordingly, to allow following heterologous expression in a bacterial host system. A pCC1FOS-based genomic library representing the complete genome of *Nocardia cyriacigeorgica* GUH-2 was constructed. Accordingly, genomic DNA *N. cyriacigeorgica* GUH-2 was isolated, processed for the genomic library as described in 2.22.2 and comprised 2069 clones in total. Perfectly matching primers were designed to screen for the up and down stream borders, as well as the ACAD gene of the BGC by PCR (Figure 39A). Hereby, in a first screening round, 48er DNA pools were screened. This screening approach revealed in total four DNA pools that contain the desired fosmid carrying the desired ACAD gene cluster, pool 13B, 20A, 20B and 21A, respectively. The positive hits were indicated by a positive PCR product for the screening primer pairs ACAD-cluster-left_f/r and ACAD-cluster-right_f/r (Figure 39B). Finally, we selected plate 21A to identify the ACAD BGC-containing fosmid by colony PCR using the same primer pairs as mentioned before. Finally fosmid 21C6 could be identified and end sequencing revealed that the fosmid insert contains genes ranging from NOCYR_5477 to NOCYR_5500 (Table 40), including the genes from position 6098,190 kb to 6120,007 kb from the *N. cyriacigeorgica* GUH-2 genome. The resulting fosmid carries a 30,8 kb insert with an average GC content of 68.4% (Figure 39C).

After computational gene cluster border analysis, using Blast alignments as well as gene cluster comparisons, we decided to assign genes NOCYR_A to NOCYR_M, as already suggested by the antiSMASH prediction software, to the biosynthetic gene cluster for the production of the putative ACAD compound (Table 35). However, if neighboring genes or genes with putative function up and downstream of the predicted cluster may play a role in the biosynthesis of the compound would need to be evaluated by border cluster gene knock-out studies.

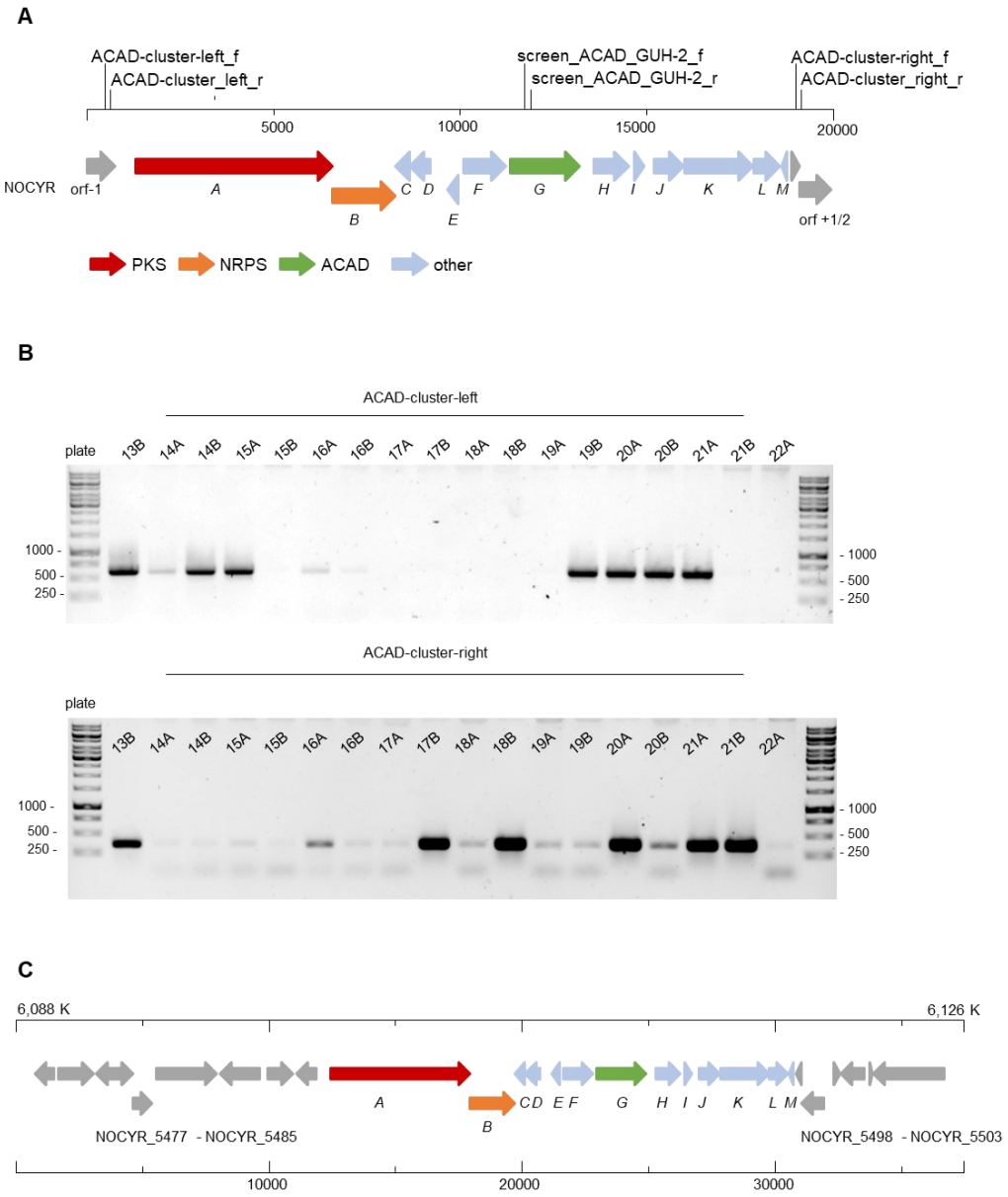


Figure 39: Screening of a fosmid library of N. cyriacigeorgica GUH-2 for the ACAD-cluster: A:Screening approach of the fosmid library with PCR primer pairs on the up and down stream border of the ACAD cluster, as well as on the ACAD gene of N. cyriacigeorgica GUH-2. B: PCR screening of 48er DNA pools with positive hits on pools 13B, 20A, 20B and 21A for both up- and downstream (ACAD cluster left and –right, respectively). C: Isolated and identified by sequencing fosmid 21C6 carries an insert size of 30.8 kb and contains genes ranging from NOCYR_5477 to NOCYR_5500.

Table 40: Gene annotation table for all genes present on fosmid 21C6:

Gene	Size (AA)	Protein homolog (Genb. Ac. No.)	ID/sim[a]	Predicted function
*NOCYR*_5477	457	*Nocardia brasiliensis* (SUB10494.1)	68/80	ubiquinone biosynthesis protein
*NOCYR*_5478	200	*Nocardia beijingensis* (WP_195043952.1)	65/75	putative lypase
*NOCYR*_5479	261	*Nocardia amikacinitolerans* (SNY87418.1)	64/76	2OG-Fe(II) oxygenase
*NOCYR*_5480	255	*Nocardia goodfellowii* (WP_209885494.1)	70/79	oxidoreductase
*NOCYR*_5481	793	*Nocardia anaemiae* (WP_062981648.1)	83/91	conserved membrane protein of unknown function
*NOCYR*_5482	523	*Nocardia aurea* (WP_109523011.1)	90/95	acyl-CoA synthetase
*NOCYR*_5483	326	*Nocardia abscessus* (WP_195100835.1)	96/97	nitronate monooxygenase
*NOCYR*_5484	262	*Nocardia vinacea* (WP_051182702.1)	80/85	hydrolase (fragment)
*NOCYR*_5485 *NOCYR*_**A**	1833	*Nocardia puris* (WP_195125031.1)	52/63	modular polyketide synthase
*NOCYR*_5486 *NOCYR*_**B**	589	*Nocardia vinacea* (WP_051182704.1)	56/65	non-ribosomal peptide synthetase NRPS5-4-3 (fragment)
*NOCYR*_5487 *NOCYR*_**C**	134	*Nocardia brasiliensis* (WP_195091945.1)	91/94	heat shock protein
*NOCYR*_5488 *NOCYR*_**D**	197	*Nocardia asteroides* (WP_022565958.1)	69/78	invasion-associated protein invasion associated endopeptidase
*NOCYR*_5489 *NOCYR*_**E**	99	*Nocardia amamiensis* (WP_067463019.1)	81/98	conserved protein of unknown function
*NOCYR*_5490 *NOCYR*_**F**	394	*Nocardia fusca* (WP_063125815.1)	95/97	ATP-binding protein
*NOCYR*_5491 *NOCYR*_**G**	651	*Nocardia anaemiae* (WP_062981751.1)	89/94	putative acyl-CoA dehydrogenase (ACAD)
*NOCYR*_5492 *NOCYR*_**H**	332	*Nocardia tenerifensis* (WP_040740175.1)	78/85	putative non-heme chloroperoxidase
*NOCYR*_5493 *NOCYR*_**I**	96	*Nocardia paucivorans* (WP_040793067.1)	86/93	protein of unknown function
*NOCYR*_5494 *NOCYR*_**J**	277	*Nocardia amamiensis* (WP_195127503.1)	89/95	conserved membrane protein of unknown function
*NOCYR*_5495 *NOCYR*_**K**	643	*Nocardia pneumoniae* WP_040784950.1	91/95	succinate dehydrogenase flavoprotein subunit
*NOCYR*_5496 *NOCYR*_**L**	289	*Nocardia brasiliensis* WP_167465999.1	92/97	fumarate reductase iron-sulfur subunit
NOCYR_5497 *NOCYR*_**M**	52	*Nocardia arizonensis* WP_054815263.1	84/93	protein of unknown function
*NOCYR*_5498	77	*Nocardia donostiensis* WP_077116011.1	83/89	protein of unknown function
*NOCYR*_5499	296	*Nocardia anaemiae* WP_062981606.1	72/87	putative DNA-binding protein
*NOCYR*_5500	74	*Nocardia inohanensis* WP_067826245.1	85/90	protein of unknown function

[a] amino acid sequence homology [%] from blastp analysis
AA protein length in amino acid
ID amino acid identity
Sim amino acid similarity

4.3 Heterologous expression of the ACAD bionsynthetic gene cluster in *Streptomyces albus* J1087 and *Amycolatopsis japonicum* M6417-CF17

Heterologous pathway expression allows straight forward expression of a complete gene cluster of interest in a heterologous host system. This technique comes in very useful to confirm the identity of gene clusters and enables easy genetic engineering of the pathway in a robust and well-established *E. coli*-based system.[219]

In order to investigate whether the genes found on fosmid 21C6 are encoding for the biosynthesis of a novel compound, we intended to express the fosmid heterologously. Furthermore, this heterologous expression system provides a biotechnological platform with a S1 biosafety level for better production and isolation conditions of secondary metabolites from the human pathogen and S2 organism *N. cyriacigerogica* GUH-2. For this purpose, the chloramphenicol resistance cassette (cat) on the backbone of 21C6, isolated from *E. coli* EPI300™ originating from the genomic library of *N. cyriacigeorgica* GUH-2, was replaced with an int_neo integration cassette using Red/ET-mediated recombination in *E. coli* BW25113 (Figure 40).[178, 220]

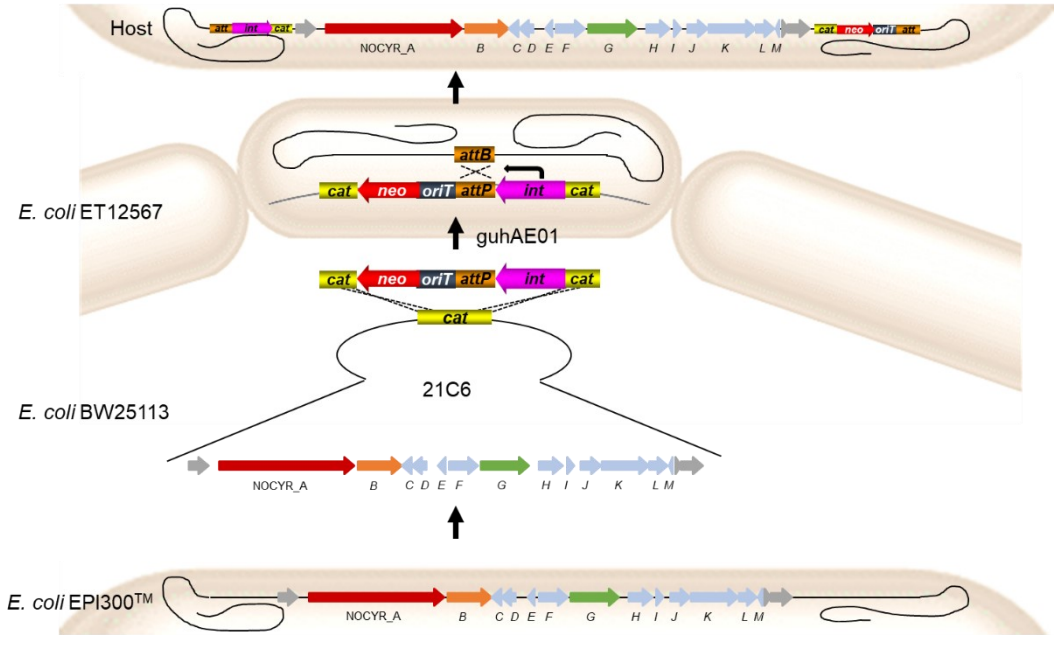


Figure 40: Generation of heterologous expression mutant strains: The fosmid guhAE01, isolated from *E. coli* EPI300™ carries the complete ACAD GUH-2 BGC and was refurbished in *E. coli* BW25113 by Red/ET mediated recombination. The cat gene was exchanged with the int_neo cassette (attP, int, neo, oriT). After transformation into *E. coli* ET12567, triparental conjugation into the host and successful integration between the two attachment sites attP and attB, positive mutants were selected by their kanamycin resistance.

This cassette contains the attP attachment site, the integrase gene (int) of phage ΦC31 that allow stable and site-specific chromosomal integration in the host strain as well as a kanamycin resistance gene (neo) and an origin of transfer (oriT). The refurbished fosmid guhAE01 was then transferred to the heterologous hosts by intergeneric conjugation. Positive heterologous expression mutants were selected by their kanamycin resistance (Figure 40). For a more detailed cloning description, refer to 2.22.4.

In a first attempt we introduced the generated fosmid guhAE01 into the well-established heterologous host strain *S. albus* J1074 and three individual kanamycin resistance clones were selected for ongoing cultivation. Extracts of cultures of the *S. albus* J1074 wild type without an introduced cluster and the mutant strains were applied to LC-ESI-MS (Figure 41A). When comparing the total ion chromatogram (TIC) traces of the heterologous host, *S. albus* J1074 without an introduced cluster (*S. albus* w/o cluster, Figure 41A *(iii)*) and the TIC after introduction of the cluster *S. albus* J1074/ ACAD GUH-2 (*S. albus* + cluster, Figure 41A *(i)*), we noticed the appearance of an additional new peak in the heterologous mutant at a retention time of around 10 min. The new peak displayed a mass of *m/z* 487.1599 $[M+H]^+$ in high resolution (HR) LC-MS (Figure 41A *(ii)*). Moreover, the new peak, possibly representing a putative novel compound with EIC *m/z* 487.1599 $[M+H]^+$ was not detectable in the host strain without an introduced ACAD GUH-2 cluster (Figure 41A *(iv)*).

Equivalently, we also decided to heterologously express the ACAD GUH biosynthetic gene cluster in the host strain *Amycolatopsis japonicum* M6417-CF17. We chose this host, as *Amycolatopsis* is phylogenetically quite closely related to *Nocardia* and was previously shown to be a good heterologous expression strain for the production of *Nocardia* derived natural products.[186, 221] After introduction of the gene cluster into the host, as described above, we analyzed culture extracts by LC-MS. In comparison to the wild type host strain *A. japonicum* M6417-CF17 (*A. japonicum* w/o cluster, Figure 41B *(iii)*) we could detect the appearance of an additional new peak in the heterologous expression strain *A. japonicum* M6417-CF17/ ACAD GUH-2 (*A. japonicum* +cluster), Figure 41B *(i)*) at a retention time of around 10 min. This novel peak also revealed an EIC *m/z* 487.1599 $[M+H]^+$ (Figure 41B *(ii)*), analogous to the expression of the gene cluster in *S. albus* J1074. Additionally, the control strain *A. japonicum* w/o cluster did not show the presence of *m/z* 487.1599 $[M+H]^+$ at a retention time of 10 min (Figure 41B *(iv)*).

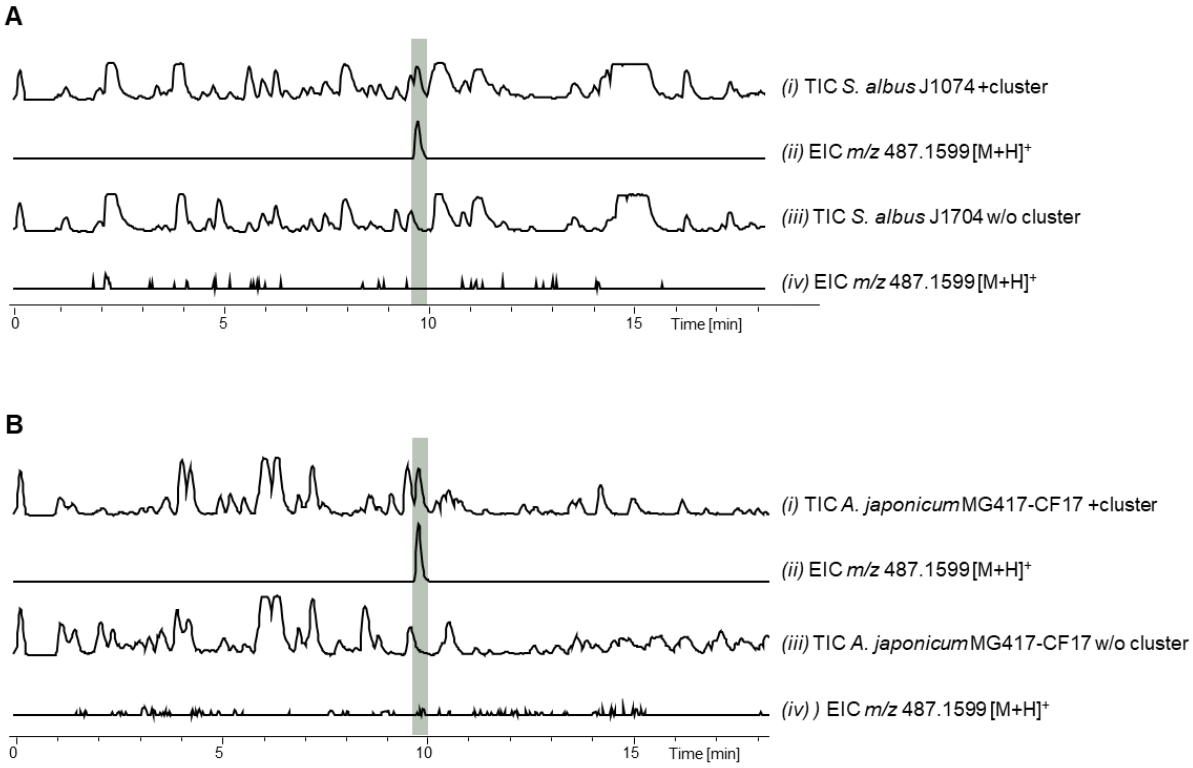


Figure 41: Heterologous expression of the ACAD gene cluster:A: Expression of the ACAD GUH-2 gene cluster in *S. albus* J1704: *(i)* TIC of culture extracts from *S. albus* J1074 with introduced cluster (+cluster), *(ii),* EIC *m/z* 487.1599 $[M+H]^+$ of *(i)*, *(iii)* TIC of culture extracts from *S. albus* J1074 without the cluster, (iv) EIC *m/z* 487.1599 $[M+H]^+$ of *(iii)*. TIC = total ion chromatogram; EIC = extracted ion chromatogram. **B:** Expression of the ACAD GUH-2 gene cluster in *A. japonicum* M6417-CF17: *(i)* TIC of culture extracts from *A. japonicum* M6417-CF17 with introduced cluster (+cluster), *(ii)* EIC *m/z* 487.1599 $[M+H]^+$ of *(i)*, *(iii)* TIC of culture extracts from *A. japonicum* M6417-CF17 without the cluster, (iv) EIC *m/z* 487.1599 $[M+H]^+$ of *(iii)*. TIC = total ion chromatogram; EIC = extracted ion chromatogram S. = *Streptomyces;* A. = *Amycolatopsis, w/*o = with out

In order to get a first insight about the structural composition and if the expressed gene cluster really encodes for a novel compound, we conducted first computational analysis. Using smartFormula (Bruker Analysis), a unique automated tool for molecular formula generation capabilities through combining accurate mass and isotopic pattern information in MS and MS/MS spectra, several chemical formulas were calculated (Appendix, Table 41). First of all, we conducted an automated database search for the proposed chemical formulas (In collaboration with Martin Konnerth, AG Grond, University of Tübingen). But anyway, our search in established natural product libraries like Dictionary of Natural Products® (DNP)[222] or MarinLit[223], did not led to any known characterized compounds.

The structural formula #1 (Appendix, Table 41) showed a hit for suggested total formula in the Dictionary of Natural Products for the NP aquayamycin.[224] However, this hit also stated

that aquayamycin has a UV absorption maximum at 430 nm, which we could not correlate for our ACAD GUH-2 component (Data not shown).

Using in-silico fragmentation approaches of structure candidates using web tools like MetFrag as well as analyzing the HR MS/MS spectra of this compound to get a first insight into the structural architecture of this novel compound did not lead to satisfactory results, as the ACAD compound, at least in this standard analysis setting, fragmented insufficiently. Following, we decided to upscale the fermentation production, as well as purify this compound in order to elucidate its structure with NMR.

4.4 Purification of the putative new ACAD GUH-2 compound

In a first attempt for purification of the ACAD GUH-2 compound, it was necessary to up scale the culture production volume in order to be able to obtain higher production yields out of the culture crude extract for downstream purification steps. In this context, we increased the culture volume of *S. albus* J1074/ ACAD GUH-2 (*S. albus* +cluster) from 50 ml to 2 L R5 production medium and subsequently analyzed the production of the ACAD GUH-2 compound by LC-MS. The measurement showed, that in the higher culture volume of 2 L the compound was still produced as seen in the 50 ml standard cultivation, regarding the composition of the BPC chromatogram, as well as retention time and intensity of the EIC 487.1599 $[M+H]^+$ of the produced novel compound (data not shown). Following on, in total a culture volume of 6 L (3 times 2 L cultures) was cultivated, as described in 2.22.5, extracted with one volume of EtoAc and evaporated under reduced pressure. In the first cultivation round, 5,43 g crude extract were isolated and further process in VLC on a C18 RP column (Figure 42A). All VLC fractions were evaporated and analyzed with LC-MS for the presence of the ACAD compound. The ACAD GUH-2 compound eluted under the selected conditions after 10 minutes, as seen in the *S. albus* J1074/ ACAD GUH-2 (*S. albus* +cluster) control (Figure 41A *(i)*) and could be detected by its EIC 487.1599 $[M+H]^+$. Fractions with gradients ranging from 50 % MeOH to 100 % MeoH contained the ACAD GUH-2 compound with varying proportions of impurities (Figure 42A,Figure 43).

After the first VLC purification run we decided, to further continue the purification porcess with fractions C and D from the 70 % and 80 % MeOH VLC fractions with 220 mg and 280 mg dry weight, respectively. Taken into account the amount of starting material of 500 mg and maximum injection concentration of 3 mg/ml on a semi preparative column, we decided to conduct a preparative RP-HPLC guided purification step as described in 2.23.5. Therefore, fraction C and D were loaded on a Unisol C18 21.2x250 mm RP-HPLC column on a waters instrument. Applying a linear gradient from 30 % to 100 % acetonitrile, fractions were manually collected every 2 min. After drying the fractions down, we analyzed them with LC-MS (Figure 44). This analysis showed that the ACAD GUH-2 compound could be detected from the collected fraction min 18 to fraction min 28 at a retention time of around 10 min as seen in the control strain, *S. albus* J1074/ ACAD GUH-2 (*S. albus* +cluster, ctrl, Figure 44).

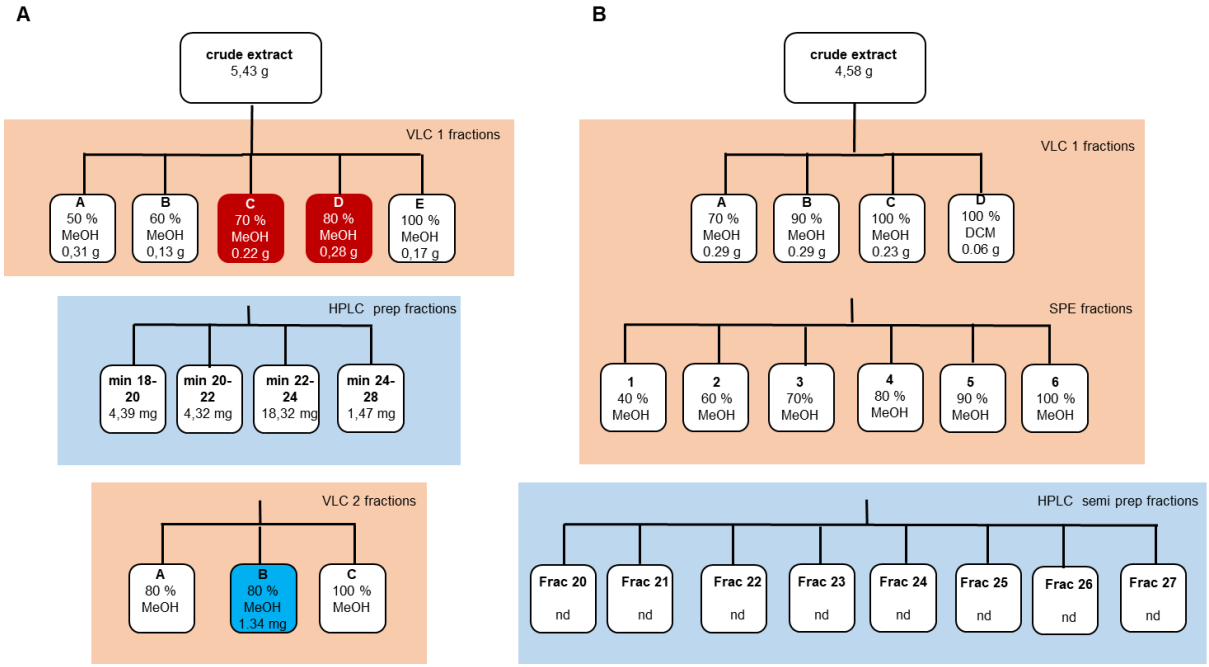


Figure 42: Schematic overview of the purifcation processes of the ACAD compound: A: In the first purification round the crude extract was further processed and purified using VLC, resulting in fractions A to E, containing the ACAD GUH-2 compound. In the next step VLC fractions C and D were purified in an preparative RP-HPLC approach. The ACAD GUH-2 compound was eluted and collected at retention time min 18 to 28, in 2 min fractions. After preparative RP-HPLC all four fractions were subjected to an another round of VLC, resulting in fraction A to C at 80 % and 100 % MeOH, respectively. Fraction B was further analyzed in NMR. **B:** In a second purification attempt the curde extract was processed in VLC, followed by SPE. Those SPE fractions were finally purified using semi preperative RP-HPLC. VLC = vaquum liquid chromatography, SPE = solid phase extraction, nd = not detectable

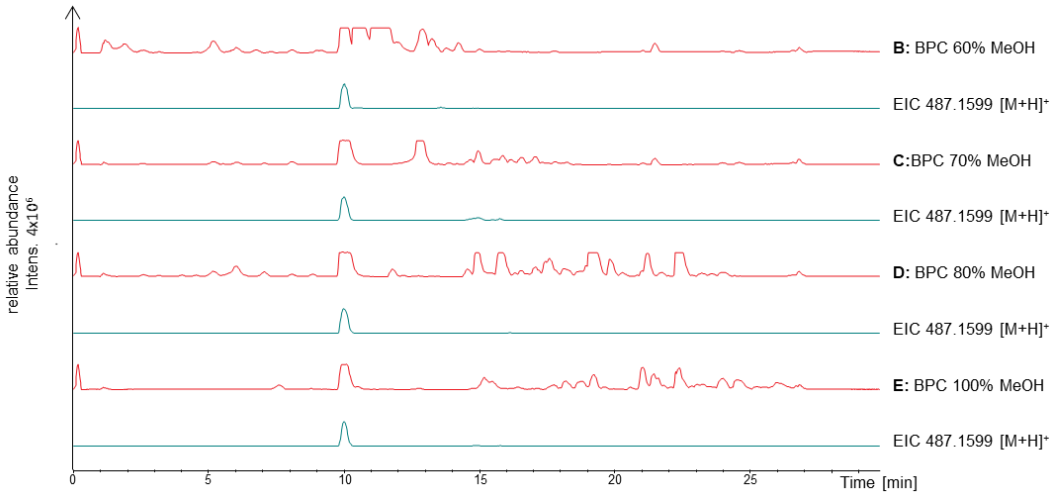


Figure 43: Fractions of the first VLC purification step of the ACAD compound:VLC fractions B to E, containing the ACAD GUH-2 compound. BPC chromatograms of extracts from 60 % to 100 % MeOH fractions and their belonging EIC 487.1599 $[M+H]^+$.

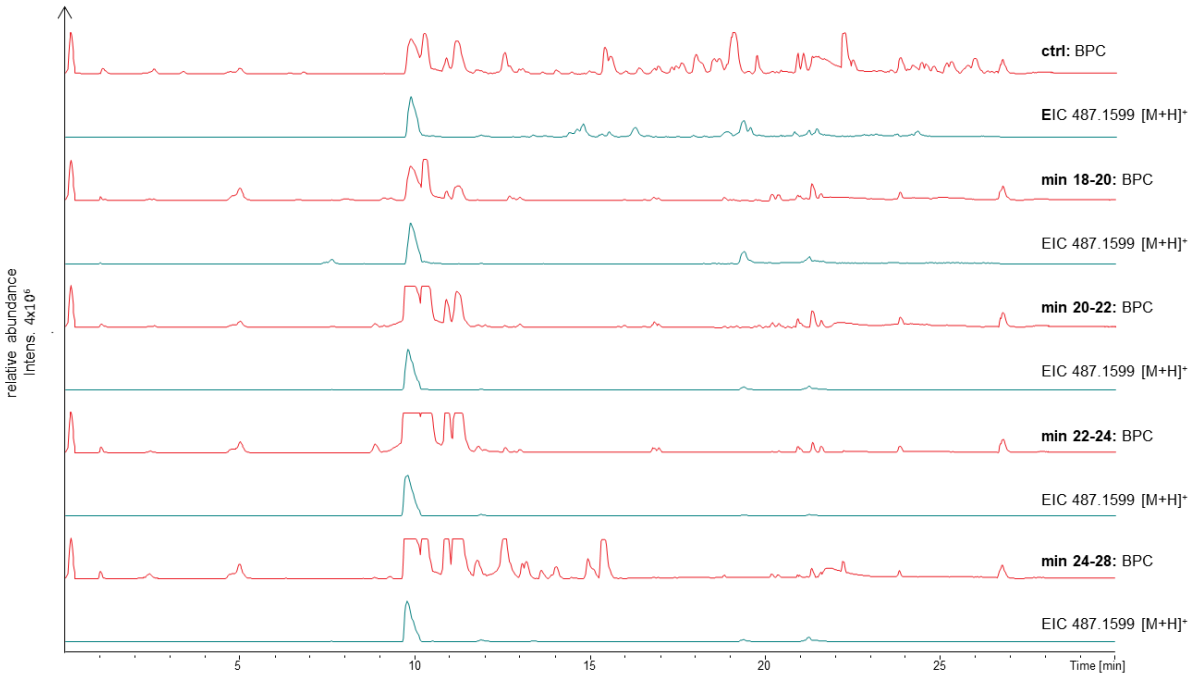


Figure 44: Fractions of the preparative purification of the ACAD compound: Preparative RP-HPLC fractions manually collected in 2 min time intervalls from min 18 to 24 and min 24 to min 28, as well as the curde extract of the control strain *S. albus* J1074/ ACAD GUH-2 (ctrl). All chromatograms are BPCs and their corresponding EICs 487.1599 $[M+H]^+$at 10 min retention time are shown.

Unfortunately, even though preparative RP-HPLC separation helped to further purify the compound it seems that especially the more polar components of our extracts seem to stick to the C18 column and so purification is rather not as straight forward as anticipated. Hence, we decided to pool all fractions from min 18 to min 28 and performed an additional round of VLC (Figure 42A, VLC 2). After all fractions were dried down, we measured them in LC-MS. Thereby, the VLC2 fraction at 80 % MeOH (second fraction 80b) showed really good purity, represented by an almost single clean peak with high intensity at 10 min in the BPC as well as in the EICs 487.1599 $[M+H]^+$ (Figure 45). We further lyophilized this fraction by freeze drying. The final dry weight of the purified compound was determined at 1.34 mg (Figure 42A, A, VLC2 fraction B). Moving on we performed a first Proton NMR analysis on the purified ACAD GUH-2 compound (Appendix, Figure 58). This proton NMR resulted in clearly distinguishable peaks with good intensity, so we decided to proceed with carbon NMR (^{13}C-NMR). Unfortunately, this measurement did not result in an evaluable spectrum, as no distinct peaks could be detected and there was a lot of background noise detectable.

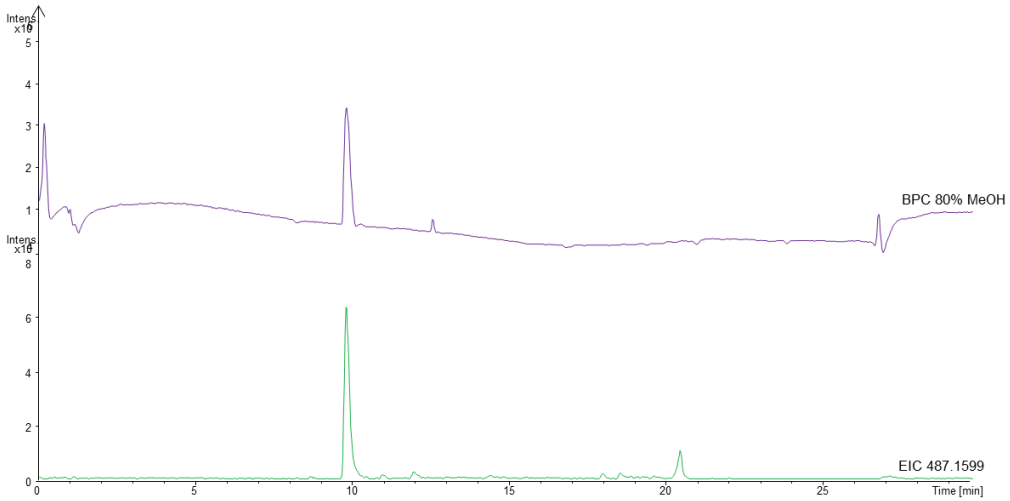


Figure 45: Fraction at 80 % MeoH (80b) from the second round of VLC (VLC2): The figure shows the BPC of the 80 % MeOH (80b) fraction of the second round of VLC of the ACAD GUH-2 compound and the associated EIC 487.1599 $[M+H]^+$at 10 min retention time.

Hence, more pure compound was needed, we decided to do a second round of cultivation and compound purification as depicted in Figure 42B. As during normal VLC, regarding the size of the stationary phase as well as the elution volume, we faced huge loss of compound, we tried to incorporate a second step purification by SPE (solid phase extraction) in 500 mg stationary C18 and 6 ml elution volume-based system (Strata C18-U, Phenomenex). Anyway, SPE purification was not satisfactory regarding separation of the compounds, so we continued with a further process taking advantage of semi preparative RP-HPLC. In total we separated around 500 mg of crude extract on a C18 250 x 10 mm column applying a linear gradient from 40 % to 100 % in 35 min at a flowrate of 3 ml/min. In order to isolate the pure ACAD GUH-2 compound, we collected fractions from retention time min 6 to min 15 every 0.25 min. Out of 36 fractions, fraction 20 to 27 contained the ACAD GUH-2 compound. After determining the dry weight of these collected fractions, they were all in the sub mg range or not able to be detected (Figure 42B, RP-HPLC semi prep fractions). Those fractions should be pooled and probably more compound needs to be purified for further NMR analysis.

5 Discussion

I Studies on the biosynthesis of the cyclopropylalanine (Acpa) moiety of the proteasome inhibitor belactosin A

The proteasome inhibitor molecules belactosins and cystargolides are built of a linear dipeptide backbone, which is *N*-acylated with a unique trans-disubstituted β-lactone moiety. Belactosin A and C are produced by *Streptomyces* sp. UCK14 and it could be shown for both molecules that they inhibit the 26S proteasome in the sub-mircomolar range. Besides their interesting bioactivity they display unprecedented chemistry. Belactosin A contains a unique amino acid moiety 3-(*trans*-2'-aminocyclopropyl) alanine (Acpa) for which the biosynthesis was unknown at the beginning of this study.

My thesis describes the identification and analysis of the belactosin biosynthetic gene cluster from *Streptomyces* sp. UCK14. Heterologous expression of the gene cluster allowed us to analyze the gene cluster borders in more detail and define the operon responsible for production of the Acpa moiety. By precursor feeding and chemical complementation studies, we could further shed light on the postulated biosynthesis route from L-lysine as a precursor towards Acpa and final incorporation into the belactosin A molecule. Feeding of racemic and double deuterium-labeled 3-(*trans*-2'-nitrocyclopropyl)alanine (d_2-Ncpa) and 3-(*trans*-2'-aminocyclopropyl)alanine (d_2-Acpa) in marker less and double crossover gene cluster knock-outs led to chemical complementation of UCK14/Δ*belK,* UCK14/Δ*belL* and UCK14/Δ*belK,* UCK14/Δ*belL* and UCK14/Δ*belN*, respectively. This gave us further proof that the putative oxidoreductases BelK and BelL are essential for the formation of Acpa by generating Ncpa as a cryptic nitro intermediate. Furthermore I could show that putative molybdopterin-dependent oxidoreductase BelN is the responsible enzyme for catalyzing the reduction from Ncpa to Acpa in the biosynthesis of belactosin A.

In further experiments, we tried to enlighten the biochemical function of the iron- and α-ketoglutarate-dependent dioxygenases BelL in an in vitro conversion assay guided approach. Following we performed chemical feeding studies with a chemical synthesized 5,5-d_2-4-hydroxylysine lactone. Eventually, in a collaborative study with the Abe lab from the University of Tokyo, Japan, we showed that both BelK and HrmI oxidize L-lysine to afford 6-nitronorleucine and that furthermore, BelL and HrmJ then catalyze the cyclization towards the Ncpa moiety of Ncpa in belactosin A and hormaomycin in a sterodivergent manner, respectively.

5.1 Heterologous expression and cluster border elucidation of the belactosin biosynthetic gene cluster

Before I took over this project, my colleague Felix Wolf could successfully construct a heterologous expression strain for the *bel* cluster in *Streptomyces albus* J1074 (see results section 3.1). Within our department we established and routinely apply a robust heterologous expression platform which provides excellent expression of diverse natural product derived biosynthetic gene clusters. Therefore, the fosmid carrying the gene cluster of interest, in the case of the belactosin BGC fosmid got redesigned for stable integration into *Streptomyces* chromosomes via attB/attP-site-specific recombination.[225] Within this thesis we were able to successfully detect two new peaks, corresponding to the signals produced by *Streptomyces* sp. UCK14 for the production of belactosin A and C. With these results we were confident that the introduced BGC encodes all elements required for the production of both belactosins, including the Acpa moiety of belactosin A. Using protein similarity models, operon organization analysis and comparison with characterized and uncharacterized homologous gene clusters; we initially assigned 22 genes to the *bel* cluster (*orf-1- belV*).

Besides heterologous expression of belactosins, we wanted to examine the borders of the *bel* BGC by the process of elimination of genes up- and downstream of the cluster. For systematic elimination of single or a group of genes, we were using a well-established PCR targeting method[183, 199, 220, 226] and could indicate the up- and downstream boundaries of the *bel* cluster. We were defining the *bel* BGC cluster from *belF* to *belV* (see results section 3.2). However, when analyzing the heterologous expression strains, it became obvious that the production of belactosin A and belactosin C was low in *S. albus* J1074 bel01 *ΔbelA-E*, in comparison to the *S. albus* J1074 bel01. A detailed comparison of MS spectra of *S. albus* J1074 bel01 *ΔbelA-E*, at 7.4 min and *Streptomyces* sp. UCK14 at 7.3 min (belactosin C) and 7.4 min (belactosin A) and of *S. albus* J1074 (without the cluster), *S. albus* J1074 bel01 (with intact cluster) showed convincing MS spectra for both compounds (Appendix, Figure 49). But it should be considered that the signals in the Δ*belA-E* mutants are quite low and thus difficult to analyze by MS/MS. In this context, we can not completely rule out partial complementation of the knock-out by the host strain, and interpretation of these results should be treated cautiously.[195]

The results of the border cluster analysis were not surprising to us, as we were recently able to identify the homologs of the putative amide synthetase BelU and AMP-dependent synthetase BelV in the BGC of cystargolides. We assume that BelU/BelV participate in the formation of

the peptidic backbone of belactosin A and belactosin C.[77, 227] We propose that that the belactosin cluster likely covers in total 17 genes from *belF* to *belV*. BelF is a 3-isopropylmalate dehydratase, and we thus speculate that this enzyme is likely involved in lactone warhead biosynthesis and therefore should be part of the belactosin cluster.

During analysis of the *S. albus* J1074 bel01 heterologous expression strains, we noticed the appearance of an additional peak in the EIC trace near 6.5 min. This extra peak shows a similar mass and is eluting near the retention time of belactosin A from the wild type producer strain *Streptomyces* sp UCK14. A detailed examination showed the characteristic MS/MS fragmentation of belactosin A (Appendix, Figure 50). This additional peak may contain an possible stereoisomer or conformer of belactosin A. However, further analysis, e.g. by NMR, should be conducted to shed light on the actual produced compound.

Now that we were confident, that the introduced *bel* cluster really encodes for belactosin A and C in the heterologous host and could further confirm, that genes *belA-E* are not part of the cluster, we aimed to prove that the four gene operon *belK-belN* is involved in the biosynthesis of the Acpa intermediate. To this end, we deleted the genes belK, belL and belN individually and transferred the mutant fosmids to *S. albus* J1074 bel01 for heterologous expression, with the help of the PCR targeting system as described before (see results section 3.3). LC-MS and MS/MS analysis of the individual knock-out strains showed complete abolishment in production of belactosin A. Production of belactosin C on the other hand was not impaired. As only belactosin A contains an Acpa moiety we are quite confident that the enzymes encoded in this operon are responsible for building the Acpa moiety.

Actinomycetes, such as *Streptomyces sp.* UCK14 are arguably the best studied producers of secondary metabolites in general. However, low production rates, silent gene clusters and laborious extraction of compounds for structure elucidation, are still major issues for these bacteria. A well-established strategy to overcome those burdens is to express the desired biosynthetic gene cluster in a model *Streptomyces* heterologous host strain. Also, since genetic manipulation of natural producer wild type strains can be difficult and time-consuming, heterologous expression of entire biosynthetic gene clusters and gene cluster knock-outs in this strain background is much more amenable and fully sequenced host strains are desirable. Furthermore genetic manipulation can be conducted in *E. coli* and after conjugation constructs can be expressed by the host strain.[219, 228, 229]

The most common used models for the heterologous expression of bioactive natural product gene clusters are host strains that have been developed from the well characterized actinomycetes strains *S. coelicolor* M145. *S. coelicolor* M512, M1146 and M1154 were developed by stepwise deleting different endogenous secondary metabolite gene clusters.[230] Further potential heterologous host are *S. lividans*[231], *S. avermitilis*[232] and *S.venezuelae*[233, 234]. These popular producer strains are well characterized, genetically determined and their genome sequences are stored and available in public databases.

Our first choice host strain *Streptomyces albus* J1074 was kindly provided by the John-Innes-Centre, Norwich, UK[235] and represents a very gifted host for the expression of diverse secondary metabolites.[236] There are numerous attempts to further engineer this strain for improved heterologous production.[236] The working group around Andriy Luzhetskyy spent a lot of afford and research into minimizing the genome of *S. albus* J1074. This leads to a fast-growing strain, as well as quick and high level of expression of the introduced a gene cluster. Furthermore, less additional host specific compounds are produced, which is huge advantage for identification of unknown compounds, as well as the downstream purification processes and following compound characterization studies.[189, 237] However, at the beginning of my study this strain was not available to the public.

5.2 Hypothesis for a possible biosynthesis route towards Acpa

A striking structural feature of belactosin A is the β-lactone warhead, which facilitates proteasome inhibitory activity and is conjugated to the atypical amino acid unit 3-(2-aminocyclopropyl)alanine (Acpa). In belactosin C this special cyclopropyl moiety is replaced with L-ornithine. In hormaomycins an analogous amino acid, 3-(*trans*-2'-nitrocyclopropyl)alanine (Ncpa) is present (chemical structures depicted in Figure 46A).[117, 195]

When we started this project the precise biochemical biosynthesis mechanisms of the route towards the Ncpa as well as the Acpa moiety, were still not understood. Based on the most recent research at the starting point of this thesis and previous stable isotope labeled precursor feeding studies by collaborative work of the Zeeck, de Meijere and Piel research groups[117, 173], we suggested a first potential biosynthesis model, as described in results section 3.4. In this work we explored the proposed route and did not find any refuting facts for the time being. However our collaborative work with the research working group of Prof. Abe (Graduate School of Pharmaceutical Sciences, The University of Tokyo, Japan) regarding the formation of Ncpa resulted in a different pathway. Both routes are depicted in Figure 46.

Because the unusual cyclopropylalanine moieties are found in both belactosin A and hormaomycin as structural features, we considered 3-(trans-2'-nitrocyclopropyl)alanine as a putative common intermediate for both compounds. Furthermore, comparison of the *bel* and *hrm* gene cluster showed that both share three pairs of homologous genes *belKLM*/*hrmIJT* (see results section 3.4 and Figure 46A) within two distinct operons, respectively. Based on the fact that both, Ncpa, which is incorporated in hormaomycin, and the belactosin A intermediate Acpa, contain *trans*-disubstituted cyclopropanes, we assumed that they could be constructed via the same biosynthetic strategy, presumably via the generation of a common *N*-oxidized intermediate (Figure 46B).

In this thesis, we wanted to elucidate the roles of those encoded homologous genes *belKLM* in the biosynthesis towards Acpa in more detail. Based on feeding studies conducted by Zeeck and Piel, we proposed a biosynthetic route to Ncpa starting from the amino acid L-lysine.[117, 173] From L-lysine a possible oxidized lactone intermediate could be formed, that allows following cyclopropanation gaining Ncpa (Figure 46B).

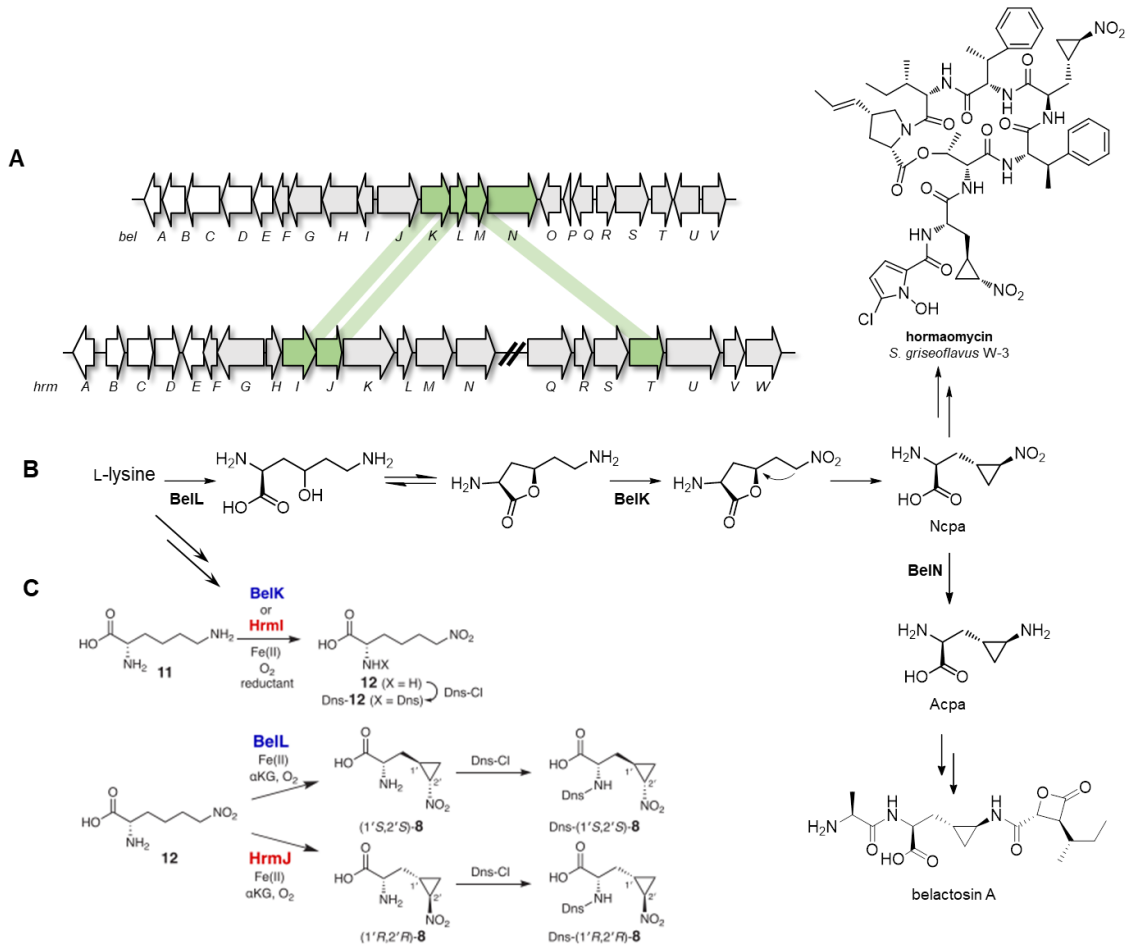


Figure 46: Biosynthesis model towards belactosin intermediates Ncpa and Acpa: A: Belactosin and hormaomycin biosynthetic gene clusters, homologous genes are indicated in green. **B:** Postulated biosynthesis hypothesis towards Ncpa and belactosin via a *N*-oxidized lactone intermediate, including the reduction of Ncpa to Acpa. **C:** Biosynthesis as postulated and verified by Abe et al., via BelK/HrmI produced 6-nitronorleucine and following stereodivergent cyclization by BelL/HrmJ towards Ncpa.

In parallel to our studies, we started a close collaboration with the research group around Prof. Abe from the University of Tokyo. In agreement with our initial in vivo feeding studies (see next chapter 5.3), they discovered an alternative route towards the formation of Ncpa. In complete accordance with our and Zeecks in vivo feeding results they also proposed that biosynthesis starts from L-lysine, which then is converted to an *N*-oxidized intermediate to afford Ncpa. They could show that BelK or HrmI are able to produce 6-nitronorleucine from oxidized L-Lysine. The cyclization towards the Ncpa intermediate in a stereodivergent manner intermediate is then further catalyzed by BelL/HrmJ (Figure 49C). More detailed findings will be further described in section 5.5 of this discussion.

Obviously, to afford the terminal amine in the Acpa moiety of belactosin A, the generated Ncpa would then subsequently undergo reduction to Acpa. Given the co-localization with the

genes encoding for BelK, BelL and BelM, a compelling candidate to catalyze this reduction reaction would be the putative molybdopterin-dependent oxidoreductase BelN, possibly via sequential two-electron, either by one six-electron reduction or three two electron reduction of the nitro group (Figure 46B).

This class of enzymes is known to catalyze two-electron reductions of diverse anorganic and organic substrates including nitrate, sulfate, DMSO and arsenate.[200] Therefore, the enzymes coordinate molybdenum by the two thiolates of a pyranopterin co-factor enabling the metal to cycle through the IV, V, and VI oxidation states. Sequential reduction cycles have also been reported, e.g. for perchlorate reductases which catalyze the two-step conversion of ClO_4^- to ClO_2^-.[238] However, molybdopterin oxidoreductases are rarely found in natural product biosynthesis e.g. in the formation of polyoxin and the benzodipyrrole CC-1065.[239, 240]

In the heterologous expression and knock-out studies of the *belKLN* operon, we excluded *belM.* We did so because we assume that this enzyme is not directly participating in Acpa biosynthesis, but rather plays a role in precursor supply. The HrmT homolog BelM is a putative diaminopimelate epimerase typically required for lysine biosynthesis. We speculate that it might have a supportive role in the formation of Acpa by ensuring sufficient supply of the amino acid precursor.[119]

5.3 Feeding of stable isotope labeled precursors – the Acpa moiety derives from L-lysine

Isotopic labeling methods continue to be an enlightening tool to source useful information in biosynthesis research and feeding of stabled isotope-labeled precursors is a very common and efficient approach to identify or characterize biosynthetic pathways or linking precursor molecules to end products.[241] Preliminary results of precursor feeding experiments by the Zeeck group suggested that the Ncpa in hormaomycins derives from L-lysine.[173] We wanted to test our hypothesis that the Acpa and Ncpa moieties of belactosin A and hormaomycin originate from the same precursor, so we performed analogous precursor feeding studies in the belactosin producer (see results 3.4). In this context, we supplemented separate cultures of the belactosin wild type producer strain with stable isotope labeled $^{13}C_5$-L-ornithine and $^{13}C_6$-L-lysine precursor molecules, respectively. These feeding studies clearly revealed that all five carbon atoms from $^{13}C_5$-L-ornithine are incorporated into the central amino acid in belactosin C. On the other hand, the feeding of $^{13}C_6$-L-lysine resulted in incorporation of all six ^{13}C-carbons from the precursor into the Acpa moiety. This very selective incorporation was further reinforced by the fact that the production of belactosin C was greatly diminished in the L-lysine supplemented cultures. Our results of the precursor feeding studies clearly demonstrate that the Acpa moiety is formed from L-lysine. This strongly supports a model in which the same biochemical machinery may direct the Acpa pathway in belactosin A and the Ncpa pathway in hormaomycin.

To further evaluate and find proof for our proposed hypothesis, we supplemented cultures of *Streptomyces* sp. UCK14 with synthetic and deuterium-labeled 3,3-d_2-Ncpa. We could successfully demonstrate, that Ncpa is a genuine intermediate of the biosynthetic pathway of belactosin A. Thereby, LC-MS and MS/MS analysis showed excellent incorporation of labeled 3,3-d_2-Ncpa in belactosin A as well as a very prominent shift towards the production of the belactosin A derivative in comparison to the C derivative. This knowledge motivated us to investigate the reduction of Ncpa to the Acpa intermediate in more detail.

5.4 Chemical complementation studies reveal reductase function of BelN

The identified Ncpa intermediate has to be reduced to afford the terminal amine in the Acpa moiety. To get further insights which enzyme is responsible for catalyzing this reaction, we aimed to conduct chemical complementation studies with labeled synthetic precursor molecules in a series of gene knock-out strains.

Our first feeding attempts in generated marker less knock-out mutants of genes *belK-N* in the heterologous expression strain *S. albus* J1074 bel01 (data not shown in this thesis) lead to disappointing results, as incorporation patterns with fed d_2-Ncpa in *S. albus* J1074 bel01 were not consistent. Out of this failed attempt, we decided to generate marker less mutant strains in the wild type producer strain *Streptomyces* sp. UCK14. We deleted genes *belK*, *belL*, *belM* and *belN*, individually (see results section 3.5) to be able to identify and assign a possible role of the encoding enzymes in the biosynthesis towards Acpa and the reduction from Ncpa to Acpa in particular, by following chemical complementation studies.

LC-MS analysis revealed that the deletion strains *Streptomyces* sp. UCK14/Δ*belK,* UCK14/Δ*belL,* UCK14/Δ*belN* still produce belactosin C, but completely abolished belactosin A production. However, the knock-out strain *Streptomyces* sp. UCK14/Δ*belM* still produced both belactosin derivatives at WT level. These results further undermine our hypothesis made during heterologous expression studies and strongly suggested that enzymes BelL and BelK and BelN, but not BelM, are crucial for the formation of belactosin A.

During feeding, and chemical complementation studies, the formation of an additional peak with *m/z* 370.2 $[M+H]^+$ and *m/z* 358.2 $[M+H]^+$, respectively, when analyzing the samples with LC-MS analysis, for both belactosin A and C, drew our attention. In the extracted ion chromatogram for the mass *m/z* 370.2 $[M+H]^+$ of belactosin A in *Streptomyces* sp.UCK14 the occurrence of such a double peak was most prominent (Appendix, Figure 53). The appearance and intensity of those peaks also increased when cultured in solid medium in comparison to liquid cultures. However, a consistent pattern in the appearance and correlating intensity of those peaks could not be observed. Obviously, as the cultivation set up was changed to gain better production yields, the formation of the double peak, probably was not previously noted because of the very low production titer of belactosins in liquid media, in general. If those additional peaks contain stereoisomers or conformers of the belactosin A and C molecules or

if an another, structurally different compound is produced, could not be finally resolved within the scope of this thesis. However, when we analyzed the two distinct peaks for belactosin A with MS/MS, we could identify very similar fragmentation patterns for both peaks (Appendix, Figure 53). However, to get an insight into the actual structural composition of those compounds one needs to conduct NMR analysis studies. In the course of this thesis, we already started a larger scale cultivation process of belactosins on solid medium. Therefore, in total over 10 L of agar plates were extracted and further purified by VLC, preparative HPLC and semi-preparative HPLC. To date several semi pure fractions are available and should be further purified and analyzed by NMR.

Interestingly, when we analyzed the LC-MS data of strain *Streptomyces* sp. UCK14/Δ*belN* in more detail, we were able to reveal a mass signal with *m/z* 175.06 $[M+H]^+$ matching Ncpa (data not shown because of very low signal). In order to allow chromatographic coupled mass analysis extracts of *Streptomyces* sp. UCK14 and *Streptomyces* sp. UCK14/Δ*belN* were deivatized with dansyl chloride. This derivatization method is a very useful tool to detect amino acids in complex matrices or if standard analytical detection is difficult. Dansyl chloride forms stable products with primary and secondary amino groups and can then be detected by HPLC couple mass analysis.[242]

Moving on, the specific accumulation of Ncpa in *Streptomyces* sp. UCK14/Δ*belN* could be further confirmed by feeding $^{13}C_6$-L-lysine to the mutant cultures. The integrity of the new build up peak was also corroborated by a positive control, which we conducted by spiking d_2-Ncpa to growing cultures of *Streptomyces* sp. UCK14/Δ*belN*. Feeding of $^{13}C_6$-L-lysine resulted in the production of $^{13}C_6$-Ncpa. Additionally, production was significantly higher than in the wild type strain (Appendix, Figure 55). These results strongly indicate that the Ncpa intermediate is indeed a substrate for the BelN enzyme, as $^{13}C_6$-Ncpa accumulated and no Acpa could be detected.

Subsequently, we supplemented cultures of *Streptomyces* sp. UCK14/Δ*belK,* UCK14/Δ*belL,* UCK14/Δ*belN* with either synthetic 3,3-d_2-Ncpa or 3,3-d_2-Acpa for chemical complementation (see results section 3.6). Feeding of 3,3-d_2-Ncpa did lead to complementation of the UCK14/Δ*belK,* UCK14/Δbel*L* mutants, but as we expected, did not lead to complementation of the UCK14/Δ*belN* mutant. As we postulated, the chemical complementation with the Ncpa intermediate, could not restore the *belN* mutant, as this enzyme most probably produces the Acpa intermediate for the final production of belactosin

A. However, when we were feeding 3,3-d_2-Acpa to the respective cultures, we were able to compensate the deletion of all three genes, including the *belN* knock-out, to generate d_2-belactosin A. Taken together these findings and additional successful genetic complementation of the *Streptomyces* sp. UCK14/Δ*belN* mutant (see results section 3.7) support the hypothesis that BelN is the enzyme responsible for the reduction of Ncpa to Acpa in the belactosin A biosynthetic pathway.

At a first glance, it appears quite metabolically inefficient to fully oxidize the γ-nitro group, only to reduce it immediately in the next biosynthesis step to the original oxidation state. One possible explanation could be, that the 6-nitro intermediate shows enhanced acidity of 6-H and seems to be mechanistic prerequisite to generate the key cyclopropane ring. Further findings by our close collaboration partners from the working group of Prof. Dr. Ikuro Abe (The University of Tokyo, Japan), will be elaborated in the next chapter.

Interestingly, the putative Mo-dependent reductase BelN shows the highest identity (27% i.d) to PcrA from *Azospira suillum* PS in a protein structure prediction alignment we created, using the web protein alignment tool Phyre2.0[243] (Data not shown). However, they have not yet been described as nitroreductases. To date not much has been enlightened on molybdopterin oxidoreductases in the context of the reduction of such cyclopropyl containing moieties, as this type of enzymes are rarely found in natural product biosynthesis. There are limited findings, for example in the formation of polyoxin, where it could be shown, that the molybdopterin oxidoreductase PolF is putatively involved in dehydrogenation of L-isoleucine. Also, in the pathway of the benzodipyrrole CC-1065, a pair of molybdopterin oxidoreductases C10B/C10C, are included, which likely catalyze the dehydroxylation of an aromatic ring moiety.[239, 240]

However, the determination of the actual enzymatic mechanism and potential accessory biocatalysts for this reaction will be an intriguing subject for further investigations.

5.5 Nitrocyclopropane formation by BelK/HrmI and BelL/HrmJ

In this project we closely collaborated with the working group of Prof. Dr. Ikuro Abe, from the University of Tokyo, on the Ncpa formation in hormaomycins and belactosin A. Our in vivo knock-out and chemical complementation experiments (see results section 3.5 and 3.6) clearly demonstrated that BelL and BelK are essential for Ncpa formation, and that biosynthesis is starting from L-lysine. Additionally, similar feeding experiments were carried out in the hormaomycin producer *Streptomyces griseoflavus* strain W 384, and they could also show that the knock-out strains Δ*hrmJ* and Δ*hrmI* could successfully be chemically complemented by feeding with d_2-Ncpa and thereby could show that these enzymes are clearly involved in Ncpa biosynthesis.[244] In alignment with these in vivo feeding experiments, the Abe group carried out enzymatic in vitro experiments in order to comprehend the exact biosynthesis mechanism and involved enzymes towards the Ncpa moieties of belactosin A and hormaomycins in more detail.[201]

In order to convert L-lysine to Ncpa, dehydrogenation at C4 and C6 position, as well as *N*-oxidation are required. We speculated that BelK (HrmI) and BelL (HrmJ) could catalyze such reactions. BelK and HrmI (51% identity), show low to moderate homology with heme oxygenase (HO)-like diiron enzymes. BelL and HrmJ (49% identity) share homology with αKG-dependent nonheme iron enzymes and are able to perform a diverse range of different oxidation reactions.[245-247]

Astonishingly, the Abe group could show that when recombinant produced and purified BelK was incubated with L-lysine 6-nitronorleucine was detected. Analogous incubation of L-Lysine with HrmI in place of BelK produced a similar result. BelK is a putative iron-containing redox enzyme of the haem oxygenase family. Belk and HrmJ show low to moderate similarity to heme oxygenases (HO)-like diiron enzymes. This family includes diverse biocatalysts such as PqqC, which mediates ring cyclization and eight-electron oxidation in the biosynthesis of pyrroloquinoline quinone (PQQ), and class I ribonucleotide reductases that generate deoxyribonucleotides for DNA synthesis.[248] Strikingly, a homology model of BelK revealed highest resemblance to the *N*-terminal and central domains of SznF from the streptozotocin BGC. SznF was recently shown to catalyze consecutive *N*-oxygenation of L-arginine followed by oxidative rearrangement to form an *N*-nitrosourea product with the help of an additional C-terminal cupin domain.[249-252] Furthermore, with a sequence alignment between HrmI (BelK) and SznF, they could identify conserved residues

E194, H204, E258, H288, D292, and H295 for diiron cofactor coordination in HrmI (BelK). When they substituted any of those residues with alanine, the activity of HrmI was completely abolished, stating their importance for iron binding.[201]

Now that nitronorleucine was identified to be produced by BelK/HrmI as an intermediate towards Ncpa, the Abe group wanted to investigate if nitronorleucine could by cyclized to Ncpa.Actually, they could show, that when BelL is incubated with 6-nitronorleucine, Ncpa with a trans-cyclopropane ring is formed. Furthermore, the stereogenic centers were unambiguously determined to be (1′S,2′S) which is fully consistent with the stereochemistry of belactosin A (Figure 47).[201]

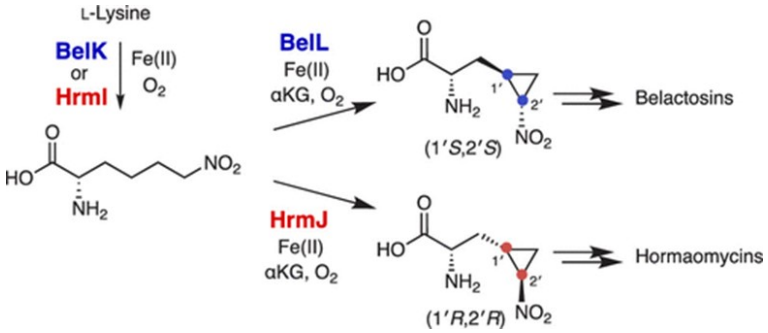


Figure 47: Biosynthesis towards Ncpa by BelK/HrmI and BelL/HrmJ: In vitro analysis indicated that both BelK and HrmI oxidize L-lysine to afford 6-nitronorleucine. Furthermore, BelL and HrmJ then catalyze the cyclization towards the Ncpa. Figure taken from Abe et al.[201].

Very strikingly, Acpa and Ncpa do not only exhibit different oxidation states of the 2′-nitrogen substituent but Ncpa also possess opposite stereoconfigurations at C1′ and C2′ in the cyclopropane ring, namely (1′S,2′S)-3-(2-aminocyclopropyl)alanine and (1′R,2′R)-3-(2-nitrocyclopropyl)alanine in belactosin A and hormaomycin, respectively. BelL and HrmJ are closely related in sequence (49% identity). However, whether they produce the same stereochemistry, (1′S,2′S) or the opposite (1′R,2′R) stereochemistry in Ncpa was unknown. When HrmJ was incubated with 6-nitronorleucine HPLC and NMR analyses suggested that a diastereomer of the BelL product is present, which was confirmed to be (1′R,2′R)-Ncpa in complete agreement with the stereochemistry of hormaomycin. HrmJ also showed much higher selectivity to only generate (1′R,2′R)-Ncpa. Since L-lysine was not accepted by either BelL or HrmJ, the nitro group formed by BelK or HrmI appears to be critical for cyclopropanation.[201]

5.6 4-hydroxy L-lysine or its 5,5-d_2-4-hydroxylysine lactone: intermediates of the biosynthetic pathway towards Acpa?

During this study we could successfully demonstrate, that BelN is responsible for the reduction of Ncpa to Acpa. However, in parallel investigations to the Abe group, we were curious about further shedding light on the earlier biosynthesis steps towards the Acpa moiety and intermediate Ncpa where we initially explored the postulated pathway by Zeeck et al.[173]

Zeeck postulated that L-lysine might be activated by hydroxylation, most likely at the C4 position, creating hydroxy-L-lysine to afford the following cyclopropanation reaction step, probably via a lactone intermediate. Brandl et al. conducted feeding studies in the hormaomycin producer strain, where feeding with the racemic 5,5-[D2]-4-hydroxylysine or its' synthetically more easily accessible lactone, was of prime interest for them. They observed the spontaneous cyclization of 4-hydroxylsine to its' γ-lactone. Indeed, feeding of the lactone led to deuterium-labeled hormaomycin as proved by ESI mass and D NMR spectra of the isolated product. The fact that the lactone was accepted as a substrate in the biosynthesis of hormaomycin made it a highly probable intermediate of the biosynthesis towards Ncpa. Additionally they observed, the loss of one deuterium label from C-4 of 4,4-[D2]-lysine, which makes 4-hydroxylysine together with its lactone a highly probable intermediate in hormaomycin and belactosin biosynthesis. Based on these findings they proposed a biosynthesis route towards Ncpa as shown in Figure 59 (see Appendix). [173]

Following, and also in alliance with our biosynthesis suggestion Höfer et al. followed on the proposed biosynthesis model by Brandl et al. towards Ncpa in hormaomycin. In summary Ncpa is generated from L-lysine via 4-hydroxylysine or its lactone, leading to the assumption, that cyclopropanation might be achieved by intramolecular nucleophilic substitution at the activated position 4 (proposed biosynthesis pathway shown in Figure 48).[117]

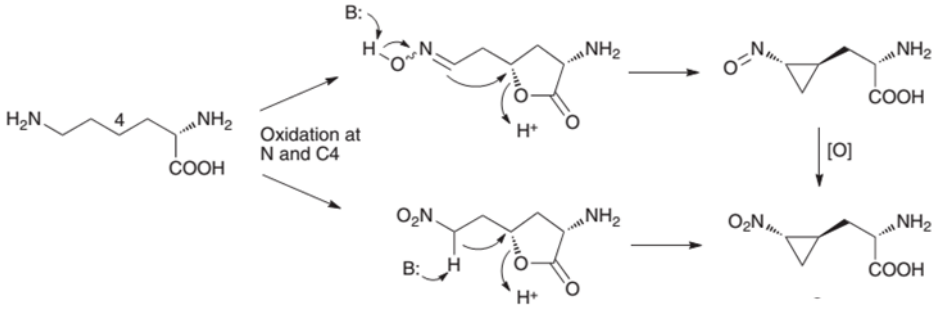


Figure 48: Proposed Ncpa Biosynthesis: Schematic representation of the proposed biosynthesis route towards Ncpa (Figure taken and adjusted from Hofer et al.)[117].

A compelling candidate for the activation of L-lysine by C4-hydroxylation would be the HrmJ homologue BelL (49 % identity). Both enzymes, HrmJ and BelL, contain a conserved protein domain (PF10014) of the former DUF 2257 family. This distinct enzyme family compromises characteristic iron- and α-ketoglutarate-dependent dioxygenases, which have lately been explored as a source for amino acid hydroxylases.[253, 254] Fe/αKGs enzymes represent a versatile and intriguing enzyme family by their ability to directly functionalize inactivated C–H bonds at the cost of αKG and O_2. They play an important role in the biosynthesis of natural products with valuable biologically activities.[204] Very recently, the BelL homolog GlbB (37% similarity on amino acid sequence level) from the glidobactin gene cluster has been characterized as a regio- and stereoselective lysine 4-hydroxylase by Renata et al. It could be shown in vitro that this enzyme is able to regio- and stereo-selectively hydroxylate L-lysine at is C4 position.[191] Based on the findings of Abe et al. and Renata et al., as well as the feeding studies by Brandl et al., we postulated that BelL might catalyze a similar hydroxylation reaction as described for GlbB to afford the proposed 4-hydroxy lysine intermediate, to allow further cyclopropanation reaction towards the Acpa moiety of belactosin A.

As favorable reaction conditions as well as co-factors were described for the GlbB enzyme, we wanted to investigate if we could facilitate a similar reaction in an in vitro assay using recombinant expressed and purified BelL enzyme and L-lysine as a substrate. Recombinant production and purification of BelL was very straight forward and could be received in good amounts. However, even though we tested several different reaction conditions, in vitro hydroxylation of L-lysine could not be achieved (see results section 3.9).

As an alternative, we debated if we could feed GlbB generated 4-hydroxy lysine in chemical complementation studies in *Streptomyces* sp. UCK14/Δ*belL* to possibly restore belactosin A production. Employing GlbB in an enzyme assay, we could successfully generate 4-hydroxy-

L-lysine. Using $^{13}C_6$-L-lysine as a substrate in GlbB assay we could also generate stable isotope labeled $^{13}C_6$-4-hydroxy-L-lysine, which could allow us to follow successful incorporation of the precursor into belactosin A by LC-MS. First we analyzed possible incorporation of $^{13}C_6$-4-hydroxy-L-lysine in the *Streptomyces* sp. UCK14 wild type. We could successfully detect a new peak for EIC *m/z* 376.2 $[M+6+H]^+$ for belactosin A at the corresponding retention time. Unfortunately, we could not chemically complement the UCK14/Δ*belL* mutant with feeding $^{13}C_6$-4-hydroxy-L-lysine to the culture. We were not able to detect $^{13}C_6$-labeled belactosin A with EIC *m/z* 376.2 $[M+6+H]^+$.

However, based on the findings by Brandl et al. regarding the feeding of the lactone, we were motivated to carry out feeding and chemical complementation studies with a synthesized lactone intermediate in the belactosin producer strains, as well. When we fed the racemic deuterium-labeled 5, 5-d_2-4-hydroxylysine lactone to *Streptomyces* sp. UCK14 wild type and *Streptomyces* sp. UCK14/Δ*belL*. To our satisfaction, we could detect a mass shift of 2 Da in the wild type and the *Streptomyces* sp UCK14/Δ*belL* supplemented culture and MS/MS analysis showed quite convincing fragmentation patterns for d_2-belactosin A. However, intensities were rather low in comparison to feeding studies with Ncpa or Acpa. The positive control Acpa matched the retention time of the new produced peak. Unfortunately, out of time reasons and because we needed to synthesize larger amounts of the 5,5-d_2-4-hydroxylysine lactone, within this thesis we were not able to repeat this chemical complementation study. One should definitely try to feed the lactone in higher final concentrations as well as actually trying to purify or derivatized the produced compound for actual structure elucidation.

When the Abe group tested different amino acids as substrates for the BelK homologue HrmI, none of them were accepted as substrates, implying that HrmI strictly recognizes the carbon chain length as well as the α-carbon stereochemistry in L-lysine. Nevertheless, when they incubated HrmI with N^6-hydroxy-L-lysine, it was converted to 6-nitro-norleucine consistent with a reaction mechanism involving N^6-hydroxyl-L-lysine as an on-pathway intermediate during the oxidation of L-lysine to 6-nitro-norleucine.[201]

Those findings actually provide explanation why BelL could not convert L-lysine to 4-hydroxy-L-lysine, as this is most likely the wrong substrate for this enzyme. However, there were different possibilities regarding the biosynthesis of Ncpa that we explored together with the working groups of Abe and Hughes. However the inconsistency with the published catalyzation there is also the possibility that both mechanisms are compatible and 4-hydroxy-

L-lysine. and the 4-hydroxy-lactone may still be possible intermediates of this biosynthesis pathway towards the Ncpa intermediate of belactosin A. Further investigations will have to prove this.

5.7 Newest discoveries on cystargolide biosynthetic gene cluster and the biosynthesis of the β-lactone warhead in both cystargolides and belactosins

Our research group recently identified the cystargolide as well as the belactosin BGCs and investigated the biosynthesis with gene knock-out strains in the cystargolide producer strain *Kitasatospora cystarginea* NRRL-B16505. We could show that both cystargolides and belactosins are not synthesized by NRPS or PKS guided machineries, but instead are constructed by rare single enzyme amino acid ligases. Furthermore, we could show that the biosynthetic pathway towards the β-lactone warhead is reminiscent of leucine biosynthesis.[84] However, a genetic platform for detailed and comprehensive knock-out studies performed in a heterologous expression system was not feasible (work conducted by my colleagues Felix Wolf and Daniel Männle, data not shown), nor stable for the cystargolide BGC of *K. cystarginea*.

In our newest study, led by my colleague Patrick Beller, we were able to establish a heterologous expression system and a series of gene cluster knock-outs within the BGC in the homologous strain *Streptomyces durhamensis* NRRL-B3309 by a TAR guided cloning approach.[255] The gene cluster of interest contains genes *cysA* to c*ysH*, which are homologues of the original cystargolide gene cluster from *Kitasatospora cystarginea* NRRL-B16505. The results and outcomes of this gene cluster knock-out analysis are summarized in a manuscript, that is currently in preparation by Beller et al.[255]

However, knock-out Δ*cysA* did not affect cystargolide production, and its function is most likely compensated by the putative IPMS (SCO2528) from the *S. coelicolor* host strain (ID 57.64% on protein level). The putative DHA2 family efflux MFS transporter permease subunit CysB and the putative LysR family transcription factor CysH also did not show an effect on the production of cystargolides. Their activity can probably be compensated by the other intrinsic unspecific transporter systems and endogenous transcription factors which interact with the *cys* BGC, respectively.[255]

At a first glance, CysG and CysE have no apparent function in cystargolide biosynthesis, however in Δ*cysG* and Δ*cysE* cystargolide production was completely abolished. Intriguingly, a similar methyltransferase/methylesterase pair is encoded in the belactosin BGC (BelI and BelR). Leading to the speculation that CysG/BelI and CysE/BelR might be involved in the

formation of the β-lactone warhead.[255] In an in vitro approach, we could show, that CysG in the presence of SAM, is able to act on 3-IPM to generate mono-methylated 3-IPM. As in the final cystargolide product a methyl group is absent, methyltransferase CysG might act on the free β-lactone carboxylic acid building block or a precursor of this intermediate.[255] We postulate that likely CysG plays a similar role in the biosynthesis of the β-lactone warhead. Structural characterization by NMR studies (AG Groß, Irina Helmle) revealed that the methylation of 3-IPM takes place at the 1-carboxylic moiety of 3-IPM. One possible hypothesis could be that CysG acts as a diverter, relegating the primary metabolite 3-IPM to cystargolide biosynthesis or CysG eventually masks the 1-carboxylic acid group for the following biotransformations.[255]

Interestingly, the methyltransferase/methylesterase pair BioC/BioH which can be found in the biotin synthesis pathway of proteobacteria[256256256], exhibit conserved protein domains that are found in CysG and CysE. BioC catalyzes SAM-dependent methyl esterification of malonyl-ACP allowing recognition by house-keeping charge-sensitive fatty acid synthetases as a starter unit.[256] After two reiterations of the fatty acid elongation cycle the generated pimeloyl-ACP methyl ester is hydrolyzed to pimeloyl-ACP and methanol by BioH. Pimeloyl-ACP is then further converted to biotin. A similar role for CysG and CysE in cystargolide biosynthesis would be imaginable.[255, 256] In this hypothetical scenario the O-methyl 3-IPM product of CysG would undergo lactonization between the 4-carboxylic acid and the 2-hydroxy group, affording the β-lactone methyl ester building block. Prior to amidation the 1-carboxyl group would have to be liberated by demethylation by the putative methylesterase CysE.[255] However further investigations are have to be made to elucidate the exact functions of 3-IPM methylation.

5.8 Outlook and further investigations in the belactosin biosynthesis project

In this thesis we could shed light on the biosynthesis of the unique 3-(*trans*-2'-aminocyclopropyl)alanine moiety of the β-lactone proteasome inhibitor Belactosin A, from *Streptomyces* sp. UCK14. In means of heterologous pathway expression, constructed of gene deletion mutants and feeding studies for chemical complementation with stable isotope labeled precursors. We could show that, a cryptic nitrocyclopropylalanine intermediate is generated from L-lysine and that the subsequent reduction of the *N*-oxygenated precursor Ncpa to the corresponding amine containing precursor Acpa is mediated by the molybdopterin-dependent enzyme BelN. In a cooperative study with the working group of Prof. Abe from the University of Tokyo, they could show that starting from L-Lysine BelK/HrmI are able to produced 6-nitronorleucine, which is then further catalyzed in a stereodivergent cyclization by BelL/HrmJ to form Ncpa.

For future research it would be very intriguing to get insights into the actual enzyme mechanism of BelN in an in vitro model. Regarding the putative 4-OH-L-lysine and the 5,5-d_2-4-hydroxylysine lactone chemical complementation studies, it would be very beneficial to further analyze the compound produced by the chemical complemented *Streptomyces* sp. UCK14/ΔbelL mutant as well as testing those compounds as substrates in in vitro assays.

Furthermore, the detailed lactonization reaction of the β-lactone warhead still is an very interesting subject for ongoing research

We would also be very interested which enzyme or enzymes are responsible for the linkage of the peptidic backbone to the Acpa as well as the β-lactone warhead.

II Studies on the role of proteasome inhibitors from *Nocardia* in pathogen survival

Nocardia spp. are filamentous Actinobacteria of the order Corynebacteriales and mostly known for their ability to cause localized and systemic infections in humans. However, the virulence mechanisms of human pathogenic *Nocardia* and the onset and progression of nocardiosis are only poorly understood until today. Because *Nocardia* have become more and more relevant in the clinics, recent genome sequencing has revealed an extraordinary capacity for the production of diverse and specialized small molecules.

The production of secondary metabolites is often crucial for microbes to survive the challenges of different environmental conditions. This project within this thesis was part of the research training group RTG1708 “Molecular principles of bacterial survival strategies” which was granted by the DFG in 2011 and launched in April 2012. This RTG addressed the question, how bacteria maintain viability in adverse environments. Facing those challenges, bacteria, have acquired elaborated strategies during evolution to withstand and overcome unfavorable conditions in order to protect their niches and colonize new habitats. Thirteen different projects were devoted to the investigation of bacterial survival strategies involving maintenance-metabolism, detoxification, repair pathways and protective substances and structures.

At the main part of this project within this thesis we wanted to investigate natural products of *Nocardia* on their virulence-associated importance. A particular gene cluster in *N. cyriacigeorgica* GUH-2 genome sparked our interest, because this BGC encodes a putative acyl-CoA dehydrogenase (ACAD) with homology to the epoxyketone synthase from the proteasome inhibitor-associated biosynthetic pathways of epoxomycin.[168] We expressed this ACAD gene cluster in a heterologous host system and were able to identify a putative new compound. Further on we aimed to purify this compound for ongoing structure elucidations.

5.9 Virulence potential of *Nocardia cyriacigeorica* GUH-2 derived natural products

Our work emphasizes that *Nocardia* represent a prolific source for natural products rivaling better-characterized genera such as Streptomyces or *Amycolatopsis*.[22, 142] Within this thesis we wanted to explore *Nocardia* derived natural products with potential bioactivity or interesting chemistry. We further wanted to investigate if those natural products can be linked to the virulence of human pathogenic *Nocardia* strains.

In order to access the biosynthetic potential of the specific pathogenic *Nocardia cyriacigeorgica* GUH-2 strain, we conducted a genome wide analysis with antiSMASH[85] to unravel interesting biosynthetic possibly virulence associated gene clusters. This analysis revealed a plethora of putative biosynthetic gene clusters of various classes, including PKS, NRPS, and terpenoid pathways. In previous studies within our working group, led by Daniel Männle in association with the working group of Prof. Dr. Ziemert, University of Tuebingen, with the help of sequence similarity networks generated by BiG-SCAPE (Biosynthetic Gene Similarity Clustering and Prospecting Engine) we could show that several distinct gene cluster families (GCF) are present distributed over several *Nocardia* strains. In general, all data was collected from *Nocardia* deriving from different origins, like soil, water or from human sampling. However, all those organisms from the different niches were more or less evenly distributed among the six clades that were identified in this analysis. But so far, we couldn't correlate human pathogenic *Nocardia* with a distinct clade. Nevertheless, we were very interested if *Nocardia* produced natural compounds can be related to pathogenicity.[22]

Interestingly we could identify a number of GCFs comprise clusters, which are highly abundant in many *Nocardia* genomes. One striking GCFs includes pathways with mono-modular type I PKS-like ketosynthases. They are potentially involved in the modification of fatty acid precursors and can for example be part of highly conserved cluster, which for example direct the synthesis of mycolic acids, a characteristic and essential cell wall components of mycobacteria and related genera like *Corynebacterium*, *Rhodococcus*, or *Nocardia*.[22, 257] Taking a closer look, interestingly, we could discover a PKS containing cluster of *N. cyriacigeorgica* GUH-2 that shows an analogous PKS which is also present in the mycolic acid BGC from *M. tuberculosis*. Such surface-exposed glycolipids including mycolic acids, mycocerosic acids and pthiocerols, besides various siderophores and the toxin mycolactone, are known to act as virulence factors in *Mycobacteria*.[156, 258, 259] [156, 260-263]

Thinking of further *Nocardia* associated pathogenic mechanisms, virulence may be promoted by favored uptake into macrophages by activating the alternative complement pathway without triggering respiratory burst. In an in vitro system it could be shown that the glycolipid TDM potentially inhibits membrane fusion between the phagosome containing the bacterial cell and lysosomes of the host. Additionally, TDM is able to stimulate essential effector cells in innate, early adaptive and both humoral and cellular adaptive immunity.[158, 159] Furthermore Daniel Männles' analysis showed, that this cluster family seems to be highly conserved in almost all *Nocardia* genomes.[22, 85] Moreover, we were able to identify PKS type I enzymes from two clusters that show pthiocerol and pthioceranic acid synthases PpsC and Pks2,significant homology (48%) to the respectively, suggesting that these BGCs could be involved in glycolipid biosynthesis, as well.[85, 142, 264] Remarkably, most sequenced *Nocardia* strains contain at least one such monomodular type I PKS cluster in addition to the mycolic-acid PKS. Furthermore, *Nocardia* genomes are rich in monomodular type I PKS. The clinically important pathogen *N. cyriacigeorgica* GUH-2, for example, harbors five biosynthetic gene clusters with this organization (out of 19 predicted BGCs in total).[141, 265]

5.10 Heterologous expression of the *N. cyriacigeorgica* GUH-2 derived ACAD GUH-2 biosynthetic gene cluster

Nocardia are getting more and more into focus as promising producers of novel natural products with interesting biological functions or chemistry. However, because most of *Nocardia* spp. produce secondary metabolites only in trace amounts, it was found difficult to produce and isolate the desired compounds. Furthermore, many biosynthetic gene clusters are silent under lab conditions or are described as orphan BGCs as they can not be assigned with a respective compound. In addition, most *Nocardia* spp. are biosafety level 2 organisms and genetic manipulation in the native producer can thus be tedious and time consuming. Therefore, heterologous expression of BGCs in heterologous has emerged as a key strategy to gain access to the promising secondary metabolite gene cluster and to reach titers of industrial levels.[266-271]

In order to gain access to the ACAD cluster and the *N. cyriacigeorgica* GUH-2 genome in general, we constructed a fosmid based genomic library and subsequently screen this library for the complete ACAD cluster. Such fosmid based genetic libraries are rather laborious to construct, but they have the advantage to provide more genetic information per vector, because of their ability to carry larger DNA inserts and their assigned functional genes. Furthermore, large inserts favor the identification and expression of natural product encoding biosynthetic gene clusters. Cosmids and fosmids can accommodate DNA inserts from 25 up to 50 kb in size. Bigger insert sizes (up to 300 kb) can be cloned into bacterial artificial chromosomes (BACs). Cosmids are artificially constructed vectors containing the Cos site, which permits packaging of DNA into phage lambda for transfection. BACs are able to introduce large DNA inserts into *E. coli* and are based on the single-copy F plasmid. The inserted DNA is present in low copy number and is thus more stable. Fosmids are cosmid-based vectors containing the replication origin of the *E. coli* F plasmid as well. They thus combine the stability-favoring properties of BACs, but they ae more accessible to easier genetic manipulation. Those fosmids are also inducible for high copy number production, which make them an ideal tool for ongoing cloning procedures in *E. coli.*[272-274] However, if biosynthetic pathways reach a size up to 100 kb, other techniques can be used for cloning large BGCs into suitable expression vectors. Recently, several whole-pathway direct cloning approaches have been developed, including full-length RecE-mediated recombination in *E. coli*[275], Cas9-assisted in vitro assembly[276], and transformation-associated recombination (TAR) in *Saccharomyces cerevisiae.*[277] Such cloning platforms can accelerate the process of

discovering new natural products, help to elucidate their biosynthesis and allow engineering of biosynthetic pathways.

For the genome of *N. cyriacigeorgica* GUH-2 we could successfully construct a fosmid libray, including a fosmid carrying the complete ACAD gene cluster. This fosmid was refurbished for heterologous expression as described before (see 2.22.4) and the ACAD cluster was introduced into the heterologous hosts *S. albus* J1074 and *Amycolatopsis japonicum* M6417-CF17, which were already successfully employed for the production of *Nocardia* derived natural products in past studies.[186, 221]

Heterologous expression of the ACAD cluster produced a novel peak in both *S. albus* J1074 and *A. japonicum* M6417-CF17. In order to characterize the novel compound we analyzed and compared in-silico fragmentation patterns as well as HR MS/MS spectra. However, this first insights did not lead to identification of a possible compound candidate. A detailed search within different natural product data bases such as the Dictionary of Natural Products® (DNP)[222] or MarinLit[223] suggested that the produced compound is a novel and has not been described in literature so far. On the basis of this knowledge, we decided to upscale the production of this compound for downstream purification, to enable structure elucidation by NMR.

In total we cultivated around 12 L of *S. albus* J1074/ ACAD GUH-2 in two cultivation rounds, which resulted in around 10 g crude extract in total. Production of the novel *m/z* 487.1599 $[M+H]^+$ compound was reliably transferred to the upscaled production volume from 50 ml to 1 L cultivations in shake flasks. In following purification steps by VLC, preparative HPLC and semi-preparative HPLC we could obtain 1.34 mg pure compound. First, ^{1}H-NMR analysis of the purified compound showed a very distinct spectrum with prominent peaks, so we tried to further conduct ^{13}C-NMR analysis. However, the amount of compound was not sufficient to receive conclusive results. Further purification of the compound is ongoing in our lab.

In general, it would be very interesting and essential to investigate the virulence mechanism of *Nocardia* in general in more detail. In first attempts we already established and primarily developed eukaryotic cell based assays to study immune evasion mechanisms or immunoproliferative effects. We also conducted cytotoxic assays by treating immune cells with cell extracts or purified compounds to gain first insights of their effect on immune cells. In this asset we also very successful established a robust protocol for the genetic manipulation

of *Nocardia* strains, which can be used for knock-out studies or competition assays in comparison with wild type strains.

5.11 Outlook and further investigations on the "natural products from *Nocardia* and associated pathogen survival" project

Within this thesis and in collaboration with our colleagues we were able to conduct a genome wide analysis of the human pathogenic *N. cyriacigeorgica* GUH-2 strain, which shows a huge potential for the production of a plethora of natural products. We strongly focused on the PKS-NPRS derived putative ACAD compound, regarding heterologous expression and purification of a novel compound. We could obtain first data for the structure elucidation of this compound, but more compound needs to be purified to actually solve the chemical structure of the putative novel ACAD compound.

However, this strain and other pathogenic *Nocardia* strains inhabit the potential for the discovery of other natural products molecules with putative interesting bioactivity or new structural diversity. In further investigations it would be very worthwhile to explore other biosynthetic gene clusters or gene cluster families. In particular a highly conserved mycolic acid associated PKS cluster seems to be conserved in all analyzed strains. The very comprehensive and detailed study by Daniel Männle within our group could be very supportive in such investigations by analyzing particular clades and gene cluster families of the generated sequence similarity networks in more detail.

Finally, we would like to gain further insights into virulence mechanism of *Nocardia* strains that are potentially linked to natural products, produced by these pathogenic strains. One could think of different in vitro cell based assays or using cultured *Nocardia* wild type cells or gene cluster knock-out strains to challenge immune cells.

6 List of Tables

7 List of Figures

8 References

1. Atanasov, A. G.; Zotchev, S. B.; Dirsch, V. M.; Supuran, C. T., Natural products in drug discovery: Advances and opportunities. *Nature Reviews Drug Discovery* **2021,** *20*, 200-216.
2. Katz, L.; Baltz, R. H., Natural product discovery: past, present, and future. *Journal of Industrial Microbiology and Biotechnology* **2016,** *43*, 155-176.
3. Berdy, J., Bioactive microbial metabolites. *The Journal of antibiotics* **2005,** *58*, 1-26.
4. Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J., The re-emergence of natural products for drug discovery in the genomics era. *Nature reviews drug discovery* **2015,** *14*, 111-129.
5. Newman, D. J.; Cragg, G. M., Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *Journal of Natural Products* **2020,** *83*, 770-803.
6. Christoffersen, R. E., Antibiotics−an investment worth making? *Nature Biotechnology* **2006,** *24*, 1512-1514.
7. Kumar, K.; Waldmann, H., Synthesis of natural product inspired compound collections. *Angewandte Chemie International Edition* **2009,** *48*, 3224-3242.
8. Dye, C., After 2015: infectious diseases in a new era of health and development. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2014,** *369*, 20130426.
9. Barka, E. A.; Vatsa, P.; Sanchez, L.; Gaveau-Vaillant, N.; Jacquard, C.; Klenk, H.-P.; Clément, C.; Ouhdouch, Y.; van Wezel, G. P., Taxonomy, physiology, and natural products of Actinobacteria. *Microbiology and Molecular Biology Reviews* **2016,** *80*, 1-43.
10. Lee, L.-H.; Goh, B.-H.; Chan, K.-G., Actinobacteria: Prolific producers of bioactive metabolites. *Frontiers in Microbiology* **2020,** *11*.
11. Mann, J., Natural products as immunosuppressive agents. *Natural product reports* **2001,** *18*, 417-430.
12. Medema, M. H.; Kottmann, R.; Yilmaz, P.; Cummings, M.; Biggins, J. B.; Blin, K.; De Bruijn, I.; Chooi, Y. H.; Claesen, J.; Coates, R. C., Minimum information about a biosynthetic gene cluster. *Nature chemical biology* **2015,** *11*, 625-631.
13. Walsh, C. T.; Fischbach, M. A., Natural products version 2.0: connecting genes to molecules. *Journal of the American Chemical Society* **2010,** *132*, 2469-2493.
14. Forchhammer, K., Editorial for Article Collection on "Bacterial Survival Strategies". *Microbial Physiology* **2021**, 1-3.
15. Netzker, T.; Flak, M.; Krespach, M. K.; Stroe, M. C.; Weber, J.; Schroeckh, V.; Brakhage, A. A., Microbial interactions trigger the production of antibiotics. *Current opinion in microbiology* **2018,** *45*, 117-123.
16. Choi, S.-S.; Katsuyama, Y.; Bai, L.; Deng, Z.; Ohnishi, Y.; Kim, E.-S., Genome engineering for microbial natural product discovery. *Current opinion in microbiology* **2018,** *45*, 53-60.
17. Nguyen, C. T.; Dhakal, D.; Pham, V. T. T.; Nguyen, H. T.; Sohng, J.-K., Recent Advances in Strategies for Activation and Discovery/Characterization of Cryptic Biosynthetic Gene Clusters in Streptomyces. *Microorganisms* **2020,** *8*, 616.
18. Rutledge, P. J.; Challis, G. L., Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nature Reviews Microbiology* **2015,** *13*, 509-523.
19. Zhang, M. M.; Qiao, Y.; Ang, E. L.; Zhao, H., Using natural products for drug discovery: the impact of the genomics era. *Expert opinion on drug discovery* **2017,** *12*, 475-487.
20. Albarano, L.; Esposito, R.; Ruocco, N.; Costantini, M., Genome mining as new challenge in natural products discovery. *Marine drugs* **2020,** *18*, 199.

21. Challis, G. L., Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology (Reading)* **2008,** *154*, 1555-1569.
22. Männle, D.; McKinnie, S. M.; Mantri, S. S.; Steinke, K.; Lu, Z.; Moore, B. S.; Ziemert, N.; Kaysser, L., Comparative Genomics and Metabolomics in the Genus Nocardia. *Msystems* **2020,** *5*.
23. Wong, C. C.; Cheng, K. W.; He, Q. Y.; Chen, F., Unraveling the molecular targets of natural products: Insights from genomic and proteomic analyses. *PROTEOMICS‐Clinical Applications* **2008,** *2*, 338-354.
24. Fricker, L. D., Proteasome inhibitor drugs. *Annual review of pharmacology and toxicology* **2020,** *60*, 457-476.
25. Bochtler, M.; Ditzel, L.; Groll, M.; Hartmann, C.; Huber, R., The proteasome. *Annual review of biophysics and biomolecular structure* **1999,** *28*, 295-317.
26. Ciechanover, A., The ubiquitin-proteasome proteolytic pathway. *Cell* **1994,** *79*, 13-21.
27. Finley, D., Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annual review of biochemistry* **2009,** *78*, 477-513.
28. Ravid, T.; Hochstrasser, M., Diversity of degradation signals in the ubiquitin‐proteasome system. *Nature reviews Molecular cell biology* **2008,** *9*, 679-689.
29. Kleiger, G.; Mayor, T., Perilous journey: a tour of the ubiquitin‐proteasome system. *Trends in cell biology* **2014,** *24*, 352-359.
30. Nandi, D.; Tahiliani, P.; Kumar, A.; Chandu, D., The ubiquitin-proteasome system. *Journal of biosciences* **2006,** *31*, 137-155.
31. Pickart, C. M.; Cohen, R. E., Proteasomes and their kin: proteases in the machine age. *Nature reviews Molecular cell biology* **2004,** *5*, 177-187.
32. Bedford, L.; Paine, S.; Sheppard, P. W.; Mayer, R. J.; Roelofs, J., Assembly, structure, and function of the 26S proteasome. *Trends in cell biology* **2010,** *20*, 391-401.
33. Coux, O.; Tanaka, K.; Goldberg, A. L., Structure and functions of the 20S and 26S proteasomes. *Annual review of biochemistry* **1996,** *65*, 801-847.
34. Murata, S.; Yashiroda, H.; Tanaka, K., Molecular mechanisms of proteasome assembly. *Nature reviews Molecular cell biology* **2009,** *10*, 104-115.
35. Schmidt, M.; Hanna, J.; Elsasser, S.; Finley, D., Proteasome-associated proteins: regulation of a proteolytic machine. **2005**.
36. Arendt, C. S.; Hochstrasser, M., Eukaryotic 20S proteasome catalytic subunit propeptides prevent active site inactivation by N-terminal acetylation and promote particle assembly. *The EMBO Journal* **1999,** *18*, 3575-3585.
37. Groll, M.; Heinemeyer, W.; Jäger, S.; Ullrich, T.; Bochtler, M.; Wolf, D. H.; Huber, R., The catalytic sites of 20S proteasomes and their role in subunit maturation: a mutational and crystallographic study. *Proceedings of the National Academy of Sciences* **1999,** *96*, 10976-10983.
38. Balch, W. E.; Morimoto, R. I.; Dillin, A.; Kelly, J. W., Adapting proteostasis for disease intervention. *science* **2008,** *319*, 916-919.
39. Schwartz, P., MD, Alan L; Ciechanover, M., PhD, Aaron, The ubiquitin-proteasome pathway and pathogenesis of human diseases. *Annual review of medicine* **1999,** *50*, 57-74.
40. Drexler, H. C., Activation of the cell death program by inhibition of proteasome function. *Proceedings of the National Academy of Sciences* **1997,** *94*, 855-860.
41. Lopes, U. G.; Erhardt, P.; Yao, R.; Cooper, G. M., p53-dependent induction of apoptosis by proteasome inhibitors. *Journal of Biological Chemistry* **1997,** *272*, 12893-12896.
42. Nalepa, G.; Rolfe, M.; Harper, J. W., Drug discovery in the ubiquitin‐proteasome system. *Nature reviews Drug discovery* **2006,** *5*, 596-613.

43. Génin, E.; Reboud-Ravaux, M.; Vidal, J., Proteasome inhibitors: recent advances and new perspectives in medicinal chemistry. *Current topics in medicinal chemistry* **2010,** *10*, 232-256.
44. Gersch, M.; Kreuzer, J.; Sieber, S. A., Electrophilic natural products and their biological targets. *Natural product reports* **2012,** *29*, 659-682.
45. Gräwert, M. A.; Groll, M., Exploiting nature's rich source of proteasome inhibitors as starting points in drug development. *Chemical Communications* **2012,** *48*, 1364-1378.
46. Kaysser, L., Built to bind: biosynthetic strategies for the formation of small-molecule protease inhibitors. *Natural product reports* **2019,** *36*, 1654-1686.
47. Kisselev, A. F.; Goldberg, A. L., Proteasome inhibitors: from research tools to drug candidates. *Chemistry & biology* **2001,** *8*, 739-758.
48. Chen, D.; Frezza, M.; Schmitt, S.; Kanwar, J.; P Dou, Q., Bortezomib as the first proteasome inhibitor anticancer drug: current status and future perspectives. *Current cancer drug targets* **2011,** *11*, 239-253.
49. Kane, R. C.; Bross, P. F.; Farrell, A. T.; Pazdur, R., Velcade®: US FDA approval for the treatment of multiple myeloma progressing on prior therapy. *The oncologist* **2003,** *8*, 508-513.
50. Hanada, M.; Sugawara, K.; Kaneta, K.; Toda, S.; Nishiyama, Y.; Tomita, K.; Yamamoto, H.; Konishi, M.; Oki, T., Epoxomicin, a new antitumor agent of microbial origin. *The Journal of antibiotics* **1992,** *45*, 1746-1752.
51. Sugawara K, H. M., Nishiyama Y, Tomita K, Kamei H, Konishi M, Oki T. , Eponemycin, a new antibiotic active against B16 melanoma. I. Production, isolation, structure and biological activity. *The Journal of antibiotics* **1990,** *43*, 8-18.
52. Khan, M. L.; Stewart, A. K., Carfilzomib: a novel second-generation proteasome inhibitor. *Future oncology* **2011,** *7*, 607-612.
53. Kuhn, D. J.; Chen, Q.; Voorhees, P. M.; Strader, J. S.; Shenk, K. D.; Sun, C. M.; Demo, S. D.; Bennett, M. K.; Van Leeuwen, F. W.; Chanan-Khan, A. A., Potent activity of carfilzomib, a novel, irreversible inhibitor of the ubiquitin-proteasome pathway, against preclinical models of multiple myeloma. *Blood, The Journal of the American Society of Hematology* **2007,** *110*, 3281-3290.
54. Curran, M. P.; McKeage, K., Bortezomib. *Drugs* **2009,** *69*, 859-888.
55. McCormack, P. L., Carfilzomib. *Drugs* **2012,** *72*, 2023-2032.
56. Glenn, R. J.; Pemberton, A. J.; Royle, H. J.; Spackman, R. W.; Smith, E.; Rivett, A. J.; Steverding, D., Trypanocidal effect of α′, β′-epoxyketones indicates that trypanosomes are particularly sensitive to inhibitors of proteasome trypsin-like activity. *International journal of antimicrobial agents* **2004,** *24*, 286-289.
57. Czesny, B.; Goshu, S.; Cook, J. L.; Williamson, K. C., The proteasome inhibitor epoxomicin has potent Plasmodium falciparum gametocytocidal activity. *Antimicrobial agents and chemotherapy* **2009,** *53*, 4080-4085.
58. Robinson, S. L.; Christenson, J. K.; Wackett, L. P., Biosynthesis and chemical diversity of β-lactone natural products. *Natural product reports* **2019,** *36*, 458-475.
59. Christenson, J. K.; Richman, J. E.; Jensen, M. R.; Neufeld, J. Y.; Wilmot, C. M.; Wackett, L. P., β-Lactone synthetase found in the olefin biosynthesis pathway. *Biochemistry* **2017,** *56*, 348-351.
60. Weibel, E.; Hadvary, P.; Hochuli, E.; Kupfer, E.; Lengsfeld, H., Lipstatin, an inhibitor of pancreatic lipase, produced by Streptomyces toxytricini I. Producing organism, fermentation, isolation and biological activity. *The Journal of antibiotics* **1987,** *40*, 1081-1085.
61. Wyatt, M. A.; Ahilan, Y.; Argyropoulos, P.; Boddy, C. N.; Magarvey, N. A.; Harrison, P. H., Biosynthesis of ebelactone A: isotopic tracer, advanced precursor and genetic studies

reveal a thioesterase-independent cyclization to give a polyketide β-lactone. *The Journal of Antibiotics* **2013,** *66*, 421-430.
62. Compton, C. L.; Schmitz, K. R.; Sauer, R. T.; Sello, J. K., Antibacterial activity of and resistance to small molecule inhibitors of the ClpP peptidase. *ACS chemical biology* **2013,** *8*, 2669-2677.
63. Feling, R. H.; Buchanan, G. O.; Mincer, T. J.; Kauffman, C. A.; Jensen, P. R.; Fenical, W., Salinosporamide A: a highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus Salinospora. *Angewandte Chemie International Edition* **2003,** *42*, 355-357.
64. Stadler, M.; Bitzer, J.; Mayer-Bartschmid, A.; Müller, H.; Benet-Buchholz, J.; Gantner, F.; Tichy, H.-V.; Reinemer, P.; Bacon, K. B., Cinnabaramides A− G: Analogues of lactacystin and salinosporamide from a terrestrial streptomycete. *Journal of natural products* **2007,** *70*, 246-252.
65. C Potts, B.; X Albitar, M.; C Anderson, K.; Baritaki, S.; Berkers, C.; Bonavida, B.; Chandra, J.; Chauhan, D.; C Cusack, J.; Fenical, W., Marizomib, a proteasome inhibitor for all seasons: preclinical profile and a framework for clinical trials. *Current cancer drug targets* **2011,** *11*, 254-284.
66. Roth, P.; Mason, W. P.; Richardson, P. G.; Weller, M., Proteasome inhibition for the treatment of glioblastoma. *Expert Opinion on Investigational Drugs* **2020,** *29*, 1133-1141.
67. Wu, L.; Ye, K.; Jiang, S.; Zhou, G., Marine Power on Cancer: Drugs, Lead Compounds, and Mechanisms. *Marine Drugs* **2021,** *19*, 488.
68. Böttcher, T.; Sieber, S. A., β-Lactams and β-lactones as activity-based probes in chemical biology. *MedChemComm* **2012,** *3*, 408-417.
69. Asai, A.; Hasegawa, A.; Ochiai, K.; Yamashita, Y.; Mizukami, T., Belactosin A, a Novel Antituor Antibiotic Acting on Cyclin/CDK Mediated Cell Cycle Regulation, Produced by Streptomyces sp. *The Journal of antibiotics* **2000,** *53*, 81-83.
70. Gill, K. A.; Berrué, F.; Arens, J. C.; Carr, G.; Kerr, R. G., Cystargolides, 20S proteasome inhibitors isolated from Kitasatospora cystarginea. *Journal of natural products* **2015,** *78*, 822-826.
71. Asai, A.; Tsujita, T.; Sharma, S. V.; Yamashita, Y.; Akinaga, S.; Funakoshi, M.; Kobayashi, H.; Mizukami, T., A new structural class of proteasome inhibitors identified by microbial screening using yeast-based assay. *Biochemical pharmacology* **2004,** *67*, 227-234.
72. Tello-Aburto, R.; Hallada, L. P.; Niroula, D.; Rogelj, S., Total synthesis and absolute stereochemistry of the proteasome inhibitors cystargolides A and B. *Organic & biomolecular chemistry* **2015,** *13*, 10127-10130.
73. Groll, M.; Larionov, O. V.; Huber, R.; de Meijere, A., Inhibitor-binding mode of homobelactosin C to proteasomes: new insights into class I MHC ligand generation. *Proceedings of the National Academy of Sciences* **2006,** *103*, 4576-4579.
74. Groll, M.; C Potts, B., Proteasome structure, function, and lessons learned from beta-lactone inhibitors. *Current topics in medicinal chemistry* **2011,** *11*, 2850-2878.
75. Rachid, S.; Huo, L.; Herrmann, J.; Stadler, M.; Köpcke, B.; Bitzer, J.; Müller, R., Mining the cinnabaramide biosynthetic pathway to generate novel proteasome inhibitors. *ChemBioChem* **2011,** *12*, 922-931.
76. Umezawa, H.; Aoyagi, T.; Uotani, K.; Hamada, M.; Takeuchi, T.; Takashi, S., Ebelactone, an inhibitor of esterase, produced by actinomycetes. *The Journal of antibiotics* **1980,** *33*, 1594-1596.
77. Wolf, F.; Bauer, J. S.; Bendel, T. M.; Kulik, A.; Kalinowski, J.; Gross, H.; Kaysser, L., Biosynthesis of the beta-Lactone Proteasome Inhibitors Belactosin and Cystargolide. *Angew Chem Int Ed Engl* **2017,** *56*, 6665-6668.

78. Cho, S. W.; Romo, D., Total synthesis of (−)-belactosin C and derivatives via double diastereoselective tandem Mukaiyama aldol lactonizations. *Organic letters* **2007,** *9*, 1537-1540.
79. Armstrong, A.; Scutt, J. N., Total synthesis of (+)-belactosin A. *Chemical communications* **2004**, 510-511.
80. Korotkov, V. S.; Ludwig, A.; Larionov, O. V.; Lygin, A. V.; Groll, M.; de Meijere, A., Synthesis and biological activity of optimized belactosin C congeners. *Organic & biomolecular chemistry* **2011,** *9*, 7791-7798.
81. Vanier, S. F.; Larouche, G.; Wurz, R. P.; Charette, A. B., Formal Synthesis of Belactosin A and Hormaomycin via a Diastereoselective Intramolecular Cyclopropanation of an α-Nitro Diazoester. *Organic letters* **2010,** *12*, 672-675.
82. Larionov, O. V.; de Meijere, A., Enantioselective total syntheses of belactosin A, belactosin C, and its homoanalogue. *Organic letters* **2004,** *6*, 2153-2156.
83. Niroula, D.; Hallada, L. P.; Le Chapelain, C.; Ganegamage, S. K.; Dotson, D.; Rogelj, S.; Groll, M.; Tello-Aburto, R., Design, synthesis, and evaluation of cystargolide-based β-lactones as potent proteasome inhibitors. *European journal of medicinal chemistry* **2018,** *157*, 962-977.
84. Wolf, F.; Bauer, J. S.; Bendel, T. M.; Kulik, A.; Kalinowski, J.; Gross, H.; Kaysser, L., Biosynthesis of the β-Lactone Proteasome Inhibitors Belactosin and Cystargolide. *Angewandte Chemie International Edition* **2017,** *56*, 6665-6668.
85. Blin, K.; Shaw, S.; Kautsar, S. A.; Medema, M. H.; Weber, T., The antiSMASH database version 3: increased taxonomic coverage and new query features for modular enzymes. *Nucleic acids research* **2021,** *49*, D639-D643.
86. De Carvalho, L. P. S.; Blanchard, J. S., Kinetic and chemical mechanism of α-isopropylmalate synthase from Mycobacterium tuberculosis. *Biochemistry* **2006,** *45*, 8988-8999.
87. Textor, S.; De Kraker, J.-W.; Hause, B.; Gershenzon, J.; Tokuhisa, J. G., MAM3 catalyzes the formation of all aliphatic glucosinolate chain lengths in Arabidopsis. *Plant Physiology* **2007,** *144*, 60-71.
88. Wolf, F., Untersuchungen zur Biosynthese von Proteaseinhibitoren aus filamentösen Actinobakterien. *Eberhard Karls Universität Tübingen* **2018**.
89. Steffensky, M.; Li, S.-M.; Heide, L., Cloning, overexpression, and purification of novobiocic acid synthetase from Streptomyces spheroides NCIMB 11891. *Journal of Biological Chemistry* **2000,** *275*, 21754-21760.
90. Dawlaty, J.; Zhang, X.; Fischbach, M. A.; Clardy, J., Dapdiamides, tripeptide antibiotics formed by unconventional amide ligases. *Journal of natural products* **2010,** *73*, 441-446.
91. Noike, M.; Matsui, T.; Ooya, K.; Sasaki, I.; Ohtaki, S.; Hamano, Y.; Maruyama, C.; Ishikawa, J.; Satoh, Y.; Ito, H., A peptide ligase and the ribosome cooperate to synthesize the peptide pheganomycin. *Nature chemical biology* **2015,** *11*, 71-76.
92. Ogasawara, Y.; Kawata, J.; Noike, M.; Satoh, Y.; Furihata, K.; Dairi, T., Exploring Peptide Ligase Orthologs in Actinobacteria・ Discovery of Pseudopeptide Natural Products, Ketomemicins. *ACS chemical biology* **2016,** *11*, 1686-1692.
93. Ooya, K.; Ogasawara, Y.; Noike, M.; Dairi, T., Identification and analysis of the resorcinomycin biosynthetic gene cluster. *Bioscience, biotechnology, and biochemistry* **2015,** *79*, 1833-1837.
94. Ziemert, N.; Ishida, K.; Weiz, A.; Hertweck, C.; Dittmann, E., Exploiting the natural diversity of microviridin gene clusters for discovery of novel tricyclic depsipeptides. *Applied and environmental microbiology* **2010,** *76*, 3568-3574.

95. Parry, R.; Nishino, S.; Spain, J., Naturally-occurring nitro compounds. *Natural product reports* **2011,** *28*, 152-167.
96. Pietruszka, J., Synthesis and properties of oligocyclopropyl-containing natural products and model compounds. *Chemical reviews* **2003,** *103*, 1051-1070.
97. Chen, D. Y.-K.; Pouwer, R. H.; Richard, J.-A., Recent advances in the total synthesis of cyclopropane-containing natural products. *Chemical Society Reviews* **2012,** *41*, 4631-4642.
98. Fan, Y.-Y.; Gao, X.-H.; Yue, J.-M., Attractive natural products with strained cyclopropane and/or cyclobutane ring systems. *Science China Chemistry* **2016,** *59*, 1126-1141.
99. Sun, M.-R.; Li, H.-L.; Ba, M.-Y.; Cheng, W.; Zhu, H.-L.; Duan, Y.-T., Cyclopropyl Scaffold: A Generalist for Marketed Drugs. *Mini Reviews in Medicinal Chemistry* **2021,** *21*, 150-170.
100. Talele, T. T., The "cyclopropyl fragment" is a versatile player that frequently appears in preclinical/clinical drug molecules. *Journal of medicinal chemistry* **2016,** *59*, 8712-8756.
101. Thibodeaux, C. J.; Chang, W.-c.; Liu, H.-w., Enzymatic chemistry of cyclopropane, epoxide, and aziridine biosynthesis. *Chemical reviews* **2012,** *112*, 1681-1709.
102. Ebner, C.; Carreira, E. M., Cyclopropanation Strategies in Recent Total Syntheses. *Chem Rev* **2017,** *117*, 11651-11679.
103. Pellissier, H.; Lattanzi, A.; Dalpozzo, R., *Asymmetric synthesis of three-membered rings*. John Wiley & Sons2017.
104. Grogan, D. W.; Cronan Jr, J. E., Cyclopropane ring formation in membrane lipids of bacteria. *Microbiology and Molecular Biology Reviews* **1997,** *61*, 429-441.
105. Wessjohann, L. A.; Brandt, W.; Thiemann, T., Biosynthesis and metabolism of cyclopropane rings in natural compounds. *Chemical reviews* **2003,** *103*, 1625-1648.
106. Li, Y.; Feng, L.; Kirsch, J. F., Kinetic and spectroscopic investigations of wild-type and mutant forms of apple 1-aminocyclopropane-1-carboxylate synthase. *Biochemistry* **1997,** *36*, 15477-15488.
107. Vaillancourt, F. H.; Yeh, E.; Vosburg, D. A.; O'connor, S. E.; Walsh, C. T., Cryptic chlorination by a non-haem iron enzyme during cyclopropyl amino acid biosynthesis. *Nature* **2005,** *436*, 1191.
108. Vaillancourt, F. H.; Yin, J.; Walsh, C. T., SyrB2 in syringomycin E biosynthesis is a nonheme FeII α-ketoglutarate-and O2-dependent halogenase. *Proceedings of the National Academy of Sciences* **2005,** *102*, 10111-10116.
109. Broberg, A.; Menkis, A.; Vasiliauskas, R., Kutznerides 1− 4, Depsipeptides from the Actinomycete Kutzneria sp. 744 Inhabiting Mycorrhizal Roots of Picea a bies Seedlings. *Journal of natural products* **2006,** *69*, 97-102.
110. Fujimori, D. G.; Hrvatin, S.; Neumann, C. S.; Strieker, M.; Marahiel, M. A.; Walsh, C. T., Cloning and characterization of the biosynthetic gene cluster for kutznerides. *Proceedings of the National Academy of Sciences* **2007,** *104*, 16498-16503.
111. Gu, L.; Wang, B.; Kulkarni, A.; Geders, T. W.; Grindberg, R. V.; Gerwick, L.; Håkansson, K.; Wipf, P.; Smith, J. L.; Gerwick, W. H., Metamorphic enzyme assembly in polyketide diversification. *Nature* **2009,** *459*, 731.
112. Kelly, W. L.; Boyne, M. T.; Yeh, E.; Vosburg, D. A.; Galonić, D. P.; Kelleher, N. L.; Walsh, C. T., Characterization of the aminocarboxycyclopropane-forming enzyme CmaC. *Biochemistry* **2007,** *46*, 359-368.
113. Garneau, S.; Dorrestein, P. C.; Kelleher, N. L.; Walsh, C. T., Characterization of the formation of the pyrrole moiety during clorobiocin and coumermycin A1 biosynthesis. *Biochemistry* **2005,** *44*, 2770-2780.

114. Thomas, M. G.; Burkart, M. D.; Walsh, C. T., Conversion of L-proline to pyrrolyl-2-carboxyl-S-PCP during undecylprodigiosin and pyoluteorin biosynthesis. *Chemistry & biology* **2002,** *9*, 171-184.
115. Zlatopolskiy, B. D.; de Meijere, A., First total synthesis of hormaomycin, a naturally occurring depsipeptide with interesting biological activities. *Chemistry–A European Journal* **2004,** *10*, 4718-4727.
116. Andres, N.; Wolf, H.; Zähner, H., Hormaomycin, a new peptide lactone antibiotic effective in inducing cytodifferentiation and antibiotic biosynthesis in some Streptomyces species. *Zeitschrift für Naturforschung C* **1990,** *45*, 851-855.
117. Hofer, I.; Crusemann, M.; Radzom, M.; Geers, B.; Flachshaar, D.; Cai, X.; Zeeck, A.; Piel, J., Insights into the biosynthesis of hormaomycin, an exceptionally complex bacterial signaling metabolite. *Chem Biol* **2011,** *18*, 381-91.
118. Otoguro, K.; Ui, H.; Ishiyama, A.; Arai, N.; Kobayashi, M.; Takahashi, Y.; Masuma, R.; Shiomi, K.; Yamada, H.; Omura, S., In vitro antimalarial activities of the microbial metabolites. *The Journal of antibiotics* **2003,** *56*, 322-324.
119. Höfer, I.; Crüsemann, M.; Radzom, M.; Geers, B.; Flachshaar, D.; Cai, X.; Zeeck, A.; Piel, J., Insights into the biosynthesis of hormaomycin, an exceptionally complex bacterial signaling metabolite. *Chemistry & biology* **2011,** *18*, 381-391.
120. Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* **1997,** *25*, 3389-3402.
121. Wheeler, D.; Bhagwat, M., BLAST quickstart. In *Comparative genomics*, Springer2007; pp 149-175.
122. Medema, M. H.; Takano, E.; Breitling, R., Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Molecular biology and evolution* **2013,** *30*, 1218-1223.
123. Dhakal, D.; Sohng, J. K., Laboratory maintenance of Nocardia species. *Current protocols in microbiology* **2015,** *39*, 10F. 1.1-10F. 1.8.
124. McNeil, M. M.; Brown, J. M., The medically important aerobic actinomycetes: epidemiology and microbiology. *Clinical microbiology reviews* **1994,** *7*, 357-417.
125. Brown-Elliott, B. A.; Brown, J. M.; Conville, P. S.; Wallace Jr, R. J., Clinical and laboratory features of the Nocardia spp. based on current molecular taxonomy. *Clinical microbiology reviews* **2006,** *19*, 259-282.
126. Luo, Q.; Hiessl, S.; Steinbüchel, A., Functional diversity of N ocardia in metabolism. *Environmental microbiology* **2014,** *16*, 29-48.
127. Nocard, E., Note sur la maladie des boeufs de la Guadeloupe connue sous le nom de farcin. *Ann L'Inst Pasteur* **1888,** *2*, 293-302.
128. Beaman, B. L.; Beaman, L., Nocardia species: host-parasite relationships. *Clinical microbiology reviews* **1994,** *7*, 213-264.
129. Fatahi-Bafghi, M., Nocardiosis from 1888 to 2017. *Microbial pathogenesis* **2018,** *114*, 369-384.
130. Steinbrink, J.; Leavens, J.; Kauffman, C. A.; Miceli, M. H., Manifestations and outcomes of nocardia infections: comparison of immunocompromised and nonimmunocompromised adult patients. *Medicine* **2018,** *97*.
131. Wilson, J. W. In *Nocardiosis: updates and clinical overview*, Mayo Clinic Proceedings2012 Elsevier; pp 403-407.
132. Conville, P. S.; Brown-Elliott, B. A.; Smith, T.; Zelazny, A. M., The complexities of Nocardia taxonomy and identification. *Journal of clinical microbiology* **2018,** *56*.
133. Mehta, H. H.; Shamoo, Y., Pathogenic Nocardia: A diverse genus of emerging pathogens or just poorly recognized? *PLoS pathogens* **2020,** *16*, e1008280.
134. Saubolle, M. A.; Sussland, D., Nocardiosis review of clinical and laboratory experience. *Journal of clinical microbiology* **2003,** *41*, 4497-4501.

135. Conville, P. S.; Witebsky, F. G., Organisms designated as Nocardia asteroides drug pattern type VI are members of the species Nocardia cyriacigeorgica. *Journal of clinical microbiology* **2007,** *45*, 2257-2259.
136. Corti, M. E.; Fioti, M. E. V., Nocardiosis: a review. *International journal of infectious Diseases* **2003,** *7*, 243-250.
137. Rafiei, N.; Peri, A. M.; Righi, E.; Harris, P.; Paterson, D. L., Central nervous system nocardiosis in Queensland: a report of 20 cases and review of the literature. *Medicine* **2016,** *95*.
138. Takiguchi, Y.; Ishizaki, S.; Kobayashi, T.; Sato, S.; Hashimoto, Y.; Suruga, Y.; Akiba, Y., Pulmonary nocardiosis: a clinical analysis of 30 cases. *Internal Medicine* **2017,** *56*, 1485-1490.
139. Ji, X.; Zhang, X.; Sun, L.; Hou, X.; Song, J.; Tan, X.; Song, H.; Qiu, X.; Li, M.; Tang, L., Mce1C and Mce1D facilitate N. farcinica invasion of host cells and suppress immune responses by inhibiting innate signaling pathways. *Scientific reports* **2020,** *10*, 1-13.
140. Vera-Cabrera, L.; Ortiz-Lopez, R.; Elizondo-Gonzalez, R.; Ocampo-Candiani, J., Complete genome sequence analysis of Nocardia brasiliensis HUJEG-1 reveals a saprobic lifestyle and the genes needed for human pathogenesis. *PLoS One* **2013,** *8*, e65425.
141. Zoropogui, A.; Pujic, P.; Normand, P.; Barbe, V.; Belli, P.; Graindorge, A.; Roche, D.; Vallenet, D.; Mangenot, S.; Boiron, P., The Nocardia cyriacigeorgica GUH-2 genome shows ongoing adaptation of an environmental Actinobacteria to a pathogen’s lifestyle. *BMC genomics* **2013,** *14*, 1-18.
142. Engelbrecht, A.; Saad, H.; Gross, H.; Kaysser, L., Natural Products from Nocardia and Their Role in Pathogenicity. *Microbial Physiology* **2021**, 1-16.
143. Maglangit, F.; Yu, Y.; Deng, H., Bacterial pathogens: threat or treat (a review on bioactive natural products from bacterial pathogens). *Natural Product Reports* **2020**.
144. Dhakal, D.; Rayamajhi, V.; Mishra, R.; Sohng, J. K., Bioactive molecules from Nocardia: diversity, bioactivities and biosynthesis. *Journal of industrial microbiology & biotechnology* **2019,** *46*, 385-407.
145. Mikami, Y., Biological work on medically important Nocardia species. *Actinomycetologica* **2007,** *21*, 46-51.
146. Shigemori, H.; Komaki, H.; Yazawa, K.; Mikami, Y.; Nemoto, A.; Tanaka, Y.; Sasaki, T.; In, Y.; Ishida, T.; Kobayashi, J. i., Brasilicardin A. A Novel Tricyclic Metabolite with Potent Immunosuppressive Activity from Actinomycete Nocardia b rasiliensis. *The Journal of organic chemistry* **1998,** *63*, 6900-6904.
147. Miethke, M.; Marahiel, M. A., Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol Rev* **2007,** *71*, 413-51.
148. Sritharan, M., Iron as a candidate in virulence and pathogenesis in mycobacteria and other microorganisms. *World Journal of Microbiology and Biotechnology* **2000,** *16*, 769-780.
149. Arai, M. A.; Ebihara, I.; Makita, Y.; Hara, Y.; Yaguchi, T.; Ishibashi, M., Isolation of nocobactin NAs as Notch signal inhibitors from Nocardia farcinica, a possibility of invasive evolution. *The Journal of Antibiotics* **2021,** *74*, 255-259.
150. Chen, J.; Frediansyah, A.; Männle, D.; Straetener, J.; Brötz-Oesterhelt, H.; Ziemert, N.; Kaysser, L.; Gross, H., New Nocobactin Derivatives with Antimuscarinic Activity, Terpenibactins A‐C, Revealed by Genome Mining of Nocardia terpenica IFM 0406. *ChemBioChem* **2020,** *21*, 2205.
151. Quadri, L. E.; Sello, J.; Keating, T. A.; Weinreb, P. H.; Walsh, C. T., Identification of a Mycobacterium tuberculosis gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. *Chemistry & biology* **1998,** *5*, 631-645.
152. Hoshino, Y.; Chiba, K.; Ishino, K.; Fukai, T.; Igarashi, Y.; Yazawa, K.; Mikami, Y.; Ishikawa, J., Identification of nocobactin NA biosynthetic gene clusters in Nocardia farcinica. *Journal of bacteriology* **2011,** *193*, 441-448.

153. Arai, M. A.; Ebihara, I.; Makita, Y.; Hara, Y.; Yaguchi, T.; Ishibashi, M., Isolation of nocobactin NAs as Notch signal inhibitors from Nocardia farcinica, a possibility of invasive evolution. *The Journal of antibiotics* **2020**, 1-5.
154. Robinson, S. L.; Terlouw, B. R.; Smith, M. D.; Pidot, S. J.; Stinear, T. P.; Medema, M. H.; Wackett, L. P., Global analysis of adenylate-forming enzymes reveals β-lactone biosynthesis pathway in pathogenic Nocardia. *Journal of Biological Chemistry* **2020,** *295*, 14826-14839.
155. Mikami, Y.; Yazawa, Y.; Tanaka, Y.; Ritzau, M.; Gräfeb, U., Isolation and structure of nocardiolactone, a new dialkyl-substituted β-lactone from pathogenic Nocardia strains. *Natural Product Letters* **1999,** *13*, 277-284.
156. Matsunaga, I.; Sugita, M., New insights into lipidic secondary metabolites in mycobacteria. *Current Chemical Biology* **2011,** *5*, 52-63.
157. Silva, C. L. d., Inflammation induced by mycolic acid-containing glycolipids of Mycobacterium bovis (BCG). *Brazilian Journal of Medical and Biological Research= Revista Brasileira de Pesquisas Medicas e Biologicas* **1985,** *18*, 327-335.
158. Beaman, B. L., Mechanisms for the virulence of Nocardia. In *Molecular Mechanisms of Bacterial Virulence*, Springer1994; pp 561-572.
159. Ryll, R.; Kumazawa, Y.; Yano, I., Immunological Properties of Trehalose Dimycolate (Cord Factor) and Other Mycotic Acid-Containing Glycolipids‐-A Review. *Microbiology and immunology* **2001,** *45*, 801-811.
160. Rosas-Taraco, A. G.; Perez-Liñan, A. R.; Bocanegra-Ibarias, P.; Perez-Rivera, L. I.; Salinas-Carmona, M. C., Nocardia brasiliensis induces an immunosuppressive microenvironment that favors chronic infection in BALB/c mice. *Infection and immunity* **2012,** *80*, 2493-2499.
161. Komaki, H.; Nemoto, A.; Tanaka, Y.; Takagi, H.; Yazawa, K.; Mikami, Y.; Shigemori, H.; Kobayashi, J. i.; Ando, A.; Nagata, Y., Brasilicardin A, a new terpenoid antibiotic from pathogenic Nocardia brasiliensis: fermentation, isolation and biological activity. *The Journal of antibiotics* **1999,** *52*, 13-19.
162. Komatsu, K.; Tsuda, M.; Tanaka, Y.; Mikami, Y.; Kobayashi, J., SAR studies of brasilicardin A for immunosuppressive and cytotoxic activities. *Bioorg Med Chem* **2005,** *13*, 1507-13.
163. Usui, T.; Nagumo, Y.; Watanabe, A.; Kubota, T.; Komatsu, K.; Kobayashi, J. i.; Osada, H., Brasilicardin A, a natural immunosuppressant, targets amino acid transport system L. *Chemistry & biology* **2006,** *13*, 1153-1160.
164. Hideyuki, S.; Yasushi, T.; Katsukiyo, Y.; Yuzuru, M.; Jun'ichi, K., Brasilinolide A, new immunosuppressive macrolide from actinomycete Nocardia brasiliensis. *Tetrahedron* **1996,** *52*, 9031-9034.
165. Hara, Y.; Arai, M. A.; Toume, K.; Masu, H.; Sato, T.; Komatsu, K.; Yaguchi, T.; Ishibashi, M., Coculture of a Pathogenic Actinomycete and Animal Cells To Produce Nocarjamide, a Cyclic Nonapeptide with Wnt Signal-Activating Effect. *Org Lett* **2018,** *20*, 5831-5834.
166. Kuo, J.; Lynch, S. R.; Liu, C. W.; Xiao, X.; Khosla, C., Partial in vitro reconstitution of an orphan polyketide synthase associated with clinical cases of nocardiosis. *ACS chemical biology* **2016,** *11*, 2636-2641.
167. Yuet, K. P.; Liu, C. W.; Lynch, S. R.; Kuo, J.; Michaels, W.; Lee, R. B.; McShane, A. E.; Zhong, B. L.; Fischer, C. R.; Khosla, C., Complete Reconstitution and Deorphanization of the 3 MDa Nocardiosis-Associated Polyketide Synthase. *Journal of the American Chemical Society* **2020,** *142*, 5952-5957.
168. Schorn, M.; Zettler, J.; Noel, J. P.; Dorrestein, P. C.; Moore, B. S.; Kaysser, L., Genetic basis for the biosynthesis of the pharmaceutically important class of epoxyketone proteasome inhibitors. *ACS Chem Biol* **2014,** *9*, 301-9.

169. Zettler, J.; Zubeil, F.; Kulik, A.; Grond, S.; Kaysser, L., Epoxomicin and Eponemycin Biosynthesis Involves gem-Dimethylation and an Acyl-CoA Dehydrogenase-Like Enzyme. *Chembiochem* **2016,** *17*, 792-8.
170. Leipoldt, F.; Santos-Aberturas, J.; Stegmann, D. P.; Wolf, F.; Kulik, A.; Lacret, R.; Popadic, D.; Keinhorster, D.; Kirchner, N.; Bekiesch, P.; Gross, H.; Truman, A. W.; Kaysser, L., Warhead biosynthesis and the origin of structural diversity in hydroxamate metalloproteinase inhibitors. *Nature communications* **2017,** *8*, 1965.
171. Barry, D. P.; Beaman, B. L., Nocardia asteroides strain GUH-2 induces proteasome inhibition and apoptotic death of cultured cells. *Research in microbiology* **2007,** *158*, 86-96.
172. Loeffler, D. A.; Camp, D. M.; Qu, S.; Beaman, B. L.; LeWitt, P. A., Characterization of dopamine-depleting activity of Nocardia asteroides strain GUH-2 culture filtrate on PC12 cells. *Microbial pathogenesis* **2004,** *37*, 73-85.
173. Brandl, M.; Kozhushkov, S. I.; Zlatopolskiy, B. D.; Alvermann, P.; Geers, B.; Zeeck, A.; de Meijere, A., The Biosynthesis of 3-(trans-2-Nitrocyclopropyl) alanine, a Constituent of the Signal Metabolite Hormaomycin. *European journal of organic chemistry* **2005,** *2005*, 123-135.
174. Kozhushkov, S. I.; Zlatopolskiy, B. D.; Brandl, M.; Alvermann, P.; Radzom, M.; Geers, B.; de Meijere, A.; Zeeck, A., Hormaomycin Analogues by Precursor-Directed Biosynthesis‐Synthesis of and Feeding Experiments with Amino Acids Related to the Unique 3-(trans-2-Nitrocyclopropyl) alanine Constituent. Wiley Online Library2005.
175. Russell, D. W.; Sambrook, J., *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Cold Spring Harbor, NY2001; Vol. 1.
176. Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A., *Practical streptomyces genetics*. John Innes Foundation Norwich2004; Vol. 291.
177. Jez, J. M.; Bowman, M. E.; Noel, J. P., Structure-guided programming of polyketide chain-length determination in chalcone synthase. *Biochemistry* **2001,** *40*, 14829-14838.
178. Kaysser, L.; Bernhardt, P.; Nam, S.-J.; Loesgen, S.; Ruby, J. G.; Skewes-Cox, P.; Jensen, P. R.; Fenical, W.; Moore, B. S., Merochlorins A‐D, cyclic meroterpenoid antibiotics biosynthesized in divergent pathways with vanadium-dependent chloroperoxidases. *Journal of the American Chemical Society* **2012,** *134*, 11988-11991.
179. MacNeil, D. J.; Gewain, K. M.; Ruby, C. L.; Dezeny, G.; Gibbons, P. H.; MacNeil, T., Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **1992,** *111*, 61-68.
180. Hanahan, D., Studies on transformation of Escherichia coli with plasmids. *Journal of molecular biology* **1983,** *166*, 557-580.
181. Durfee, T.; Nelson, R.; Baldwin, S.; Plunkett III, G.; Burland, V.; Mau, B.; Petrosino, J. F.; Qin, X.; Muzny, D. M.; Ayele, M., The complete genome sequence of Escherichia coli DH10B: insights into the biology of a laboratory workhorse. *Journal of bacteriology* **2008,** *190*, 2597-2606.
182. Datsenko, K. A.; Wanner, B. L., One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences* **2000,** *97*, 6640-6645.
183. Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F., PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proceedings of the National Academy of Sciences* **2003,** *100*, 1541-1546.
184. Cherepanov, P. P.; Wackernagel, W., Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **1995,** *158*, 9-14.

185. Flett, F.; Mersinias, V.; Smith, C. P., High efficiency intergeneric conjugal transfer of plasmid DNA from Escherichia coli to methyl DNA-restricting streptomycetes. *FEMS microbiology letters* **1997,** *155*, 223-229.
186. Schwarz, P. N.; Buchmann, A.; Roller, L.; Kulik, A.; Gross, H.; Wohlleben, W.; Stegmann, E., The immunosuppressant brasilicardin: determination of the biosynthetic gene cluster in the heterologous host Amycolatopsis japonicum. *Biotechnology journal* **2018,** *13*, 1700527.
187. Nishikiori, T.; Okuyama, A.; Naganawa, H.; Takita, T.; Hamada, M.; Takeuchi, T.; Aoyagi, T.; Umezawa, H., Production by actinomycetes of (S, S)-N, N-ethylenediamine-disuccinic acid, an inhibitor of phospholipase c. The Journal of antibiotics **1984,** *37*, 426-427.
188. Chater, K. F.; Wilde, L. C., Streptomyces albus G mutants defective in the SalGI restriction-modification system. *Microbiology* **1980,** *116*, 323-334.
189. Zaburannyi, N.; Rabyk, M.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A., Insights into naturally minimised Streptomyces albus J1074 genome. *BMC genomics* **2014,** *15*, 1-11.
190. Schorn, M.; Zettler, J.; Noel, J. P.; Dorrestein, P. C.; Moore, B. S.; Kaysser, L., Genetic basis for the biosynthesis of the pharmaceutically important class of epoxyketone proteasome inhibitors. *ACS chemical biology* **2014,** *9*, 301-309.
191. Amatuni, A.; Renata, H., Identification of a lysine 4-hydroxylase from the glidobactin biosynthesis and evaluation of its biocatalytic potential. *Organic & biomolecular chemistry* **2019**.
192. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Wilkins, M. R.; Appel, R. D.; Bairoch, A., Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook* **2005**, 571-607.
193. Laemmli, U. K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *nature* **1970,** *227*, 680-685.
194. Epicentre, CopyControl™ Fosmid Library Production Kit. In *CCFOS110, Cat NoCCFOS059, Cat No*
195. Engelbrecht, A.; Wolf, F.; Esch, A.; Kulik, A.; Kozhushkov, S. I.; de Meijere, A.; Hughes, C. C.; Kaysser, L., Discovery of a Cryptic Nitro Intermediate in the Biosynthesis of the 3-(trans-2′-Aminocyclopropyl) alanine Moiety of Belactosin A. *Organic Letters* **2022,** *24*, 736-740.
196. Kozhushkov, S. I.; Zlatopolskiy, B. D.; Brandl, M.; Alvermann, P.; Radzom, M.; Geers, B.; de Meijere, A.; Zeeck, A., Hormaomycin Analogues by Precursor-Directed Biosynthesis‐Synthesis of and Feeding Experiments with Amino Acids Related to the Unique 3-(trans-2-Nitrocyclopropyl) alanine Constituent. *European journal of organic chemistry* **2005,** *2005*, 854-863.
197. Kozhushkov, S. I.; Zlatopolskiy, B. D.; Brandl, M.; Alvermann, P.; Radzom, M.; Geers, B.; de Meijere, A.; Zeeck, A., Hormaomycin Analogues by Precursor-Directed Biosynthesis - Synthesis of and Feeding Experiments with Amino Acids Related to the Unique 3-(trans-2-Nitrocyclopropyl)alanine Constituent. *European Journal of Organic Chemistry* **2005,** *2005*, 854-863.
198. Brandl, M.; Kozhushkov, S. I.; Loscha, K.; Kokoreva, O. V.; Yufit, D. S.; Howard, J. A.; de Meijere, A., Synthesis of trans-(2-Aminocyclopropyl) alanine-A Key Constituent of the Novel Antitumor Antibiotic Belactosin A. *Synlett* **2000,** *2000*, 1741-1744.
199. Gust, B., Cloning and analysis of natural product pathways. *Methods in enzymology* **2009,** *458*, 159-180.
200. Hille, R.; Hall, J.; Basu, P., The mononuclear molybdenum enzymes. *Chem Rev* **2014,** *114*, 3963-4038.
201. Shimo, S.; Ushimaru, R.; Engelbrecht, A.; Harada, M.; Miyamoto, K.; Kulik, A.; Uchiyama, M.; Kaysser, L.; Abe, I., Stereodivergent Nitrocyclopropane Formation during

Biosynthesis of Belactosins and Hormaomycins. *Journal of the American Chemical Society* **2021**.
202. Bräuer, A.; Beck, P.; Hintermann, L.; Groll, M., Structure of the Dioxygenase AsqJ: Mechanistic Insights into a One-Pot Multistep Quinolone Antibiotic Biosynthesis. *Angewandte Chemie International Edition* **2016,** *55*, 422-426.
203. Kodera, T.; Smirnov, S. V.; Samsonova, N. N.; Kozlov, Y. I.; Koyama, R.; Hibi, M.; Ogawa, J.; Yokozeki, K.; Shimizu, S., A novel L-isoleucine hydroxylating enzyme, L-isoleucine dioxygenase from Bacillus thuringiensis, produces (2S, 3R, 4S)-4-hydroxyisoleucine. *Biochemical and biophysical research communications* **2009,** *390*, 506-510.
204. Zwick, C. R.; Renata, H., Harnessing the biocatalytic potential of iron-and α-ketoglutarate-dependent dioxygenases in natural product total synthesis. *Natural product reports* **2020,** *37*, 1065-1079.
205. Zwick III, C. R.; Renata, H., Remote C‑H hydroxylation by an α-ketoglutarate-dependent dioxygenase enables efficient chemoenzymatic synthesis of manzacidin C and proline analogs. *Journal of the American Chemical Society* **2018,** *140*, 1165-1169.
206. Hausinger, R. P., Fe (II)/α-ketoglutarate-dependent hydroxylases and related enzymes. *Critical reviews in biochemistry and molecular biology* **2004,** *39*, 21-68.
207. Yang, Z.; Chi, X.; Funabashi, M.; Baba, S.; Nonaka, K.; Pahari, P.; Unrine, J.; Jacobsen, J. M.; Elliott, G. I.; Rohr, J., Characterization of LipL as a non-heme, Fe (II)-dependent α-ketoglutarate: UMP dioxygenase that generates uridine-5′-aldehyde during A-90289 biosynthesis. *Journal of Biological Chemistry* **2011,** *286*, 7885-7892.
208. Hughes, C. C., Chemical labeling strategies for small molecule natural product detection and isolation. *Natural Product Reports* **2021**.
209. Gilar, M.; Jaworski, A., Retention behavior of peptides in hydrophilic-interaction chromatography. *Journal of Chromatography A* **2011,** *1218*, 8890-8896.
210. Huang, Y.; Li, W.; Keller, A.; Anumol, T.; Bivens, A., Quantitative Analysis of Underivatized Amino Acids in Plant Matrix by Hydrophilic Interaction Chromatography (HILIC) with LC/MS Detection. *Application Note* **2018**.
211. Challinor, V. L.; Bode, H. B., Bioactive natural products from novel microbial sources. *Annals of the New York Academy of Sciences* **2015,** *1354*, 82-97.
212. Navarro-Muñoz, J. C.; Selem-Mojica, N.; Mullowney, M. W.; Kautsar, S. A.; Tryon, J. H.; Parkinson, E. I.; De Los Santos, E. L.; Yeong, M.; Cruz-Morales, P.; Abubucker, S., A computational framework to explore large-scale biosynthetic diversity. *Nature chemical biology* **2020,** *16*, 60-68.
213. Schorn, M.; Zettler, J.; Noel, J. P.; Dorrestein, P. C.; Moore, B. S.; Kaysser, L., Genetic basis for the biosynthesis of the pharmaceutically important class of epoxyketone proteasome inhibitors. *ACS chemical biology* **2013,** *9*, 301-309.
214. Liu, J.; Zhu, X.; Zhang, W., Identifying the minimal enzymes required for biosynthesis of epoxyketone proteasome inhibitors. *Chembiochem* **2015,** *16*, 2585-2589.
215. Barry, D. P.; Beaman, B. L., Nocardia asteroides strain GUH-2 induces proteasome inhibition and apoptotic death of cultured cells. *Research in microbiology* **2007,** *158*, 86-96.
216. Ji, X.; Tan, X.; Hou, X.; Si, C.; Xu, S.; Tang, L.; Yuan, X.; Li, Z., Cloning, expression, invasion, and immunological reactivity of a mammalian cell entry protein encoded by the mce1 operon of Nocardia farcinica. *Frontiers in microbiology* **2017,** *8*, 281.
217. Ji, X.; Zhang, X.; Li, H.; Sun, L.; Hou, X.; Song, H.; Han, L.; Xu, S.; Qiu, X.; Wang, X., Nfa34810 facilitates Nocardia farcinica invasion of host cells and stimulates TNF-α secretion through activation of the NF-κB and MAPK pathways via TLR4. *Infection and Immunity* **2020**.

218. Ott, L.; Höller, M.; Gerlach, R. G.; Hensel, M.; Rheinlaender, J.; Schäffer, T. E.; Burkovski, A., Corynebacterium diphtheriae invasion-associated protein (DIP1281) is involved in cell surface organization, adhesion and internalization in epithelial cells. *BMC microbiology* **2010,** *10*, 1-9.
219. Zhang, J. J.; Tang, X.; Moore, B. S., Genetic platforms for heterologous expression of microbial natural products. *Natural product reports* **2019,** *36*, 1313-1332.
220. Gust, B.; Chandra, G.; Jakimowicz, D.; Yuqing, T.; Bruton, C. J.; Chater, K. F., λ Red-mediated genetic manipulation of antibiotic-producing Streptomyces. *Advances in applied microbiology* **2004,** *54*, 107-128.
221. Schwarz, P. N.; Roller, L.; Kulik, A.; Wohlleben, W.; Stegmann, E., Engineering metabolic pathways in Amycolatopsis japonicum for the optimization of the precursor supply for heterologous brasilicardin congeners production. *Synthetic and systems biotechnology* **2018,** *3*, 56-63.
222. from, T. D. o. N. P. d. i. a.; http://www.crcpress.com/, C. H. C. a. U.
223. Blunt, J.; Munro, M.; Upjohn, M., The role of databases in marine natural products research. *Handbook of marine natural products* **2012,** *1*, 389-422.
224. Sezaki, M.; Kondo, S.; Maeda, K.; Umezawa, H.; Ohno, M., The structure of aquayamycin. *Tetrahedron* **1970,** *26*, 5171-5190.
225. Bierman, M.; Logan, R.; O'brien, K.; Seno, E.; Rao, R. N.; Schoner, B., Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp. *Gene* **1992,** *116*, 43-49.
226. Gust, B.; Kieser, T.; Chater, K., PCR targeting system in Streptomyces coelicolor A3 (2). *John Innes Centre* **2002,** *3*, 1-39.
227. Baulig, A.; Helmle, I.; Bader, M.; Wolf, F.; Kulik, A.; Al-Dilaimi, A.; Wibberg, D.; Kalinowski, J.; Gross, H.; Kaysser, L., Biosynthetic reconstitution of deoxysugar phosphoramidate metalloprotease inhibitors using an N-P-bond-forming kinase. *Chemical science* **2019,** *10*, 4486-4490.
228. Eustáquio, A. S.; Gust, B.; Galm, U.; Li, S.-M.; Chater, K. F.; Heide, L., Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters. *Applied and environmental microbiology* **2005,** *71*, 2452-2459.
229. Huo, L.; Hug, J. J.; Fu, C.; Bian, X.; Zhang, Y.; Müller, R., Heterologous expression of bacterial natural product biosynthetic pathways. *Natural product reports* **2019,** *36*, 1412-1436.
230. Gomez-Escribano, J. P.; Bibb, M. J., Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters. *Microbial biotechnology* **2011,** *4*, 207-215.
231. Ahmed, Y.; Rebets, Y.; Estévez, M. R.; Zapp, J.; Myronovskyi, M.; Luzhetskyy, A., Engineering of Streptomyces lividans for heterologous expression of secondary metabolite gene clusters. *Microbial cell factories* **2020,** *19*, 1-16.
232. Ikeda, H.; Shin-ya, K.; Omura, S., Genome mining of the Streptomyces avermitilis genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters. *Journal of Industrial Microbiology and Biotechnology* **2014,** *41*, 233-250.
233. Park, S. R.; Park, J. W.; Jung, W. S.; Han, A. R.; Ban, Y.-H.; Kim, E. J.; Sohng, J. K.; Sim, S. J.; Yoon, Y. J., Heterologous production of epothilones B and D in Streptomyces venezuelae. *Applied microbiology and biotechnology* **2008,** *81*, 109-117.
234. Park, D.; Swayambhu, G.; Pfeifer, B. A., Heterologous biosynthesis as a platform for producing new generation natural products. *Current Opinion in Biotechnology* **2020,** *66*, 123-130.
235. Chater, K. F.; Wilde, L. C., Streptomyces albus G mutants defective in the SalGI restriction-modification system. *J Gen Microbiol* **1980,** *116*, 323-34.

236. Kallifidas, D.; Jiang, G.; Ding, Y.; Luesch, H., Rational engineering of Streptomyces albus J1074 for the overexpression of secondary metabolite gene clusters. *Microbial cell factories* **2018,** *17*, 1-14.
237. Myronovskyi, M.; Rosenkränzer, B.; Nadmid, S.; Pujic, P.; Normand, P.; Luzhetskyy, A., Generation of a cluster-free Streptomyces albus chassis strains for improved heterologous expression of secondary metabolite clusters. *Metabolic engineering* **2018,** *49*, 316-324.
238. Youngblut, M. D.; Tsai, C.-L.; Clark, I. C.; Carlson, H. K.; Maglaqui, A. P.; Gau-Pan, P. S.; Redford, S. A.; Wong, A.; Tainer, J. A.; Coates, J. D., Perchlorate reductase is distinguished by active site aromatic gate residues. *Journal of Biological Chemistry* **2016,** *291*, 9190-9202.
239. Chen, W.; Huang, T.; He, X.; Meng, Q.; You, D.; Bai, L.; Li, J.; Wu, M.; Li, R.; Xie, Z., Characterization of the polyoxin biosynthetic gene cluster from Streptomyces cacaoi and engineered production of polyoxin H. *Journal of Biological Chemistry* **2009,** *284*, 10627-10638.
240. Wu, S.; Jian, X.-H.; Yuan, H.; Jin, W.-B.; Yin, Y.; Wang, L.-Y.; Zhao, J.; Tang, G.-L., Unified biosynthetic origin of the benzodipyrrole subunits in CC-1065. *ACS chemical biology* **2017,** *12*, 1603-1610.
241. Rinkel, J.; Dickschat, J. S., Recent highlights in biosynthesis research using stable isotopes. *Beilstein journal of organic chemistry* **2015,** *11*, 2493-2508.
242. Alterman, M. A.; Hunziker, P., *Amino acid analysis*. Springer2012.
243. Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J., The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols* **2015,** *10*, 845-858.
244. Cai, X., Molecular Analysis of Genes Involved in the Biosynthesis and Regulation of Hormaomycin, an Exceptionally Complex Bacterial Signaling Metabolite. **2013**.
245. Gao, S.-S.; Naowarojna, N.; Cheng, R.; Liu, X.; Liu, P., Recent examples of α-ketoglutarate-dependent mononuclear non-haem iron enzymes in natural product biosyntheses. *Natural product reports* **2018,** *35*, 792-837.
246. Hausinger, R. P., Biochemical diversity of 2-oxoglutarate-dependent oxygenases. **2015**.
247. Wu, L.-F.; Meng, S.; Tang, G.-L., Ferrous iron and α-ketoglutarate-dependent dioxygenases in the biosynthesis of microbial natural products. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2016,** *1864*, 453-470.
248. Irvin, J.; Ropelewski, A. J.; Perozich, J., In silico analysis of heme oxygenase structural homologues identifies group-specific conservations. *FEBS Open Bio* **2017,** *7*, 1480-1498.
249. Ng, T. L.; Rohac, R.; Mitchell, A. J.; Boal, A. K.; Balskus, E. P., An N-nitrosating metalloenzyme constructs the pharmacophore of streptozotocin. *Nature* **2019,** *566*, 94-99.
250. He, H.-Y.; Henderson, A. C.; Du, Y.-L.; Ryan, K. S., Two-enzyme pathway links L-arginine to nitric oxide in N-Nitroso biosynthesis. *Journal of the American Chemical Society* **2019,** *141*, 4026-4033.
251. McBride, M. J.; Sil, D.; Ng, T. L.; Crooke, A. M.; Kenney, G. E.; Tysoe, C. R.; Zhang, B.; Balskus, E. P.; Boal, A. K.; Krebs, C., A peroxodiiron (III/III) intermediate mediating both N-hydroxylation steps in biosynthesis of the N-nitrosourea pharmacophore of streptozotocin by the multi-domain metalloenzyme SznF. *Journal of the American Chemical Society* **2020,** *142*, 11818-11828.
252. McBride, M. J.; Pope, S. R.; Hu, K.; Okafor, C. D.; Balskus, E. P.; Bollinger, J. M.; Boal, A. K., Structure and assembly of the diiron cofactor in the heme-oxygenase‐like domain of the N-nitrosourea‐producing enzyme SznF. *Proceedings of the National Academy of Sciences* **2021,** *118*.

253. Smirnov, S. V.; Sokolov, P. M.; Kodera, T.; Sugiyama, M.; Hibi, M.; Shimizu, S.; Yokozeki, K.; Ogawa, J., A novel family of bacterial dioxygenases that catalyse the hydroxylation of free L-amino acids. *FEMS microbiology letters* **2012,** *331*, 97-104.
254. Xu, Q.; Grant, J.; Chiu, H. J.; Farr, C. L.; Jaroszewski, L.; Knuth, M. W.; Miller, M. D.; Lesley, S. A.; Godzik, A.; Elsliger, M. A., Crystal structure of a member of a novel family of dioxygenases (PF10014) reveals a conserved cupin fold and active site. *Proteins: Structure, Function, and Bioinformatics* **2014,** *82*, 164-170.
255. Patrick Beller Phillipp J. Fink, I. H., Daniel Männle, Daniel Unterfrauner, Felix Wolf, Alicia Engelbrecht, Nicole Staudt, Andreas Kulik, Harald Groß and Leonard Kaysser, Characterization of the Cystargolide Biosynthetic Gene Cluster. *Manuscript in preparation* **2023**.
256. Cao, X.; Zhu, L.; Hu, Z.; Cronan, J. E., Expression and activity of the BioH esterase of biotin synthesis is independent of genome context. *Scientific reports* **2017,** *7*, 1-12.
257. Marrakchi, H.; Lanéelle, M.-A.; Daffe, M., Mycolic acids: structures, biosynthesis, and beyond. *Chemistry & biology* **2014,** *21*, 67-85.
258. Garcia-Vilanova, A.; Chan, J.; Torrelles, J. B., Underestimated manipulative roles of Mycobacterium tuberculosis cell envelope glycolipids during infection. *Frontiers in immunology* **2019,** *10*, 2909.
259. Ly, A.; Liu, J., Mycobacterial virulence factors: Surface-exposed lipids and secreted proteins. *International Journal of Molecular Sciences* **2020,** *21*, 3985.
260. Chopra, T.; Banerjee, S.; Gupta, S.; Yadav, G.; Anand, S.; Surolia, A.; Roy, R. P.; Mohanty, D.; Gokhale, R. S., Novel intermolecular iterative mechanism for biosynthesis of mycoketide catalyzed by a bimodular polyketide synthase. *PLoS Biol* **2008,** *6*, e163.
261. Portevin, D.; de Sousa-D'Auria, C.; Houssin, C.; Grimaldi, C.; Chami, M.; Daffé, M.; Guilhot, C., A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. *Proceedings of the National Academy of Sciences* **2004,** *101*, 314-319.
262. Stinear, T. P.; Mve-Obiang, A.; Small, P. L.; Frigui, W.; Pryor, M. J.; Brosch, R.; Jenkin, G. A.; Johnson, P. D.; Davies, J. K.; Lee, R. E., Giant plasmid-encoded polyketide synthases produce the macrolide toxin of Mycobacterium ulcerans. *Proceedings of the National Academy of Sciences* **2004,** *101*, 1345-1349.
263. Yu, J.; Tran, V.; Li, M.; Huang, X.; Niu, C.; Wang, D.; Zhu, J.; Wang, J.; Gao, Q.; Liu, J., Both phthiocerol dimycocerosates and phenolic glycolipids are required for virulence of Mycobacterium marinum. *Infection and immunity* **2012,** *80*, 1381-1389.
264. Rousseau, C.; Sirakova, T. D.; Dubey, V. S.; Bordat, Y.; Kolattukudy, P. E.; Gicquel, B.; Jackson, M., Virulence attenuation of two Mas-like polyketide synthase mutants of Mycobacterium tuberculosis. *Microbiology* **2003,** *149*, 1837-1847.
265. Komaki, H.; Ichikawa, N.; Hosoyama, A.; Takahashi-Nakaguchi, A.; Matsuzawa, T.; Suzuki, K.-i.; Fujita, N.; Gonoi, T., Genome based analysis of type-I polyketide synthase and nonribosomal peptide synthetase gene clusters in seven strains of five representative Nocardia species. *BMC genomics* **2014,** *15*, 323.
266. Wohlleben, W.; Mast, Y.; Muth, G.; Röttgen, M.; Stegmann, E.; Weber, T., Synthetic biology of secondary metabolite biosynthesis in actinomycetes: engineering precursor supply as a way to optimize antibiotic production. *FEBS letters* **2012,** *586*, 2171-2176.
267. Katz, L.; Chen, Y. Y.; Gonzalez, R.; Peterson, T. C.; Zhao, H.; Baltz, R. H., Synthetic biology advances and applications in the biotechnology industry: a perspective. *Journal of Industrial Microbiology and Biotechnology* **2018,** *45*, 449-461.
268. Brown, P.; Dawson, M. J., A perspective on the next generation of antibacterial agents derived by manipulation of natural products. *Progress in medicinal chemistry* **2015,** *54*, 135-184.

269. Luo, Y.; Enghiad, B.; Zhao, H., New tools for reconstruction and heterologous expression of natural product biosynthetic gene clusters. *Natural product reports* **2016,** *33*, 174-182.
270. Baltz, R. H., Genetic manipulation of secondary metabolite biosynthesis for improved production in Streptomyces and other actinomycetes. *Journal of Industrial Microbiology and Biotechnology* **2016,** *43*, 343-370.
271. Dhakal, D.; Kumar Jha, A.; Pokhrel, A.; Shrestha, A.; Sohng, J. K., Genetic manipulation of Nocardia species. *Current protocols in microbiology* **2016,** *40*, 10F. 2.1-10F. 2.18.
272. Shizuya, H.; Birren, B.; Kim, U.-J.; Mancino, V.; Slepak, T.; Tachiiri, Y.; Simon, M., Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. *Proceedings of the National Academy of Sciences* **1992,** *89*, 8794-8797.
273. Wang, L.; Nasrin, S.; Liles, M.; Yu, Z., Use of bacterial artificial chromosomes in metagenomics studies, overview. *Encyclopedia of metagenomics. New York, USA: Springer* **2014**.
274. Martin, M.; Vandenbol, M., The hunt for original microbial enzymes: an initiatory review on the construction and functional screening of (meta) genomic libraries. *Biotechnologie, Agronomie, Société et Environnement* **2016,** *20*.
275. Fu, J.; Bian, X.; Hu, S.; Wang, H.; Huang, F.; Seibert, P. M.; Plaza, A.; Xia, L.; Müller, R.; Stewart, A. F., Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nature biotechnology* **2012,** *30*, 440-446.
276. Lee, N. C.; Larionov, V.; Kouprina, N., Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast. *Nucleic acids research* **2015,** *43*, e55-e55.
277. Zhang, J. J.; Yamanaka, K.; Tang, X.; Moore, B. S., Direct cloning and heterologous expression of natural product biosynthetic gene clusters by transformation-associated recombination. *Methods in enzymology* **2019,** *621*, 87-110.

9 Appendix

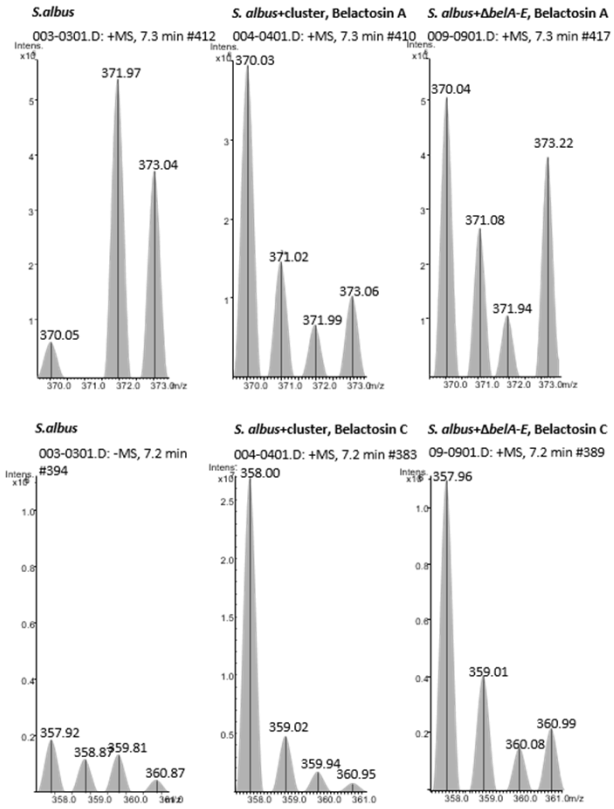
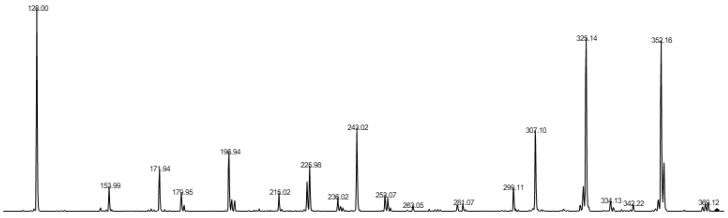


Figure 49: MS analysis of *S. albus* J1074 bel01 extracts: Detailed analysis of the *S. albus* J1074 bel01 Δ*belA-E*: mass spectra at 7.3 min (belactosin A) and 7.2 min (belactosin C) in the *S. albus (S. albus* J1074 bel01*)*, *S. albus+cluster (S. albus* J1074 bel01*)* and *S. albus+ΔbelA-E (S. albus* J1074 bel01 Δ*belA-E)*. Figure taken and adjusted from Engelbrecht et al.[195]

(i) MS/MS analysis of belactosin A from *S S. albus* J107 *bel01*; retention time 6.7-7.0 min, precursor ion *m/z* 370.2 $[M+H]^+$



(ii) MS/MS analysis of belactosin A from *S. albus* J107 *bel01*; retention time 7.0-7.3 min, precursor ion *m/z* 370.2 $[M+H]^+$

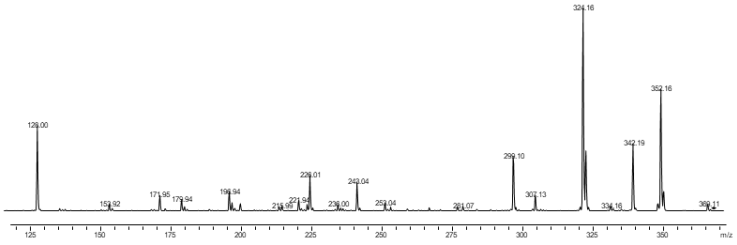


Figure 50: MS/MS analysis of belactosin A in the heterologous host strain *S. albus bel01*: Fragments for belactosin A (370.2 $[M+H]^+$) in *S. albus* J1074 *bel01* at retention time 6.7-7.0 (i) and 7.0-7.3 min (ii). Precursor mass is marked with a black diamond.

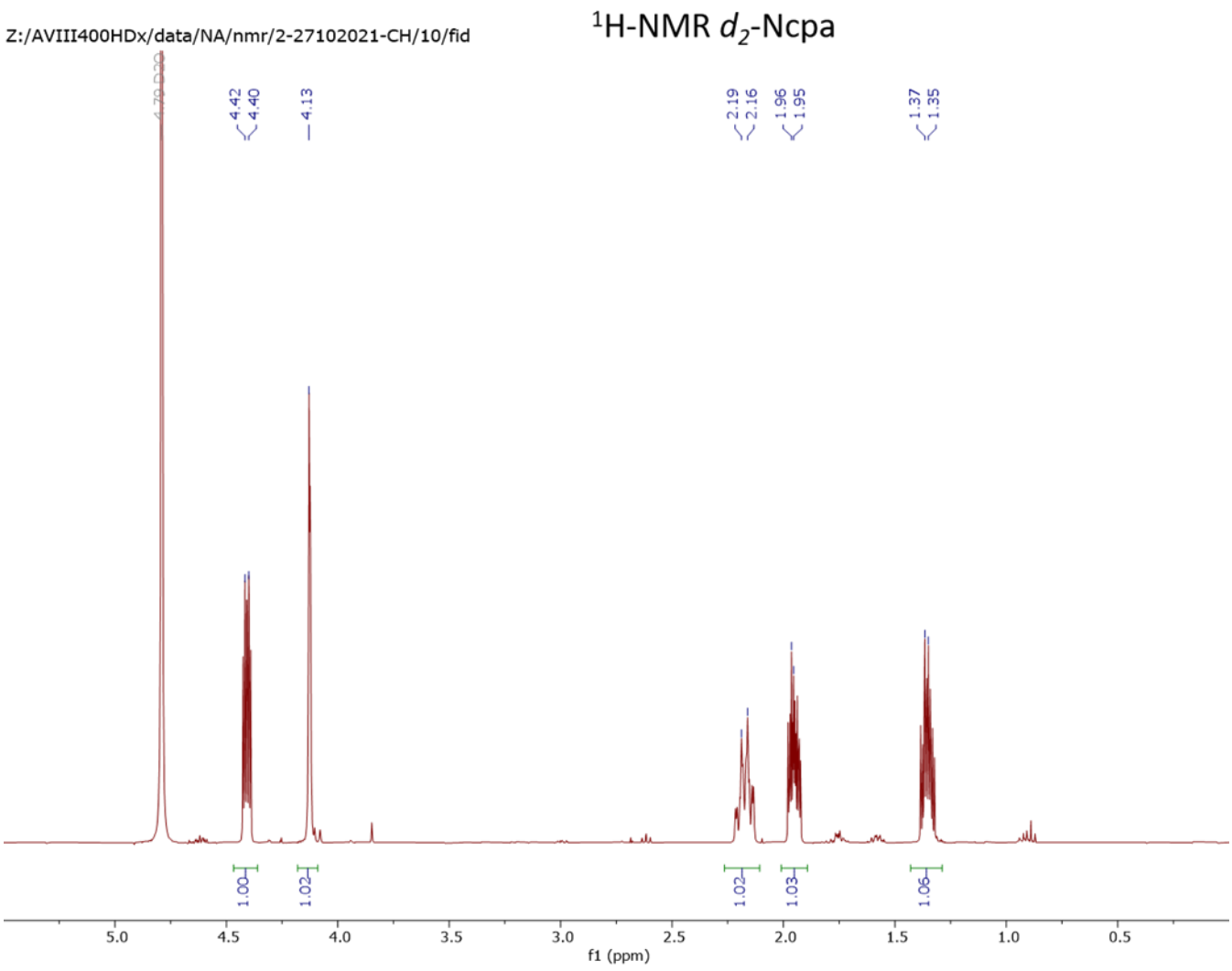


Figure 51: ^{1}H-NMR analysis of synthetic d2-Ncpa.

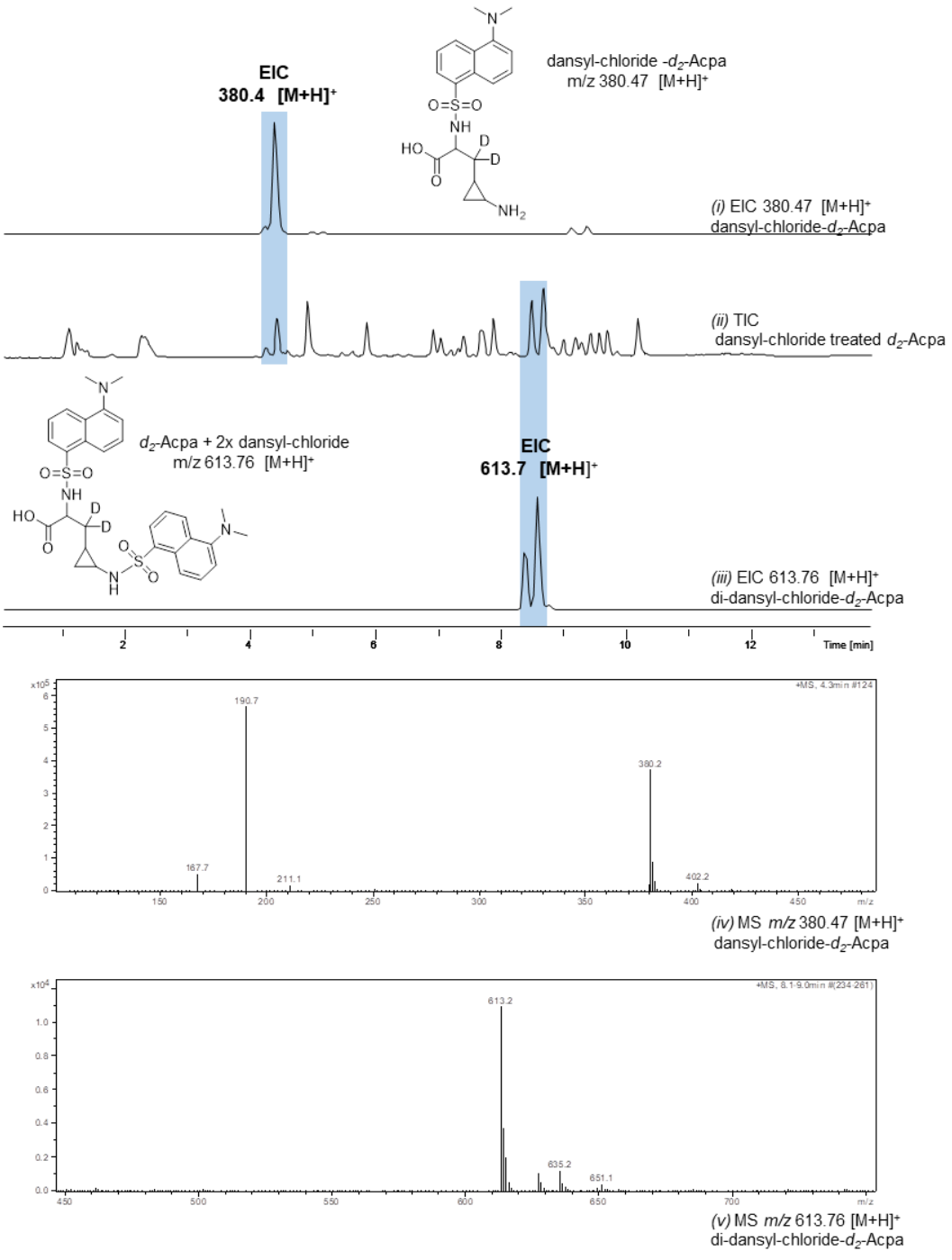


Figure 52: LC-MS analysis of single and doubled dansyl chloride derivatized d_2-Acpa: TIC of d_2-Acpa fed cultures extratcs treated with dansyl-chloride *(i)*, EIC *m/z* 380.4 $[M+2+H]^+$ dansyl-chloride-d_2-Acpa *(ii)* and EIC *m/z* 613.7 $[M+2+H]^+$ di-dansyl-chloride-d_2-Acpa *(iii)*. MS spectrum of *m/z* 380.47 $[M+2+H]^+$ dansyl-chloride-d_2-Acpa and *m/z* 613.76 $[M+2+H]^+$ di-dansyl-chloride-d_2-Acpa *(iv,v)*.

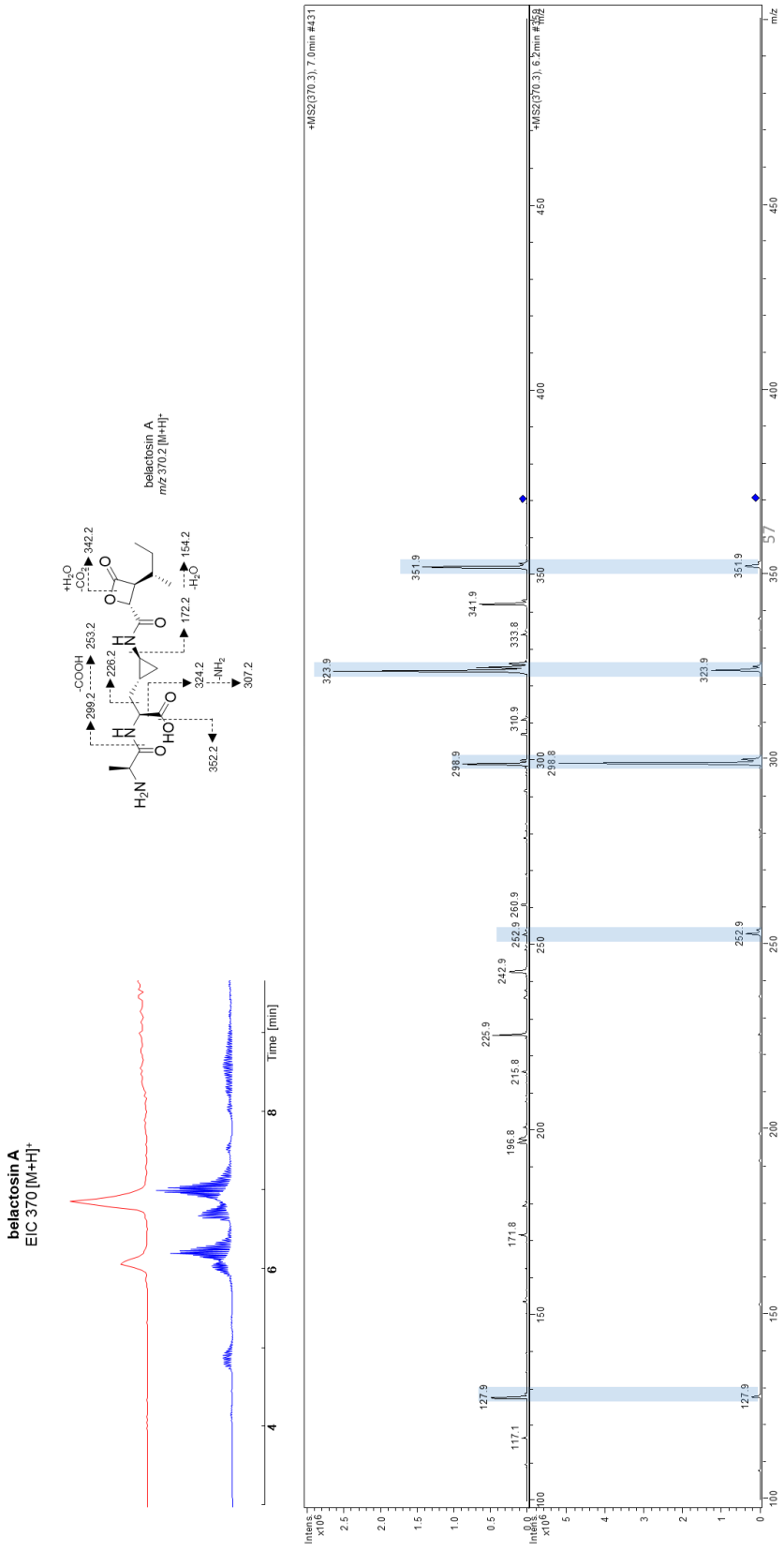


Figure 53: LC-MS and MS/MS analysis of belactosin A in *Streptomyces* sp. UCK14:*S.* sp. UCK14 shows production of two peaks for EIC 370.2 $[M+H]^+$ at retention time 6.2 min and 7.0 min. The related MS/MS spectra are shown at the bottom.

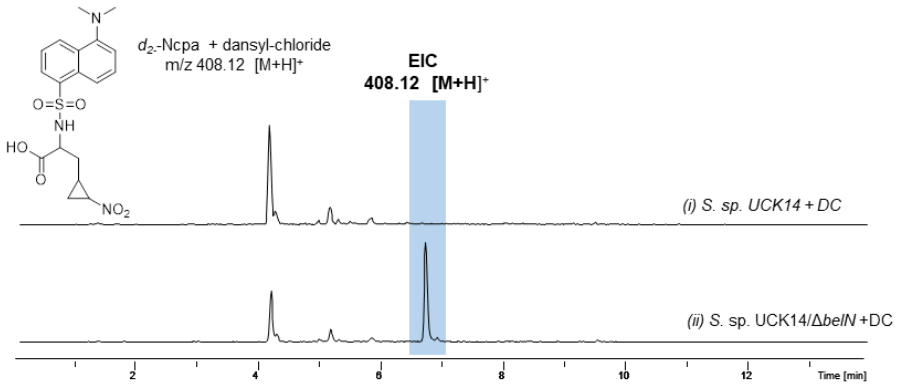


Figure 54: LC-MS analysis of dansyl chloride derivatized Ncpa:EIC *m/z* 408.12 $[M+H]^+$ in *S. sp.* UCK14 *(i)* and *S. sp.* UCK14/ Δ*belN (ii)*. DC = dansyl chloride

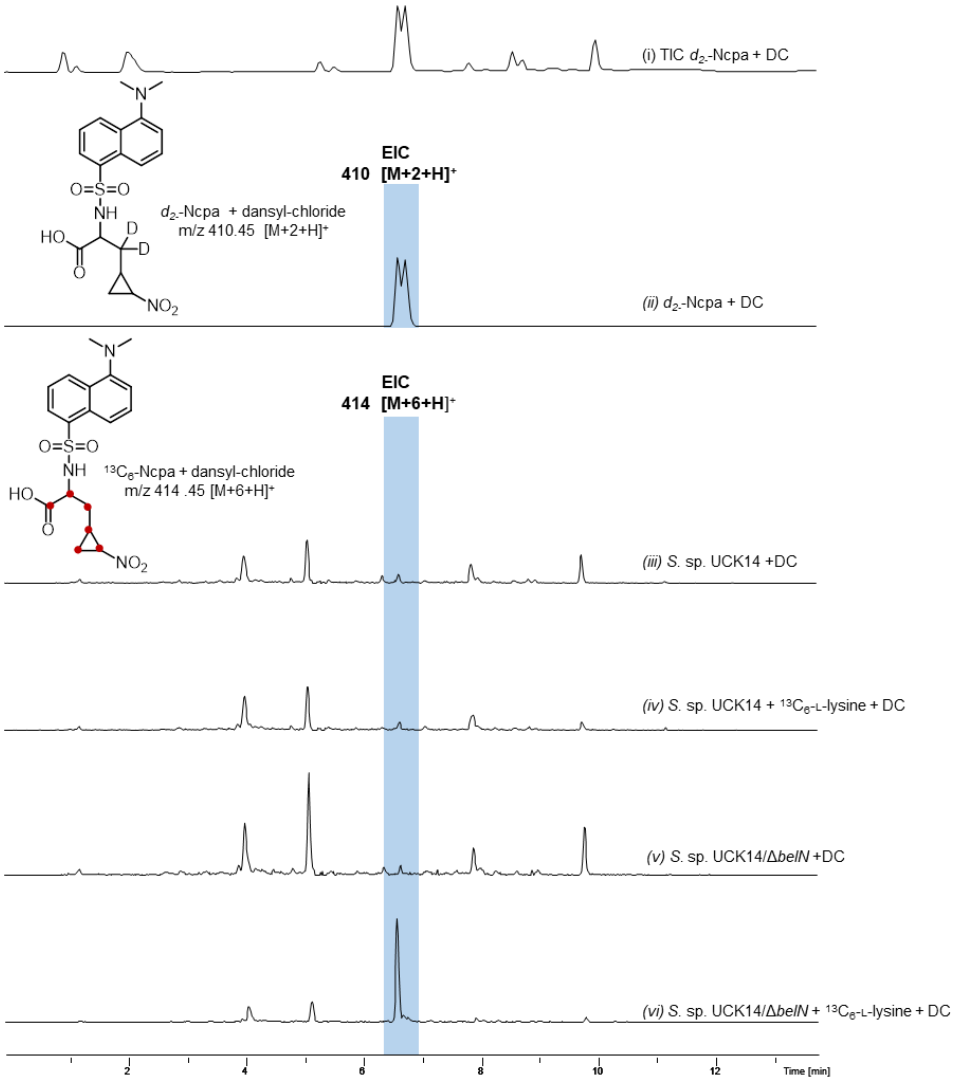


Figure 55: LC-MS analysis of dansyl chloride derivatized d_2-Ncpa:TIC *(i)* and EIC *m/z* 410 $[M+2+H]^+$ *(ii)*. LC-MS analysis of culture extracts of *S. sp.* UCK14 and *S. sp.* UCK14/Δ*belN* which were unfed *(iii, v)* or supplemented with $^{13}C_6$-L-lysine *(iv,vi)*. Accumulation of $^{13}C_6$ labeled and dansyl chloride derivatized Ncpa intermediate was followed by EIC 414 $[M+6+H]^+$. DC = dansyl chloride

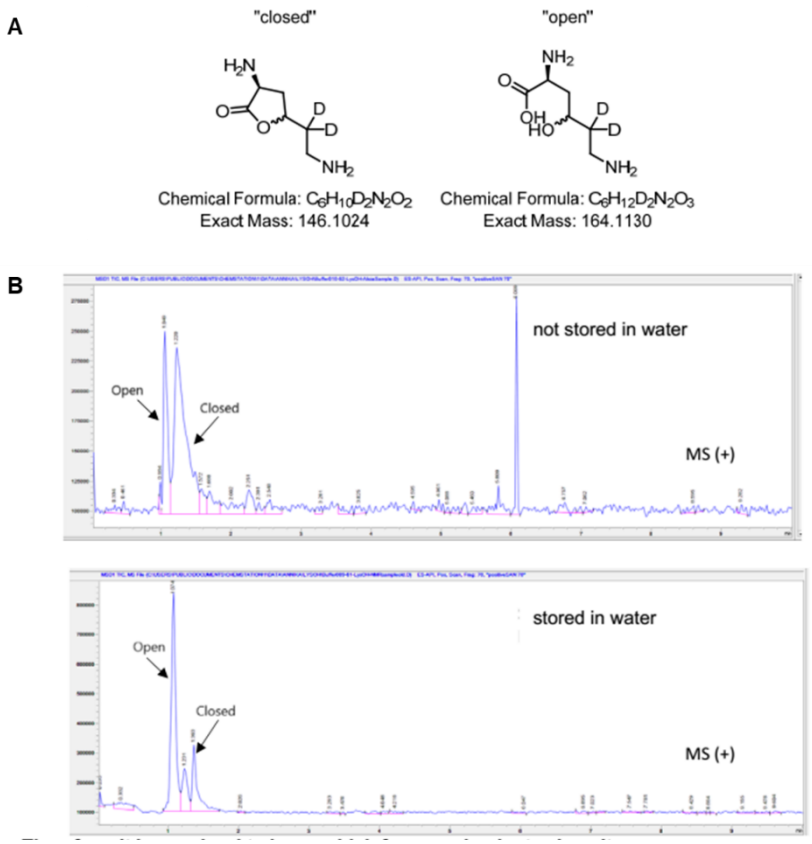


Figure 56: LC-MS anaylsis of syntheized 5,5-*d_2*-4-hydroxylysine lactone Both "open" and "closed" form are originally produced. If the compound is kept dry and stored in the freezer, a nearly 50:50 mixture is observed. If left in H_2O for several weeks the equilibrium seems to shift towards open form. Observed masses: 147 ([M+H]+ closed), 165 ([M+H]+ open) LC-MS data using a gradient of 10-100% MeCN in water (10 mM NH4OAc): Therefore, it is very hard to know which form predominates in culture. Figure was adapted from a template generated by Annika Esch and Chamber Hughes (Interfaculty Institute of Microbiology and Infection Medicine,

Microbial Bioactive Compounds).

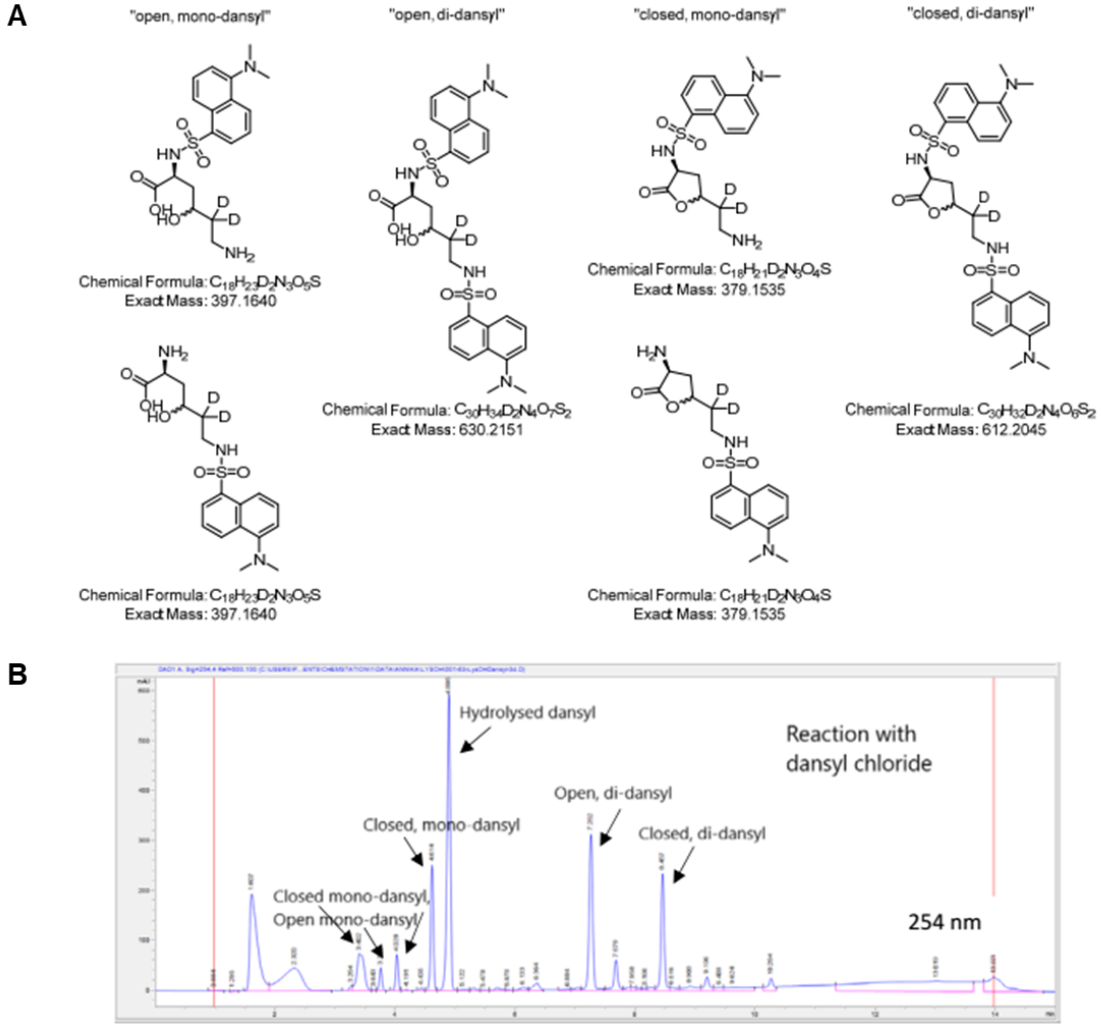


Figure 57: Derivatization of 5,5-d_2-4-hydroxylysine lactone with dansyl chloride:A: Structural formulas of possible derivatization products **B:** LC-MS analysis of dansyl chloride derivatization gave favorable retention time. Figure was adapted from a template generated by Annika Esch and Chamber Hughes (Interfaculty Institute of Microbiology and Infection Medicine, Microbial Bioactive Compounds).

Table 41: Automated molecular formula generation by SmartFormula (Bruker Analysis)

Meas. *m/z*	#	Ion formula	*m/z*	err [ppm]	Mean err [ppm]	rdb	mSigma
487.1594	1	C25H27O10	487.1599	1.0	1.0	12.5	3.0
487.1594	2	C22H19N10O4	487.1585	-1.8	-2.4	18.5	3.8
487.1594	3	C23H15N14	487.1599	1.0	0.0	23.5	13.5
487.1594	4	C8H11N26O	487.1604	2.0	1-1.7	16.5	59.0
487.1594	5	C10H23N12O11	487.1604	2.0	0.5	5.5.	70.4
487.1594	6	C7H15N22O5	487.1590	-0.7	-3.7	11.5	71.3
487.1594	7	C9H27N8O15	487.1590	-0.7	-1.9	0.5	83.6

Meas. = measured, err = error, rdb = rings and double bonds

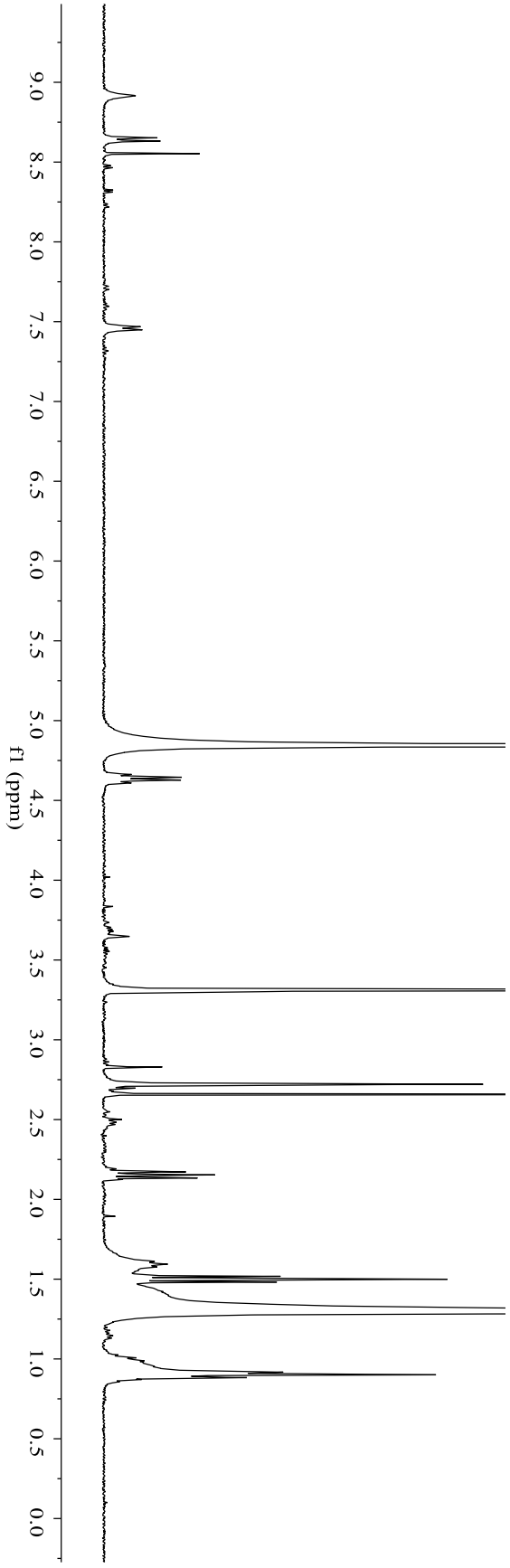


Figure 58: ^{1}H-NMR analysis of the ACAD compound

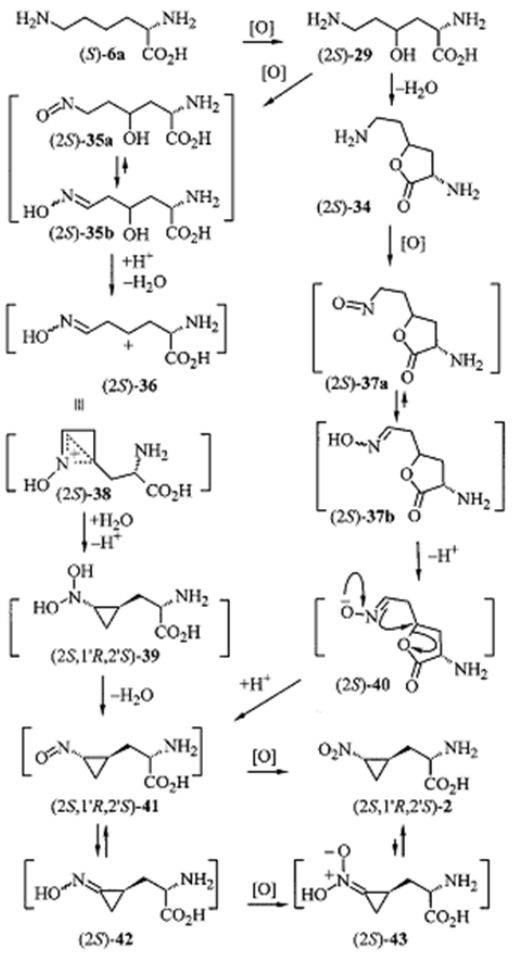


Figure 59: Proposed biosynthesis of Ncpa by Brandl et al.: Figure taken and adapted from Brandl et al.[173]